E-Infrastructures

H2020-INFRAEDI-2018-1

INFRAEDI-2-2018: Centres of Excellence on HPC

EoCoE-II

Energy oriented Center of Excellence : toward exascale for energy

Grant Agreement Number: INFRAEDI-824158

D1.3

Final scientific results and exascale tools delivery

Project and Deliverable Information Sheet

EoCoE-II	Project Ref:	INFRAEDI-824158
	Project Title:	Energy oriented Centre of Excellence: towards exascale for energy
	Project Web Site:	http://www.eocoe.eu
	Deliverable ID:	D1.3
	Deliverable Nature:	Report
	Dissemination Level:	PU*
	Contractual Date of Delivery:	M18 30/06/2022
	Actual Date of Delivery:	30/06/2022
	EC Project Officer:	Matteo Mascagni

* - The dissemination level are indicated as follows: PU – Public, CO – Confidential, only for members of the consortium (including the Commission Services) CL – Classified, as referred to in Commission Decision 2991/844/EC.

Document Control Sheet

Document	Title :	Final scientific results and exascale tools delivery
	ID :	D1.3
	Available at:	http://www.eocoe2.eu
	Software tool:	LaTeX
Authorship	Written by:	Edoardo Di Napoli (FZJ)
	Contributors:	Herbert Owen (BSC), Frédéric Blondel (IFPEN), Gerard Guillamet (BSC), Ani Anciaux-Sedrakian (IFPEN), Garrett Good (IEE), Yen Sen Lu (FZJ), Philipp Franke (FZJ), Florian Dupoy (MF), Francesco Buonocore (ENEA), Alessandro Pecchia (CNR), Mathieu Salanne (CEA-MdlS), Alison Walker (UBAH), Irene Aguilera (FZJ), Bibi S. Naz (FZJ), Bruno Majone (UNITN), Andrea Galletti (UNITN), Sophie Ricci (CERFACS), Thanh Huy Nguyen (CERFACS), Andrea Piacentini (CERFACS), Johanna Fink (RWTH), Yanick Sarazin (CEA-IRFM), Virginie Grandgirard (CEA-IRFM), Eric Sonnendruecker (MPG-IPP).
	Reviewed by:	Edouard Audit, Bruno Raffin

Document Keywords: Computational Fluid Dynamics, Large Eddy Simulation, Complex Terrain, Full Rotor, Density functional theory, Molecular dynamics, Quantum Transport, Photovoltaics, Quantum Monte-Carlo, Capacitance, Seebeck coefficient, Thermo-electrochemical devices, Multiscale Modeling, Monte Carlo, TerrSysMp, Reservoir, Hydropower, Geothermal, ITER, Nuclear Fusion, Wind Turbine.

Contents

List of Figures

List of Tables

1 Executive summary

This report describes the final scientific results of the Work Package 1 (WP1) and the outcome of the corresponding flagship codes towards the goal of exascale computing. The scope of the EoCoE-II project is to build on its unique, established role at the crossroads of HPC and renewable energy to accelerate the adoption of production, storage and distribution of clean electricity. We realize such a target by developing a sustainable structure able to develop state-of-the-art numerical tools and promote the usage of HPC in the energy domain. At the core of this structure there are five Energy Science Challenges (SC) addressing the most important and diverse area in the energy domain. The scientific payload of these challenges is the target of the WP1. Each challenge is a task labeled by one letter (**T** for task) and two numbers, the first indicating the work package and the second specific to the challenge.

Label	Energy Science Challenge
T1.1	Wind for Energy
T1.2	Meteorology for Energy
T1.3	Materials for Energy
T1.4	Water for Energy
T1.5	Fusion for Energy

Each of these challenges focuses on a selected number of main tasks whose payload provides significant advances in its respective domain. These advances will trigger innovation in fields such as energy production by wind turbines, photovoltaic cells, and hydroelectric and geothermal stations; prediction of weather forecast and cloud cover for the better placement of renewable power stations; simulations and validation of water resources at the surface and subsurface level for better water management; up until future energy production from nuclear fusion. In addition, advances in these energy domains prove the benefit of exascale computing for the growing low-carbon energy community as well as pave the way to exascale for other related energy sectors such as energy transportation and storage.

Wind. Significant advances have been made in the simulation of the wind over complex terrain using Large Eddy Simulations (LES) that have positioned Alya among the best available codes worldwide as well as the European alternative to US-developed open source exascale codes that are part of the Exawind and A2E projects. In addition, significant advances have been made in the mesoscale coupling which went beyond the goals described in the original proposal, making the implementation in Alya among the best available. Alya now includes thermal coupling to treat stable and unstable conditions with the possibility of simulating diurnal cycles using the tendencies approach for mesoscale coupling. Alya's LES simulation can include: Coriolis forces. Modeling of the forest using a canopy model. Inclusion of wind turbines with an actuator disc model. Participating in several Benchmarks has enhanced Alya's visibility within the EU wind community.

Alya's capabilities for the fully resolved simulation of wind turbines have improved significantly during EoCoE-II. The sliding mesh approach used to incorporate the rotation of the wind turbine blades involves the coupling between a fixed grid for most of the domain and a rotating grid around the turbine blades. The coupling algorithm was relatively new at the beginning of EoCoE-II, and it soon became apparent that it required significant rewriting since it had become too complex and error-prone. The re-writing has involved significant debugging and testing, heading to a better structured and easier to develop code. We have simulated a 1.5 MWatt wind turbine including mast and nacelle. The end result is a simulator that goes beyond the results of the ExaWind US project, which does not include the mast. Aeroelastic simulations have also been tackled during EoCoE-II. In this case, fluid-structure interaction (FSI) simulations using wall model LES have been conducted on the whole blade of the aforementioned wind turbine. The results show that large deformations are predicted for the small pitch angle. When the pitch angle is +87.4º (a typical angle in service conditions), the maximum deflection predicted is 0.2m with a wind speed of 24m/s.

Meteo. The scientific challenge Meteorology for Energy centered around the flagship code ESIAS for ultra-large ensembles of numerical weather predictions and their application to renewables forecasting. In EoCoE-II, ESIAS was integrated with the MELISSA middle-ware (part of Work Package 5) to demonstrate ultra-large simulations as laid out in the roadmap. This framework could be used to perform an extensive sensitivity study on the performance of thousands of physics combinations in ESIAS-Met (WRF), finding the best combinations for the European domain. ESIAS-Met was also coupled to ESIAS-Chem and a study completed on its ensemble performance and the effect on irradiance.

The optimized ESIAS-Met configuration was used to generate a year of ensemble data for renewables forecasting of solar and wind feed-in. This representative data could be used to calibrate the solar and wind power models of electric feed-in in Germany to the WRF output. In the end, a similar quality to operational weather services was achieved with the research model, though some unexplained features in the irradiance modeling remained.

Satellite observations of cloud positions and optical flow solutions of their motions and the underlying geostrophic structures can be used to validate ensemble members. For solar we showed that a 512-member ensembles consistently yielded a few solutions similar to the observation, which could be quite useful in the application of grid congestion.

The results highlight some divergent goals in the application of ultra-large ensembles and ideas for future work. EoCoE-I was concerned with costs to large-scale energy markets due to extreme weather, but EoCoE-II showed how calibrating the meteorological outputs for regional forecasts or performing particle filtering improved accuracy at the cost of ensemble resolution. Future work can still address both sides. Not all events are captured, whether by a lack of variance in the modeling or inaccuracy of initial condition. The general improvement of ESIAS-Met for the domain would addresses issues of accuracy. There is also growing interest in "long-term" forecasts, which could certainly be a more dynamic application of ultra-large ensembles.

Regarding short-term accuracy and spatial variance, as opposed to rare weather events, a very practical and immediate implementation of ultra-large ensembles would be the very uncertain high-resolution details of clouds, which although unremarkable for the aggregate national feed-in or energy market, could have value on a daily basis for grid operation, especially as PV and smart metering dominate the low voltage grids of the future.

Materials. In the challenge “Materials for Energy”, we combined ReaxFF MD simulations and ab initio calculations to investigate the intra-gap states of the a-Si:H/c-Si interfaces at different annealing temperatures. To this end, we represented the Hamiltonian in the more appropriate basis set of Wannier functions so as to make future transport calculations of the c-Si/a-Si:H interface fully ab initio: the WANTRANS code has been implemented to interface the transport code libNEGF and the results from a DFT/wannier90 calculation. Since WANTRANS interface needs some further development before to be applied to the cSi/aSi/cSi structure, we resolved to compute the transport properties across the heterojunction using the density-functional tight binding semi-empirical approach (DFTB). The I-V currents across the junctions for holes and electron currents have been computed using LIBNEGF.

Initially, the flagship code of choice was PVNEGF, developed within the IEK-5 institute part of the FZJ partner. Due to several drawbacks of the PVNEGF code (advanced functionalities have been developed based on a simplified one-dimensional geometry and the lead developer left the project at its very beginning) that would not easily allow the simulation of interfaces between amorphous and crystalline silicon (as stated in the task T1.3.1-3 of D1.1), we changed the flagship code to LIBNEGF developed within the CNR. LIBNEGF already implemented the full 3D multiband treatment. The extension of LIBNEGF by including non-ballistic scattering has been developed within the scope of the EoCoE project.

We performed high-level Quantum Monte-Carlo (QMC) simulations of adsorbed molecules/ions at the surface of carbon materials to drive the choice of the most accurate exchange-correlation functional to be used for the parameterization of the potential for the interactions. The vdW-DF-C09 functional better catches the interactions. The potentials have been used in classical DFT (using MDFT) and molecular

dynamics (using METALWALLS) to simulate, respectively: a) the adsorption free energy of the lithium ion on the carbon surface; and b) a system in which the ionic liquid is confined between a pair of planar electrodes held at constant electrical potential, which fluctuates due to the thermal motion of the ions in the liquid.

Algorithmic optimization of the FMM-KMC code for the KMC simulator has been developed and applied to the study of doped organic semiconductors demonstrating the scaling of the electrostatics library, and to the simulation of all-solid-state batteries. The BOLTMC code for the DMC simulator, which solves the Boltzmann transport equation, has been optimized and used to investigate the polaronic effects on scattering and mobilities of charge carriers in halide perovskites. The PPMD Python framework has been used for the optimization of the two codes.

Water. The "Water for Energy" challenges tackles a number of diverse issues dealing with water management and energy production from surface water usage. This challenge is divided in five main tasks, each with its own objective both in terms of scientific payload and code scaling.

In the first task ParFlow hydrologic model is used to simulate hydrologic states and fluxes relevant to the energy sector. ParFlow is a massively parallel, physics-based integrated hydrologic model and simulates fully coupled dynamic 2D/3D hydrological, groundwater and land-surface processes for large scale problems. Using ParFlow-GPU version (see details in Section 6.6), the model was setup at 3 km resolution over Europe and an explicit simulation of lateral groundwater flow, groundwater discharge and recharge was performed for 10 years of time period (1997 - 2006). In addition, a comprehensive evaluation of hydrologic states and fluxes was performed using in-situ and remote sensing observations including discharge, surface soil moisture, evapotranspiration and water table depth (Task T1.4.2). Overall, the uncalibrated ParFlow model shows good agreement in simulating river discharge for 176 gauging stations across Europe. Comparison with satellite-based datasets of soil moisture (SM) shows that ParFlow performs well in semi-arid and arid regions, but simulates overall higher SM in humid and cold regions. We conclude that the addition of alluvial aquifers in a continental scale hydrologic model can illuminate the complex relationship of water transfer to and from river systems and also significantly affect the transportation of water through the continent.

In the second task, HYPERstreamHS hydrological model was refactored in order to allow explicit representation of hydropower systems while preserving good computational performance overall. The Human System modules have been tuned with reference to the Adige river basin, in order to ensure reliable prediction basing the modules solely on publicly available information: this allows to apply the framework to other areas, provided that the required information is available. The model showed great performances in reproducing historical streamflows and hydropower production, highlighting some shortcomings of the main approaches to large scale hydropower modeling that are present in literature [53]. After compiling an extensive dataset containing the relevant information, the model was then applied to other large Alpine catchments, showing satisfactory performances.

In order to improve streamflow reproduction, we carried on a benchmarking exercise in the Adige river basin aimed at the identification of suitable modifications of the Community Land Model v3.5 (CLM3.5) runoff outputs. This activity highlighted that the inclusion of some physical processes, as well as replacing the grid based routing scheme of CLM3.5 with a more accurate scale-independent routing scheme can greatly improve the ability of CLM3.5 to reproduce observed streamflows, particularly by improving its reproduction of sub-surface flow component. Overall, we developed useful tools for highly reliable streamflow and hydropower production modeling, as well as setting them up for future coupling with hyper-resolved products.

In the third task a mixture of experts surrogate models has been used to carry out a global sensitivity analysis to classify sources of uncertainty to explain water level variance over the 2D domain. This work was carried out for stationary flow and led to the estimation of Sobol indices maps. The cost reduction allowed by the surrogate model used in place of the direct solver allows to compute Sobol indices and rank the sources of uncertainties over the Garonne Marmandaise catchment. Analysis of the first order Sobol

indices reveals the large influence of the discharge and smaller influence of friction. It should be noted that the surrogate model was established for stationary flow, assuming the uncertainties relate to friction and upstream forcing. The original plan was to force the surrogate T2D model with ParFlow discharge outputs at the local Telemac model over 50km of the Garonne river. Yet, the upgrade of the surrogate model previously described from stationary flow to non-stationary flow was not completed. ParFlow was thus chained to the direct model Telemac only. While the use of a surrogate model would have lowered the computational cost of the chained simulation, the IO interfaces for the direct and the surrogate remain similar. The ParFlow-Telemac chained hydrology-hydraulic model was implemented and tested over 2003-2004 with a focus on 2 significant flood events. It was shown that ParFlow provide good discharge time series, yet imperfect with underestimation of the flow and a 5 day delay. These uncertainties translate into similar errors in the outputs of the local model outputs with Telemac. Several strategies are possible for improvement ranging from ParFlow calibration of friction, ParFlow off-line rescaling for discharge or error correction in Telemac with Data Assimilation in the local hydrodynamics model only. These are possible leads for further research.

In the last two tasks we elaborated an OED workflow that is applicable to geothermal reservoir models with the open-source code SHEMAT-Suite for solving different OED problems in context of geothermal exploration. Several OED problems were simulated on synthetic models for investigating the influence of prior assumptions regarding estimation of permeability and thermal conductivity. In addition, optimal borehole locations were defined for a realistic 3D reservoir model in order to estimate thermal conductivity or basal heat flow with least uncertainty. Due to reduced work efficiency and missing child care during the pandemic, some research topics could not be addressed as deeply as intended initially. Therefore, the test model suite is smaller than intended initially (T1.4.4-6). Transient test models or an advective 3D model are missing. Furthermore, the OED functionality has not been extended to optimizing not only borehole location but also borehole depth. Finally, T1.4.5 was adapted in order to be feasible within the remaining time, resulting in a quasi-synthetic scenario based on a realistic 3D structural model and realistic rock properties.

Fusion. We aimed at addressing the issue of plasma confinement from the very hot core to the unconfined peripheral region of tokamaks—the so-called scrape-off layer (SOL)—in ITER-relevant configurations and parameter ranges. In order to achieve this target we push the development of the flagship 5-dimensional GYSELA code – with the support of satellite codes – towards the exascale limit.

Major physical upgrades have been implemented and both numerical and HPC bottlenecks have been alleviated, while continuously improving numerical performance. In particular, (i) GYSELA can now handle plasma-wall interactions with adiabatic electrons; the case of kinetic electrons has been studied with a low-dimensional version of GYSELA , namely VOICE, as a preliminary step. (ii) Non-circular poloidal cross-sections of the magnetic surfaces can now be considered, together with non axi-symmetric perturbations of the equilibrium magnetic field. The treatment of the X-point has led to dedicated studies – although not yet implemented in GYSELA – involving satellite codes and on-purpose developments of reduced models; importantly, this analysis has allowed us to identify the optimal strategy for GYSELA . (iii) Variations of several orders of magnitude in plasma temperature from core to edge can now be handled at an affordable memory cost thanks to the development of non-equidistant splines. (iv) The implemented electromagnetic effects (Maxwell-Ampère equation) are currently being benchmarked.

GYSELA is now one of the rare gyrokinetic codes worldwide capable of modeling ion turbulence and collisional transport from the core to the SOL in the relevant flux-driven regime of tokamak plasmas. These cutting edge upgrades have led to decisive breakthroughs, as attested most notably by the two highlighted papers published in top rank peer-reviewed scientific journals, namely Physical Review Letters (American Physical Society) and Communications Physics (Nature). In particular, we derive a simple criterion to decide whether edge flows are governed by turbulence or collisions when the axisymmetry of the magnetic configuration in tokamaks is weakly broken, as is usually the case. Last but not least, we predict that pressure inhomogeneities and finite Larmor radius effects are key in the development of large scale flows at

the tokamak edge, and ultimately suspected to be critical in the triggering of bifurcated states of enhanced confinement, which are the reference scenarios to achieve the most performant discharges in ITER. These findings have attracted the attention—in view of possibly resolving the misunderstood mismatch of flow measurements with earlier theories-—of experimentalists running adequate diagnostics on the Spanish Heliac TJ-II.

2 Acronyms

Table 1: Acronyms for the partners and institutes therein.

Acronym	Partner and institute
AMU:	Aix-Marseille University
BSC:	Barcelona Supercomputing Center
CEA:	Commissariat à l'énergie atomique et aux énergies alternatives
CERFACS:	Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique
CIEMAT:	Centro De Investigaciones Energeticas, Medioambientales Y Tecnologicas
CoE:	Center of Excellence
CNR:	Consiglio Nazionale delle Ricerche
EDF:	Électricité de France
ENEA:	Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile
FAU:	Friedrich-Alexander University of Erlangen-Nuremberg
FSU:	Friedrich Schiller University
FZJ:	Forschungszentrum Jülich GmbH
IBG-3:	Institute of Bio- and Geosciences Agrosphere
IEK-8:	Institute for Energy and Climate Research 8 (troposhere)
IEE:	Fraunhofer Institute for Energy Economics and Energy System Technology
IFPEN:	IFP Énergies nouvelles
INAC:	Institut nanosciences et cryogénie
INRIA:	Institut national de recherche en informatique et en automatique
IRFM:	Institute for Magnetic Fusion Research
ISMN:	Istituto per lo Studio dei Materiali Nanostrutturati
MdlS:	Maison de la Simulation
MF:	Meteo France
MPG:	Max-Planck-Gesellschaft
RWTH:	Rheinisch-Westfälische Technische Hochschule Aachen, Aachen University
UBAH:	University of Bath
UNITN:	University of Trento

Table 2: Acronyms of software packages

Acronym	Software and codes
CLM:	Community Land Model
EFCOSS:	Environment For Combining Optimization and Simulation Software
ESIAS:	Ensemble for Stochastic Interpolation of Atmospheric Simulations
EURAD-IM:	EURopean Air pollution Dispersion-Inverse Model
HYPERstreamHS:	Dual-layer MPI large scale hydrological model including Human Systems
ICON:	Icosahedral Nonhydrostatic model
MDFT:	Molecular Density Functional Theory
MELISSA:	Modular External Library for In Situ Statistical Analysis
Meso-NH:	Mesoscale Non-Hydrostatic model
Nemo5:	NanoElectronics MOdeling Tools 5
neXGf:	non-equilibrium eXascale Green's functions library
OpenFOAM:	Open Source Field Operation and Manipulation
ParFlow:	PARallel Flow

PF-CLM:	Parflow-Community Land Model
PPMD:	Performance Portable Molecular Dynamics
ReaxFF:	Reactive Force Field
SHEMAT:	Simulator of HEat and MAss Transport
SOWFA:	Simulator fOr Wind Farm Application
SPS:	Solar Prediction System
TELEMAC:	TELEMAC-MASCARET system
TerrSysMP:	Terrestrial Systems Modeling Platform
WaLBerla:	A Widely Applicable Lattice-Boltzmann Solver
WanT:	Wannier Transport
WPMS:	Wind Power Management System
WRF:	Weather Research and Forecast model

Table 3: Acronyms for the Scientific Terms used in the report.

Acronym	Scientific Nomenclature
2D:	Two-Dimensional
3D:	Three-Dimensional
ABL:	Atmospheric Boundary Layer
AD:	Automatic Diffentiation
ALEX17:	Alaiz Experiment 2017
AOT:	Aerosol Optical Thickness
PBE:	Perdew-Burke-Ernzerhof functional
BLYP:	Becke-Lee-Yang-Parr functional
COT:	Cloud Optical Thickness
COVID-19:	Corona Virus Disease 2019
COSMO-REA6:	Convective-scale Regional Reanalysis System 6 Km
CPU:	Central Processing Units
CSP:	Concentrated Solar Power
CUDA:	Compute Unified Device Architecture
DA:	Data Assimilation
DEM:	Digital Elevation Model
DFT:	Density Functional Theory
DMC:	Dynamic Monte Carlo
EnKF:	Ensemble Kalman Filter
E-RUN:	European Runoff
ESACCI:	European Space Agency Climate Change Initiative
ESSMRA:	European Surface Soil Moisture reanalysis
ET:	Evapotranspiration
FSI:	Fluid-Structure Interaction
FLUXNET:	Flux Network
GMTED2010:	Global Multi-resolution Terrain Elevation Data 2010
GPU:	Graphical Processing Unit
HLST:	High Level Support Team
HPC:	High Performance Computing
IEA-Wind:	International Energy Agency - Wind
ITER:	International Thermonuclear Experimental Reactor
ISMN:	International Soil Moisture Network
KMC:	Kinetic Monte Carlo

LES:	Large Eddy Simulations
LiDAR:	Light Detection and Ranging
MD:	Molecular Dynamics
MODIS:	Moderate Resolution Imaging Spectroradiometer
MPI:	Message Passing Interface
NEGF:	Non-Equilibrium Greens functions
NEWA:	New European Wind Atlas
NREL:	National Renewable Energy Laboratory
NWP:	Numerical Weather Prediction
OED:	Optimal Experimental Design
OpenMP:	Open Multi-Processing
PBC:	Periodic Boundary Conditions
PDAF:	Parallel Data Assimilation Framework
pdf:	probability density functions
PDOS:	Projected Density of States
PF-CLM:	Parflow-Community Land Model
PRUDENCE:	Prediction of Regional Scenarios and Uncertainties for Defining European Climate Change Risks and Effects
QMC:	Quantum Monte Carlo
QM:	Quantum Mechanics
SHJ:	Silicon HeteroJunction
SM:	Soil Moisture
SOL:	Scrape-Off Layer
WP:	Work Package
WTD:	Water Table Depth

Figure 1: Word cloud of the Work Pakage 1.

3 Wind for Energy (T1.1)

T1.1 is divided into two main tasks:

T1.1.1	Development, Verification, and Validation for Complex terrain
T1.1.2	Development, Verification, and Validation for Full rotor

T1.1.1 focused on flow over Complex terrain has improved wall modeling, and inflow boundary conditions in Alya's low dissipation Large Eddy Simulation (LES) formulation for Atmospheric Boundary Layer (ABL) flows. The model now includes thermal coupling, Coriolis forces, canopy, and the actuator disc. Moreover, important advances in the coupling with mesoscale models have been performed. Alya has been compared against Meso-NH. Moreover, Alya has participated in several community benchmarks. The model has been tested on real wind farms as part of an ongoing collaboration with Iberdrola.

T1.1.2 has enhanced Alya's full rotor model, where the actual geometry of the wind turbine blades and tower is modeled exactly. A sliding mesh approach has been used to incorporate the rotation of the blades. Their deformation has been considered using a fluid-structure interaction (FSI) approach. A comparison between the actuator line (IFPEN) model in WaLBerla and the full rotor in Alya has been performed.

3.1 Task T1.1.1 - Scientific results

The main task, T1.1.1, is subdivided into five subtasks.

T1.1.1-1	**Improve wall modeling for Atmospheric Boundary Layer of Large Eddy Simulation in the code Alya developed at BSC**.
T1.1.1-2	**Improve the inflow boundary conditions for the Atmospheric Boundary Layer of Large Eddy Simulation for the Alya code**.
T1.1.1-3	**Including thermal coupling, Coriolis forces, canopy, and the actuator disc**.
T1.1.1-4	**Benchmarking against the Meso-NH code**.
T1.1.1-5	**Test the new model against a realistic set up**.

T1.1.1-1 - Improve wall modeling for Atmospheric Boundary Layer of Large Eddy Simulation in the code Alya developed at BSC

Task T1.1.1-1 was completed during the project's first half and was reported in deliverable D1.2. It is interesting to note that the Ph.D. Student, Sarath Radhakrishnan, has been working on extending the work from Task T1.1.1.1 to cases with adverse pressure gradients using artificial intelligence. Oriol Lehmkuhl directs him, and he has also received significant informal support from Herbert Owen on wall modeling and finite element for flow problems.

T1.1.1-2 - Improve the inflow boundary conditions for the Atmospheric Boundary Layer of Large Eddy Simulation for the Alya code

As an alternative to prescribing boundary conditions at the lateral boundaries, which are usually not known for cases of industrial relevance, we have opted to treat the lateral boundaries as periodic boundaries and force the flow with results from a mesoscale model. This mesoscale coupling methodology was initially developed during the NEWA project. Its implementation in Alya has been improved during EoCoE-II, and it has been applied successfully to several wind farm benchmarks.

The results we sent for the Hornamossen test case during the first half of EoCoE-II were improved by coupling them to mesoscale results from WRF. Herbert Owen and Matias Avila have participated in most of the Hornamossen benchmark meetings. They submitted results using LES and RANS turbulence models for the two suggested wind directions. Participation in the benchmark was very enriching because it generated a strong interaction with other European wind energy groups. The Benchmarks has suffered some delays due to the paternity leave of the main organizer, Johan Arnqvist, but we expect the final results should be published soon. Figure 2 shows preliminary results for the wind velocity under stable conditions at different masts from the Hornamossen Benchmark presented at the IEA-Task 31 Wakebench in 2021. It includes results with Alya using both RANS and LES.

T1.1.1-3 - Including thermal coupling, Coriolis forces, canopy, and the actuator disc.

Thermal coupling, Coriolis forces, canopy models, and an actuator disc model have been introduced and tested in the LES version of Alya for wind problems.

Thermal coupling needed to simulate the diurnal cycle and couple to the mesoscale models proved particularly challenging. The Boussinesq forcing used to model the thermal effects on the Navier Stokes equations led to instabilities close to the ground when using a Fractional Step method to stabilize the pressure and allow for equal order interpolation. The previous deliverable found a solution for the problem of flat terrain cases. The method has been extended to complex terrain during the second half of EoCoE-II. It has removed the instabilities and has been a critical enabler in thermally coupled simulations using the fractional step scheme.

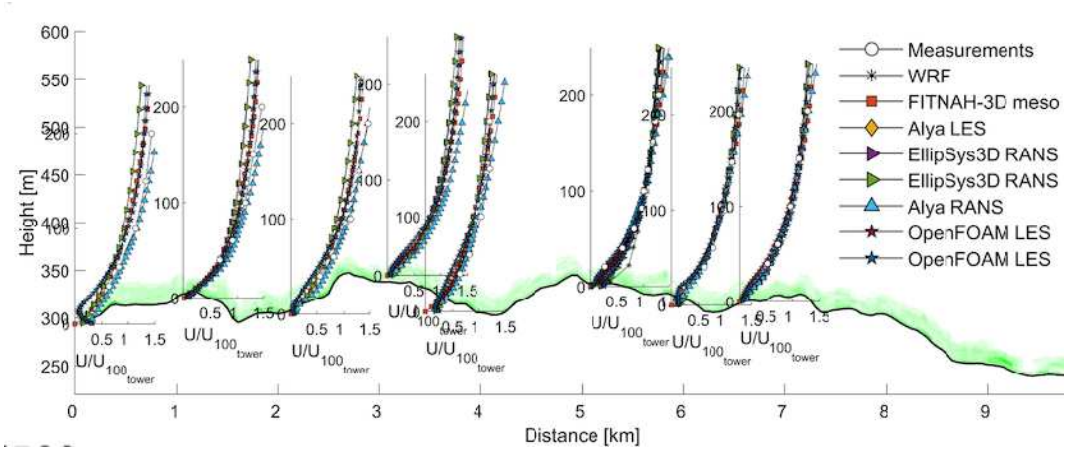


Figure 2: Wind velocity under stable conditions - Hornamossen Benchmark.

Regarding the actuator disc model, we have directed the Master's student, Eva Peinado Montoya, in her thesis in Aerospace Engineering entitled, 'Large Eddy Simulations of wind turbines using the actuator disc model implemented in Alya' at the BarcelonaTech university.

T1.1.1-4 - Benchmarking against the Meso-NH code.

The ALEX17 benchmark was chosen to benchmark Alya against Meso-NH. Moreover, it would allow comparing against several other codes. In the previous deliverable (D1.2), an introduction to the benchmark was provided together with initial results from Meso-NH.

The ALEX17 campaign is the last of a series of experiments carried out in the New European Wind Atlas (NEWA) project1. A benchmark has been launched by the IEA-Wind Task 31 Wakebench to characterize the wind conditions in the Alaiz mountains near Pamplona, Spain, and to compare the numerical modeling of participants to the field measurements. In the original benchmark, four days have been selected (from 30/09/18 to 04/10/18) corresponding to inflow conditions from the north. The focus is on the diurnal variability of mesoscale drivers and their impact on microscale wind conditions relevant for turbine siting.

This experiment has been modeled using the mesoscale numerical weather prediction model WRF on a regional domain with a resolution of 27km and two nested domains with resolutions of 9 and 3km. The final area of interest is around the valley. The initialization and the forcing term for the WRF simulation come from ERA5 reanalysis datasets. No recirculation and blockage zone has been observed in the mesoscale WRF simulation hence the need to investigate meso-micro modeling. Details on the Meso-NH code and the setup used in this study are presented in the previous deliverable. Three simulations with different domains and meshes have been performed with Meso-NH. Table 4 summarises the properties of each case.

A first simulation with a coarser grid has covered the whole 4-day period of the Benchmark. This simulation is called alex17_001 in the figures below and is represented by the green color. A second one, with a finer resolution, only runs on the first day (September 30, 2018) (alex17_002, red color). A third one (alex17_003, purple color) has a fine resolution in the vertical direction, but some parts of the orography have been smoothed to help the solver.

Alex17_001 simulation has a horizontal grid resolution of 100m and a first vertical grid level of 10m. In contrast, the alex17_002 simulation has a 25m-horizontal resolution and a first vertical level at 5m in the area of interest. This second mesh is a lot heavier, so a grid nesting has been implemented to have a horizontal grid resolution of 25m in the area of interest (Domain 2) and a coarser one (50m) elsewhere (Domain 1). The downscaling flow uses the coarse mesh value (of Domain 1) as boundary conditions for

	Alex17_001	Alex17_002	Alex17_003
Domain size (km)	30x30x7	30x30x7	20x20x7
Nesting	no	yes	no
Horizontal resolution (m)	50	50 and 25	50
1st vertical level	10	5	5
Orography filter	no	no	yes
Time step (s)	0.75	0.75-0.125	0.25
Core-hours for 1h simulation	1400	max 40000	1150

Table 4: Parameters of the mesh and the domain for the three simulations.

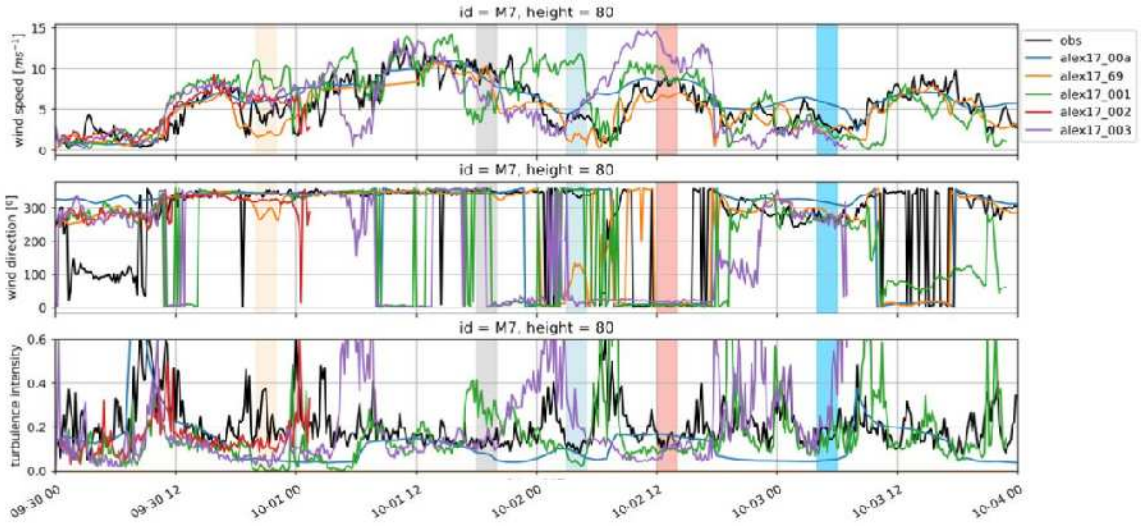


Figure 3: Temporal evolution of the wind speed, wind direction, and turbulence intensity at mast M7 and 80m-high.

the fine mesh domain (Domain 2). In contrast, the upscaling flow relaxes the coarse mesh fields towards the fine mesh spatial average on the coarse grid size in the overlapping area. The vertical mesh grid is the same between Domain 1 and Domain 2. This grid nesting procedure made it possible to run the first day of the simulation on 1764 processors. The simulation crashed after the first day: it is suspected that some slopes of the orography locally reach too high values in the domain.

Alex17_003 simulation has a horizontal grid resolution of 50m and a first grid level at 5m, and the results are plotted in purple. The orography has been smoothed locally over the slopes (along the horizontal directions); it is an automatic preprocessing step that can be activated in Meso-NH. The domain is smaller to limit the computational time. This simulation has run on 360 cores with a time step of 0.25s, leading to a CPU cost of 1150 core- hours for a one-hour flow simulation. Alex17_003 advanced significantly more than Alex17_002 but did not finish either. Results for wind velocity and turbulence intensity at mast M7 are shown in Figure 3 . The IFPEN team was not satisfied with the results on the coarser and smoothed meshed and could not manage to converge Meso-NH on a fine grid. Therefore, they decided not to include their results in the Benchmark.

Alya had no convergence issues independently of the mesh size used. Results were submitted to the Benchmark and appeared in [44]. Alya is much better suited for complex terrain cases than Meso-NH because it uses an unstructured grid. It has been used as a production tool by the wind energy company Iberdrola on cases of similar complexity for several years.

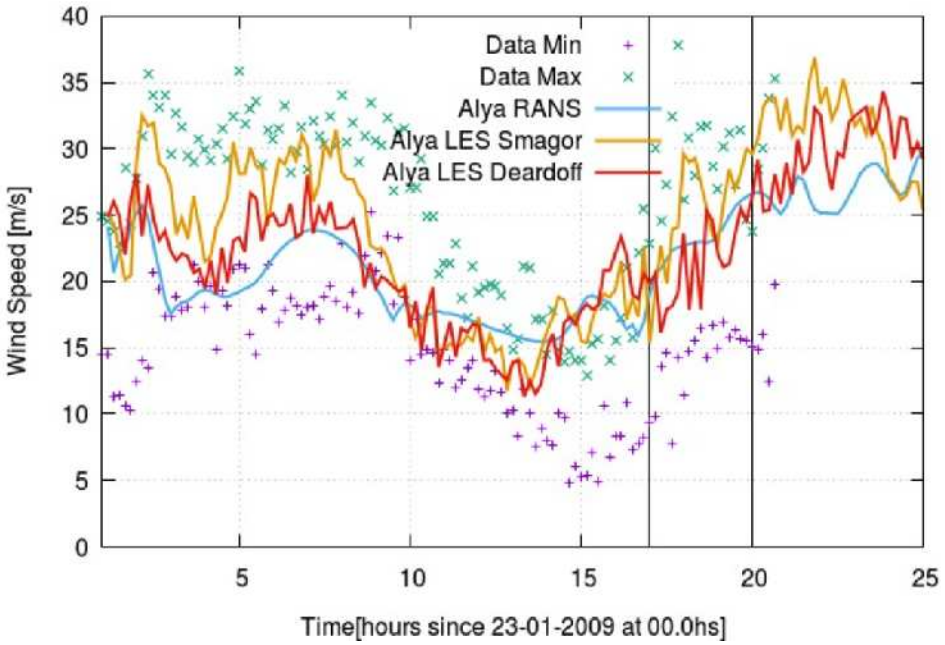


Figure 4: Simulated averaged wind speed at the collapsed WT and data measurements.

RANS and LES simulations have been submitted using the reference WRF simulation as input data following the 1D tendencies approach. This approach uses horizontally averaged (45x45 km) mesoscale pressure gradient and advection terms from WRF as volumetric forcing terms to drive the flow. The surface temperature drives the energy equation. It is inferred using the Monin-Obukhov similarity theory from the 2 m temperature extracted from WRF. Periodic boundary conditions are imposed laterally, wall modelling based on Monin-Obukhov is used at the terrain and symmetry boundary conditions are applied at the top of the domain.

The topography has been simulated over a 16.5x15 km area, using a horizontal resolution of 50 m in the central part for RANS and 35 m for LES. The topography surface mesh is extended to 33x43 km through a buffer mesh to reach a uniform height at the lateral boundaries. The domain height is 8 km. The elements closer to the ground have a vertical length of 2 m for the RANS simulations and 10 m for the LES simulation. The results with Alya and their comparison with several other codes can be found in [44].

T1.1.1-5 - Test the new model against a realistic set up.

We have applied the tendencies downscaling strategy to analyze extreme weather events that produced the collapse of wind turbines in a wind farm from Iberdrola. Results have been presented both at Wind Europe and Eccomas2022.

We show that the tendencies coupling has excellent potential for understanding transient events under extreme weather conditions in very complex terrain. The wind industry can use such simulations as a tool to enhance forensic analysis in cases of accidents.

We use URANS and LES closures coupled to mesoscale flow through the same coupling methodology. Simulation results using URANS and LES closures agree reasonably well with observations. However, the use of LES obtains results that are closer to observations. Fig. 4 shows the simulated averaged wind speed at the collapsed WT and data measurements. The wind speed values obtained using the Deardorff model agree better with the minimum and maximum data measurements range. The wind turbine collapse occurred between 17hs and 20hs, marked with two vertical black lines in Fig. 4. Fig. 5 shows TKE contour plots at 80m height above the entire wind farm at 18hs. It is observed that the collapsed WT (red dot) is exposed to higher TKE values than the other WTs (blue dots), explaining its unique collapse.

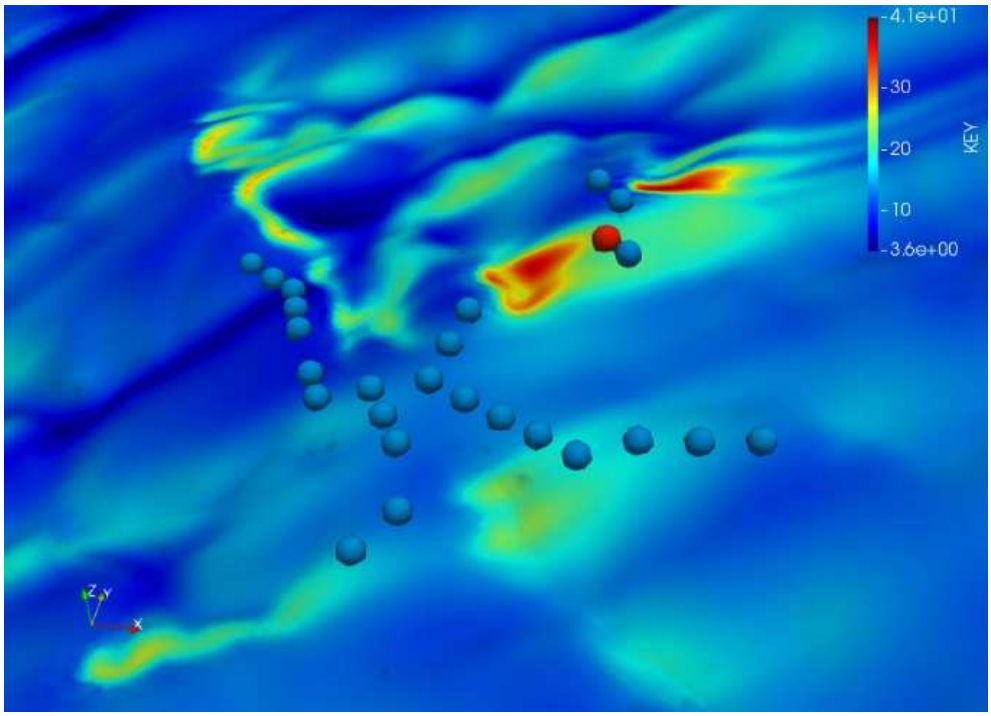


Figure 5: TKE contour plots at 80m height above the entire wind farm at 18hs.

3.2 Task T1.1.2 - Scientific results

The main task, T1.1.2, is subdivided into five subtasks.

T1.1.2-1	**Develop and test the sliding mesh approach for rotating blades in Alya using rigid blades**.
T1.1.2-2	**Adapt shall elements to turbine blades**.
T1.1.2-3	**Merge the results of the previous two subtasks**.
T1.1.2-4	**Compare Alya with FLOWer**.
T1.1.2-5	**Comparing models**.

T1.1.2-1 - Develop and test the sliding mesh approach for rotating blades in Alya using rigid blades

Task T1.1.2.1 was completed during the project's first half and was reported in deliverable D1.2. An implicit sliding mesh coupling for large-scale industrial applications has been developed that allows us to simulate problems where one part of the mesh is fixed and the other one is rotating. The wind turbine blades are enclosed in a cylinder that rotates inside the fixed mesh covering the whole wind farm.

T1.1.2-2 - Adapt shell elements to turbine blades

In this sub-task the continuum shell finite element from Alya code has been adapted to use isotropic and orthotropic materials which are common materials used in wind turbine blades. Continuum shell elements are 3D elements for the modeling of structures that are generally slender, with a shell-like response but continuum element topology. The proposed element type is a first-order solid shell finite element with

Enhanced Assumed Strains (EAS) and Assumed Natural Strain (ANS) methods to remedy the locking pathologies of solid brick elements. The details of this formulation can be found in [127]. For the case of wind turbine blades, two constitutive material law formulations are implemented in Alya code [1]: isotropic and orthotropic linear elastic material laws. The relation between the stress and strain field is carried out through the constitutive tensor, Kirchhoff-Saint Venant constitutive model $\mathbf{S} = \mathcal{C} : \mathbf{E}$. In order to verify and validate the new implemented material laws, two benchmark problems are carried out.

Annular plate subjected to lifting force

This example consists of a ring plate with a slit cut along the radial direction, as shown in Fig.6a. The material modeled here is isotropic linear elastic. The ring is clamped at one side of the slit cut, while the other end has a prescribed distributed load of $0.8N/m$. The ring internal radius is $6\,\mathrm{m}$, the external radius is $10\,\mathrm{m}$, the plate thickness is $0.03\,\mathrm{m}$. The material properties corresponding to the isotropic material law are the following: an elastic Young modulus of $21\,\mathrm{MPa}$ and a Poisson ratio of 0.0. The mesh consists of 10 element along the width direction and 80 through the circumferential direction and one element through the thickness.

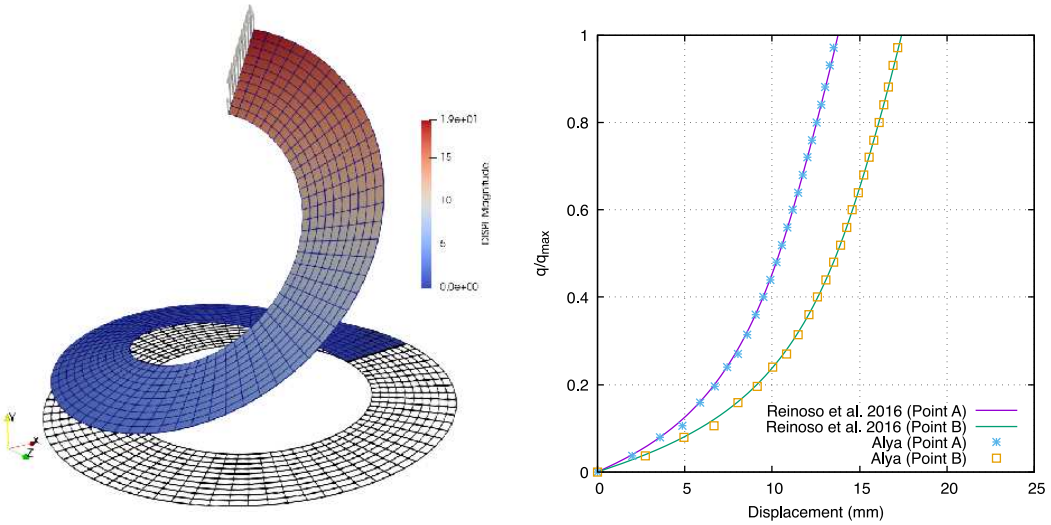


Figure 6: Slit annular plate subjected to a lifting force using isotropic material law.

Fig.6b depicts the normalized distributed load at the free tip versus the vertical displacement. We can see that a good agreement is achieved comparing the present formulation with the results from [127].

Multilayer composite with ply drops

This example demonstrates the applicability of this element type for the mechanical behaviour of composite structures (orthotropic material law). The stacking sequence of the plydrop is $[45/-45]_3$, with 0° reference along X axis. The plate length is 12 and the width is 5. It has a total of six layers at the thick end, which is fully clamped. The free thinner end has a transverse normal load along the free edge of $6 \times 10^9 h^3$. The location (in length direction) of the ply dropoffs are at $X = 4$ and $X = 8$ removing the top two layers for each dropoff. The unidirectional material properties for each layer are $E_{11} = 25 \times 10^9$, $E_{22} = E_{33} = 10 \times 10^9$, $\nu_{12} = \nu_{13} = \nu_{23} = 0.2$ and $G_{12} = G_{13} = G_{23} = 0.5 \times 10^9$. The ply thickness is $h = 0.1$. The model is meshed with 12 elements in the length direction, 6 elements along the width and each ply has one element through the thickness, see Fig. 7.

The resulting free-tip transverse displacement measured along the Z axis is 6.61 and according to Quoc and Tan [160] is 6.72 (relative error of 1.6%). This example demonstrates that using a coarse mesh the geometrically non-linear response is well captured.

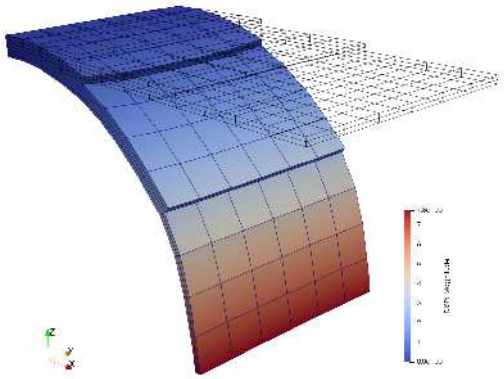
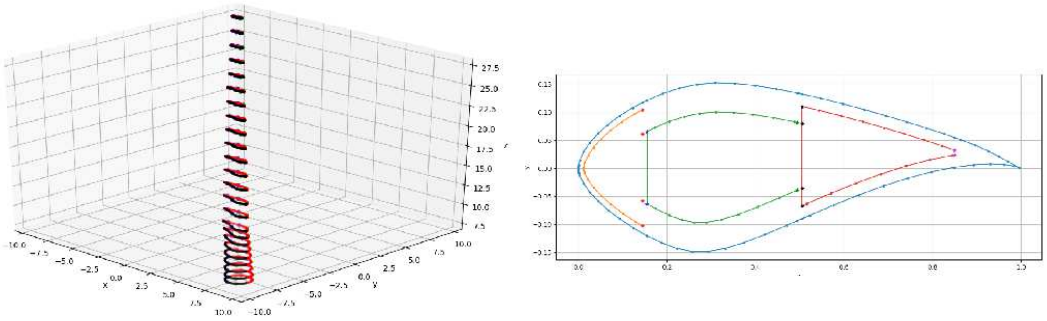


Figure 7: Deformed mesh of the multilayer composite cantilever plate using orthotropic material law.

Wind turbine blade

The selected case study corresponds to the WindPACT 1.5 MW horizontal-axis wind turbine blade (Malcolm and Hansen, 2002 [94]). This blade has been selected because it includes information of the internal structural parts for the structual analysis. This blade is representative of megawatt-class horizontal-axis wind turbines and its length is $33\,\mathrm{m}$ long. Due to the lack of geometry details and CAD models of real wind turbine blades in the literature, we have programmed a Python script to generate all the airfoils and internal parts using the available information in the literature. Fig. 8a and 8b show respectively the necessary points to define the surfaces of the wind turbine blade and also, the internal surfaces to define internal structural parts at each airfoil station. It is worth mentioning that this script serves as general tool to generate the CAD model of other blade designs including shear webs.



(a) (b)

Figure 8: (a) Points for the generation of the CAD model. (b) Internal points for the definition of thicknesses and shear webs.

The blade consists of 24 airfoils and two shear webs that start at station No. 6 to station No. 24. The blade has 5 different materials which their material properties are summarized in 6. The composite stacks and their thicknesses are different depending on the blade region. The reader is referred to [94] for more details on the thicknesses at the blade root and shell.

With regards to the generation of the mesh, the wind turbine blade model has been fully discretized with hexahedrons and prisms using ANSA pre-processor [2]. The generation of this mesh with 3D elements has been very complex and challenging due the wide variety of materials, different thickness sections and the intersection of the internal shear webs with the blade skin, see Fig.9. The resulted mesh has a total of

Material	$E_{xx}(GPa)$	$E_{yy}(GPa)$	$E_{xy}(GPa)$	$\nu_{xy}(GPa)$	$\rho(kg/m^3)$
Gel coat	3.44	3.44	1.38	0.3	1230
Random material	9.65	9.65	3.86	0.3	1670
CDB340 triaxial fabric	8.97	8.97	4.97	0.39	1700
Balsa	2.07	2.07	0.14	0.22	144
Spar cap mixture	8.35	8.35	4.7	0.37	1700

Table 6: Materials used in the wind turbine blade [94].

$4\,303\,554$ elements which is approximately 14.6M dof.

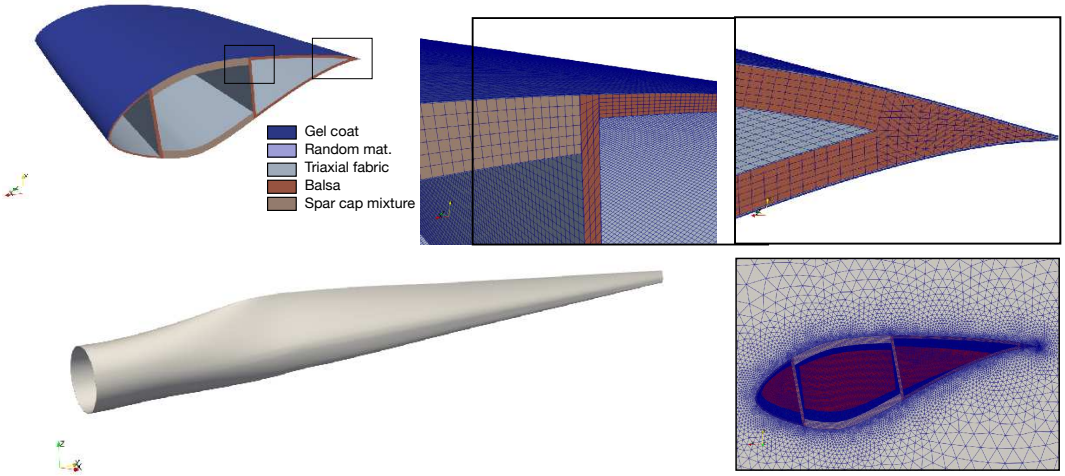


Figure 9: Mesh details of the wind turbine blade.

Difficulty on the use of continnum shell elements with the proposed geometry

One of the main requirements on the use of continuum shell elements is that this element technology requires to be oriented according to what is called the element normal or thickness direction. The stacking of the elements and the local material coordinate systems (when modeling the fibers) is really important to capture the through-thickness response for composite laminate structures. This stacking direction is based on the nodal connectivity of the element, see Fig. 10.

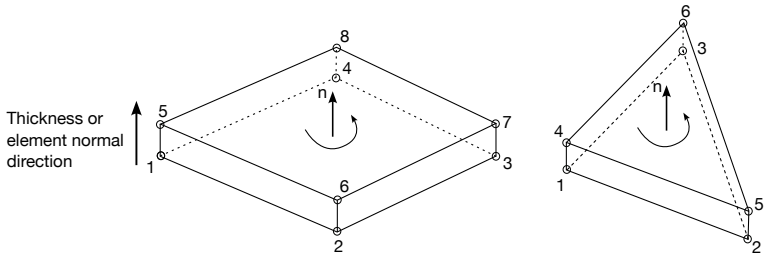


Figure 10: Thickness or element normal direction in continuum shell elements.

Due to the difficulties on orienting the elements (proper element normal direction) encountered with ANSA pre-processor for the proposed geometry, all the numerical results from now on has been conducted using solid brick element types. With the lesson learned, two approaches have been concluded that would definitively simplify the problems on the mesh generations for future analysis: (i) the use of continuum shell

elements with the single layer approach (homogenization of lamina properties) together with simplified mesh with few elements through the skin thickness or (ii) the use of conventional 3D shell elements with six degrees of freedom for each node.

Structural analysis

A preliminary structural analysis has been conducted on the wind turbine blade to evaluate maximum load before buckling. The blade has been clamped at the root and a prescribed displacement of 10m is applied at the end tip. Fig. 11 shows the deformed shape of the wind turbine blade using the previous prescribed displacement. The resulted load for a 10m displacement is 80kN, which provokes the buckling at the root of the blade.

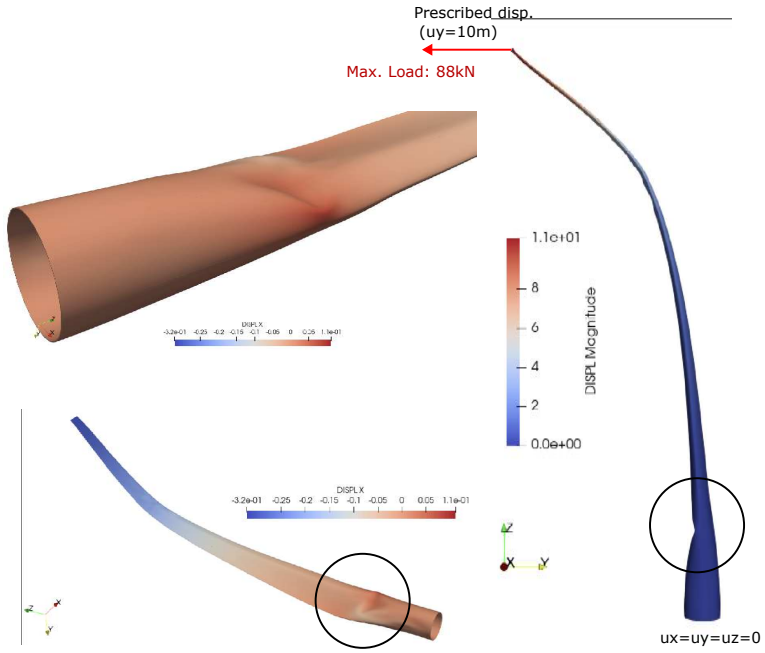


Figure 11: Wind turbine blade deflection with a prescribed displacement of 10m.

T1.1.2-3 - Merge the results of the previous two subtasks

This sub-task consists of performing Fluid-Structure Interaction (FSI) on the wind turbine blade and then apply the sliding mesh method developed on T1.1.2-1 for the full rotor including nacelle. The first part has been achieved successfully, however the second part was not conducted due to the difficulties encountered on the mesh discretisation of the structure with 3D elements and the large computational cost that would be required by using the modeling methods stated in T1.1.2-2. At the end of this section we justify the limitations that we have encountered (lesson learned) and we propose a solution that could be applied for future analysis.

Wind tunnel model and model setup

For the FSI simulation a wind tunnel has been created for the evaluation of different flow conditions and pitch angles of the blade, as shown in Fig. 12. This wind tunnel is discretized with an hybrid mesh of tetrahedrons and prisms. The total number of elements is $6\,446\,653$. The blade is located at the end of the wind tunnel and it is fully clamped at its root end. A free-stream air is set at the inflow face of the wind tunnel as well as the top, bottom, right and left faces of the tunnel. The air is considered incompressible and due to this, the fluid density is assumed constant with a value of $1.225kg/m^2$. The viscosity is also considered to be constant at $1.7894 \times 10^{-5}kg/ms^{-1}$. The flow is solved with a Wall Model Large Eddy Simulation (WMLES) framework with an explicit SGS (typically Vreman).

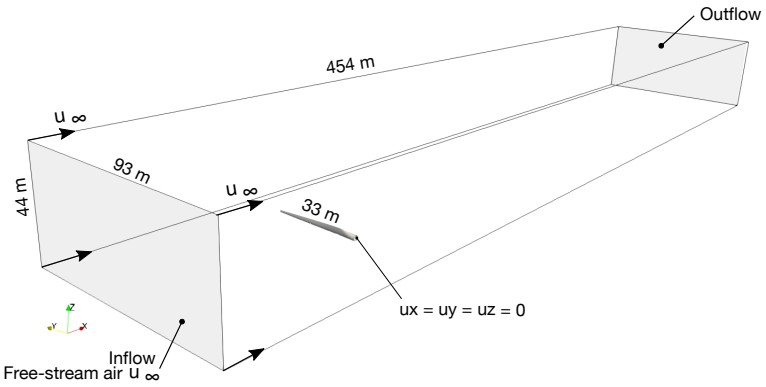
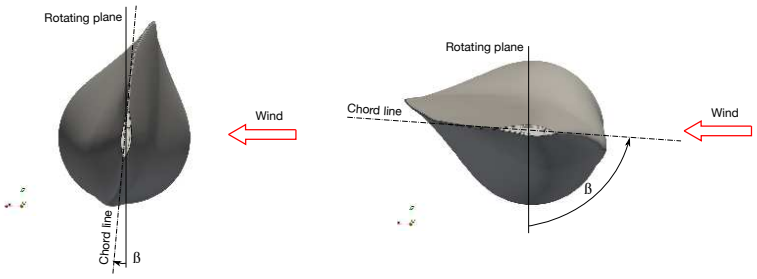


Figure 12: Wind tunnel modelled for the turbine blade.

Aeroelasticity test cases and results

Several aeroelasticity simulations have been carried out for the wind turbine blade. These case studies consider two pitch angles of -2.6º and +87.4 (see Fig. 13), each of them with different wind speed velocities, 8m/s, 16 m/s and 24 m/s.



(a) (b)

Figure 13: Different pitch angles for the wind turbine blade.

The results obtained are depicted in Fig 14 and Fig. 15 for the -2.6º and 87.4º respectively. These results, show that for small pitch angles -2.6º, large deformations are predicted, which are around 1m. When the pitch angle is +87.4º (a typical angle in service conditions) the maximum deflection predicted is 0.2m with a wind speed of 24m/s.

Parallel performance

The parallel performance of the fluid-structure interaction simulation has also been analysed. The flow solver uses a fractional time step for the time integration, which is very cheap in terms of computational cost. On the other hand, the solid solver uses and implicit time integration scheme and the algebraic system is solved with an iterative solver. We have seen that the major computational cost is derived from the algebraic solver used for the solid domain. Based on that, we have investigated the use of a direct/iterative solvers and also the application of Dynamic Load Balancing (DLB) method [55] to better manage the computational resources for the Gauss-Seidel method used for the FSI simulation.

The MUMPS direct solver [148] has been implemented in Alya code [1] and tested for the resolution of the FSI problem of the wind turbine blade. According to different tests that we have carried out, we have concluded that the use of direct solvers for large-scale and multiphysics problems (FSI) is not competitive respect to iterative solvers such as Conjugate Gradient or GMRES. MUMPS only provides a significant advantage if the factorization governs the solution of the problem. Other strategies such as the use of MUMPS direct solver in the RAS preconditioner is a promising method for further computational analysis.

The aerolastic simulation uses a Gauss-Seidel method for the coupling between the solid and the

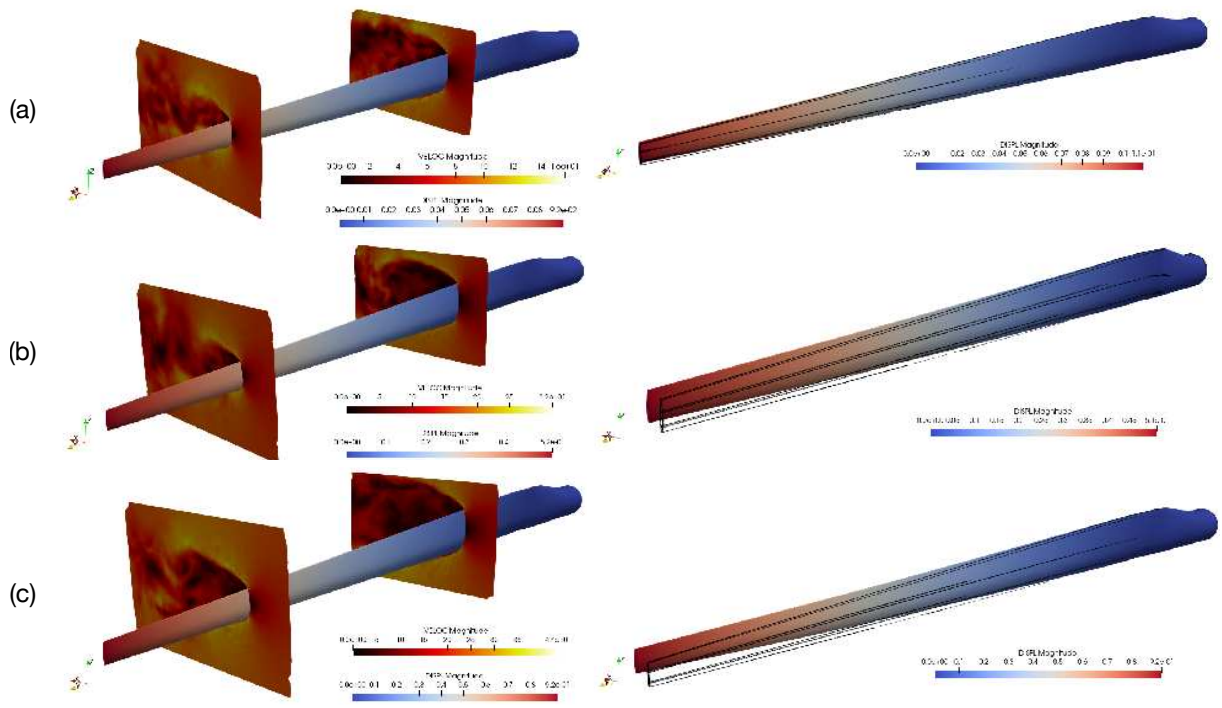


Figure 14: Blade with a pitch angle of -2.6º. (a) 8m/s (b) 16 m/s and (c) 24 m/s. Deflections are represented with a scale factor of 1.

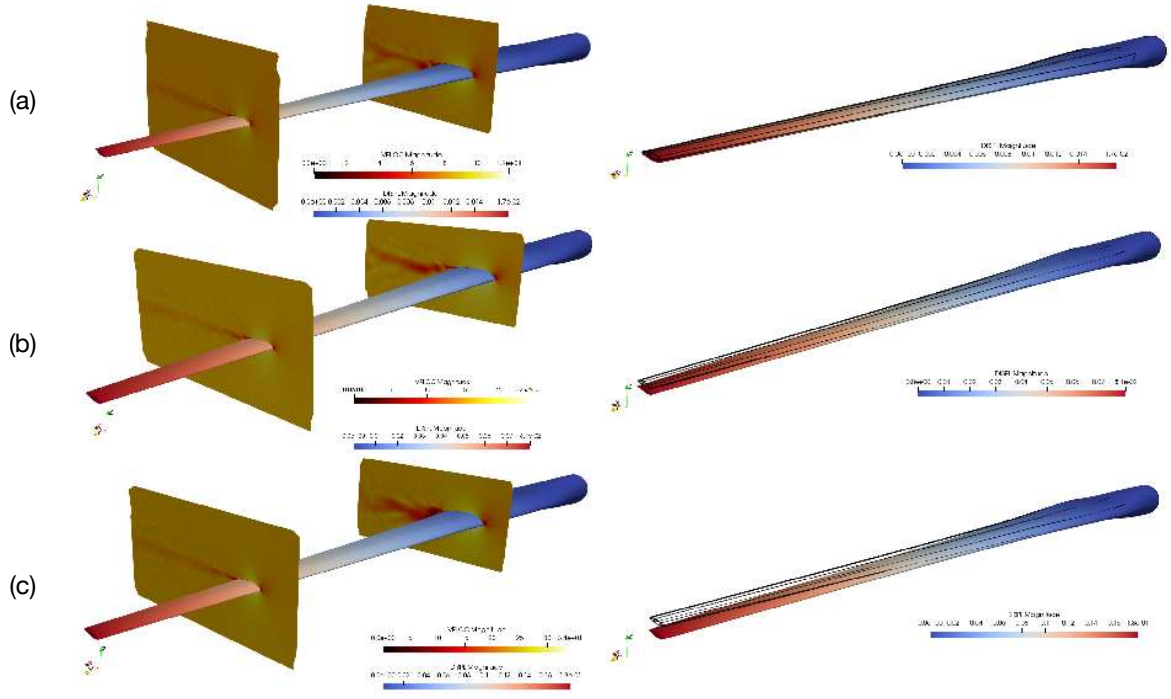
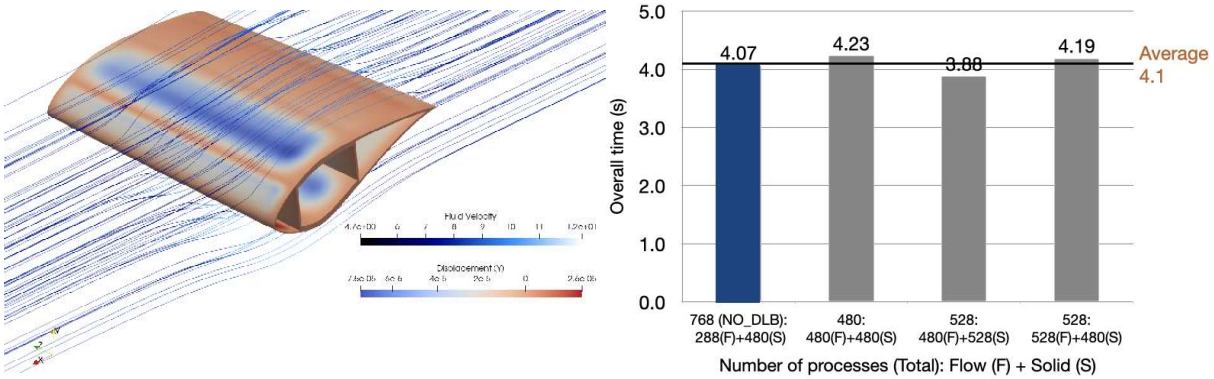


Figure 15: Blade with a pitch angle of +87.4º. (a) 8m/s (b) 16 m/s and (c) 24 m/s. Deflections are represented with a scale factor of 2.

fluid domains. This method is well-known to have better convergence of the solution and less numerical instabilities than the Jacobi algorithm. However, it works in a "serial manner", as one of the physics runs,

the cores belonging the other physics are idle. To circumvent this problem, one possibility is to overload the cores of both physics using Dynamic Load Balancing (DLB). This functionality has been applied to the FSI simulation of the wind turbine blade, see Fig. 16a. We have demonstrated that we can obtain the same wall-clock simulation time per time step overloading the cores with DLB (using a total 528 cores) or using 768 cores without DLB, as shown in Fig. 16b.



(a) (b)

Figure 16: Preliminary analysis blade tip. (a) Flow around blade tip (b) Overall computational cost with the use Dynamic Load balancing (DLB)

T1.1.2-4 - Compare Alya with FLOWer

In EoCoE-II, we proposed to set up a collaboration with Professor Thorsten Lutz, probably the top European expert in full rotor simulations. Edouard Audit, the project coordinator, said it was impossible to include Thorsten as a partner in the project. Instead, Edouard proposed that we could allocate some money to subcontract him. Unfortunately, during the development of EoCoE-II subcontracting conditions at Stuttgart University, where Thorsten Lutz works, changed. Stuttgart University started charging a 50% fee on subcontracts. Under those conditions, the money we had reserved was insufficient for Professor Thorsten Lutz, and he declined the offer. We reported the situation to the European Commision officer in 2021. The money reserved to subcontract Professor Thorsten Lutz was reallocated to pay the salary of Herbert Owen during the 6-month extension of the project. Although he has been working at BSC since 2009, he still has a temporary position, and his salary depends on money from projects. Therefore, it was essential to find some money to pay his salary during the project extension.

T1.1.2-5 - Comparing models

Although the collaboration with Professor Thorsten Lutz was not possible, the BSC team made significant progress in full rotor simulations that will be described in the current subsection. The sliding mesh approach described in task T1.1.2-1 is used for the simulations in this section. The wind turbine blades are enclosed in a cylinder that rotates inside the fixed mesh covering the whole wind farm. During the first half of the project, we concentrated on the NREL Phase VI wind turbine without including the nacelle and mast. Simulations used a hybrid mesh formed by 50 million tetrahedral, pyramidal, and prismatic elements. A case with twice smaller elements leading to a total of 400 million elements was also performed using an automatic mesh subdivision algorithm available in Alya.

