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

D1.1

Scientific objectives and roadmap

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.eocoe2.eu
	Deliverable ID:	D1.1
	Deliverable Nature:	Report
	Dissemination Level:	PU*
	Contractual Date of Delivery:	M6 30/06/2019
	Actual Date of Delivery:	16/07/2019
	EC Project Officer:	Andréa Feltrin

* - 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 :	Scientific objectives and roadmap
	ID :	D1.1
	Available at:	http://www.eocoe2.eu
	Software tool:	LaTeX
Authorship	Written by:	Edoardo Di Napoli
	Contributors:	Herbert Owen (BSC), Hendrik Elbern (FZJ), Francesco Buonocore (ENEA), Irene Aguilera (FZJ), Alison Walker (UBAH), Matthiew Wolf (UBAH), Massimo Celino (ENEA), Rafael Mayo-García (CIEMAT), Pablo García Müller (CIEMAT), Marie Cathelain (IFPEN), Daniel Borgis (CEA-MdlS), Stefano Mossa (CEA-INAC), Carlo Pierleoni (CEA-MdlS), Mathieu Salanne (CEA-MdlS), Bibi S. Naz (FZJ), Bruno Majone (UNITN). Alberto Bellin (UNITN), Sophie Ricci (CERFACS), Andrea Piacentini (CERFACS), Siham El Garroussi (CERFACS), N. Goutal (CERFACS), Johanna Bruckmann (RWTH), Yanick Sarazin (CEA-IRFM), Virginie Grandgirard (CEA-IRFM), Philipp Franke (FZJ)
	Reviewed by:	PEC, PBS

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

Contents

List of Figures

List of Tables

1 Executive summary

This report describes the roadmap leading towards the completion of the objectives of the Work Package 1 (WP1) outlined in the proposal of the EoCoE-II project. 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. Such a target will be realized 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. As such each, challenge can be seen as a task labelled 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 focus on a selected number of main tasks whose payload will provide significant advances in its respective domain. The two main tasks of Wind for Energy (i.e. T1.1.1 and T1.1.2) deals with wind turbine modeling, from the full understanding of single rotor dynamics to the prototyping of the inflow across entire wind farms in complex terrain. In Meteorology for Energy there are three main tasks addressing the need for probabilistic forecasting to predict the production of solar and wind parks and their impact on energy trading across the grid. The generation of these probabilistic power forecasts will be achieved by perform ultra-large $\mathcal{O}(1000)$ ensemble calculations. Two of the three tasks of Materials for Energy focus on the design and study of new materials for photovoltaic cells including hetero-interfaces in silicon cells, and both organic and perovskites solar cells. The third focal point of this SC is on modeling the possible exploitation of salinity and temperature gradients for electricity production. The fourth challenge, Water for Energy, takes on continental-scale hydrological simulations in order to enable the management of geothermal and hydropower energy sources. The Fusion for Energy challenge tackles the modeling of self-consistent gyrokinetic plasma in magnetically confined and unconfined regions inside a tokamak. The aim is to simulate plasma turbulence and transport from the core to the edge of complex tokamak magnetic geometries.
After two preliminary sections, the report is structured in five main sections, each describing one of the five SC that are part of WP1. Each of the sections is split into subsections describing the main tasks (e.g. T1.1.1 and T1.1.2 for the Wind for Energy SC) of each challenge and their relative subtasks (e.g. T1.1.1 is divided in 5 subtasks, from T1.1.1-1 to T1.1.1-5). The SC are color coded while their corresponding main tasks inherit the color coding of the challenge and are distinguished by distinct shades of the same color. Every subsection has a rigid structure; a description of all the subtasks followed by a timeline and concluded by a discussion on risks and mitigations. The SC sections are drawn to a close with an outline of the partners and team members involved in each of the substasks that are part of the challenge.

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
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
MdIS:	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
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
Nemo5:	NanoElectronics MOdeling Tools 5
neXGf:	non-equilibrium eXascale Green's functions
OpenFOAM:	Open Source Field Operation and Manipulation
ParFlow:	PARallel Flow
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
ABL:	Atmospheric Boundary Layer
AD:	Automatic Diffentiation
AOT:	Aerosol Optical Thickness
PBE:	Perdew-Burke-Ernzerhof functional
BLYP:	Becke-Lee-Yang-Parr functional
COT:	Cloud Optical Thickness
CLM3.5:	Community Land Model version 3.5
CPU:	Central Processing Units
CSP:	Concentrated Solar Power
DA:	Data Assimilation
DFT:	Density Functional Theory
DMC:	Dynamic Monte Carlo
FSI:	Fluid-Structure Interaction
GPU:	Graphical Processing Unit
HLST:	High Level Support Team
HPC:	High Performance Computing
ITER:	International Thermonuclear Experimental Reactor
KMC:	Kinetic Monte Carlo
LES:	Large Eddy Simulations
MD:	Molecular Dynamics
MPI:	Message Passing Interface
NEGF:	Non-Equilibrium Greens functions
NREL:	National Renewable Energy Laboratory
NWP:	Numerical Weather Prediction
OED:	Optimal Experimental Design
PBC:	Periodic Boundary Conditions
PDAF:	Parallel Data Assimilation Framework
pdf:	probability density functions
PF-CLM:	Parflow-Community Land Model
QMC:	Quantum Monte Carlo
QM:	Quantum Mechanics
SHJ:	Silicon HeteroJunction
SOL:	Scrape-Off Layer
WP:	Work Package

Figure 1: Word cloud of the WP1 deliverable D1.1.