To test the suitability of the proposed approach for more realistic problems, we switched to the WindPACT 1.5 MW horizontal-axis wind turbine(Malcolm and Hansen, 2002 [94]), which is a representative of megawatt-class horizontal-axis wind turbines. This is the same turbine used in task T1.1.2-2 for the solid mechanics problem. We have included not only the rotating blades but also the nacelle and mast for this

wind turbine. This makes meshing of the geometry much more challenging. The main difficulty arises from the small distance between the rotating blades and the mast. The mesh around the blades must be contained in the rotating part domain, but the mast belongs to the fixed part. The successful simulation involved several iterations between mesh generation with ANSA and the actual simulation with Alya to obtain a correct mesh. However, the meshing learning process with this wind turbine should also apply to other wind turbines.

A mesh formed by linear tetrahedra, pyramids, and prisms, with a total of 168 million elements and 30M nodes, was used. The simulation ran for a total of 15.74 seconds with an average time step of 3.148E-5 s and 500000 steps. Fifty Marenostrum IV nodes (Intel Skylake) with 48 cores each were used for the simulation. The case ran for a total of 18 days. Thus, the total computational cost was 1036800 CPU hours.

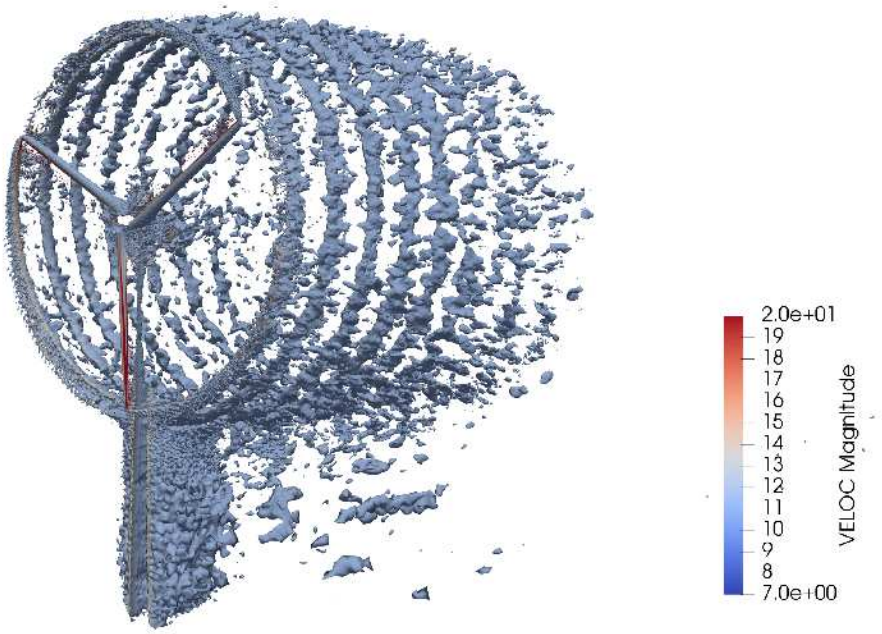


Figure 17: Constant axial velocity iso-surface colored with velocity magnitude for the WindPACT turbine.

Figure 17 uses an iso-surface of constant axial velocity colored with velocity magnitude to display the complex wake structure behind the rotating wind turbine blades. Since the mast is also included, its wake can also be observed. Including the mast generates a more interesting simulation that captures the interaction between the mast and the blade passing in front of it. The perturbation of the flow around the blade when it passes in front of the mast can lead to fatigue effects.

Figure 18 shows the velocity magnitude over a vertical cut through the center of the wind turbine. The interface between the fixed and rotating meshes is marked in orange. We observe that the sliding mesh algorithm performs satisfactorily since the perturbation to the velocity field is negligible at the interface.

To have some idea about the computational cost comparison of both approaches as implemented in Alya and WaLBerla we can calculate the solution time per time step and degree of freedom for both approaches.

As mentioned earlier, the Alya run performed 500000 times steps in 18 days, leading to an average CPU time of 3.11s. The mesh has 168 M elements. The case is run in 2400 cores. The average load per core is thus 70000 elements per core. The computational cost is then 44428 nanoseconds per element. It is interesting to note that the average time step for this case is 3.148E-5 s.

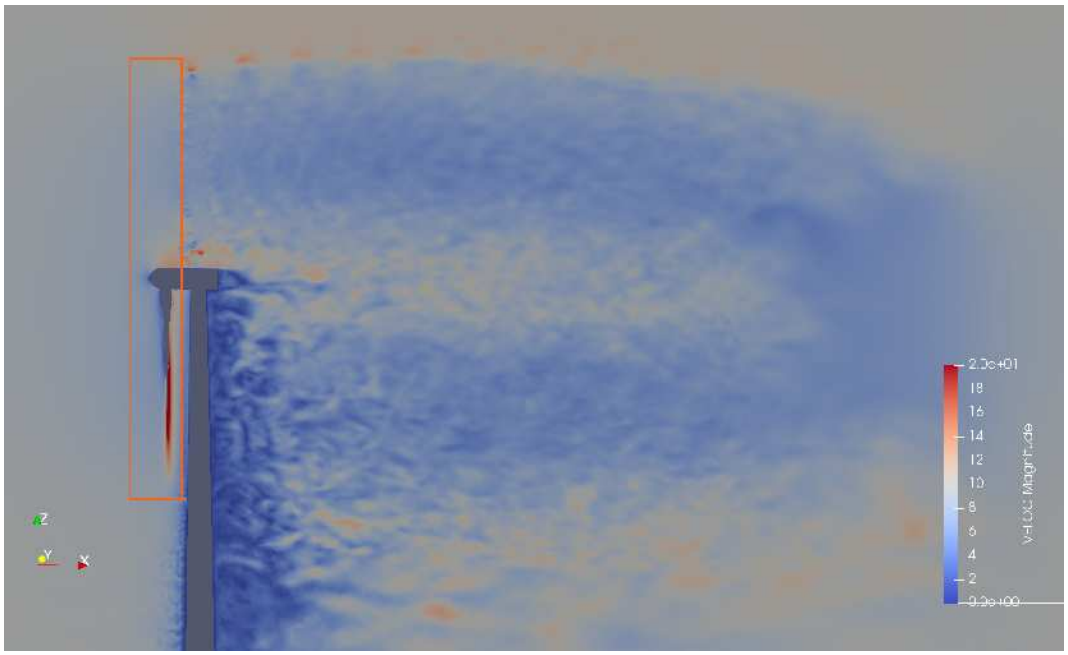


Figure 18: Velocity magnitude for a vertical plane passing through the center of the WindPACT turbine.

Node	GPU	CPU	Cores	waLBerla GPU	waLBerla CPU
2	8	4	256	1351	9652
5	20	10	640	852	5368
10	40	20	1280	864	2734
15	60	30	1920		1597
20	80	40	2560		1390

Table 7: Simulation times in seconds for the strong scaling runs for waLBerla.

For WaLBerla the results presented in Table 7 from [138] are used to estimate costs. The case with 2560 cores takes 1390s of computational time on the Topaze supercomputer with AMD Milan@2.45GHz (AVX2) CPUs with 64 cores per CPU. The grid has 163 840 000 cells. The timing corresponds to 1200s of real time with a time step of 0.026s leading to a total of 46135 time steps. Thus, the computational time per time step is 0.030 second per time step. Since the average number of cells per core is 64000, we get the computational cost is 470 nanoseconds per cell. We can compare this value to the 44428 nanoseconds per element obtained with Alya to say that WaLBerla is 95 times faster per timestep. Moreover, it is important to remark that the time step size use used by WaLBerla is 0.026s while the one used by Alya, 3.148E-5 s, is 825 times smaller. This means that if one wants to compare the computational cost of both approaches to simulate a certain amount of time using the same amount of cells/elements per core WaLBerla would not only be 95 times faster per time step, but it would need 825 times fewer time steps. Leading to an advantage of more than 78000 times.

The main conclusion from this comparison is that the wide difference in computational cost is clearly related to comparing methodologies that are hardly comparable. Moreover, in our comparison, not only does the method used to model the turbine differ but also the numerical implementations in Alya and WaLBerla are significantly different. Alya solves the incompressible Navier Stokes equations using and implicit time discretization for the pressure. WaLBerla does not rely on the Navier-Stokes equations, but rather on the discrete Boltzman approach. Moreover, Alya uses and unstructured grid approach with great flexibility in grid size while WaLBerla only allows to use isotropic cartesian grids with local refinement.

The high difference in time step size, related to the high difference in grid resolution required for both methodologies is something that also make the comparison very difficult.

The computational requirements for both approaches indicate that using the Actuator Line method it should be possible to run production full wind farm simulations in the medium to short term. Fully resolved simulations of a wind farm will only be possible in the medium to long term. They will probably require the use of a full exascale supercomputer. Production level whole wind farms simulations we will probably need to wait until the zetascale era.

3.3 Code demonstrator

Alya

Alya is a high-performance computational mechanics code that solves complex coupled multi-physics BSC problems, mostly coming from the engineering realm. As mentioned earlier, Alya uses a Fractional Step scheme to solve the incompressible Navier Stokes equations. A Runge Kutta explicit time discretization is used for temporal discretization. Thus, from a computational point of view, the two main kernels are the assembly of the right-hand side vector for the momentum equation and the solution of a pressure Poisson equation for the pressure.

Alya is a relatively mature code that is used in several CoEs. Moreover, it is part of the Unified European Applications Benchmark Suite for CPU and GPU. It has been tested on most European Supercomputers with scalability results up to more than 100.000 cores. When EoCoE-II started, Alya already had a basic GPU implementation, and progress was made towards a co-execution approach that took advantage of both CPU and GPU. However, initially, it was not clear how efficient the GPU implementation was. Thanks to the collaboration with the FAU node level optimization team, we discovered that there was still a vast potential for improvement. We concentrated on obtaining an optimal GPU implementation for the most consuming kernel, the volumetric element assembly. A specialized version of the kernel was sped up 80 times, bringing the performance close to 50% of the maximum peak floating-point performance on an A100 Nvidia GPU. This is an outstanding result compared to the GPU progress in other CoEs in which Alya participates. It has been an inflection point, and now it is clear that Alya can be much more energy-efficient on the GPU than on the CPU. As stated in deliverable D2.3, Alya assembly is 8.5 times more energy efficient on the GPU than on the CPU.

Scientific simulations towards exascale

Alya is one of the two CFD codes in the Unified European Applications Benchmark Suite for CPU and GPU. It has been tested on most European Supercomputers. Also in the US and Japan. It has performed production runs (24 hours) on 100.000 cores of the Marenostrum IV Supercomputer.a

`https://alya.gitlab.bsc.es/alya/open-alya/-/releases/eocoe2-v1.0.0-pre`
EoCoE II demonstrator v1.0.0-pre

3.4 Summary

Significant advances have been made in the simulation of the wind over complex terrain using LES that have positioned Alya among the best available codes, as shown in the participation in several Benchmarks. Alya is now positioned as the European alternative to US-developed open source exascale codes within Exawind and A2E. All the proposed activities in this task have been fulfilled. Moreover, we have advanced significantly in the mesoscale coupling, although it was not included in the original proposal, and our implementation is clearly among the best ones available. Alya can now include thermal coupling to treat stable and unstable conditions with the possibility of simulating diurnal cycles using the tendencies approach for mesoscale coupling. Alya's LES simulation can include: Coriolis forces; Modeling of the forest using a

canopy model; Inclusion of wind turbines with an actuator disc model. Participating in several Benchmarks has enhanced Alya's visibility within the EU wind community.

Alya's capabilities for the fully resolved simulation of wind turbines have improved significantly during EoCoE-II. The sliding mesh approach used to incorporate the rotation of the wind turbine blades involves the coupling between a fixed grid for most of the domain and a rotating grid around the turbine blades. The coupling algorithm was relatively new at the beginning of EoCoE-II, and it soon became apparent that it required significant rewriting since it had become too complex and error-prone. Moreover, one of the initial developers, Juan Carlos Cajas, left BSC and returned to Mexico, where he was offered a much more stable position. Guillaume Houzeaux, Alya's leading developer, took care of the new implementation, which we had not accounted for at the beginning of EoCoE-II. As happens with every new implementation, it has involved significant debugging and testing, but not we have something that is better structured and easier to develop. We have simulated a 1.5 MWatt wind turbine including mast and nacelle. During the ExaWind US project review, which is open to the public, one of the main criticisms from the reviewers was that their full rotor simulations did not include the mast. The interaction when the blade passes close to the mast is important in analyzing vibrations and fatigue. Our achievements are significant considering that the budget for Wind within EoCoE is less than 10% of the ExaWind budget. Aeroelastic simulations have also been tackled during EoCoE-II. In this case, fluid-structure interaction (FSI) simulations using wall model LES have been conducted on the whole blade of the aforementioned wind turbine. The aeroelastic case studies consider two pitch angles of -2.6º and +87.4º, each with different wind speed velocities, 8, 16, and 24 m/s. The results show that large deformations are predicted for the small pitch angle. When the pitch angle is +87.4º (a typical angle in service conditions), the maximum deflection predicted is 0.2m with a wind speed of 24m/s. In terms of computational performance, we have investigated using a direct solver (MUMPS) for the solid domain, which is the most expensive part of the multi-code coupling. The results showed that MUMPS only provides a significant advantage if the factorization governs the solution of the problem. As explained earlier, the comparison with Professor Thorsten Lutz could not be performed due to administrative issues. A comparison between the fully resolved simulation with Alya and an Actuator Line model in WaLBerla was performed.

4 Meteorology for Energy (T1.2)

T1.2 comprises of three partners with three main tasks.

T1.2.1	Continuous probabilistic short-term prediction of optical thickness and wind.
T1.2.2	Wind and solar power calculation for meteorological ensembles.
T1.2.3	Calibration of ensemble prediction of wind and cloud optical thickness.

Task 1.2.1 involved developements of the flagship codes ESIAS-Met and ESIAS-Chem and simulations for the scientific use cases. This included demonstration on supercomputers, including HAWK, JUWELS, JURECA, and EAGLE, with at least 512 members and stochastic schemes or multi-physics for probabilistic simulations.The multi-physics simulation used ESIAS as a framework for a very large-scale sensitivity study to establish the best of thousands of possible physics combinations for the model prediction of cloud fraction. ESIAS-Chem was tested using Saharan Dust Event simulations, showing that the assimulation of meteorological data actually decreases the ensemble spread and that stochastic input data (e.g. GEFS) may be preferable for achieving e.g. better horizontal variance in particulates in the atmosphere. The data assimilation techniques and particle filter could be performed by an online simulation procedure with the MELISSA middle-ware supporting ESIAS-Met (WRF3.7.1). The particle filter can at present increase the accuracy of the simulation by filtering outliers, though this is observed to decrease the ensemble spread.

Task 1.2.2 applied the ultra-large meteorological ensembles from T.1.2.1 to the renewables forecasting use case. In the first part of this task, the wind and solar forecasting systems were calibrated and optimized to electric meter data for the ESIAS-Met weather model based on an entire year of simulations. While the

ensemble irradiation data was quite biased overall, a novel calibration method could successfully improve the point forecasts. This however came at the cost of reducing the overall resolution and probablistic quality of the ensemble, such that its benefit depends on the use case. The second part of task 1.2.2 used satellite data to attempt to validate ensemble members in real time for short-term forecasting and feed-back to the particle filter, both by their clouds and cloud motion. It could be observed that such large ensemble groups tend to form a few members with accurate representations of highly variable clouds, a great advantage over ordinary ensemble systems. As for the motion, good agreement was seen between the simulations and optical flow algorithm, but the synoptic motion of the ensemble members were too similar to one another and too uncorrelated to their turbine height winds to benefit the wind power forecast.

Finally, Task 1.2.3 employed non-parametric techniques and CNN neural networks for the probabilistic calibration of point forecasts (quantiles) of the ESIAS wind speeds and irradiance. The results are compared between WRF and the French PEARP model. While the raw forecasts were biased and underdispersed, this could be improved for the solar irradiation and corrected to a flat ranked histogram for the turbine height wind speeds.

4.1 Task 1.2.1: Scientific results

Task 1.2.1 is divided into 3 subtasks:

T1.2.1-1	**Probabilistic wind prediction**. For the meteorological simulations the Weather Research and Forecast (WRF) model is adopted to predict winds typically at rotor hub heights, commonly taken at 100 m height.
T1.2.1-2	**Probabilistic cloud optical thickness (COT) prediction**. In addition, the radiative impact on solar energy by aerosol-induced turbidity (aerosol optical thickness, AOT) on solar energy production will be forecasted by the EURopean Air pollution Dispersion-Inverse Model (EURAD-IM).
T1.2.1-3	**Joint simulations**. In both cases, non-Gaussian data assimilation by particle filter and smoother methods will be applied, with remote sensing data processing of satellite images as a prominent data source, combined with big data analytics, based on suitably selected metrics. A prominent objective includes a middleware based flexible and non-synchronous hyper-ensemble operation by MELISSA middleware. This model will also be integrated in stochastic mode by ESIAS-chem and operated in hyper-ensemble mode. The result will be evolving pdfs as an approximation to the corresponding Fokker-Planck equation.

T1.2.1-1 - Probabilistic wind prediction

For the project mid-term (D1.2) we reported on a sensitivity analysis of the effect of the model physics options on weather variables to optimize the simulation of wind and clouds. There were six test cases using 48 hours of day-ahead simulation. The wind results were compared with observatins (in the x and y directions). The observation data, NCEP ADP Global Observation [111], was acquired from the NCAR Research Data Archive. The stations are shown in the Figure 19 as red dots.

Figure 20 and Figure 21 show the Taylor diagrams of wind in the x and y directions, respectively. The diagrams show the correlation between the simulation and the observation and the relative standard deviation between simulations and observations to show whether the dynamics of the simulations are similar to those of the observations. This diagram is suitable for measuring which simulation best fits the observation, especially for sensitivity analysis.

Both figures shows the simulation cases with different options of microphysics and planetary boundary layer (PBL) physics by different colors and marker shapes, respectively. Cases 20150413 and Case 2015-0515 show a large variability in wind speed during the springtime. The correlations have more variability

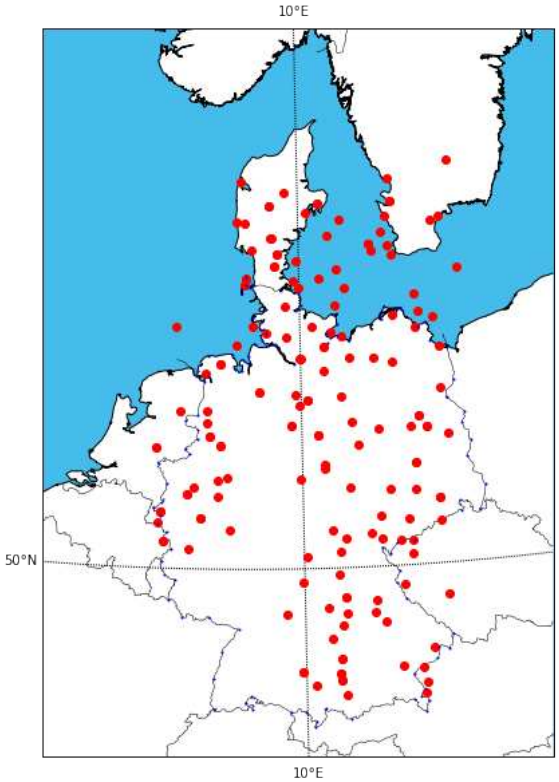


Figure 19: The station locations of observed wind. The red dots indicate the location using in the sensitivity analysis results

than in the cases 20150617 and 20150823. For the PBL option, the QNSE (Q) is the least dynamic options. The MYNN2 (MN2) and ACM2 (A2) are the closest results to the observation (marked as 1.0 in the standard deviation). Therefore, to simulate the wind condition in the European domain, we recommand the choice of MYNN2 and ACM2 for the best results.

While applying the same PBL physics, the Taylor diagrams show that using WSM6 and Goddard microphysics can have greater variability than the other microphysics, regardless of accuracy. In report D1.2 we showed that WSM6 and Goddard microphysics are also better choices for simulating the cloud cover condition when comparing to the observed cloud fraction.

To investigate the possible effect of multi-physics simulation on the wind field, we use an Saharan dust event from 2018-10-28 to 2018-10-29 to investigate the simulation performance and compare the results to observation. The observed wind speed components are based on data from the NCEP ADP Global Observation.

Figure 22 shows quiver plots for wind directions and their magnitudes at the stations in Italy, which are highly affected by the Saharan dust event. In the plots, the simulated wind directions to not differ much. This might be due to the factor that the input of simulation inputs are the same, especially the bounday layer conditions and therefore the same input pressure field and boundary conditions keep the wind field very similar despite different physics setups. This can also explain the fast transportation of Saharan dust during the simulation period in section 4.1

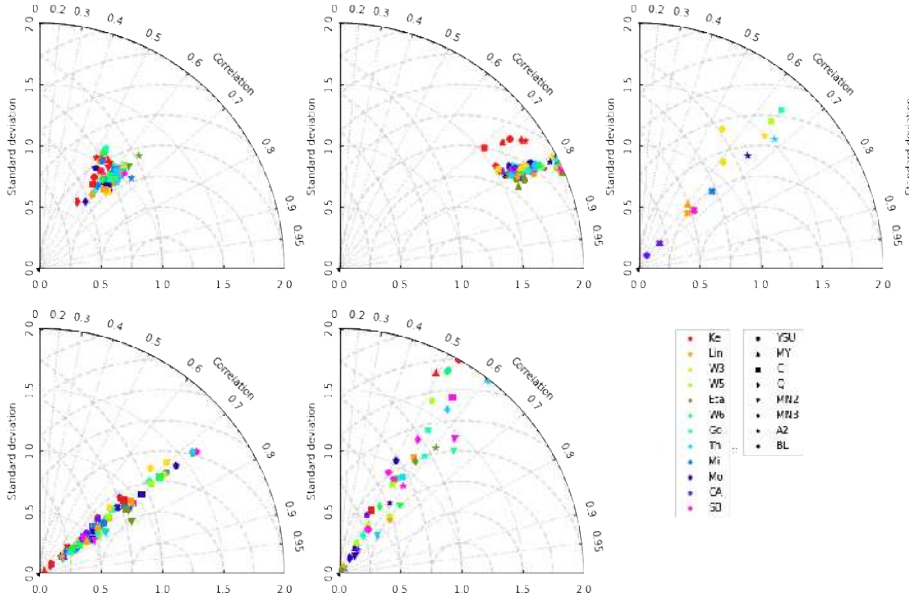


Figure 20: Taylor diagram of wind speed in x direction for case 2015-04-13, 2015-05-15, 2015-06-17, 2015-08-23, and 2015-09-21. The different color codes and marker shape represent the usage of different microphysics and planetary boundary layer physics, respectively.

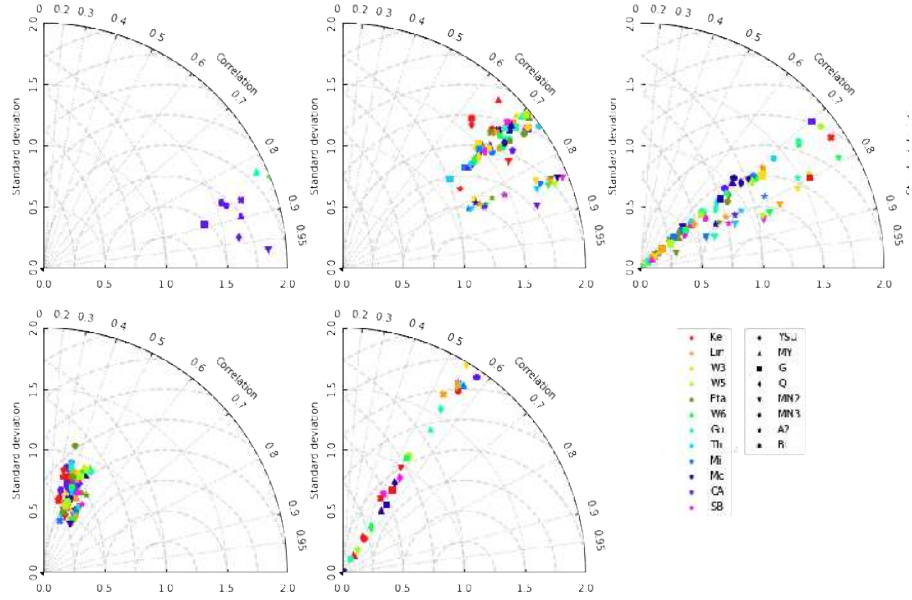


Figure 21: Taylor diagram of wind speed in y direction for case 2015-04-13, 2015-05-15, 2015-06-17, 2015-08-23, and 2015-09-21. The different color codes and marker shape represent the usage of different microphysics and planetary boundary layer physics, respectively.

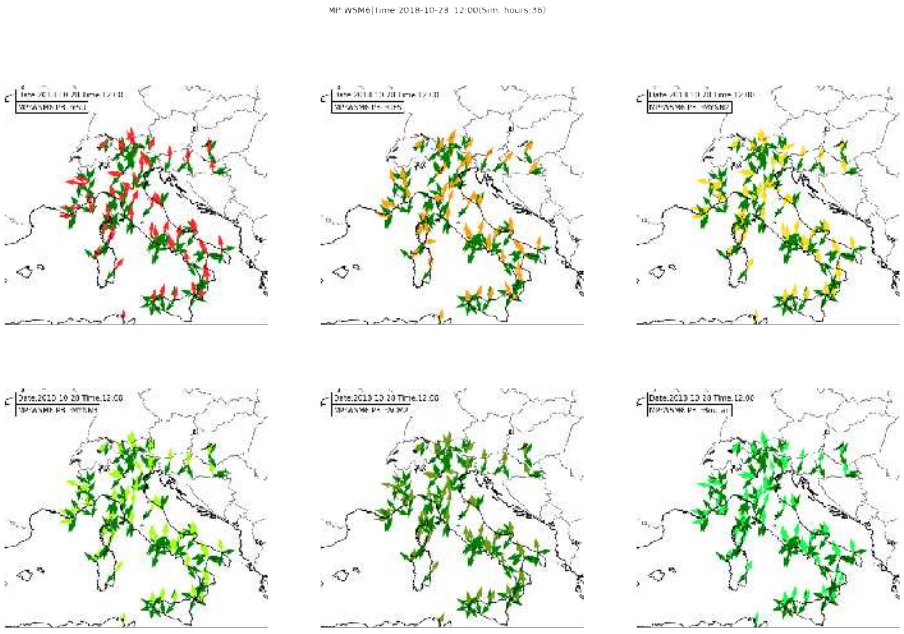


Figure 22: Taylor diagram of wind speed in y direction for case 2015-04-13, 2015-05-15, 2015-06-17, 2015-08-23, and 2015-09-21. The different color codes and marker shape represent the usage of different microphysics and planetary boundary layer physics, respectively.

T1.2.1-2 - Probabilistic cloud optical thickness prediction

Deterministic weather predictions have limited accuracy [87]. Accurate weather forecasting by deterministic simulation requires very detailed work in model calibration to optimize the choice of physics, numerous input parameters, and the spatial resolution of the modelling domain [143]. These works also require deep knowledge about the various physics and parameterizations, as well as great computational expense. Convection-permitting simulations are possible at an increase to computational demands by a factor of four to ten time non-convective simulations[113, 11], but the resulting high-resolution details cannot be deterministically accurate anyway due to nonlinearity and imperfect initial conditions. Ensemble simulations thus give realistic solutions as a basis for a probablistic description and the individual members need not be as computationally expensive as the most accurate possible deterministic model.

In this section we demonstrate the readiness of ESIAS-Met for ultra-large ensemble simulations in the study of a Saharan dust event between 2018-10-28 and 2018-10-29. The source of the Saharan dust comes from the lee side of Altas Mountains [158]. ESIAS-met, as the meteorology component of the large ensemble system, can produce large ensemble simulation up to 1,024 members. Here, we perform a large simulation with 32 ensemble members $\times$ 6 physics settings, which generates a total of 192 members. The results are primarily validated against the clear-or-cloudy cloud mask observation MSGCLMK from EUMETSAT. The simulation starts from 2018-10-27 for a spin-up of at least 24 hours to reach equilibrium before producing results. The six setups are used to see the effect of different physics options on the simulation, including WSM6, Goddard, and Thompson microphysics and MYNN2 and MYNN3 PBL physics. For each physics setup model, 31 ancillary members are generated using the SKEBS scheme for perturbating the weather fields to produce probabilistic simulations.

Figure 23 shows an example of simulated cloud fraction compared to the cloud mask from EUMETSAT. The six plots on the right show the probability of clouds by grid cells for different setups. The panels show the probability of a correctly simulated cloud, i.e. the percentage of the 32 members that match the observation. During the Saharan dust event, both the simulated and observed clouds covered most of the sky over central and western Europe. Amongst the six physics setups in our numerical experiments,

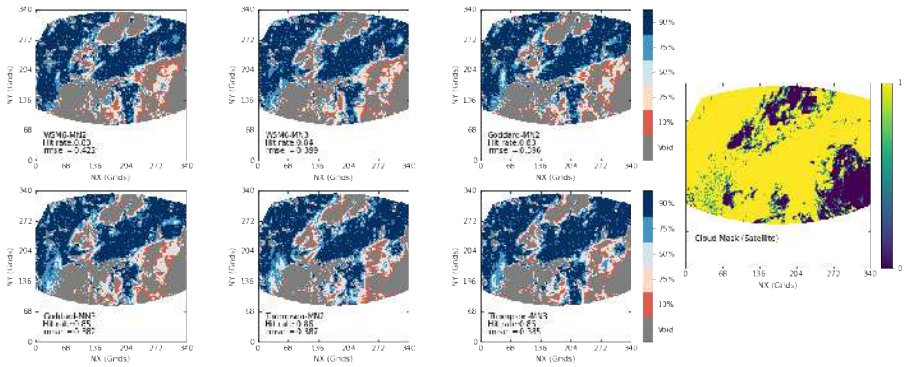


Figure 23: The probability analysis of cloud mask on 2018-10-28 at 12:00:00 for six simulation physics setups and the cloud mask from satellite image.

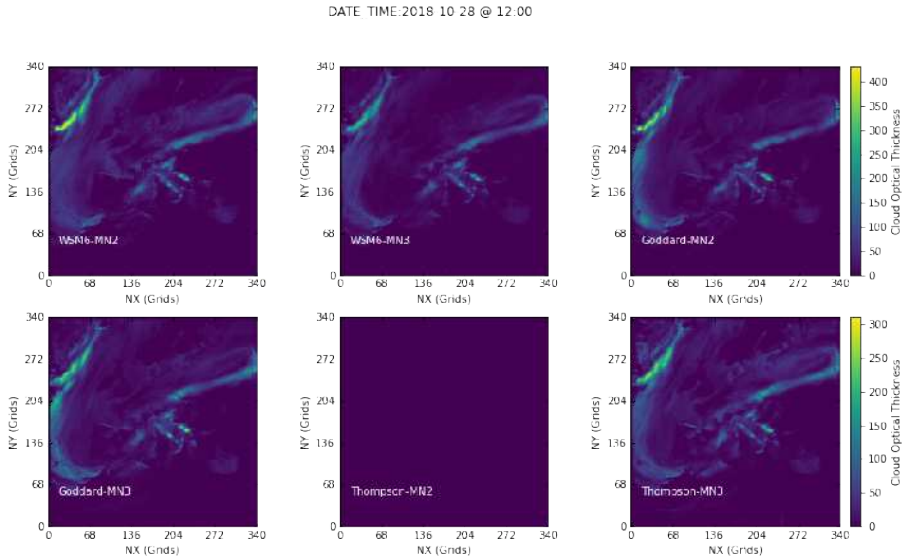


Figure 24: The probability analysis of cloud optical thickness on 2018-10-28 at 12:00:00 for six simulations

Thompson performed better than the other microphysics.

An example result of cloud optical thickness (COT) is also shown in Figure 24 for 12:00 on 2018-10-28. (Of the six setups, one failed to produce COT). The plot shows the average COT from the 32 ensemble members. The different setups gave similar cloud patterns. The difference in the simulation cases was in the magnitude of the COT. Figure 25 shows the standard deviations. The difference in the COT variability results from the different microphysics, with Goddard-MN3 showing the most.

One important process for Saharan dust emissions is the uptake of dust by high wind speeds at the ground infront of a long wave trough reaching into the Saharan desert in the West of Algeria. The interaction between Saharan dust an meteorological fields is diverse. Initially, Saharan dust emissions are triggered by saltation, which is the bombardment of the ground with larger grains of sand. Thus, uncertainties in the wind field, especially the distribution and strength of wind gusts, is a large error source of Saharan dust emissions. Further, the transport of Saharan dust is controled by the large scale circulation. As the Saharan dust release is often associated with frontal zones and synoptic disturbances (e. g. [50]), which are one of the key sources of uncertainty in the large scale circulation, the uncertainties in wind fields contribute to the uncertainty of the Saharan dust dispersion. In addition, the present of dust in the atmosphere is likely to trigger cloud formation, which also influences the persistence of the clouds and

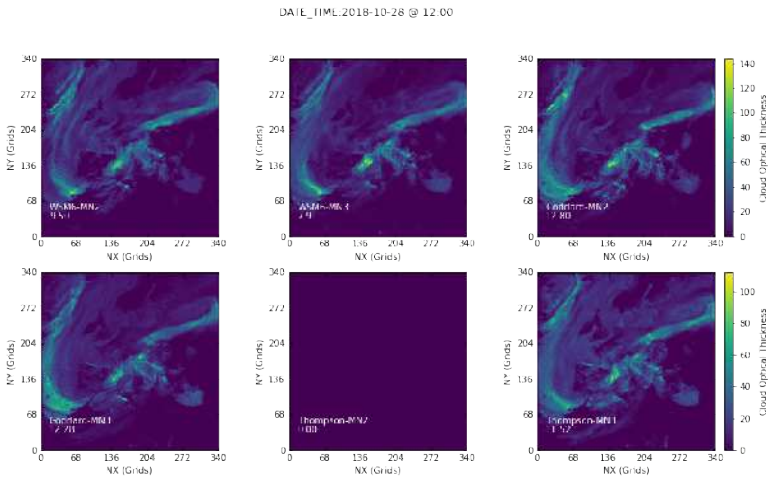


Figure 25: The probability analysis of cloud optical thickness on 2018-10-28 at 12:00:00 for six simulations

the cloud droplet size, also commonly know as aerosol indirect radiative effect. Thus, the presence of dust does not only reduce the solar irradiation by absorption and scattering itself, but also by altering the clouds properties and existence. In summary, the interaction between Saharan dust and meteorological fields is bi-directional. In this report, only the influence of the meteorological fields on the Saharan dust is discussed. The influence of the Saharan dust on clouds is leveraged as this demands an online coupling of the meteorological and atmospheric chemical submodels of ESIAS. This basis for an easy and efficient coupling within ESIAS is set within the project time by the utiliziation of MELISSA, as described in the respective deliverables. However, the online coupling of the two ESIAS compartments was not the scope of this project.

As a demonstration of the capabilities of the full offline coupled ESIAS system to analyze uncertainties in Saharan dust emissions, a case study has been performed analysing the influence of uncertainties in the meteorological fields on Saharan dust emissions (SDE). Therefore, the Saharan dust event initiated on 27. October, 2018, ([158]) was analysed. This Saharan dust event was characterized by strong mineral dust wet deposition in Southeast Europe. For analyzing the impact of meteorologically induced uncertainties on the SDEs, the model setup as described in the following was used.

To simulate the uncertainty incorporated in the Saharan dust event in October 2018 a large ensemble simulation was performed with ESIAS. Two different parameterizations with different realizations have been combined to a total of six different model setups for ESIAS-met. These combinations are listed in Tab. 8. The special characteristics of the Saharan dust event in October 2018 are accounted for by the choice

	Microphysics	PBL
1	W6	MYNN2
2	W6	MYNN3
3	Goddard	MYNN2
4	Goddard	MYNN3
5	Thompson	MYNN2
6	Thompson	MYNN3

Table 8: Parameterization setup for the ensemble analysis of the Saharan dust event in October 2018. Here, PBL refers to parameterizations of the planetary boundary layer. References are: W6 ([69, 91]), Goddard ([147]), Thompson ([150]), MYNN2 ([108]), MYNN3 ([109])

of the different parameterization schemes. The perturbations in the microphysics induce perturbations of the cloud formation and cloud properties, while perturbations of the planetary boundary layer influence the winds close to the ground. In addition to the pertubations in the parameterization schemes used for the analysis, each parameterization setup was stochastically perturbed using the SKEBS scheme (Stochastic Kinetic Energy Backscatter Scheme, [14]) leading to one control run and 31 stochstically perturbed ensemble members for each parameterization setup. Thus, in total 192 ensemble members have been simulated from 27. October, 2018, until 3. November, 2018, in a single long ensemble integration. The domain covered large parts of the Saharan desert on a 15 km horizontal resolution to include the dust source regions in the model. As the Saharan dust is transported towards the Southeast of Europe, Northern Europe is excluded from the model domain.

The meteorological fields of the ESIAS-met ensemble are used as input for the ESIAS-chem ensemble simulation for perturbing the Saharan dust emissions. The impact of the perturbed meteorological fields on the SDE is shown in Fig 26. In comparison with the variability of SDE induced by perturbations of the soil texture and desert fraction, which are parameters controling the dust uptake, the influence of the meteorological perturbations is generally one order of magnitude larger. Thus, in selected areas as the Northeast of Algeria the emission uncertainty of Saharan dust induced by the uncertainty of the meteorological fields (especially in the wind fields) is up to 100% or even larger.

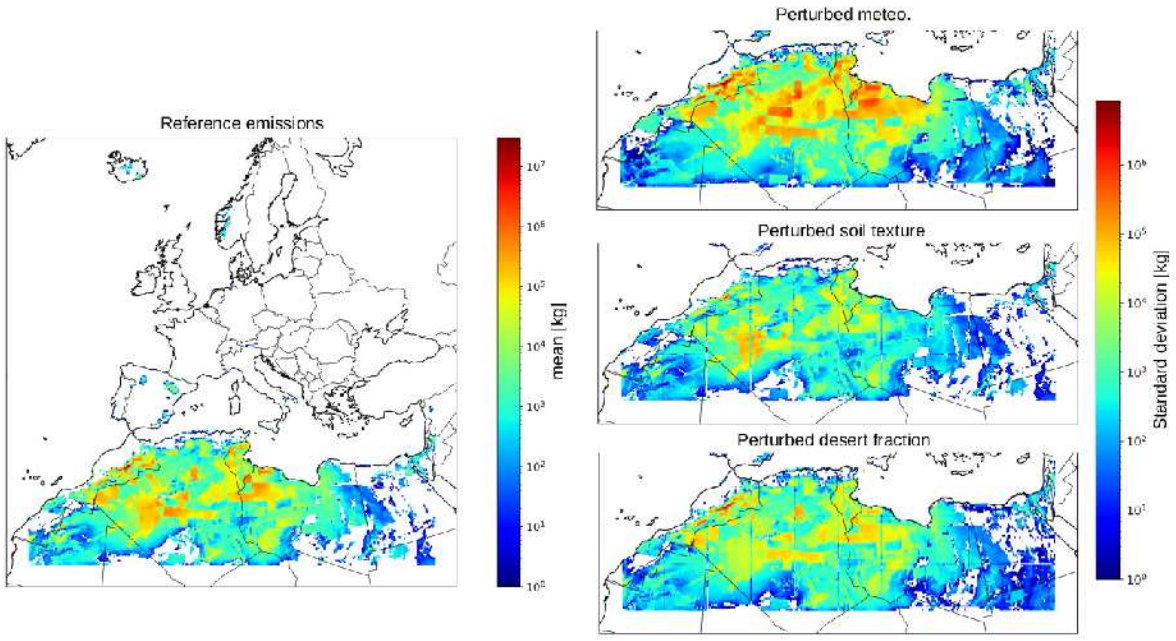


Figure 26: Comparison of the reference emissions of Saharan dust for the Sahran dust event initiated on 27. October, 2018, and the variability of Saharan dust emissions induced by meteorological uncertainties as well as by perturbations of the soil texture and desert fraction.

Further perturbations of the emission parameters is planned to also test the combined effect of joint perturbations of different parameters. However, even with simultaneous perturbations of different emission parameters the effect of uncertainties in the meteorological fields is expected to remain with the largest impact.

Fig. 27 shows the aerosol optical thickness (AOT) of the different subensembles with each using a different parameterizations setup on 29. October, 2018, 03 UTC. Although the different combinations of parameterizations produce large differences in wind speeds at the ground (not shown), the averaged AOT show only few differences in the horizontal distribution of the AOT. However, large differences in the emission strength (see also Fig. 26) lead to different aerosol loadings in the atmosphere. Thus, the ensemble setup is sufficient to model the variable amount of Saharan dust in the atmosphere but it lacks in

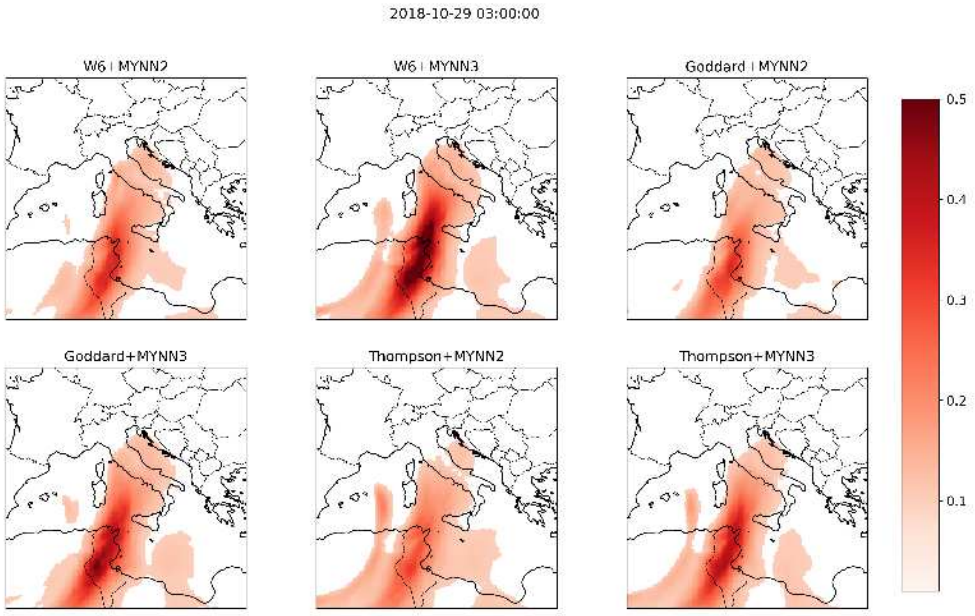


Figure 27: Ensemble mean of AOT on 29. October, 2018, 03 UTC, for each parameterization setup including the control and 31 perturbed ensemble members.

simulating uncertainties in the horizontal distribution of the Saharan dust. This suggests to include further perturbations of the parameters in the dust emission module or even the use of different dust emission parameterizations.

The lack of horizontal variance of Saharan dust in the atmosphere is also illustrated in Fig. 28, which compares the observed PM10 concentration in central Westitaly (upper panel) and Southeast Italy (lower panel) with the ensemble simulation. While the modeled advent and strength of the dust aerosols in central Westitaly is in good agreement with the observations, the ensemble fails to simulate both in Southeast Italy. The simulated dust at the surface in Southeast Italy is by a factor of 3 to low, although selected ensemble members are able to simulate the large peak concentrations of approximately 150 $\mu g/m^3$. However, nearly all ensemble members simulate the peak aerosol concentration at around 29. October, 2018, 21 UTC, which is about 24 hours earlier than the observed peak aerosol concentrations.

The results suggest the use of a particle filter to constrain the ensemble spread by observations. However, although the ESIAS-chem model framework is now able to utilize MELISSA for ensemble generation, the extension of the model towards particle filtering applications needed to be postponed due to critical model errors that hampered further progress. These errors as well as the performance analysis of ESIAS-chem using MELISSA is described in more detail in Deliverable D1.2.

T1.2.1-3 - Joint simulations

The particle filter data assimilation Deterministic simulations create a single prediction with the best accurate physics, parameters, and solutions to the governing equations. However, the limitation of deterministic simulation results in a high consumption of computational resource (e.g. very high necessary resolutions of the modelling domain) and cost of man power to improve the simulation models, especially the complex Earth System Modelling that couple with atmospheric circulation, soil-vegetation-atmospere transfering, soil hydrology, and human-nature interaction. This highly non-linear system requires great effort to improve its accuracy in order to perform a good deterministic simulation.

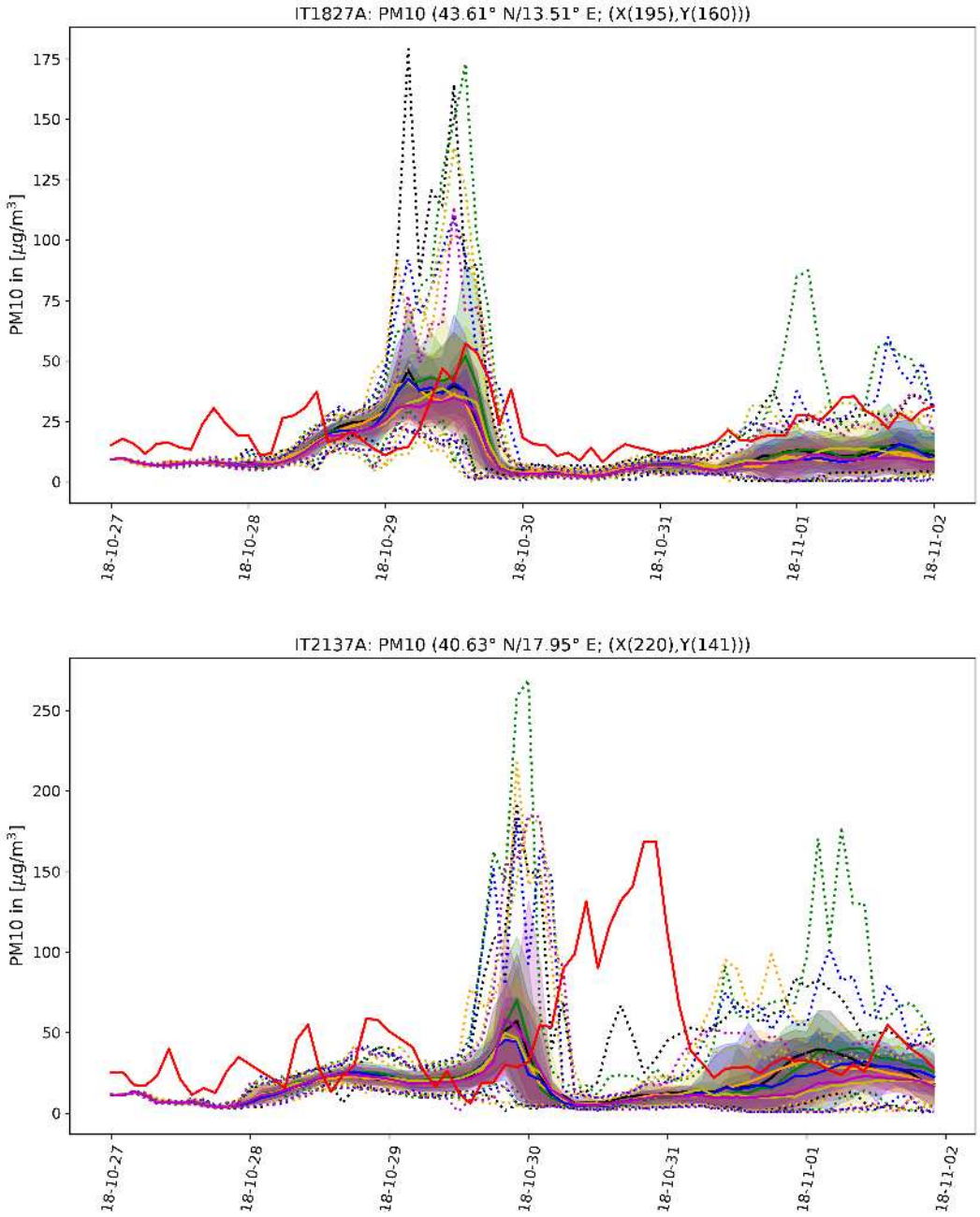


Figure 28: Comparison of the PM10 observations in central West Italy (upper panel) and Southeastitaly (lower panel) with the simulated PM10 concentrations. The different colors indicate different parameterization setup in ESIAS-met, where the solid line shows the subensemble mean and the shading illustrated the range of the ensemble simulations.

Data assimilation is one of the mathematical method to increase the accuracy of deterministic simulation [8]. Novel observation techniques that increase the temporal and spatial availability of observation data, such that researchers can adjust simulation results by assimilating such observations. The adjusted simulation results will be closer to the observation without yielding inaccurate fields of physics variables that can cause simulations to fail due to the convergence problems or inconsistencies in mass or energy conservation.

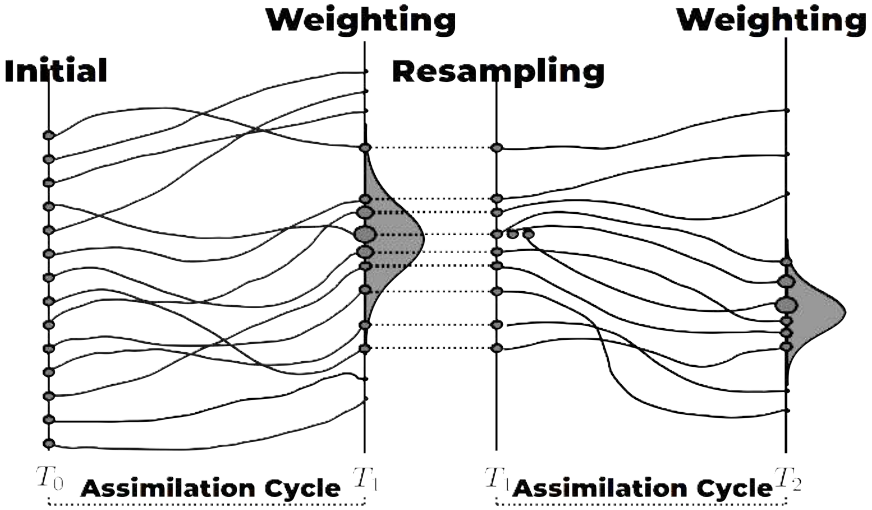


Figure 29: The concept of Particle Filter (adapted from [155])

Probabilistic simulation can play a key role here. With high performance computers and large resource allocating on pre-exascale computers, researchers can perform very large ensemble simulations [42, 117, 32, 61, 87]. Large ensembles can result in a good sample size for statistical analysis. Moreover, large ensemble simulations can be used to estimate the uncertainty in the numerical solution and the physics in order to create probabilities.

In our research aims, the particle filter is the main algorithm for probabilistic simulation and data assimilation [155]. The particle filter is a data assimilation method based on the concept of the Bayes equation for large ensembles of highly nonlinear simulations. The probability of a model can be calculated by its simulated results as compared to the observation. This representation of an ensemble member is called a particle. The weighting of the particles shows the importance of different particles and thus the probability density of the model is calculated. This is sketched in Figure 29.

The density of the particle is given as

$$p(\psi) = \frac{1}{N}\sum_{i=1}^{N}\sigma(\psi - \psi_i) \tag{1}$$

,

where ψ is the system state of study target, $p\psi$ is its probability density, and any ψ_i the system state based on the particle i.

We follow the van Leeuwen's procedure to have importance sampling to filter the particles using following euqation:

$$p(\psi|d) = \sum_{i=1}^{N} w_i\sigma(\psi - \psi_i) \tag{2}$$

where d is the observation and the weight w_i is given by:

$$w_i = \frac{p(d|\psi_i)}{\sum_{j=1}^{N} p(d|\psi_j)} \tag{3}$$

The $p(d|\psi_i)$ is often taken as a Gaussian to estimate the importance of a particle. The selected particles are simulated further according to their importance and/or also with an increased number of

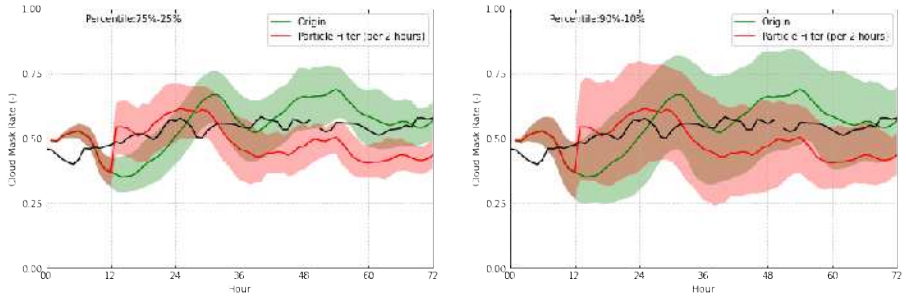


Figure 30: The cloud mask rates (covering rate) of by the large ensemble simulations with (red) and without particle filter (green) comparing to the observation (black). The right plot shows the resulting cloud mask rates between 75th to 25th percentile, and the left plot shows the one between 90th to 10th percentile. The particle filter exclude the particles with less weight of importance.

simulations pertubated around it using a stochastic scheme. The original ESIAS-Met wasn't able to use the stochastic scheme to restart a simulation. However, the removal of particles decreases the ensemble spread and thus decreases the capability of capturing the probability of a simulation output. The technical issue is solved by inserting a function for forcing restarts with which ESIAS is able to restart the simulation while reallocating rather than losing particles. Without this step, the importance sampling reduces the number of particles with each sampling step. This development thus facilitates the use of the particle filter for data assimilation.

The first particle filtering is performed for the European domain using a cloud mask with probability densities from a large ensemble simulation with 128 particles. The particle filtering is applied after 12 hours of simulation time to allow the simulation to reach equilibrium and the ensemble to have enough spread to capture outliers. After the first filtering (12:00 on 2018-10-10) the sampling is performed again after every 2 hours of simulation. Figure 30 shows the resulting 75th to 25th and 90th to 10th percentile bands for the simulated cloud in the left and right plots, respectively. The plot of 90th to 10th percentiles captures most of the spread and shows how the particle filter excludes outliers to focus on the closest particles. We see however how fewer base particles means less possibility of spread in the future. The cloud mask rate initially matches better to the observation, but loses accuracy over the long term after 60 hours of simulation, as the ensemble no longer contains enough independent samples to calculate future probabilities.

Figure 31 shows another particle filter result, but in the summertime (2018-08-09). The simulation uses the same European domain, but with a different forecast length (72 hours of simulations). The result of 72 hours of simulation is shown and compared to the observation. Here, due to the lack of strong sea-atmospheric interaction in the Mediterranean Sea, the cloud is not well formed and thus the simulated cloud cover is less than in the observation. The particle filter forces restarts to maintain the size of the ensemble and thus the particles with lower weighting are restarted with new perturbations in the weather field to increase the variability of the resulting simulation.

The simulation shows a complete situation where the prediction fails to capture the major cloud formation. The particle filter works with two setups, (a) with the particle filter scheme weighting the particles for each hour after the spin-up and also (b) every 6 hours to increase the particle spread. The particles with less cloud cover are excluded and thus we can see the improvement in performance of the median. Moreover, due to the intensive filtering, the particle size decreases since the spread is not large enough for the one-hour restarts. The weighting and filtering after 6 hours or simulation increase the particle spread (green), but the accuracy is not as good as for the hourly filtering, so there is a trade off.

This experiment shows the contradiction of the ensemble simulation with data assimilation. When en-

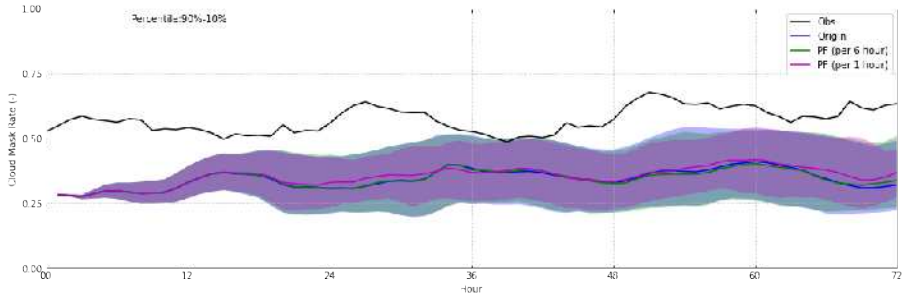


Figure 31: The cloud mask rates (covering rate) over Europe for the large ensemble simulations (1) without the particle filter (blue), (2) with particle filtering every six hours, and (3) every hour, all compared to the observation (black). The color blocks indicate the range between 90th to 10th percentiles. The particle filter in this case does not reject any particles and maintains the ensemble size of 128 members.

semble simulation produces a large ensemble size with great variability from the model uncertainty, the data assimilation method will decrease the variability of the model output by reducing its uncertainty. The spread of particles from the large ensemble simulation therefore decreases but the accuracy can increase, though this effect is very small at the studied filtering rates. The data assimilation for the highly non-linear model is however essential, because the particles with higher weighting can remain within the simulation and then the accuracy of the whole system can be improved.

MELISSA middleware In order to apply the more advanced particle filtering scheme (e.g. localization or Gaussian resampling) we also integrated ESIAS with MELISSA-DA to utilize the advantages of an elastic ensemble simulation framework. Work Package 5 instrumented the MELISSA-DA [51] functions into the WRF3.7.1 for constructing ESIAS-Met. The intrumentation of codes and the prototype of the ESIAS-met-MELISSA is successful, though more optimization of the model is required for simulating ultra-large ensembles. WP5 has also finished instrumenting the EURAD-IM (the core of ESIAS-Chem) with the MELISSA-DA API for further application. More detail of the experiment by MELISSA-DA for the meteorology issue over the European domain can be found in the D5.3 report.

In the future, the WP5 partners should improve ESIAS for two-way coupling of the meteorology and chemistry components. This two-way feedback would include any effect of aerosols on the cloud formation to capture more physically accurate and thus realistic cloud simulation.

4.2 Task 1.2.2: Scientific results

Task 1.2.2 is divided into 2 subtasks:

T1.2.2-1	**Power calculation from ensemble winds and COT**[1]. Training of the IEE wind power management system reference plants and ESIAS ensembles. Optimization of the IEE solar prediction system for varying modeling conditions in a two-step process: 1) Optimization of power model to German TSO electric meter data based on live satellite measurements of COT. 2) Spatially resolved calibration of a WRF input time series to the satellite COT grid. Solar and wind power calculations for WRF meteorological ensembles of COT and 100-m wind will be provided as cumulative distribution functions.

[1] Description edited from D1.1 for accuracy differentiating wind and solar training

T1.2.2-2	**The confluence of ultra-large ensembles with IEE's satellite-based cloud-tracking system for short-term forecasting**. Wind velocity gradients at cloud heights from a Taylor-based cloud-tracking system can identify flow structures and dynamically weight the intra-day wind power ensemble, as can satellite measurements of cloud cover for the solar power forecast. This ensemble member validation will provide feedback to the particle filter by weighting the ensemble member performance.

T1.2.2-1: Power calculation from ensemble winds and COT

In the first EoCoE project, both the wind and irradiation data from WRF in Europe were initially quite biased, leading to biased estimations of wind and solar feed-in in Germany. Much progress could be made for the wind bias at hub height in the boundary layer, particularly through the surface parameterization. The solar bias nevertheless remained and while the ensemble distributions could be indicative of extreme events through e.g. their skewness or higher order moments, they were not good estimates of the actual German feed-in according to meter data.

In EoCoE-II, focus has been turned to addressing the PV bias, both by optimizing COT in T1.2.1 and by calibrating the solar power model to the WRF irradiation data in this task. Additionally, the wind model has been changed from a physical model to an operational model based on machine learning to power measurements. In both cases, the goal is to achieve large-scale ESIAS simulations from WRF that have an accuracy more similar to the operational weather models maintained by European weather services and currently used in grid operation.

Solar Prediction System (SPS) The mid-term deliverable demonstrated a preprocessing procedure for a numerical weather prediction's irradiation input to SPS. This calibrated the prediction to spatially resolved satellite irradiation measurements. The resulting map of scaling factors showed good correlation to German topography, indicating topology and surface features as key to the difference between the simulated and measured irradiation.

Since the mid-term, T1.2.1 supplied WRF training data in the form of 512 ensemble simulations for every other day in a full year, representing all seasons and conditions. Despite consisting of 4 model physics combinations ($\times 128$) optimized for cloud optical thickness, the resulting irradiation data remained somewhat overestimated. This could indicate a possible issue with WRF's irradiation schemes downstream of the cloud prediction. Nevertheless, an improved calibration technique could be developed to compensate for this, though such manipulations of the raw ensemble members negatively affected the ensemble probablistic.

Various improvements upon the linear correction model were tested, including a quadratic correction, using neighboring pixels, and experiments with extreme learning machines (ELM), but the best result came from the unique approach of a linear model of the original prediction $I(\vec{x}, t)$ and the additional inputs of percentiles $P_i(t)$ of all irradiation values in Germany. The model thus responds to the correlation of any particular location with the overall conditions in Germany. (We use 11 percentile inputs, i.e. the percentiles $i \in 0, 10, ..., 100$). The individual irradiation percentile time series are smooth and robust, as they result from o(10^5) grid points. It is important to note that these percentiles have nothing to do with probabilistic simulation - they are simply percentiles of the irradiation map applied to the same simulation and timestep.

The improvement is summarized in table 10, which scores the results against the satellite irradiation measurements. The raw ensemble clusters have a large bias in the range of -12 to -24 Wm^{-2}, resulting in an RMSE of typically 114 Wm^{-2}. A scale correction improves these to a bias and RMSE of around 4 and 96 Wm^{-2}, respectively. The new calibration model improves these to about -1 and 79 Wm^{-2}, a significant improvement and a value more typical of an operational forecast like ICON.

Table 10: Table of irradiation model errors before and with two different calibration algorithms for the four 128-member ensemble clusters, i.e. physical schemes, of the full ensemble.

(Wm^{-2})	Goddard-MYNN2	Goddard-MYNN3	WSM6-MYNN2	WSM6-MYNN3
original bias	-24	-18	-24	-12
scale corrected bias	4.3	4.5	4.3	4.3
percentile correction bias	-0.7	-1.7	-1.1	-1.1
original RMSE	114	112	114	114
scale correction RMSE	96	99	96	96
percentile correction RMSE	79	80	78	79

Graphically, the result is somewhat harder to understand than the simple linear model in the mid-term deliverable, which directly showed whether the irradiation was corrected up or down. Here, particular quantiles may result in an increase or decrease of the local irradiation value only on average.

Figure 32 shows the overall result. The weighting of the raw $I(\vec{x}, t)$ is quite low, especially in central and eastern Germany. The second panel shows that the low raw weights are balanced by the contributions of the quantiles. This unfortunately indicates that the calibration comes at the cost of reducing the effective spatial resolution of the forecast, as the quantiles are overall values for Germany. (Ideally, numerical weather predictions would of course be perfect and the original resolution could be preserved). The third panel gives a rough illustration of what quantiles contribute most at each location on average. This is generally around the median, with higher radiations on the coast and elevated alps and lower values common throughout central Germany, where connected cloud systems typically cross the country to the east-southeast.

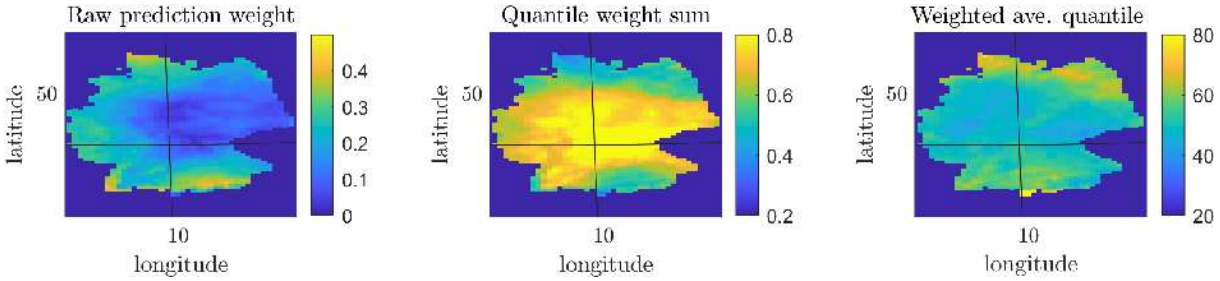


Figure 32: Training weight maps of the percentlie calibration model $y(x,t) = \vec{b} \times [GHI(x,t), P_0(t), P_{10}(t), ..., P_{100}(t)]$ with the raw prediction weight (left) and sum of all percentile P_i weights (center). The right panel shows the (weighted) average percentile for the percentile contributions.

Figure 33 gives an example from Goddard-MYNN2 (ensemble members 1-128) of the percentile weight maps using the the 0 and 100% percentiles, showing some regional preferences. The exact maps and typical mix of the $P_i(t)$ as shown in the last panel varies for each ensemble cluster. The weighting is trained independently for each satellite target pixel ($\vec{x}$) according to around half a million samples (128 members $\times$ 182 days $\times$ 24 hours), though there appears to be room for improvement regarding the number of percentile inputs and overfitting.

The German photovoltaic power feed-in was calculated using the year of preprocessed, day-ahead meteorological data and evaluated against electric meter data. Figure 34 shows the ensemble forecasts as probablity distributions alongside meter data for the last two weeks of October, 2018[2]. The actual feed-

[2]The solar power is generally low in October, but this timeframe contains familiar weather events of interest for T.1.2.2-2

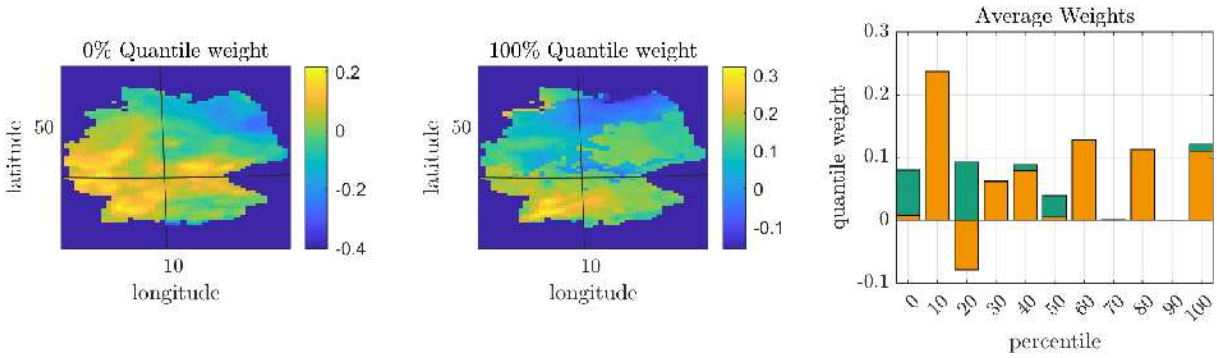


Figure 33: While figure 32 shows average weights over all percentiles, this figure shows sample weight maps for individual percentiles like the min (left) and max (center). The bar chart shows map averages across all percentiles.

in typically lies within the distribution, with the exception of October 19th. Systematically we can only note that the sunrise and sunset can vary from day to day compared to the meter data. The size of the shift can be puzzling, even often appearing as a time zone or summer time error, but these are mixed in with days with no apparent shift and may differ by location, such that this seems to be coincidental and the reason for the shifts is not yet understood. The irradiation calibration is effective in using the quantiles also to partially correct this feature, as there is some variance in the time of dawn and dusk in the domain.

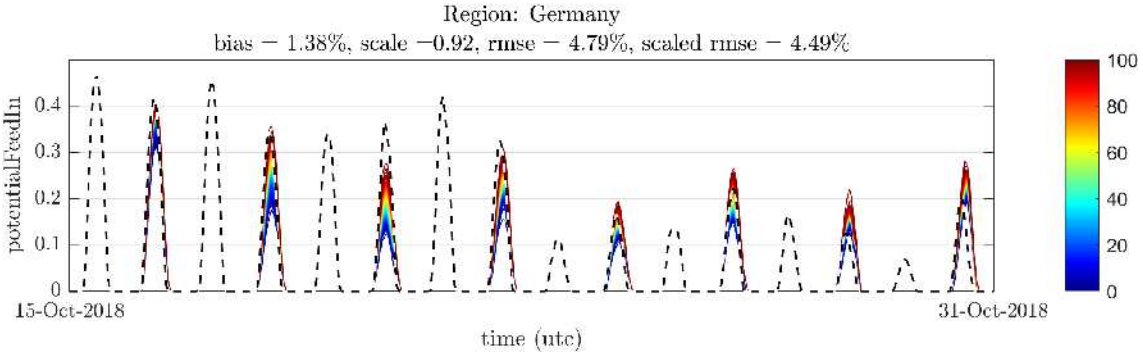


Figure 34: German photovoltaic feed-in time series for the second half of October, 2018 from the 12 months and 512 members of ensemble training data. The day-ahead forecasts are available every other day.

A quantitative evaluation of the probablistic forecast is shown in figure 35 and reported in table 11, which shows that the results and time series errors are similar for the different physics setups. The continuous ranked probability score (CRPS) has been decomposed into its *potential* and *reliability* components according to [67], where the reliability is related to the ranked histogram and bias of members and the potential relates to the uncertainty and behavior of outliers. The figure decomposes these scores further by showing the contributions of each percentile of the ensemble distribution. Like for forecast error, lower values are better.

The largest CRPS contributions result from the lower part of the ensemble distribution. This makes sense, given the general bias as well as the fact that extreme solar errors typically result from overestimates due to e.g. unexpected fog. The major contribution comes from the potential CRPS, especially at the extremes below the 20th and above the 90th percentiles, indicating still too little spread and missing outliers. From the CRPS standpoint, the clusters with WSM6 are better, mainly from their superior reliability.

In the last column of table 11, values are shown from a year of the forty-member opreational ICON-EU-EPS ensemble maintained by the German weather service. The calibration improves the WRF member bias and RMSE to values far closer to this operational model's. The ICON model is also better than the

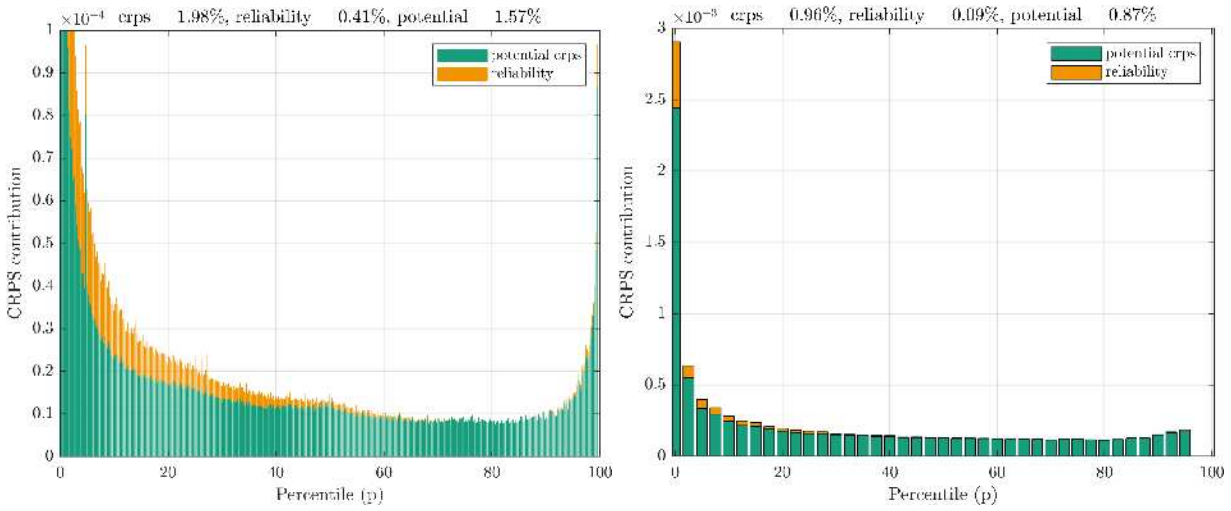


Figure 35: The left panels shows the distribution of the CRPS contributions for every percentile of the ESIAS German PV forecast, (broken down by potential and realiability), calculated over the entire year of day-ahead training data and compared to meter data. For comparison on the right is the distribution for ICON-EU-EPS.

Table 11: Deterministic and probablistic total German PV feed-in errors for the training ensemble year. This is shown for all 512 members as well as the four 128-member clusters. The calibrated model errors are followed by the (raw errors) in parentheses. All values are percentages of the installed PV capacity. The last column shows the operational, forty-member ICON-EU-EPS model for comparison.

(%)	All	Goddard-MYNN2	Goddard-MYNN3	WSM6-MYNN2	WSM6-MYNN3	ICON-EU-EPS
bias	1.4 (4.3)	1.8 (3.9)	1.8 (4.2)	1.1 (4.3)	0.8 (4.6)	0.4
sRMSE	4.5 (14.6)	4.4 (14.6)	4.4 (14.6)	4.5 (14.6)	4.5 (14.6)	3.3
crps	2.0 (1.3)	2.2 (1.6)	2.3 (1.4)	1.9 (1.3)	1.8 (1.1)	1.0
reliability	0.4 (0.4)	0.7 (0.5)	0.7 (0.4)	0.5 (0.4)	0.4 (0.3)	0.1
potential	1.6 (0.9)	1.5 (1.1)	1.5 (1.0)	1.4 (1.0)	1.4 (0.9)	0.9

research model regarding its probablistic scores for the total German power feed-in, though as will be shown in T1.2.2-2, the ultra-large ESIAS ensemble consistently manages to generate a few members that better depict the observed clouds.

Wind Power Management System (WPMS) For EoCoE-II, the operational WPMS was implemented instead of the physical grid model used in the first project in order to improve the accuracy as compared to the meter data. The physical model used typical power curves to directly estimate power from the 100-meter wind, which was aggregated according to maps of installed capacity. WPMS utilizes reference wind parks. A typical strategy in grid operation is to use measured reference parks to extrapolate the normalized power at locations with unmeasured parks. For forecasting, a similar extrapolation can be used and the problem of forecasting regional power is changed to training forecasts for specific reference parks using e.g. machine learning. (This works better for wind than for solar because solar plants behave more uniquely due e.g. to differing orientations). While the physical model assumed an accurate wind input, the trained park models implicitly adjust to the target.

An example of a resulting reference plant forecast is shown in figure 36. The large number of members

provide a good distribution that well captures the measured value. In this figure, the full 48 hours of each simulation are shown, such that there are no gaps as in figure 34. It is however visible how the ensemble distribution begins again on each odd day in October and takes time to spread. We note that the errors often present as phase shifts between the time series and measurement, i.e. where there is some delay in the expected behavior at 100 meters.

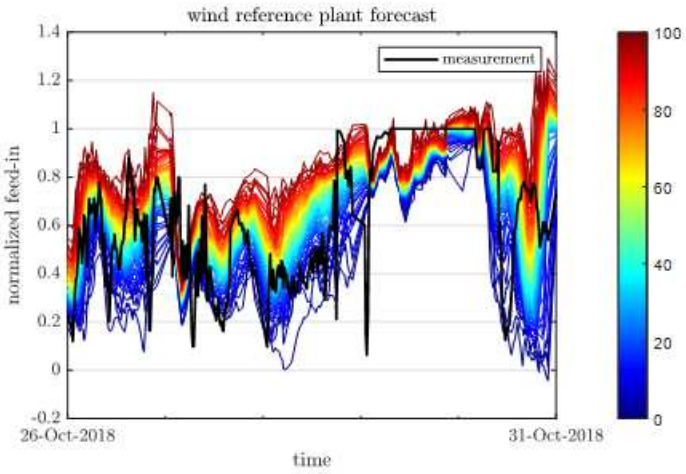


Figure 36: Power forecast for an anonymous wind reference plant according to the ELM model and training data.

The resulting ensemble distributions of the German wind power feed-in are shown for the same two example weeks as before in figure 37. The last five days of October, 2018 were a period of strong cyclogenesis in the western Mediteranian resulting in the storm Adrian that was deadly in northern Italy. On the 24th, 26th, and 28th, the German wind power feed-in was captured by the tail and upper outliers of the day-ahead ensemble distribution. On the 30th, as the eye of the cyclone crossed Germany, the forecast is overestimated or rather a half day premature in the peak wind power.

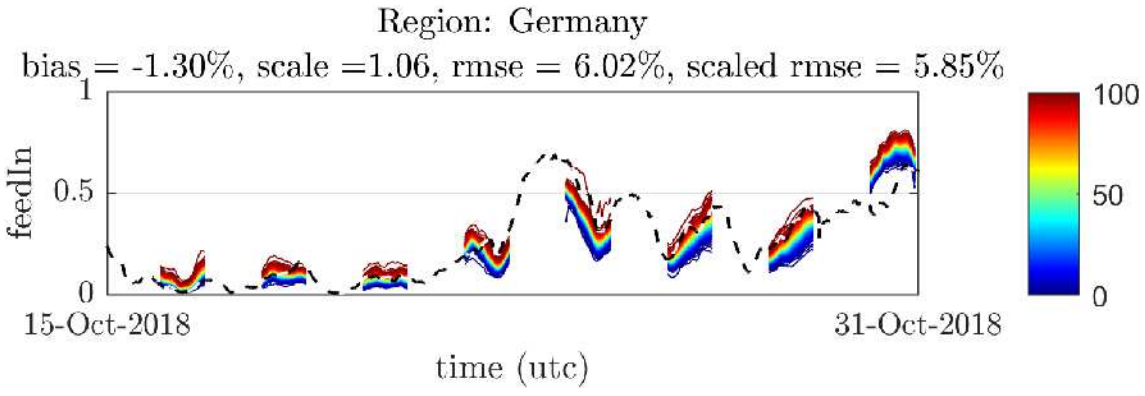


Figure 37: German wind power feed-in time series for the second half of October, 2018 from the 12 months and 512 members of ensemble training data. The day-ahead forecasts are available every other day.

Regarding the general performance of the probablistic forecast over one year, figure 38 and table 12 show the same analysis that was done for PV. We note that the overall errors as a percentage of installed capacity are expectedly larger for wind than for PV, 6% vs. 4.5% RMSE, as a greater amount of the maximum feed-in is achieved on average for a wind turbine than for a solar module. The overall wind power CRPS is 3% as compared to 2% for PV. The contributions are higher at the upper end of the distribution

for wind. As illustrated by the events in figure 37, extreme wind outliers tend to be unexpectedly stormy conditions resulting in higher feed-in. The wind power reliability is an even much smaller fraction of the total CRPS as was for PV, consistent with WRF's wind data and their ranked histograms being better than for the irradiation (see also T1.2.3). Goddard-MYNN3 generally showed the best wind power distribution in Germany, also considering the individual contributions of Germany's four transmission grid regions (not shown).

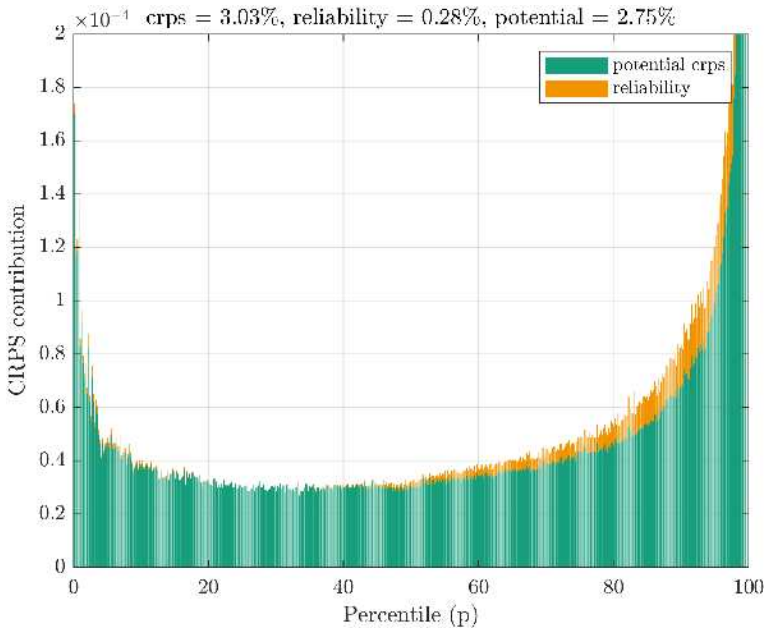


Figure 38: Distribution of the ESIAS German wind power CRPS contributions for every percentile, (broken down by potential and realiability), calculated over the entire year of day-ahead training data and compared to meter data.

Table 12: Deterministic and probabilistic errors of the German total wind power feed-in for the full and sub-ensemble groups. The values are percentages of the installed wind capacity.

(%)	All	Goddard-MYNN2	Goddard-MYNN3	WSM6-MYNN2	WSM6-MYNN3
bias	-1.3	-1.2	-1.3	-1.3	-1.4
sRMSE	5.9	5.9	5.8	5.9	5.8
crps	3.0	2.7	2.4	2.7	2.6
reliability	0.28	0.33	0.18	0.24	0.25
potential	2.8	2.4	2.2	2.4	2.4

T1.2.2-2: Confluence of ultra-large ensembles with satellite-based cloud-tracking

This task aims to combine ultra-large ensemble simulations from ESIAS with real-time satellite data for short-term forecast updates. The satellite data gives not only a current map of clouds and irradiation, but also some information on the synoptic motion using optical flow. Both the current cloud cover and motion are possible observations for the a best-member selection or feedback to the particle filter in ESIAS to generate a better distribution.

Cloud motion tracking The mid-term deliverable showed the development of Fraunhofer IEE's cloud motion system to constrain and classify flow topologies based on invariants of the velocity gradient tensor. The optical flow algorithm optimizes a Taylor approximation of the flow field, thereby creating a continuous description of synotic structures. The velocity gradient data from the numerical weather prediction however proved to be too obscured by turbulence to serve as a good basis of comparison, even at 500 hPa heights.

Unlike the instantaneous velocity, the geopotential output of the weather simulations is better representative of the synoptic scale. While the initial optical flow algorithm created Taylor approximations for both u and v, a new version creates a single approximation of a scalar geopotential whose associated geostrophic wind

$$\vec{u} = \frac{\hat{k}}{f} \times \nabla\Phi \tag{4}$$

best describes the cloud motion, (with f being the local rotation rate).

Figure 39 illustrates a cloud motion optimization from cloud albedo satellite images. In the image, the clouds are shaded with the geopotential height, whose contours are parallel to the motion. Such solutions are used for the geopotential correlation scores in the following section.

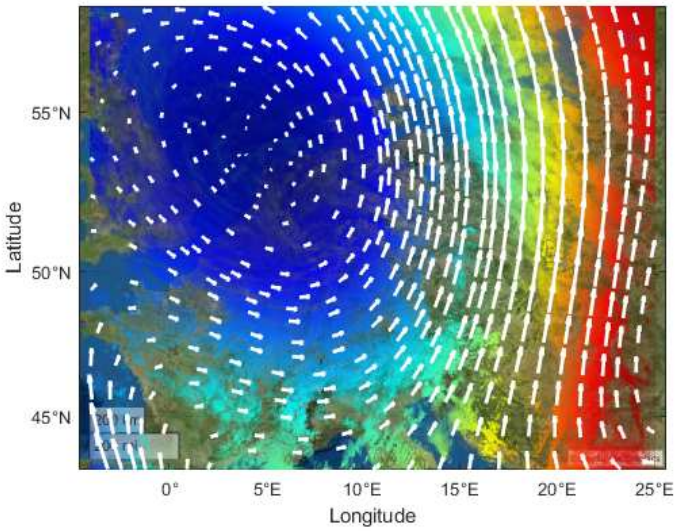


Figure 39: Demonstration of the new optical flow solution during Storm Adrian, centered in the North Sea at 12 pm UTC on October 30, 2018. The velocity vectors follow the gepotential solution, whose values color satellite cloud image.

Ensemble scoring The ensemble members are scored both directly based on cloud cover (GHI RMSE) as well as from the correlation coefficient of the 500 hPa geopotential with the cloud motion solution. Both scores are evaluated for the 512 day-ahead simulations in October, 2018, shown in figure 40. Despite the overall bias found in T1.2.2-1, we see that some ensemble members can match the satellite irradiation far better than most (excluding on a simple clear-sky day). On some days, particular physics clusters (e.g. 1-128, 129-256, etc...) generated more accurate clouds overall, while on others, just a few members manage the particular details of the clouds that occured. An example of this is shown in figure 41, where a few of the 512 members give an accurate representation of the actual clouds. The impressive aspect of this result was its consistency - on nearly every day a member of the ultra-large ensemble could be found that contained the major details of the occuring clouds. This was not observed in a smaller, twenty-member operational ensemble like COSMO-D2.

Clouds are one of the most difficult and uncertain parts of numerical weather predictions. While focus in this Scientific Challenge has previously been on the potential of ultra-large simulations to capture extreme but rare synoptic level events, the results here highlight the potential of large ensembles to provide at least some correct realizations of the small-scale structure that is too uncertain to simulate deterministically. In grid operation, such data could be used daily to plan switching states robust to all eventualities of congestion and line failures.

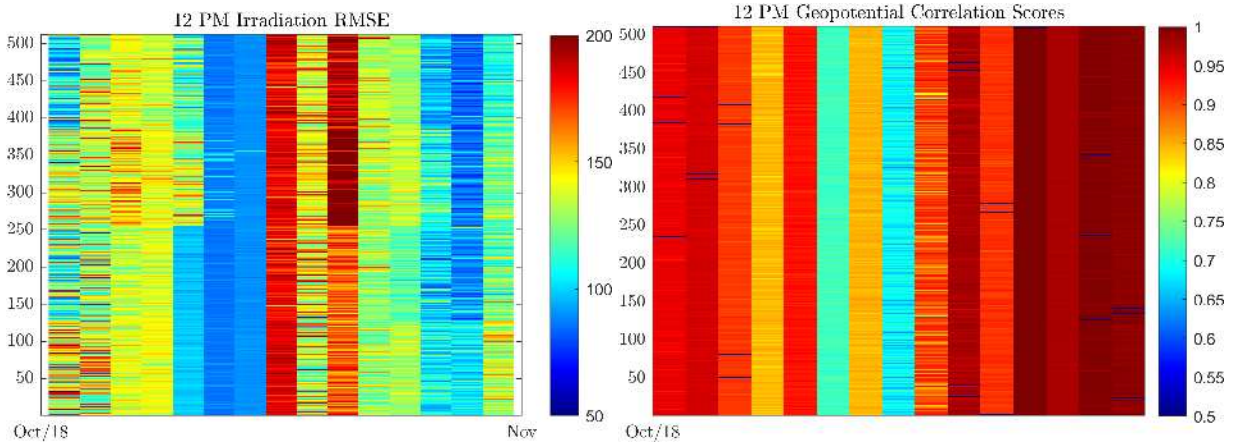


Figure 40: Scores for all 512 ensemble members and even days in October, 2018 at 12 pm UTC according to the clouds, i.e. GHI RMSE (left) and geopotential correlation scores (right, scaled from min to max). Members 1-128, 129-256, 257-384, and 385-512 use physics Goddard-MYNN2, Goddard-MYNN3, WSM6-MYNN2, and WSM6-MYNN3, respectively.

For the motion or geopotential, the ensemble selection is less clear. The new optical flow works more stably and the simulations match very well with the observation (also with COSMO-D2). It only fails to perform well on a few days with little or very incomplete cloud cover. However, there is very little difference in geopotential within the ensemble groups themselves, such that the scores are not helpful for differentiating the simulations. This is illustrated by showing the nearly identical fields of the best and worst geopotential matches on a day in figure 42. It would seem that this large-scale structure does not vary much from the global model initial condition, such that outliers in the synoptic structure are rare. Accordingly, no relationship was found between the cloud motion score and wind power prediction errors, which result closer to the ground and are some physical steps removed. Perhaps more variance could be created in the future by using a global EPS model for multiple initial conditions instead of using the ERA5 reanalysis data. It's unclear if generating more variance is however at cross-purposes with more accurate simulations, as it was not clear that the ensemble group was necessarily innacurate in its geopotential predictions.

4.3 Task 1.2.3: Scientific results

Task 1.2.3 is divided into two subtasks:

T1.2.3-1	**Calibration of improved probabilistic wind ensemble predictions**. Exploitation of ultra-large probabilistic point forecasts using the optimized physical schemes from D1.2. Development of statistical non-parametric calibration adapted to 100 m turbine hub-height winds for lead times from 3 hours to one day.

T1.2.3-2	**Calibration probabilistic downward solar radiation ensemble predictions**. Exploitation of the ultra-large ensembles of forecasts in order to produce improved power predictions with focus on downward solar radiation for photovoltaic, for lead times from 3 hours to one day. Extension of the method of statistical non-parametric calibration for clouds.

This last task of the SC concerns methods for probablistic (quantile) calibrations of ultra-large ensemble meteorological data. The study utilizes machine learning and the 512 member ensemble of year-long training data provided in T.1.2.1, reduced to 100 quantiles. The calibration is developed for both solar and wind, i.e. global horizontal irradiation (GHI, subtask T1.2.3-2) as well as 100-meter wind speeds (subtask T1.2.3-1). As target data we use the AGATE dataset of the CAMS radiation service (Copernicus Atmosphere Monitoring Service, freely available at `http://www.soda-pro.com`) for solar and the ARPEGE analysis data from Meteo France [29] for wind.

Methods Many statistical methods exist to model the probability distribution of one predictand knowing some predictors. Here we use quantile regression (QR) methods which describe this distribution by providing a discrete subset of its quantiles.

In this study, we used a U-Net architecture [128], which is a fully convolutional network (CNN) that generates images from images in order to take advantage of information contained by spatial structures. Its name comes from its U-shaped architecture in which convolutional layers are separated first with pooling layers and then with transposed convolutional layers. The first phase, with pooling layers, reduces the size of images, which is known to capture the context of input images. The second phase, with transposed convolutional layers, increases the size of the contracted images, enabling precise localization. These particularities fit the needs of forecast correction.

The architecture used is the same one that is described in [40]. We used a padding of 1 in order to have the same resolution for inputs and outputs of the U-Net. Adding a padding generates inconsistencies on the boundaries of the patches. The input patches are then overlapped and the outputs are cropped to remove the boundaries of the output patches, resulting in 48×48 output patches from the 64×64 input patches. There is no activation function after the final 1×1 convolutional layer in order to produce a regression. The complete architecture of the CNN is depicted in figure 43.

The output layer is composed of 9 neurons that are expected to reproduce 9 equally spaced quantiles (from $\tau = 0.5/9$ to $\tau = 34.5/35$). The loss function used is the quantile regression error function (or Pinball function), defined for the quantile τ as:

$$\mathcal{L}_\tau = \frac{1}{N} \sum_{t=1}^{N} \rho_\tau \left(y\left(t\right) - \hat{y}_\tau \left(t\right) \right) \tag{5}$$

with the function ρ_τ defined as:

$$\rho_\tau(\epsilon) = \begin{cases} \tau\epsilon & \text{if } \epsilon \geq 0 \\ \left(\tau - 1\right)\epsilon & \text{else.} \end{cases} \tag{6}$$

The complete loss function of the NN $\mathcal{L}$ is the sum of the functions of the M quantiles:

$$\mathcal{L} = \sum_{\tau \in M} \mathcal{L}_\tau \tag{7}$$

We faced the well known quantile crossing issue, which happens when the ordered quantiles do not increase monotonically. Therefore, we simply rearranged the quantiles afterward.

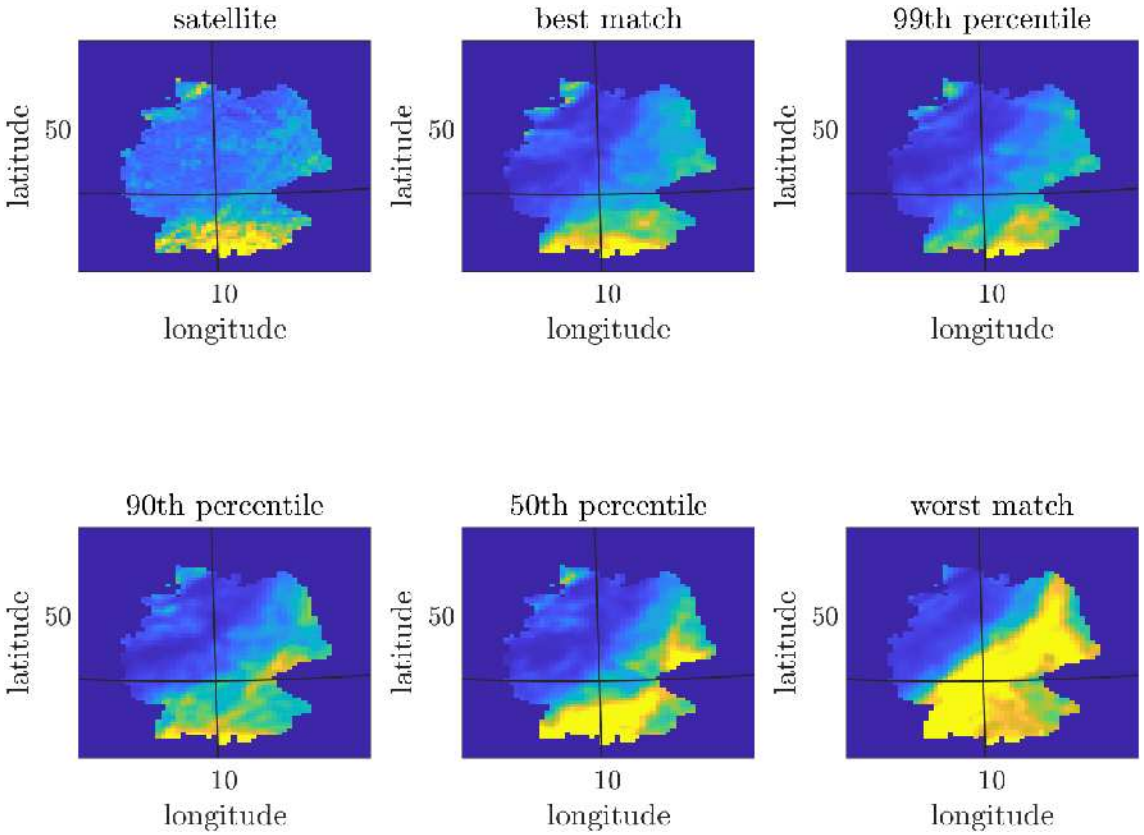


Figure 41: Irradiation maps of cloud cover on October 2nd, 2018 from the satellite observation as well as different ensemble members ordered according to the score in figure 40.

Forecast evaluation There are two main requirements for a high quality probabilistic forecast. The primary requirement is the distribution correctness, which is known as calibration or reliability [74]. A reliable forecast is obtained when there is statistical consistency between the *a priori* predicted probabilities and the *a posteriori* observed frequencies. In that sense, a climatological forecast would be highly reliable, demonstrating that high reliability is a necessary but not a sufficient condition for a high quality probabilistic forecast. That is why there is a second requirement, called resolution [74], which is the ability of a forecast system to distinguish among different forecast situations. A high quality probabilistic model provides reliable forecasts with high resolution.

Following the recommendations given by [84] on the evaluation of the quality of solar irradiance probabilistic forecasts, we used the continuous ranked probability score and rank histograms in order to assess both reliability and resolution.

CRPS The most popular metric in the weather forecasting verification community to quantitatively assess probabilistic forecasts is the continuous ranked probability score (CRPS, [162]), which simultaneously assesses calibration and resolution of probabilistic forecasts. The CRPS measures the difference between the predicted and observed cumulative distributions functions (CDF), and it is defined as:

$$\text{CRPS} = \int_{\mathbb{R}} \left[F(x) - \mathbb{1}_{(x \geq y)}\right]^2 \text{dx} \tag{8}$$

with F the predicted cumulative distribution function (CDF) and $\mathbb{1}_{(x \geq y)}$ the empirical CDF of the scalar observation y ($\mathbb{1}$ is the indicator function). Therefore, a good prediction corresponds to a predicted CDF concentrated around the observed value. The CRPS is negatively oriented, meaning that smaller values indicate better predictions. Additionallz, the continuous ranked probability skill score (CRPSS) assesses

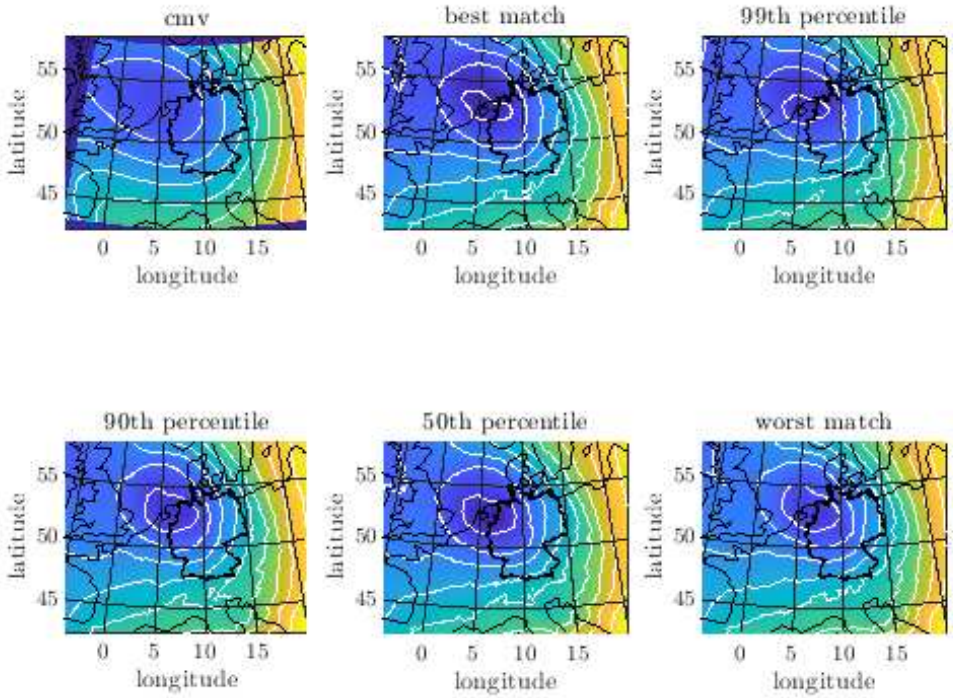


Figure 42: Geopotential maps during Storm Adrian on October 30th, 2018 from the cloud motion vector observation as well as different ensemble members ordered according to the scores in figure 40.

improvements between a distribution forecast F and a reference distribution forecast F_{ref}:

$$\text{CRPSS} = 1 - \frac{\text{CRPS}_{(F)}}{\text{CRPS}_{(F_{\text{ref}})}} \tag{9}$$

Using skill scores is also motivated by a desire to equalize the effects of intrinsically more or less difficult forecasting situations (meridian differences, low values near sunset and sunrise versus high values of the middle of the night), when comparing forecasters or forecast systems.

Rank histogram The reliability of ensemble forecast systems can also be assessed visually using rank histograms [65]. Considering a reliable probabilistic forecast, the CDF values of the predictive distributions for the observations should be uniformly distributed, resulting in a flat histogram for these CDF values. Likewise, for a reliable K-member ensemble forecast, the ranks of observations according to the corresponding ordered members should be uniformly distributed between 1 and $K + 1$, resulting on a flat rank histogram. Otherwise, rank histograms can help to detect deficiencies in the probabilistic forecast. For instance, a U-shaped rank histogram generally indicates an under-dispersion. However, uniformity of the overall distribution is necessary but not sufficient for reliability. Ideally the distribution of the ranks should be uniform, conditional on different forecast scenarios [22] in order to ensure that there is no bias compensation.

Although the rank histogram was originally designed to assess ensemble forecasts, it can also be applied to a set of evenly spaced quantiles assuming that they form an ensemble. Finally, the significance of the results was evaluated with the same method as for the CRPS.

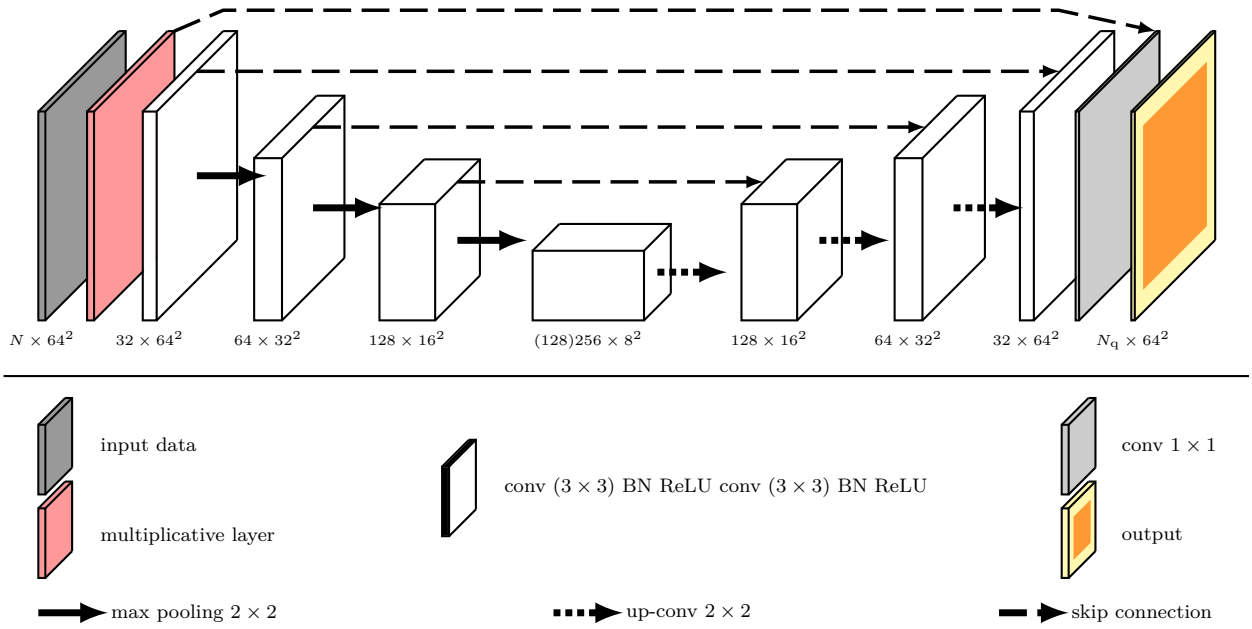


Figure 43: Schematic illustrating the architecture of the CNN used in the study. N represents the number of predictors. N_q represents the number of quantiles calculated (35). BN stands for batch normalization. The numbers under the different blocks indicate the shape of data on the output of the block of calculation at different stages of the network. On the output, the orange part represents the crop (48×48) from the yellow part (64×64).

4.3.1 T1.2.3-2: Calibration of probabilistic downward solar radiation ensemble predictions

Details on the method Due to the coarse horizontal resolution $(0.2° \times 0.2°)$ and the limited extent of the domain (moreover, data are only available over land) and time period, we chose to gather the lead times of the whole period (from lead time 1 to 48) in order to increase the size of the dataset, which is a key feature when training deep learning algorithms. Hence, we made a unique calibration model for the whole period, the whole domain and all the lead times.

Results The rank histogram for the WRF forecast (figure 44a) has an asymmetric U-shape, with higher frequencies associated to the first rank. The asymmetry indicates a positive bias in the forecast while the U-shape, which is common for raw ensemble forecasts, is generally considered as an under-dispersion marker of the probabilistic forecast. After calibration (figure 44b), the histogram is closer to the desired flatness, indicating an improvement in the reliability of the probabilistic forecast. However, there is a slight bell-shape that is generally considered as an over-dispersion marker of the probabilistic forecast.

In figure 45a, the CRPS is stratified by lead time in order to appreciate its evolution as forecast time increases. The calibration reduces the CRPS (improvement of the performance) for all lead times. For the two days forecasted, the improvement is higher (maxima of CRPSS) in the morning, and then decreases along the day to reach a minima in the evening (figure 45b), which is consistent with the study of [7].

The spatial comparison based on the CRPS and the CRPSS (figure 46) shows that the calibration reduces the CRPS relatively to the WRF forecast over the whole domain. It results that CRPSS are positive for the whole domain (figure 46c calculated using WRF as reference in the equation 9). Generally, maxima of CRPS have the same location in WRF and in the CNN (same thing for the minima). Maximum values of CRPS are obtained over regions with complex topography (mountains), where minimum values of CRPSS are also experienced.

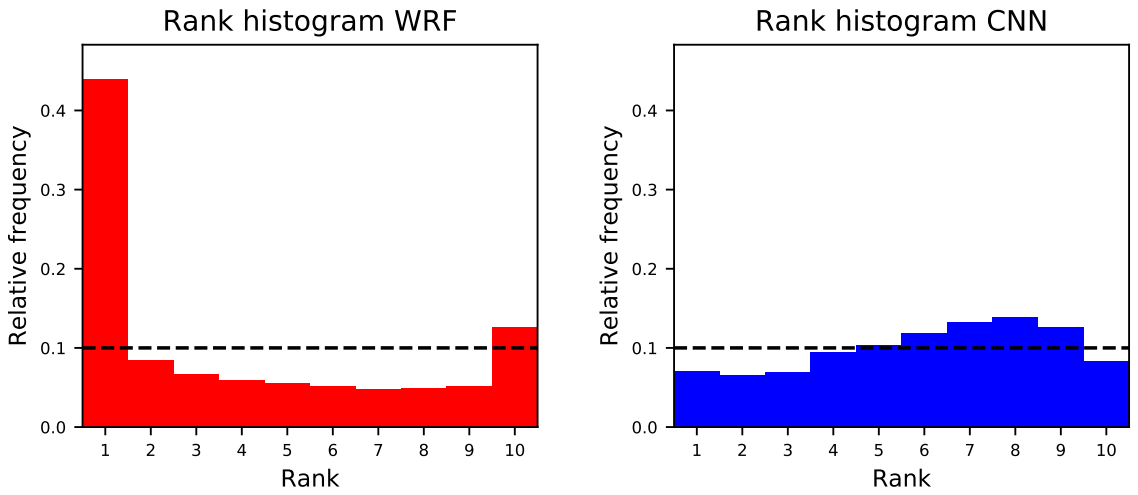


Figure 44: Rank histograms for the WRF (left) and CNN (right) probabilistic forecasts of solar irradiance.

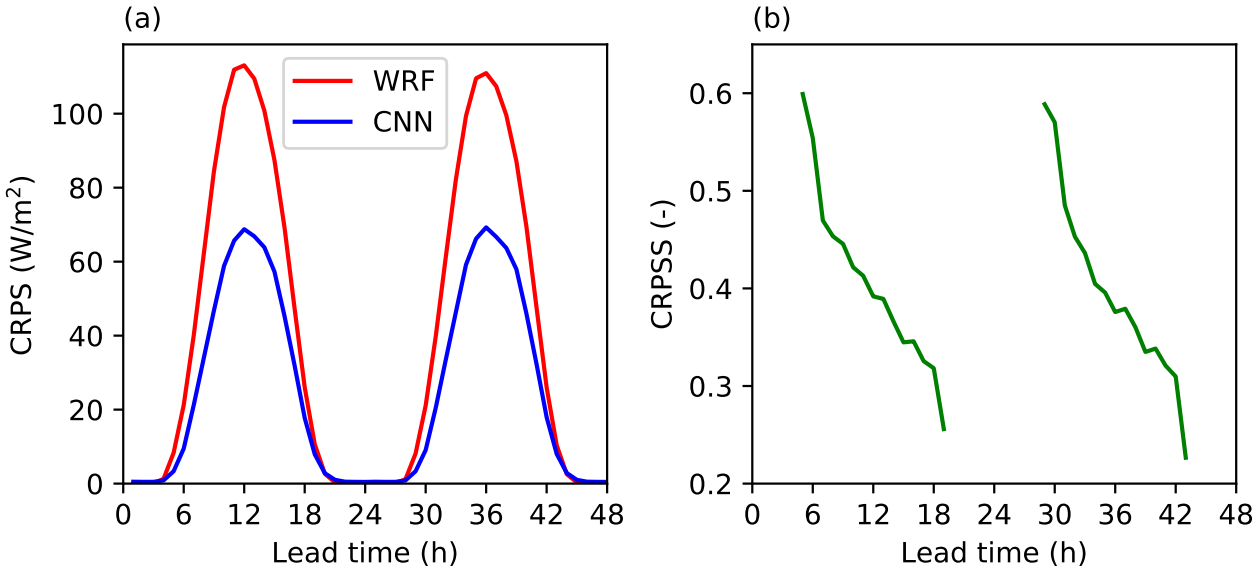


Figure 45: Evolution of CRPS for WRF (in red) and the CNN (in blue) in function of the lead time.

4.3.2 T1.2.3-1: Calibration of probabilistic wind ensemble predictions

Details on the method The finer horizontal resolution ($0.1° \times 0.1°$) and the larger domain extent give a larger dataset by lead time than for the solar radiation. Therefore, we chose to train a new model for each lead time (from lead time 24 to 48). Hence, we made 25 calibration models, each one trained on the whole period and the whole domain.

Results Like for the solar radiation forecast, the rank histogram associated to the wind speed forecast for WRF (figure 47a) has an asymmetric U-shape, with higher frequencies associated to the first rank (negative bias and marker of under-dispersion of the probabilistic forecast). After calibration (figure 47b), contrary to the post-processing of the GHI, there is no bell-shape and the histogram is nearly flat. This corresponds to the desired result, which indicates an improvement in the reliability of the probabilistic forecast.

The stratification of the CRPS by lead time demonstrates that there is an improvement for all the lead times (figure 48a). Moreover, in the CNN, there is a slight increase of CRPS throughout the day, which is the sign of a reduction of the forecast quality with the increase in forecast horizon. In WRF, this increase of

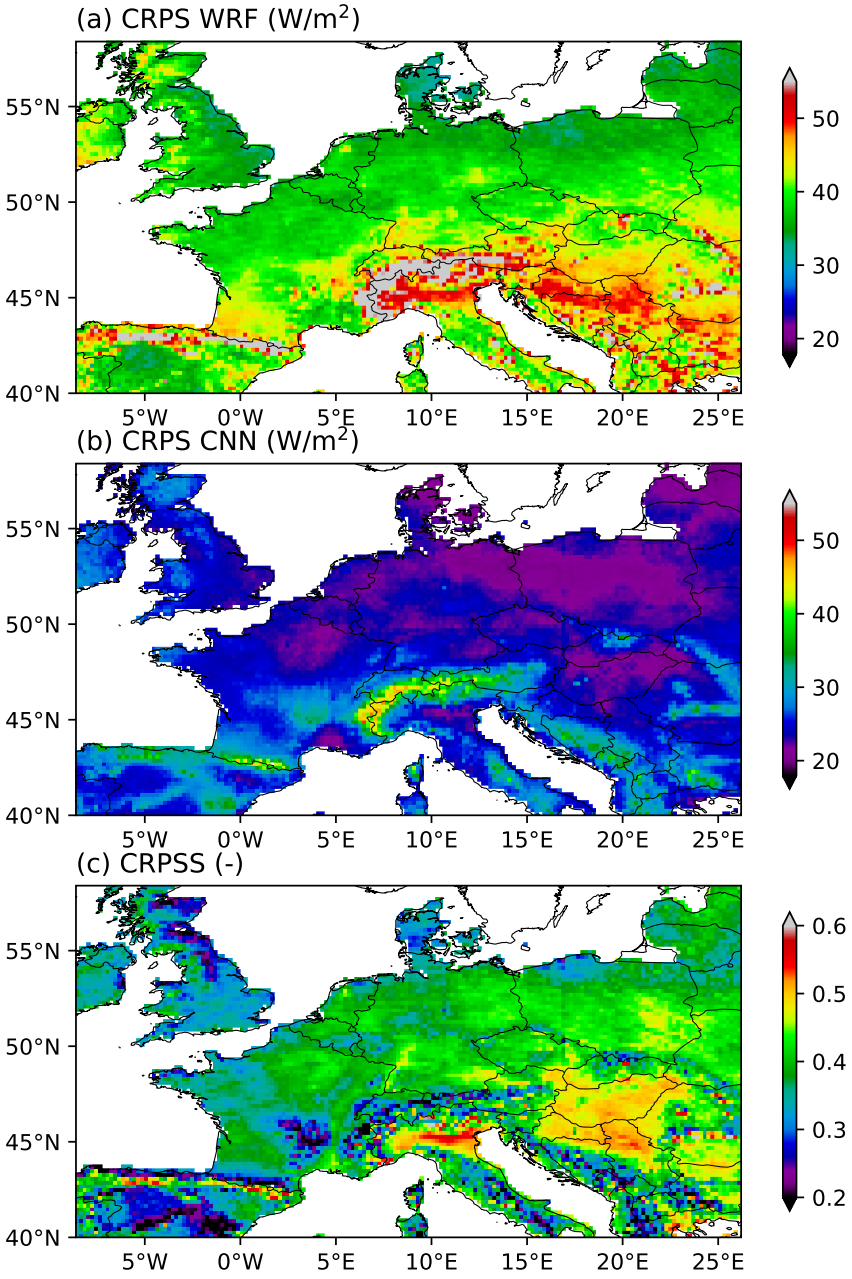


Figure 46: Comparison of CRPS values for WRF (a) and the CNN (b), and CRPSS for the relative improvement between the CNN and WRF (c).

CRPS is less visible because of a larger augmentation in the morning and then a decrease in the afternoon, possibly caused by a diurnal cycle in the wind speed forecast quality.

The spatial comparison based on the CRPS and the CRPSS (figure 49) shows that the calibration reduces the CRPS relative to the WRF forecast over a large part of the domain, the only exceptions being the Gulf of Lion and some parts of the Atlantic ocean, which is also visible on the CRPSS map (figure 49c calculated using WRF as reference in the equation 9).

In WRF, maxima of CRPS are observed over mountainous regions and over Mediterranean coastal regions as well as Norge coastal regions. In the CNN, the coastal maxima are still present, but mountainous regions do not exhibit maxima anymore. It results that maximum values of CRPSS are obtained over mountainous regions. Finally, higher CRPS and smaller CRPSS are generally obtained over land in

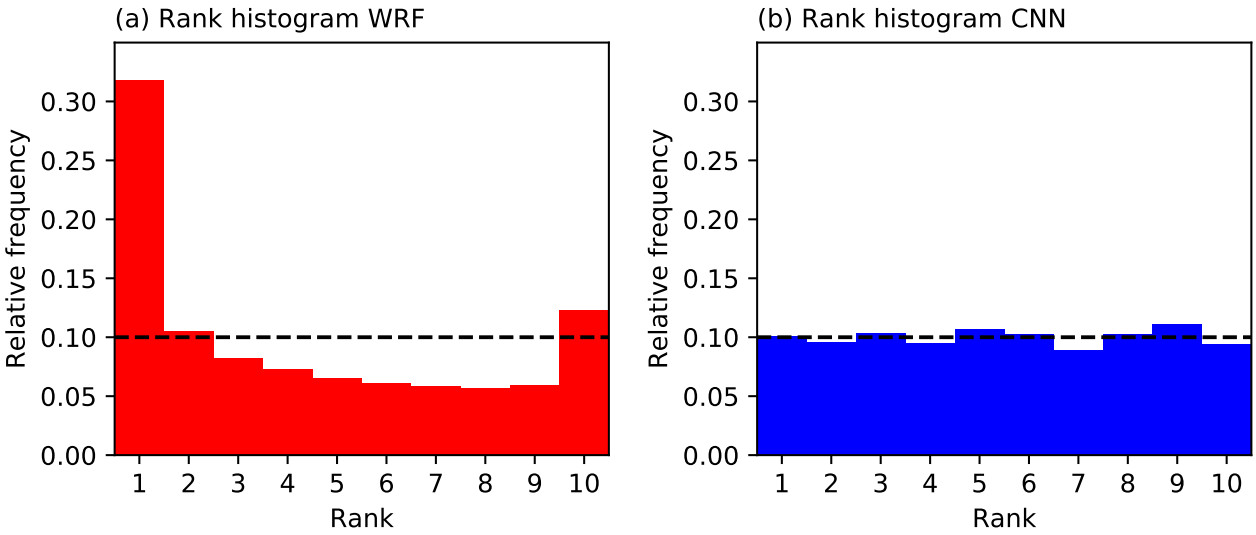


Figure 47: Rank histograms for the WRF (a) and CNN (b) probabilistic forecasts.

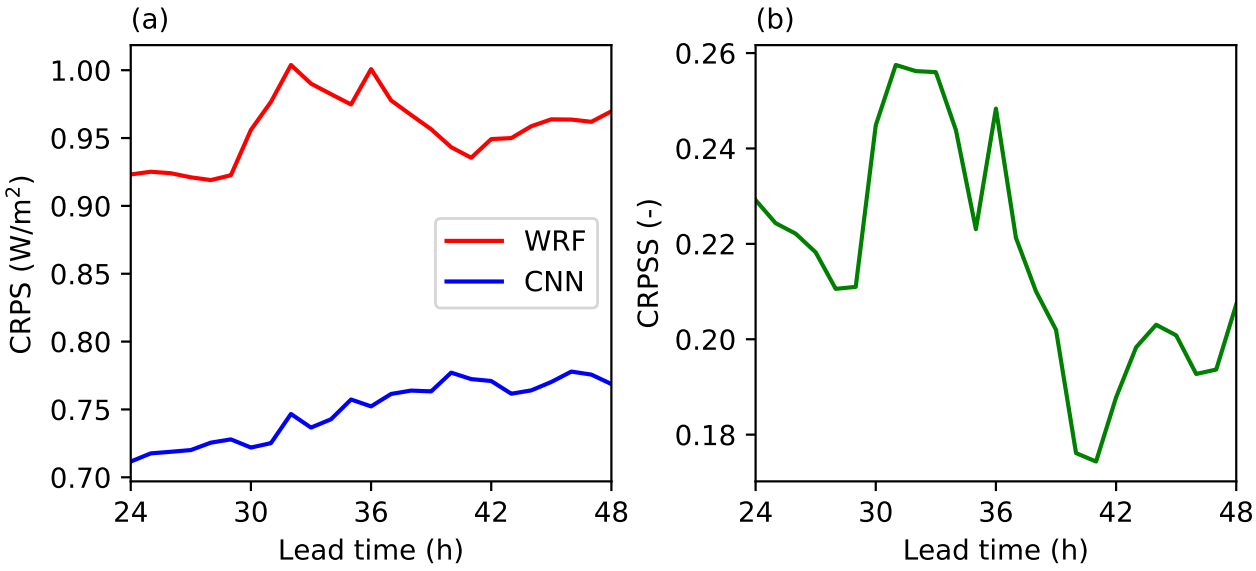


Figure 48: Evolution of CRPS for WRF (in red) and the CNN (in blue) in function of the lead time.

comparison with seas.

4.3.3 T1.2.3 Conclusions

We used convolutional neural networks in order to calibrate a 512-member ensemble of solar radiation and wind speed forecasts produced by ESIAS-Met with the WRF model. The CNNs were designed to describe the distributions of these parameters by providing a discrete subset of their quantiles through quantile regressions.

The raw forecast was biased and under-dispersed, both for the solar irradiation and the wind speed. After post-processing, the forecast is not biased anymore. The rank histograms suggest that the post-processed wind speed is well calibrated while the solar radiation is slightly over-dispersed, while being more reliable than the WRF forecast. However, both for the solar irradiance and wind speed, the CRPS was reduced for the whole domain and for all lead times.

It would be valuable to compare these results with post-processed forecasts resulting from other (and more traditional) methods. Note that concerning the GHI, the CNN reaches better performances than a

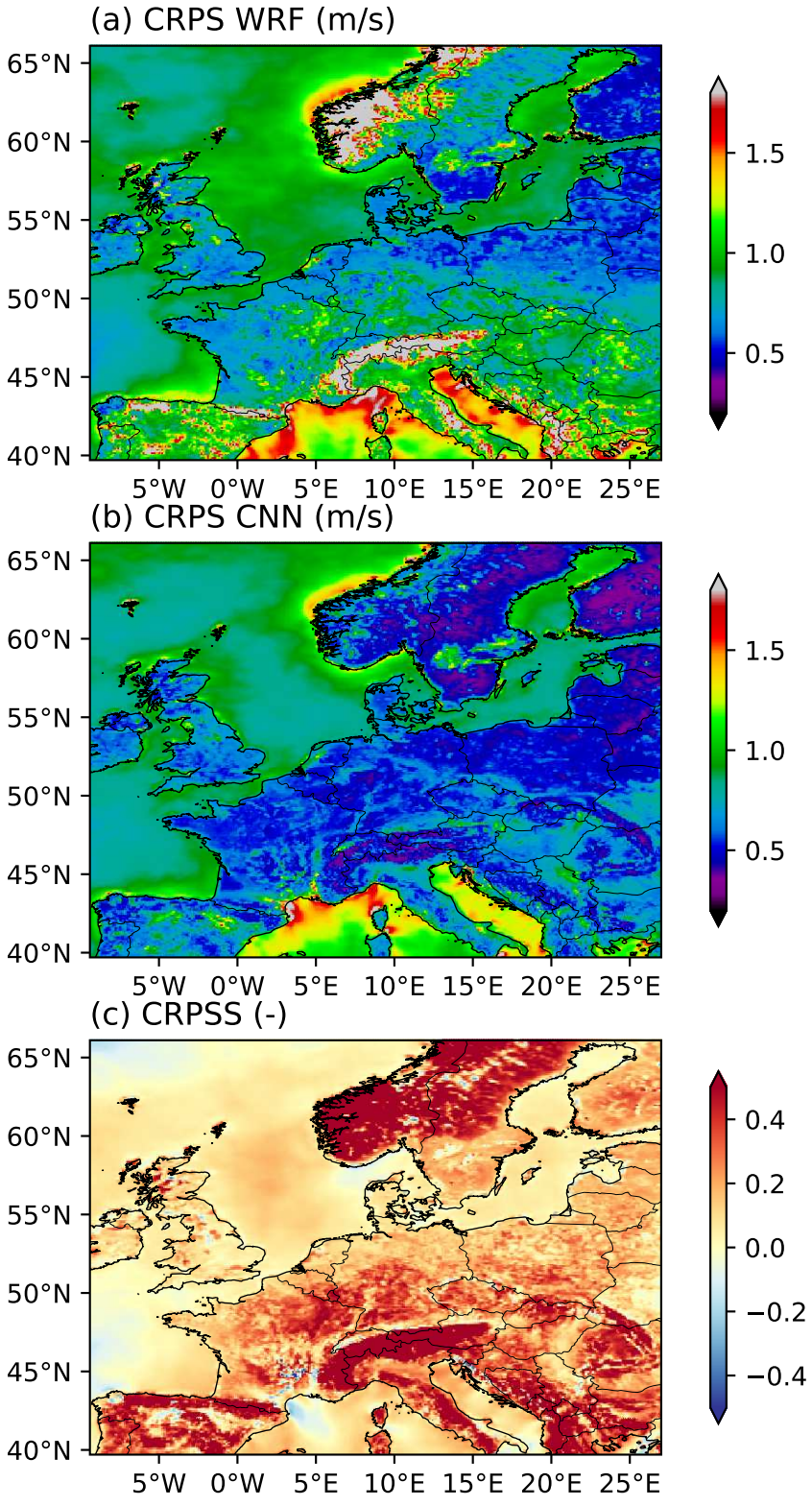


Figure 49: Comparison of CRPS values for WRF (a) and the CNN (b), and CRPSS for the relative improvement between the CNN and WRF (c).

simpler neural network, which is yet among the better methods for this task [39].

A complementary study was performed on the solar radiation using an other dataset (larger spatial and

temporal domain, less members) based on PEARP forecasts (ARPEGE ensemble forecasts), the results of which will be submited soon to the journal Solar Energy [39].

4.4 Code demonstrator

The demonstrator for the ESIAS framework for performing ensemble simulations was detailed in the mid-term D1.2 and is provided again below. This uses common KSH code for user-friendliness and quick deployment, compatible with different versions of the WRF preprocessing system for ESIAS-Met. The full instructions for installing ESIAS-Met and executing the components of the flagship codes are documented in the included README.md.

`https://gitlab.maisondelasimulation.fr/eocoe-ii/code-demonstrators`
Code demonstrators accompanying the D1.2 report of the EoCoE-II project.

A second demonstrator is available via the SaaS portal. This includes both a "toy domain" case for quick testing and the EU domain with a nested grid for Germany. The user can run simulations without difficulty by first building the environment with the included code to ensure optimal usage of the CPU resource for the given ensemble size. The two schochastic schemes, SKEBS and SPPT, and six different physics are all included in this demonstrator in the SaaS portal.

`https://eocoe.psnc.pl/`
Toy and EU domain demonstrators of the SaaS portal.

4.5 Summary

The scientific challenge Meteorology for Energy centered around the flagship code ESIAS for ultra-large ensembles of numerical weather predictions and their application to renewables forecasting. In EoCoE-II, ESIAS could be integrated with the MELISSA middle-ware to demonstrate ultra-large simulations as laid out in the roadmap. This framework could be used to perform an extensive sensitivity study on the performance of thousands of physics combinations in ESIAS-Met (WRF), finding the best combinations for the European domain. ESIAS-Met was also coupled to ESIAS-Chem and a study completed on its ensemble performance and the effect on irradiance. (A software analysis of the alternative model ICON could not be completed due to over a year hiring delay after the departure of the original Scientific Challenge leader and ESIAS staff).

The optimized ESIAS-Met configuration was used to generate a year of ensemble data for renewables forecasting of solar and wind feed-in. This representative data could be used to calibrate the solar and wind power models of electric feed-in in Germany to the WRF output. In the end, a similar quality to operational weather services was achieved with the research model, though some unexplained features in the irradiance modelling remained. This was also apparent in the non-parametric calibration of wind and solar point forecasts, which improved both probabilistic distributions, but was still more successful for wind than for solar.

Satellite observations of cloud positions and optical flow solutions of their motions and the underlying geostrophic structures could be used to validate ensemble members. For solar it was shown that ensembles of this size consinstently yielded a few solutions similar to the observation, which could be quite useful in the application of grid congestion.

The results highlight some divergent goals in the application of ultra-large ensembles and ideas for future work. Target outliers, or accuracy? Costs to large-scale energy markets due to extreme weather, or small-scale accuracy for grid operation? EoCoE-I was concerned with the former, but EoCoE-II showed how calibrating the meteorological outputs for regional [3] forecasts or performing particle filtering improved

[3] Calibrating quantiles for point forecasts as in T1.2.3 for e.g. plants is valid, as no spatial relationships from the ensembles are needed.

the accuracy of the ensemble mean only at the expense of the probabilistic quality of the ensemble distribution. (EoCoE-I used higher statistical moments of large-scale quantities like national feed-in to warn of extreme events. Such moments were mentioned in the EoCoE-II proposal but were in the end not relevant).

Future work can still address both sides. Not all events are captured, whether by a lack of variance in the modelling or uncertainty in the initial condition. Global ensemble models could be used to generate more variance in the initial condition, more accurate or not. The general improvement of ESIAS-Met for the domain would address issues of accuracy. There is also growing interest in "long-term" forecasts, which could certainly be a more dynamic application of ultra-large ensembles. (Regarding renewables, it must also be considered in the future how machine learning models of plant forecasts suppress outliers as a means to improve their cost-function optimization, contrary to our efforts on the meteorological side to capture them in the first place).

Regarding short-term accuracy and spatial variance, as opposed to rare weather events, a very practical and immediate implementation of ultra-large ensembles would be the very uncertain high-resolution details of clouds. Although unremarkable for the aggregate national feed-in or energy markets, these could have value on a daily basis for grid operation, especially as PV and smart metering dominate the low voltage grids of the future.

5 Materials for Energy (T1.3)

T1.3 is divided in three main tasks:

T1.3.1	Shedding light on carrier dynamics at hetero-interfaces in silicon solar cells
T1.3.2	Harvesting electricity from salinity or temperature gradients
T1.3.3	Organic and perovskite solar cells

T1.3.1 has the final objective of optimizing silicon solar cells to increase in performance and extension of lifetime. Amorphous-crystalline heterointerfaces play a crucial role in the photovoltaic operation of silicon heterojunction (SHJ) technology, but the microscopic mechanisms of transport and recombination mechanisms at the interface are still poorly understood. In the present task the transport mechanisms underlying photovoltaic devices based on SHJ technology have been investigated by simulating at atomistic resolution amorphous-crystalline heterointerfaces. Medium and large c-Si/a-Si:H interface models have been built up from classic molecular dynamics (MD) simulations and first-principles calculations. Ab initio electronic properties of the c-Si/a-Si:H interfaces have been calculated. Starting from the first-principles calculations, tight-binding Hamiltonians are represented in a basis of localized Wannier functions. Next, non-equilibrium Green's functions (NEGF) calculations have been performed using the LIBNEGF code to analyse the effect of interfaces on the carrier transport in silicon solar cells.

T1.3.2 focuses on optimizing capacitive blue energy electrodes and thermo-electrochemical devices. The first objective of this project has been to simulate the adsorption free energy of the lithium ion on the carbon surface in order to ascertain the best electrode structure which optimizes blue energy production. The second objective of this task has been to simulate a system in which the ionic liquid is confined between a pair of planar electrodes held at constant electrical potential, which fluctuates due to the thermal motion of the ions in the liquid to find compositions that will enable optimal performances of thermo-electrochemical devices. In both cases, a fundamental understanding of the cation and anion adsorption at the surface of the electrodes has been achieved. The challenge for simulation is that simulating the interfaces requires to rigorously account for the interactions between the electrode atoms and the adsorbed species. Due to the large size of the simulated systems for the final application, it is not possible to use electronic density functional theory (DFT) for such calculations. We therefore developed new force fields for classical molecular simulations. The parameterization of these force fields has been made based on a series of electronic DFT calculations. However, the commonly used exchange-correlation functionals may yield very different results for the adsorption energy of the molecules. We overcame this problem by performing

a series of QMC reference calculations in order to benchmark them on the adsorption energies. Once the DFT functional was benchmarked on the QMC reference, a large amount of calculations has been performed to fine-tune force fields for classical molecular simulations with Metalwalls/MDFT.
T1.3.3 deals with the development of a flexible and modular scheme for the multiscale modelling of electronic and ionic transport in materials for next generation photovoltaic devices. This has been achieved using a (augmented versions of) pre-existing, MPI parallelised Python framework, namely PPMD. The KMC simulator (FMM-KMC code) has been developed and applied to the study of doped organic semiconductors and to the simulation of all-solid-state batteries. The DMC simulator (BOLTMC code), which solves the Boltzmann transport equation, has been optimized and used to investigate the polaronic effects on scattering and mobilities of charge carriers in halide perovskites.

5.1 Task 1.3.1: Scientific results

The main task T1.3.1 is subdivided in three subtasks:

T1.3.1-1	**Classical Molecular Dynamics (MD) simulations of c-Si/a-Si:H interface**. We have used ReaxFF (Reactive Force Field) molecular dynamics to efficiently simulate the thermalisation, quenching, and equilibration processes involving thousands of atoms forming realistic a-Si:H/c-Si interface structure (large interface model), during up to ten nano-seconds at T=300, 500, 700 and 900 K. Snapshots of the equilibrated c-Si/a-Si:H interface atom configurations have been generated to be characterized in T1.3.1-2.
T1.3.1-2	**First-principles electronic properties of c-Si/a-Si:H interface**. A workflow has been used for the first-principles characterization of snapshots of classical MD simulations (T1.3.1-1) to monitor the electronic properties of the c-Si/a-Si:H interface. We used the PWscf code of the Quantum Espresso suite to perform the first-principles calculations. The evolution of the intragap states at different temperatures is monitored by analyzing projected and local density of states, electrostatic potential, and charge density. This all will allow to select the atomic configurations that will be used in the electronic transport model, in order to design more efficient silicon solar cells belonging to the silicon heterojunction technology.
T1.3.1-3	**Non-equilibrium Green's functions transport properties c-Si/a-Si:H interface**. In order to analyze the effect of interfaces on the carrier transport and dynamics in silicon solar cells we use the results from T1.3.1-2 as input for a quantum trasport code based on the NEGF formalism. To this end, we adopted a new code, libNEGF, in substitution to PVnegf, as explained in D1.2. This includes the treatment of the contact self-energies, the coupling to a solver for the self-consistent Poisson equation, and the inclusion of electron-photon and electron-phonon scattering processes via the corresponding self-energies. The input for these calculations is a tight-binding Hamiltonian that reproduces the ab initio electronic structure of the heterostructure of T1.3.1-2. For the time being we used the relaxed and quenched structured provided by the T1.3.1-1 and generate a tight-binding Hamiltonian using DFTB+. At the same time, we are in the process of representing such an Hamiltonian in the more appropriate basis of localized Wannier functions that constitutes a very natural and very accurate basis for extended bulk states. The application of the NEGF formalism to the large interfaces from T1.3.1-2 requires a high-parallel and scalable code. Therefore, the optimization of the libNEGF code will need to be carried on in parallel to this task in the WP2 Task 2.4.

T1.3.1-1: Classical Molecular Dynamics (MD) simulations of c-Si/a-Si:H interface

Methods. The two a-Si:H/c-Si interfaces are made of hydrogenated amorphous silicon (a-Si:H) between two crystalline silicon (c-Si) slabs. The relaxed p(2$\times$1) symmetric reconstruction of the Si(001) surface constitutes the c-Si side of the interface. It is formed by 576 Si atoms, 16 layers of silicon with 36

atoms each. The a-Si:H side of the system is generated by cutting the a-Si:H structure, built as in reference [30], such that the surface area is equal to the c-Si side and the thickness is about 16 Å. It is composed of 512 Si atoms and 64 H atoms. Periodic boundary conditions are imposed in all directions. The total length of the periodic cell is L_z = 46.44 Å, while in the x- and y-direction the system has $L_x = L_y$ = 23.22 Å. The described a-Si:H/c-Si interface structures, comprising 1,152 atoms, is indicated in the following as the large heterojunction (LHJ) model. A similar study has been performed in a full ab initio approach by some of us [31] for a smaller a-Si:H/c-Si interface formed by 336 atoms.

Starting from the DFT optimized LHJ model (Figure 50a), a subsequent classical MD analysis of the final ab initio configuration has been carried out by means of LAMMPS [121]) using the ReaxFF training set parametrization previously employed for the simulation of H bombardment of Si, Ge and SiGe (100) surfaces [122, 49]. This training set parametrization includes the dissociation of Si-Si bonds in the Si_2H_6 and Si_2H_4 molecules, therefore taking into account single and double silicon bonds, as well as Si-H bond dissociation of the SiH_4 molecule. The entire dissociation energy landscape is obtained by means of adiabatic energy DFT calculations vs. bond length and covers from the equilibrium distance to the dissociation limit. The energy dependence on valence and torsion angles in the Si-H force field is also incorporated. For example, for silicon bonds, the angular dependence is included in the form of the adiabatic DFT energy of the Si_3H_8 molecule as a function of the Si-Si-Si bond angle, and torsional terms are adapted by including energy differences between chair, boat, and planar conformations of c-$(SiH_2)_6$ six-member rings. Several other reactions are included, such as conversion of Si_2H_4 to H_3Si-SiH. Finally, the force field accurately takes into account also the cohesive energy and the equation-of-state DFT predictions for various silicon crystal phases, including the simple cubic Si, α-Si, and β-Si phases.

The complete MD analysis starts with T= 0 K minimization: first, fixing all cell dimensions (L_x= L_y= 23.22 Å and L_z= 2$\times$ L_x), and then, with L_x= L_y still fixed but with L_z varying as an additional degree of freedom. The resulting geometry (see Figure 50b) is used as the initial condition for a subsequent quenching-thermalisation process. Initially, the system is heated up to 1100 K at zero pressure with a NPT (Nose-Hoover thermostat and barostat) [70] for 325 ps (Figures 50c and 50d) and next it is cooled down to the desired final temperatures of 300, 500, 700 and 900 K in 325 ps. In Figure 50d the final configuration of the quenching process is shown. A final thermalisation procedure is applied during 10 ns keeping a constant temperature with a csld [25] thermostat to avoid the flying ice cube artifact. A constant 1 fs integration time step is employed in the 300 and 500 K cases whereas at 700 and 900 K a smaller 0.5 fs time step was used to maintain the numerical stability of the simulations. The pressure is controlled along the z coordinate exclusively, keeping L_x and L_y box sizes fixed and allowing L_z to evolve isobarically (P= 0).

Scientific results. The ab initio relaxed system with a double a-Si:H/c-Si interface of the LHJ model has been used as starting configuration for MD simulations. The minimization at T= 0 K produces a shrinking of the system along the z direction of about 6 Å. The quenching process, in which the temperature raised up to 1100 K and cooled down to the thermalisation temperature, has been repeated for each of the thermalisation temperatures of T= 300, 500, 700 and 900 K. We have observed a displacement of the hydrogen atoms towards the interfaces. This trend is maintained during each of the subsequent thermalisations of 10 ns (Figure 50). Indeed the concentration of hydrogen atoms near the interfaces remains almost constant for all temperatures. The full description of the scientific results of this sub-task, completed on M18, has been reported in the deliverable D1.2.