3 Wind for Energy (T1.1)

T1.1 is divided in two main tasks:

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

The goal of T1.1.1. is to improve wall modeling and inflow boundary conditions of the low dissipation Large Eddy Simulation (LES) formulation for Atmospheric Boundary Layer (ABL) flows in the code Alya. The model will include thermal coupling, Coriolis forces, canopy, and the actuator disc. Benchmarking against the SOWFA package will be performed. The model will be tested on real wind farms as part of an ongoing collaboration with Iberdrola.
Within T1.1.2 we will develop a full rotor model where the actual geometry of the wind turbine blades and tower is modeled exactly. A sliding mesh approach will be used to incorporate the rotation of the blades. Their deformation will be considered using a Fluid-Structure Interaction (FSI) approach. The model will be compared against the full rotor model available in the code FLOWer. Comparison between the actuator line model developed by IFPEN and full rotor simulations will also be performed.

3.1 Task T1.1.1

The main task T1.1.1 is subdivided in five subtasks:

T1.1.1-1	**Improve wall modeling for Atmospheric Boundary Layer of Large Eddy Simulation in the code Alya developed at BSC**. Implementation of wall modeling normally used in a finite element setting will be improved to increase its accuracy for high Reynolds numbers typical of ABL flows. Moreover, a new methodology that provides increased robustness for complex geometries shall be implemented. This methodology modifies how the wall law is applied to avoid problems that can appear at sharp edges.
T1.1.1-2	**Improve the inflow boundary conditions for the Atmospheric Boundary Layer of Large Eddy Simulation for the Alya code**. We will evaluate the effect of turbulent inflow boundary conditions on the flow across the wind farm region. Improved inflow boundary conditions will be proposed. These may incorporate data from mesoscale simulations from codes such as WRF.
T1.1.1-3	**Including thermal coupling, Coriolis forces, canopy, and the actuator disc**. The Alya code will be adapted to include the previously mentioned terms in the thermal and Navier Stokes equations for the LES simulations. The results shall be validated for LES problems and tested in complex terrain cases.
T1.1.1-4	**Benchmarking against the SOWFA code**. A couple of selected cases will be run with both SOWFA (by IFPEN), and Alya (by BSC) and a comparison between both codes will be performed to analyze their accuracy for similar computational costs.
T1.1.1-5	**Test the new model against a realistic set up**. Test the model on real wind farms as part of an ongoing collaboration with Iberdrola.

Timeline

For the milestone in month 12 (MS1) we expect to have advanced results for T1.1.1-1 and preliminary results for T1.1.1-2. It is expected that by the second deliverable (D1.2) T1.1.1-1 is completed and partial results are available for tasks T1.1.1-2 and T1.1.1-3. The remaining subtasks will be completed for the final deliverable (D1.3).

WIND	**Project Months**											
T1.1.1		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.1.1-1												
T1.1.1-2												
T1.1.1-3												
T1.1.1-4												
T1.1.1-5												

Risk and mitigation

The validation of the results will depend on the availability of experimental data from other projects, such as NEWA (New European Wind Atlas). The degree of uncertainty in the numerical results obtained with Alya or SOWFA will depend on the degree of uncertainty of the inputs provided by the experimentalists. In other to minimize the uncertainty, we shall try to focus on simple problems where the uncertainty in the inputs is more limited. It is interesting to note that even in simple benchmark problems for complex terrain flows that have been highly accepted by the wind community the uncertainty in the inputs from experimental measurements is still quite high [1].

In order to minimize the impact of such uncertainty, we will collaborate with the European wind energy community to have access to the best possible experimental data.
The relative importance of T1.1.1-1 and T1.1.1-2 will depend on the influence that boundary conditions at the ground and inflow boundary conditions have on the parameters of interest within the wind farm region. Therefore, numerical experimentation will help to guide where to put more effort. We expect that in the next six months a better identification of their relative importance will allow us to concentrate our efforts where they are most needed.
Including thermal coupling, Coriolis forces, canopy and the actuator disc for LES ABL flow over complex terrain using unstructured grids is a relatively unexplored area of research within the wind energy community. Therefore unexpected difficulties might appear. The objective, within the available resources, will be to include all three effects. If this would not be possible, we will adjust our objective by including at least two of them by the end of the project. The interest of Iberdrola will be taken into account when deciding what to prioritize.
When testing the model on real wind farms, we shall be limited in the results that can be published by what Iberdrola allows us to make public. We expect to continue and improve our collaboration with Iberdrola so that we can have access to their data.
From IFPEN side, the main risk is about the size of the domain of the complex terrain and its topography as the meshing is not an easy task in OpenFOAM. A solution could be to work on a simpler terrain (e.g., a hill or an island), or to use another code.
In the timeline, we left some margin, at the end of the project, to deal with the possible difficulties we encounter.

3.2 Task T1.1.2

The main task T1.1.2 is subdivided in five subtasks:

T1.1.2-1	**Develop and test the sliding mesh approach for rotating blades in Alya using rigid blades**. To start testing the methodology, we will use the NREL Phase VI wind turbine, which is a relatively small turbine that has been tested extensively in the literature. The first step will be to test the sliding mesh approach in a problem where only the wind turbine blades are included. Later the tower, and probably the nacelle, shall be included to obtain a more realistic scenario. Further wind turbines, such as Mexico or NREL 5MW, will be explored as the project advances.
T1.1.2-2	**Adapt shall elements to turbine blades**. We will adapt continuum shell elements available in Alya to wind turbine blade simulations and test it in a fixed blade subject to external forces. The results will be validated with data from the literature.
T1.1.2-3	**Merge the results of the previous two subtasks**.The outcome of the previous two subtasks will be used to simulate the rotating wind turbine blades and their deformation using the fluid-structure interaction (FSI) approach available in Alya. The fluid part will be treated with sliding mesh approach from subtask T1.1.2-1 while the solid part will be treated using continuum shell elements.
T1.1.2-4	**Compare Alya with FLOWer**. We plan to compare the model developed in Alya against the full rotor model available in the code FLOWer that is developed at Stuttgart University by Professor Thorsten Lutz and his group. The objective is to take advantage of the know-how from Stuttgart University to improve Alya's model. It is expected that people from BSC will visit Stuttgart University to learn from their experience. German researchers may also visit BSC within this collaboration.