T1.3.1-2: First-principles electronic properties of c-Si/a-Si:H interface

Methods. We use the PWscf (Plane-Wave Self-Consistent Field) code of the Quantum ESPRESSO suite [58, 57] to relax the final snapshot of the a-Si:H/c-Si systems from the MD simulations. Si and H ultrasoft pseudopotentials available in the Quantum ESPRESSO library with Perdew-Burke-Ernzerhof (PBE) [119] GGA exchange-correlation potential. The electronic wave functions were expanded in a plane-wave basis set with a kinetic energy cut-off equal to 40 Ry (the charge density cut-off was 240 Ry). The Brillouin zone integration for the self-consistent calculation is restricted to the Γ-point, which is justified by the

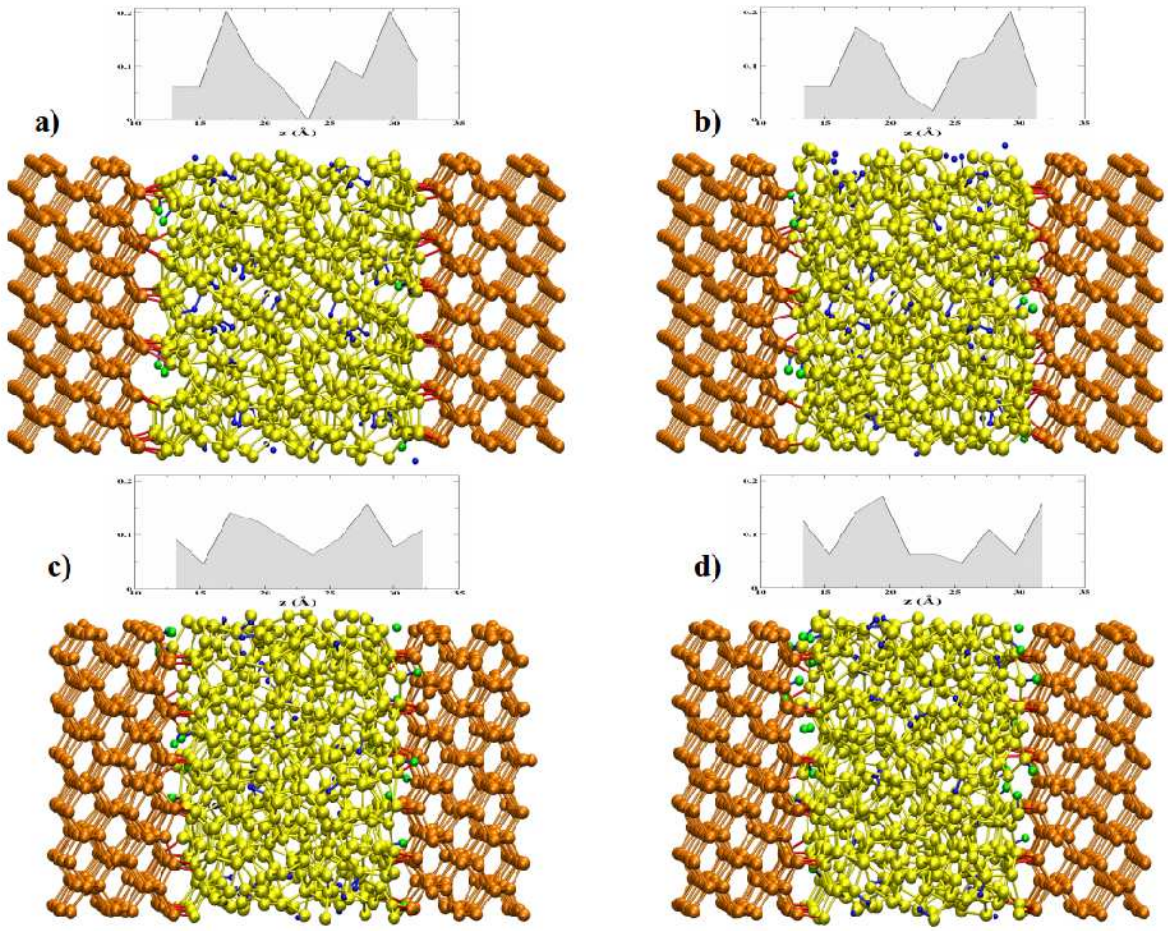


Figure 50: Snapshots of the large heterojunction (LHJ) model of the a-Si:H/c-Si interfaces. a) The ab initio relaxed system; b) the system at the end of the minimization at T= 0 K; c) the system in the middle of the quenching process at T = 1100 K; d) the system at the end of the quenching process. The silicon atoms and their bonds are shown in orange in the c-Si side and in yellow in the a-Si:H side, hydrogen atoms and bonds with silicon atoms are shown in blue. Bonds between c-Si and a-Si are shown in red and the hydrogen atoms close to the interfaces (with a distance less than 3.4 Å to the c-Si sides) are in green. Above the configurations, the distribution of the hydrogen atoms along the z direction are reported.

sufficiently large super cell. Gaussian smearing of 0.08 Ry is needed to reach convergence due to defect states at the Fermi level. All the parameters are chosen by checking the convergence of the total energy of the system. The energy minimization is performed by using the conjugate gradient (CG) minimization energy method, with the convergence threshold for self-consistency equal to 10^{-6} Ry. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-newton algorithm is used to perform geometry optimization. Ionic relaxation is stopped when the two following conditions are satisfied: energy changes less than 10^{-4} Ry between two consecutive self-consistent field (SCF) steps and all components of all forces are smaller than 5×10^{-4} Ry/Bohr. Each thermalisation process of the large heterojunction (LHJ) model (defined in the section T1.3.1-1) during 10 ns is followed by a final ab initio relaxation and the resulting structures are analyzed by ab initio calculations. The non-self-consistent calculation of the electronic states is performed on a $2\times2\times1$ k-point grid, which was found to yield a sufficiently accurate representation of the relevant quantities (density of states, electron localization function, and charge density).

We introduce an other atomistic model of the hydrogenated amorphous silicon (a-Si:H) between two crystalline silicon (c-Si) slabs, shown in Figure 51a. The relaxed p(2×1) symmetric reconstruction of the Si(001) surface constitutes the c-Si side of the interface. It is formed by 256 Si atoms, 16 layers of silicon with 16 atoms each. The a-Si:H side of the system is generated by cutting the a-Si:H structure, built as in reference [30], such that the surface area is equal to the c-Si side and the thickness is about 20 Å. It is composed of 256 Si atoms and 32 H atoms. Periodic boundary conditions are imposed in all directions. The total length of the periodic cell is L_z = 43.58 Å, while in the x- and y-direction the system has $L_x = L_y$ = 15.55 Å. The described the a-Si:H/c-Si interface structures constituted by 544 atoms is indicated in the following as the medium heterojunction (MHJ) model. The MHJ model will be also used in the following.

First-principles electronic properties calculations. We combined ReaxFF MD simulations and ab initio calculations to investigate the time evolution of the intra-gap states of a large a-Si:H/c-Si interface system. Therefore, the electronic structure is calculated and analyzed with a focus on the identification and characterization of the intragap states, which have a crucial impact on the device performance due to their role as recombination centers. Throughout the annealing process, we monitor the evolution of the relevant structural and electronic properties. We have found that at the end of the equilibration process of 10 ns of the Projected Density of States (PDOS) related to intragap states is decreased as well as the number of the electronic states into the gap. Nonetheless, the defects states are still localized both in the bulk of a-Si that at the interface with c-Si, until to be formed in the few c-Si layers nearest to a-Si:H. However, the system is quite ordered after the quenching process with a high percentage of fourfold coordinated Si atoms, and it keeps this condition during the final thermalisation. By increasing the thermalisation temperature the number of fourfold coordinated Si has a monotonous increase, with a decrease of the defects. We found that electron charge is accumulated along the c-Si/aSi:H interface while it is depleted from the nearby c-Si and a-Si:H surfaces. The two opposite pointing dipoles that are formed give rise to different barriers to the each type of carriers favoring the hole transport while hindering the transport of electrons. However, the energy barrier at the interface decreases by increasing the thermalisation temperature.

The last step of the thermalization process during 10 ns is followed by a final ab initio relaxation. The PDOS of each of the resulting structures are shown in Figure 52. We observe that the PDOS of the a-Si layers in the configuration before the start of the MD simulation is modified in different ways by changing the thermalisation process. The broad peaks in the band gap associated to a-Si before the MD simulation is started are modified by changing the temperature of the thermalisation process. Narrowed peaks are evident for T=300K, and increasing the T of thermalisation the number of peaks decreases. Increasing T to 700 and 900 K, the intensity of the intragap states decreases too. However, intense peaks are always found nearby the Fermi energy level at 0 eV.

It is useful to use an accurate approach to evaluate the atomic coordination of the relaxed heterostructures. We introduced a method in Ref. [31] to analyze the formation of covalent bonds based on the ab initio calculation of the electronic localization function (ELF). The method is based on the observation that the existence of a local attractor, i.e. a maximum of the ELF, between two atoms indicates the formation of the covalent bond. By comparing the values, it is possible to see that the 3-fold (5-fold) coordination is

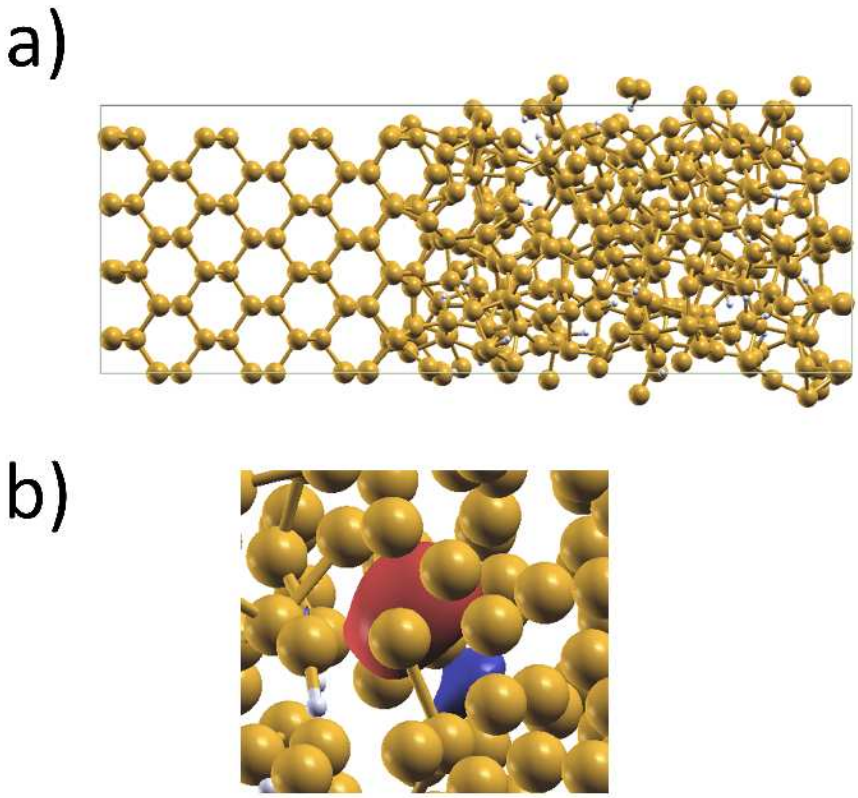


Figure 51: Medium heterojunction (MHJ) atomistic model of the a-Si:H/c-Si interfaces. a) The ab initio relaxed system; b) example of sp^3 maximally localized Wannier function.

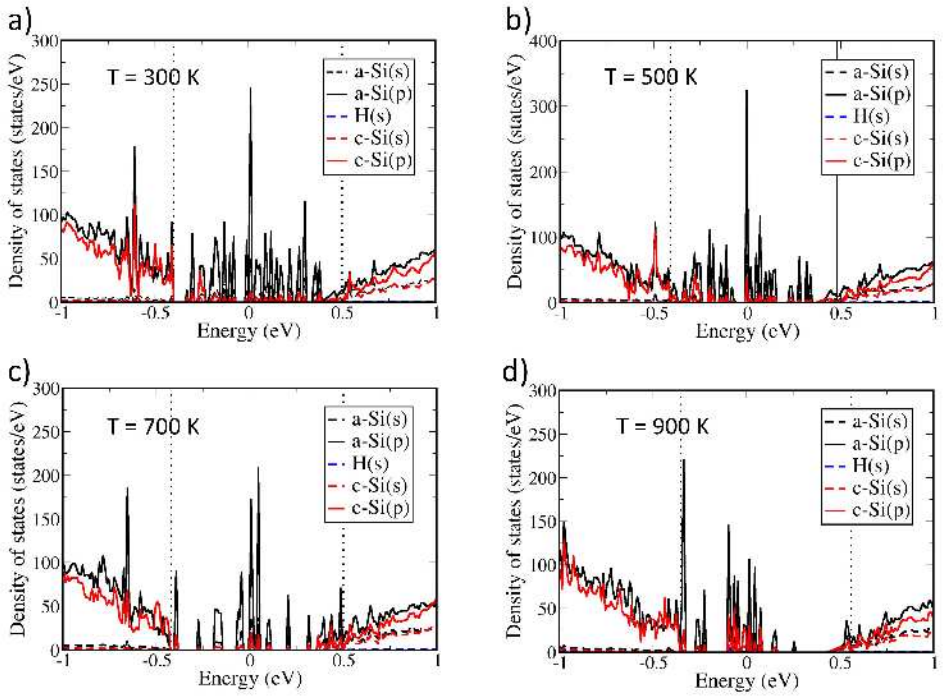


Figure 52: Projected density of states of the relaxed heterostructure after the thermalisation at a) T= 300K, b) T=500 K, c) T=700 K and d) T=900 K. The vertical dotted lines represent the band-gap edges. The Fermi energy is at 0 eV.

underestimated (overestimated) by using the cutoff method, but the discrepancy decreases by increasing the temperature.

The PDOS of 4-fold coordinated atoms close to the mobilty gap energy range are shown in in Figure 53 for the different thermalisations, where we take into account both the Si atoms in the amorphous region with Si4 environment and all the Si atoms with Si3H environment. The latter have very low weight when compared to the Si atoms with Si4 environment forming distorted bonds. However by increasing the T of the thermalisation the intensity of the Si4 states around the mid-gap decreases while their PDOS around the band-gap edges increases. On the other hand, the dangling and floating bonds defects become dominant in the mid-gap for high T of thermalisation. Therefore, the common view [146] where band tail states are mainly distorted tetrahedral bonds and mid-gap states are mainly under- and over-coordinated defects is confirmed in our simulation only increasing the T of thermalisation.

The coordination in the different regions of the heterostructure after the relaxation as calculated with the ELF analysis are reported in Table 14, and the spatial distribution of the defects featuring dangling and floating bonds is plotted in Figure 54. The number of dangling and floating bonds distributed in the a-Si region decreases by increasing the temperature of thermalization. Simultaneously, the number of 4-fold coordinated a-Si atoms increases. Indeed, the 64 H atoms diffuse in the a-Si and to the interfaces to favor the creation of H bond defects. We found that the number of 4-fold coordinated Si bound to one H atom (no di-hydrides have been found) is 60, 63, 60 and 60 at 300, 500, 700 and 900 K, respectively. Therefore, H plays an important role in the passivation of the dangling bonds. H diffuses to the c-Si region to form floating bond 4Si+H coordinated atom at 900 K. Just one 2-coordinated dangling bond is found in the a-Si region, not far from the interface, that is removed by increasing the temperature of thermalisation above 500K.

The PDOS on each of the 2-, 3-, and 5-coordinated defect in Figure 55 shows that the stronger contribution is from 3-fold coordinated Si around the mid-gap. Two-coordinated Si states are in the gap and

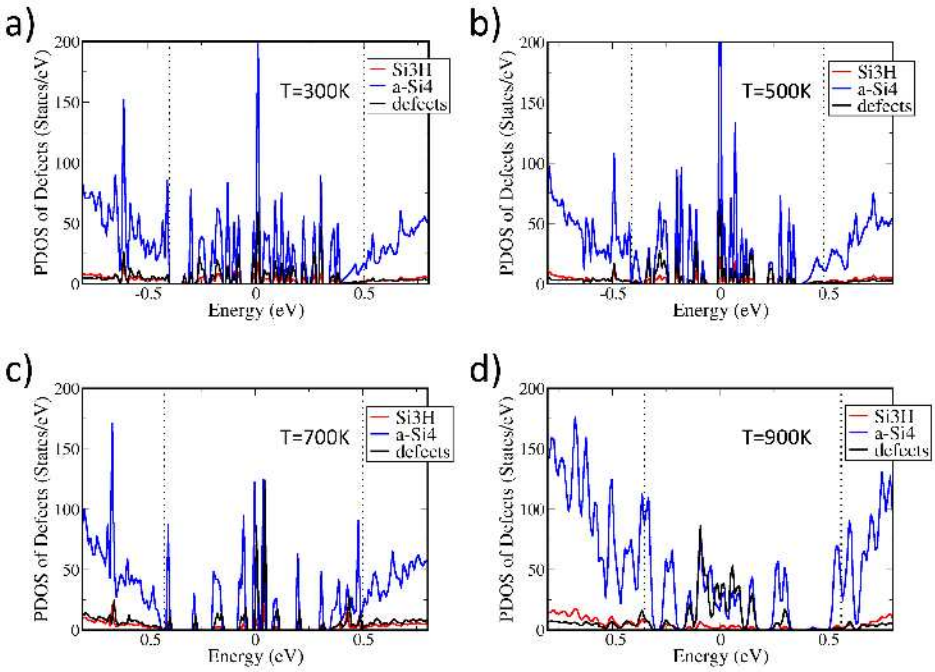


Figure 53: Projected density of states of 4-coordinated Si atoms (a-Si4 and Si3H), as well as dangling and floating bond defects for thermalisation at a) T=300K, b) T=500K, c) T=700K and d) T=900K. The vertical dotted lines represent the band-gap edges. The Fermi energy is at 0 eV.

Table 14: Coordination computed with the ELF method in a-Si and c-Si of the relaxed heterostructure after thermalisation. The number in brackets gives the percentage in the respective region.

Coord.	Environment	300K	500K	700K	900K	region
3	Si_3	1 (0.17)	1 (0.17)	1 (0.17)	1 (0.17)	c-Si
3	Si_2H	0	0	0	0	c-Si
4	Si_4	571 (99.13)	570 (98.96)	573 (99.48)	568 (98.61)	c-Si
4	Si_3H	4 (0.69)	5 (0.87)	2 (0.35)	6 (1.04)	c-Si
5	Si_5	0	0	0	0	c-Si
5	Si_4H	0	0	0	1 (0.17)	c-Si
2	Si_2	1 (0.20)	1 (0.20)	0	0	a-Si
3	Si_3	13 (2.54)	10 (1.95)	8 (1.56)	6 (1.17)	a-Si
3	Si_2H	0	0	0	0	a-Si
4	Si_4	429 (83.79)	435 (84.96)	438 (85.55)	446 (87.11)	a-Si
4	Si_3H	56 (10.94)	58 (11.33)	58 (11.33)	54 (10.55)	a-Si
5	Si_5	10 (1.95)	6 (1.17)	6 (1.17)	5 (0.98)	a-Si
5	Si_4H	3 (0.59)	2 (0.39)	2 (0.39)	1 (0.20)	a-Si

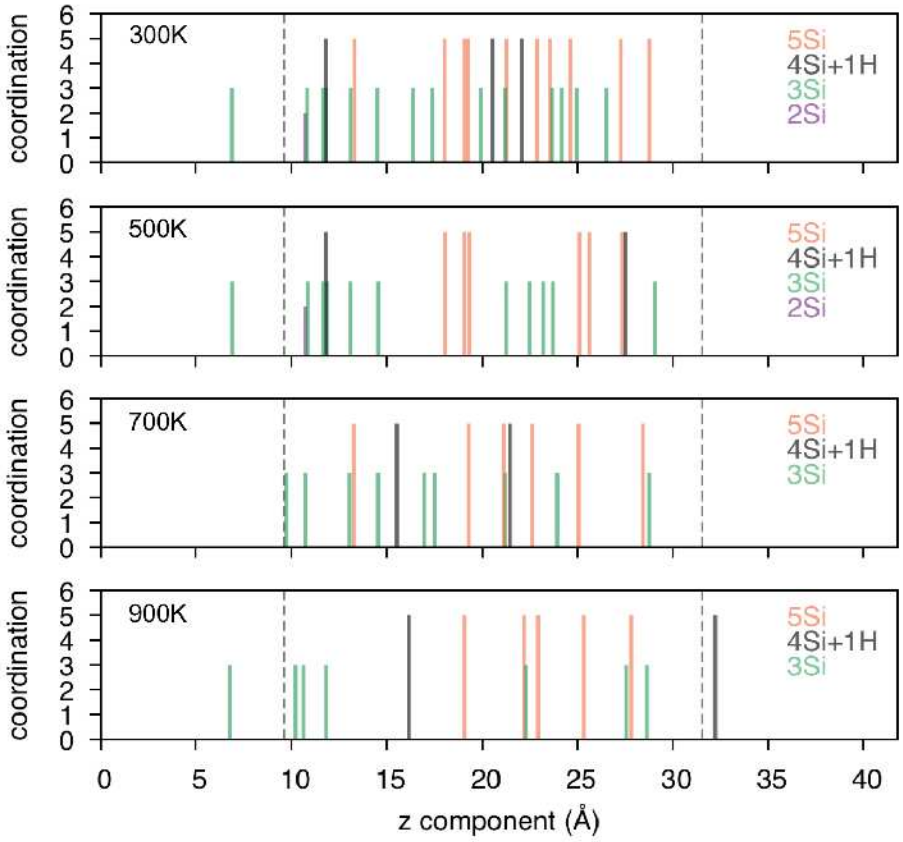


Figure 54: Spatial distribution of the 2-, 3-, and 5-coordinated defects across the interface. The coordination is shown both by the y axis and by the color scheme for comparison with Fig. 55. The different panels show results for 300, 500, 700, and 900 K (top to bottom). The vertical dashed lines demarcate the a-Si region.

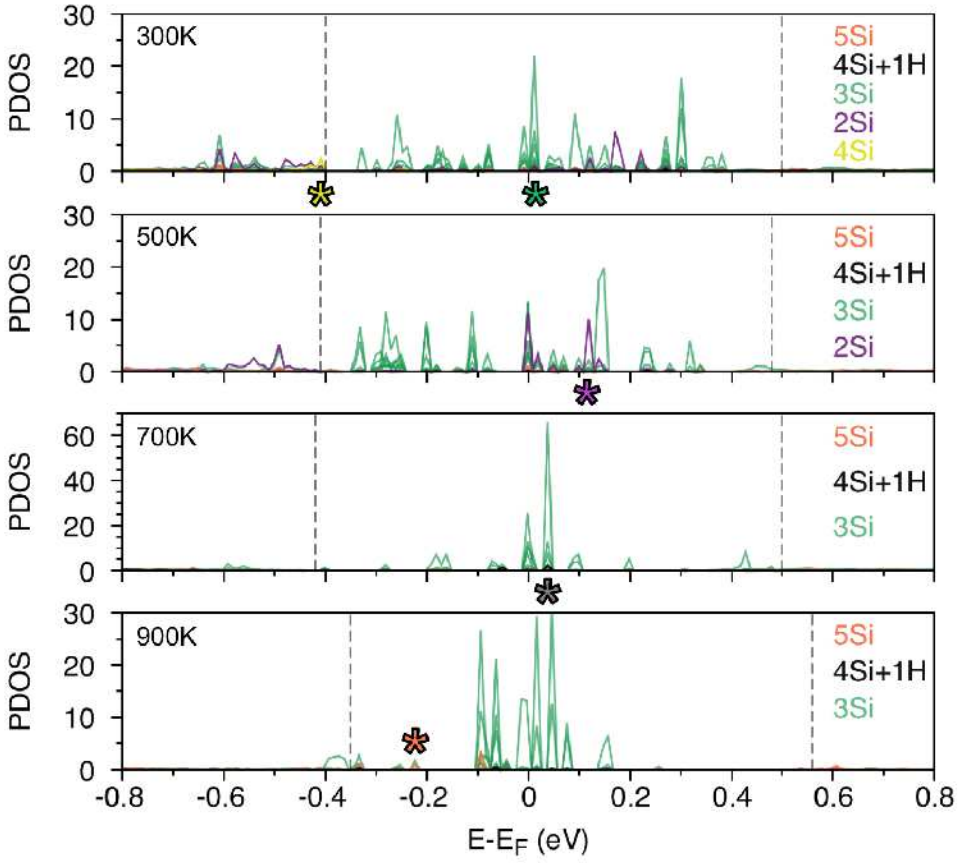


Figure 55: Projected density of states on each defect atom for 2-, 3-, and 5-coordinated defects. The different panels show results for 300, 500, 700, and 900 K (top to bottom). The vertical dashed lines represent the band-gap edges. The asterisks mark the states for which the LDOS and geometry is represented in Fig. 56. For 300 K, we also show one example of a 4-coordinated defect (yellow) for which we plot the LDOS in Fig. 56.

below the valence band edge at T= 300 and 500 K. On the other hand, there is no evidence of floating bonds unless one peak barely visible at -0.22 eV at T=900 K.

In order to elucidate the spatial localization we calculated the local DOS (LDOS) of the intragap energy levels. This analysis allows to visualize both the localization and the weight of the states in the gap. We found that high value LDOS isosurfaces are close to distorted 4-fold coordinated Si and dangling bonds. The LDOS analysis confirms that the latest are more important in the mid-gap for the T= 700 and 900 K cases. Double dangling bonds related to 2-fold coordinated Si are the ones contributing the most to the valence band tail and mid-gap levels at low temperature, while floating bonds have low weight. Examples of LDOS and geometrical parameters for each type of defect are shown in Figure 56. Three-fold coordinated defects are forming distorted trigonal pyramid. Five-fold coordinated defects lead to trigonal bipyramid-like configuration with slight deviations from the 90° and 120° bond angles of the regular trigonal bipyramid. Distorted 4-fold coordinated atoms have strong deviations from the the average bond length and angle of 2.37 Å and 109.1°, respectively, in the c-Si region.

Maximally Localized Wannier Functions. The localized Wannier functions constitute a very natural and very accurate basis for extended bulk states. We used the wannier90 code (http://www.wannier.org/) to compute the maximally-localised Wannier functions (MLWF) following the method of Marzari and Vanderbilt [98]. For entangled energy bands, the method of Souza, Marzari and Vanderbilt [145] is used. Full details can be found in Ref. [106].

First-principles codes typically solve the electronic structure of periodic materials in terms of Bloch

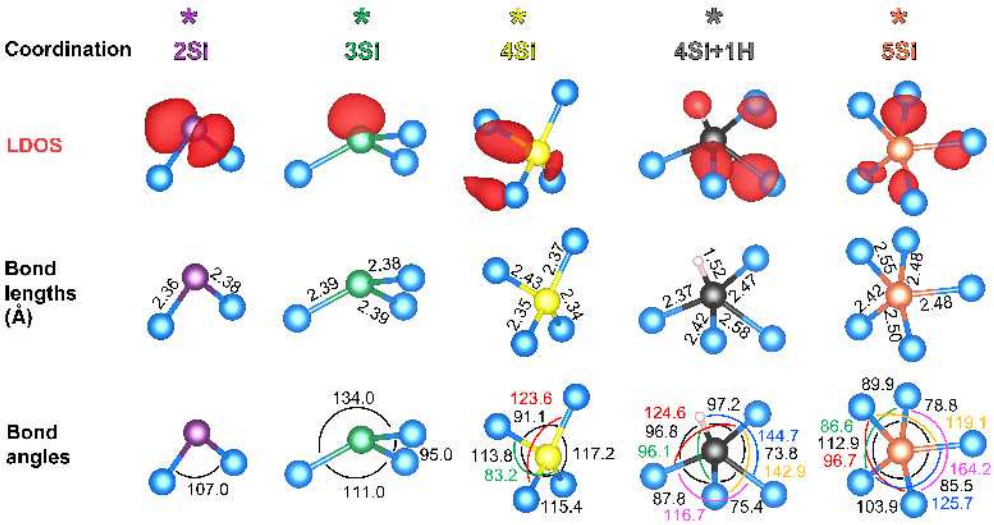


Figure 56: Example LDOS and geometrical parameters for each type of defect. The examples correspond to the states marked with asterisks in the PDOS panels of Fig. 55.

states. An alternative representation can be given in terms of spatially localised functions known as Wannier functions (WFs). The WFs are centred on each lattice site and are a unitary transformation of the set of Bloch states. The code wannier90 has been used in combination with the Quantum ESPRESSO electronic code based on plane-waves and pseudopotentials (PP). The use of norm-conserving PP turned out to produce more localized WFs and also allowed faster computations compared to ultrasoft PP as a very time consuming projection step was avoided.

In order to obtain MLWF for entangled energy bands we use the “disentanglement” procedure introduced in Ref. [145]. We define an energy window (the “outer window”). At a given k-point k, N(k) states lie within this energy window. We obtain a set of N Bloch states by performing a unitary transformation amongst the Bloch states which fall within the energy window at each k-point. The unitary transformation is obtained by minimising the gauge invariant spread within the outer energy window. The Marzari-Vanderbilt procedure [98] can then be used to minimise and hence obtain MLWF for this optimal subspace. It should be noted that the energy bands of this optimal subspace may not correspond to any of the original energy bands (due to mixing between states). In order to preserve exactly the properties of a system in a given energy range (e.g., around the Fermi level) we introduce a second energy window. States lying within this inner, or “frozen”, energy window are included unchanged in the optimal subspace.

We applied the procedure to calculate the MLWF to the LHJ model of the hydrogenated amorphous silicon (a-Si:H) between two crystalline silicon (c-Si) slabs, formed by 1152 atoms. The number of WFs to be calculated is equal to the number of electrons, i.e. 4,416.

In order to have suitable WFs for transport calculations to have spreads as low as possible is highly desirable. For c-Si, 1.4-2.4 Ang^2 is a satisfying range. By setting the number of bands in such a way that we have 4 bands/Si atom (and 1 band/H atom), i.e., 4,416 bands results in an average spread of 4.8 Ang^2 that is too large. Indeed, 4 bands/Si atom are not sufficient to get suitable WFs, because the disentanglement requires additional bands. This enlarges the energy window requested. In this respect, according to the tests we made on c-Si periodic cells from 2 to 64 atoms, and a-Si:H periodic cell of 72 atoms, 6 bands/Si atom seems a good choice. Indeed, by setting 6 bands/Si atom (1.5 bands/H atom) we have calculated an average spread for the test systems in the range of 2.1-2.2 Ang^2, that is satisfying. This choice leads to set the calculation of 6,624 bands. Unfortunately, we have found that this setting is quite challenging for the Quantum ESPRESSO code, and it will become feasible only in some future, once the code will be optimized and improved in parallel scalability. Since Quantum Espresso is under the competence of the MaX-II CoE, a Memorandum of Understanding between EoCoE-II and MaX-II has been subscribed on the development of the code optimization. By means of this collaboration, the wannier90 code performance

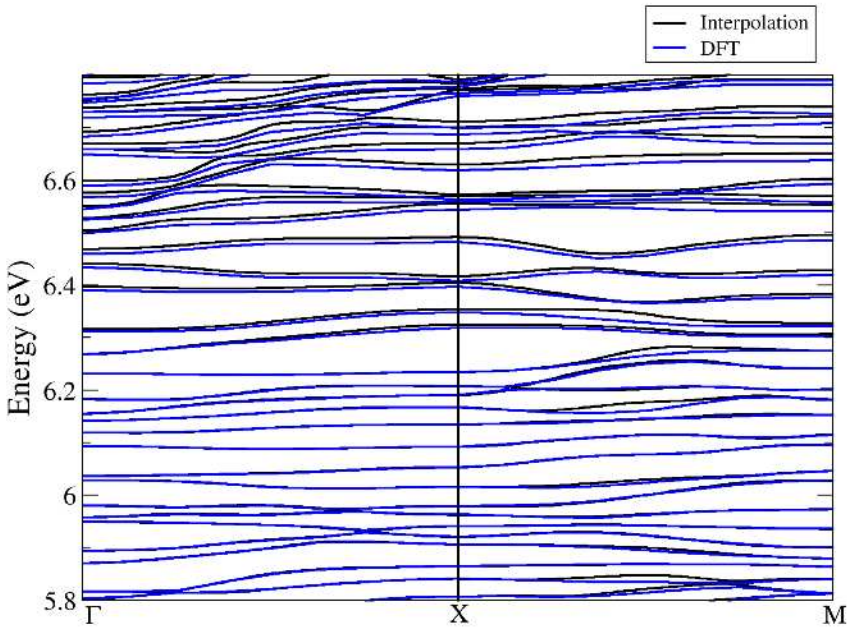


Figure 57: Electronic band structure calculated within the density functional theory and bands interpolation using the maximally localized Wannier functions basis set.

has been improved enough to make possible the calculations for the MHJ atomistic model.

Therefore, we used the MHJ model of the hydrogenated amorphous silicon (a-Si:H) between two crystalline silicon (c-Si) slabs, formed by 544 atoms. The number of WFs to be calculated is equal to the number of electrons, i.e. 2,080. Setting 6 bands/Si atom we have a number of bands equals to 3,120 that leads to an average spread of 2.8 Ang^2, that is acceptable.

We set the guess of the projections to four sp^3 orbitals for each Si atom and one s orbital for each H in the unit cell. The maximum energy for the frozen window was set just above the highest occupied energy level and the maximum energy for the disentanglement to an energy large enough so as to contain enough bands for each k point; 15.6 eV was a reasonable value.

In figure 57 we show the band structure calculated by the DFT method and the bands interpolation in the MLWF basis set. The very good agreement has been obtained after the application of 600 disentanglement iterations and 500 maximization of localization steps.

In figure 51b an example of sp^3 orbital of MLWF is shown. The orbital is in the expected shape.

Next, the Hamiltonian matrix in the calculated MLWF basis set of Wannier functions has been generated and it can be used to apply the NEGF formalism to the heterojunctions.

In conclusion, we combined ReaxFF MD simulations and ab initio calculations to investigate the intragap states of the a-Si:H/c-Si interfaces at different annealing temperatures. The distributions of the microscopic defects in the energy gap as well as along the heterostructure have been investigated. The intragap states play a key role, indeed they have a crucial impact on the device performance due to their role as recombination centers. We have found that the annealing process at 900K reduces the defects in the a-Si energy gap. The localized WFs constitute a very natural and very accurate basis for extended bulk states. We represent the Hamiltonian in the more appropriate basis set of WFs in such a way that it is possible to apply the NEGF formalism to the heterojunctions. The localization of the WFs and the interpolation of the DFT bands demonstrates that the maximally localized WFs calculated with wannier90 are a good representation of the electronic structure.

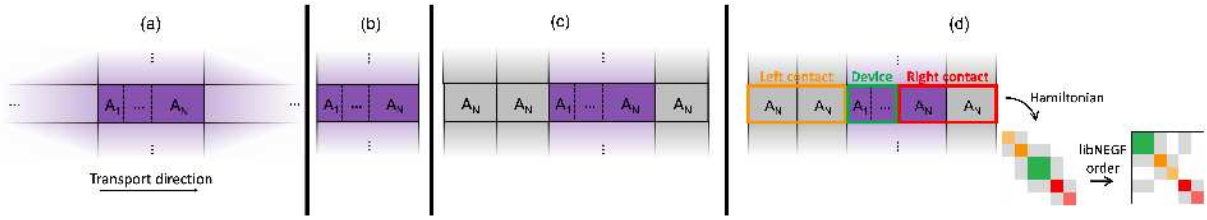


Figure 58: Schematic of the procedure of the WANTRANS interface.

T1.3.1-3: Non-equilibrium Green's functions transport properties c-Si/a-Si:H interface

Transport calculations using fully-ab initio Wannier tight binding. To make possible future transport calculations of the c-Si/a-Si:H interface made fully ab initio, the WANTRANS code has been implemented to interface the transport code libNEGF and the results from a DFT/wannier90 calculation. WANTRANS takes a *bulk* tight-binding Hamiltonian in the format of the wannier90 code and constructs a *device* Hamiltonian that can be used by the libNEGF library for a transport calculation.

The starting point is an ab initio calculation of a bulk system. The unit cell for this calculation [dark purple in Fig. 58(a)] contains all the interfaces for which transport properties will be calculated. This unit cell can contain domains of several materials, materials A_1 to A_N, where material A_N will form the contacts. The ab initio calculation has to be performed for a bulk system (periodic in all three dimensions).

WANTRANS first removes the periodicity in the transport direction [Fig. 58(b)], while keeping the periodicity in the transversal directions, i.e., perpendicular to transport. The region of material A_N from the unit cell is used as the 1st principal layer (PL) of the right contact and it is duplicated to construct the 2nd PL of the right contact and the two PLs of the left contact [Fig. 58(c)]. The device is then formed by the domains of materials A_1 to A_{N-1} and a Hamiltonian for device and contacts is constructed [Fig. 58(d)]. The periodicity in the direction perpendicular to transport is maintained via the dependence of the Hamiltonian on the real-space Wigner-Seitz grid-points used in the Wannier interpolation technique.

LibNEGF requires the Hamiltonian to be constructed in a specific order for the transport calculations: Device + 1st PL of the left contact + 2nd PL of the left contact + 1st PL of the right contact + 2nd PL of the right contact. Therefore, WANTRANS does not construct the Hamiltonian in the form of the left inset of Fig. 58(d), but rather in the form of the right inset. Starting from the 4 PLs of the contacts, libNEGF can later construct the infinite contacts with the iterative Green's functions method.

To test the implementation of the WANTRANS interface, we have constructed a heterostructure between GaAs and AlAs (Fig. 59). The unit cell of the calculation consists of 48 atoms of AlAs sandwiched between two layers of GaAs (each of them with 48 atoms). The GaAs layer on the right is used to construct the semi-infinite contacts. Using WANTRANS and libNEGF we are able to perform a transport calculation for this system. The results for the transmission in this interface is shown in the middle bottom panel in Fig. 59. The other two panels represent the transmission in the infinite bulk GaAs (left) and AlAs (right) for $k_\perp = 0$, where $k_\perp$ represents a k-vector in the plane perpendicular to transport direction (in some literature this is referred as $k_\parallel$, meaning that the k-vector is parallel to the heterostructure planes). As expected, the transmission shows the integer values of the conductance steps.

However, when trying to apply this method to the c-Si/a-Si:H interface, we faced another challenge. To construct the off-diagonal blocks (gray) in the insets of Fig. 58(d), the contact PL needs to be internally organized in equivalent layers. In terms of atomic positions, this can be easily done, and one can define layers of atoms that should be equivalent inside the c-Si PL. But we have realized that the Wannier functions of equivalent atoms are not necessarily equivalent. This is because in the maximally-localization procedure in wannier90, Wannier functions of one atom can acquire a phase different from that of an equivalent atom. In practice this means that matrix elements that should be identical or very similar might even have a different sign. This gives rise to unphysical results in the construction of the crystalline contacts. We are considering two solutions for this issue. Either new developments need to be made in the wannier90 code,

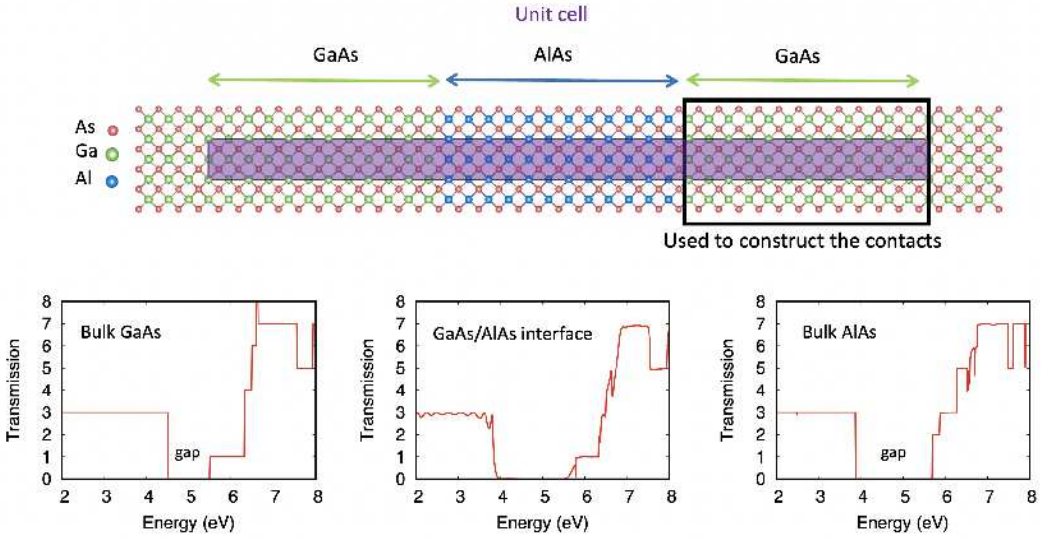


Figure 59: GaAs/AlAs heterostructure calculated with WANTRANS+libNEGF.

or WANTRANS needs to make an automatic search of Wannier functions that should be equivalent and are not. In the latter case, WANTRANS could then change the sign of the corresponding matrix elements accordingly.

Transport calculations using Density Functional Tight Binding. Since WANTRANS interface needs some further development before to be applied to the c-Si/a-Si/c-Si structure, we resolved to compute the transport properties across the heterojunction system using the density-functional tight binding semi-empirical approach (DFTB). DFTB constructs the system Hamiltonian in terms of 2-center matrix elements, starting from rigorous all-electrons atomic calculations. Contrary to empirical TB, where matrix elements are typically fitted in order to reproduce bulk bandstructures and becomes unreliable when transferred to different materials, the DFTB approach give in general less precise bandstructures (due to less adjustment freedom), but the parameters are transferrable across allotropes. The set of interactions between Si, O and H atoms has been constructed paying a special attention to reproduce the c-Si bands using sp^3d^5 orbitals, corresponding to 9 orbitals per atoms.

Furthermore, the dftb+/libNEGF interface is at a mature state of development and this allows to perform transport calculations with a reliable exisiting infrastructure able to handle a complex calculation.

Transmission calculations. The four relaxed geometries at the temperatures of 300 K, 500 K, 700 K and 900 K studied in the previous sections were also analyzed in order to compute the transmission across the a-Si layer. The structures where first prepared for transport by replicating the existing Si layers on both sides and partitioning the central region into 4 layers in order to exploit the recursive algorithm discussed in D2.3.

The results of these calculations are shown in Fig. 60 with corresponding density of states shown in the same Figure. Here it is possible to see how the gap defects of the a-Si introduce features in the transmission in the form of peaks corresponding to defect states. As a reference we have reported the transmission of the ideal cristalline Si with a black dashed line. This is also useful in order to identify the band-edges and gap in the range -4.2 to -3.1 eV. We observe that the transmission within the energy gap should be rigorously 0, but here a finite small value is the result of the numerically finite value of $\delta \approx 10^{-}5$ Hartree, set in defining the Green's functions. The meaning of this finite transmission becomes relevant when a bias is imposed across the junction and a net current can develop also assisted by the defect states.

Unfortunately, because of the different underlying approach, it is not possible to correlate 1:1 the peaks shown in Figures 52 and 53 with those of the transmissions shown here. Only a qualitative comparison

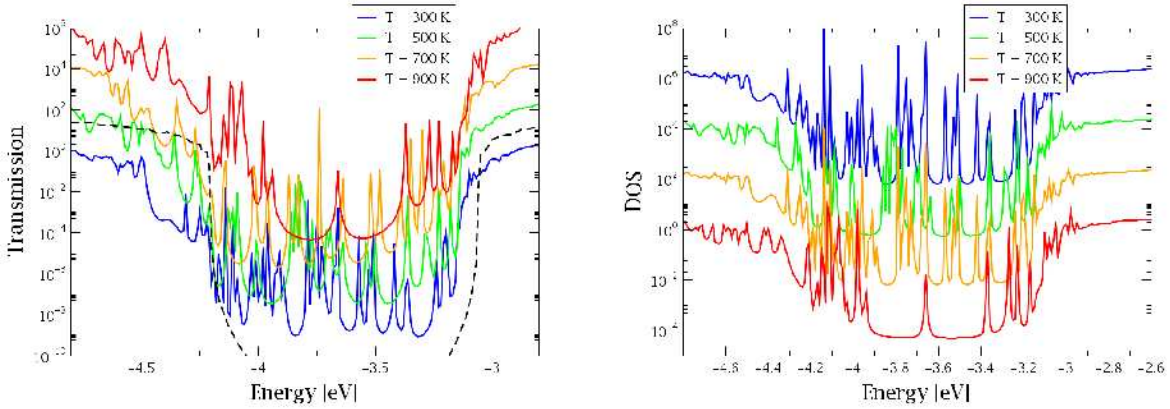


Figure 60: (*left)* Transmissions across structures annealed at 300 K, 500 K, 700 K and 900 K. The black dashed line refers to the ideal Si crystal. (*right*) Density of states for the same annealing temperatures. A constant shift factor of 100 is applied to the curves for clarity.

between the DFTB and DFT can be made. Comparing the gap defects with Figure 52, we observe a certain degree of similarity, especially looking at the structure annealed at 300 K, where in both cases the density of gap defect spreads almost uniformily within the gap. Increasing annealing temperatures, the defect states within the gap progressively decrease and move near the band-edges or within the bands. Perhaps the most evident difference between the two approaches is observed for the case of T=900K, where an isolated peak is found at mid-gap in the DFTB DOS, compared to the increase of mid-gap DOS found in the DFT calculations. Analyzing the special resolution of the projected DOS, shown in Figure 61, we conclude that the defect states originates from the bulk of the a-Si and not at the interfaces. Looking closely at the PDOS we also see that the 3-fold coordinated Si give the largest contribution to gap states, quite consistently to the DFT calculations.

The 2.5 nm long a-Si layer reduces the transmission by more than 4 orders of magnitude near the valence band (VB) edge and about 1 order of magnitude near the conduction band (CB) edge. A reduction is expected due to the disorder of the a-Si layer, however the value can be hardly obtained using a simple effective medium model.

The transmission functions comprise 480 energy points and a Mokhorst-Pack mesh of 4x4 k-points (8 points in total). The calculation is parallelized on 480 nodes running 8 MPI tasks each and 6 OpenMP threads for a total of 3840 processes and 23040 cores employed. The wall time clock of each calculation is 16 mins, corresponding to a total of 6144 core-hours.

The developments on libNEGF that made this possible has been carried out in the EoCoE-II project and are all explained on D2.3. Before EoCoE-II the calculation was not parallelized across the k-grid (hence a factor 8 slower) and the openMP performance was bad because of the usage of non-threaded functions (factor of about 6 slower).

I-V current calculations. The computed I-V currents across the junctions are shown in Figure 62 for holes and electron currents. In these calculations we have considered two slightly different cases of the hole and electron quasi Fermi levels (μ_p and μ_n, respectively). The knowledge of the exact position of the quasi Fermi levels requires an electrical simulation of the whole solar cell device under illumination and open-circuit or short-circuit operation, including TCO or metallic electrodes and band alignments to them, which is completely out of reach to the scale of atomistic simulations. A typical result of such simulation is shown in Figure 62a, depicting the band edges and quasi Fermi levels of a SHJ solar cell under illumination. It is possible to see that across the a-Si layers the μ_p and μ_n are close to the respective band edges of c-Si. For this reason we have considered few cases in which the quasi Fermi levels are moved slightly with

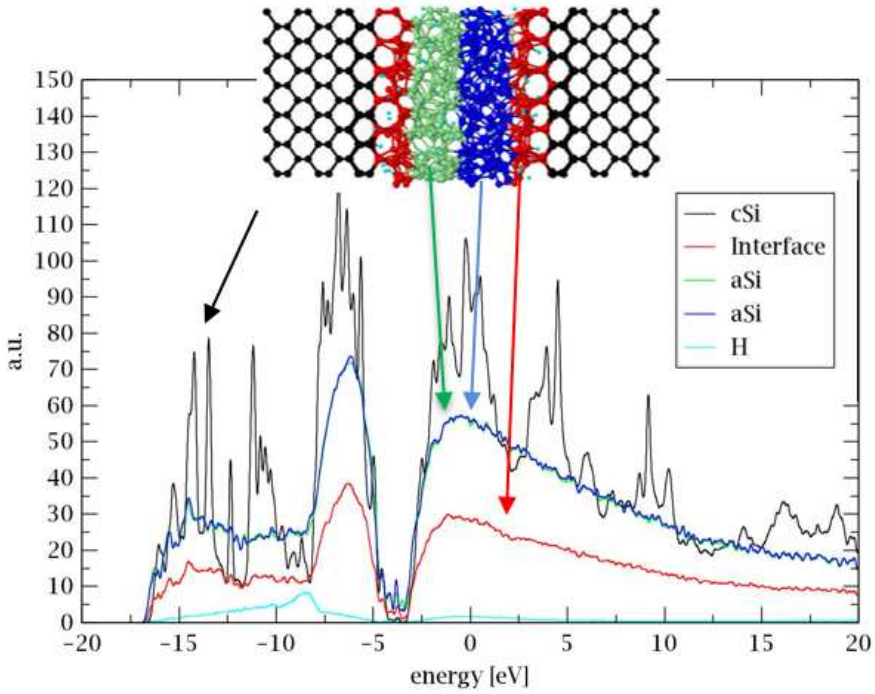


Figure 61: Density of states for the case of 300 K projected into 4 spatial regions of the junction: c-Si, interfaces and a-Si.

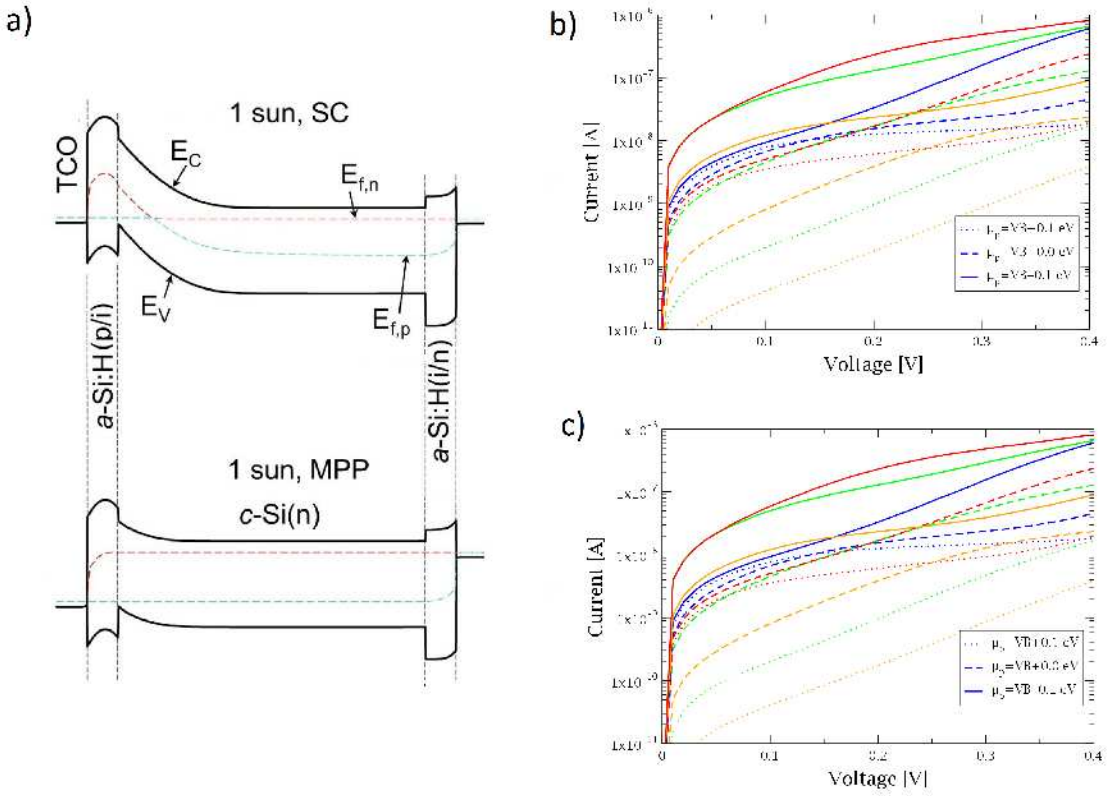


Figure 62: a) Typical band-diagram of a SHJ solar cell at short-circuit (SC) and maximum power point (MPP). The I-V characteristics of the c-Si/a-Si/c-Si junctions for b) hole current and c) electron current.

respect to the Si band edges. For holes we have considered the values of μ_p=-4.3 eV, μ_p=-4.2 eV and μ_p=-4.1 eV, corresponding to μ_p varying from 100 meV below the VB bandedge of c-Si, to 100 meV above it. For electrons we have considered the cases μ_n=-3.2 eV and μ_n=-3.0 eV, corresponding to the quasi Fermi level 200 meV below and at the CB edge. The corresponding I-V plots are shown in Figures 62b and 62c. The main conclusion of these calculations is that the annealing process can be quite beneficial to the hole current across the a-Si, but it is actually decreasing the electron current, albeit only slightly. The increase of hole current only occurs at low bias voltage across the junction, which is true under typical operating conditions, as also seen in Figure 62a. The reason for the increase of hole current can also be understood looking at the transmission plots showing an increase near the VB edge as the annealing temperature is increased. To the contrary, the electron transmission decreases at the CB edge, expecially for the case of T=900 K. However, the slight decrease of electron current is more than compensated by the increase of hole current, showing that the annealing process is indeed beneficial to the solar cell operation. The increase of current across the junction reduces the chance of non-radiative recombination losses, providing a significant improvement in short-circuit current of the device and PV conversion efficiency.

5.2 Task 1.3.2: Scientific results

The main task T1.3.2 is subdivided in four subtasks:

T1.3.2-1	**Electronic structure calculations**. QMC reference calculations will be performed on a series of reference systems. A planar graphite electrode and/or a coronene molecule will be put in contact with several adsorbed molecules: Water, sodium, chloride, 1-ethyl-3-methyl-imidazolium and tetrafluoroborate. Then we will perform DFT calculations on the same systems in order to benchmark the various functionals available (either in the gradient-generalized approach, such as BLYP or PBE, or the more costly hybrid functionals such as HSE06). Once the best functional is chosen, we will perform a large series of DFT calculations on larger systems containing several layers of water or ionic liquids. These will produce a reference data set of forces, energies, and multipoles for T1.3.2-2.
T1.3.2-2	**Force-field developments**. The simulations of T1.3.2-3 and T1.3.2-4 require a realistic representation of the interatomic interactions, but on large systems, which puts them outside the range of what is currently feasible using brute-force DFT. We will therefore introduce physically-motivated model potentials for the interactions, in which additional degrees of freedom are introduced to account for the response of the electronic structure of the molecules and the electrode to their changing environments, namely induced dipoles for the former and atomic charge fluctuations for the latter. These potentials will be parametrized by fitting the reference data set gathered in T1.3.2-1 using a generalized force and multipole-fitting procedure that is now well-established.
T1.3.2-3	**Capacitances of carbon materials for blue energy production**. The force fields develop in T1.3.2-2 will be used in large-scale simulations. Molecular dynamics (Metalwalls) and molecular density functional theory (MDFT) techniques will be used in this task. MDFT will allow to screen a large amount of carbon electrode materials. The most promising ones will then be thoroughly studied using Metalwalls in order to determine precisely the capacitance of the device as well as the charging time, the structure of the adsorbed species, etc.

T1.3.2-4	**Seebeck coefficients for redox active species in thermo-electrochemical cells**. The force fields developed in T1.3.2.2 will be used to study redox active species, such as ferrocene, dissolved in ionic liquids. By using Metalwalls, we will determine the free energy profile for electron transfer in the bulk liquid and in the vicinity of electrodes. This will provide the redox potential; by varying the temperature we will be able to extract the Seebeck coefficient for ranking the systems in terms of performance for thermo-electrochemical devices.

T1.3.2-1: Electronic structure calculations

Short overview of the results presented in D1.2. In previous work, to evaluate the importance of the exchange-correlation functional choice in DFT calculations of adsorbed molecules/ions at the surface of carbon materials, we calculated the adsorption of acetonitrile on a surface made of two graphene layers with PBE and SCAN, which are two of the currently most employed functionals (with and without the D3 correction for dispersion effect). It was clear that the difference between each choice is very large, and that dispersion interactions play an important role. Since there is no experimental data allowing to discriminate easily between the methods, we had therefore decided to perform high-level Quantum Monte-Carlo (QMC) simulations of such systems.

Methods. QMC simulations are a set of stochastic computational methods for the evaluation of observables of quantum systems. All QMC computations in the present work were performed using the QMCpack software [46]. The computational cost of the technique is very large, so we had to apply for a "grand challenge" on the GPU partition of Jean Zay, for which we were attributed 400,000 GPU hours.

We used QMC calculations to determine the adsorption energy of Li atom adsorbed on a graphite substrate as a function of the separation between the Li atom from the surface. If z is the distance between the Li atom and the carbon surface the adsorption energy profile $E_{ads}(z)$ is defined as

$$E_{ads}(z) = E_{Li+C}(z) - (E_{Li} + E_C) \tag{10}$$

where $E_{Li+C}(z)$ is the energy of a system made of a graphite substrate with a Li atom at a distance z and E_{Li} and E_C are the energy of the isolated atom and graphite respectively. The latter is modeled using two graphene layers made of 50 C atoms each, with an AB stacking (which corresponds to a $5\times5\times1$ supercell). The energy profile was computed for three different setups: with the Li atom lying above the centre of a C hexagon (hollow site), above a C atom (top site) and above a C–C bond (bridge site).

The height of the simulation box was selected after systematically testing the convergence of the total energy as a function of the amount of vacuum between periodic images in the z direction of the graphite bi-layer using DFT with the PBE functional. Dispersion corrections [63] were then used to relax the carbon bilayer structure for the selected box height of 30 Å, yielding a distance between the two planes of 3.47 Å which was used for the QMC computations. This value is slightly larger than the one provided in previous QMC studies (3.43 Å), [105, 141] but this should not impact the computed adsorption energies. Previous investigations systematically studied convergence with the size of the supercell and the number of twists for single and a bilayer systems [105, 141]. Here for computational reasons, we assumed the mentioned 5x5x1 supercell and we used a 4x4x1 twist grid (corresponding to eight nonequivalent twists) [90] based on the convergence of the corresponding DFT energy. Although convergence of the DFT total energy with twist grid is reached only for a 8x8x1 k-points grid, the residual effect is less than 1mHa. Moreover in the property we are aiming, namely the absorption energy of the Lithium atom, size effects should largely cancel since the same simulation cell and twist grid were used in all the simulations, including the ones of the isolated atom and substrate.

In all QMC simulations we used trial wave function with a Slater–Jastrow form

$$\Psi_T(\mathbf{p};\mathbf{R}) = J(\mathbf{p};\mathbf{R})D(\mathbf{R}) \tag{11}$$

where $J(\mathbf{p};\mathbf{R})$ is a Jastrow term describing electronic correlation, with one and two body terms, and $D(\mathbf{R})$ is a Slater determinant, ensuring the correct fermionic antisymmetry. The single particle orbitals used in the Slater determinant were evaluated using DFT with a PBE[119] functional. The orbital calculations were performed using the QUANTUM ESPRESSO software [45]. In the DFT calculations a plane wave basis set was used, with a cutoff at 150 Ry, using norm conserving pseudopotentials for both the Li and C atoms. In QMC simulations the Burkatzki–Dolg–Filippi set of pseudopotentials was used [23]. Wave function optimization is performed using the Linear method [153], and iterated until the optimized energy converges. Only the Jastrow part of the trial wave function is optimized. We report DMC results obtained using an imaginary time step of $\tau = 0.01$ Ha^{-1} and a population of 12000 to 16000 random walkers.

Electronic DFT calculations were performed using the QUANTUM ESPRESSO electronic structure code.[45] To be consistent with the QMC result, an identical simulation cell was considered, consisting of one hundred carbon atoms divided amongst two graphite layers with AB stacking (which corresponds to a 5×5×1 supercell). A kinetic energy cutoff of 40 Ry was used.

We compared the adsorption energy profiles for two series of XC functionals. In the first series we used the LDA, [26] PBE [119] and BLYP [12, 85] which neglect the London dispersion interaction. Then we included the latter using either the D2 correction parameterized by Grimme, [62] or through the use of the vdW-DF-C09 functional [152, 151, 13] implemented in the Libxc library. [86] Rappe-Rabe-Kaxiras-Joannopoulos ultrasoft (rrkjus) pseudopotentials [124] were used for both carbon and lithium atoms.

Several uniform Monkhorst-Pack grids of 1×1×1, 2×2×1, 3×3×1 and 5×5×1 k-points were tested for a single Li distance of 2.4 Å from the graphite surface (see Supporting Information Section S3). The difference between the total energy of the 1×1×1 and 2×2×1 grids is roughly 26 meV and thus the 1×1×1 grid was used for the calculations throughout this work. Additionally, an extended system consisting in four carbon layers instead of two was simulated to check the effect of the number of graphite layers; almost no difference was observed for the absorption energy profile as shown in Section 4 of the Supporting Information.

Like for the QMC calculations, the lithium atom was systematically above the three adsorption sites. The energies were converged at each step to an accuracy of 1×10^{-6} Ry. To align the various curves, the non-interacting systems (i.e. graphite and lithium atom separately) were computed for each XC functional, and the binding energy was obtained according to Equation 10.

Results. We first discuss the QMC results for the adsorption energies, which are reported in Table 16. These energies are clearly larger for the top and bridge sites than for the hollow site for distances of 1.5 and 2.0 Å, in good agreement with previous DFT results from the literature and from the current study. At a larger distance of 3.0 Å the three sites display similar energies, which shows that the difference between them has a short-range character. Due to the high computational cost of QMC, further lithium-carbon distances were only considered for the hollow site. We obtained a binding energy E_b of -1.08 eV for a lithium-surface distance of 1.8 Å. We also observe a somewhat peculiar behavior since the adsorption energy is rather similar for distances of 3 and 4 Å. By analyzing the corresponding electronic densities as shown on Figure 63A, we observe that this correspond to the region in which the electron transfer occurs. For distances lower than 3 Å the energies correspond to the adsorption of a lithium ion on a polarized carbon surface while for distances greater than 4 Å the system corresponds to neutral lithium atom and carbon material.

The adsorption profiles obtained for the hollow site for the LDA, BLYP, PBE, BLYP+D2, PBE+D2 and vdW-DF-C09 functionals are compared with the QMC benchmark on Figure 63B. The LDA results in a strong overbinding, which is expected. The comparison with the other functionals is more surprising. The QMC results lie between the BLYP and the BLYP+D2, while all the other functionals predict too low

	E_{QMC} (Ha)	E_{ads} (eV)
Graphite	-568.531(1)	–
Li atom	-0.198314(3)	–
Hollow site		
1.0	-568.693(1)	0.98(4)
1.5	-568.763(1)	-0.92(5)
1.8	-568.769(1)	-1.08(5)
2.0	-568.768(1)	-1.06(4)
3.0	-568.742(1)	-0.35(4)
4.0	-568.741(1)	-0.32(5)
8.0	-568.730(2)	-0.03(6)
Top site		
1.5	-568.705(2)	0.65(5)
2.0	-568.761(1)	-0.87(4)
3.0	-568.740(2)	-0.31(4)
Bridge site		
1.5	-568.716(2)	0.36(4)
2.0	-568.760(1)	-0.83(3)
3.0	-568.742(1)	-0.35(3)

Table 16: QMC energies for the adsorption of lithium on graphite, for the three different sites for several lithium-carbon distances reported in the first column (in Å). Second column are QMC results while the last column gives the adsorption energies computed using equation 10.

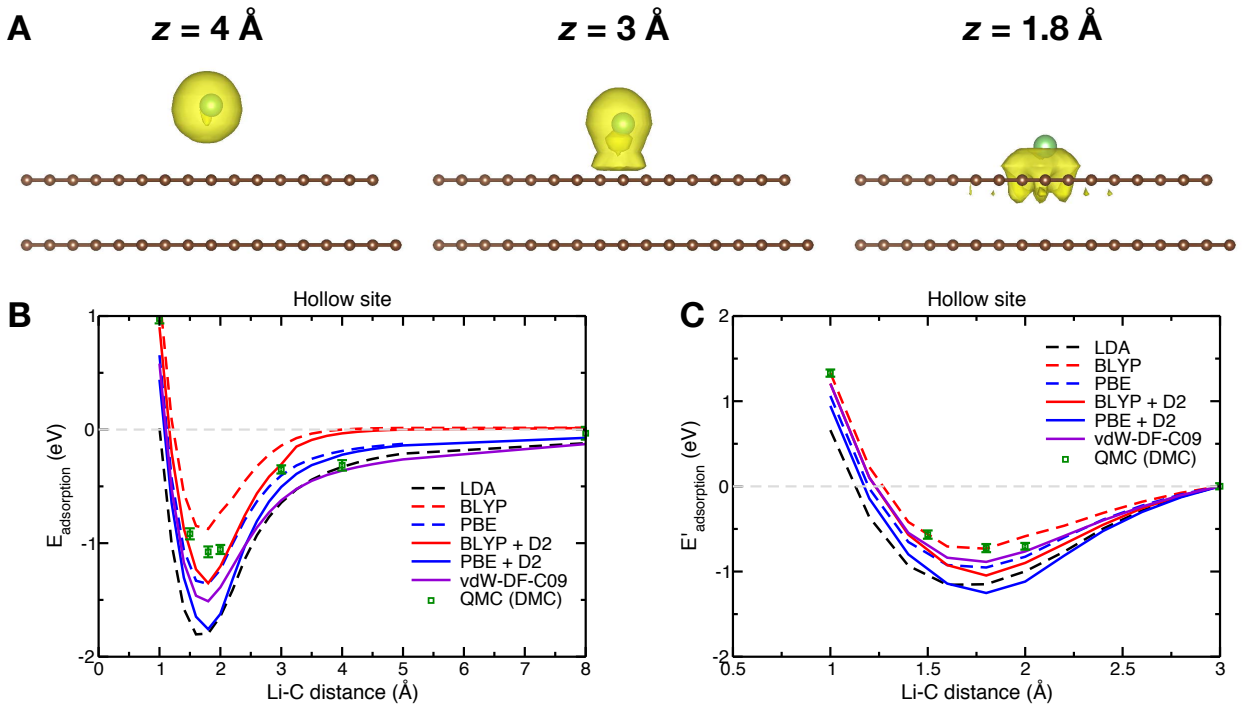


Figure 63: A) Electron density around a Li atom adsorbed on a graphite substrate, computed via quantum Monte Carlo, for different atom–substrate distances. The densities were obtained by computing the overall electronic density of a system with a Li atom adsorbed on graphite, and subtracting the density of the isolated substrate, in absence of the adsorbed atom. All shown isosurfaces correspond to a density of 6 $\cdot 10^{-4}$ electrons/Å^3. B) Comparison of the adsorption energies obtained with various DFT functionals and QMC for the adsorption of the lithium on the hollow site of graphite. C) Same as B) but substracting the adsorption energy at a distance of 3 Å.

adsorption energies. However, as noted by Valencia *et al.*, the binding energy (defined as the minimum of the adsorption energy profile) should be approximately given by [154]

$$E_b \approx E_b(\mathrm{Li}^+) - (IP[\mathrm{Li}] - W_f[\mathrm{Graphite}]) \tag{12}$$

where $E_b(\mathrm{Li}^+)$ is the binding energy of the lithium *ion*, IP its ionization potential and W_f the work function of graphite. The observed variation may therefore be due to different values for IP and W_f from the various functionals.

By visualizing the electronic density of the system using various DFT functionals, we also observe that they yield similar results as the QMC calculations up to 3 Å while some discrepancies are observed at a larger distance of 4 Å. In the present work, we are mostly interested in developing an accurate potential for the adsorption of the lithium ion. Consequently, we performed a second comparison of the various functionals in which the energy at z = 3 Å is subtracted. The results are shown on Figure 63C. The discrepancy between the various functionals is somewhat lower. The best agreement is now obtained with BLYP, PBE and the vdW-DF-C09 functionals, while the others predicts overbinding. For the two former approximations, the good agreement may be fortuitous since they do not account for the dispersion effects. However, this points towards an overestimation of the dispersion effects when using the D2 correction. Indeed, in the case of the lithium ion, only two semi-core electrons take part in the interaction, which should result in a very weak dispersion term. The vdW-DF-C09 functional, which accounts for these effect explicitly and not through a parameterised term, seems to better catch the interactions.

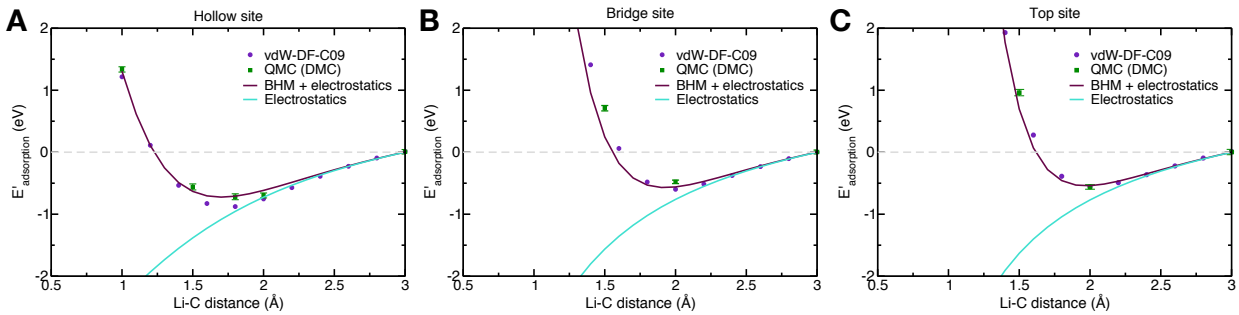


Figure 64: Comparison of the fitted potential with the QMC and vdW-DF-C09 energies.

T1.3.2-2: Force-field developments

Short overview of the results presented in D1.2. In the previous deliverable we had discussed the development of polarizable force fields for redox-active species in acetonitrile solvent using force-fitting methodologies.

Results. In order to develop accurate classical interaction potentials, it is not possible use directly the QMC calculations described above due to the limited number of data. Instead, we pick the most accurate functional, vdW-DF-C09, and calculate the Li–graphite binding energy for a much larger number of distances. The intermolecular interaction should in principle account for four different effects: electrostatics, polarization, short-range repulsion and dispersion. In our electrostatic model, the two former effects are explicitly introduced through the use of a +1 point charge on the lithium and of the calculation of partial charges on the carbon atoms. Indeed, the electrode charges are represented with Gaussian charge distributions with a width of 0.40 Å. These partial charges are calculated by enforcing a uniform potential within the whole carbon electrode, with an overall electroneutrality constraint (hence forcing the total charge on the carbon to be equal to -1).[142]. Electrode charges are optimized self consistently with the functional minimization through an iterative scheme. This part of the potential therefore does not need to be parameterized further.

Concerning the short-range repulsion and the dispersion effects, the two main potential forms used in the literature are the Lennard-Jones and the Born-Huggins-Mayer (BHM) ones. However, it appears that the electrostatic interaction was sufficient to account for the attractive part of the binding energy, as shown on Figure 64. The fitted potential should therefore add very few, if no contribution for the dispersion interaction, which agrees with the previous observation on the use of dispersion-corrected functionals. A well-known drawback of the Lennard-Jones potential is that it is not possible to fit the short-range repulsion and the dispersion term separately since they both involve the same parameters. We have therefore chosen a BHM potential instead, which analytical form is:

$$V_{\mathrm{BHM}}(r) = A\exp\left(-br\right) - \frac{C_6}{r^6} \tag{13}$$

where A, b are the parameters describing the intensity and the range of the repulsion interaction, while C_6 is the dipole-dipole dispersion interactions. In principle higher order terms could be included for dispersion, but as discussed above this term is almost negligible in the case of the lithium ion. The fitted potential reproduces with a very high accuracy the vdW-DF-C09 for the three types of adsorption sites as shown on Figure 64. It also matches very well with the BLYP (not shown on the figure) as well as the QMC results (on which it was not fitted) adsorption curves. The corresponding parameters are A = 91.17, b = 2.518 and C_6 = 1.107 (all numbers are given in atomic units). In future work we will use the same procedure for other ionic species.

T1.3.2-3 Capacitances of carbon materials for blue energy production

Methods. Solvation free energies were computed with MDFT[72] while the polarisability of the graphite sheets was handled using fluctuating Gaussian charges method[142, 136]. MDFT is a flavor of classical density functional theory (cDFT) developed to study the solvation properties of molecular solutes into molecular solvents such as water or acetonitrile. The solvent is described by its density field $\rho(\mathbf{r},\omega)$ which measures the average number per unit volume of molecules with an orientation ω at a given position $\mathbf{r}$. The solute acts a perturbation through an external potential $V_{ext}(\mathbf{r},\omega)$ causing the solvent to deviate from the homogeneous bulk fluid.