T1.1.2-5	**Comparing models.** We will compare the results from the full rotor model developed in Alya against the actuator line model developed by IFPEN using WaLBerla.

Timeline

For the milestone MS1 we expect to have advanced results for T1.1.2-1 and preliminary results for T1.1.2-2 We expect to complete T1.1.2-1 and T1.1.2-2 before D1.2. The remaining tasks will be completed by D1.3.

WIND	Project Months											
T1.1.2		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.1.2-1												
T1.1.2-2												
T1.1.2-3												
T1.1.2-4												
T1.1.2-5												

Risk and mitigation

T1.1.2 is expected to be more challenging than T1.1.1 since the experience in full rotor simulations in our group is much more limited than in flow over complex terrain. Nevertheless, we expect that the challenge is worth the effort since it could allow us to expand our services to a wider range of industrial customers within the wind sector. We are currently working with Iberdrola (a wind energy producer) on wind resources assessment (related to T1.1.1). Full rotor simulations will enable us to extend our services to wind turbine producers. We have already obtained a support letter for EoCoE-II from Vestas (one of the biggest wind turbine producer), and we expect to develop technologies that might be of their interest. If necessary, additional resources will be devoted to this task.

Our targets are ambitious, and it is not clear if we will be able to fulfill all of them by the end of the project. We have already started working on T1.1.2-1, and we have a methodology that is already working with encouraging results. Some improvements are underway to enhance the robustness, and it is expected that we can shortly test it on the NREL Phase VI wind turbine. We expect to collaborate with Stuttgart University, that is an expert in the field and can provide us with good advise to overcome the difficulties that appear. For the Computational Solids Dynamics part, we do not expect much difficulties, but some might appear since we have not worked with wind turbine blades before. If necessary, additional resources will be devoted to this task.

Coupling the fluid dynamics and solid dynamics together might be the biggest challenge, but we have already obtained good results in the past on a similar problem for the SME Vortex Bladeless, and we hope we can also obtain a satisfactory scheme for Horizontal Axis Wind Turbine. Fluid-Structure Interaction is a hot topic for Alya, and we expect to find synergies with other projects within our department to make it work. The final challenge will be to produce something that can be of interest for wind turbine producers. As mentioned before, we have a support letter from Vestas, and we are also approaching Siemens Gamesa. IFPEN and BSC are two partners that have close contact with industry, and they will make their best effort in this direction.

Subtasks 1.1.2-3 and 1.1.2-4 are probably the riskiest since it is where we have less experience. To compensate for this, we have allocated a portion of our budget to subcontract an expert. We expect it to be Professor Thorsten Lutz from Stuttgart University that is one of the best European experts in the subject, but any other European expert in the subject would also be suitable.
In the timeline, we left some margin, in the end, to deal with the possible difficulties we encounter.

3.3 Partners and resources

BSC will be involved in all subtasks (T1.1.1-1 to T1.1.1-5 and T1.1.2-1 to T1.1.2-5). IFPEN will participate in in T1.1.1-4, T1.1.2-5, and probably in T1.1.1-5. No new hiring is expected by either partner.

Partner	People	Role	Task
BSC	Herbert Owen	Responsible for the Alya team within EoCoE-II	T1.1.1, T1.1.2
BSC	Oriol Lehmkuhl	Leader of the Computational Fluid Dynamics group within BSC	T1.1.1, T1.1.2
BSC	Matias Avila	Wind Energy expert within BSC	T1.1.1
BSC	Guillaume Houzeaux	Alya's main developer	T1.1.2
BSC	Ricard Borrell	Alya's HPC expert	T1.1.2
IFPEN	Marie Cathelain	In charge of the SOWFA runs	T1.1.1-4, T1.1.1-5
IFPEN	Frederic Blondel	In charge of the WaLBerla runs	T1.1.2-5
IFPEN	Pauline Bozonnet	Responsible for the IFPEN team within EoCoE-II	
IFPEN	Ani Anciaux	In charge of the WaLBerla runs	T1.1.2-5

4 Meteorology for Energy (T1.2)

This Science Challenge addresses the need for improved weather predictions, emerging from a growing contribution of weather dependent renewable energy sources to the European electrical power supply, which challenges the existing energy system in many ways. Resilient power grid and power plant management as well as trading at power stock exchanges by sound price fixing are the two most obvious and known concerns to be satisfied under the conditions of volatile and insufficiently predictable wind and solar power. There are also similar needs for optimized operation of Concentrated Solar Power plants (CSP) and other linked power storage systems, which will also be treated.
In practice, the joint practical requirement of the efficient operation of the power system both from economic and managerial viewpoint calls for improved forecasts augmented with uncertainty information in terms of probability density functions (pdf) for parameters of interest like wind, solar insolation, and the related wind and solar power production. Integrating developments from WPs 2,4, and 5, this WP will take advantage by skillful forecasts with error estimates. While today's weather forecasts of leading national and international centres are issued in combination with typically normally distributed uncertainties, unstable weather situations and cloud effects pose notoriously special computational challenges. In particular, the problem of assessing the pdf for cloud and aerosol modulated solar power and unstable weather situations induces exceptionally poor

predictions skills. The computational approach within this project rests on ultra-large ensemble sizes with $\mathcal{O}(1000)$ model runs.
The Flagship codes are ESIAS for running the hyper-ensemble, controlling $\mathcal{O}(1000)$ WRF runs, and EURAD-IM for aerosol and mineral dust simulations. The Satellite codes are Weather Research and Forecast model (WRF), Wind Power Management System (WPMS) and Solar Prediction System (SPS), and ICON. T1.3 is divided in thre 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.

In T1.2.1 we will provide pdfs construction from stochastic integration of the ultra-large ensemble by the ESIAS hyper-ensemble system. WRF-solar (for wind and cloud optical thickness) and EURAD-IM (for aerosol optical thickness) will mainly be used to achieve this but a software analysis for ICON will be produced as well. In this task the focus is placed on cases with poor predictability for both solar and wind power, which remain a coupled problem from the practical viewpoint of volatile and uncertain power source integration. From a meteorological viewpoint, cases of convective instability for cloud formation and dynamic instabilities of both barotropic and baroclinic types for wind evolution are important. Based on developments in EoCoE, the related pdfs are constructed from stochastic integration of the ultra-large ensemble by the ESIAS hyper-ensemble system.
T1.2.2 focuses on power calculation for meteorological ensembles, that is the simulation of the estimated energy obtained from ensemble member results in terms of 100-m winds, Cloud Optical Thickness (COT) and Aerosol Optical Thickness (AOT) at site locations. Envisaged developments include (1) the optimization of the IEE Solar Prediction and Wind Power Management System for varying conditions, e.g. for variable resolution grids, and (2) the confluence of ultra-large ensembles with IEE's satellite-based cloud-tracking system for short-term forecasting. Here, wind vectors at cloud heights can be used to dynamically weight the intraday 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 ensemble member performance. The general output will be pdfs of power production, including inference of higher statistical moments.
T1.2.3 deals with the production and calibration of improved probabilistic wind ensemble predictions with additional focus on downward solar radiation, for photovoltaics and wind velocity for lead times from 3 hours to one day. CNRS-CNRM will exploit the ultra-large ensembles of forecasts in order to produce improved power predictions with additional focus on downward solar radiation, for photovoltaics, and wind velocity (for wind power), for lead times from 3 hours to one day. Meteo France (MF) will work on the development of statistical non-parametric calibration. In addition, methods for the identification of cloud fields along their type of formation.

4.1 Task T1.2.1

The main task T1.2.1 is subdivided in three 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 middle ware. 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.

Timeline

METEO	Project Months											
T1.2.1		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.2.1-1												
T1.2.1-2												
T1.2.1-3												

Risk and mitigation

Due to their proven concept, methods to be applied are expected to pose only marginal to no risks. However, staff recruitment risks are imminent. Due to the leave of the main developper and ensuing recruitment problems to fill the position, a delay of at least 6 months plus a period of vocational adjustment must be addressed, which must be reprocessed. As a consequence the attribution of work will be redistributed by an additional involvement of permanent personnel.

4.2 Task T1.2.2

The main task T1.2.2 is subdivided in two subtasks:

T1.2.2-1	**Power calculation from ensemble winds and COT**. Optimization of the IEE wind power prediction and wind power management system for varying modelling 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 probability density distributions.

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 intraday 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. The general output will be pdfs of power production, including inference of higher statistical moments.

Timeline

METEO	Project Months											
T1.2.2		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.2.2-1												
T1.2.2-2												

Risk and mitigation

Methods to be applied are expected to pose only marginal to no risks. However, as a consequence of staff recruitment risks at the FZJ-IEK-8 institute, delays of input data delivery pose a threat. As a consequence the attribution of work will be redistributed by an additional involvement of personnel. This risk apply to all subtasks. If necessary, NWP data from EoCoE-I could be repurposed to initially test power model calibration and cloud-tracking ensemble validation.

4.3 Task T1.2.3

The main task T1.2.3 is subdivided in 2 subtasks:

T1.2.3-1	**Calibration of improved probabilistic wind ensemble predictions**. CNRS-CNRM will exploit the ultra-large ensembles of forecasts in order to produce improved power predictions and wind velocity (for wind power), for lead times from 3 hours to one day. MF will work on the development of statistical non-parametric calibration. The method of non-parametric calibration will be adapted to 100 m winds at the site locations of wind farms.
T1.2.3-2	**Calibration probabilistic downward solar radiation ensemble predictions**. In this subtask CNRS-CNRM will exploit the ultra-large ensembles of forecasts in order to produce improved power predictions with focus on downward solar radiation, for photovoltaics, for lead times from 3 hours to one day. MF will work on the extension of the method of statistical non-parametric calibration for clouds. In addition, methods for the identification of cloud fields along their type of formation will be included.

Timeline

METEO T1.2.3	Project Months											
		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.2.3-1												
T1.2.3-2												

Risk and mitigation

Methods to be applied are expected to pose only marginal to no risks. However, as a consequence of staff recruitment risks at the FZJ-IEK-8 institute, delays of input data delivery pose a threat. As a consequence the attribution of work will be redistributed by an additional involvement of internal personnel. This risks apply to all subtasks.

4.4 Partners and resources

Partners involved in the Scientific Challenge are FZJ-IEK-8, FRAUNHOFER IEE, CNRS-CNRM (Meteo France)
M6: We expect to recruit 1 postdoc. Recruitment efforts for such a position are currently in progress at FZJ-IEK-8.

Partner	People	Role	Task
FZJ	NN	SC lead meteo for energy	
FZJ	NN	postdoc, main researcher	T1.2.1
FZJ	Philipp Franke	consultant EURAD-IM	T1.2.1
FRAUNHOFER IEE	Garrett Good	main researcher	T1.2.2
CNRS-CNRM	Matthieu Plu	main researcher	T1.2.3

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 highlights the scientific objectives and roadmap for 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. The purpose of the present task is to understand the transport mechanisms underlying photovoltaic devices based on SHJ technology by simulating at atomistic resolution amorphous-crystalline heterointerfaces. Medium and large c-Si/a-Si:H interface models will be build up from classic molecular dynamics (MD) simulations and first-principles calculations. Ab initio electronic properties of the c-Si/a-Si:H interfaces will be calculated. Starting from the first-principles calculations, tight-binding Hamiltonians will be represented in a basis of localized Wannier functions.