According to the cDFT principles[48, 66], there exists a unique functional, F, of the solvent density, ρ, that is equal to the solvation free energy at its minimum which is reached for the equilibrium solvent density. To find an expression for the functional, a common practice is to start by splitting it into the following sum

$$F[\rho(\mathbf{r},\omega)] = F_{id}[\rho(\mathbf{r},\omega)] + F_{exc}[\rho(\mathbf{r},\omega)] + \int\int \rho(\mathbf{r},\omega)V_{ext}(\mathbf{r},\omega)d\mathbf{r}d\omega. \tag{14}$$

In Equation 14, the first term of the rhs is called ideal and corresponds to the entropic contribution of a non-interacting fluid with the same density. The second term is due to solvent-solvent interaction and is often called the excess term while the last term is due to solute-solvent interaction and thus called the external term.

Exact expressions exist for the ideal and external functionals that can be computed numerically. The excess part, however, requires approximations. It can be expressed as an infinite Taylor expansion around the homogeneous bulk solvent density ρ_b,

$$F_{excess}[\rho] = -\frac{k_BT}{2}\int\int\int\int \Delta\rho(\mathbf{r},\omega)c^{(2)}(\mathbf{r}-\mathbf{r}',\omega,\omega')\Delta\rho(\mathbf{r}',\omega')d\mathbf{r}d\omega d\mathbf{r}'d\omega' + F_B[\rho] \tag{15}$$

In Equation 15, $\Delta\rho(\mathbf{r},\omega) = \rho(\mathbf{r},\omega) - \rho_b$, k_B is the Boltzmann constant, T is the temperature and $c^{(2)}$ is the two-body direct correlation function of bulk solvent. F_B is so-called bridge functional that collects all the terms higher than quadratic, involving many-body direct correlation functions of the bulk solvent. A common way to approximate the excess functional is to ignore the bridge functional, i.e. $F_B = 0$, resulting to the "HNC" functional because it is equivalent to using the hypernetted chain (HNC) closure for the solute-solvent correlations in the molecular Ornstein-Zernike equation[37]. In this work, we use a very simple bridge functional [18, 19] based on weighted density approximation that is known to correct well for the dramatic pressure overestimation of the HNC approximation. Water is modeled with the SPC/E force field while the external potential V_{ext} is created by the graphite electrodes and the lithium ion whose interactions are described as the sum of Lennard-Jones and electrostatic interactions. The Lennard-Jones and charge parameters of the 4 types of atoms involved are collected in Table 17. For the carbon atoms, the choice of the force field of Werder *et al.* [161] was made based on a previous QMD study, in which it was shown to provide a good estimate of the energy of adsorption of a water molecule on a graphene surface. [3]

The electrode charges are represented with Gaussian charge distributions with a width of 0.40 Å. These partial charges are calculated by enforcing a uniform potential within the whole carbon electrode, with an overall electroneutrality constraint (hence forcing the total charge on the carbon to be equal to -1).[142]. Electrode charges are optimized self consistently with the functional minimization through an iterative scheme[72]. In the first step, the functional of Equation 14 is minimized with no charges on the lithium and the carbon atoms. Then, carbon charges are optimized in the presence of the inhomogeneous water charge density and of the lithium cation. The functional is minimized again but in the presence of lithium charge and the previously determined electrode charges. The process is repeated until it converges, with a convergence criterion of 5×10^{-4} on the relative change in solvation free energy between two consecutive steps.

MDFT calculations were performed using an in-house Fortran code and electrode charges are optimized using the constant potential molecular dynamics software MetalWalls[96] (which was partly devel-

Atoms	σ (Å)	ϵ (kJ/mol)	charge (e)
O	3.166	0.65	-0.8476
H	0	0	0.4238
Li[4]	2.216	0.07648	1
C[161]	3.214	0.2364	Fluctuating

Table 17: Force-field parameters used in the MDFT simulations. Mixed parameters are computed using the Lorentz-Berthelot rules (except for the C-Li interaction, which does not affect the MDFT results).

oped within the EoCoE project). We use a $24.672 \times 21.366 \times 40$ Å^3 simulation box (the unit cell used in QMC is replicated twice in x and y directions) with $74 \times 64 \times 120$ grid nodes and an angular grid of 196 orientations per grid node. We run calculations for a distance z between the electrode plane and the lithium varying between $z = 1.0$ Å and $z = 10$ Å with an increment of 0.2 Å between 1 Å and 6 Å and of 0.5 Å otherwise.

Results. The fitted potential obtained in the previous subtask can directly be used in any classical molecular simulation, such as MD. Since we focus here on the adsorption free energy of the lithium ion on the carbon surface we prefer to use MDFT which is a computationally more efficient alternative. The solvation free energy of a single system can be computed within a few minutes on a single CPU while it would require tens of CPU hours with MD. The free energy profile obtained for the three adsorption sites in the presence of liquid water are shown on Figure 65A. The profiles are very different from the gas phase results. The minimum at $\approx$ 1.8 Å completely disappears and is replaced by a strongly repulsive wall. This shows that there is no preferential adsorption of lithium on the graphite surface in aqueous phase.

This result is in qualitative agreement with a recent MD study on the adsorption of hydrated ions on graphene, [92] which provided a repulsive free energy profile over the whole range of considered distances. Yet, the latter study did not include any Coulombic interaction between the ion and the carbon surface, which is the main driving force for adsorption in the gas phase as discussed above. It is thus interesting to examine the various contributions to the total free energy, which are provided on Figure 65B. We observe that the electrostatic attraction between Li and C is in fact counterbalanced by the solvation free energy. The latter contains the electrostatic interactions of the water molecules with both the lithium and the graphite surface, which results in strong screening effect. The extent of this screening effect was studied in details in a recent study focused on gold surfaces: [120] The presence of the water molecules strongly impacts the polarization of the surface. Consequently, the total free energy is almost equal to the BHM contribution over the whole range of distances, except between 2 and 5 Å where the solvation free energy overcomes the ion-surface Coulombic interaction, resulting in a more repulsive potential.

The effect of the solvent can be further analyzed by plotting the density profiles for various distances between the ion and the surface (Figure 66). At z = 8 Å, two regions with larger densities emerge, corresponding to the surface adsorbed water molecules at a distance of 3 Å from the surface [72] on the one hand and to the lithium ion first solvation shell on the other hand. At a distance of 5 Å the solvation shell starts to overlap with the adsorbed layer at close contact to the electrode, which results in small depletion zones in the latter. These depletion zones remain observable at smaller distances, but the impact on the free energy becomes negligible w.r.t the large short-range repulsion between the carbon and the lithium.

Now that we have applied the method on a simple system, we will now focus on more complex nanoporous carbon materials. The computations are much more involving, and we need to solve memory issue in the MDFT program in order to tackle the large grids which are involved. However we have recently been able to simulate a system consisting of one electrode in contact with a salt, represented using a

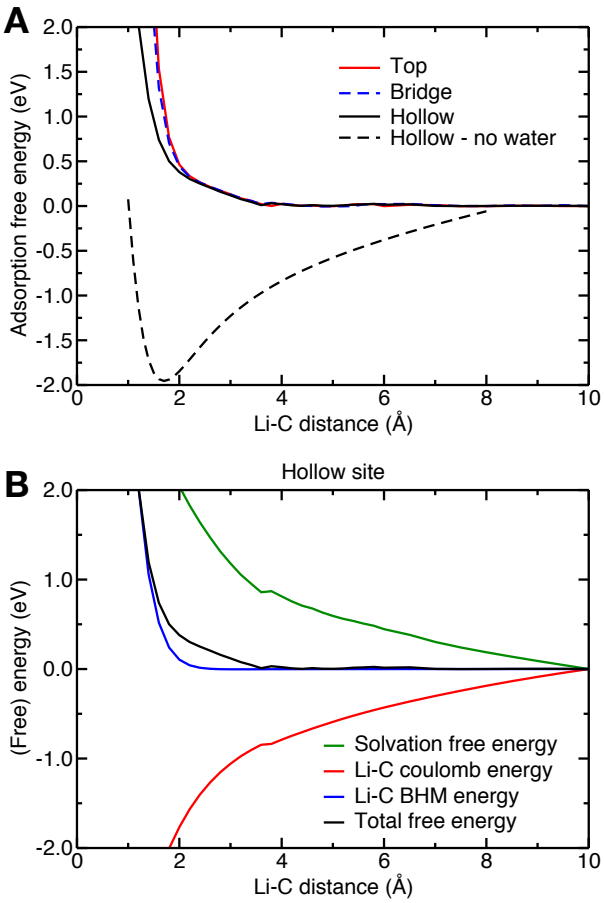


Figure 65: A) Adsorption free energy for a lithium ion on the graphite surface in the presence of water, computed using MDFT, for the three adsorption sites. The energy variation in the absence of water is also shown for comparison. B) Contributions to the total free energy for the hollow adsorption site in the MDFT calculation.

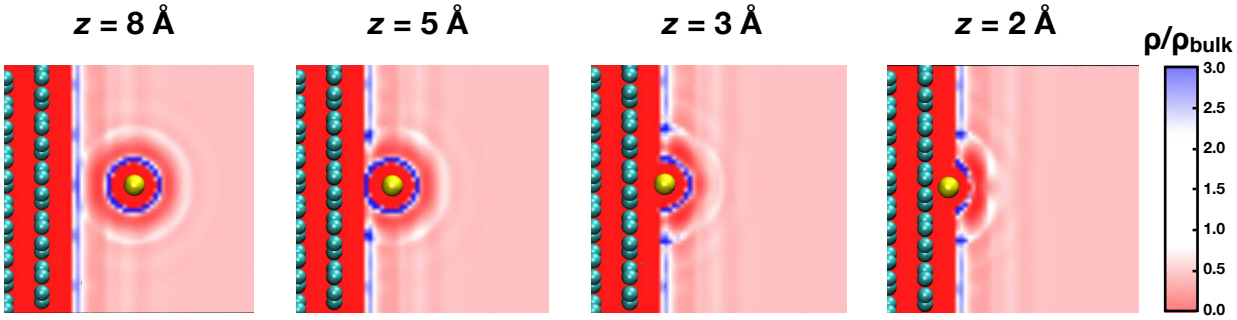


Figure 66: Projection of the solvent densities computed using MDFT for various lithium–carbon distances.

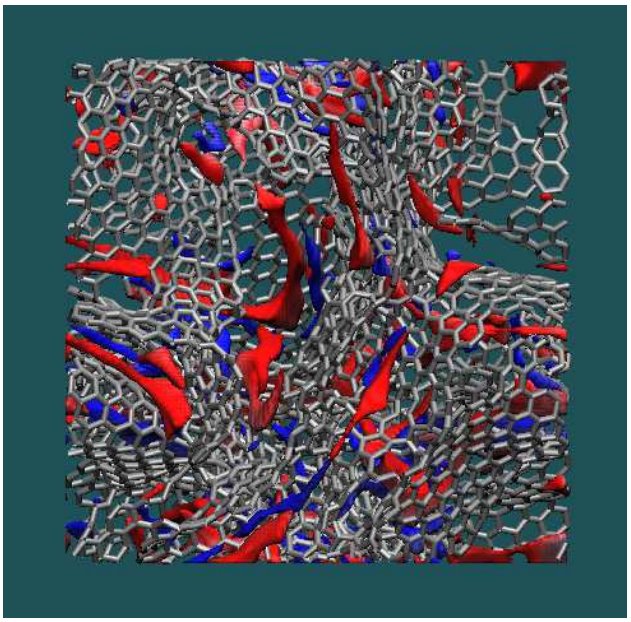


Figure 67: Ionic densities for a NaCl salt (represented with a primitive model) inside a carbide-derived carbon electrode. C–C bonds are shown in gray, and the blue and red isosurface correspond to isovalues of 0.01 $Å^{-3}$ for Na^+ and Cl^- ions, respectively.

primitive model. The result remain quantitative so far, but the MDFT calculation allows us to determine the regions of the electrode in which ionic species display preferential adsorption, as shown with isosurfaces on Figure 67.

T1.3.2-4 Seebeck coefficients for redox active species in thermo-electrochemical cells

Short overview of the results presented in D1.2. In the previous deliverable we had discussed an approach based on machine-learning to accelerate the calculation of redox properties of molecules using molecular dynamics trajectories. Our objective was to apply this to *ab initio* MD simulations performed at various temperatures, in order to track variations in the redox potential that would lead to interesting Seebeck coefficients. However our further calculations using this tool did not display such an effect since the potentials did not change with temperature. After looking at the litterature, we found some experimental results showing that interesting thermoelectric effects could be obtained using ionic liquids based capacitors [17], so in the last period of the project we focused on such systems.

Methods. We have simulated a system in which the ionic liquid is confined between a pair of planar electrodes held at constant electrical potential and studied the induced charge on the electrodes, which fluctuates due to the thermal motion of the ions in the liquid. The specific system we consider is a molecular simulation model of the RTIL butylmethylimidazolium hexafluorophosphate (BMI-PF_6) bounded by planar, constant voltage graphite electrodes, as shown in Fig. 68.

The model we use to simulate a metallic electrode maintained at a constant potential follows from the work of Siepmann and Sprik [142], later adapted by Reed *et al.* to the case of electrochemical cells [126]. Several potentials ranging between -1 V and 1 V were applied to the electrodes. All molecules are represented by a coarse-grained model in which the forces are calculated as the sum of site-site Lennard-Jones potentials and Coulombic interactions. Parameters for the ions and carbon atoms are the same as in our previous works [104, 102, 103, 101]. Three sites are used to describe the cation, while the anions are treated as spheres [129].

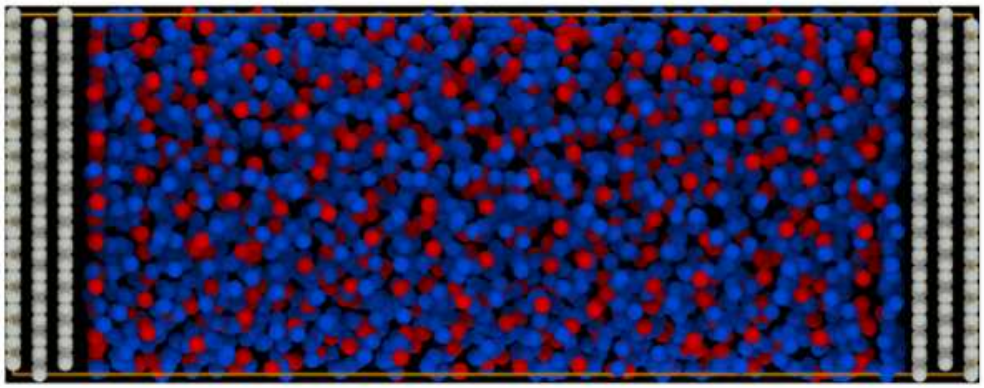


Figure 68: Typical snapshot of the simulated thermoelectrochemical cell. The ionic liquid butylmethylimidazolium hexafluorophosphate (red and blue spheres) is put in contact with two graphite electrodes (gray spheres) held at constant potential.

Each electrode is modelled as three fixed graphene layers, with a distance between carbon atoms within each layer of 1.43 Å and a distance between layers of 3.38 Å. The simulation cell lengths are 4.3, 4.7 and 12.4 nm in the x, y and z directions, respectively.

In order to mimic a thermo-electrochemical cell, in which the two electrodes are in contact with materials at different temperatures, a heat flux was applied to the simulation cell according to the algorith of Wirnserger *et al.* [163]. Several heat flux values were tested: 0.0157×10^{-18}, 0.0314×10^{-18}, 0.0471×10^{-18} and 0.0628×10^{-18} W m^{-2}. The time step of the simulations was set to 2 fs.

Results. In a first step, it is necessary to check whether the heat flux induces a temperature gradient as expected, since the algorithm has only been applied to 3D-periodic systems so far. It is indeed the case as shown on Figure 69: We obtain respective gradients of 1.9, 3.9, 5.9 and 8.3 K Å^{-1} with increasing heat fluxes. Such numbers are of course very large compared to the experimental cases, but it is necessary to obtain a visible effect in the simulations.

We then analyze the impact of the temperature gradients on the structure of the liquid. As shown on the left panel of Figure 70 (black line), at equilibrium the density profiles of the ions across the cell are characterized by strong oscillations that arise due to Coulombic ordering. Due to the applied potential on the electrode, the first adsorbed layer is enriched in counterions, which charge is then counterbalanced in the second layer by ions of the opposite sign, and so on until the effect dampens and the isotropic bulk density (and structure) is recovered. For small temperature gradients (red and green plots), this structure is not much affected but for the larger ones we observe an important decrease in the layering effect. The cold interface remains structured since the density profiles are still characterized by the presence of peaks (albeit less resolved than for the equilibrium case), while the hot interface becomes much less structured, with only one large adsorption peak followed by quickly dampened oscillations. This result is consistent with the obtained temperature: One interface becomes glass-like while in the other one the large temperature leads to a system dominated by entropic effects.

We now turn to the electrochemical properties of the system. When applying a constant potential between the electrodes, after a transient regime the total charge of the electrodes stabilize to an equilibrium value. The variation of this total charge q with ΔV for the various temperature gradients is shown on Figure 71. We observe that the largest the gradient, the smaller the charge accumulated at the surfaces of the electrodes. In practice, this would correspond to a negative Seebeck coefficient which means that we could not retrieve electricity from such a device. However, as discussed above, the temperature gradients are large, which results in particular in a glassy behavior at the cold electrode in some of the cases. It is therefore possible that the sampling of the simulations is not enough converged. Running further these

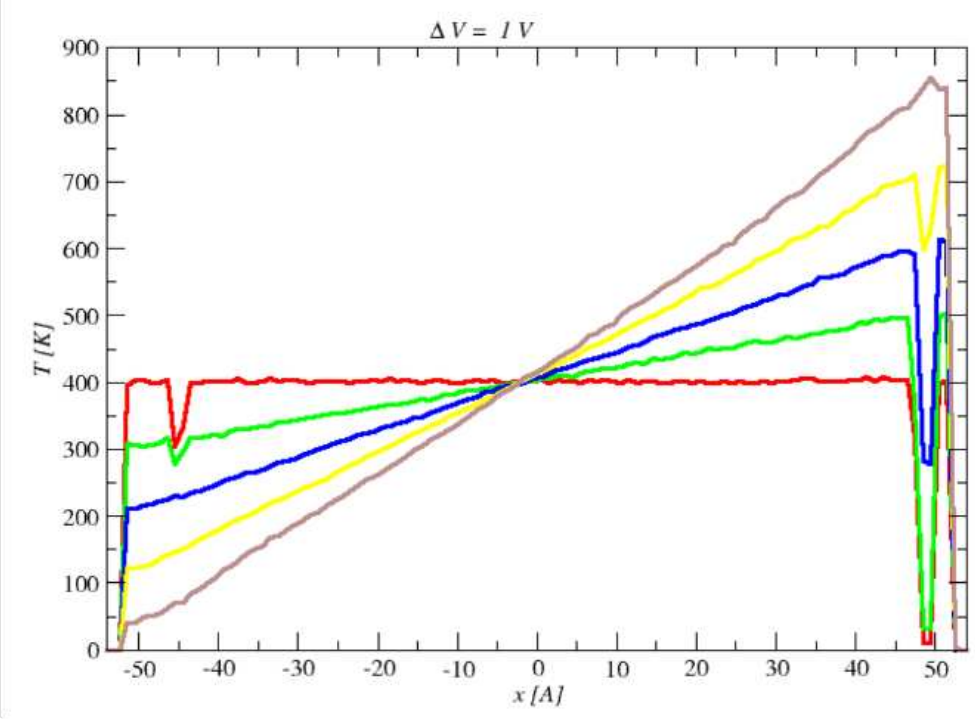


Figure 69: Temperature profiles across the simulation cell, for an applied potential of 1 V and the four different applied heat fluxes (red: without heat flux, green: 0.0157×10^{-18} W m^{-2}, blue: 0.0314×10^{-18} W m^{-2}, yellow: 0.0471×10^{-18} W m^{-2} and marroon: 0.0628×10^{-18} W m^{-2}.

simulations would not be very helpful since in experiments such effects should not occur. If we then focus on the smaller temperature gradient, we observe that the variation of the charge is very similar to the equilibrium case. Since this corresponds to more realistic conditions, this leads us to conclude that this regime should be explored further. In practice, future work should involve more elaborate systems, such as the 1-ethyl-3-methylimidazolium tetrafluorosulfonylimide, which was used in the experiments. It is then necessary to use all atom models, which are much more expensive computationally than the coarse-grained models involved here.

5.3 Task 1.3.3: Scientific results

The main task T1.3.3 is subdivided in four subtasks:

T1.3.3-1	**Development of KMC simulator**. Algorithmic optimisation of the FMM-KMC code for the KMC simulator.
T1.3.3-2	**Application of KMC simulator to electron transport in organic materials**. Preliminary results applying the FMM-KMC code to the study of doped organic semiconductors have been published in a paper demonstrating the scaling of the electrostatics library [135]. The FMM-KMC code has been used to simulate all solid state batteries [35].
T1.3.3-3	**Development of DMC simulator**. Development of the BOLTMC code for the DMC simulator, which solves the Boltzmann transport equation, using the PPMD Python framework.

T1.3.3-4	**Application of DMC simulator to electron transport in in perovskites**. The BOLTMC code has been used to investigate the polaronic effects on scattering and mobilities of charge carriers in halide perovskites [71].

Tasks T1.3.3-1 –T1.3.3-4 describe multiscale simulation methodologies which are amenable to studying phenomena at a broad range of length and time scales. These methodologies are shown for lead halide perovskite solar cells in Fig. 72.

To aid scalability, the Kinetic Montecarlo (KMC) and the Device Monte Carlo (DMC) simulators employed PPMD code generation [133]. This framework allows scalable code to be written using a domain-specific language written in Python. It is easy to run on any HPC architecture. It uses MPI distributed memory parallelization. OpenMP could be used for shared memory.

T1.3.3-1: Development of KMC simulator

As stated in the Executive summary, KMC simulations track the paths of localised charges hopping to unoccupied sites, illustrated in Fig. 73 [135]. In D1.2 results were presented on simulations of doped organic semiconductors. Here the UBAH KMC code was modified such that the electrostatic potential due to static charges is calculated just once at the beginning of a simulation, rather than at every re-evaluation of the electrostatic energy over its course. This significantly decreases the time required for simulations involving ionised dopants, for example. We have also implemented Dirichlet boundary conditions in the electrostatic solvers so that the effects of metallic contacts may be included in the simulation. Furthermore, algorithmic developments are being implemented so as to enable injection and extraction events to take place from and to those contacts, and for creation and annihilation of electron–hole pairs due to photoexcitation and radiative recombination in the bulk of the material.

T1.3.3-2: Application of KMC simulator to electron transport in organic materials

A new multilevel method which reduces the computational complexity to $\mathcal{O}(\log(N))$ per Metropolis-Hastings step, where N is the number of particles, while maintaining errors which are comparable to direct Ewald summation [134], has been developed. Fig 74 illustrates the exascalability of the code on different high performance computing processors [135].

We used this code in D1.2 to model the hole mobility variation with ionised dopant concentrations down to 0.01% in an off-lattice model that accounts for amorphous molecular packing, which allowed us to investigate up to $\sim$ 20,000 charge carriers in a system containing 510,300 molecular sites. Preliminary results applying the KMC simulator to the study of doped organic semiconductors were published in a paper demonstrating the scaling of the electrostatics library [135].

Since M18, this code has been used to simulate all-solid-state batteries, and the results have been published in [35]. These batteries are safer, easier to process, higher achievable power density and cyclability. The task has demonstrated collaboration between (i) Materials science (heterogeneity of solid state electrolyte at grain boundaries); (ii) Mathematics (Fast Multipole Model); (iii)Statistics (Quantify space charge profiles by maximum likelihood sampling; (iv) Exascale computing (code designed for exascale computers).

Our computational model consists of a 3D-periodic simple-cubic lattice of $m \times m \times m$ sites, populated by a fixed number, N, of mobile point defects with +1 charge (Fig 75) [35]. The simulated space-charge profile presented in Fig. 76 shows a large space-charge width—i.e., long decay lengths at a relative permittivity ϵ_r of 1. This is the opposite behavior to that predicted by the classic space-charge treatment, wherein the Debye length decreases with decreasing relative permittivity. The observation of a decay length that is larger than the classical Debye length mirrors the behavior of concentrated liquid electrolytes, where this phenomenon is termed underscreening.

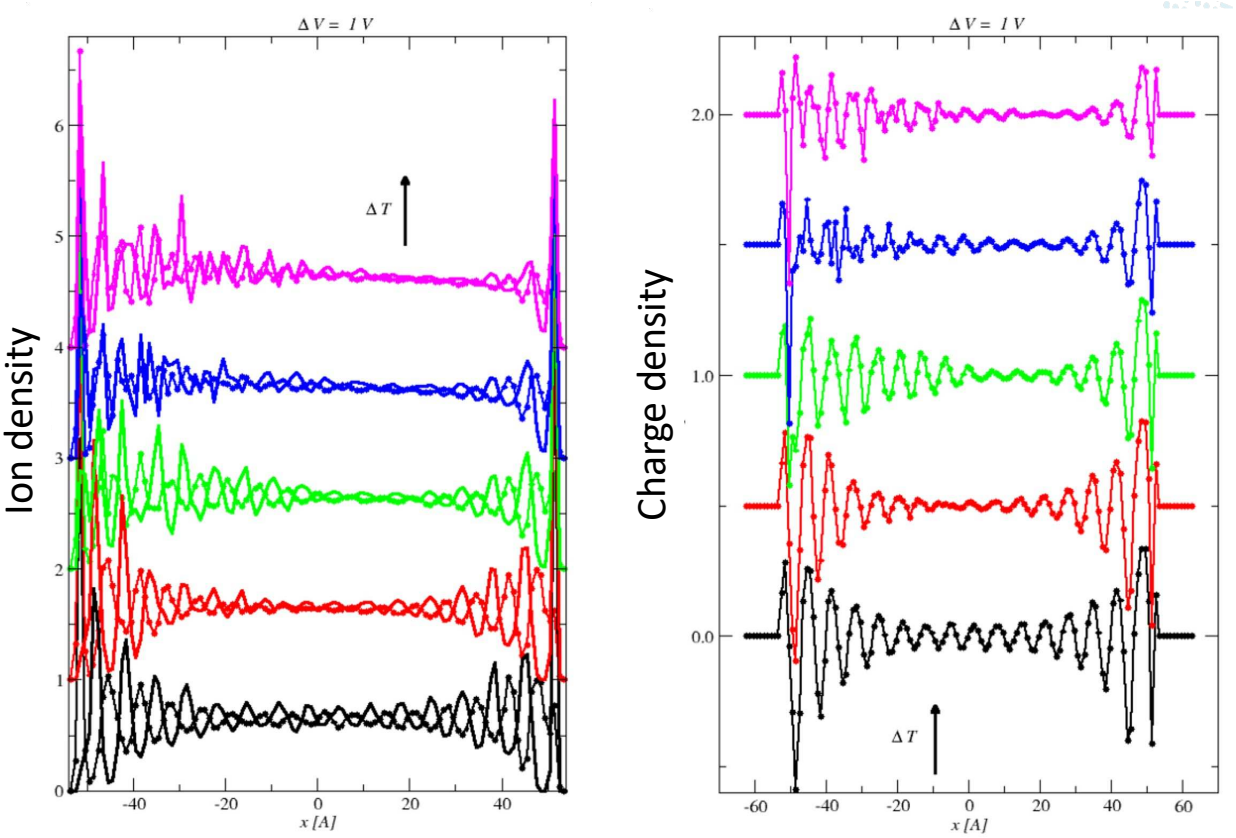


Figure 70: Ionic (left panel) and charge (right panel) densities across the simulation cell, for an applied potential of 1 V and for the various heat fluxes (from bottom to top on each figure, the black plot corresponding to the equilibrium case with no heat flux for comparison). The plots are shifted upwards for each case for visibility purposes.

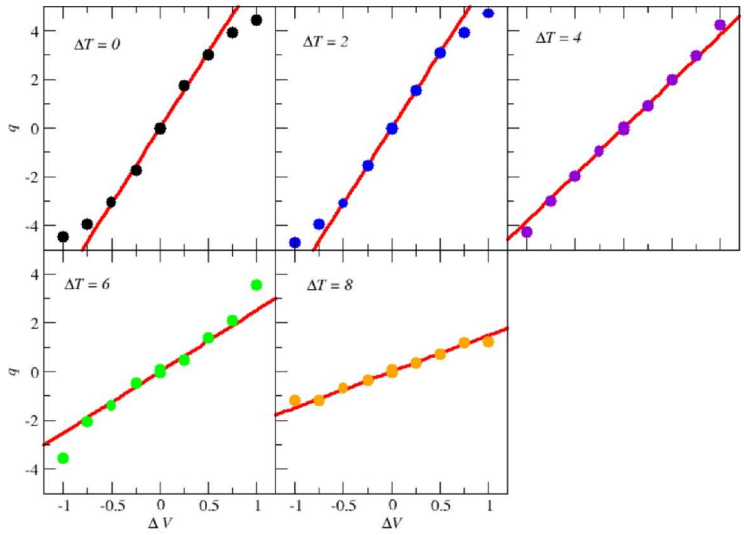


Figure 71: Accumulated charge on the electrodes with respect to the applied potential, for the various temperature gradients (given in K $Å^{-1}$).

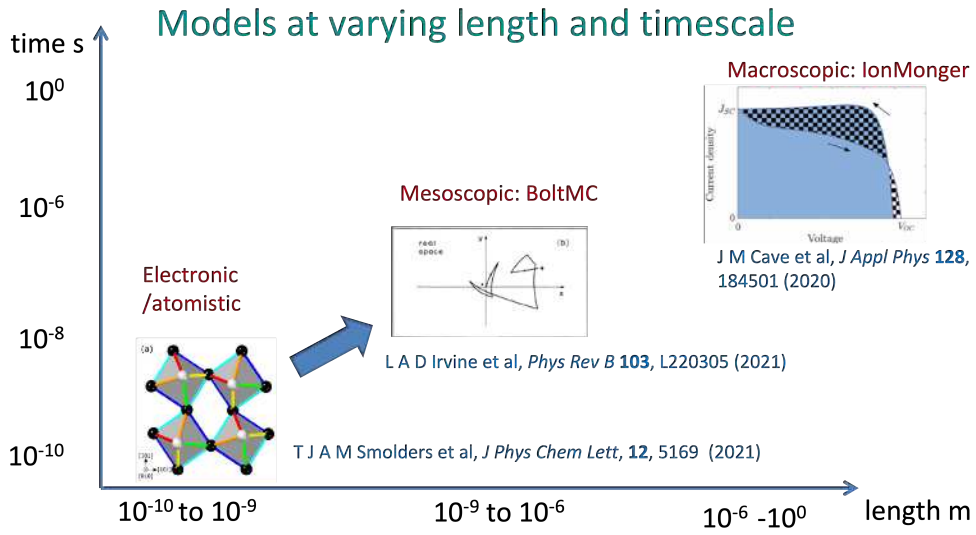


Figure 72: Diagram showing length- and time-scales over which electronic/atomistic, mesoscopic and macroscopic models developed by UBAH operate.The figure shows example publications by UBAH and the open source codes BOLTMC and IonMonger.

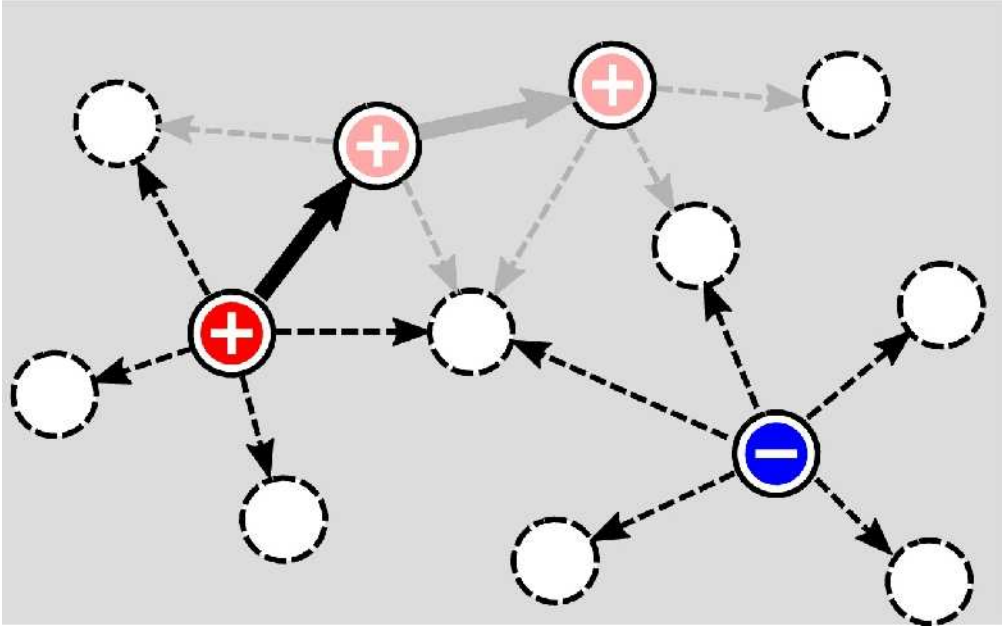


Figure 73: Schematic illustration of possible hops of a charged particle (positive (negative charge is a + (-) sign on a red (blue) disc) to neighbouring empty sites (white discs).

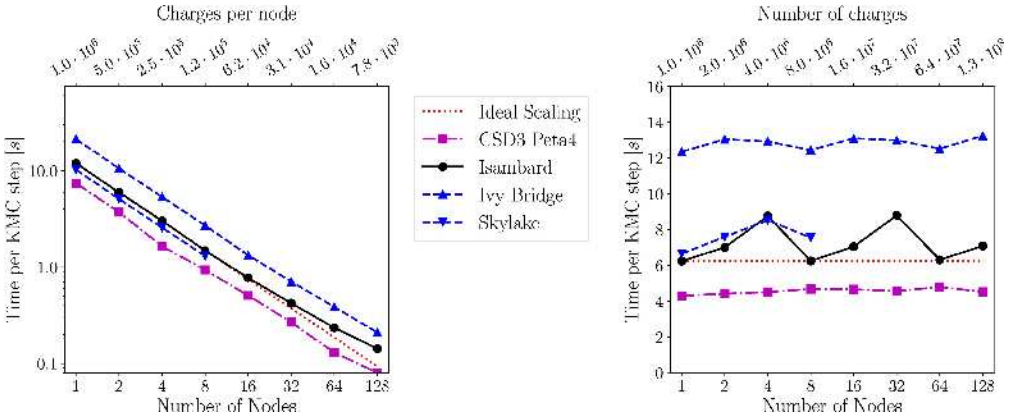


Figure 74: Demonstration of scalability of FMM-KMC code by showing how the time per KMC step (s) varies with (left panel) the number of charges per node; (right panel) the number of nodes, or equivalently the number N of charges).

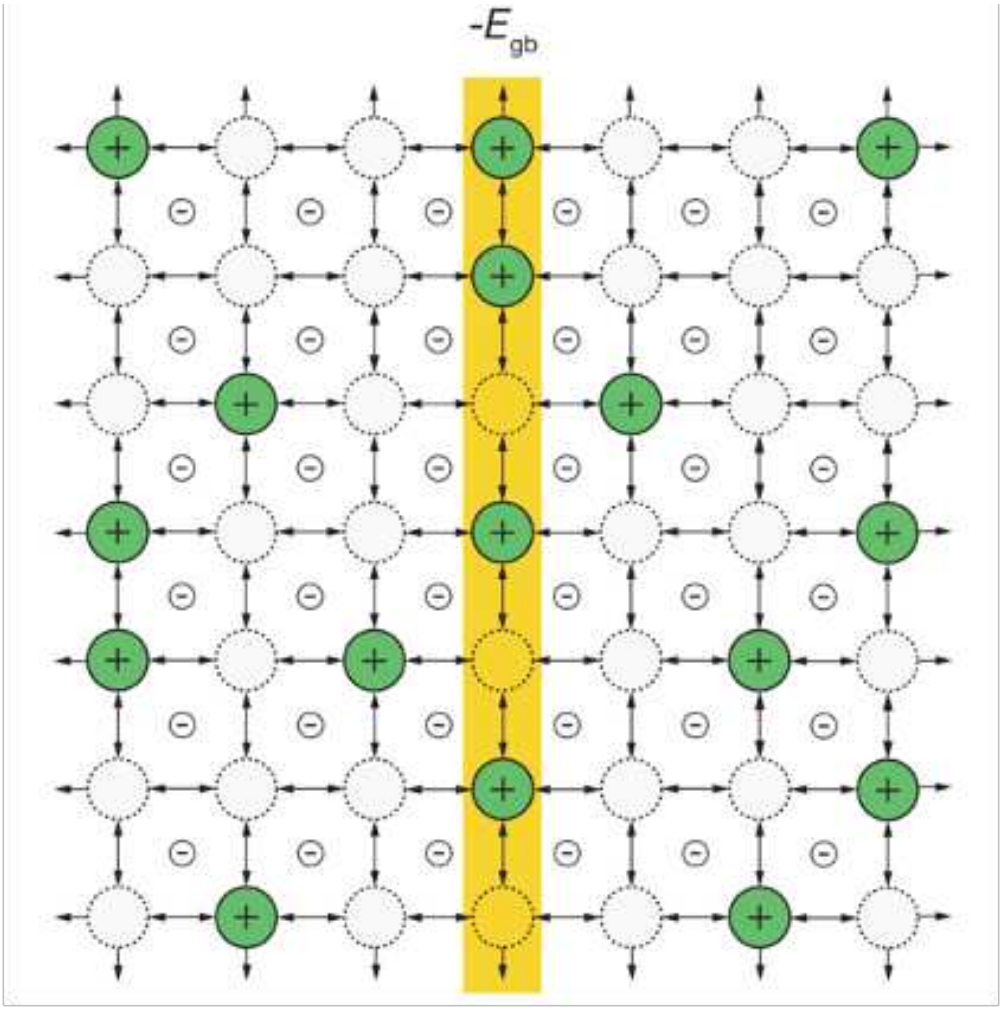


Figure 75: Two-dimensional schematic of the kinetic Monte Carlo model used in this work. Lattice sites (circles) are either vacant (dashed circles) or occupied by positively charged "defects" (solid green circles). Arrows indicate allowed site-site moves. Interstitial positions are assigned partial negative charges. All lattice sites in the central plane (yellow) are assigned an on-site occupation energy of $-E_{gb}$).

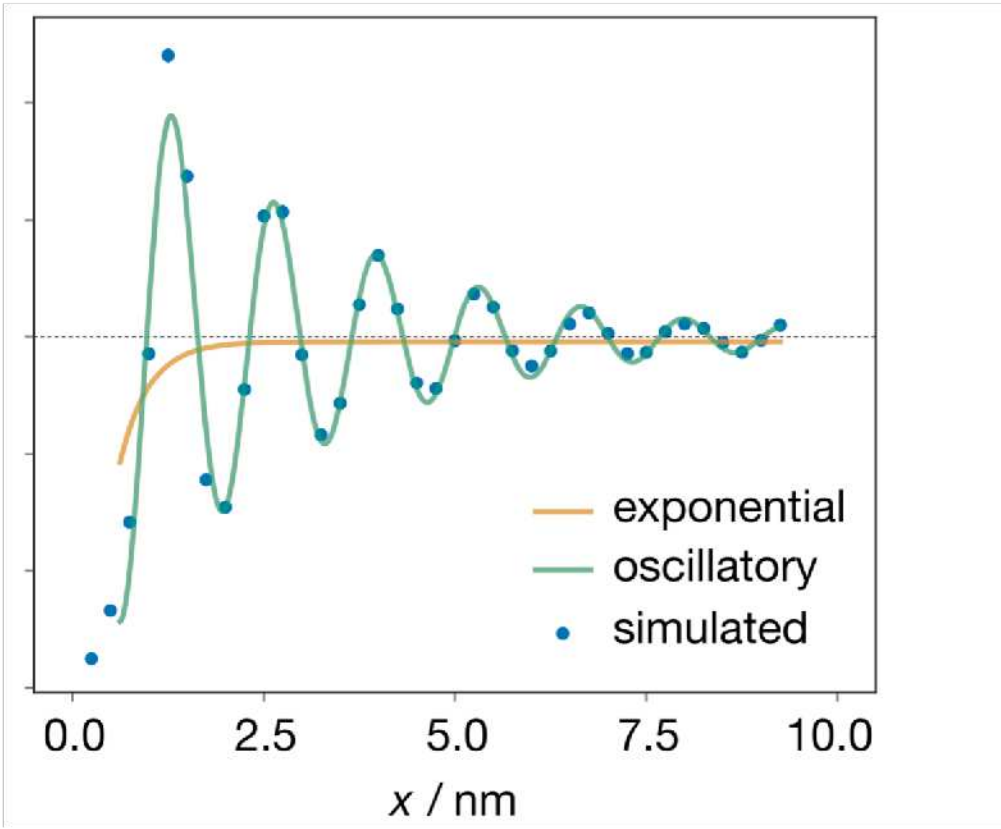


Figure 76: One-dimensional time-averaged mobile-defect distribution for ϵ_r of 1. The point x=0 corresponds to the grain boundary plane. For each set of simulation data, we also plot the maximum likelihood exponential and oscillatory models.

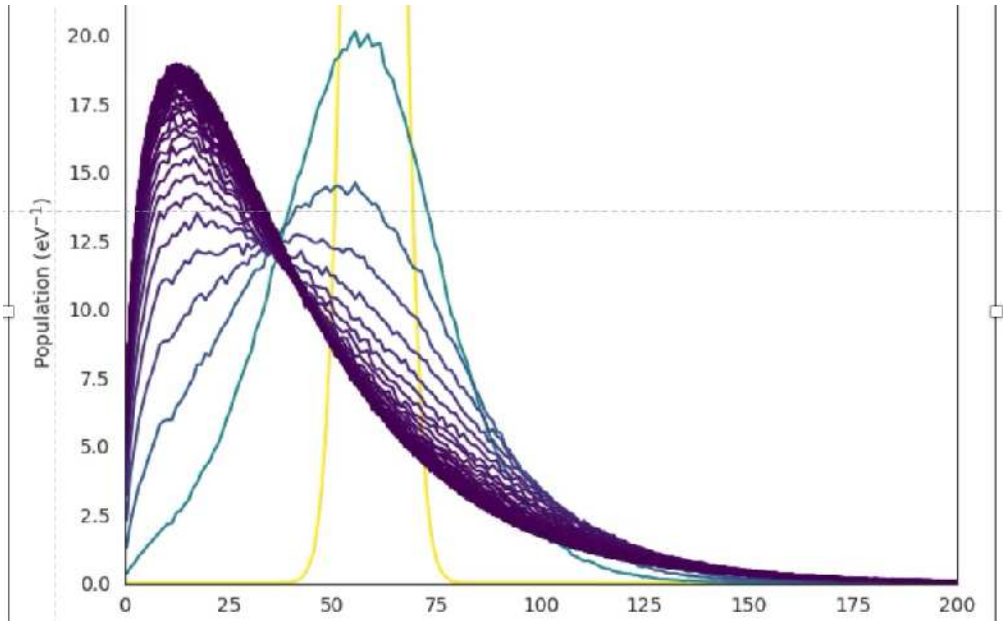


Figure 77: Carrier distribution per energy vs carrier energy at times from 0 to 10 fs.

T1.3.3-3: Development of DMC simulator

The BOLTMC code was validated against the version used to produce the results in [71]. BoltMC calculates the charge transport for charges delocalised throughout the perovskite layer by solving the Boltzmann transport equation. The PPMD code allowed to scale BoltMC by using a domain-specific language written in Python.

T1.3.3-4: Application of DMC simulator to electron transport in perovskites

As stated in the Executive Summary, we showed polaronic effects on scattering and mobilities of charge carriers are more limited than previously claimed concurring with recent angle-resolved photoemission measurements [130]. The results were published in [71]. Another recent paper citing our work noted that polaronic effects have been associated with the mechanism of slow hot carrier cooling in perovskites; however it is not clear if both the polaronic and potential charge effects can operate concurrently [95].

BoltMC is also being used to explore how photo-excited charge carriers in perovskite solar cells reach equilibrium following above band gap excitation in a collaboration with the spectroscopy group of Deschler at Heidelberg who undertake 2-dimensional electron spectroscopy measurements of evolution of athermal carrier distribution functions on timescales $\sim$ 10 fs. Fig. 77 shows the evolution of carrier distribution at times 0 to 10 fs. Cooling over 10-100 fs dominated by carrier-carrier scattering. As the distribution rapidly spreads out and cools, the carrier–carrier scattering rate decreases and phonon scattering becomes more important. This project is motivated by experimental observations [93] showing nanoscale phase impurities are degradation sites.

5.4 Code demonstrators

libNEGF $\rightarrow$ neXGf

Scientific simulations towards exascale

The demonstrator for NEXGF can be found on the gitlab repository under the `neXGf` subfolder. The git repository contains two subfolders, `dftbplus` and `interface`. The folder `dftbplus` contains a copy of the DFTB+ code available at `https://github.com/dftbplus/dftbplus.git` , modified from the release version 22.1. The NEGF library developed in the project can be found in `external/libnegf/`.

This is a modified version of the LIBNEGF released version 1.0.3 available at `https://github.com/libnegf/libnegf.git` .

The code contains all the dependencies (submodules) required for transport calculations. It compiles including or not GPU support. In order to compile the code follow the README file in the main `neXGf` folder. Compilation has been tested on the Juwels (CPU) and Juwels Booster (GPU) machines in Jülich. The modules to be loaded for the two compilations are listed in the aforementioned README file.

The folder `interface` contains example calculations used to produce the figures of the Report. In particular, `elastic_cpu` contains an example of the cSi/aSi/cSi interface studied in the project with the computation of the Transmission function across junction, including scattering due to elastic electron-phonon scatterings, e.g. due to acoustic phonons. The folder `inelastic_cpu` contains an example of the inelastic scattering in a 5 nm long Si system with lateral periodic boundary conditions (supercell) comprising 4×4 conventional cells of Si. The folder `gpu` contains the coherent transmission across the 5 nm long Si system with lateral periodic boundary conditions (supercell) comprising 6×6 conventional cells of Si including GPU acceleration (tested on NVIDIA Ampere 100 cards). The folder `slako` contains materials parameters needed by the DFTB+ code.

All tests run by submitting the batch file `slurmbatch` to the Slurm queue system.

https://gitlab.maisondelasimulation.fr/eocoe-ii/code-demonstrators.git
NEXGF *demonstrator release v1.0.0-pre*

KMC/DMC

Scientific simulations towards exascale

Detailed information on the FMM-KMC code is available from [134] and the code can be downloaded from [132]. This code fulfilled the goals set at the beginning of the project by providing a code suitable for the exascale. We demonstrated that the algorithm scales linearly with the number of charges. This was confirmed numerically by measuring the total time per KMC step, t_{step}, for systems with up to $1.3 \cdot 10^8$ charges. Running in parallel on 8192 cores we find t_{step} = 7.09s.

https://doi.org/10.5281/zenodo.3873308
Version 1

The following is a version of BoltMC to demonstrate the calculation of charge carrier mobilities in semiconductors, in particular polaron mobilities in lead–halide perovskites. The results are shown in [71].

https://gitlab.com/ABW_bath_group/boltmc-eocoe-demo
Version 1

5.5 Summary

We combined ReaxFF MD simulations and ab initio calculations to investigate the intra-gap states of the a-Si:H/c-Si interfaces at different annealing temperatures. We represent the Hamiltonian in the more appropriate basis set of Wannier functions and to make possible future transport calculations of the c-Si/a-Si:H interface made fully ab initio, the WANTRANS code has been implemented to interface the transport code libNEGF and the results from a DFT/wannier90 calculation. Since WANTRANS interface needs some further development before to be applied to the cSi/aSi/cSi structure, we resolved to compute the transport properties across the heterojunction using the density-functional tight binding semi-empirical approach (DFTB). The I-V currents across the junctions for holes and electron currents have been computed using LIBNEGF.

Initially, the flagship code of choice was PVNEGF, developed within the IEK-5 institute part of the FZJ partner. PVNEGF provides photocarrier dynamics (generation, transport and recombination) and the coupling of the charge carriers to photons and phonons. The drawback is that PVNEGF advanced functionalities have been developed based on a simplified geometry. Namely, PVNEGF targets quasi one-dimensional (1D) systems in a typical simplified tight-binding approximation with only two bands. In order to simulate interfaces between amorphous and crystalline silicon (as stated in the task T1.3.1-3 of D1.1), what is needed is a full 3D multiband treatment. Unfortunately, the main developer and scientific lead of the code, Dr. Urs Aeberhard, just left FZJ to work for the industry. Carrying out code development (reworking of the mathematical formalism and implementation of new parallelization scheme) and validation without the scientific lead is unthinkable. Therefore, we changed the flagship code to LIBNEGF developed within the CNR. LIBNEGF already implemented the full 3D multiband treatment and the extension of LIBNEGF by including non-ballistic scattering has been developed within the scope of the EoCoE project.

We performed high-level Quantum Monte-Carlo (QMC) simulations of adsorbed molecules/ions at the surface of carbon materials to drive the choice of the most accurate exchange-correlation functional to be used for the parameterization of the potential for the interactions. The vdW-DF-C09 functional seems to better catch the interactions. The potentials have been used in classical DFT (using MDFT) and molecular dynamics (using METALWALLS) to simulate, respectively: a) the adsorption free energy of the lithium ion on the carbon surface; and b) a system in which the ionic liquid is confined between a pair of planar electrodes held at constant electrical potential, which fluctuates due to the thermal motion of the ions in the liquid.

Algorithmic optimisation of the FMM-KMC code for the KMC simulator has been developed and applied to the study of doped organic semiconductors demonstrating the scaling of the electrostatics library. Since M19, FMM-KMC has been used to simulate all-solid-state batteries. These batteries are safer, easier to process, higher achievable power density and cyclability. We found a decay length that is larger than the classical Debye length mirrors the behavior of concentrated liquid electrolytes, where this phenomenon is termed underscreening.The BOLTMC code for the DMC simulator, which solves the Boltzmann transport equation, has been optimized and used to investigate the polaronic effects on scattering and mobilities of charge carriers in halide perovskites. BoltMC is also being used to explore how photo-excited charge carriers in perovskite solar cells reach equilibrium following above band gap excitation. We observed as the distribution rapidly spreads out and cools, the carrier–carrier scattering rate decreases and phonon scattering becomes more important. The PPMD Python framework has been used for the optimization of the two codes.

6 Water for Energy (T1.4)

T1.4 is divided in five main tasks:

T1.4.1	Exascale hyper-resolution hydrologic simulations
T1.4.2	Hydropower modelling in the Italian Alpine region
T1.4.3	Reduced model for high fidelity hydrodynamics simulation.
T1.4.4	Experimental design for geothermal modelling
T1.4.5	Modelling of geothermal reservoir system

T1.4.1 deals with the modeling of shallow subsurface flow which is of major importance in order to accurately simulate and predict the exchange of groundwater with streams under low-flow conditions, and the transport of energy. The major challenge is the representation of topographically driven groundwater convergence and streamflow generation, and of the geological heterogeneity across a number of space scales ranging from centimeters to thousands of kilometers in case of continental river systems. Predicting hydrologic states and fluxes for hydropower production requires a balanced combination of computational

power to run climatic and hydrological models at the proper scale and enough information concerning the hydropower systems (i.e. location of withdrawals and redistribution, such as reservoir volume and maximum water discharge that can be derived). This is extremely challenging at the pan-European scale. For example, predicting water cycle processes for scientific and applied assessment of the terrestrial water cycle requires a high-resolution modeling framework on the order of 1 to 3km. Thus, HPC technologies are required to enable hyper-resolution simulations at continental scales. This task provide one of the first physics-based hyper-resolution hydrological modelling systems, which provides the current hydrologic status of the terrestrial system and predictions of all pertinent states and fluxes at higher resolution relevant to the energy sector.

The main focus of T1.4.2 was to develop and set up a modelling framework for the simulation of hydropower production in the Italian Alpine Region by means of the HYPERstreamHS hydrological model. To achieve this objective the activities have been organized in several complementary branches. Firstly, a refactoring of HYPERstreamHS code was performed in order to implement MPI parallel computing features and thus improve simulations speed-up. The model was also developed to explicitly represent human infrastructures connected to hydropower systems (e.g., storage reservoir, diversion etc) such as those typically present in Alpine catchments. An intrinsic challenge to this goal lies in the fact that information concerning large hydropower infrastructures is very sparse and often not made publicly available from hydropower companies; for this reason, the explicit modelling of hydropower systems included in HYPERstreamHS is based on publicly available information, which should be easier to retrieve, ensuring wide applicability of the developed framework to other domains. HYPERstreamHS was also modified to ease the coupling with PF-CLM high-resolution gridded outputs. The framework was tested with application to the Adige river basin, a large Alpine catchment (10500 km^2) hosting 40 large hydropower systems. HYPERstreamHS was able to reproduce the historical observed hydropower production at the catchment scale, validating the developed framework. A benchmarking exercise in the Adige catchment aimed at identifying the most suitable framework for streamflow modelling analyzed a variety of combinations between several flow component modules pertaining to both HYPERstreamHS and PF-CLM. An optimal setup was found coupling HYPERstreamHS deep infiltration module with PF-CLM surface flow components. Finally, the modelling framework presented in Subtask T1.4.2-2 was applied to other large Alpine catchments allowing to validate the framework at the scale of the Italian Alpine region.

The main focus of T1.4.3 was to chain hydrodynamic modelling at local scale and medium scale forcing a fine resolution and full hydrodynamics processes hydrodynamic model with Parflow discharge computed on a larger grid. The Telemac model implemented in 2D over the Garonne Marmandaise catchment provides a fine solution with a spatial and temporal resolution on grill cells of about 100m and time steps of 5 minutes. This model was forced at his upstream boundary condition with the discharge computed by Parlow simulation at a 10km resolution, with hourly or daily output. The closest neighbouring point was selected to provide discharge at Tonneins. Chained simulations were carried out over the selected period 2003-2004, with a focus on flood events in February and December. These flood events are overflowing while not reaching extreme values such as those reached for historical events in 1981 or 2021. We noted that ParFlow model provides a nice description of the water cycle at the surface at local scale while it is not calibrated for this purpose. Yet the flow is underestimated, especially near the flood peaks, a five day delay in the time series was also noted in ParFlow outputs. This underestimation at medium scale propagates to local scale in Telemac simulations with underestimated water surface elevation, for instance in the urban area of Marmande that is subject to flooding. This was noted with a comparison to *in-situ* water level observed time series. The simulated water extents in Telemac are also significantly smaller than those simulated with observed forcing at Tonneins. Preliminary conclusions advocate for a local calibration of the friction in ParFlow that would imply to run ParFlow simulations again. An off line solution to rescale the ParFlow discharge with information on the river width may also be worth investigating. Finally, the correction of inflow and friction in Telemac alone, thanks to the assimilation of observed water levels at Tonneins and Marmande should be investigated. While, this solution would improve the simulation in re-analysis, the improvement in forecast relies on the stationary of the error in the forcing. Since this error seems to be

quite stationary over the 1997-2006 simulated period in ParFlow for the Garonne catchment, this strategy seems promising.

T1.4.4 deals with experimental design for predicting borehole locations in geothermal reservoirs in order to decrease the uncertainty in parameter estimation, which, in turn, improves the quality of the reservoir simulation.Drilling boreholes during exploration and development of geothermal reservoirs not only involves high cost, but also bears significant risks of failure. In geothermal reservoir engineering, techniques of optimal experimental design (OED) have the potential to improve the decision making process. Earlier publications explained the formulation and implementation of this mathematical optimization problem and demonstrated its feasibility for finding borehole locations in two- and three-dimensional reservoir models that minimize the uncertainty of estimating hydraulic permeability of a model unit from temperature measurements. Subsequently, minimizing the uncertainty of the parameter estimation results in a more reliable parametrization of the reservoir simulation, improving the overall process in geothermal reservoir engineering. Various OED techniques are implemented in the Environment for Combining Optimization and Simulation Software (EFCOSS). To address problems arising from geothermal modeling, this software framework links mathematical optimization software with SHEMAT-Suite, our geothermal simulation code for fluid flow and heat transport through porous media.
Within EoCoE-II we want to improve the computational performance of the existing OED workflow for geothermal modelling. This will enable us to apply it to more detailed numerical models of geothermal reservoir systems. Higher spatial resolution of numerical models, in turn, improves validity of OED results. Moreover, we aim at extending the OED workflow by additional functionalities, such as predicting optimal borehole depth or length of temperature logs or applying it for inverting parameters other than permeability (e.g., thermal conductivity). By simulating different model scenarios, we analyze the sensitivity of the OED approach to factors such as measurement errors, prior data density or quality. Our results improve the understanding of the presented workflow and show that the application of OED techniques to practical problems in geothermal reservoir engineering is feasible.
In Task 1.4.5 we carry out the application of optimal experimental design (OED) on a real world geothermal reservoir system model. The aim is to predict optimal locations for additional exploration wells for temperature measurements. We define the optimal location as the one, where the information content of the data (i.e temperature) is highest with respect to certain reservoir rock parameters (e.g. permeability, thermal conductivity); i.e, the rock parameters will be estimated with least uncertainty when inverting data from the optimal location. The quality of geothermal reservoir models and hydrothermal flow simulations highly depends on the quality of the subsurface parametrization. Reliable determination of physical rock properties such as porosity, permeability or thermal conductivity is crucial for the simulation results as they are directly linked to the observations (e.g., temperature, pressure, flow rate). This task will apply the OED workflow, which is improved and extended in Task 1.4.4. On the one hand, its results will demonstrate the feasible and successful applicability of the OED workflow to geothermal reservoir modeling. On the other hand, the results will provide new insights for the simulated geothermal reservoir system.

6.1 Task 1.4.1: Scientific results

The main task T1.4.1 is subdivided in three subtasks:

T1.4.1-1	**Model Setup for hyper-resolution hydrologic simulations**. This task aims to demonstrate the application of physics-based fully distributed hydrological ParFlow model with explicit groundwater representation over pan-European domain at 3km resolution.Hydrological simulations were performed using the integrated Terrestrial Systems Modeling Platform (TerrSysMP), consisting of the three-dimensional surface-subsurface model ParFlow, and the Community Land Model, CLM3.5 (CLM).

T1.4.1-2	**Model Evaluation and validation**. The model results with various *in-situ* and remote sensing observations were evaluated to investigate sources of uncertainties and identify areas for improvements, focusing on the impacts of groundwater representation on the spatial variability and dynamics of hydrological variables such as soil moisture, water table depth and its effects on ET fluxes and river flows.
T1.4.1-3	**Uncertainty Quantification using Data Assimilation**. To provide reliable hydrologic data and information for management of water resources, this task focuses on data assimilation (DA) techniques which combine observations and model to improve model estimates.

Task 1.4.1-1: Model Setup for hyper-resolution hydrologic simulations

This task uses the ParFlow which is an integrated groundwater-surface water model [79] in which ParFlow simulates 3D groundwater, surface and unsaturated flows and CLM (common land model) simulates the land surface energy balance [33]. ParFlow is a computationally advanced code and has been extended by a terrain-following grid capability enabling continental scale simulations [100].
In Deliverable 1.2, we described the implementation of standalone CLM model for the CORDEX European model domain with a spatial resolution of 0.0275^o $(\tilde{3}km)$ in details. Here we provide the details of the Parflow model setup. To setup the model, we generated the land surface static input data which consists of topography, soil properties (soil color, percentage sand and clay), dominant land use types, dominant soil types in the top layers, dominant soil types in the bottom layers, subsurface aquifer and bedrock bottom layers and physiological vegetation parameters (Figure 78). Digital elevation model (DEM) data were acquired from the 1 km Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) [34] as shown in Figure 78a. Using the 1-km DEM and a pan-European River and Catchment Database available from Joint Research Center (CCM; [159]), a hydrologically consistent DEM was generated as input to calculate D4 slopes (in x and y directions) from topography information using the stream following algorithm developed by [9], which are used to specify the connected drainage network in the ParFlow model.
The land cover data was based on the Moderate Resolution Imaging Spectroradiometer (MODIS) data set [52] (Figure 78b). The properties of individual sub-grid tiles, such as leaf area index, stem area index, and the monthly heights of each land cover, were calculated based on the global CLM3.5 surface data set [116]. The aquifer network was added to the ParFlow model in order to better model the relationship between the surface and subsurface water flow where the aquifer network serves as a conduit for lateral groundwater transport through the continent. The subsurface aquifer information was derived from the BGR International Hydrogeological map of Europe (IHME; [41]. For ParFlow, bedrock geology was developed by combining the IHME hydrogeological information with the CCM river database as a proxy for the alluvial aquifer system, where the river database is converted from D8 to D4 flow in order to be compatible for the ParFlow overland flow (Figure 78c). We assumed that alluvial aquifers underlay or are in close proximity to existing rivers. To provide soil texture data in the model (Figure 78d–Figure 78f), sand and clay percentages were prescribed based on pedotransfer functions from Schaap and Leij (1998) for 19 soil classes derived from the FAO/UNESCO Digital Soil Map of the World [10].
In addition to the above static input data, the high-resolution atmospheric reanalysis COSMO-REA6 dataset [16] from the Hans-Ertel Center for Weather Research (HErZ; [144]) for the time period from 1997 to 2006 was used as the atmospheric forcing for ParFlow. The essential meteorological variables applied in this study, such as barometric pressure, precipitation, wind speed, specific humidity, near surface air temperature, downward shortwave radiation and downward longwave radiation were downloaded at 1-h temporal resolution from the German Weather Service (DWD; `https://opendata.dwd.de/climate_environment/REA/COSMO_REA6/`). The COSMO-REA6 reanalysis is based on the COSMO model and

available at 0.055° (about 6 km) covering the CORDEX EUR-11 domain [64] and was produced through the assimilation of observational meteorological data using the existing nudging scheme in COSMO with boundary conditions from ERA-Interim reanalysis data.

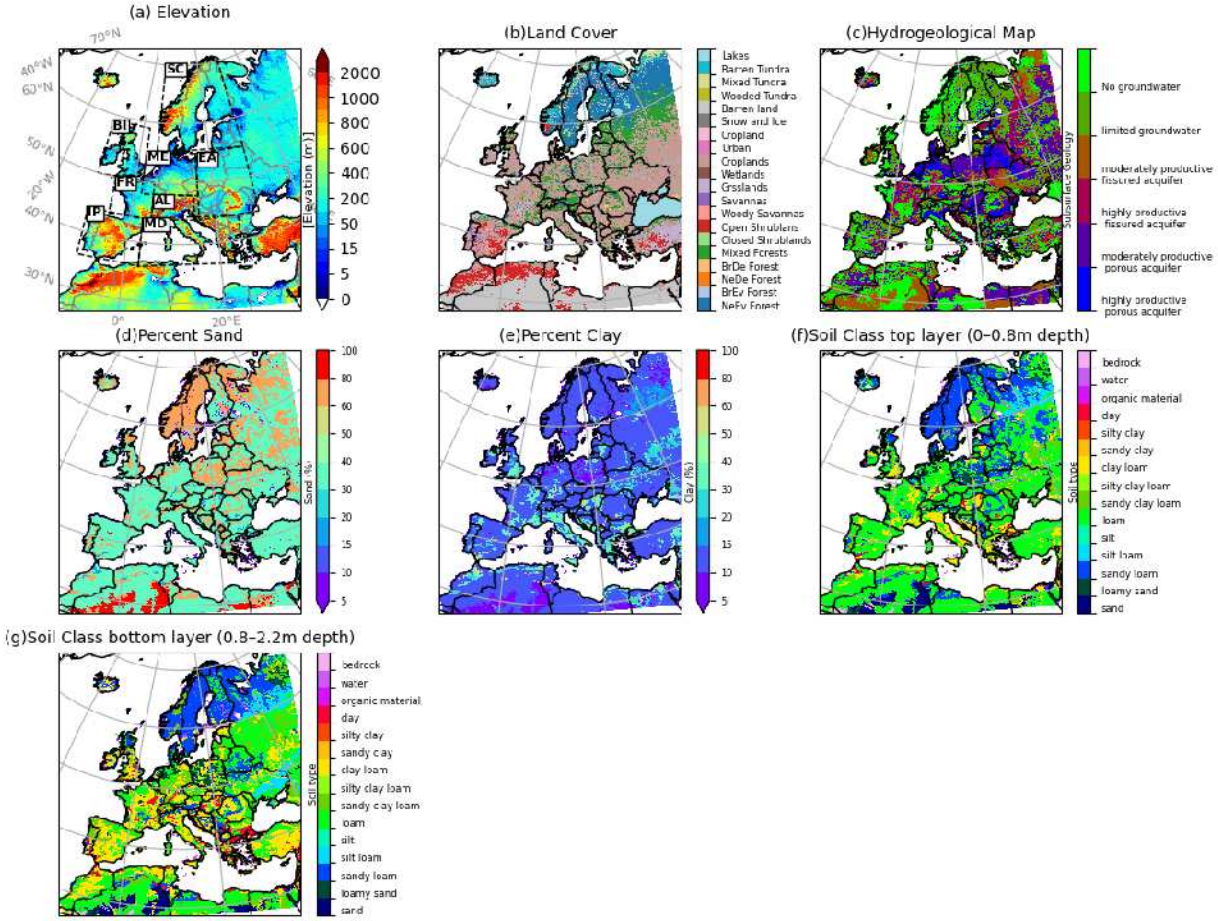


Figure 78: Model surface input data: a) USGS GMTED2010 DEM, b) dominant land use type based on MODIS data, c) percent sand content, and d) percent clay content based on global FAO soil database. The inner boxes in (a) show the boundaries of the PRUDENCE regions (FR: France, ME: mid-Europe, SC: Scandinavia, EA: Eastern Europe, MD: Mediterranean, IP: Iberian Peninsula, BI: British Islands, AL: Alpine region; Christensen et al., 2007).

We performed a 10-year simulation using the ParFlow model to evaluate the model performance of hydrologic states and fluxes over the EURO-CORDEX domain (Figure 78). The model was run at an hourly time step and at a horizontal resolution of 3 km resulting in 1592 x 1540 grid cells. Vertically, the model consists of 15 layers (upper 10 soil and bottom 5 bedrock layers) of variable depths with a total depth of 60 m. Distributed parameters describing the soil properties, saturated hydraulic conductivity, van Genuchten parameters, and porosity were assigned to each soil class and were based on the pedotransfer functions from [137]. Using this modeling setup, a steady state simulation of the hydrological variables of ParFlow was first conducted (spinup run) to reach a dynamic equilibrium. A spinup of nine years, by simulating the year 1997 nine times, was performed in order to obtain a stable and reasonable distribution of the initial state variables. The steady-state initial conditions were then used for model simulations over the period from 1997 to 2006. To tackle the computational challenge of simulating 3-D subsurface flow, ParFlow is designed for high-performance computing infrastructures with demonstrated performance (e.g., [24, 80]), where the 3-D variably saturated subsurface and lateral groundwater flow is simulated using a parallel Newton-Krylov nonlinear solver [5], [75] and multigrid-preconditioners. In this application, ParFlow was applied in its GPU-parallel version[68] (See Section 6.6).

Task 1.4.1-2: Model evaluation and validation

This task evaluate the model results with various *in-situ* and remote sensing observations to investigate sources of uncertainties and identify areas for improvements, focusing on the impacts of groundwater representation on the spatial variability and dynamics of hydrological variables such as soil moisture, water table depth and its effects on ET fluxes and river flows. The main features of this task include: (1) explicit simulation of lateral groundwater flow, groundwater discharge and recharge, (2) extensive evaluation of simulated riverflow, evapotranspiration (ET), and surface soil moisture (SM) and water table depth (WTD) using *in-situ* and remote sensing data observations , and (3) to provide higher resolution hydrological states fluxes over continental Europe for longer time scale which are useful for understanding the long-term climatic changes and can serve as a benchmark for evaluating hydrological extremes in future studies.
In Deliverable 1.2, we reported the evaluation of community land model (CLM3.5), which was also selected

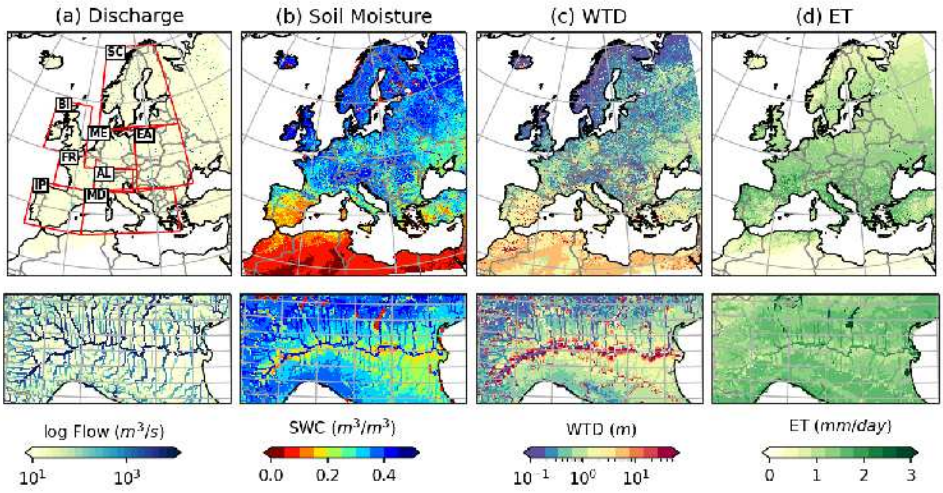


Figure 79: (a) Maps of Euro-CORDEX domain (1544 x 1592 grid cells) showing the spatially average distribution of (a) discharge, (b) surface soil moisture, (c) water table depth and (d) evapotranspiration (1997 - 2006) and close-up over Alpine (AL) region estimated by Parflow model. The inner boxes correspond to PRUDENCE regions with abbreviated letters indicating names of the regions (FR: France, ME: Mid-Europe, SC: Scandinavia, EA: Eastern Europe, MD: Mediterranean, IP: Iberian Peninsula, BI: the British Isles, AL: Alpine region.

to provide a flexible framework for generating hydrologic predictions for water resource management applications. Both models were compared to each other to evaluate their strengths and weaknesses and use these results to improve and calibrate the hydrological models. In particular, the influence of groundwater representation on soil moisture, evapotranspiration, total water storage, water table depth and groundwater recharge/discharge were explored through the comparison of multi-model simulations using the standalone Community Land Model (CLM) and the ParFlow hydrologic model.
In this deliverable, we described in details more the ParFlow model performance for water balance components such as streamflow, SM, ET and WTD. Because of the explicit lateral groundwater and surface flow representation, the ParFlow model is able to resolve small-scale spatial variability in hydrological states and fluxes which is strongly related with the river network and topography as shown in Figure 79. We compared these variables with both *in-situ* observation, remote sensing and reanlaysis datasets to discuss the model performance at different spatial and temporal scales and for different regions as described below in more detail. For the regional analysis, the results are presented for eight predefined analysis regions from the “Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks

and Effects" (PRUDENCE) project ([27]) as shown in Figure 79a. We referred to these regions as the "PRUDENCE" regions.

Streamflow evaluation

Daily river flow observations over Europe were obtained from the Global Runoff Data centre (GRDC, obtained via `https://www.bafg.de/GRDC/EN/Home/homepage_node.html`) for more than 2000 gauging stations. For model validation of river flow, predicted streamflow may be extracted at the grid cell location of the gauging station where discharge measurements are available. However, because of the relatively coarse resolution of the model with respect to the river network, the gauging station locations might be slightly off with respect to the modelled river network. Therefore, these locations were adjusted to the nearest locations on the model river network (centre of the 0.0275o cell) through comparison of the actual drainage areas with the modelled drainage areas. Only those stations were selected for model validation where drainage area differences were less than 20% and more than 50% of data is available for the time period of 1997–2006. Additionally, we only selected stations where the upstream drainage area is greater than 1000 km^2. This resulted in a selection of 176 gauging stations located along many rivers and mostly concentrated in central Europe (Figure 80).

In evaluating model performance pertaining to mean flow, comparison of the observed and simulated mean flow in the simulation period showed that ParFlow appropriately reproduced the mean flow, where the percentage bias (PBIAS) is below 20 % for 48 % of stations and only 8 stations show a higher bias (PBIAS > 50 %) between the observed and simulated monthly river flow (Figure 80a). To better understand the seasonal variability of the simulated streamflow, 16 stations along large rivers across different climatic zones, with a total drainage area upstream of the gauging station greater than 5000 km^2, were selected and compared with monthly observed streamflows for the simulation period (Figure 80b).

Overall, the comparison shows that the streamflow dynamics are well captured for the selected 16 large rivers, however, there is an overestimation of the winter flow by the model and an underestimation of summer flow for most gauging stations. The overestimation of peak flow is more pronounced in wet years (for example years 2001 and 2002), whereas low flows in summer are mostly underpredicted in dry years (for example, years 2003 and 2004). The discrepancy between the simulated and observed flow may be related to the following: coarse river resolution in the model, human impacts on discharge regimes – particularly for highly regulated rivers through reservoir regulations, and power generation or groundwater extraction (e.g. in the case of Rhine, Elbe and Danube rivers). In addition, the simulated flow is overpredicted for both River Kemijok (Finland) and Nemunas (Lithuania) in northeastern Europe across all years (Figure 80a).

Soil Moisture evaluation

The simulated surface soil moisture (SM) from ParFlow model was evaluated by comparing with the global satellite observations of SM from the European Space Agency Climate Change Initiative (ESACCI; [38]). The globe ESACCI SM product was created at 0.25^o resolution by combining the active and passive microwave sensors providing a homogeneous and the longest time series of SM data to date, starting from 1979. The model results of surface SM were also evaluated with the 3 km European surface SM reanalysis (ESSMRA) datasets [110] which was created through assimilation of the ESACCI data into the land surface model CLM3.5 (Community Land Model, version 3.5[116], driven with the same meteorological forcing and static model inputs as used for ParFlow.

Figure 81 shows the large-scale spatial patterns of surface SM over the study domain simulated by ParFlow and compared with ESSMRA [110] and ESACCI datasets. Since ESSMRA data is available from the year 2000 onwards, the comparison of mean surface SM from ParFlow with ESSMRA and ESACCI were made for the period of 2000–2006. As shown in Figure 81, ParFlow shows slightly higher SM than both ESSMRA and ESACCI over most parts of Europe, except in the southern parts of the domain. The difference is explained by the shallow groundwater system simulated only by ParFlow, which contributes to the saturation of the deeper soil layers leading to higher soil water content, whereas the standalone CLM3.5

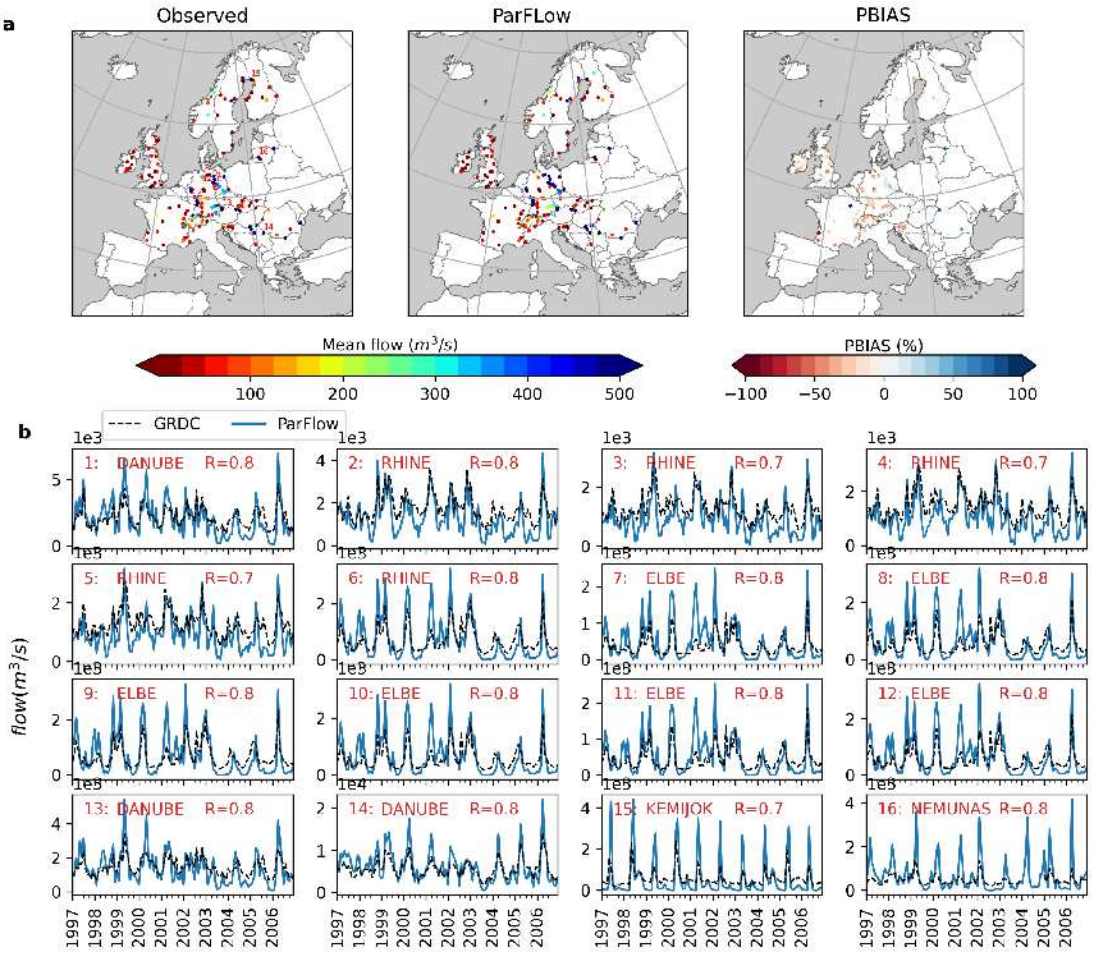


Figure 80: Comparison of observed and simulated average discharge and the percentage bias in monthly discharge (PBIAS) for 176 gauging stations. (b) Comparison of time series of observed and simulated discharge for selected large rivers with drainage areas greater than 50,000 km^2. The selected gauges in (b) have red labels in the left panel of (a).

model applies a simple approach to simulate groundwater recharge and discharge processes in a single column and neglects explicit lateral groundwater flow. Furthermore, Figure 81b shows the comparison of the spatial distribution of SM simulated by ParFlow with ESACCI and ESSMRA as violin plots. The spatial distributions of SM simulated by ParFlow over PRUDENCE regions shows consistently higher mean SM than both CLM3.5 and ESACCI except for the IP region where SM simulated by ParFlow is lower than both datasets (Figure 81b). We observe that the distribution range of ParFlow simulated SM in most regions is quite wide when compared to both ESSMRA and ESACCI, indicating higher spatial variability is simulated by ParFlow.

To evaluate the model performance in simulating average, wet and dry periods, a comparison of monthly time series of SM anomalies at an aggregated regional scale is undertaken. The SM standardized monthly anomalies are calculated by subtracting the long-term mean of the complete time series from each month and then dividing by the long-term standard deviation for the period of 2000–2006. Our results show that ParFlow agrees well with both CLM3.5 and ESACCI anomalies over the simulation period (Figure 81c).By looking at the correlation coefficient (R) values for different regions, the results show that the correlation pf ParFlow with ESSMRA is higher than ESACCI (i.e. 0.65<R>0.85 and 0.18<R>0.77 for ESSMRA and

ESACCI, respectively), primarily due to the direct impact of identical forcings used for both modeling setups.

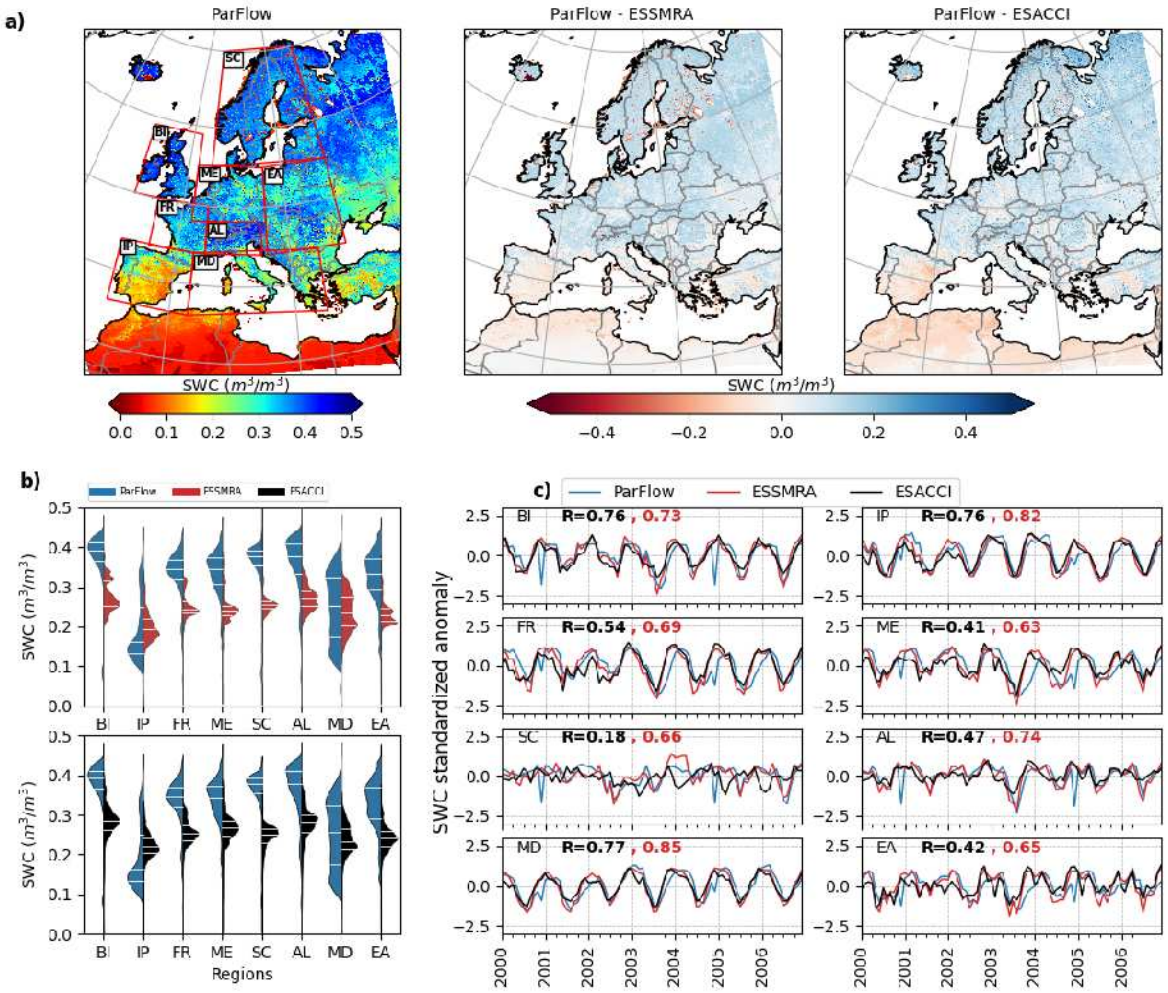


Figure 81: (a) Evaluation of time averaged surface soil moisture (SM) simulated by ParFlow with ESSMRA and ESACCI datasets over the time period of 2000-2006. (b) Violin plots showing comparison of spatial distribution of time averaged surface SM simulated by ParFlow with ESSMRA (upper plot) and ESACCI (lower plot) over PRUDENCE regions. The violin plots show the estimated kernel density distribution as well as the median, the lower and upper quartile (white lines). (c) Comparison of spatially aggregated surface SM monthly anomalies estimated by ParFlow with ESSMRA and ESACCI datasets for PRUDENCE regions. The SM standardized monthly anomalies in (c) were calculated by subtracting the long-term mean of the complete timeseries from each month and then dividing by long-term standard deviation for the period of 2000–2006.

Evapotranspiration evaluation

Figure 82 compares the simulated monthly ET from ParFlow with observed ET from 60 eddy covariance tower stations from the FLUXNET database [118] in order to evaluate the model's ability to capture seasonal ET dynamics. The ParFlow model performs well and shows reasonable consistency for all stations with respect to monthly ET, with R values greater than 0.6 (Figure 82a) for all stations. To better understand the agreement between seasonal dynamics of simulated ET with observations, we compared the cumulative distribution of monthly ET for different seasons with observations over all stations in Figure 82b. The

differences between ParFlow simulated ET and FLUXNET are smaller for winter (DJF), spring (MAM) and autumn (SON) seasons (on average 0.11 mm d^{-1}, 0.18 mm d^{-1}, 0.13 mm d^{-1}, respectively) but larger for summer (JJA) season (0.39 mm d^{-1}) over most stations.

While ParFlow shows acceptable performance for all stations, the relatively small number of stations limits

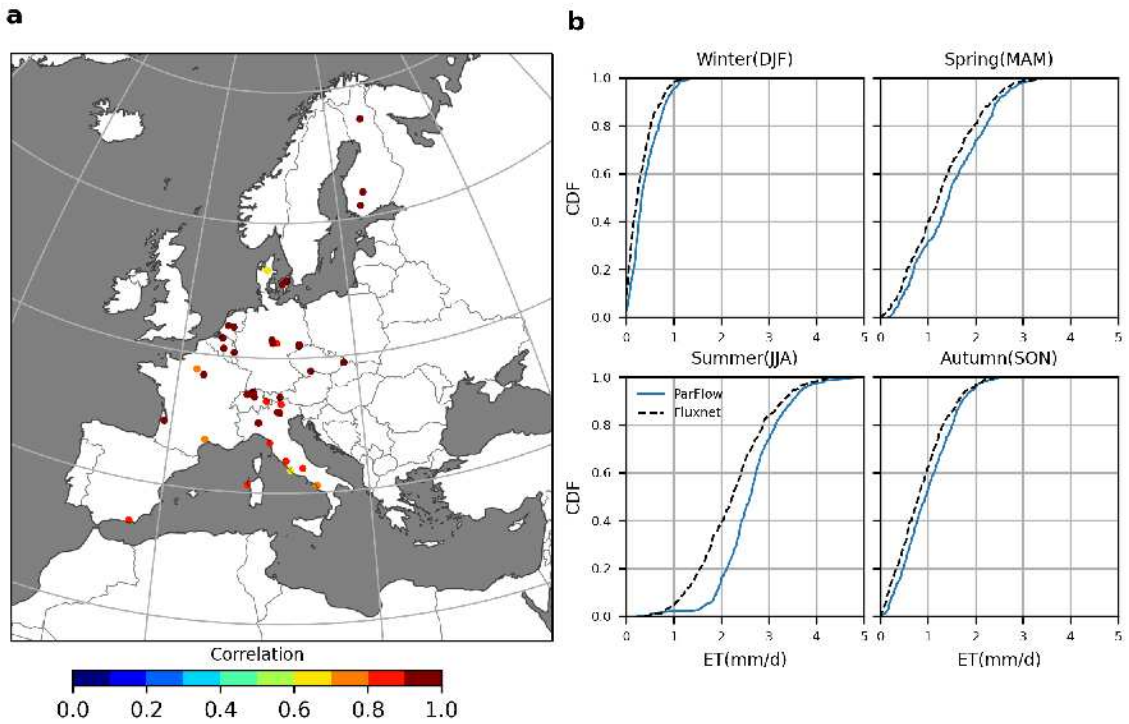


Figure 82: Evaluation of ParFlow simulated monthly evapotranspiration (ET) with ground-based observation from 60 eddy-covariance FLUXNET stations. (b) Comparison of Cumulative distribution of seasonal ET estimated by ParFlow with FLUXNET stations.

a comprehensive evaluation of model performance over the study domain. Therefore, ParFlow performance in simulating the spatial variation in ET is further evaluated with the remotely sensed Global Land Surface Satellite (GLASS, [88] and Global Land Evaporation Amsterdam Model (GLEAM; [97] datasets. The ET data from GLASS is calculated by a multimodel ensemble approach merging five process-based ET datasets [89], while GLEAM is based on water balance method and uses Priestley–Taylor equation and set of algorithms to estimate ET separately for both soil and vegetation [97]. The spatially distributed ET simulated by ParFlow and its difference with both GLASS and GLEAM estimated ET are shown in Figure 83. The ParFlow simulated ET is lower than both GLASS and GLEAM ET over most areas in the EURO-CORDEX domain. However, the difference is smaller between ParFlow and GLEAM ET (i.e. average difference is -0.09 mm d^{-1}), than with the GLASS ET (i.e. the average difference is about -0.30 mm d^{-1}) over the study domain.

Despite the differences in spatial patterns, the time series of spatially aggregated ET simulated by ParFlow over PRUDENCE regions is highly correlated with both GLASS and GLEAM dataset (R > 0.9) as shown in Figure 83b. The main differences in ET are mostly detected in summer where GLASS estimated ET is larger than both GLEAM and ParFlow simulated ET. But the fact that GLASS has large positive bias over summer when compared with FLUXNET data (not shown here) suggests that GLASS ET data has relatively large uncertainties. We also noted relatively large negative differences with GLEAM in areas of complex topography which may be partly caused by the downscaling of GLEAM data from coarse spatial resolution ($0.25°$) to 3km resolution.

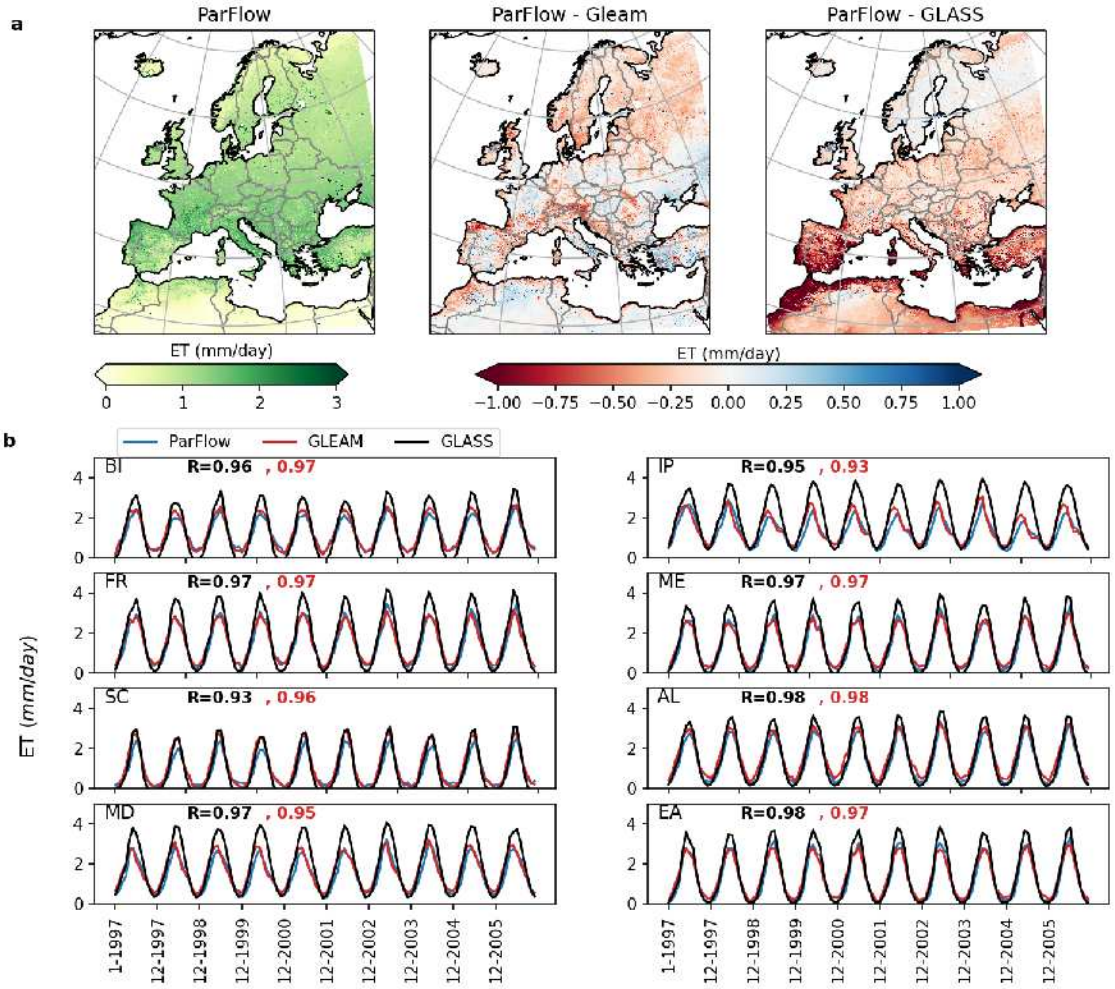


Figure 83: (a) Evaluation of time averaged surface evapotranspiration (ET) simulated by ParFlow-EU3km with GLEAM and GLASS datasets over the time period of 1997-2006. (b) Comparison of spatially aggregated monthly ET estimated by ParFlow-EU3km with GLEAM and GLASS datasets for PRUDENCE regions. R values in red color show the correlation of ParFlow with GLEAM and in black color R values represent correlation between ParFlow and GLASS dataset.

Water table depth evaluation

To validate the model outputs for water table depth (WTD), we collected monthly well observations at 5,075 groundwater monitoring wells distributed over Europe from 1997 to 2006. The WTD measurements were obtained either from web services or by request from governmental authorities in eight countries (France, Spain, Portugal, the Netherlands, the UK, Sweden, Denmark and Germany) with most stations concentrated in Germany. The WTD measurements were first converted to 3 km gridded WTD data by averaging WTD data from all the wells that lie within the same 3 km grid cell. Additionally, we selected only those grid cells where ParFlow simulated WTD < 10 m. This resulted in 2,346 grid cells which were then used to evaluate the ParFlow results. It should be noted that the reference surface elevations provided with the groundwater observation data used in this study were not consistent across regions which makes it difficult to derive the absolute values of WTD for comparison with the model simulated WTD. Therefore, standardized anomalies were calculated from observed groundwater data in order to reduce errors related to inconsistencies in the observations. Figure 84 shows the temporal correlation coefficients between the monthly time series of WTD anomalies from ParFlow and observations over Europe. Overall 80 % of grid

cells show R values above zero and 20 % result in R > 0.5 with the simulated anomalies (inset Figure 84) indicating that in general ParFlow model appropriately captures the seasonal cycles. Performance of ParFlow in simulating WTD anomalies also varies across PRUDENCE regions, with an average R value ranging between 0.21 to 0.34. As an example of ParFlow performance with highest and lowest R values across different regions, we show the time series comparison of selected individual stations. This comparison indicates that the weaker correlation in WTD anomalies by ParFlow for some grid cells are related to less fluctuations in the observed WTD anomalies than ParFlow. These discrepancies might be related to uncertainties in aquifer parameterization used in the ParFlow or the limitations in model resolution such that local aquifers in areas with complex topography cannot be captured. Additionally, model evaluation can be hampered by the challenges associated with groundwater monitoring (e.g. [60]). For example, the observations might be biased if they are located towards river valleys, in low elevations, in areas with confined or perched aquifer systems or in coastal areas. In addition, the comparison of the resolved simulated head, averaged across 3 km, with the point scale observation head, which is highly governed by local surface elevation, can bring about misleading results and amplify inaccuracies. Water table depth observations can also be impacted by pumping which may not be known for many locations.

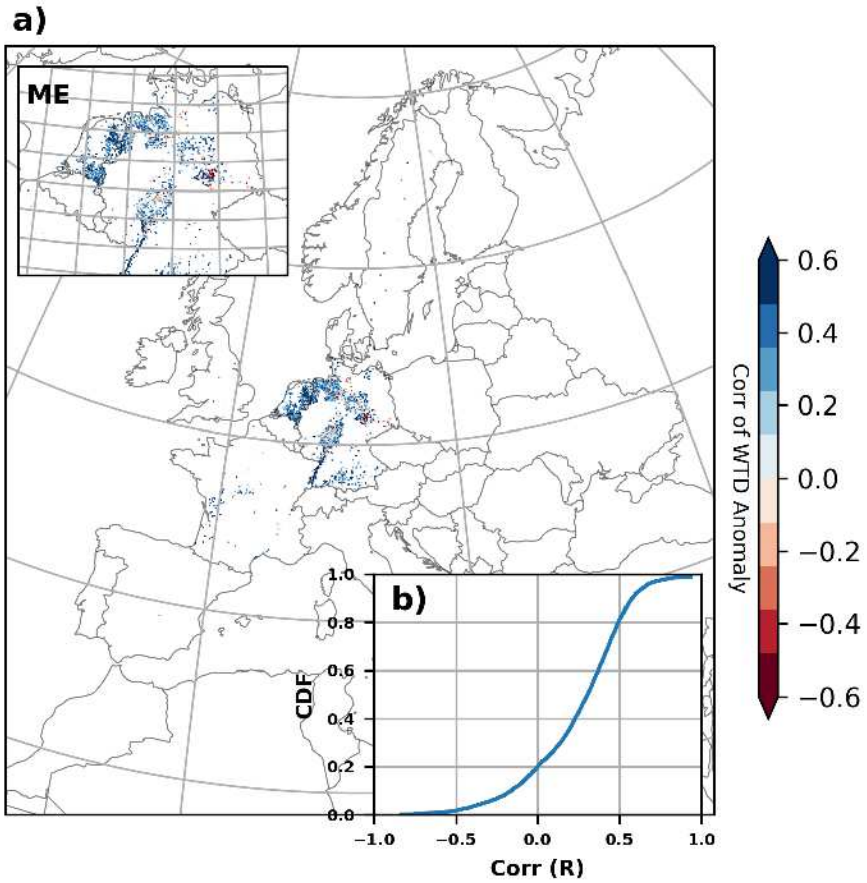


Figure 84: (a) Correlation map between *in-situ* water table depth (WTD) anomalies and ParFlow-EU3km model. (b) Cumulative distribution function (CDF) of correlation coefficient of ParFlow with observed WTD anomalies. The inset in (a) shows a zoom of the Mid-Europe (ME) region.

Task 1.4.1-3: Uncertainty Quantification using Data Assimilation

This task focuses on using the data assimilation approach to improve surface soil moisture estimates from the model. In deliverable 1.2, we reported the results of this task more in details. Here we only provide the summary of our results.
The standalone Community Land Model(CLM) (V3.5) [116] coupled with Parallel Data Assimilation Framework (PDAF; [112]) was setup to provide hydrologic predictions for water resource management applications with higher accuracy. With data assimilation (DA), simulated soil moisture distributions from the land surface models which provide spatially and temporally continuous information can be improved by ingesting observations either from satellite or *in-situ* observations. Since we performed our simulations at high spatial resolution and at continental scale, we selected the CLM–PDAF modeling framework [83], which is designed for high-performance computing infrastructures and can efficiently cope with the high computational burden of ensemble-based data assimilation. [83] showed the efficient use of parallel computational resources by CLM–PDAF, which is needed to simulate predicted states and fluxes over large spatial domains and long simulations. Using the CLM-PDAF setup, we generated a 16 years (2000–2015) European high-resolution surface soil moisture reanalysis (ESSMRA) dataset. Satellite derived soil moisture data were assimilated into the land surface model using an ensemble Kalman filter data assimilation scheme, producing a 3 km daily soil moisture reanalysis dataset. The assimilated surface soil moisture was compared with other global gridded products and *in-situ* station data from the International Soil Moisture Networks to evaluate the assimilated SM data in capturing daily, inter-annual, intra-seasonal patterns, and extreme events under different climatic conditions. This product overcomes the shortcomings of sparse spatial and temporal datasets and provides a better estimate of SM than obtained only by modeling or by sparse observations alone. Comparison with the existing reanalysis products and independent *in-situ* soil moisture observations showed that data assimilation performed well in capturing daily, inter-annual, intra-seasonal patterns, and extreme events under different climatic conditions. The dataset produced in this task provides daily surface soil moisture at a high spatiotemporal resolution which is important not only for research in agriculture, flood and drought forecast, land cover changes, and modeling of the regional carbon and water cycles, but can also be useful to validate soil moisture estimates from other modeling studies.

6.2 Task 1.4.2: Scientific results

The main task T1.4.2 is subdivided in four subtasks:

T1.4.2-1	**Refactoring of HYPERstreamHS model**. In this task the HYPERstreamHS hydrological model is substantially refactored in order to achieve the two main goals of increasing its efficiency by implementing dual-layer MPI parallel coding, and to include Human Systems (HS) modules aimed at the explicit representation of hydropower systems in hydrological modelling.
T1.4.2-2	**Set-up of the model in the Adige river basin**. The framework developed in T1.4.2-1 was set-up and tested in the Adige river basin, a large alpine watershed whose streamflows are greatly affected by hydropower activities. The model was able to reliably reconstruct observed time series of daily streamflow at selected nodes and monthly hydropower production at the catchment scale.
T1.4.2-3	**Hydrological benchmarking exercise in the Adige river basin**. In view of a coupling between CLM3.5 gridded runoff products and HYPERstreamHS, this subtask investigated the best parametrization to correctly reproduce observed streamflows, based on the prediction of surface and sub-surface flow components.

T1.4.2-4	**Set-up of the model in the Italian Alpine Region**. The framework developed in T1.4.2-1 and tested in T1.4.2-2 was then extended to the Italian Alpine Region, collecting critical information on the existing hydropower systems and allowing to reproduce historical observations of hydropower production.

T1.4.2-1: Refactoring of HYPERstreamHS model

The model was refactored to implement MPI features for parallel computing and to include specific modules dealing with the detailed simulation of Human Systems (HS) including those related to hydropower production. In particular, the model adopts a dual-layer MPI parallelization that splits among the available processors the workload involved in the hydrological kernel computations finalized at generating runoff at modelling nodes (first layer) and that of the multiple runs required for inverse modelling and calibration procedures aimed at identifying the optimal set of parameters to be employed in the hydrological kernel (second layer).

Scalability tests (i.e., decrease in computational time gained by increasing the number of processors used, where the ideal ratio is 1:1) performed on multiple model configurations including different numbers of simulation nodes and of HS included into the conceptual model, highlighted that an increased number of nodes negatively affects the model's scalability, hence decreasing the overall speed-up gained by using an increased number of processors. This latter effect is due to a larger amount of time being spent in communication among the processors during the execution of the river routing module of the model. Moreover, the presence of an increased number of HS further decreases the overall speed-up, due to the mandatory serial computations required by HS module. This showed that the achievable speed-up is strictly dependent on the modelled system, however good scalability could be achieved up to 8 to 16 processors, depending on the modelled configuration. Results are summarized in Figure 85. The second MPI layer adopted to

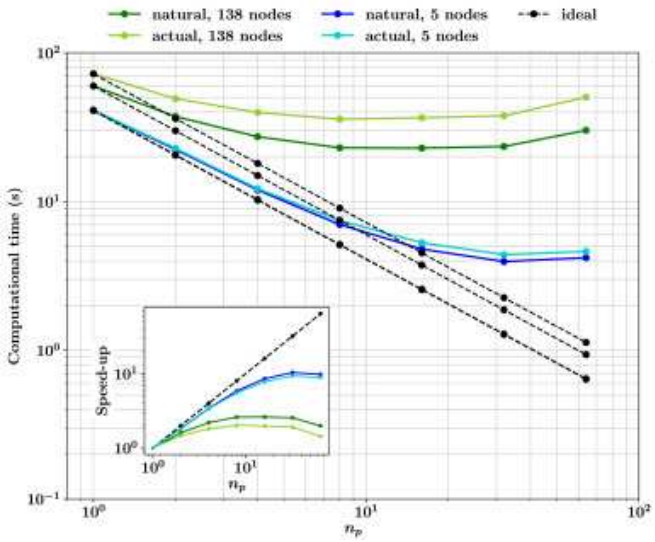


Figure 85: Computational time of first MPI layer considering 5 and 138 network nodes under both natural and actual (i.e., considering the presence of hydropower systems) conditions, respective, with reference to the study conducted in the Adige river basin. The inset in the lower-left corner shows the corresponding speed-up. Ideal computational time and speed-up are represented with dashed lines. *(reproduced with permission from [6])*

subdivide the workload during model calibration allows for further exploitation of the available resources, by

subdividing the several forward executions of the model required by using different calibration algorithms (i.e., PSO and LHS) among the available processors. In this case, as it can be observed in Figure 86, good scalability could be kept up to 64 processors for both calibration algorithms, gaining more than 40x speed up.

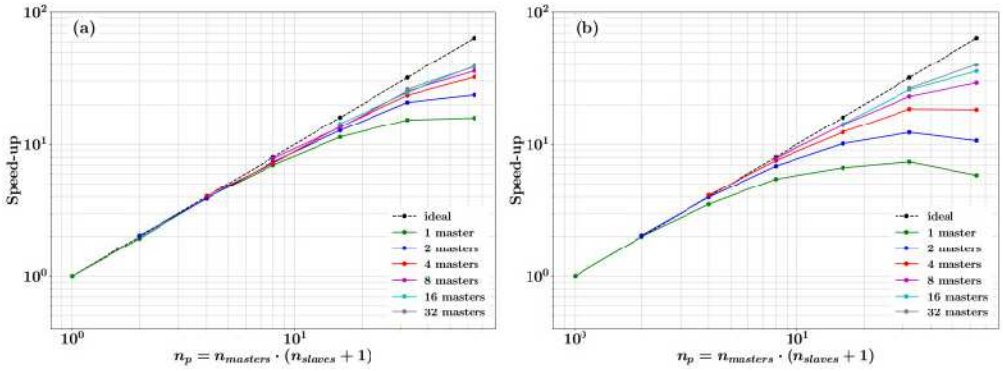


Figure 86: MPI dual-layer speed-up of parallel LHS scheme for 5 (a) and 138 network nodes (b), respectively, with reference to study conducted in the Adige river basin. n_p represents the total number of processors as given by different combinations of processors assigned to the second layer ($n_{masters}$) and to the first layer (n_{slaves}). Ideal speed-up is represented with dashed lines. *(reproduced with permission from [6])*

A description of HYPERstreamHS hydrological model together with its main features and capabilities has been included in the publication by [6].

T1.4.2-2: Set-up of the model in the Adige river basin

HYPERstreamHS capabilities to reproduce observed streamflows and hydropower production in a complex domain were tested with reference to the Adige river basin. After extensive data collection to characterize the 40 large hydropower systems present in the basin, the model was set up for simulating streamflows and hydropower production over the 1989-2013 time window. A multi-site calibration adopting the PSO algorithm was performed to find the optimal set of parameters for the hydrological kernel. The chosen calibration sites are three stream gauging stations located at the end of undisturbed headwater catchments: Cadipietra ($149km^2$), Gadera ($290km^2$) and Vermiglio ($79km^2$). This step was crucial to maximize the capability of the model to simulate inflows to storage reservoirs which are typically located in such sub-catchments. The performances of the model were satisfactorily and led to an overall NSE index of 0.63 in calibration. On the other hand, validation of the model was performed at the downstream gauging stations of Vandoies ($1917km^2$), Mezzolombardo ($1356km^2$), Bronzolo ($7400km^2$) and Trento ($9600km^2$), which streamflows are heavily impacted from hydropower activities. Streamflow and hydropower production were both computed at the computational $1h$ time step: subsequently, streamflows were aggregated at the daily scale and compared with observations, whereas hydropower production was aggregated over the entire river basin, in order to be compared with the monthly time series made available by the public authorities. Streamflows were computed both in fully natural conditions (i.e., neglecting HS) and in actual condition (i.e., considering HS in the model). Results obtained indicate that the inclusion of HS in the modelling framework adds constructive information to the model's prediction. Streamflow time series and flow duration curves at the validation sites depicted in Figure 87 confirm that the inclusion of HS modules improves modelled streamflows by i) improving the timing of the predicted peak flows and ii) greatly improving the accuracy of the predicted low flows, which a natural model isn't able to grasp at impacted sites.

Hydropower production at the catchment scale was computed and compared against observations, which were available for the 2000-2013 time window. The model achieved outstanding results, by pre-

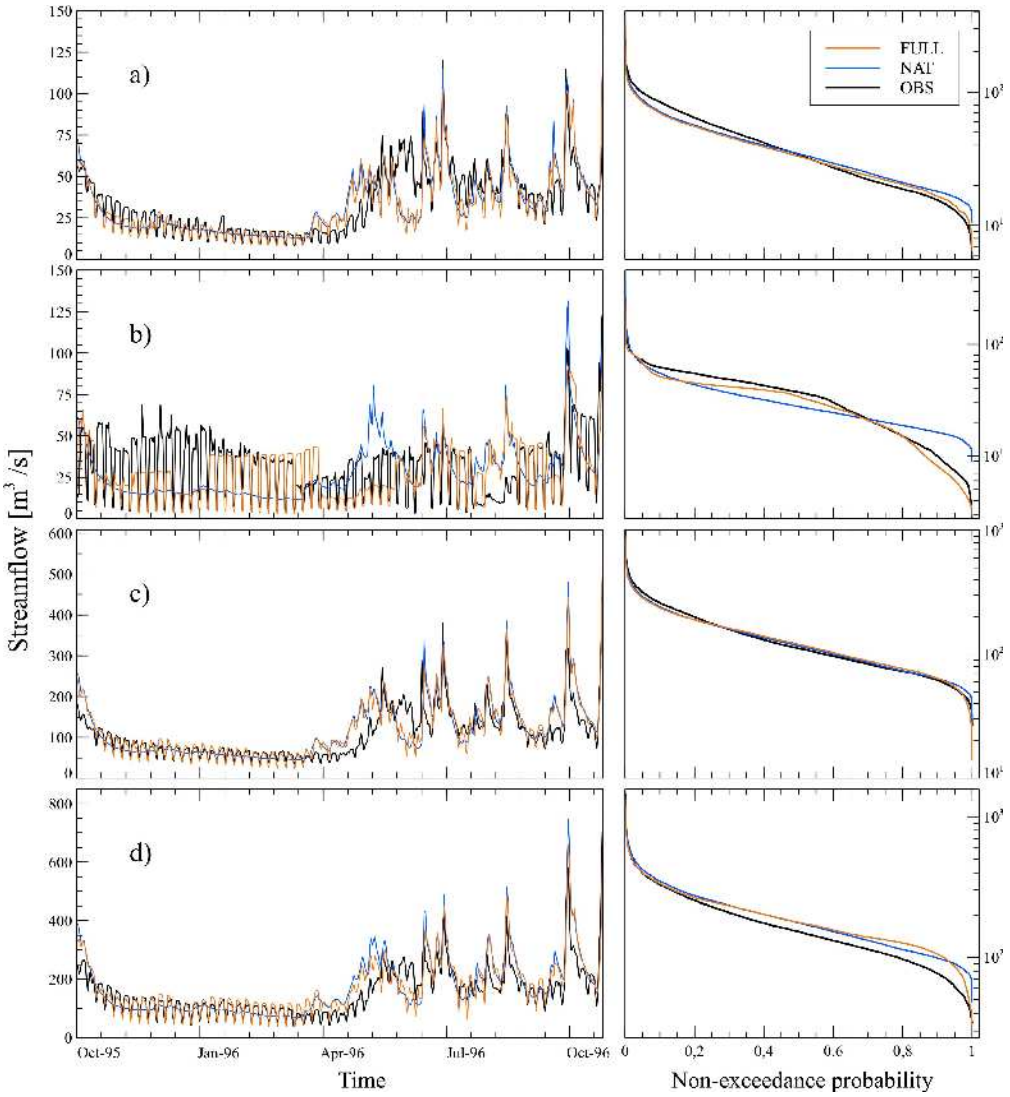


Figure 87: Comparison between observed and simulated streamflow time series (left column) and flow duration curves (right column) at the validation sites: a) Vandoies, b) Mezzolombardo, c) Bronzolo and d) Trento stream gauging stations. To better showcase the effect of modelling hydropower activities, which introduce periodic oscillations, streamflow is shown only for one hydrological year starting from October 1995, while the FDCs represent the entire simulation period 1991–2013. *(reproduced with permission from [53])*

dicting a mean annual production of 6449 GWh/y which differs by only -4% from the observed value of 6717 GWh/y. Notably, also the monthly timing of modelled hydropower production followed that of the observation closely, as depicted in Figure 88. The modelling experiment conducted in the Adige catchment proved that HYPERstreamHS is a reliable tool for hydropower production modelling over complex alpine catchments and that the framework can be extended to other areas. This case study, together with some additional analyses showcasing the main advantages of this framework compared to some traditional approaches in large scale hydropower modelling, were published in [53].

T1.4.2-3: Hydrological benchmarking exercise in the Adige river basin

Subtask 1.4.2-3 evaluated suitable modifications of the Community Land Model v3.5 (CLM3.5) runoff outputs, to improve streamflow reproduction in the Adige river basin. The goal of this activity was to identify

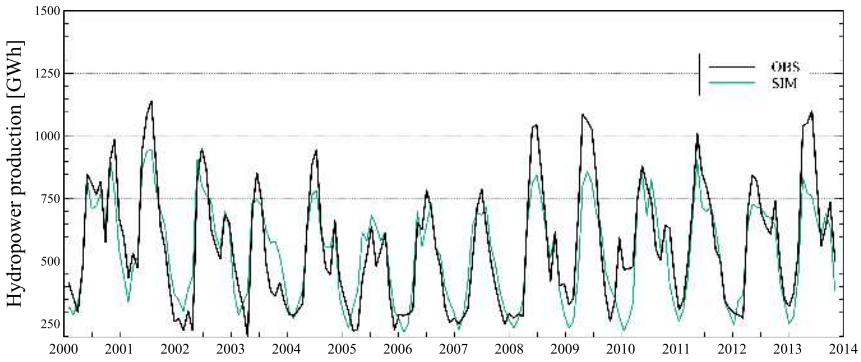


Figure 88: Comparison between modelled (blue) and observed (black) monthly hydropower production in the Adige catchment within the 2000-2013 time window. *(reproduced with permission from [53])*

which physical processes can be responsible of a significant improvement in the capability of CLM3.5 to reproduce observed streamflow. The performance of a multi-site calibrated run of HYPERstreamHS stand-alone model in reproducing observed streamflow has been evaluated to provide a hydrological benchmark, showing a satisfactorily average NSE index of 0.65 in the Trento and Bronzolo streamflow gauge stations. For comparison, the uncalibrated CLM3.5 model run provided a NSE of -0.05. A correlation analysis between surface and sub-surface flow time series produced by HYPERstreamHS and CLM3.5 showed that CLM3.5 surface flows are temporally aligned (maximum of 1 day of delay) with those generated by HYPERstreamHS, whereas the sub-surface counterparts are shifted by 27 days, revealing that CLM3.5 return flow component is too fast if compared to observations. To address such an issue we decided to test the inclusion of different modules using as input the subsurface runoff generated by CLM3.5 and mimicking with different degree of complexity the deep infiltration and return flow component of the streamflow response. Results indicate that the inclusion and calibration of key physical processes related to subsurface flow improved significantly the performance of the modified version of CLM3.5 model. Furthermore, results indicate that increased performances of large hydrological models in reproducing observed streamflow time series can be achieved by replacing the grid based routing scheme of CLM3.5 with a more accurate scale independent routing scheme.

T1.4.2-4: Set-up of the model in the Italian Alpine Region

The goal of Subtask T1.4.2-4 was to extend the modelling framework developed in T1.4.2-1 and tested in T1.4.2-2 to the entire Italian Alpine Region, representing large hydropower systems therein located. In order to do so, the first step has been to create a comprehensive dataset containing all the information concerning large hydropower systems located in the Italian Alpine Region that is required for running HYPERstreamHS simulations. As anticipated in D1.2, not only the collection but also the validation of information concerning large hydropower systems in Italy is critically connected to the responsiveness of the energy companies and public energy agencies, as well as their willingness to share or confirm relevant information concerning hydropower production, as well as to that of Regional Environmental Protection Agencies, who are responsible for determining the Minimum Environmental Flow that each hydropower system should abide by, which also is an essential input information for running HYPERstreamHS simulations. Despite the aforementioned limitations, a very extensive (possibly complete) registry of Italian Alpine hydropower systems was compiled, however not all of the information for some individual systems could be double checked and/or validated by the relevant agencies. To date, the dataset includes information on 370 large hydropower systems (i.e, by Italian regulation, systems with installed power larger than 3 MW) including storage- and run-of-the-river hydropower systems. The dataset includes information on 241 large dams, 207 of which were built in the 20th century with the main purpose of hydropower production. Maps and technical drawings were also retrieved, detailing the schematic structure of the connections of the dam

with the local hydropower plant, which also allowed to extract and validate information regarding reservoir location and shape (e.g., stage-storage relationships specific for each dam, linking water level to stored water volume), as well as their spillway characteristics. An overview of the GIS database created with the collected information is provided in Figure 89.

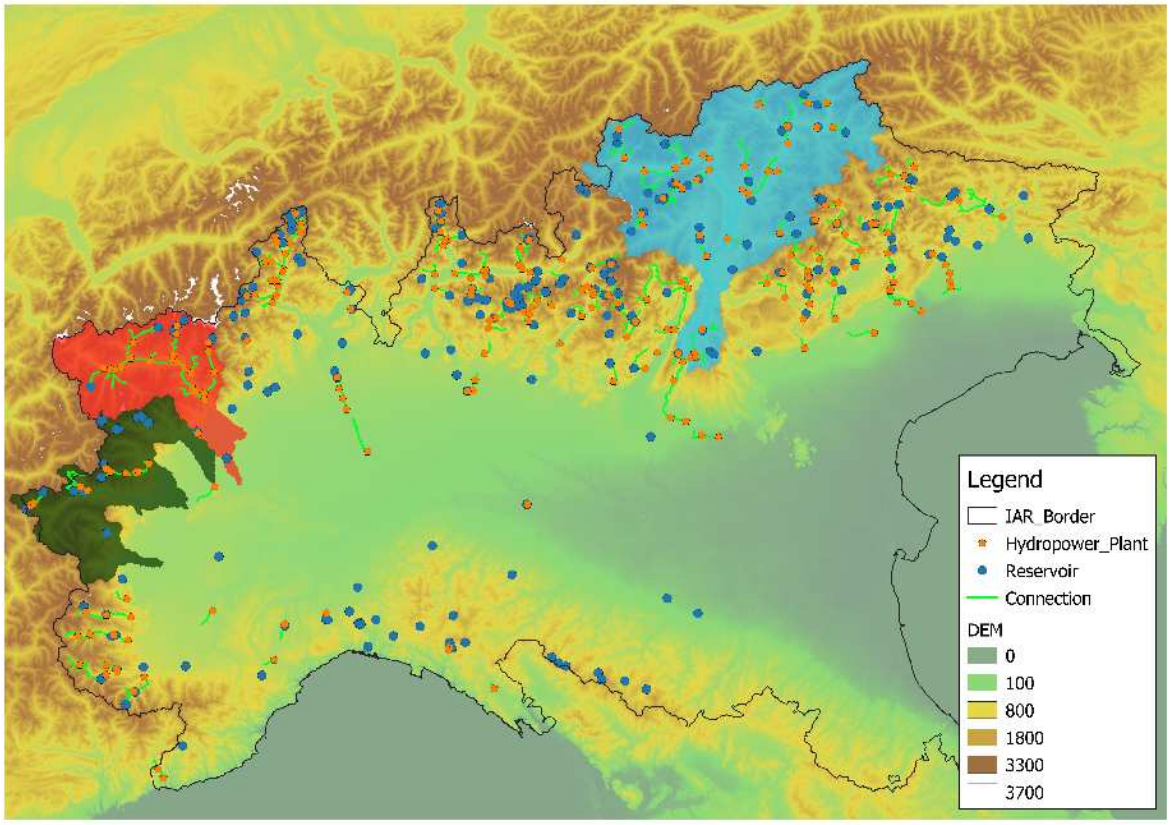


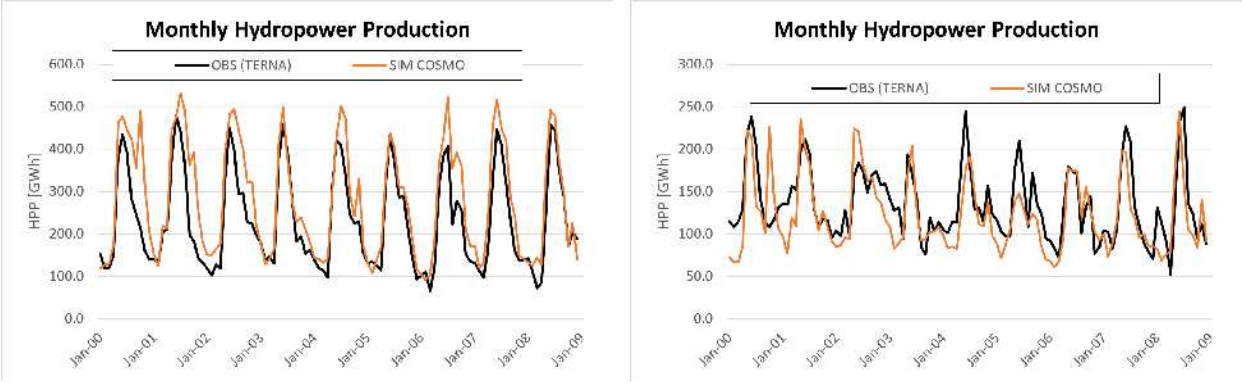
Figure 89: Overview of the Human Systems database developed for the Italian Alpine Region (IAR). Shaded in color are the three case studies described in this Deliverable: Adige (blue), Dora Baltea (red) and Piedmont (green).

The extension of the HYPERstreamHS framework to the Italian Alpine Region and its validation are herein exemplified with reference to two case studies located in the Western Italian Alps: the Dora Baltea river basin (red-shaded area in Figure 89) and a set of four Po tributaries, referred to as the Piedmont case study (green-shaded area in Figure 89). Notice that the validation in the Adige river basin (blue-shaded area in Figure 89) has been already presented in T1.4.2-2.

Hydrological calibration and validation of the model followed the same logic presented in T1.4.2-2, calibrating the model in multiple upstream undisturbed catchments while performing validation in the downstream reaches. Hydrological calibration of the Dora Baltea case study was performed in two upstream gauging stations, achieving an overall NSE of 0.63. Validation performed at one downstream gauging station yielded an NSE index of 0.66, confirming the spatial reliability of the adopted parameters. Hydropower production was also validated in the catchment, showing an average annual value of 3118 GWh/y over the 2000-2008 time window, which exceeded by +8.0% the observed value of 2886 GWh/y. The modelled hydropower production monthly values are shown in Figure 90a.

Hydrological calibration of the Piedmont case study achieved an overall NSE of 0.66. Validation performed at four downstream gauging stations yielded an average NSE index of 0.64, again confirming the spatial reliability of the adopted parameters. Hydropower production was also validated in the catchment, showing an average annual value of 1644 GWh/y over the 2000-2008 time window, which exceeded by

+2.1% the observed value of 1610 GWh/y. The modelled hydropower production monthly values are shown in Figure 90b.



(a) (b)

Figure 90: Comparison of observed (black) and modelled (orange) monthly hydropower production time series in the a) Dora Baltea catchment and b) Piedmont river basins.

Overall, the validation results in several river basins are encouraging, and suggest that the modelling framework is well-suited for simulating hydropower production in the Italian Alpine Region; moreover, it is possible to apply the same framework to wider domains, provided that all the hydropower systems are fully and properly characterized, as highlighted in the results presented in this Deliverable.

6.3 Task 1.4.3: Scientific results

The main task T1.4.3 is subdivided in three subtasks:

T1.4.3-1	**Set up of the hydrodynamic surrogate model over Garonne river in France**. The meta modeling exercise has been carried out over the Garonne river test case between Tonneins and La Réole, focusing on uncertainties related to friction and upstream forcing. The surrogate stands for replacement of the hydrodynamics solver TELEMAC (OpenSource software developed by EDF) using the OpenTURNS UQ dedicated library (Open Source software developed by EDF, AirbusGroup, Phimeca); a Galerkin projection (Polynomial Chaos Expansion) strategy has been favored. The construction of the learning data base relies on the develoment of Python classes for computationally efficient ensemble integration with TELEMAC.
T1.4.3-2	**Sensitivity analysis using the surrogate model**. The surrogate model implemented in the T1.4.3-1 has been used to carry out a global sensitivity analysis to classify sources of uncertainty to explain water level variance over the 2D domain. This work was carried out for stationary flow and led to the estimation of Sobol indices maps.
T1.4.3-3	**Chain PF-CLM hydrological modeled discharge with Telemac hydraulic simulation**. The direct and surrogate models will be forced with observed discharge at upstream location of the estuary or the river. The use of Parflow simulated discharge at neighboring points to these inputs will be investigated following T1.4.1 if this data is available.

Task 1.4.3-1: Set up of hydrodynamic surrogate model over the Garonne river in France

The first part of of Task 1.4.3-1 on the construction and use of a surrogate for Telemac 2D (T2D) hydrodynamic model over the Garonne Marmandaise catchment was detailed in D1.2 report. The catchment is presented in Fig. 91.

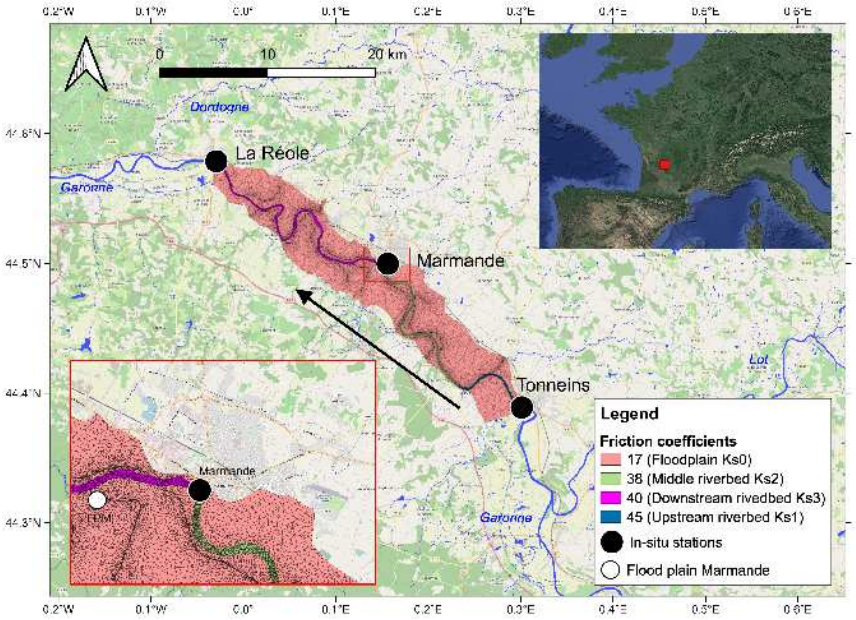


Figure 91: Study area of the Garonne River (southwest France, as shown in the upper-right corner inlet figure) 50-km reach between Tonneins (upstream) and La Reole (downstream). The black arrow indicates the flow direction. The black solid circles represent the in-situ Vigicrue observing stations. The inlet figure at lower-left corner magnifies the area around Marmande. The white solid circle indicates a diagnosis location on the flood plain near Marmande (FPM). The friction coefficient Ks is uniform over 4 zones: upstream, middle, and downstream river bed and flood plain. Background image: Map data ©OpenStreetMap contributors and available from `https://www.openstreetmap.org`. Source: [114]

A surrogate model with a Polynomial Chaos surrogate model was implemented. This work was continued in order to deal with non-linearities and a surrogate base on a mixture of polynomial Chaos experts was implemented in the context of S. EL Garroussi PhD at CERFACS. This work was published in [43]. A surrogate model was developed to accurately approximate a two-dimensional hydrodynamics numerical solver in order to conduct a reduced-cost variance-based global sensitivity analysis of the hydraulic state. The impact of uncertainties in river bottom friction and boundary conditions on the simulated water depth is analyzed for quasi-unsteady flows. An autoencoder technique adapted to non-linear variable dimension reduction is used to reduce the multi-dimensional model output so that the formulation of the surrogate remains computationally parsimonious. In addition, following the divide-and-conquer principle, a mixture of local polynomial chaos expansions is proposed to deal with non-linearity in the hydraulic state with respect to uncertain inputs. Machine learning techniques are used to automatically partition the input space into clusters that are not affected by non-linearities and support accurate surrogates. This combined strategy is applied to a reach of the Garonne River where river and floodplains dynamics are simulated by the numerical solver Telemac-2D. The merits of this strategy are highlighted when the flood front reaches regions where the topography features a strong gradient and where, consequently, strong non-linearities occur between the water depth and friction as well as hydrologic input forcing. By applying this strategy, the Q2 metric improves by 90% compared to a classical polynomial chaos expansion surrogate as illustrated in Fig. 92, resulting in a much more reliable sensitivity analysis (see Task 1.4.3-2). This is particularly important in floodplain areas where human and economic activities are at stake.

The surrogate model was established for stationnary flow, assuming the uncertainties relate to friction

and upstream forcing. The output quantity of interest is the 2D water level map, considered at a given time step during the flood' s rising part. At this simulation time, the classical single Polynomial Chaos leads to poor results ([56]) and the Mixture of Experts solution over comes the limits due to the presence of non-linearities between the input and the output spaces. For this work, CERFACS's cluster, Nemo, has been used to run T2D simulations. The Nemo cluster includes 6912 cores distributed in 288 compute nodes. The ECU power peak is 277 Tflop/s. The computational cost of T2D solver is reduced thanks to the parallel computing (single simulation lasts 6 min using 24 processors instead of 20 min using one processor). GSA based on a large set of T2D simulations is too costly. Hence the need for surrogate model formulation.The mixture of experts surrogate model proposed in this study is based on algorithms from different Python libraries. The first step uses AE from Keras Tensorflow with a graphics processing unit (GPU) support Python package to reduce the dimension of the output space. The second step of this algorithm involves clustering and classifying data using a GMM and SVM algorithms from the Scikit-Learn library In the final step, the algorithm constructs a local regression model within the cluster; for this purpose, PCE of the Open-TURNS library is used.The meta-model learning stage is moderately costly: the tuning of the AE parameters takes about 3 h and the construction of the PCE takes about 15 min. The computational cost of the prediction stage is then drastically reduced, e.g., predicting 500 simulations takes 470 s.

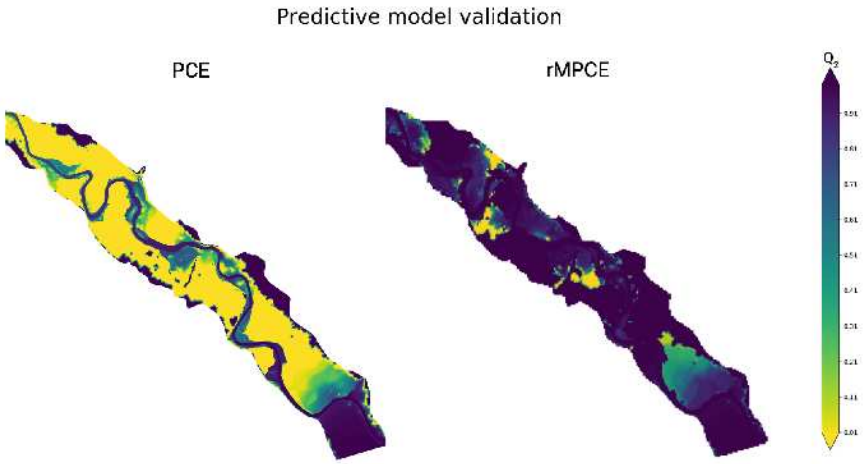


Figure 92: Spatialized predictive coefficient computed between the validation database and the surrogate prediction at T: classical PCE (left) and rMPCE (right). [43].

In D1.2, additional information were provided on the development of peripheral functionalities for data driven modeling framework for Telemac. A short summary is provided here for clarity purpose as well as addtional details on more recent advances. Python-API classes were developed for data-driven integration of the model by A. Piacentini. The main class allow to handle ensemble runs with perturbed inputs, working from a single instance of TELEMAC2D and using MPI through mpi4py to launch several members on available resources by batches. A second class is specific to the applicative case and defines which inputs to the model could be modified and how the perturbations are applied to these inputs that may vary in time and space. Additional classes are dedicated to the management for ensemble output fields or output values. Further developments on the Python classes dedicated to Data Driven Modeling were achieved by A. Piacentini for the purpose of ensemble runs, coupling and mostly data assimilation. It should be noted that a significant effort was made to develop observation operators that allow to compare T2D simulation outputs with remote sensing observation derived from Sentinel SAR and Optical images, post-processes by Machine Learning algorithms to derive flood extents. This specific work was carried out, complementary to EoCoE efforts, with additional funding from CNES through the Space Climate Observatory (SCO) and TOSCA programs. These developments mostly stand in Python classes that are added to the existing Python classes developed in the Telemac API and available in the SVN server for the Open Source Telemac software (`https://gitlab.pam-retd.fr/otm/telemac-mascaret`), in a dedi-

cated branch named fishstick. These developments also stand in Python classes dedicated to application and studies available on gitlab:`https://gitlab.com/cerfacs/tools4telemac`. These functionalities were used for a publication on the improvement of flood extent representation with remote sensing data and data assimilation ([114]).

Task 1.4.3-2: Sensitivity analysis using the surrogate model

The details of Task 1.4.3-2 on the implementation of a sensitivity analysis with a single Polynomial Chaos surrogate were given in D1.2 report. The results for the sensitivity analysis from the use of the Mixture of Expert surrogate model were also published in [43], and are reported here.

A variance-based global sensitivity analysisis is based on Saltelli's method for the estimation of Sobol indices using the rMPCE surrogate model. The main goal of GSA is to rank the uncertain parameters according to their influence on the variance of the QoI, here, the water depth 2D field. Figure 93 displays the first-order (left panels) and total order (right panels) Sobol indices for the four Strickler friction coefficients and discharge (from top to bottom) at time T. Analysis of the first order Sobol indices reveals the large influence of the discharge as this uncertain variable explains about 80% of the water depth variance on the overall domain. The Strickler friction coefficient associated to the floodplain area influences by 9% the water depth variance upstream and in some dyked areas. The influence of the Strickler coefficients associated with the river bed remains weak or slightly significant in a few places; for example, Ks4 influences the water depth variance by 82% locally in a dyked zone downstream of the river. The analysis of the total Sobol indices indicates that while the friction coefficients have a low first order Sobol index, they are not negligible as they have a significant influence through their interactions with other variables. Yet, the discharge remains by far the most influencing variable when it interacts with the other variables as shown in the right-bottom plot. It should be noted that the GSA results depend on the hypothesis on the input random variables distributions. For instance, the significant influence of the floodplain Strickler friction coefficient com- pared to that of the river bed coefficients may be due to the large uncertainty translated by the large range of Ks1's uniform distribution.

Task 1.4.3-3: Chain PF-CLM hydrological modeled discharge with Telemac hydraulic simulation

Delivarable D1.3, task 1.4.3-3, focuses on the chain of Parflow and Telemac simulations. Parflow solves the water cycle dynamics at large scale with all physical processes that are resolved but simplified while only the hydrodynamics is solved in Telemac. Indeed, ParFlow is a numerical model that simulates the hydrologic cycle from the bedrock to the top of the plant canopy. It solves saturated and variably saturated flow in three dimensions using an orthogonal or terrain-following, semi-structured mesh that enables fine vertical resolution near the land surface and deep confined and unconfined aquifers. ParFlow models dynamic surface and subsurface flow solving the simplified shallow water equations implicitly coupled to Richards' equation. TELEMAC-MASCARET is an integrated suite of solvers for use in the field of free-surface flow. Having been used in the context of many studies throughout the world, it has become one of the major standards in its field. It solves the Shallow water equations either in 2D or 3D over an unstructured mesh with an explicit, implicit or semi-implicit scheme. Shallow-water models are widely used in the field of rivers and maritime hydraulics. These equations are derived from the Navier-Stokes equations for shallow-flows assuming hydrostatic pressure and low variation in bathymetry. Hydrodynamic models use the amount of water entering the river system to compute the water level and velocity in the river network, and when the storage capacity of the river is exceeded, in the flood plain. These models are used to predict river water surface elevation (WSE) and velocity from which flood risk can be assessed for lead times that range from a couple of hours to several days. However, these numerical codes are imperfect as uncertainties inherently existing in the models and in the inputs (model parameters, boundary conditions (BCs), geometry) translate into uncertainties in the outputs. The performance of the hydrodynamic model is limited by the amount and quality of available data. Model parameters such as friction coefficients are usually calibrated for significant flood events with respect to the observational data. As a result, the model

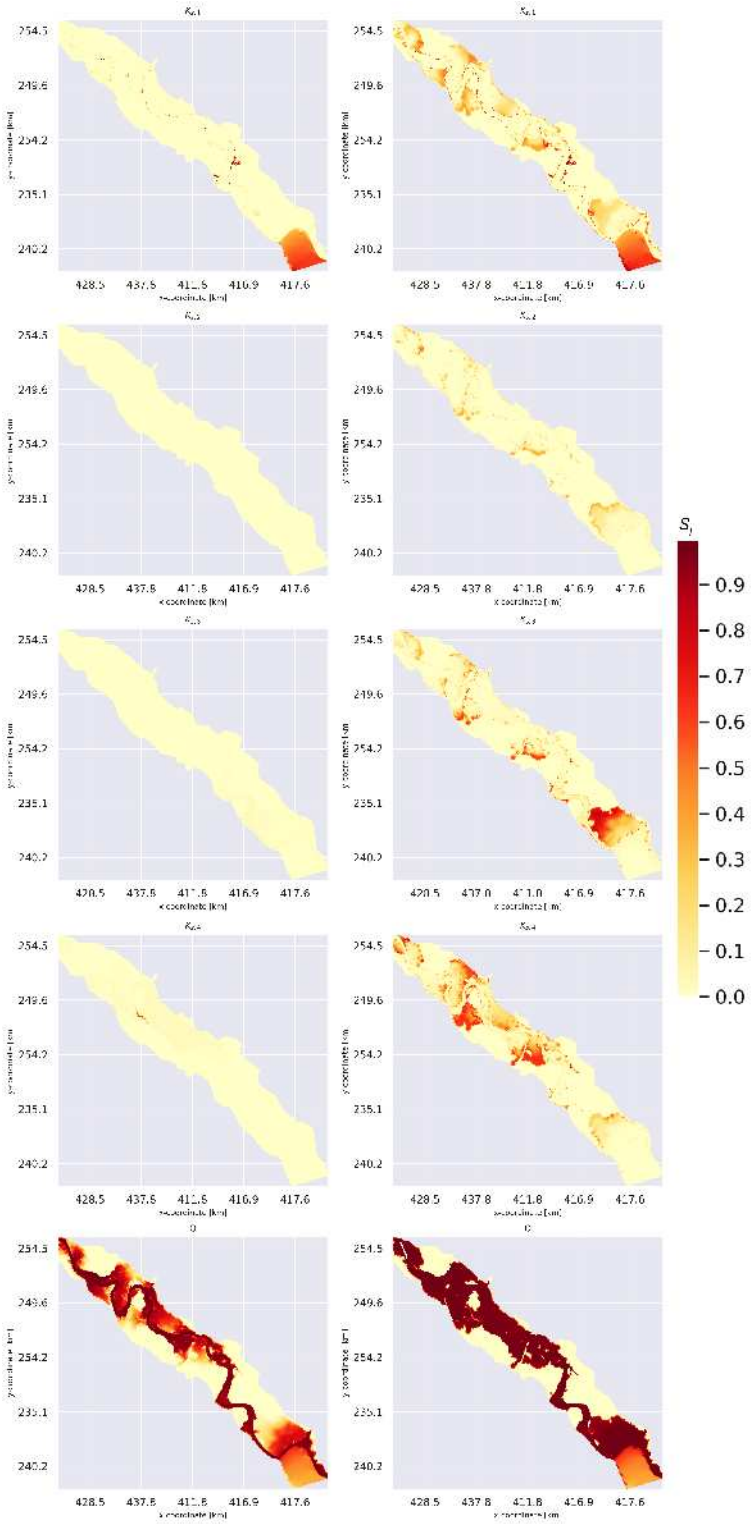


Figure 93: Sobol indices of the hydraulic input variables estimated using Saltelli's method based on rMPCE for the simulated water depth at time T = 95,000 s. First-order indices are plotted on the left panels and total order on the right panels for Ks;1 (floodplain), Ks;2 (upstream river bed), Ks;3 (middle river bed), Ks;4 (downstream river bed), and Q (upstream forcing) from top to bottom. [43].

can only be calibrated and validated as finely as the available data allow it, stressing out the need for a time and space densified observing network. In this framework, Parflow outputs are used as boundary conditions for the Telemac local model. It should be noted that the upgrade of the surrogate model previously described from stationnary flow to non-stationnary flow was not acheived in the context of the PhD, as planned. Indeed, the PhD was terminated by the student for personal reasons in early 2021. Parflow was thus chained to the direct model Telemac only. While the use of a surroagate model would have lowered the computational cost of the chained simulation, the IO interfaces for the direct and the surrogate remain similar. As a consequence, the hydrology-hydraulic chained model presented here is representative of the dynamics of the expected flow (even with a surrogate), and the solution for technical aspects treated here stil valid when the surrogate model is used.