Next, non-equilibrium Green's functions (NEGF) formalism will be used to analyse the effect of interfaces on the carrier transport and dynamics in silicon solar cells. Electron-photon and electron-phonon scattering processes will be taken into account.
T1.3.2 focuses on optimizing capacitive blue energy electrodes and thermo-electrochemical devices. Electric power production from salinity gradients harvests the free energy lost during the mixing of river with sea water in estuaries. The main technologies developed for this purpose to date exploit the electric potential differences applied across membranes, but another approach based on capacitive mixing was recently proposed. The first objective of this project will be to ascertain the best electrode structure which optimizes such a blue energy production. Thermo-electrochemical devices employ the variation of the redox potential of an active species with temperature to convert a gradient into electricity. Ionic liquids were recently proposed as optimal media for performing such an energy harvesting, and the second objective of this task will be to find compositions that will enable optimal performances. In both cases, a fundamental understanding of the cation and anion adsorption at the surface of the electrodes is essential. The challenge for this task is that simulating the interfaces requires the rigorous accounting for the interactions between the atoms of the electrodes 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 aim at developing new force fields for classical molecular simulations. The parameterization of these force fields can be made based on a series of electronic DFT calculations. However, it was shown recently that the commonly used exchange-correlation functionals may yield very different results for the adsorption energy of the molecules. We will overcome this problem by performing a series of Quantum Monte Carlo (QMC) reference calculations in order to benchmark them on the adsorption energies. Once the DFT functional is benchmarked on the QMC reference, a large amount of calculations will be performed to fine-tune force fields for classical molecular simulations with Metalwalls/MDFT codes. These two codes aim at simulating electrochemical systems with explicit electrodes, using either molecular dynamics or classical density functional theory to sample the configurational space of the solvent.
T1.3.3 deals with the development of a flexible and modular scheme for the multiscale modeling of electronic and ionic transport in materials for next generation photovoltaic devices. This will be built on (augmented versions of) pre-existing, MPI parallelised Python frameworks, namely Firedrake and PPMD. The scientific goals of this project, as stated in the EoCoE proposal, are:

- Simulate organic photovoltaic cells of 10 nm size and study interfaces on the nm length scale to refine models of charge generation and recombination (Kinetic Monte Carlo, KMC code).
- Understand the complex processes of charge transport in a perovskite solar cell thanks to the implementation of a semiclassical approach based on solving the Boltzmann transport equation in submicron inorganic semiconductors (Device Monte Carlo, DMC code). Both KMC and DMC codes are exascale flagship codes are part of tasks in Work Package 3 (WP3) of EoCoE-II.

5.1 Task T1.3.1

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

T1.3.1-1	**Classical Molecular Dynamics (MD) simulations of c-Si/a-Si:H interface**. By means of the ReaxFF (Reactive Force Field) approach which casts the empirical interatomic potential within a bond-order formalism, therefore implicitly describing chemical bonding without expensive quantum mechanics (QM) calculations, the thermalization, quenching and equilibration processes involving thousands of atoms interactions during hundreds of pico-seconds will be efficiently simulated in few hours of CPU time. Snapshots of the equilibrated c-Si/a-Si:H interface atom configurations at room temperature will be provided for T1.3.1-2.
T1.3.1-2	**First-principles electronic properties of c-Si/a-Si:H interface**. The electronic properties of the c-Si/a-Si:H interface are analyzed by first-principles using the PWscf code of the Quantum Espresso suite. The models with periodic boundary conditions (PBC), meant to mimic an infinitely extended system for medium and large interfaces, will be built with the aid of the classical MD simulations (T1.3.1-1) and refined within ab initio approach. A workflow will be used for the analysis of snapshots of MD simulations. Ab initio optimized coordinates, band structure, total and projected density of states, charge density, and total potential will be produced as input of T1.3.1-3.
T1.3.1-3	**Non-equilibrium Green's functions transport properties c-Si/a-Si:H interface**. To use the results from T1.3.1-2 to analyze the effect of interfaces on the carrier transport and dynamics in silicon solar cells, the NEGF formalism will be applied. To this end, the PVnegf code will be first adapted to treat multi-band Hamiltonians. 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 needed for these calculations is a tight-binding Hamiltonian that reproduces the ab initio electronic structure of the heterostructure of T1.3.1-2. This Hamiltonian will be represented in a 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 evolution of the PVnegf code to a multi-band code will need to be integrated in the fully parallelized and optimized neXGf code (WP2, Task 2.4), which explains the late appearance of some parts of this subtask in the timeline below.

Timeline

MATERIALS	Project Months											
T1.3.1		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.3.1-1			■	■	■	■						
T1.3.1-2				■	■	■	■		■			
T1.3.1-3			■	■	■	■			■	■		

For the milestone MS1 we will have preliminary results and resizing/reorienting of scientific objectives. Early check on feasibility of scientific results at scale will be carried out. First scientific results and progresses on exascale tools are expected by the deliverable D1.2. Final scientific results and exascale tools delivery will be accomplished for the final deliverable D1.3.

Risk and mitigation

T1.3.1-1: Although the ReaxFF potential for hydrogenated Silicon that will be used for the classical MD simulations [2] has been developed by fitting against a training set of DFT data of several Si-H bonding environments, a potential risk exits that will not be able to fully resemble a 'first principles' model. Therefore, a DFT accuracy check should be applied to the atomic configurations of the deliverables of T1.3.1-1.
T1.3.1-2: Optimization of hetero-interfaces models with more than one-thousand atoms could be limited to a few cycles of optimization because the number of electron becomes very large. The mitigation path is to use, only for geometry optimization, alternative DFT codes such as CP2K.
T1.3.1-3: The validation of the ballistic version of the multi-band code can be achieved by comparing results against simulations from existing codes such as Nemo5 or WanT. The validation of the fully non-ballistic multi-band code is, instead, risky and its delivery might be subject to delays. To the best of our knowledge, there is no other existing multi-band NEGF code with the rich physics of the one we plan here. This is because of the complexity of the scattering events (electron-photon and electron-phonon) involved in recombination processes. The mitigation path could be to use the parallel 1-band neXGf code, employing all the available information extracted from the accurate DFT simulations of T1.3.1-2: effective masses, band gaps, band offsets, position of defect/interface states, lattice vibrations modes, dielectric functions, etc.