The local model of the Garonne Marmandaise was considered here (see Fig. 91). It was forced at its upstream boundary condition at Tonneins by Parflow timeseries issued by Julich (Collaboration with Bibi Naz). Parflow timeseries where made available over the 1997-2006 period. Year 2003 was selected for chained simulation as significant flow events are observed over this period. First, Parflow discharge timeseries were compared to observed discharge time series at Tonneins Observing station (Vigicrue network https://www.vigicrues.gouv.fr/) as well as at La Réole (which is the downstream location of the local Telemac model) over 2003 and 2004. As shown in Fig. 94 and Fig. 95, the hourly timeseries from Parflow (blue curve) tend to underestimate the observed flood peaks (orange curve), especially at Tonneins, with an underestimation of about 50%. At La Réole, only the time shift is visible as the Parflow discharge is plotted along with the Observed water level (this choice was made in order to get rid of the possible error in the local rating curve at La Réole). Also, there seems to be a 5-day lag in time Parflow simulations that represents the flooding with a significant. Yet, the over all signal of the flow is in good coherence with the observed dynamics, thus allowing for chaining with Telemac. The time delay in Parflow could likely be corrected with additional calibration efforts over this specific area of the Manning coefficients, this requires further investigation. It should also be noted that Parflow tends to simulate really dry periods in summer. Yet, this is not a limiting aspects in the context of flood simulation.

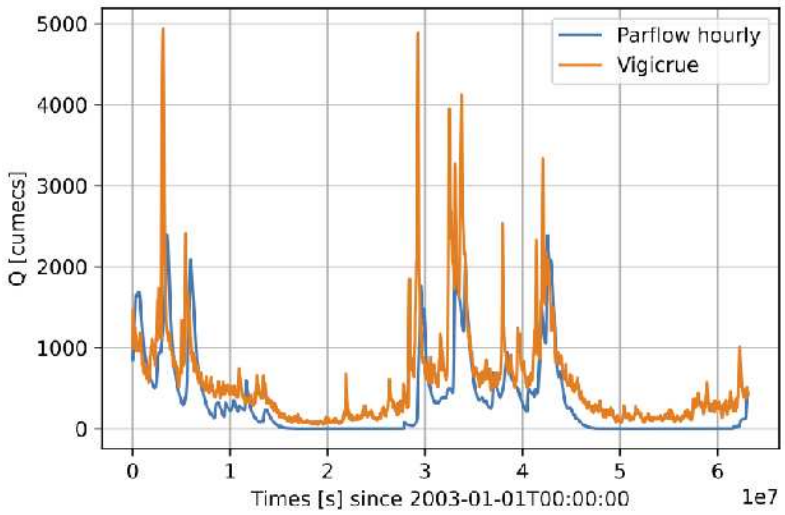


Figure 94: Comparison of the Parflow hourly simulated discharge (blue line) with the observed discharge at Tonneins (orange curve) for 2003-2004.

Figure 96 displays the results of the simulation from Telemac over year 2003, forced either by the Vigicrue observation and Parflow hourly outputs. The water depth is shown at Marmande where flood frequently occurs over an urban area as well as cultivated flood plain. For the two top panels, the water depth observed at Marmande (Vigicrue network) is plotted in orange. On the top panel, the water depth

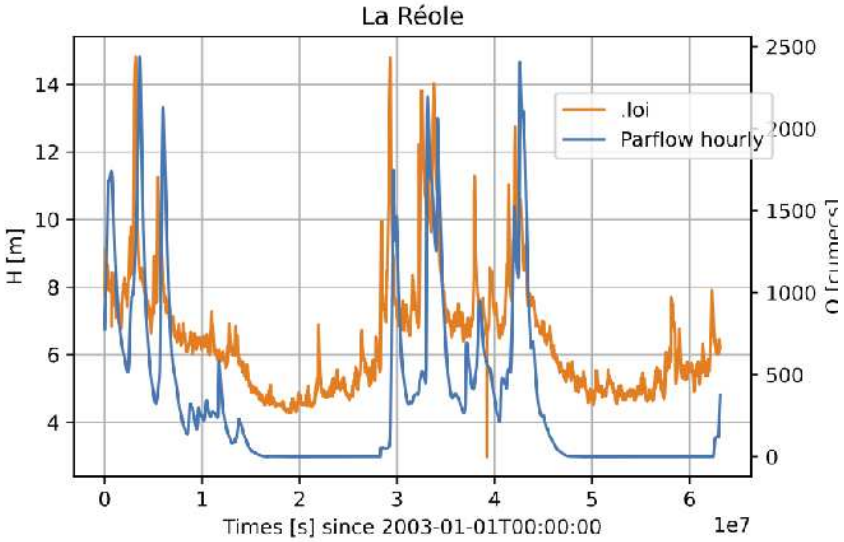


Figure 95: Comparison of the Parflow hourly simulated discharge (blue line, right y-axis) with the observed water level at La Réole (orange curve, left y-axis) for 2003-2004.

simulated by Telelamac, when forced by the upstream observed timeseries at Tonneins is represented in blue. On the middle panel, the water depth simulated by Telemac (here hourly T2D outputs), when forced by the upstream simulated discharge by Parflow with hourly outputs, at Tonneins is represented in green. Finally, the bottom panel recalls the Vigicrue and Parflow timeseries at tonneins for 2003. The 5-day delay in the forcing naturally translates into a delay for the simualtion of the flood over the entire hydraulic network, including at Marmande. Also, the underestimation of the input forcing at Tonneins inParflow leads to an underestimation of the flood peak at Marmande. The RMSE between the observed water level and the simulated water level with T2D forced by Vigicrue timeseries is about 0.274m. The RMSE between the observed water level and the simulated water level with T2D forced by Parflow (hourly) timeseries is about 1.584m. It should be noted that negative water level observed at Marmande over summer is due to the referencing data used for low flow in summer at the observing station.

A temporal zoom was carried out over 2 short periods in 2003: 20 day of flooding in February and 20 days of flooding in December 2003. The observed and simulated water levels at Marmande are shown in Fig. 97 for February event (left column) and December event (right column). The water level observed at Marmande is shown in orange. The water level simulated by Telemac when forced by Vigicrue timeseries is plotted in blue, when forced by Parflow hourly outputs is plotted in green and when forced by Parflow daily outputs is plotted in red. The forcing timeseries at Tonneins are plotted with the same color code in the bottom panel. The peak delay and underestimation is clearly visible with respectively an RMSE of 0.344m, 2.653m and 2.565m for February event and 0.284m, 3.259m and 3.122m for December event (from top to bottom panels). The associated flood areas are plotted in Fig. 98 and Fig. 99 respectively for February and December events. The flood extent is issued from the T2D simulated 2D water level map with a threshold of 5cm for dry pixels. For each T2D simulated timeseries, the flood extent is issued at the time of the flood peak (thus eliminating the 5-d delay when simulations are forced with Parflow). For both events, it should be noted that when Telemac is forced by Parlow, the flood extent is clearly smaller, as expected since the volume of water inputed in the network is too small.

As a conclusion, it shoud be noted that parflow timseries offer a great synchronicity with the observation, but a 5-d delay should be taken care of, especially if forecasts are to be issued. Forcing a local refined model that solves the hydrodynamics of the flood plain is feasible and allows to represent the dynamics in the flood plain. The errors in the forcing from Parflow naturally translate into errors in the local simulated water level and flood extent with Telemac. For 2003, the validation of the simulated flood extent is not

possible with respect to observed flood extent from space as missions such as Sentinel 1 was launched many years later. This comparison could be carried over more recent periods after 2014. As a perspective, further investigations should be pursued for better calibration of Parflow over this area in order to improve hindcast simulation. Finally, when running T2D in parallel over 12 processors, 1 day of physical simulation takes about 3 minutes of CPU time. This performance would be greatly reduced if a surrogate model for T2D was available as we previously described for stationnary flow.

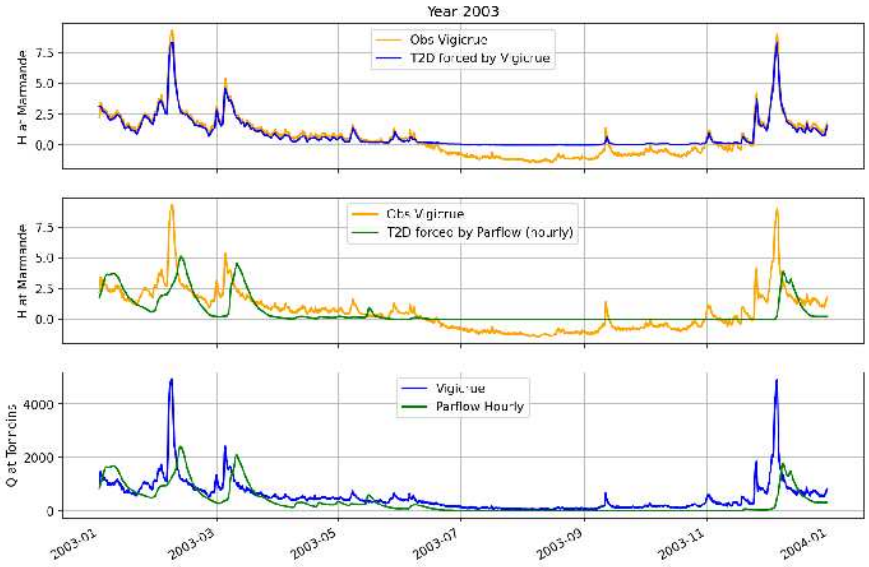


Figure 96: Two top panels : Water level at Marmande over 2003 with observation in orange, T2D forced by Vigicrue in orange and T2D forced by Parflow hourly outputs in green. Bottom panel: observed and Parflow simulation for discharge at Tonneins.

6.4 Task 1.4.4: Scientific results

The main task T1.4.4 is subdivided into seven subtasks:

T1.4.4-1	**Becoming acquainted with OED theory and the state of the art**. Literature research and study on the theory of OED which comprises the mathematical background as well as its computational implementations.
T1.4.4-2	**Software access**.The software framework EFCOSS (Environment for Combining Optimization and Simulation Software) developed by partners at FSU Jena was used for interfacing the geothermal simulation code SHEMAT-Suite with mathematical optimization software. The subtask includes getting access to the EFCOSS code, as well as training and a practical introduction to the usage of the OED workflow.
T1.4.4-3	**Learning and understanding the EFCOSS software environment and its coupling with SHEMAT-Suite**. This subtask aims at understanding the OED workflow, which involves the software framework EFCOSS and the geothermal reservoir simulator SHEMAT-Suite, by becoming acquainted with the source code.

T1.4.4-4	**Reproducing OED simulations and developing a post-processing workflow**. We reproduce an existing OED study by the FSU Jena group: a synthetic two-dimensional model [13]. In order to evaluate OED simulation results fast, efficiently, and comparably, a post-processing workflow will be developed comprising several Python-based scripts for reading, analyzing and visualizing simulation outputs.
T1.4.4-5	**Defining a research and development concept for OED for geothermal modelling**. Identifying relevant and feasible extensions of the existing OED functionalities in the context of geothermal reservoir modeling. Additionally, performance bottlenecks of the work flow will be identified and feasible optimization strategies will be defined.
T1.4.4-6	**Setting-up a testmodel suite and extending OED functionality**. A suite of synthetic 2D and 3D testmodels will be used for studying aspects defined in T1.4.4-5 and for optimizing the OED workflow.
T1.4.4-7	**Optimizing the OED workflow**. According to the outcomes of T1.4.4-5, the performance of the existing OED workflow will be increased by using reasonable programming models such as MPI parallelization, e.g. to parallelize the repetitive computation of the Fisher Matrix. This will enable us to apply the OED approach to larger model domains, which has not been computationally feasible so far.

T1.4.4-1 to T1.4.4-4 In deliverable D1.2, we reported on the first four subtasks in detail. D1.2 gives an overview on the OED theory and implementation of the approach in the Environment for Combining Optimization and Simulation Software (EFCOSS). The software framework EFCOSS [140] interfaces the geothermal simulation code SHEMAT-Suite [77] with optimization software, such as the SciPy package. During an OED simulation, EFCOSS executes SHEMAT-Suite for computing the model's Jacobi matrix using Automatic Differentiation (AD) for an exact calculation of the first order derivatives [125]. Subsequently, the Fisher information matrix is computed from which different optimal design criteria can be computed, e.g. the D-optimality. EFCOSS routines compute the D-optimality for each grid cell, which was specified in the steering script and write them to an output file. The reproduction of an existing 2D synthetic OED study (subtask T1.4.4-4) revealed several bugs in the AD code generation within SHEMAT-Suite. After bug fixing and consecutive testing in collaboration with developers in Jena, the OED simulation could be reproduced successfully. The simulation results were used for setting up Jupyter Notebooks for post-processing and visualization. Subsequently, they were extended for post-processing of 3D simulation results.

T1.4.4-5 During this reporting period, the testmodel suite has been extended by a 3D model (see Sec. 6.5). The test models were used for determining different experimental conditions such that when estimating different parameters of model units from temperature measurements in the borehole the uncertainty is minimized. By simulating different synthetic model scenarios, the sensitivity of the OED approach to prior data location and quality was investigated. In addition, the feasibility of the OED approach in context of thermal conductivity estimation has been studied.
For investigating the inversion and OED results and for obtaining synthetic observation data, we simulate synthetic true reference forward scenarios. The detailed workflow for the OED investigations is as follows:

1. Synthetic true reference scenario: forward simulation with previously defined "true" reference rock parameters

2. Generate synthetic observation data: disturbing the forward simulation results (e.g. temperature) with Gaussian white noise for obtaining synthetic data logs

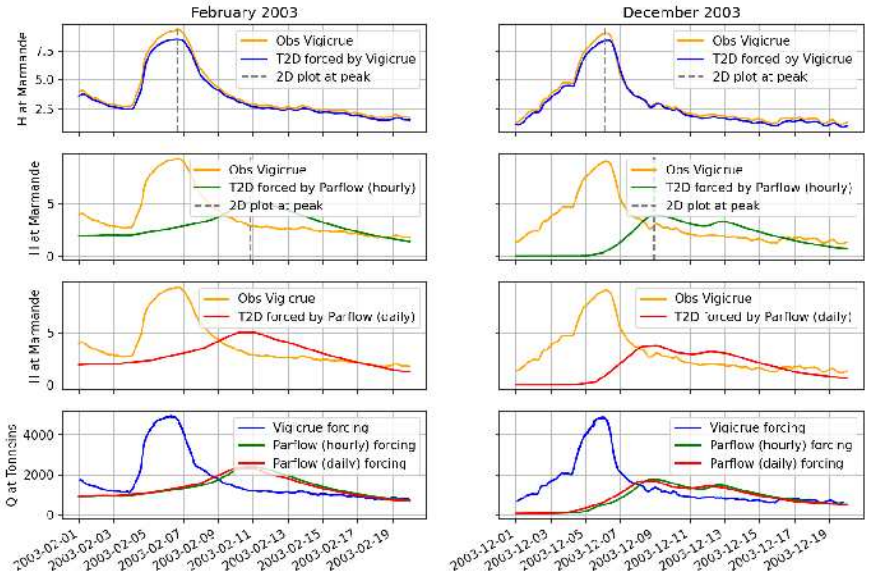


Figure 97: Three top panels : Water level at Marmande over 2003 with observation in orange, T2D forced by Vigicrue in orange, T2D forced by Parflow hourly outputs in green and by Parflow daily outputs in red. Bottom panel: observed and Parflow simulation for discharge at Tonneins. Left column : February 2003 event, Right column: December 2003 event.

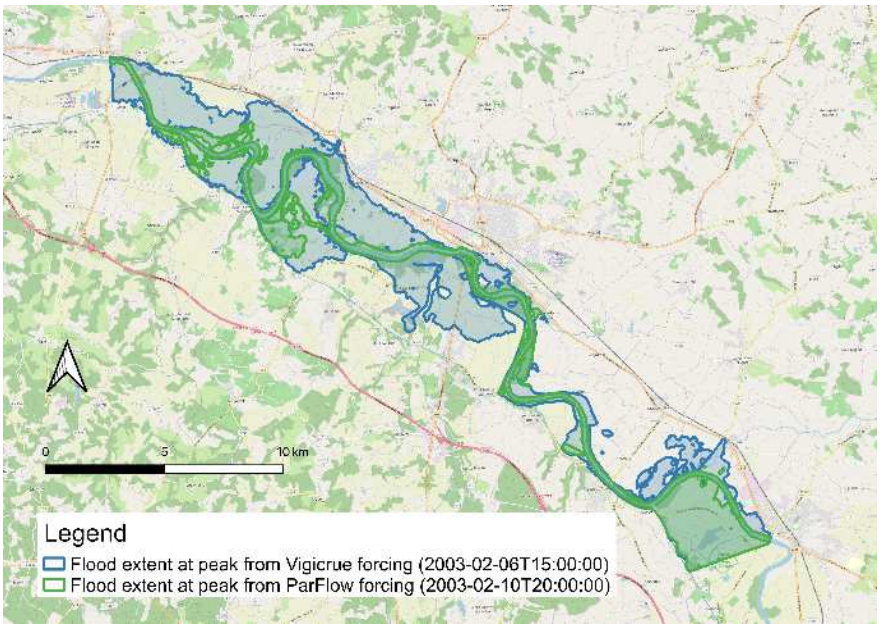


Figure 98: Flood extent for February flood event, computed from T2D water level simulated maps, with a threshold of 5cm for dry pixels. The time of the food peak is used for each timeseries.

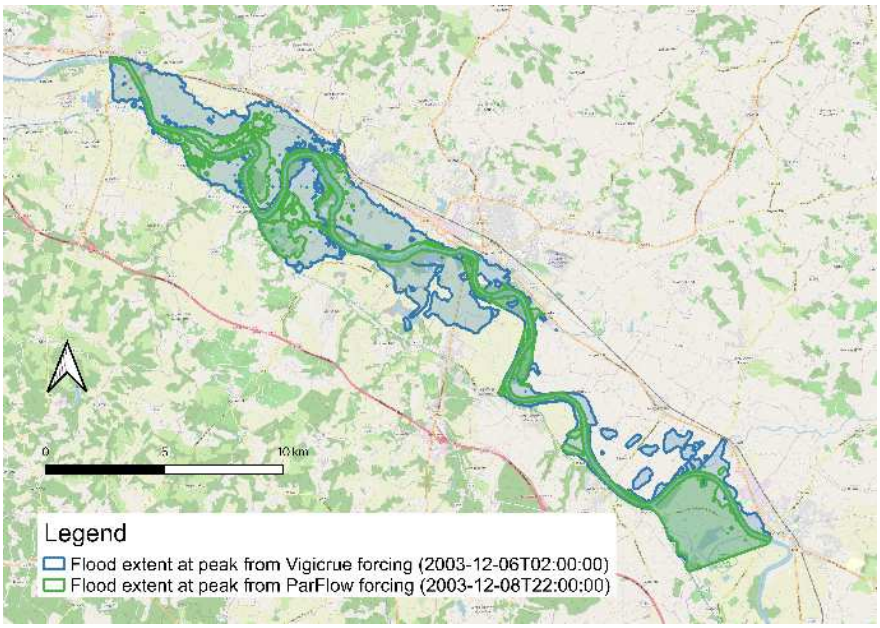


Figure 99: Flood extent for December flood event, computed from T2D water level simulated maps, with a threshold of 5cm for dry pixels. The time of the food peak is used for each timeseries.

3. Apriori forward model: one or more rock parameters differ from the true reference model (i.e., prior parameters)

4. Inverse model: invert for one or more rock parameters based on observation data from step 2. to obtain the posterior model.

5. OED run: determine optimal logging locations for solving the inverse problem from step 3. with least uncertainty.

This way, various OED problems were simulated for the synthetic 2D model described in deliverable D1.2. For instance, the OED problem to find the optimal location for an additional temperature log for inverting the fault permeabilities (see Fig. 100a, units 10 and 11) with least uncertainty was solved with different prior fault permeability assumptions. The prior fault permeabilities range from $1 \times 10^{-16}\ m^2$ to $1 \times 10^{-12}\ m^2$. The true reference fault permeability is $5 \times 10^{-14}\ m^2$. Figure 100d) shows the resulting normalized D-optimalities for the different prior assumptions. The black curve represents the true reference scenario. Scenarios with prior fault permeability lower than the true reference value reveal the minimum D-optimality, i.e. the optimal borehole location, at around x=9000 m. Scenarios with prior fault permeability larger than the true reference value reveal the minimum D-optimality at around x=3000 m. The reason for this shift of the optimal location is that high permeabilities in the faults enable convective transport of in this case cold water downwards along the red fault, which decreases the temperature gradient in this area, compared to scenarios with lower fault permability (see Fig. 100 b and c). To conclude, we see that in case of permeability estimation from temperature data the prior permeability influences the OED result. The information content of temperature data with regard to permeability is usually highest, where the temperature gradient is steep. Convective heat transport due to permeable units will affect the temperature gradient. If the prior permeability is assumed much lower or higher than the "true" permeability value, the OED result will not represent the actual optimal drilling location. This means, that the OED approach is not robust to prior permeability assumptions.

Furthermore, OED problems with regard to thermal conductivity inversion have been formulated. Figure 101 shows the resulting D-optimality for different prior assumptions. Here, the thermal conductivity of

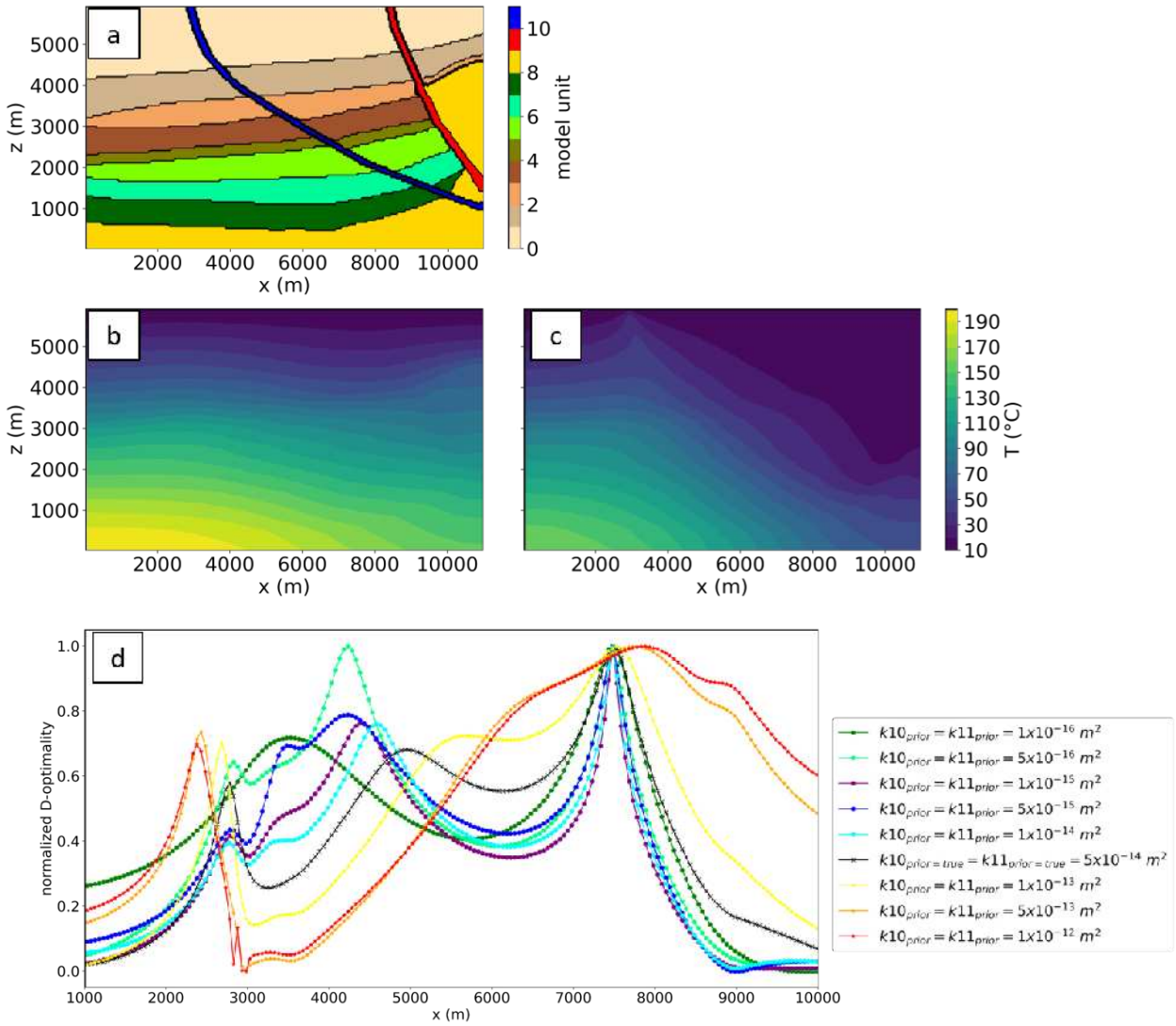


Figure 100: a) Structure of the 2d synthetic test model. The OED problem is to find the optimal borehole location for inverting the permeabilities of the two faults, units 10 (red) and 11 (blue). b) Prior temperature solution for assuming $k10_{prior} = k11_{prior} = 5 \times 10^{-16}\ m^2$. c) Prior temperature solution for assuming $k10_{prior} = k11_{prior} = 5 \times 10^{-13}\ m^2$. d) Normalized D-optimality for different prior fault permeabilities (k10, k11). The black curve represents the true reference case.

the blue fault and the dark green layer are estimated. The optimal borehole location is at 3175 m for all scenarios, regardeless of the prior assumptions. In this case, where permeability stays constant, temperature observations are directly linked to thermal conductivity. The optimal location is where temperature data can be measured from the blue fault unit directly. The OED result does not depent on prior thermal conductivity assumptions, but on the model structure.

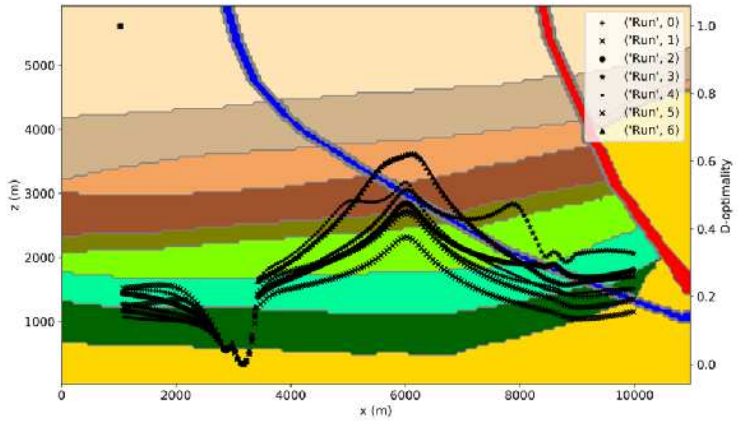


Figure 101: Normalized D-optimalities for different prior thermal conductivity assumptions. The OED problem is to find the optimal meaurement location in order to invert the thermal conductivity of units 7 (dark green) and 11 (blue fault) with least uncertainty. See Fig. 100a for legend

T1.4.4-6 Regarding the optimization of the OED workflow (subtask T1.4.4-7), adding the option to store the Jacobi matrix in a Numpy array after the first SHEMAT iteration was a simple yet effective improvement. Since for the studied OED problem, the Jacobi matrix does not change during the simulation, EFCOSS can read from the stored array when computing the D-optimality for the next grid cell. In the previous implementation [139], SHEMAT-Suite was called iteratively and the complete inverse simulation was executed repeatedly for each new grid cell. Storing the Jacobi matrix greatly reduces the computing time for the OED simulation; for one of the 2D test models, for instance, it is reduced from more than 11 hours to around 6 minutes. This makes the initially planned MPI parallelization of this task obsolete.
Moreover, the OED workflow benefits from the integration of the AGMG solver into SHEMAT-Suite (WP3-Task 3.2: Linear Algebra solvers for Water). This work has been done in collaboration with WP3 and is reported in detail in deliverable D3.4 (Final versions of LA solvers with documentation and performance evaluation). The AGMG solver speeds up the solution of the conductive heat flow problem compared to the previously used BicGStab solver, which in turn also speeds up the OED simulation. The AGMG solver performs considerably better than the BiCGStab solver in terms of required solver iterations and computing time. The speedup due to AGMG is between 2.7 and 24.6 for tested simulations. It seems that especially models with a fine spatial discretization benefit from the AGMG solver. Comparison between two 3D models with similar number of unknowns but different spatial discretization reveals that AGMG is robust to the discretization, whereas BiCG needs more time the finer the discretization. In addition, a largescale test revealed that AGMG shows a nearly constant time per unknown, which is near-optimal behavior.

6.5 Task 1.4.5: Scientific results

The main task T1.4.5 is subdivided in five subtasks:

T1.4.5-1	**Evaluation of the study area**. Evaluation of optional study areas (worldwide) which are digitally recorded in their geological structures already for defining a suitable study area and show case for the OED workflow.
T1.4.5-2	**Data collection**. Collecting available data for the study area defined in 1.4.5-1.
T1.4.5-3	**Setting-up a conceptual and numerical model**. Developing a conceptual model for the OED study of the geothermal reservoir involves defining the area to be simulated, deciding on the model discretization, defining boundary conditions, and initial hydraulic and thermal rock parameters. Subsequently, the respective subpart of the geological model needs to be gridded and the numerical model will be set-up accordingly for fluid and heat flow simulations with SHEMAT-Suite.
T1.4.5-4	**Model calibration**. The initial numerical reservoir model will be constrained to available data, e.g. temperature logs or hydraulic head observations. Thus, boundary conditions and rock parameters can be calibrated. We might use deterministic or stochastic inverse approaches depending on the data available. In case some parameters cannot be calibrated satisfactorily, we will use the resulting quasi-synthetic version of the reservoir model for applying the OED work flow.
T1.4.5-5	**OED study**. Application of the OED workflow to the geothermal reservoir model for predicting optimal borehole locations for new temperature measurements in order to deterministically invert for physical rock properties (e.g., permeability, thermal conductivity) with least uncertainty.

T1.4.5-1 This subtasks aims at applying the OED workflow to a real world geothermal reservoir system model. In the previous deliverable D1.2, we discussed why the originally proposed show case of a geothermal reservoir in Tuscany, Italy, turned out to be not a feasible candidate for an OED study with SHEMAT-Suite within the framework of the project. Instead, in view of the remaining project duration and person months, we decided for the subsurface model around the city of Geilenkirchen, which was setup and used during the first EoCoE phase already.

T1.4.5-2 to T1.4.5-4 Therefore, subtask T1.4.5-2 - collecting data for setting up the model - was obsolete. In addition, less time was required for subtasks T1.4.5-3 and T1.4.5-4, since we could use the available model and calibration results. However, a shortcoming of this choice is that no real deep exploration data, e.g. temperature logs, are available in the study area. Also, in order to simplify the model for lack of time reasons, we studied a purely conductive model. Groundwater flow was neglected, although this does not correspond to real world conditions in the study area. Therefore, we study a quasi-synthetic scenario based on a realistic 3D structural model and realistic rock properties.

The 3D model represents an area of 12 x 12.5 km in the vicinity of Geilenkirchen north of Aachen, Germany (Fig. 102 a). It is located at the German-Dutch border and partly extends into the Netherlands. It contains the settlement “Fliegerhorstsiedlung Teveren” that was of interest for the potential use of shallow geothermal energy and therefore investigated in [115]. The model thickness is around 1200 m depending on the topography. Cells above the topographic surface are represented by a dummy air unit. The discretization is 100 m x 100 m x 5 m. The model area is situated in the Lower Rhine Embayment (LRE) an active subsidence site, characterized by NW-SE striking normal faults that yielded the development of

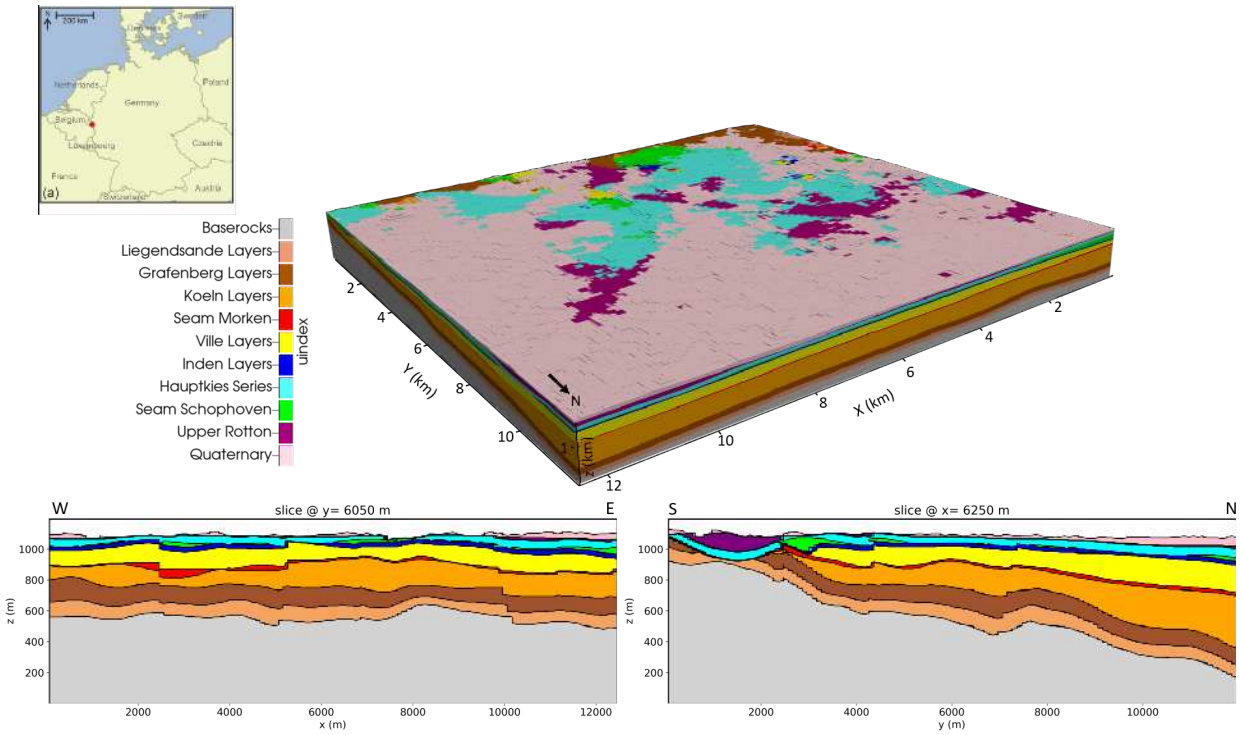


Figure 102: Structural model of the 3D application case presented as 3D block viewed from northeast and two slices through the model center in x- and y-directions, respectively. Map a) shows the location of the study area in western Germany at the Dutch-German border.

a Horst-Graben-system. The resulting accommodation space was filled with siliciclastic sediments from the Tertiary onwards in alternating sequences of relatively unconsolidated marine and continental deposits. Figure 102 shows the different Tertiary units in the study area and the Quaternary cover. Two lignite seams, represented in red and green, are present in the model. The sedimentary successions overlay the Paleozoic basement rocks unconformably. The 3D geological model focuses on the detailed representation of the Tertiary successions. Thus, the Paleozoic rocks are pooled in one single basement unit represented in grey (Fig. 102). The model represents faults as offsets of model units, as depicted in the 2d slices in Fig. 102, neglecting their representation as explicit model units with distinct properties.

T1.4.5-5 For the presented OED study, we consider only conductive heat flow in the model. Hence, we can formulate the following OED problems: find optimal drillhole locations for temperature logging in order to (i) estimate thermal conductivity of certain model units with least uncertainty and (ii) estimate the basal heat flow with least uncertainty. Table 24 lists the true reference thermal conductivities of all model units. As thermal boundary conditions, the surface temperature is constant at 10.6 $^{\circ}C$ and the basal heat flow is constant at 0.035 $W\ m^{-2}$. The two lignite seams (units 5 and 9) are characterized by low thermal conductivities below 1 $W\ m^{-1}\ K^{-1}$. The other Tertiary sediments range between 1.88 and 2.78 $W\ m^{-1}\ K^{-1}$. The consolidated basement rocks have a relatively high TC. Figure 103 shows the resulting temperature distribution for the purely conductive model along the S-N slice presented in Fig. 102. The temperatur logs from different locations illustrate the effect of the low conductive lignite seams: they act as thermal insulators and increase the thermal gradient. Moreover, the base rocks dip towards north, resulting in a thickening of the sedimentary succession, which in turn results in higher temperature with depth towards the north. This underlines that different drilling locations result in different temperatur logs which may have different effects on the inversion result.

In order to solve the inverse problem, a 500 m deep synthetic temperature log from the model center is generated by disturbing the true reference values with 0.3 K Gaussian white noise. The first OED problem

Table 24: Thermal conductivities λ of model units for the true reference case.

unit	uindex	$\lambda\ [W\ m^{-1}\ K^{-1}]$
Basement	1	3.78
Liegendsande Layers	2	2.50
Grafenberg Layers	3	2.51
Koeln Layers	4	2.51
Seam Morken	5	0.15
Ville Layers	6	2.30
Inden Layers	7	2.78
Hauptkies Series	8	2.68
Seam Schophoven	9	0.90
Upper Rotton	10	1.88
Quaternary Formations	11	2.34

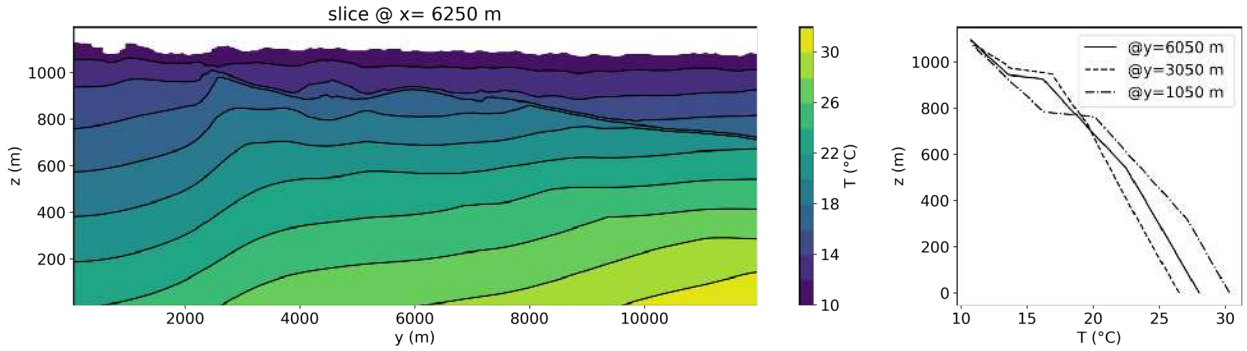


Figure 103: True reference temperature solution illustrated as slice in y-direction through the model center (left), and as temperature logs at three different positions along the slice (right). See Fig. 102 for model structure

Table 25: Inverted units and their prior thermal conductivity for the five OED-scenarios.

Scenario	inverted unit	λ_{prior} $[W\ m^{-1}\ K^{-1}]$
A	1 - Basement	3.0
B	4 - Koeln Layers	2.0
C	5 - Seam Morken	0.5
D	8 - Hauptkies	2.0
E	9 - Seam Schophoven	1.8

is to find a location for an additional 500 m deep temperature log in order to estimate thermal conductivity of a certain model unit with least uncertainty. Five OED scenarios were simulated, in which the thermal conductivities of different model units were estimated. Table 25 lists the five scenarios and the respective prior thermal conductivity in the active unit (see Table 24 for the true reference thermal conductivities). Figure 104 presents the results for all five scenarios as map views of the classified normalized D-optimality. Areas with the lowest D-optimality are the best location for additional 500 m long temperature logs in order to invert for the respective thermal conductivity with least uncertainty. Optimal areas differ for the different scenarios. For example, the optimal area for estimating thermal conductivity of the basement (unit 1, scenario A) is at the southern model boundary. Here, the basement rocks are closest to the surface (see Fig. 102) so that direct temperature data could be obtained from a 500 m deep borehole. Also for the other scenarios, the optimal position is related to the structure of the respective model unit. If the inverted unit can be drilled directly, the optimal position is where the maximum number of temperature measurements can be obtained, i.e. where the unit is thickest.

A second OED problem is to find the optimal borehole location for an additional 500 m deep temperature log in order to estimate the basal heat flow q, which is the boundary condition at the model base. Figure 105 shows the resulting D-optimality in map view and along slices through the optimal position at x=2650 m and y= 5550 m. The optimal position for the 500 m long temperature log is exactly at a fault zone. The temperature contours in Fig. 105 (grey contour lines in 2d slices) show that there is a high thermal gradient at this position, because of the offset and high thickness of Seam Morken (unit 5, red), which has a low thermal conductivity.

6.6 Code demonstrators

ParFlow

ParFlow (v3.8) is a massively parallel, physics-based integrated watershed model. It simulates fully coupled dynamic 2D/3D hydrological, groundwater and land-surface processes for large scale problems. Saturated and variably saturated subsurface flow in heterogeneous porous media are simulated in three spatial dimensions using a Newton-Krylov nonlinear solver ([5]; [75]) and multigrid-preconditioners. ParFlow also features coupled surface-subsurface flow which allows for hillslope runoff and channel routing ([78]). ParFlow is a highly scalable code enabling it to be used for high resolution watershed simulations ranging from single river catchments to continents ([99]). ParFlow is successfully run on a large range of platforms ranging from single CPU notebooks and workstations, to distributed/shared memory clusters to massively parallel systems such as IBM BlueGene P and Q.

Scientific simulations towards exascale

Rapidly changing heterogeneous supercomputer architectures pose a great challenge to many scientific communities trying to leverage the latest technology in high-performance computing. In order to adapt

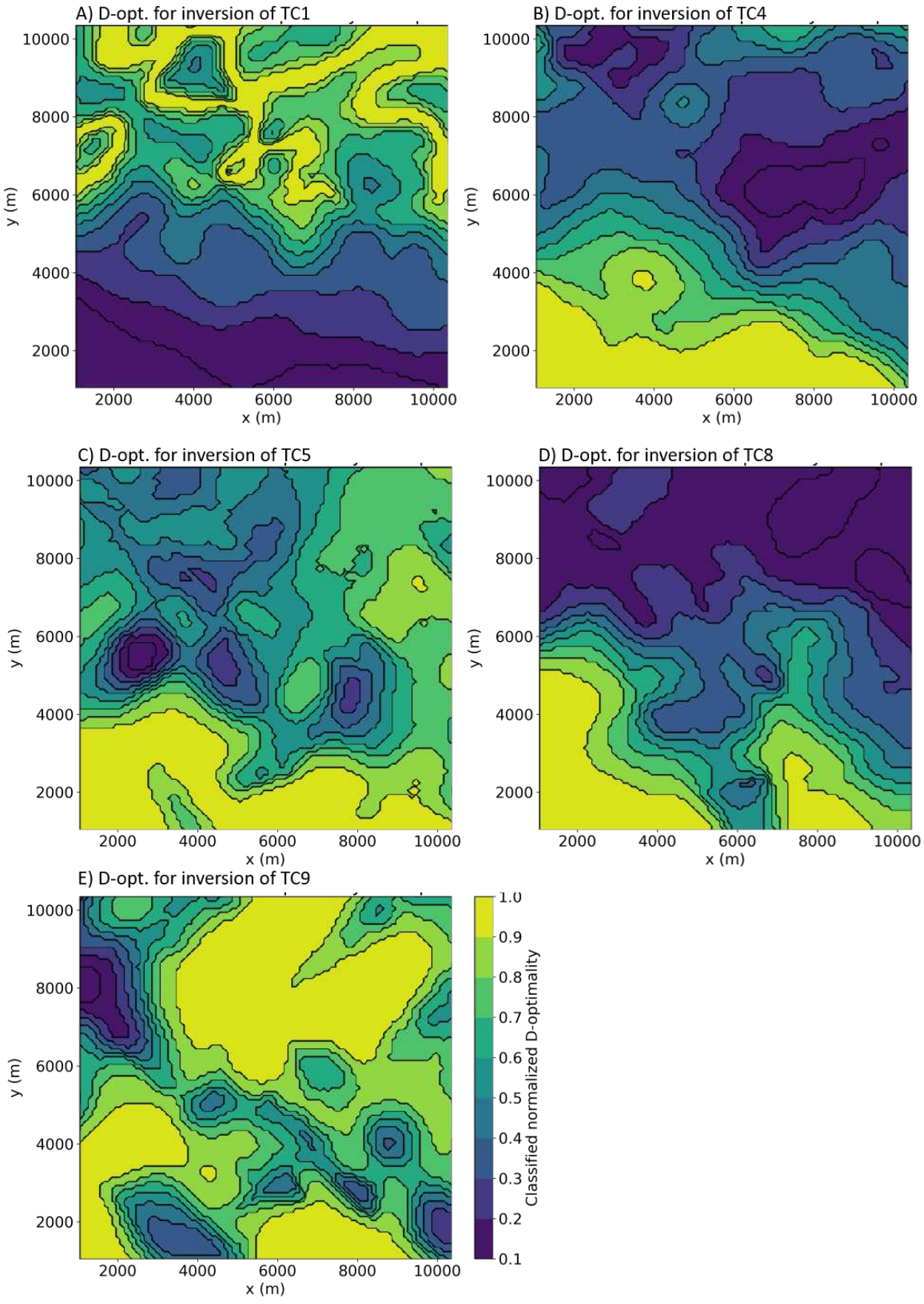


Figure 104: Classified normalized D-optimality for inverting thermal conductivity (TC) of different units in map view. See also Table 25 for scenarios A) to E)

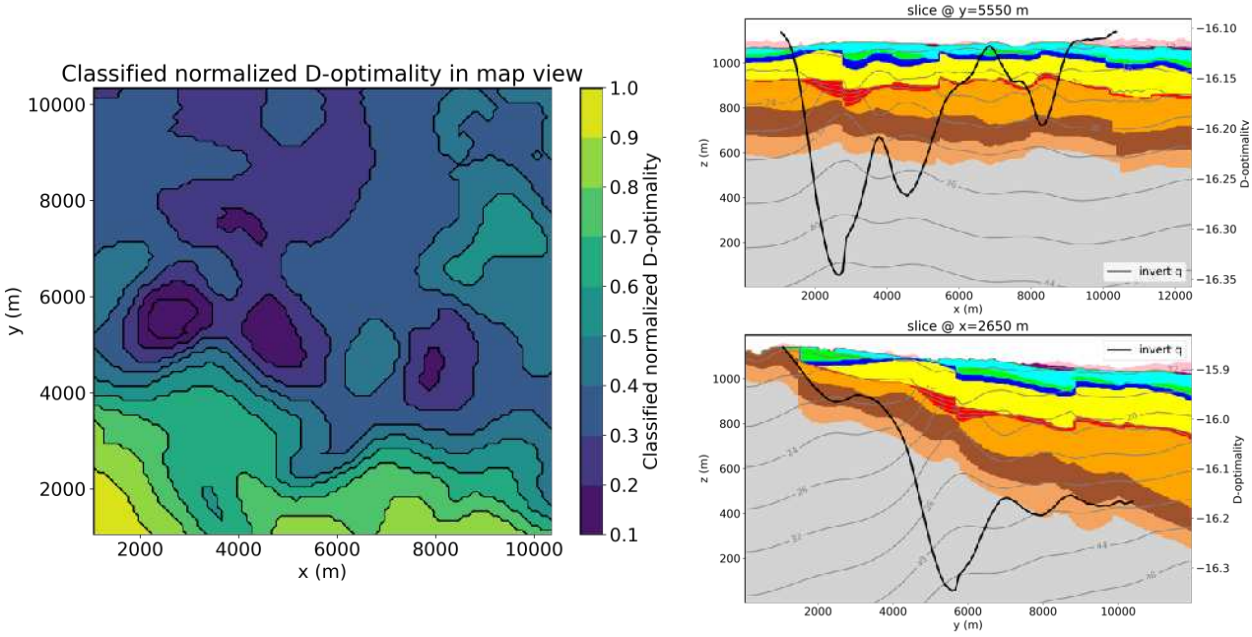


Figure 105: D-optimality for inverting basal heat flow q in map view and along slices through the optimal position. The slices additionally show temperature contours (grey contour lines).

to modern accelerator architectures, the graphics processing units (GPUs) were implemented with the ParFlow code to accelerate model performance in simulating the three-dimensional variably saturated groundwater flow and overland flow[68]. The implementation for accelerator architectures in ParFlow hydrologic model demonstrates that a significant performance gain, high developer productivity, and minimally invasive implementation are all achievable at the same time while keeping the codebase well maintainable in the long-term. A representative benchmark problem was run on the booster module1 of the JUWELS supercomputer where each utilized node is equipped with dual AMD EPYC Rome 7402 processors (2 × 24 cores @ 2.8 GHz) and 4 NVIDIA A100 40 GB GPUs. The nodes are connected through 4 HDR200-InfiniBand devices. It is expected that more HPC systems in the near future are adopting a design similar to that of JUWELS Booster (around 50 CPU cores and 4 GPU devices per node). The benchmark consist of a variably saturated infiltration problem into a homogeneous soil with a fixed water table at a depth of 6m and a constant infiltration rate of 8 x 10^{-4} m hour^{-1}. The vertical and lateral spatial discretization was 0.025 and 1 m, respectively. The time step size was 1 h. The profile was initialized with a hydrostatic profile based on a matric potential of -9m at the top resulting in a considerable initial hydrodynamic disequilibrium with respect to the water table at the bottom boundary. The number of grid cells in the lateral directions was varied to change the total number of degrees of freedom in the performance testing. The reference results were obtained without accelerator devices by launching an MPI process for each CPU core. In case of the GPU accelerated runs, 4 MPI processes per node were launched such that each process uses one GPU and GPU-GPU data transfer is handled by a CUDA-aware MPI. The current, moderately optimized ParFlow GPU version runs a representative model up to 20 times faster on a node with 2 Intel Skylake processors and 4 NVIDIA V100 GPUs compared to the original version of ParFlow, where GPUs are not used (Figure 106). These results demonstrate that the ParFlow GPU implementation may serve as a blueprint to tackle the challenges of heterogeneous HPC hardware architectures on the path to exascale.

Below is the url of ParFlow-GPU (version 3.8) repository address

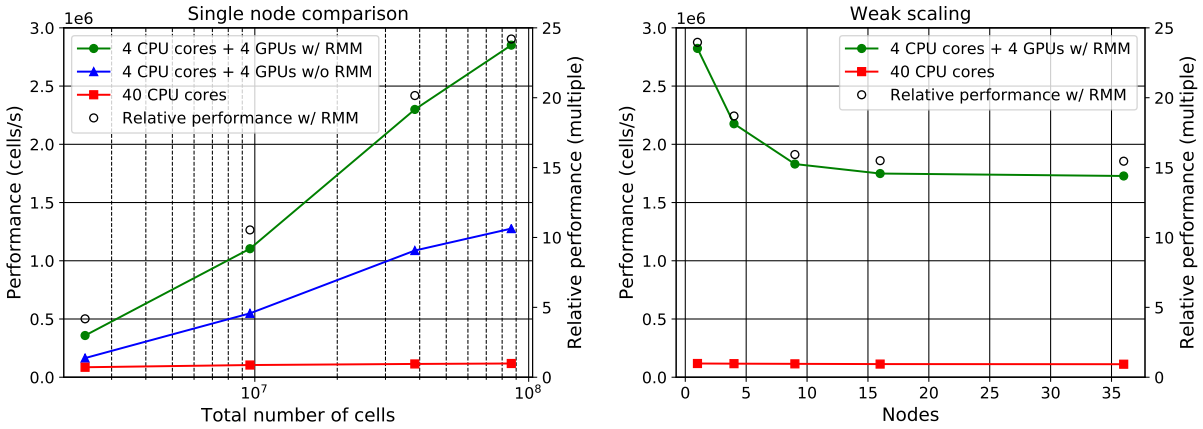


Figure 106: (a) Single node performance comparison., (b) Weak scaling comparison.

`https://github.com/hokkanen/parflow`
Demonstrator release v3.8

6.7 Summary

In the Water for Energy scientific challenge, main achievements are summarized below:

T1.4.1 This task uses ParFlow hydrologic model to simulate hydrologic states and fluxes relevant to the energy sector. ParFlow is a massively parallel, physics-based integrated hydrlogic model and simulates fully coupled dynamic 2D/3D hydrological, groundwater and land-surface processes for large scale problems. In order to allow for computationally demanding high resolution simulations and to also enable efficient (big) data handling, processing and analysis features of the integrated terrestrial modeling system, ParFlow code was adapted to modern accelerator architectures[68]. Using ParFlow-GPU version (see details in Section 6.6), the model was setup at 3 km resolution over Europe and an explicit simulation of lateral groundwater flow, groundwater discharge and recharge was performed for 10 years of time period (1997 - 2006). In addition, a comprehensive evaluation of hydrologic states and fluxes was performed using in-situ and remote sensing observations including discharge, surface soil moisture, evapotranspiration and water table depth (Task T1.4.2). Overall, the uncalibrated ParFlow model shows good agreement in simulating river discharge for 176 gauging stations across Europe. Comparison with satellite-based datasets of SM shows that ParFlow performs well in semi-arid and arid regions, but simulates overall higher SM in humid and cold regions. Comparisons with reanalysis and remotely sensed ET datasets (GLEAM and GLASS) show no significant differences, both, across the European domain and within regions. This work also demonstrates the added advantage of using a 3D parallel integrated hydrologic model to produce multi-scale processes such as lateral groundwater flow, streamflow and water level fluctuations. We conclude that the addition of alluvial aquifers in a continental scale hydrologic model can illuminate the complex relationship of water transfer to and from river systems and also significantly affect the transportation of water through the continent. While, in this task we mainly focused on evaluating overall model performance over a pan-European model domain, it should be noted that we did not address the uncertainties in the ParFlow model parameters that are required for model simulations such as hydraulic conductivity, porosity, soil and vegetation parameters which may introduce biases in our results. Because of the associated computational cost with ParFlow, sensitivity studies of water balance variables to these parameters are difficult. With the ongoing model developments and collaborative efforts to improve computational efficiency of ParFlow with its GPU version (e.g. Hokkanen et al., 2021) and ensemble-based sensitivity analysis tools (work pack-

age WP5) it will be possible in the future to also conduct continental-scale ensemble-based sensitivities analyses for quantifying model parameter uncertainties.

T1.4.2 HYPERstreamHS hydrological model was refactored in order to allow explicit representation of hydropower systems while preserving good computational performance overall: this was achieved by adopting a dual-layer parallelization strategy that preserved the model's scalability up to 64 processors, ensuring a 40x speed-up [6]. The Human System modules have been tuned with reference to the Adige river basin, in order to ensure reliable prediction basing the modules solely on publicly available information: this allows to apply the framework to other areas, provided that the required information is available. The model showed great performances in reproducing historical streamflows and hydropower production, highlighting some shortcomings of the main approaches to large scale hydropower modelling that are present in literature [53]. After compiling an extensive dataset containing the relevant information, the model was then applied to other large Alpine catchments, showing satisfactory performances. A benchmarking exercise aimed at the identification of suitable modifications of the Community Land Model v3.5 (CLM3.5) runoff outputs, to improve streamflow reproduction was conducted in the Adige river basin. This activity highlighted that the inclusion of some physical processes, as well as replacing the grid based routing scheme of CLM3.5 with a more accurate scale-independent routing scheme can greatly improve the ability of CLM3.5 to reproduce observed streamflows, particularly by improving its reproduction of sub-surface flow component. Overall, in this Task we developed useful tools for highly reliable streamflow and hydropower production modelling, as well as setting them up for future coupling with hyper-resolved products.

T1.4.3 In Task T1.4.3-2, the mixture of experts surrogate model implemented in T1.4.3-1 has been used to carry out a global sensitivity analysis to classify sources of uncertainty to explain water level variance over the 2D domain. This work was carried out for stationary flow and led to the estimation of Sobol indices maps in the framework of S. El Garroussi PhD at CERFACS. The cost reduction allowed by the surrogate model used in place of the direct solver allows to compute Sobol indices and rank the sources of uncertainties over the Garonne Marmandaise catchment. Analysis of the first order Sobol indices reveals the large influence of the discharge and smaller influence of friction. It should be noted that the surrogate model was established for stationary flow, assuming the uncertainties relate to friction and upstream forcing. The original plan was to force the surrogate T2D model with ParFlow discharge outputs at the local Telemac model over 50km of the Garonne river. Yet, the upgrade of the surrogate model previously described from stationnary flow to non-stationary flow was not achieved in the context of the PhD, as planned (the PhD was terminated by the student for personal reasons in early 2021). ParFlow was thus chained to the direct model Telemac only. While the use of a surrogate model would have lowered the computational cost of the chained simulation, the IO interfaces for the direct and the surrogate remain similar. As a consequence, the hydrology-hydraulic chained model presented here is representative of the dynamics of the expected flow (even with a surrogate), and the solution for technical aspects treated here still valid when the surrogate model is used. The ParFlow-Telemac chained hydrology-hydraulic model was implemented and tested over 2003-2004 with a focus on 2 significant flood events. It was shown that ParFlow provide good discharge timeseries, yet imperfect with underestimation of the flow and a 5 day delay. These uncertainties translate into similar errors in the outputs of the local model outputs with Telemac. Several strategies are possible for improvement ranging from ParFlow calibration of friction, ParFlow off-line rescaling for discharge or error correction in Telemac with Data Assimilation in the local hydrodynamics model only. These are possible leads for further research.

T1.4.4 and T1.4.5 The presented OED workflow is applicable to geothermal reservoir models with the open-source code SHEMAT-Suite for solving different OED problems in context of geothermal exploration. Several OED problems were simulated on synthetic models for investigating the influence of prior assumptions regarding estimation of permeability and thermal conductivity. In addition, optimal borehole locations were defined for a realistic 3D reservoir model in order to estimate thermal conductivity or basal heat flow

with least uncertainty. Due to reduced work efficiency and missing childcare during the pandemic, some research topics could not be addressed as deeply as intended initially. In addition, some time was spent on working for WP3 (integration of the AGMG solver) and the granted funding did not last for the intended 16 PMs. Therefore, the testmodel suite is smaller than intended initially (T1.4.4-6). Transient test models or an advective 3D model are missing. Furthermore, the OED functionality has not been extended to optimizing not only borehole location but also borehole depth. Regarding T1.4.4-7, the parallelism of EFCOSS has not been investigated. On the one hand, because of lack of time. On the other hand, because it was not really necessary for the performed simulations, since they ran maximum around 20 minutes serially. Finally, T1.4.5 was adapted in order to be feasible within the remaining time, resulting in a quasi-synthetic scenario based on a realistic 3D structural model and realistic rock properties.

7 Fusion for Energy (T1.5)

Task 1.5 aims at paving the way towards the exploitation of GYSELAX , and is divided in three main tasks:

T1.5.1	Prototype of GYSELAX : arbitrary magnetic equilibrium in limiter configuration
T1.5.2	Advanced GYSELAX : X-point configuration & alternative/complementary methods
T1.5.3	Core-edge-SOL physics: GYSELAX

7.1 T1.5.1: Scientific results

This task focuses on the basics with the help of prototypes, namely the derivation and implementation of the new set of equations and the identification of optimal choices for handling non-circular geometries and boundary conditions.
The main task T1.5.1 is subdivided in two subtasks:

T1.5.1-1	**Identify minimal set of equations for ion turbulence**: Extending turbulence and transport simulations towards the edge of tokamak plasmas requires accounting for electromagnetic effects and plasma-wall interaction. To this end, the Maxwell-Ampère equation and upgraded boundary conditions have to be derived within the gyrokinetic framework and implemented.
T1.5.1-2	**Comparative efficiencies of numerical strategies (flux-surface aligned and Cartesian mesh grids in poloidal plane, and of patches with regular or irregular meshes)**: In the ITER-relevant cases of non-circular poloidal cross-sections of the toroidal magnetic surfaces, the treatment of both Maxwell (Poisson and Ampère) and Vlasov equations needs to be updated and optimized as compared to the simpler circular case. Several strategies are benchmarked and their respective advantages compared.

T1.5.1-1 Minimal set of equations

In D1.2, we reported that the Ampère equation – which allows electromagnetic turbulence involving fluctuations of the magnetic field to be addressed – was implemented and benchmarks are ongoing. Regarding the Scrape-Off Layer (SOL) where magnetic field lines intercept solid materials, immersed boundary conditions were successfully implemented in GYSELAX for the ion response only. The full regime with both kinetic ions and electrons was under verification in VOICE, a low-dimensional version of GYSELAX .

Since M18, critical advances in these directions have been achieved:

1. The implementation of the Ampère equation has been tested against the tearing mode instability. This large scale instability taps energy from the gradient of the equilibrium current; it involves helical perturbations of the parallel (to the equilibrium magnetic field $\mathbf{B}_{eq}$) current $\tilde{j}_{\|}$ which develop at low-order rational surfaces of the field line helicity, called resonant surfaces. The associated perturbation

of the parallel vector potential $\tilde{A}_{\parallel}$ generates a perturbed component of the magnetic field transverse to $\mathbf{B}_{eq}$, possibly modifying the properties of the magnetic topology. Extensive electromagnetic simulations of GYSELAX have revealed that, while we recover the expected profile of $\tilde{A}_{\parallel}$ outside the resonant layer, the localized perturbed current sheet $\tilde{j}_{\parallel}$ does not exhibit the right parity with respect to the resonant surface. This is likely the reason why we do not recover the expected linear growth rate of the tearing instability. We suspect an inconsistency in the implementation of the modified particle trajectories. Dedicated tests are ongoing. The active collaboration between CEA-IRFM and MP-IPP, expert in this matter, will allow us to resolve the issue.

2. Because of their small inertia, electrons move much faster than ions along the magnetic field lines. When the plasma encounters the wall, an intense parallel electric field develops in the close vicinity of the wall to confine electrons and accelerate ions, ensuring a vanishing current into the wall. This physics, which bounds the electric potential in the SOL to the electron temperature and the electron to ion mass ratio, is an essential ingredient of the core-edge-SOL interplay. However, plasma-wall interaction develops at such a small scale – in the so-called sheath of a few Debye lengths ($\lambda_D \sim 10^{-4}$ m) – that it will not be resolved in GYSELAX simulations, where the typical grid size in the parallel direction is of about 10 cm. VOICE aims precisely at resolving the sheath physics along a magnetic field line with kinetic ions and electrons. The same penalization technique as in GYSELAX is used to treat the wall as an immersed boundary. The equations solved by the code have been upgraded to include (i) both a kinetic single species and a fluid inter species collision operator, and (ii) versatile sources and sinks to test the benefit, drawback and overall impact of several physical assumptions. From the numerical point of view, non equidistant splines have been successfully implemented, allowing for a refined grid mesh at the sheath location. The last but not least point deals with HPC: VOICE has been ported to GPU. All these critical upgrades pave the way of their implementation in GYSELAX . A paper on the numerical and HPC part is in preparation [E. Bourne et al. (2022)].

In conclusion, the initial objectives of this sub-task have been achieved: GYSELAX now runs in the electrostatic regime from the confined very core to the unconfined far edge Scrape-Off Layer, with both ions and (trapped) electrons treated kinetically in the confined region. The issue of kinetic electrons in the SOL has been addressed in depth with the VOICE code – reduced version of GYSELAX – with critical advances in physics, numerics and HPC, paving the way to the implementation in GYSELAX . Addressing the electromagnetic regime was not part of the initial targets and appears as a bonus. The Ampère's equation has been implemented and the equations of motion have been modified accordingly; the unexpected parity of the linear modes suggests that the numerical treatment of this latter part needs to be revised. This task will be one of our priorities in the coming months.

T1.5.1-2 Comparative efficiencies of numerical strategies

This subtask is divided into two activities, each dealing with one of the two main solvers: (i) the Poisson (& Ampere) solver and (ii) the Vlasov solver. In report D1.2 the work was in progress. Since M18, the work has been finalised:

- A test platform has been developed to allow the comparison of different numerical schemes for the Poisson solvers. These results are the subject of a paper currently being written [20].

- The Poisson solver based on finite elements projected into a spline base has been coupled to GYSELAX and the Vlasov equations have been modified so as to be able to consider a more realistic magnetic configuration. The GYSELAX code is now able to run both in circular or D-shape type geometries. The Culham magnetic configuration has been successfully validated on Geodesic Acoustic Mode simulations (GAM).

All these activities were carried out in close collaboration with CERFACS and IPP via WP3 workpackage, as well as in the framework of Kevin Obrejan's post-doc (funded by EoCoE-II) and Emily Bourne's PhD (funded by European NUMERICS PhD program).

Development of a test platform for 2D Poisson solvers on D-shape geometries In GYSELAX , axisymmetric assumptions allow the simplification of the 3D Poisson solver in (r, θ, φ) to N_φ Poisson equations which are solved on the (r, θ) plane describing the poloidal cross-section of the torus. In the GYSELAX code each 2D Poisson equation was solved by a Fourier projection in the periodic poloidal direction θ and by using finite differences of second order in the radial direction. This strategy is not adapted to non-circular poloidal cross-sections since Fourier modes are no longer poloidal eigenmodes of the operator in this case. The objective of this sub-task was to test different 2D solvers able to tackle D-shape geometries in terms of accuracy of the solution and computational efficiency. To this aim, we developed manufactured solutions to solve the gyrokinetic Poisson equation with homogeneous Dirichlet boundary conditions

$$\begin{aligned} Lu = -\nabla \cdot (\alpha \nabla_\perp u) + \beta u = f \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{aligned} \tag{16}$$

where $\Omega \subset \mathbb{R}^2$ is a disk-like domain, $f : \Omega \to \mathbb{R}$, and $\alpha : \Omega \to \mathbb{R}$ is a coefficient involving the density and magnetic field profiles. Moreover, $\nabla_\perp$ is the gradient in the direction orthogonal to the magnetic field. If $\mathbf{b}$ is a unit vector along the magnetic field, $\nabla_\perp u = \mathbf{b} \times \nabla u \times \mathbf{b}$.
As a reminder the three solvers chosen in D1.2 were:

- **A Finite Elements solver based on B-splines**[4] as proposed by Zoni et al [165, 164]. This solver has specially constructed basis functions around the singularity to ensure that a $\mathcal{C}^1$ smooth solution can be found. The matrix system is solved with a preconditioned conjugate gradient method using a simple Jacobi preconditioner.
- **A multigrid solver, called *gmgpolar***[5], based on an implicit extrapolation method [76] and a classical multigrid cycle where the line smoother is especially tailored for disk-like shapes. The development of such multigrid solver dedicated to curvilinear coordinates (r, θ) was part of WP3 workpackage and led to the writing of two papers [82, 81].
- **An embedded boundary solver**[6] **based on the AMReX library** [47, 149]. Unlike the first two which use curvilinear coordinates, it uses a simple Cartesian mesh where the discretization is based on a box that includes the full physical domain. Then the physical domain is cut out of the box by finding the intersections of the domain boundary and the cell boundaries (see Figure 107b). The physical domain is finally represented by the piecewise linear representation connecting the intersection points with the cell boundaries. The AMReX library implements a finite volume solver based on the work of Johansen and Colella [73].

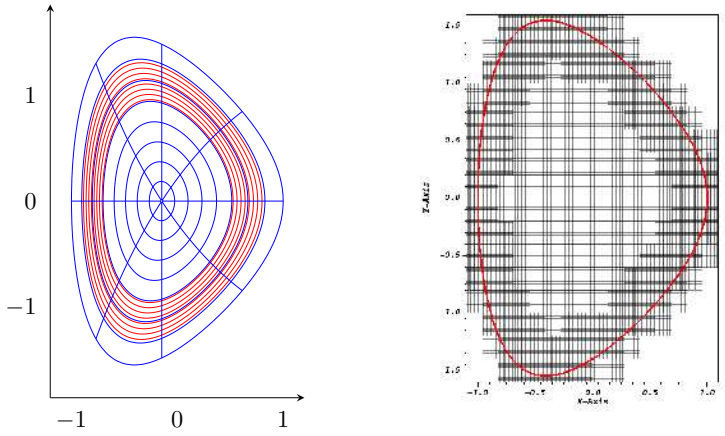
The spline FEM solver was extracted from the SELALIB[7] library developed in Fortran while the two others are written in C++. Efforts have been made to define analytically relevant test cases and to homogenize the input and domain definition in the three codes. Python scripts [8] have been developed to run and compare the three solvers. We show in Figure 107 the kind of radial refinement used for curvilinear geometry (a) and an example of refinement of 2 levels for cartesian coordinates (b) for the D-shape geometry defined in

[4] https://gitlab.maisondelasimulation.fr/ebourne/culham-metric/-/tree/master/poisson_code
[5] https://gitlab.com/mknaranja/gmgpolar
[6] https://github.com/kkormann/quasineutral
[7] https://github.com/selalib/selalib
[8] https://gitlab.maisondelasimulation.fr/ebourne/culham-metric/-/tree/master/poisson_code/tests/python_code

Czarny's paper as

$$
\begin{aligned}
x(r,\theta) &= \frac{1}{\varepsilon}\left(1-\sqrt{1+\varepsilon\left(\varepsilon+2r\cos\theta\right)}\right) \\
y(r,\theta) &= y_0 + \frac{e\xi r\sin\theta}{2-\sqrt{1+\varepsilon\left(\varepsilon+2r\cos\theta\right)}} = y_0 + \frac{e\xi r\sin\theta}{1+\varepsilon x(r,\theta)}
\end{aligned}
\tag{17}
$$

where y_0 corresponds to the mapping center, ε is the inverse aspect ratio, e the ellipticity, and $\xi = 1/\sqrt{1-\varepsilon^2/4}$. Results for a manufactured solution given by the following equations:



(a) r-refinement around $\frac{2R}{3}$ (b) Adaptive refinement of the Cartesian grid

Figure 107: Example of refined meshes used in curvilinear coordinates (a) and cartesian coordinates (b) for the analytical mapping defined by equations (17)

$$
f(r,c,w) = \exp(-(r-c)^2/w) \tag{18}
$$

$$
u(x,y) = g(r(x,y),0.45,0.02)\cos(9\theta) + g(r(x,y),0.9,0.0003)\cos(21\theta) \tag{19}
$$

$$
\begin{aligned}
\text{with } g(r,c,w) =& f(r,c,w) - r\partial_r f(0,c,w) - f(0,c,w) \\
&+ (r\partial_r f(0,c,w) + f(0,c,w) - f(1,c,w))r^2
\end{aligned}
$$

are shown for the spline FEM solver in Figure 108 where we find a numerical error of the order of 10^{-6} for 204 non-uniform points in radial direction and 512 uniform points in poloidal direction.
The choice of this manufactured solution requiring a more refined mesh at the edge of the D-shape geometry has the advantage of mimicking the type of problem encountered in the GYSELAX code. The three solvers have shown their ability to solve a Poisson-type equation on a more realistic tokamak poloidal cross-section. The work of the last few months has also shown that they can handle non-uniform meshes with refined meshes to take into account more realistic gradients at the edge of the plasma.
The advantages and disadvantages of the three solvers are described in a paper currently being written [20]. The two solvers based on curvilinear coordinates are of 4th order while the embedded boundary solver is of second order. Conversely, one of the major advantages of the Cartesian mesh-based solver is that it will be much more suitable for handling more complex geometries such as X-point geometries.

Implementation and validation of a more realistic geometry in GYSELAX At the beginning of the EoCoE-II project GYSELAX was only able to tackle circular concentric magnetic configurations. However, the geometry of the plasma – i.e. the shape of the magnetic flux surfaces – is governed by the MHD equilibrium, which can be recast in the form of the Grad–Shafranov equation. One objective of the project was to implement a Culham magnetic configuration in the GYSELAX code. Indeed, Culham equilibrium

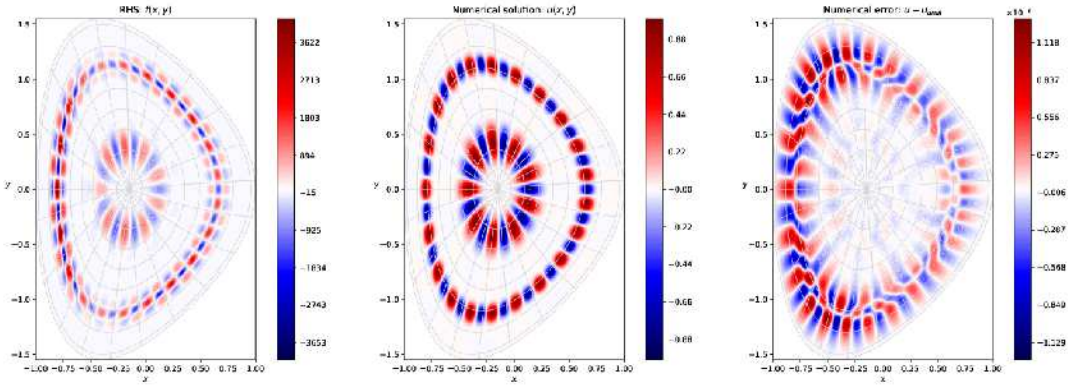


Figure 108: Manufactured test with spline FEM solver for Czarny's geometry: (a) Analytical RHS $f(r, c, w)$ given by eq. (18) and numerical results for 204 non-uniform points in radial direction and 512 uniform points in poloidal direction: (b) numerical solution $u(x, y)$ and numerical error $u(x, y) - u_{anal}(x, y)$ with $u_{anal}(x, y)$ defined by eq. (19).

[28] is a class of Grad–Shafranov solutions in the $\epsilon \equiv a/R_0 \to 0$ limit , which is valid up to $\mathcal{O}(\epsilon^2)$. The flux surfaces are parameterized in the form

$$\begin{aligned} R = &\quad R_0 + \Delta(r) + (r - E(r) - P(r)) \cos\theta + T(r) \cos(2\theta) \\ Z = &\quad (r + E(r) - P(r)) \sin\theta - T(r) \sin(2\theta) \end{aligned} \tag{20}$$

where the functions $E(r)$ and $T(r)$ relate respectively to the –local– elongation $\kappa(r) \simeq (r + E(r))/(r - E(r))$ and –local– triangularity $\delta(r) \simeq 4T(r)/r$ of the flux surface, $\Delta(r)$ is the Shafranov shift, and P is a correction term. The advantage of such an equilibrium is that the co-variant and contra-variant metric tensors can be derived analytically.

To take into account this more realistic geometry, all the main kernels of the GYSELAX code have been modified, namely the Vlasov equation, the gyroaverage operator, the quasi-neutrality solver as well as the diagnostics. For the sake of simplicity, among the three solvers that have been tested, the implemented quasi-neutrality solver is the spline-based FEM. More details regarding the implementation can be found in task T2.6.2(iii) in workpackage WP2.

Zonal Flows (ZFs) and their oscillatory components play an important role in the saturation of turbulence in tokamaks. The kinetic theory of GAMs in toroidal plasmas (see Figure 109) is well documented (see for instance a brief review in [123]). So as to validate their correct treatment in gyrokinetic codes, the analytical predictions regarding the time evolution of GAMs and ZFs have been widely compared with simulation outputs. In Biancalani's paper [15], benchmarks were done for three of the European ORB5, GENE and GYSELAX codes. In Figure 2 of this paper, the results concerning the dependence of GAM frequency and GAM damping rate according to plasma elongation were missing for GYSELAX due to the circular magnetic configuration constraint. The adding of the Culham magnetic configuration removes this constraint and GYSELAX results have been successfully validated with the two other codes as shown in Figure 110. There, the mismatch with the analytical theory developed by Gao et al. (from a few percent to about 35% for the frequency and the damping rate, respectively) is due to the simplifying assumptions required in the analytical developments. The impact of plasma shape on Ion-Temperature-Gradient turbulence is an ongoing work.

7.2 T1.5.2: Scientific results

This task deploys in the advanced version the solutions retained in the prototypes, and extends the problem to critical large temperature variation and X-point issues. The main task T1.5.2 is subdivided in two subtasks:

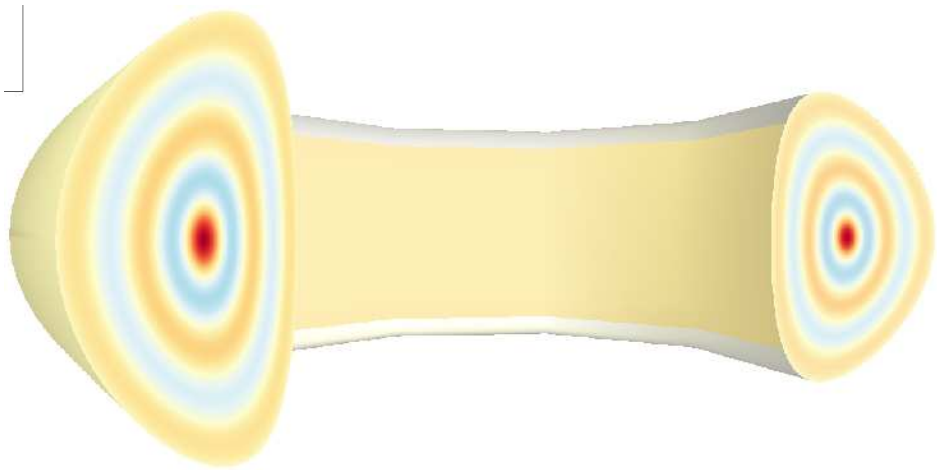


Figure 109: Amplitude of the electrostatic potential in a 3D D-shape geometry for a Geodesic Acoustic Mode (GAM) initalized with the radial mode number $k_{rn} = 5$ where k_{rn} is the number of roots of the Bessel function between the center $r = 0$ and the minor radius $r = a$.

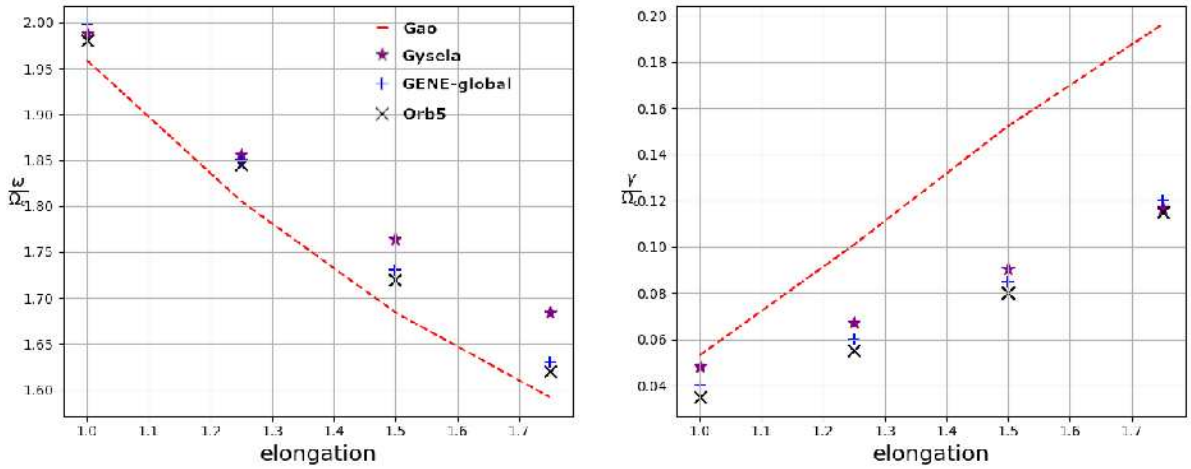


Figure 110: Frequency (left) and damping rate (right) of the radial electric field versus elongation for GYSELAX (magenta stars) compared with ORB5 (black Xs), GENE global (blue crosses) and the analytical theory of Gao-2009 [54] (dashed red line).

T1.5.2-1	**Handle the X-point of ITER-relevant magnetic configurations with multi-patch or flux coordinate independent schemes**: ITER-relevant magnetic configurations include an X-point on the outermost closed magnetic surface, characterized by the nullity of one of the magnetic field components. With the O-point located at the magnetic axis, it represents a singular point of the configuration in the usual curvilinear representation. They both require a dedicated treatment.
T1.5.2-2	**Multi-patch treatment of the Vlasov equation to handle large variations of temperature**: The plasma temperature varies by several orders of magnitude (3 to 4) from the very core to the far edge in tokamaks. A single patch and/or a uniform grid would be prohibitive in terms of numerical memory. Alternative strategies need being implemented and tested.

T1.5.2-1 Handle the X-point

Handling the X-point is one of the Holy Grails for gyrokinetic codes. It is an extremely numerically complex challenge. Advances have been made during this project to remove some of the numerical bottlenecks.

- In the first 18 months, the numerical challenge to treat spatial regions of both closed and open magnetic field lines in a single simulation was addressed by means of the gyrokinetic GENE code. To this aim, the so-called "flux coordinate independent (FCI)" approach – pioneered in the fluid turbulence code FENICIA (F. Hariri et al., Plasma Phys. Control. Fusion 57 (2015) 054001, CEA-IRFM) and later applied in GRILLIX (A. Stegmeir et al., Plasma Phys. Control. Fusion 60 (2018) 035005, MP-IPP) – was extended to GENE such that both codes now use the same FCI routines (as an API). This technique was tested thoroughly and successfully via the method of manufactured solutions.
- Since M18, we investigate numerical schemes to handle the full tokamak –including the region outside the separatrix– with a flux surface aligned grid. As mentioned in section 7.1 one of the bottlenecks for Poisson solvers based on curvilinear coordinates is the handling of X-point geometry. To overcome this problem computational domain can be decomposed into several patches that are connected at the separatrix. On each of them a mapping based on B-Splines can be constructed and the B-Splines used both for computing the fields with a Finite Element method and for interpolating the distribution function in the semi-Lagrangian method.
  - As a first step, a technique for handling the interface between the different regions based on a non-conforming Galerkin method has been analyzed and successfully implemented in the PSYDAC code[9] for the field solvers for the case of a smooth connection between patches. This enables the handling of all the interfaces except the X-point.
  - In a second phase a technique similar to the one that we developed for the magnetic axis (O-point) has been investigated, but not yet implemented.

Given the successful implementation of the FCI approach in several codes exhibiting X-points (FENICIA, GRILLIX and GENE) on the one hand, and on the other hand, the fact that the grid mesh based on flux-aligned coordinates needs being reconstructed each time the magnetic equilibrium is changed – and not mentioning the complexity of matching the patches–, the future strategy for GYSELAX is to adopt Cartesian coordinates in the poloidal cross-section combined with the FCI approach.

T1.5.2-2 Multi-patch treatment of Vlasov equation

A multi-patch strategy was initially proposed in the project to address the large temperature variations from the very hot core to the far cool edge. As already mentioned in the mid-term report, this strategy was abandoned because it would have required an almost complete rewriting of the code, which was no longer an option after the CEA-IRFM team lost one of the pillar developers of the GYSELAX code. However, GYSELAX still needs to be able to address theses large temperature variations as this represents a major issue in core-edge turbulence interplay in tokamak plasmas. To this end, we have decided to treat this intrinsic difficulty by using non-equidistant splines. At M18, semi-Lagrangian schemes based on non-uniform splines were under investigation.

Since M18, numerical methods based on non-uniform splines of varying degrees have been successfully developed in the 2-dimensional (1D-1V) VOICE kinetic code to run kinetic semi-Lagrangian simulations of the plasma sheath. The sheath describes a region of plasma in contact with a wall, acting as a heat, momentum, and particle bath. This region is particularly difficult to simulate, due to its kinetic nature and the presence of steep gradients requiring fine numerical resolution. This 2D reduced model was very

[9] PSYDAC is a Python 3 library for isogeometric analysis: `https://github.com/pyccel/psydac`

well designed to find the most suitable method to implement in the GYSELAX code. We show in a paper recently submitted [21] that non-equidistant points constructed from Greville abscissae of judiciously chosen non-uniform knots improve the simulation of the plasma-wall transition. The semi-Lagrangian scheme applied on this non-equidistant mesh presents very good conservation properties as shown in Figure 111. This choice allows a reduction of the memory requirements by 89%. This strong gain added to a GPU

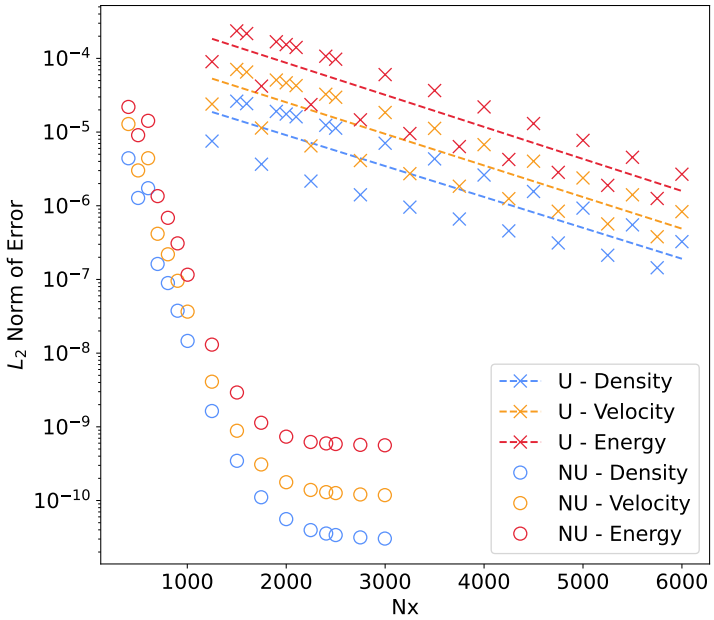


Figure 111: Conservation errors at time t=2000 with 2000 points in the velocity dimension for density (blue), velocity (orange) and energy (red). "U - X" indicates the X conservation for uniform splines of degree 3, while "NU - X" indicates the X conservation for non-uniform splines of degree 3. The saturation of the error at 10^{-9} is expected due to the truncation error in the computation of the integrals in velocity space.

parallelisation (see T2.6.2) was clearly essential for the study of the plasma-wall interactions discussed in section 1.5.3-2.

7.3 T1.5.3: Scientific results

This task takes advantage of all the developments to address first physics issues and is subdivided in two subtasks:

T1.5.3-1	**Large scale equilibrium flows expected in the SOL**: In the scrape-off layer (SOL), magnetic field lines intercept material boundaries. There and in its vicinity, large scale flows develop – transverse to the magnetic field – that strongly impact turbulent transport and in fine the overall plasma confinement in tokamaks. These require dedicated studies and benchmarks with respect to available experimental data.
T1.5.3-2	**Core, edge and SOL interplay in turbulent regime**: Extending core turbulence studies to the unconfined region of tokamak plasmas is motivated by the experimental evidence that SOL conditions strongly impact the overall plasma confinement. Understanding this interplay is of uttermost importance in view of predictive capabilities, although most challenging from both HPC and theoretical points of view. Capitalizing on developments and critical upgrades within EoCoE-II has permitted major breakthroughs in this direction.

T1.5.3-1 Large scale flows

In D1.2, we reported on the modification of edge plasma flows when comparing two simulations with adiabatic electrons, either with a homogeneous edge boundary layer or with a poloidally localized limiter. With limiter only, we observe the development of a strongly sheared poloidal flow in the vicinity of the limiter, as well as a parallel flow. The former tends to significantly increase the linear threshold of the main instability, leading to a steepening of the temperature. We suggested that the observed increase of edge turbulence resulted from the possible role of the sub-dominant parallel velocity gradient instability.

Since M18, we have extended our analysis and simulations in the following directions:

1. A detailed analysis of the charge conservation equation allowed us to diagnose the dynamical build up of the strongly sheared poloidal flow (which scales like the radial electric field E_r) in the limiter case. The shear of E_r is suspected to be key in accessing to the improved confinement regime called H-mode, one of the reference scenarios in ITER. The causality is elucidated thanks to the "transfer entropy" method imported from information theory. It appears that, close to the limiter, it is primarily driven by the diamagnetic component of the turbulence-driven Reynolds stress tensor. The poloidal inhomogeneities of the radial electric field are then propagated poloidally by the diamagnetic flow, hence contributing to making global the locally born flow. Surprisingly, the electric component of the Reynolds stress – formerly regarded as the dominant (if not the only relevant) component – becomes dominant at later time only and downstream of the limiter. Our work underlines the crucial role of pressure inhomogeneities and finite Larmor radius effects in barrier build-up and the access to bifurcated states of enhanced confinement. Two papers have been published on this issue [131, 36]. Experimentally, the power threshold to bifurcate towards the H-mode depends on whether the ion drift velocity points towards or away from the limiter. We therefore expect changes in the profile of E_r whether the limiter is on top or at the bottom of the machine. A low-resolution simulation has been run with the limiter located on top of the machine. Our preliminary analysis reveals the observed changes in the E_r profile as compared to simulations with bottom limiter, although qualitatively consistent with experimental observations, seem to remain transient and weaken in the long time limit. This work will be carried on.

2. So far, only adiabatic electrons were considered. GYSELAX has been upgraded to account for kinetic electrons in the confined plasma, while still adiabatic in the unconfined outer region[10]. In this case, an additional type of turbulence is expected, driven by a certain class of electrons – and in particular responding differently to density and temperature gradients than the ion turbulence which characterizes simulations with adiabatic electrons. The faster dynamics (resulting from the difference of inertia) and smaller characteristic transverse scales of electrons as compared to ions requires much smaller time steps and refined mesh grids. They have only been possible thanks to the critical upgrades of GYSELAX performed within EoCoE-II. Their characteristics are detailed in section7.4.

3. With the same aim to study the impact of sheared poloidal flows on turbulence, new simulations were run with a polarization source, amenable to a momentum source. Consistently with experiments, turbulence is found to be reduced in magnitude and its characteristics modified, with convective cells more elongated poloidally and shorter radially. The resulting transport is found to be drastically

[10] This last simplification will be alleviated in the future (it is part of the EoCoE-III proposal), based on the knowledge we have gained with the VOICE code in this matter (cf. Tasks 1.5.1-1 and 1.5.2-2).

reduced when the flow-induced shearing rate exceeds the linear growth rate of the underlying instability, leading to a transport barrier. The next step is to study the transport of impurities across this barrier, one of the key issues of H-mode plasmas in current and next step tokamaks.

4. Tokamaks are not exactly axi-symmetrical against the toroidal angle. In particular, the finite number of coils results in a so-called “ripple” of the main component of the magnetic field, which increases with the major radius. This 3-dimensional magnetic equilibrium leads to plasma braking which opposes to the possible source of toroidal momentum coming from turbulence. Extensive simulations have allowed us (i) to recover the theoretical predictions of the collisional toroidal viscosity in asymptotic regimes and (ii) to predict poloidal flow reversal at experimentally relevant values of collisionality and ripple magnitude [156]. Finally, we have derived a critical value of ripple amplitude – backed by simulations and consistent with Tore Supra (former tokamak at CEA Cadarache, France) data – above which friction overcomes turbulence so that the toroidal flow is set by collisional processes [157].

T1.5.3-2 Core-edge-SOL interplay in the turbulent regime

Only preliminary analyses were reported in D1.2 regarding the impact of the edge on core turbulence. The issue is the following: fusion performance is governed to a large part by the confinement time of the plasma energy stored in the core confined region. However, the edge conditions are long known experimentally to drastically influence the overall energy confinement time. However, accounting for realistic boundary conditions in core turbulence codes is extremely challenging, both from the HPC point of view and due to the rich and complex physics at the edge. GYSELAX is among the rare codes worldwide capable of addressing this issue.

Since D1.2, we have come to a clear understanding of how the core plasma interacts with the edge and the SOL (Scrape off Layer: region where magnetic field lines intercept solid materials) in the regime where electrons can be assumed adiabatic. Further directions have also been explored:

1. The influence on turbulence and transport of two different boundary conditions have been compared and confronted to experimental measurements: either a poloidally homogeneous SOL where the distribution function is forced to relax towards a Maxwellian with e-folding profiles characteristic of the SOL, or a poloidally localized limiter, Fig.112. The edge region, intermediate between the very core and the SOL, was already reported to be linearly stable in D1.2. In the first case, core turbulence is found to propagate into the edge by a mechanism bearing similarities with the “beach effect” when shallow water turbulence encounters an inclined ground. Alone, however, such an effect reveals insufficient to recover the universal experimental observation that relative density fluctuations sharply increase when approaching the last closed magnetic field surface (the separatrix). Conversely, such a trend is recovered in the limiter case. It results from the inward spreading of the turbulence which develops in the vicinity of the limiter, likely triggered by a parallel velocity gradient instability (of Kelvin-Helmholtz nature) [36]. Of course, such results will need being consolidated when kinetic electrons will be accounted for.

2. So as to prepare this next step, a huge effort has been invested in both the development (enriched physics and optimized numerical and HPC treatment[11]) and the exploitation of the VOICE code, a low dimensional version of GYSELAX also used as test-bed. In particular, plasma-wall interaction is studied along a magnetic field line with kinetic ions and electrons. The technique of immersed boundary conditions, imported from the satellite code Soledge3X (formerly Tokam3X), have allowed one to recover the expected physics of the sheath [107]. Interactions with European experts in simulations in this field have attested that VOICE results are state-of-the-art. Various scans have

[11] Cf. also Tasks 1.5.1-1 and 1.5.2-2

been performed in view of paving the path for the implementation in GYSELAX , where the mesh grid size in the direction parallel to the magnetic field will be orders of magnitude larger. A paper is in preparation [Y. Munschy et al. (2022)].

3. GYSELAX is a first principle code which models turbulence in the experimentally relevant flux-driven regime, without any scale separation between equilibrium and fluctuating quantities. Because of this peculiarity, it is particularly demanding in terms of numerical resources. Conversely, reduced quasilinear (QL) and nonlinear (gradient-driven) models with scale separations are commonly used to interpret experiments and to forecast turbulent transport levels in magnetized plasmas. The improvement of the CPU consumption of GYSELAX within EoCoE-II has allowed one to compare these various approaches. Two distinct regimes of turbulence – either far above threshold or near marginal stability – have been investigated with Boltzmann electrons. The success of reduced models especially hinges on the reproduction of nonlinear fluxes. Good agreement between models is found above threshold whilst reduced models would significantly under predict fluxes near marginality, overlooking mesoscale flow organization and turbulence self-advection. The mismatch appears larger at the edge. Constructive prescriptions whereby to improve reduced models is finally discussed. [59]

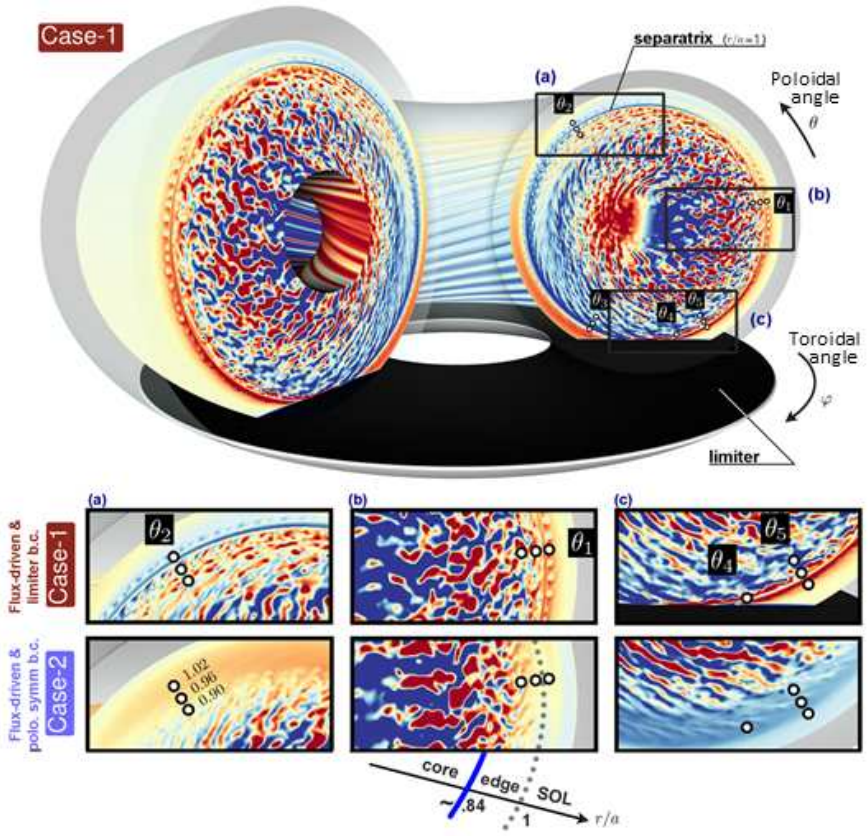


Figure 112: Snapshots of the non axi-symmetric components of the electrostatic potential fluctuations in two different configurations, at statistical equilibrium: case 1 with a limiter, case 2 with a poloidally homogeneous SOL.

7.4 Code demonstrator

GYSELAX

GYSELAX solves the 5-dimensional (3 space coordinates, 2 velocity coordinates) gyrokinetic equation of up to 3 species (main ions, electrons and an impurity) in a tokamak geometry. Self-consistency is achieved by means of the quasi-neutrality and Ampère equations, which couple the electric potential and the parallel component of the vector potential to charge and current densities, respectively. No scale separation is assumed between equilibrium and fluctuating parts of the distribution functions, which are treated on an equal footing.

An original semi-Lagrangian scheme[12] is used, taking advantage of both Lagrangian (one follows particle trajectories, along which distribution functions remain invariant – up to collisions and sources terms) and Eulerian (fixed grid in phase space) approaches. Time evolution is performed with a second-order accuracy predictor-corrector scheme, with a Strang splitting allowing for a separate treatment of several operators which govern the advection in the phase space.

The code is written in Fortran 90 and only uses HDF5 as external library. A hybrid MPI/OpenMP paradigm is used to benefit from a large number of processors while reducing communication costs. Reading and writing the checkpoint/restart and diagnostic files allow one to test the performance of the inputs/outputs on disk. The programming model using MPI/OpenMP is particularly efficient in multi-threaded environments. In addition, the proper use of SIMD vector units reveals important to gain performance on new architectures. In this framework, the recent optimized vectorization of GYSELAX allows the successful exploitation of these vector units.

Scientific simulations towards exascale

Two very large simulations – production runs – are presently running on the EXA1 machine at CEA, France, as part of a lighthouse project (see section 7.3 regarding the addressed physical challenge). These have only been affordable recently thanks to the many numerical improvements made to the code, both in terms of enrichment of the physical model, and of memory and CPU time consumption (cf. D2.3 for details).

They both use a mesh grid of $512 \times 1024 \times 64 \times 128 \times 64 \approx 2.75\,10^{11}$ points in the 5-dimensional phase space and run on 2048 nodes, involving $2048 \times 2 \times 64 = 262\,144$ cores. Per simulation, $60\,000$ iterations are foreseen, amounting to about 40 millions CPU hours and roughly 30 TB of stored data ($\sim 110\,000$ files).

`https://gyselax.github.io/` Demonstrator release v1.0.0-pre

7.5 Summary

The aim in EoCoE-II was to push the flagship 5-dimensional GYSELAX code – with the support of satellite codes – towards the exascale in view of addressing the issue of plasma confinement from the very hot core to the unconfined peripheral region of tokamaks, the so-called scrape-off layer (SOL), in ITER-relevant configurations and parameter ranges.

Major physical upgrades have been implemented and both numerical and HPC bottlenecks have been alleviated, while continuously improving numerical performance. In particular, (i) GYSELAX can now handle plasma-wall interactions with adiabatic electrons; the case of kinetic electrons has been studied with a low-dimensional version of GYSELAX , namely VOICE, as a preliminary step. (ii) Non-circular poloidal cross-sections of the magnetic surfaces can now be considered, together with non axi-symmetric perturbations of the equilibrium magnetic field. The treatment of the X-point has led to dedicated studies – although not yet implemented in GYSELAX – involving satellite codes and on-purpose developments of

[12] Cf. [V. Grandgirard et al., J. Comput. Phys. 217 (2006) 395] and [V. Grandgirard et al., Comput. Phys. Communications, 207 (2016) 35]

reduced models; importantly, this analysis has allowed us to identify the optimal strategy for GYSELAX . (iii) Variations of several orders of magnitude in plasma temperature from core to edge can be handled at an affordable memory cost thanks to the development of non-equidistant splines, and (iv) the implemented electromagnetic effects (Maxwell-Ampère equation) are currently under benchmark.

GYSELAX is now one of the rare gyrokinetic codes worldwide capable of modelling ion turbulence and collisional transport from the core to the SOL in the relevant flux-driven regime of tokamak plasmas. These cutting edge upgrades have led to decisive breakthroughs, as attested most notably by the two highlighted papers published in top rank peer-reviewed scientific journals, namely Physical Review Letters (American Physical Society) and Communications Physics (Nature). In particular, we derive a simple criterion to decide whether edge flows are governed by turbulence or collisions when the axisymmetry of the magnetic configuration in tokamaks is weakly broken, as is usually the case. Last but not least, we predict that pressure inhomogeneities and finite Larmor radius effects are key in the development of large scale flows at the tokamak edge, and ultimately suspected to be critical in the triggering of bifurcated states of enhanced confinement, which are the reference scenarios to achieve the most performant discharges in ITER. These findings have attracted the attention – in view of possibly resolving the misunderstood mismatch of flow measurements with earlier theories – of experimentalists running adequate diagnostics on the Spanish Heliac TJ-II.

References

[1] The alya system - large scale computational mechanics. https://www.bsc.es/es/computer-applications/alya-system.

[2] Ansa pre-processor. https://www.beta-cae.com/ansa.htm.

[3] Y. Al-Hamdani, D. Alfè, and A. Michaelides. How strongly do hydrogen and water molecules stick to carbon nanomaterials? *J. Chem. Phys.*, 146:094701, 2017.

[4] Johan Aqvist. Ion-water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem.*, 94(21):8021–8024, October 1990. Publisher: American Chemical Society.

[5] Steven F. Ashby and Robert D. Falgout. A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations. 124(1):145–159.

[6] Diego Avesani, Andrea Galletti, Sebastiano Piccolroaz, Alberto Bellin, and Bruno Majone. A dual-layer mpi continuous large-scale hydrological model including human systems. *Environmental Modelling & Software*, 139:105003, 2021.

[7] Kilian Bakker, Kirien Whan, Wouter Knap, and Maurice Schmeits. Comparison of statistical post-processing methods for probabilistic NWP forecasts of solar radiation. *Solar Energy*, 191:138–150, October 2019.

[8] R. N. Bannister. A review of operational methods of variational and ensemble-variational data assimilation. *Q.J.R. Meteorol. Soc*, 143(703):607–633, January 2017.

[9] Michael L. Barnes, Claire Welty, and Andrew J. Miller. Global topographic slope enforcement to ensure connectivity and drainage in an urban terrain. 21(4):06015017.

[10] N. H. Batjes. A world dataset of derived soil properties by FAO–UNESCO soil unit for global modelling. *Soil use and management*, 13(1):9–16, 1997.

[11] Peter Bauer, Peter D. Dueben, Torsten Hoefler, Tiago Quintino, Thomas C. Schulthess, and Nils P. Wedi. The digital revolution of earth-system science. *Nature Computational Science*, 1(2):104–113, 2021.

[12] A.D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098–3100, 1988.

[13] K. Berland, V. R. Cooper, K. Lee, E. Schröder, T. Thonhauser, P. Hyldgaard, and B. I. Lundqvist. Van Der Waals Forces in Density Functional Theory: A Review of the VdW-DF MethodVan Der Waals Forces in Density Functional Theory: A Review of the VdW-DF Method. *Rep. Prog. Phys.*, 78:066501, 2015.

[14] J. Berner, G. J. Shutts, M. Leutbecher, and T. N. Palmer. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ecmwf ensemble prediction system. *J. Atmos. Sci.*, 66(3):603–626, March 2009.

[15] A. Biancalani, A. Bottino, C. Ehrlacher, V. Grandgirard, G. Merlo, I. Novikau, Z. Qiu, E. Sonnendrücker, X. Garbet, T. Goerler, S. Leerink, F. Palermo, and D. Zarzoso. Cross-code gyrokinetic verification and benchmark on the linear collisionless dynamics of the geodesic acoustic mode. *Phys. Plasmas*, 24(6):062512, 2017.

[16] C. Bollmeyer, J. D. Keller, C. Ohlwein, S. Wahl, S. Crewell, P. Friederichs, A. Hense, J. Keune, S. Kneifel, and I. Pscheidt. Towards a high-resolution regional reanalysis for the European CORDEX domain. *Quarterly Journal of the Royal Meteorological Society*, 141(686):1–15, 2015.

[17] M. Bonetti, S. Nakamae, B. Tao Huang, T. J. Salez, C. Wiertel-Gasquet, and M. Roger. Thermoelectric energy recovery at ionic-liquid/electrode interface. *J. Chem. Phys.*, 142:244708, 2015.

[18] Daniel Borgis, Sohvi Luukkonen, Luc Belloni, and Guillaume Jeanmairet. Simple Parameter-Free Bridge Functionals for Molecular Density Functional Theory. Application to Hydrophobic Solvation. *J. Phys. Chem. B*, 124(31):6885–6893, August 2020. Publisher: American Chemical Society.

[19] Daniel Borgis, Sohvi Luukkonen, Luc Belloni, and Guillaume Jeanmairet. Accurate prediction of hydration free energies and solvation structures using molecular density functional theory with a simple bridge functional. *J. Chem. Phys.*, 155(2):024117, July 2021. Publisher: American Institute of Physics.

[20] Emily Bourne, Philippe Leleux, Katharina Kormann, Carola Kruse, Virginie Grandgirard, Yaman Güçluü, Martin J. Kühn, Ulrich Rüde, Eric Sonnendruücker, and Edoardo Zoni. Comparison of three solvers for the 2d gyrokinetic poisson equation. *in preparation*, 2022.

[21] Emily Bourne, Yann Munschy, Virginie Grandgirard, Michel Mehrenberger, and Philippe Ghendrih. Non-Uniform Splines for Semi-Lagrangian Kinetic Simulations of the Plasma Sheath. working paper or preprint, August 2022.

[22] Jochen Bröcker and Zied Ben Bouallègue. Stratified rank histograms for ensemble forecast verification under serial dependence. *Quarterly Journal of the Royal Meteorological Society*, 146(729):1976–1990, April 2020.

[23] M. Burkatzki, C. Filippi, and M. Dolg. Energy-consistent pseudopotentials for quantum monte carlo calculations. *The Journal of Chemical Physics*, 126(23):234105, 2007.

[24] Carsten Burstedde, Jose A. Fonseca, and Stefan Kollet. Enhancing speed and scalability of the ParFlow simulation code. 22(1):347–361.

[25] G. Bussi and M. Parrinello. Accurate sampling using Langevin dynamics. *Physical Review E*, 75:056707, 2007.

[26] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45:566, 1980.

[27] Jens Hesselbjerg Christensen and Ole Bøssing Christensen. A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Climatic change*, 81(1):7–30, 2007.

[28] J. W. Connor, S. C. Cowley, R. J. Hastie, T. C. Hender, A. Hood, and T. J. Martin. Tearing modes in toroidal geometry. *Physics of Fluids*, 31(3):577–590, 1988.

[29] P. Courtier, C. Freydier, J.-F. Geleyn, Florence Rabier, and M. Rochas. The arpege project at meteo france. In *Seminar on Numerical Methods in Atmospheric Models, 9-13 September 1991*, volume II, pages 193–232, Shinfield Park, Reading, 1991. ECMWF, ECMWF.

[30] P. Czaja, M. Celino, S. Giusepponi, M. Gusso, and U. Aeberhard. Ab initio study on localization and finite size effects in the structural, electronic, and optical properties of hydrogenated amorphous silicon. *Computational Materials Science*, 155:159–168, 2018.

[31] P. Czaja, S. Giusepponi, M. Gusso, M. Celino, and U. Aeberhard. Computational characterization of a-Si:H/c-Si interfaces. *Journal of Computational Electronics*, 17:1457–1469, 2018.

[32] Aiguo Dai, Gerald A. Meehl, Warren M. Washington, Tom M. L. Wigley, and Julie M. Arblaster. Ensemble simulation of twenty-first century climate changes: Business-as-usual versus co2 stabilization. *Bull. Amer. Meteor. Soc.*, 82(11):2377–2388, November 2001.

[33] Yongjiu Dai, Xubin Zeng, Robert E. Dickinson, Ian Baker, Gordon B. Bonan, Michael G. Bosilovich, A. Scott Denning, Paul A. Dirmeyer, Paul R. Houser, and Guoyue Niu. The common land model. *Bulletin of the American Meteorological Society*, 84(8):1013–1024, 2003.

[34] Jeffrey J. Danielson and Dean B. Gesch. Global multi-resolution terrain elevation data 2010 (GMTED2010). Technical report, US Geological Survey, 2011.

[35] Jacob M. Dean, Samuel W. Coles, William R. Saunders, Andrew R. McCluskey, Matthew J. Wolf, Alison B. Walker, and Benjamin J. Morgan. Overscreening and underscreening in solid-electrolyte grain boundary space-charge layers. *Phys. Rev. Lett.*, 127:135502, Sep 2021.

[36] G. Dif-Pradalier, Ph. Ghendrih, Y. Sarazin, E. Caschera, F. Clairet, Y. Camenen, P. Donnel, X. Garbet, V. Grandgirard, Y. Munschy, L. Vermare, and F. Widmer. Transport barrier onset and edge turbulence shortfall in fusion plasmas. *accepted in Communications Physics*, 2022.

[37] Lu Ding, Maximilien Levesque, Daniel Borgis, and Luc Belloni. Efficient molecular density functional theory using generalized spherical harmonics expansions. *J. Chem. Phys.*, 147(9):094107, September 2017.

[38] Wouter Dorigo, Wolfgang Wagner, Clement Albergel, Franziska Albrecht, Gianpaolo Balsamo, Luca Brocca, Daniel Chung, Martin Ertl, Matthias Forkel, and Alexander Gruber. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, 203:185–215, 2017.

[39] Florian Dupuy. Calibraitno of solar radiation ensemble forecasts using convolutional neural networks. *Solar Energy*, submission pending.

[40] Florian Dupuy, Olivier Mestre, Mathieu Serrurier, Valentin Kivachuk Burdá, Michaël Zamo, Naty Citlali Cabrera-Gutiérrez, Mohamed Chafik Bakkay, Jean-Christophe Jouhaud, Maud-Alix Mader, and Guillaume Oller. ARPEGE cloud cover forecast postprocessing with convolutional neural network. *Weather and Forecasting*, 36(2):567–586, April 2021.

[41] Klaus Duscher, Andreas Günther, Andrea Richts, Patrick Clos, Uta Philipp, and Wilhelm Struckmeier. The GIS layers of the “international hydrogeological map of europe 1:1,500,000” in a vector format. 23(8):1867–1875.

[42] Martin Ehrendorfer. Predicting the uncertainty of numerical weather forecasts: A review. *METEOROLOGISCHE ZEITSCHRIFT-BERLIN-*, 6:147–183, 1997.

[43] Siham El Garroussi, Sophie Ricci, Matthias De Lozzo, Nicole Goutal, and Didier Lucor. Tackling random fields non-linearities with unsupervised clustering of polynomial chaos expansion in latent space: applica tion to global sensitivity analysis of river flooding. 36(3):693–718.

[44] J. Sanz Rodrigo et al. The alex17 diurnal cycles in complex terrain benchmark. *Journal of Physics: Conference Series*, 2021.

[45] P. Giannozzi et al. Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys.: Condens. Matter*, 29:465901, 2017.

[46] P. R. C. Kent et al. QMCPACK: Advances in the development, efficiency, and application of auxiliary field and real-space variational and diffusion Quantum Monte Carlo. *J. Chem. Phys.*, 152:174105, 2020.

[47] Zhang et al. AMReX: A Framework for Block-Structured Adaptive Mesh Refinement. *Journal of Open Source Software*, 4(37):1370, 2019.

[48] R. Evans. The nature of the liquid-napour interface and other topics in the statistical mechanics of non-uniform, classical fluids. *Adv. Phys.*, 28(2):143, 1979.

[49] J. C. Fogarty, H. M. Aktulga, A.Y. Grama, A.C.T. van Duin, and S.A. Pandit. A reactive molecular dynamics simulation of the silica-water interface. *Journal of Chemical Physics*, 132:17470, 2010.

[50] Diana Francis, Ricardo Fonseca, Narendra Nelli, Deniz Bozkurt, Ghislain Picard, and Bin Guan. Atmospheric rivers drive exceptional Saharan dust transport towards Europe. *Atmospheric Research*, 266:105959, March 2022. Publisher: Elsevier BV.

[51] Sebastian Friedemann and Bruno Raffin. An elastic framework for ensemble-based large-scale data assimilation. Technical Report RR-9377, 2020.

[52] Mark A. Friedl, Douglas K. McIver, John CF Hodges, X. Y. Zhang, D. Muchoney, Alan H. Strahler, Curtis E. Woodcock, Sucharita Gopal, Annemarie Schneider, and Amanda Cooper. Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83(1-2):287–302, 2002.

[53] Andrea Galletti, Diego Avesani, Alberto Bellin, and Bruno Majone. Detailed simulation of storage hydropower systems in large alpine watersheds. *Journal of Hydrology*, 603:127125, 2021.

[54] Zhe Gao, Lili Peng, Ping Wang, Jiaqi Dong, and H. Sanuki. Plasma elongation effects on temperature gradient driven instabilities and geodesic acoustic modes. *Nuclear Fusion*, 49(4):045014, mar 2009.

[55] Corbalan J. Garcia M. and Labarta J. LeWI: A runtime balancing algorithm for nested parallelism. *Proceedings of the International Conference on Parallel Processing*, pages 526–533, 2009.

[56] Siham El Garroussi, Sophie Ricci, Matthias De Lozzo, Nicole Goutal, and Didier Lucor. Assessing uncertainties in flood forecasts using a mixture of generalized polynomial chaos expansions. In *Hyper Article en Ligne - Sciences de l'Homme et de la Société*. Hyper Article en Ligne - Sciences de l'Homme et de la Société.

[57] P. Giannozzi and et al. Quantum ESPRESSO: a modular and open-source software project for quantum simulations of materials. *Journal of Physics: Condensed Matter*, 21:395502, 2009.

[58] P. Giannozzi and et al. Advanced capabilities for materials modelling with Quantum ESPRESSO. *Journal of Physics: Condensed Matter*, 29:465901, 2017.

[59] C. Gillot, G. Dif-Pradalier, Y. Sarazin, C. Bourdelle, Y. Camenen, J. Citrin, X. Garbet, Ph. Ghendrih, V. Grandgirard, P. Manas, , and F. Widmer. The problem of marginality in model reductions of turbulence. *under review in Nuclear Fusion*, 2022.

[60] Tom Gleeson, Thorsten Wagener, Petra Döll, Samuel C. Zipper, Charles West, Yoshihide Wada, Richard Taylor, Bridget Scanlon, Rafael Rosolem, Shams Rahman, Nurudeen Oshinlaja, Reed Maxwell, Min-Hui Lo, Hyungjun Kim, Mary Hill, Andreas Hartmann, Graham Fogg, James S. Famiglietti, Agnès Ducharne, Inge de Graaf, Mark Cuthbert, Laura Condon, Etienne Bresciani, and Marc F. P. Bierkens. GMD perspective: The quest to improve the evaluation of groundwater representation in continental- to global-scale models. 14(12):7545–7571.

[61] Tilmann Gneiting and Adrian E. Raftery. Weather forecasting with ensemble methods. *Science*, 310(5746):248, October 2005.

[62] S. Grimme. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.*, 27:1787–1799, 2006.

[63] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.*, 132:154104, 2010.

[64] William J. Gutowski Jr, Filippo Giorgi, Bertrand Timbal, Anne Frigon, Daniela Jacob, Hyun-Suk Kang, Krishnan Raghavan, Boram Lee, Christopher Lennard, and Grigory Nikulin. WCRP coordinated regional downscaling experiment (CORDEX): a diagnostic MIP for CMIP6. 9(11):4087.

[65] Thomas M. Hamill. Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, 129(3):550–560, March 2001.

[66] Jean-Pierre Hansen and I.R. McDonald. *Theory of Simple Liquids, Third Edition*. Academic Press, 3 edition, April 2006.

[67] Hans Hersbach. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Wea. Forecasting*, 15(5):559–570, October 2000.

[68] Jaro Hokkanen, Stefan Kollet, Jiri Kraus, Andreas Herten, Markus Hrywniak, and Dirk Pleiter. Leveraging HPC accelerator architectures with modern techniques — hydrologic modeling on GPUs with ParFlow. 25(5):1579–1590.

[69] Song-You Hong, Jimy Dudhia, and Shu-Hua Chen. A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Monthly Weather Review*, 132(1):103–120, 01 2004.

[70] W.G. Hoover and B.L. Holian. Kinetic moments method for the canonical ensemble distribution. *Physics Letters A*, 211:253–257, 1996.

[71] Lewis A. D. Irvine, Alison B. Walker, and Matthew J. Wolf. Quantifying polaronic effects on the scattering and mobility of charge carriers in lead halide perovskites. *Phys. Rev. B*, 103:L220305, June 2021.

[72] G. Jeanmairet, B. Rotenberg, D. Borgis, and M. Salanne. Study of a water-graphene capacitor with molecular density functional theory. *J. Chem. Phys.*, 151:124111, 2019.

[73] Hans Johansen and Phillip Colella. A cartesian grid embedded boundary method for poisson's equation on irregular domains. *Journal of Computational Physics*, 147(1):60–85, 1998.

[74] Ian T. Jolliffe and David B. Stephenson, editors. *Forecast Verification: A practioner's guide in atmospheric sciences*. Wiley, December 2011.

[75] Jim E. Jones and Carol S. Woodward. Newton–Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems. *Advances in Water Resources*, 24(7):763–774, 2001.

[76] Michael Jung and Ulrich Rüde. Implicit extrapolation methods for variable coefficient problems. *SIAM Journal on Scientific Computing*, 19(4):1109–1124, 1998.

[77] J. Keller, V. Rath, J. Bruckmann, D. Mottaghy, C. Clauser, A. Wolf, R. Seidler, H. M. Bücker, and N. Klitzsch. SHEMAT-Suite: An open-source code for simulating flow, heat and species transport in porous media. *SoftwareX*, 12, 2020.

[78] Stefan J. Kollet and Reed M. Maxwell. Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources*, 29(7):945–958, 2006.

[79] Stefan J. Kollet and Reed M. Maxwell. Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. *Water Resources Research*, 44(2), 2008.

[80] Stefan J. Kollet, Reed M. Maxwell, Carol S. Woodward, Steve Smith, Jan Vanderborght, Harry Vereecken, and Clemens Simmer. Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources. 46(4).

[81] Martin J Kühn, Carola Kruse, and Ulrich Rüde. Implicitly extrapolated geometric multigrid on disk-like domains for the gyrokinetic Poisson equation from fusion plasma applications. *Journal of Scientific Computing*, 91(1):1–27, 2022.

[82] Martin Joachim Kühn, Carola Kruse, and Ulrich Rüde. Energy-minimizing, symmetric discretizations for anisotropic meshes and energy functional extrapolation. *SIAM Journal on Scientific Computing*, 43(4):A2448–A2473, 2021.

[83] Wolfgang Kurtz, Guowei He, Stefan J. Kollet, Reed M. Maxwell, Harry Vereecken, and Harrie-Jan Hendricks Franssen. TerrSysMP–PDAF (version 1.0): a modular high-performance data assimilation framework for an integrated land surface–subsurface model. *Geoscientific Model Development*, 9(4):1341–1360, 2016.

[84] Philippe Lauret, Mathieu David, and Pierre Pinson. Verification of solar irradiance probabilistic forecasts. *Solar Energy*, 194:254–271, December 2019.

[85] C. Lee, W. Yang, and R. G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37(2):785–789, 1988.

[86] S. Lehtola, C. Steigemann, M. J. T. Oliveira, and M. A. L. Marques. Recent developments in libxc — a comprehensive library of functionals for density functional theory. *SoftwareX*, 7:1–5, 2018.

[87] M. Leutbecher and T. N. Palmer. Ensemble forecasting. *Predicting weather, climate and extreme events*, 227(7):3515–3539, March 2008.

[88] Shunlin Liang, Jie Cheng, Kun Jia, Bo Jiang, Qiang Liu, Zhiqiang Xiao, Yunjun Yao, Wenping Yuan, Xiaotong Zhang, Xiang Zhao, and Ji Zhou. The Global Land Surface Satellite (GLASS) Product Suite. *Bulletin of the American Meteorological Society*, 102(2):E323–E337, February 2021.

[89] Shunlin Liang, Xiaotong Zhang, Zhiqiang Xiao, Jie Cheng, Qiang Liu, and Xiang Zhao. *Global LAnd Surface Satellite (GLASS) Products: Algorithms, Validation and Analysis*. Springer Science & Business Media.

[90] C. Lin, F. H. Zong, and D. M. Ceperley. Twist-averaged boundary conditions in continuum quantum monte carlo algorithms. *Phys. Rev. E*, 64:016702, Jun 2001.

[91] Y.-L. Lin, R. D. Farley, and H. D. Orville. Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, 22:519–542, 1983.

[92] P. Loche, C. Ayaz, A. Schlaich, D. J. Bonthuis, and R. R. Netz. Breakdown of linear dielectric theory for the interaction between hydrated ions and graphene. *J. Phys. Chem. Lett.*, 9:6463–6468, 2018.

[93] Stuart Macpherson, Tiarnan A. S. Doherty, Andrew J. Winchester, Sofiia Kosar, Duncan N. Johnstone, Yu-Hsien Chiang, Krzystof Galkowski, Miguel Anaya, Kyle Frohna, Affan N. Iqbal, Satyawan Nagane, Bart Roose, Zahra Andaji-Garmaroudi, Kieran W. P. Orr, Julia E. Parker, Paul A. Midgley, Keshav M. Dani, and Samuel D. Stranks. Local nanoscale phase impurities are degradation sites in halide perovskites. *Nature*, 2022.

[94] D. Malcolm and A. Hansen. Windpact turbine rotor design study. *National Renewable Energy Laboratory, Golden, CO*, 2002.

[95] Gabriel J. Man, Cody M. Sterling, Chinnathambi Kamal, Konstantin A. Simonov, Sebastian Svanström, Joydev Acharya, Fredrik O. L. Johansson, Erika Giangrisostomi, Ruslan Ovsyannikov, Thomas Huthwelker, Sergei M. Butorin, Pabitra K. Nayak, Michael Odelius, and Håkan Rensmo. Electronic coupling between the unoccupied states of the organic and inorganic sublattices of methylammonium lead iodide: A hybrid organic-inorganic perovskite single crystal. *Phys. Rev. B*, 104:L041302, 2021.

[96] A. Marin-Laflèche, M. Haefele, L. Scalfi, A. Coretti, T. Dufils, G. Jeanmairet, S. Reed, A. Serva, R. Berthin, C. Bacon, S. Bonella, B. Rotenberg, P. A. Madden, and M. Salanne. Metalwalls: A classical molecular dynamics software dedicated to the simulation of electrochemical systems. *J. Open Source Softw.*, 5:2373, 2020.

[97] Brecht Martens, Diego G. Miralles, Hans Lievens, Robin van der Schalie, Richard A. M. de Jeu, Diego Fernández-Prieto, Hylke E. Beck, Wouter A. Dorigo, and Niko E. C. Verhoest. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5):1903–1925, May 2017.

[98] N. Marzari and D. Vanderbilt. Maximally localized generalized Wannier functions for composite energy bands. *Physical Review B*, 56:12847, 1997.

[99] R. M. Maxwell, L. E. Condon, and S. J. Kollet. A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. *Geoscientific Model Development*, 8(3):923, 2015.

[100] Reed M. Maxwell. A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling. *Advances in Water Resources*, 53:109–117, 2013.

[101] C. Merlet, C. Péan, B. Rotenberg, P. A. Madden, P. Simon, and M. Salanne. Simulating supercapacitors: Can we model electrodes as constant charge surfaces? *J. Phys. Chem. Lett.*, 4:264–268, 2013.

[102] C. Merlet, B. Rotenberg, P. A. Madden, P.-L. Taberna, P. Simon, Y. Gogotsi, and M. Salanne. On the Molecular Origin of Supercapacitance in Nanoporous Carbon Electrodes. *Nat. Mater.*, 11:306–310, 2012.

[103] C. Merlet, M. Salanne, and B. Rotenberg. New coarse-grained models of imidazolium ionic liquids for bulk and interfacial molecular simulations. *J. Phys. Chem. C*, 116(14):7687–7693, 2012.

[104] C. Merlet, M. Salanne, B. Rotenberg, and P. A. Madden. Imidazolium ionic liquid interfaces with vapor and graphite: Interfacial tension and capacitance from coarse-grained molecular simulations. *J. Phys. Chem. C*, 115:16613–16618, 2011.

[105] E. Mostaani, N. D. Drummon, and V. I. Fal'ko. Quantum monte carlo calculation of the binding energy of bilayer graphene. *Phys. Rev. Lett.*, 115:115501, 2015.

[106] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari. wannier90: A tool for obtaining maximally-localised Wannier functions. *Computer Physics Communications*, 178:685, 2008.

[107] Y. Munschy. Kinetic modelisation and simulation of plasma-wall interaction using voice code. *Internship report (+ paper in preparation), Master 2 GI-PLATO, Paris-Saclay University*, 2021.

[108] M Nakanishi and H. Niino. An improved mellor-yamada level-3 model: its numerical stability and application to a regional prediction of advection fog. *Bound.-Layer Meteor.*, 119:397–407, 2006.

[109] M Nakanishi and H. Niino. Development of an improvedturbulence closure model for the atmospheric boundary layer. *J. Meteor. Soc. Japan*, 87:895–912, 2009.

[110] Bibi S. Naz, Stefan Kollet, Harrie-Jan Hendricks Franssen, Carsten Montzka, and Wolfgang Kurtz. A 3 km spatially and temporally consistent European daily soil moisture reanalysis from 2000 to 2015. *Sci Data*, 7(1):111, December 2020.

[111] NCEP. National centers for environmental prediction/national weather service/noaa/u.s. department of commerce. 2008, updated daily. ncep adp global upper air and surface weather observations (prepbufr format). research data archive at the national center for atmospheric research, computational and information systems laboratory., 2008.

[112] Lars Nerger and Wolfgang Hiller. Software for ensemble-based data assimilation systems—Implementation strategies and scalability. *Computers & Geosciences*, 55:110–118, June 2013.

[113] Philipp Neumann, Peter Düben, Panagiotis Adamidis, Peter Bauer, Matthias Brück, Luis Kornblueh, Daniel Klocke, Bjorn Stevens, Nils Wedi, and Joachim Biercamp. Assessing the scales in numerical weather and climate predictions: will exascale be the rescue? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 377(2142):20180148, April 2019.

[114] Thanh Huy Nguyen, Sophie Ricci, Christophe Fatras, Andrea Piacentini, Anthéa Delmotte, Emer ic Lavergne, and Peter Kettig. Improvement of flood extent representation with remote sensing data and data assimilation. Publication Title: Hyper Article en Ligne - Sciences de l'Homme et de la Société.

[115] J. Niederau, J. Fink, and M. Lauster. Connecting dynamic heat demands of buildings with borehole heat exchanger simulations for realistic monitoring and forecast. *Advances in Geosciences*, 56:45–56, 2021.

[116] K. W. Oleson, G.-Y. Niu, Z.-L. Yang, D. M. Lawrence, P. E. Thornton, P. J. Lawrence, Reto Stöckli, R. E. Dickinson, G. B. Bonan, and Samuel Levis. Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research: Biogeosciences*, 113(G1), 2008.

[117] T. N. Palmer. Predicting uncertainty in forecasts of weather and climate. *Reports on Progress in Physics*, 63(2):71–116, January 2000.

[118] Gilberto Pastorello. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. page 27.

[119] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation made simple. *Physical Review Letters*, 77:3865, 1996.

[120] G. Pireddu, L. Scalfi, and B. Rotenberg. A molecular perspective on induced charges on a metallic surface. *J. Chem. Phys.*, 155:204705, 2021.

[121] S. Plimpton. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics*, 117:1–19, 1995.

[122] G. Psofogiannakis and A. C. T. van Duin. Development of a ReaxFF reactive force field for Si/Ge/H systems and application to atomic hydrogen bombardment of Si, Ge, and SiGe (100) surfaces. *Surface Science*, 646:253–260, 2016.

[123] Z. Qiu, L. Chen, and F. Zonca. Kinetic theory of geodesic acoustic modes in toroidal plasmas: a brief review. *Plasma Science and Technology*, 20(9):094004, 2018.

[124] A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos. Optimized pseudopotentials. *Phys. Rev. B*, 41:1227–1230, 1990.

[125] V. Rath, A. Wolf, and H. M. Bücker. Joint three-dimensional inversion of coupled groundwater flow and heat transfer based on Automatic Differentiation: sensitivity calculation, verification, and synthetic examples. *Geophysical Journal International*, 167(1):453–466, 2006.

[126] S. K. Reed, O. J. Lanning, and P. A. Madden. Electrochemical Interface Between an Ionic Liquid and a Model Metallic Electrode. *J. Chem. Phys.*, 126:084704, 2007.

[127] J. Reinoso and A. Blázquez. Geometrically nonlinear analysis of functionally graded power-based and carbon nanotubes reinforced composites using a fully integrated solid shell element. *Composite Structures*, 152:277–294, 2016.

[128] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. 2015.

[129] D. Roy and M. Maroncelli. An Improved Four-Site Ionic Liquid Model. *J. Phys. Chem. B*, 114:12629–12631, 2010.

[130] Maryam Sajedi, Maxim Krivenkov, Dmitry Marchenko, Jaime Sánchez-Barriga, Anoop K. Chandran, Andrei Varykhalov, Emile D. L. Rienks, Irene Aguilera, Stefan Blügel, and Oliver Rader. Is there a polaron signature in angle-resolved photoemission of $\mathrm{cspbbr_3}$? *Phys. Rev. Lett.*, 128:176405, 2022.

[131] Y Sarazin, G Dif-Pradalier, X Garbet, P Ghendrih, A Berger, C Gillot, V Grandgirard, K Obrejan, R Varennes, L Vermare, and T Cartier-Michaud. Key impact of phase dynamics and diamagnetic drive on reynolds stress in magnetic fusion plasmas. *Plasma Physics and Controlled Fusion*, 63:064007, 2021.

[132] William Robert Saunders, James Grant, and Eike Hermann Mueller. Code and Data Release: Monte Carlo simulations with fast and accurate electrostatics, June 2020. https://doi.org/10.5281/zenodo.3873308.

[133] William Robert Saunders, James Grant, and Eike Hermann Müller. A domain specific language for performance portable molecular dynamics algorithms. *Computer Physics Communications*, 224:119–135, 2018.

[134] William Robert Saunders, James Grant, and Eike Hermann Müller. A new algorithm for electrostatic interactions in monte carlo simulations of charged particles. *Journal of Computational Physics*, 430:110099, 2021.

[135] William Robert Saunders, James Grant, Eike Hermann Müller, and Ian Thompson. Fast electrostatic solvers for kinetic monte carlo simulations. *Journal of Computational Physics*, 410:109379, 2020.

[136] L. Scalfi, D. T. Limmer, A. Coretti, S. Bonella, P. A. Madden, M. Salanne, and B. Rotenberg. Charge fluctuations from molecular simulations in the constant-potential ensemble. *Phys. Chem. Chem. Phys.*, 22:10480–10489, 2020.

[137] Marcel G. Schaap and Feike J. Leij. Database-related accuracy and uncertainty of pedotransfer functions. *Soil Science*, 163(10):765–779, 1998.

[138] Helen Schottenhamml, Ani Anciaux-Sedrakian, Frederic Blondel, Adria Borras-Nadal, Pierre-Antoine Joulin, and Ulrich Rüde. Evaluation of a lattice boltzmann-based wind-turbine actuator line model against a navier-stokes approach. *Journal of Physics: Conference Series*, 2265(2):022027, may 2022.

[139] R. Seidler, K. Padalkina, H. M. Bücker, A. Ebigbo, M. Herty, G. Marquart, and J. Niederau. Optimal experimental design for reservoir property estimates in geothermal exploration. *Computational Geosciences*, 20(2):375–383, 2016.

[140] Ralf Seidler. *Automatisierte Anbindung von Simulations- und Optimierungssoftware zur parallelen Lösung inverser Problemklassen*. PhD thesis, Friedrich-Schiller-Universität Jena, 2020.

[141] H. Shin, J. Kim, H. Lee, O. Heinonen, A. Benali, and Y. Kwon. Nature of interlayer binding and stacking of sp–sp^2 hybridized carbon layers: A quantum monte carlo study. *J. Chem. Theory Comput.*, 13:5639–5646, 2017.

[142] J. I. Siepmann and M. Sprik. Influence of Surface-Topology and Electrostatic Potential on Water Electrode Systems. *J. Chem. Phys.*, 102:511–524, 1995.

[143] Jana Sillmann, Thordis Thorarinsdottir, Noel Keenlyside, Nathalie Schaller, Lisa V. Alexander, Gabriele Hegerl, Sonia I. Seneviratne, Robert Vautard, Xuebin Zhang, and Francis W. Zwiers. Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities. *Weather and Climate Extremes*, 18:65–74, 2017.

[144] Clemens Simmer, Gerhard Adrian, Sarah Jones, Volkmar Wirth, Martin Göber, Cathy Hohenegger, Tijana Janjic, Jan Keller, Christian Ohlwein, and Axel Seifert. Herz: The german hans-ertel centre for weather research. *Bulletin of the American Meteorological Society*, 97(6):1057–1068, 2016.

[145] I. Souza, N. Marzari, and D. Vanderbilt. Maximally localized Wannier functions for entangled energy bands. *Physical Review B*, 65:035109, 2001.

[146] R. A. Street. *Hydrogenated Amorphous Silicon*. Cambridge University Press, Cambridge, 1991.

[147] W.-K. Tao and J. Simpson. The goddard cumulus ensemble model. part i: Model description. *Terr. Atmos. Oceanic Sci.*, 4:35–72, 1993.

[148] MUMPS team. Multifrontal massively parallel solver. http://mumps-solver.org.

[149] the AMReX Development Team, Ann Almgren, Vince Beckner, Johannes Blaschke, Cy Chan, Marcus Day, Brian Friesen, Kevin Gott, Daniel Graves, Axel Huebl, Maximilian Katz, Andrew Myers, Tan Nguyen, Andrew Nonaka, Michele Rosso, Sam Williams, Weiqun Zhang, and Michael Zingale. Amrex-codes/amrex: Amrex 22.06, June 2022.

[150] G. Thompson, R. M. Rasmussen, and K. Manning. Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. part i: Description and sensitivity analysis. *Monthly Weather Review*, 132:519–542, 2004.

[151] T. Thonhauser, V. R. Cooper, S. Li, A. Puzder, P. Hyldgaard, and D. C. Langreth. Van der waals density functional: Self-consistent potential and the nature of the van der waals bondvan der waals density functional: Self-consistent potential and the nature of the van der waals bond. *Phys. Rev. B*, 76:125112, 2007.

[152] T. Thonhauser, S. Zuluaga, C. A. Arter, K. Berland, E. Schröder, and P. Hyldgaard. Spin signature of nonlocal correlation binding in metal-organic frameworks. *Phys. Rev. Lett.*, 115:136402, 2015.

[153] Julien Toulouse and C. J. Umrigar. Optimization of quantum monte carlo wave functions by energy minimization. *The Journal of Chemical Physics*, 126(8):084102, 2007.

[154] F. Valencia, A. H. Romero, F. Ancilotto, and P. L. Silvestrelli. Lithium adsorption on graphite from density functional theory calculations. *J. Phys. Chem. B*, 110:14832–14841, 2006.

[155] van Leeuwen and Peter Jan. Particle filtering in geophysical systems. *Mon. Wea. Rev.*, 137(12):4089–4114, December 2009.

[156] R. Varennes, X. Garbet, L. Vermare, Y. Sarazin, G. Dif-Pradalier, V. Grandgirard, P. Ghendrih, P. Donnel, M. Peret, K. Obrejan, and E. Bourne. Impact of magnetic ripple on neoclassical equilibrium in gyrokinetic simulations. *under review in Plasma Physics and Controlled Fusion*, 2022.

[157] R. Varennes, X. Garbet, L. Vermare, Y. Sarazin, G. Dif-Pradalier, V. Grandgirard, P. Ghendrih, P. Donnel, M. Peret, K. Obrejan, and E. Bourne. Plasma rotation: when turbulent momentum drive opposes magnetic braking. *accepted in Physical Review Letters*, 2022.

[158] György Varga. Changing nature of saharan dust deposition in the carpathian basin (central europe): 40 years of identified north african dust events (1979-2018). *Environment International*, 139:105712, June 2020.

[159] Jürgen Vogt, Pierre Soille, A. De Jager, E. Rimaviciute, W. Mehl, S. Foisneau, K. Bodis, J. Dusart, M. L. Paracchini, and P. Haastrup. A pan-european river and catchment database. 22920:120.

[160] L. Vu-Quoc and X.G. Tan. Optimal solid shells for non-linear analyses of multilayer composites. i. statics. *Computer Methods in Applied Mechanics and Engineering*, 192(9):975–1016, 2003.

[161] T. Werder, J. H. Walther, R. L. Jaffe, T. Halicioglu, and P. Koumoutsakos. On the water-carbon interaction for use in molecular dynamics simulations of graphite and carbon nanotubes. *J. Phys. Chem. B*, 107:1345–1352, 2003.

[162] D.S. Wilks. Forecast verification. In *International Geophysics*, pages 301–394. Elsevier, 2011.

[163] P. Wirnsberger, D. Frenkel, and C. Dellago. An enhanced version of the heat exchange algorithm with excellent energy conservation properties. *J. Chem. Phys.*, 143:124104, 2015.

[164] Edoardo Zoni. *Theoretical and numerical studies of gyrokinetic models for shaped Tokamak plasmas*. PhD thesis, Technische Universität München, 2019.

[165] Edoardo Zoni and Yaman Güçlü. Solving hyperbolic-elliptic problems on singular mapped disk-like domains with the method of characteristics and spline finite elements. *Journal of Computational Physics*, 398:108889, 2019.