5.2 Task T1.3.2

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.

Timeline

MATERIALS	Project Months											
T1.3.2		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.3.2-1												
T1.3.2-2												
T1.3.2-3												
T1.3.2-4												

For the milestone MS1 we will have preliminary results on T1.3.2-1 and T1.3.2-2. These preliminary results will be necessary for resizing/reorienting of scientific objectives. Early check on feasibility of scientific results at scale will be carried out. Note that T1.3.2-2 can start only 6 months after the start of T1.3.2-1 since the results will be obtained for the first systems at that point. First scientific results and progresses on exascale tools are expected by the deliverable D1.2. Final scientific results and exascale tools delivery will be accomplished for the final deliverable D1.3.

Risk and mitigation

T1.3.2-1: The main risk in this task concerns the convergence of the QMC calculations. The system may be really large so that large computer resources will be necessary. This will be mitigated by simulating the coronene molecule, which is a simplified version of the graphene plane with similar electronic properties.
T1.3.2-2: The main risk of this task is to have a poor fitting of the reference data set. In such a case we can increase the level of complexity of the force fields, by using for example the aspherical ion model. In such a case the drawback will be that the simulations will become somewhat more expensive.
T1.3.2-3: The MDFT method is not yet able to treat explicitly ionic solutions. Two solutions can be used: include explicitly the ions in the system or use molecular dynamics instead. In both cases the amount of studied systems will have to be reduced.
T1.3.2-4: There is no example in the literature of determination of Seebeck coefficient. Nevertheless there is no conceptual difficulty, since electron transfer reactions are now well handled using classical

molecular dynamics.

5.3 Task T1.3.3

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

T1.3.3-1	**Development of a drift-diffusion simulator for ion motion in devices using the Firedrake Python framework**. Firedrake [3] is an MPI parallelised Python framework for solving partial differential equation using the finite element method. It is developed principally by researchers at Imperial College London, but is used extensively by colleagues in the Mathematics Department at Bath. This task will require a phase of training on the use of Firedrake, and a second phase implementing the model developed in collaboration with researchers at the University of Southampton within the same framework [4].
T1.3.3-2	**Development of a Boltzmann Transport simulator for charge carriers in devices built on the Performance Portable Molecular Dynamics (PPMD) framework**. PPMD [5] is an MPI parallelised Python framework developed by colleagues in the Mathematics Department at Bath. This task needs an initial period of training on the use of the PPMD, followed by translating a code that he has written in C for solving the Boltzmann Transport Equation for charge carriers using a Monte Carlo technique to use the PPMD framework. This will require augmentation of PPMD to treat stochastic scattering events due to electron-lattice interactions (i.e. defects and phonons) in addition to the deterministic evolution under the influence of electrostatic fields, which has been(already implemented.
T1.3.3-3	**Coupling of the two simulators, along with a pre-existing KMC simulator built on PPMD**. In addition to the two simulators that will be developed in the previous two sub-tasks, we will also make use of Coulomb KMC code [6], which is an MPI parallelised Python framework for carrying out KMC simulations including full long ranged electrostatic interactions, developed by the same colleagues who developed PPMD. Coupling of the three simulators is complicated by the different timescales of motion of the two types of particle (ions and electrons). Therefore we will use a dual timestep approach, in which for each long ionic timestep, the electrons will be evolved to steady state using many shorter timesteps which sum to a fraction of an ionic time step.
T1.3.3-4	**Application to the simulation of mixed electron-ion conduction in halide perovskites, comparison of the results to previous all drift-diffusion simulations**. We will use the coupled simulator to address the phenomenon of long lived transients in J-V curves of perovskite solar cells, including what the community has dubbed "hysteresis". We will compare our results with those in the literature that have been obtained using drift-diffusion modelling of both the electrons and ions.

Timeline

MATERIALS	Project Months											
T1.3.3		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.3.3-1												
T1.3.3-2												
T1.3.3-3												
T1.3.3-4												

For the milestone MS1 we will have preliminary results on T1.3.3-1, that will be used for resizing/reorienting of scientific objectives. Early check on feasibility of scientific results at scale will be carried out. First scientific results and progresses on exascale tools from T1.3.3-2 are expected by the deliverable D1.2. Final scientific results and exascale tools delivery will be accomplished for the deliverable D1.3.

Risk and mitigation

We consider the development of the code to be low risk: The frameworks upon which it will be built are well established and tested, and the theory is also well established. Furthermore, the risks are further mitigated by our contact with colleagues in the Mathematics department who are familiar with implementing codes within these frameworks.
Regarding the applications, the perovskite field is highly competitive and fast moving, and therefore there is a chance that it might have moved on from studying "hysteresis" by the time we are in a position to apply our code. However, since the underlying phenomenon of coupled electron and ion motion are a ubiquitous and seemingly unavoidable property of these materials, we are confident that there will continue to be relevant problems that can be addressed with our code. The organics field is less competitive, and in any case we are already, or will be very soon, able to deploy the KMC code to relevant problems in this area.

5.4 Partners and resources

T1.3.1: The ENEA team will build the models of c-Si/a-Si:H interface. S. Giusepponi (ENEA) and the resource to be hired will optimize the model by first principles. F. Buonocore (ENEA) will take care of post-processing of the ab initio charge density. M. Celino (ENEA) will coordinate the first-principles calculations of ENEA team. The CIEMAT team will carry out the MD simulations of the c-Si/a-Si:H interface and is composed of P. García Müller, who is in charge of designing and performing the MD calculations, and R. Mayo-García, who supports the former. Irene Aguilera (FZJ) will derive the methods to generalize the PVnegf code to the multi-band case and will coordinate the interface between the first-principles calculations performed by ENEA and the refactored neXGf code developed by FZJ in WP2. Irene Aguilera has been hired by FZJ for 24PMs. One Ph.D. resource of ENEA will be deployed on month 13.
T1.3.2: Daniel Borgis (CEA-MdlS) will participate to the MDFT calculations. Carlo Pierleoni (CEA-MdlS) will participate to the QMC and DFT calculations. Stefano Mossa (CEA-INAC) will participate to the Metalwalls calculations on thermo-electrochemical systems. Mathieu Salanne (CEA-MdlS) will participate to the force-fitting and the Metalwalls calculations. Two post-docs will be hired: one at month 12 for T1.3.2-1, one at month 18 for T1.3.2-4. The first post-doc will work on the QMC and DFT calculations as well as the force-fitting tasks. The second post-doc will participate to the tasks involving Metalwalls.
T1.3.3: One post-doctoral researcher (Matthew Wolf) for 30 PMs. Matthew Wolf will work on all sub tasks. He will be in contact with colleagues in the Mathematics Department at Bath, who are

familiar with Firedrake, and who are involved in the development of PPMD. Namely Drs. Eike Muller, and Will Saunders. These tasks will be coordinated by Alison B. Walker.

Partner	People	Role	Task
ENEA	Francesco Buono-core	Permanent staff (Physics)	T1.3.1
ENEA	Massimo Celino	Permanent staff (Physics)	T1.3.1
ENEA	Simone Giusepponi	Expert (Physics)	T1.3.1
FZJ	Irene Aguilera	Expert (Physics)	T1.3.1
CIEMAT	Pablo García Müller	Researcher (MD simulations)	T1.3.1
CIEMAT	Rafael Mayo-García	Liaison with EERA JPs related to materials	T1.3.1
CEA-MdlS	Daniele Borgis	Permanent Staff (Physics)	T1.3.2
CEA-MdlS	Carlo Pierleoni	Expert (Physics)	T1.3.2
CEA-INAC	Stefano Mossa	Permanent Staff (Physics)	T1.3.2
CEA-MdlS	Mathieu Salanne	Permanent Staff (Physics)	T1.3.2
UBAH	Alison Walker	Permanent Staff (Physics)	T1.3.3
UBAH	Matthew Wolf	Research Associate (Physics)	T1.3.3

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

In T1.4.1, the Parflow model [7] will be setup for hydropower simulations over the European continent which will provide the current hydrologic status of the terrestrial system and predictions of all pertinent states and fluxes relevant to the energy sector. The objective is to improve the continental-scale hydrologic simulations at higher spatial and temporal resolution.

T1.4.2 aims at simulating hydropower generation in a meso-scale watershed by coupling information on main hydropower systems over the Italian Alpine region and runoff time series produced by hyper-resolved hydrological model (T1.4.1). The coupling with a model accounting explicitly for spatial and temporal varying water usage and treating storage capacities and diversions within a river basin, is indeed a relevant topic in the hydrologic field which contributes directly to the first Milestone (MS1) of the project. The added value of these simulations will be shown by feeding coupled ParFlow, and the Community Land Model (PF-CLM) [8] gridded runoff time series into the operational hydropower model that will be specifically developed over the Italian Alpine region by means of HYPEstreamHS hydrological model.

T1.4.3 deals with forecast floods and resources management. Water inflow and river characteristics are uncertain inputs to the numerical models which can be improved by reducing uncertainties using ensemble-based data assimilation algorithms such as Ensemble Kalman filter. However, full 2D simulations are computationally expensive for large catchments and the cost of the ensemble integration that allows to propagate uncertainty from the input space to the output space is often too expensive for use in an security or productive operational context. Thus, it is convenient to replace the numerical model by an approximate and less expensive surrogate model built from a learning data base. The scientific challenge in T1.4.3 is to reduce the dimension of the large input field (discretized in space and/or time) to allow for the formulation of the surrogate model. In this task a surrogate model will be formulated and exploited to carry out a global sensitivity analysis that leads to classifying sources of uncertainty at a reasonable computational cost.

T1.4.4 deals with experimental design for geothermal modelling. Optimal experimental design (OED) is a mathematical optimization method [9]. The general approach is to find optimal experimental conditions for constraining model parameters, where optimal means to predict parameters with minimal uncertainty. In a former research project [10], this approach was applied to geothermal reservoir models to seek the optimal position of a borehole for temperature measurements in order to estimate permeability with minimal uncertainty. In this approach, permeability is estimated from temperature data by deterministic inversion. Both, the geothermal forward model as well as the inverse model are computed with our in-house code SHEMAT-Suite. The answer to the question of where to place a borehole for temperature measurements in order to invert permeability with minimum uncertainty depends on the sensitivity of the model with respect to the unknown parameters. This sensitivity is mathematically described by the Fisher Matrix. It contains the first order derivative of the model output towards the parameters. SHEMAT-Suite uses automatic differentiation (AD) for calculating first order derivatives which is a fast and flexible method for computing the exact derivative [11]. For evaluating the information contained in the Fisher matrix, OED criteria are formulated, for example the D-optimal design criterion which is based on the determinant of the Fisher Matrix. This whole "OED workflow" of geothermal forward modelling, inversion, computation of the Fisher matrix and of the optimal design criterion can be executed with the software framework EFCOSS: Environment for Combining Optimization and Simulation Software [12]. EFCOSS controls and combines the execution of SHEMAT-Suite and the optimization algorithm. EFCOSS is developed at the Chair of Advanced Computing, FSU Jena (group around Professor H. Martin Bücker). The coupling to SHEMAT-Suite was implemented during the MeProRiskII project and OED has been applied successfully to two- and three-dimensional geothermal reservoir models [13], [14]. However, one identified constraint of this method is its relatively high computational cost. This high computational cost limited this approach to a numerical model with significantly reduced number of unknowns. 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 this existing 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).

In Task 1.4.5 we carry out the application of optimal experimental design (OED) to a real world geothermal reservoir system model. The aim is to predict optimal borehole locations for temperature measurements in order to estimate rock parameters such as permeability or thermal conductivity with least uncertainty. The quality of geothermal reservoir models and hydrothermal flow simulations highly depends on the quality of the subsurface parameterization. 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). Additionally, drilling boreholes during exploration and development of geothermal reservoirs not only involves high costs, but also bears significant risks of failure. OED can substantially improve the decision making process and reduce the uncertainty in parameter estimation. An OED workflow for geothermal reservoir modelling will be further developed and optimized in T1.4.4. A description of the OED workflow and of T1.4.4 can be found in the deliverable for this respective task. One goal of it is to optimize the computational performance of the OED workflow for making its application computationally feasible for detailed numerical models of geothermal reservoir systems with high spatial resolution. On the one hand, T1.4.5 aims at demonstrating the feasible and successful applicability of the OED workflow developed in T1.4.4 to large-scale geothermal reservoir models. On the other hand, it will provide new insights for the simulated geothermal reservoir system. The large historical data available in Tuscany, Italy, coupled to the OED workflow developed in T1.4.4 will allow the setup of a detailed reservoir model of a geothermal reservoir system and assess the optimal location of an exploration borehole. This reservoir system is in production for several decades, providing useful data for model calibration. Information on specific heat flow and temperature from boreholes will be used for calibrating the detailed reservoir model. All available information on reservoir permeability will be used for setting up reasonable distributions of potential permeability distributions in the reservoir system. We will apply the OED workflow for assessing the optimal location of an additional borehole for estimating permeability with the minimum uncertainty and thus reducing overall spatial uncertainty of permeability in the remaining reservoir system.

6.1 Task T1.4.1

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

T1.4.1-1	**Model Setup for hyper-resolution hydrologic simulations**. Hydrological simulations will be performed using the integrated Terrestrial Systems Modeling Platform (TerrSysMP; [15]), consisting of the three-dimensional surface-subsurface model ParFlow and the Community Land Model CLM3.5 (CLM). In EoCoE-I, PF-CLM has been setup at a spatial resolution of 3km over Europe [16]. The meteorological forcing variables (variables such as temperature, precipitation, wind, vapor pressure, and downward longwave and shortwave radiation) will be used as input to the PF-CLM model for the time period of the spinup and subsequent years simulated. In order to obtain an adequate representation of the natural variability, a higher spatial resolution is required in the vertical and lateral directions respectively. We will increase the resolution of the model from 3km to 1km, first over smaller scale (i.e. Alpine region) and then will be implemented over Europe.
T1.4.1-2	**Model Evaluation and validation**. In this task, model simulations will be performed using PF-CLM and then validated the model outputs with observations to identify sources of uncertainties. In this task, the validation process will be performed in two stages: (1) conduct two experiments with differing subsurface geology and compare the results in order to show that our geological improvements lead to an improved continental groundwater flow, (2) compare the observed streamflow data to the simulated streamflow and water table depths. Additionally, lateral groundwater flow at coastal locations will be compared with observations of submarine discharge to demonstrate the advantage of adding alluvial aquifers to the model geology.

T1.4.1-3	**Uncertainty Quantification using Data Assimilation**. To provide reliable hydrologic data and information for management of water resources, we will use data assimilation (DA) techniques which combine observations and model to improve model estimates. In EoCoE-I, a series of data assimilation experiments were conducted using Parallel Data Assimilation Framework (PDAF) that has been merged into CLM3.5 and ParFlow models. We incorporated the satellite-based soil moisture observation into CLM3.5 model to improve soil moisture predictions at 3km resolution over Europe. We will extend this setup for longer time scale to generate a high resolution reanalysis of model states and fluxes. This will allow to quantify uncertainties and also to improve model predictions by updating simultaneously model states and parameters utilizing observational information. Three long-term model simulations will be performed for the European domain. In addition to the above model runs, several experiments will be performed to conduct sensitivity analysis to quantify uncertainties in model parameterization with 40 ensemble members. Previously we used 20 ensemble members to characterize uncertainties in precipitation and soil properties. Previous studies show that increase in ensemble members lead to improved model simulations. However, the computational burden can increase linearly with an increasing number of ensembles in the data assimilation system. To improve model computational efficiency, WP5 (Ensemble Runs;T5.1 and T5.4) will develop a new version of the PDAF, which will be tested with increasing number of ensembles in comparison to the current version of PDAF.

Timeline

For T4.1.1-1, the meteorological forcings has already been obtained and post processed for the time period of 2000-2015. This forcing dataset will be used to drive the PF-CLM model and deliver early results for the first Milestone (MS1). For T4.1.1-2, the validation data, such as soil moisture, evapotranspiration, total water storage from satellite-based observation and discharge information at various gauge stations and measured ground water table depths across Europe have been retrieved and pre-processed. Using these validation datasets, PF-CLM coupled model simulations will be assessed to identify sources of uncertainties by the second deliverable (D1.2).

WATER	Project Months											
T1.4.1		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.4.1-1												
T1.4.1-2												
T1.4.1-3												

Risk and mitigation

In Task 1.4.1-1, simulation of PF-CLM over Europe at 1 km is highly dependent on the modernization of PF code activities (development of new linear algebra solvers for parflow and refactoring and optimization of the Parflow routines where much of the compute time is invested) in other technical work packages ((T3.2.1 in WP3 and T2.5.1 in WP2) to make the PF code more computationally efficient. Implementation of the PF-CLM on a smaller scale (for example only at Alpine region) at higher resolution could be one option to mitigate the risk.

6.2 Task T1.4.2

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

T1.4.2-1	**Refactoring of HYPERstreamHS model**. Refactoring of the HYPERstreamHS source code is needed in order to apply techniques for fast computations by distributing wokload by an MPI interface. Such modifications will allow both to assess the uncertainty related to hydrological predictions (both parametric and epistemic) and to decrease substantially computational time in the view of the meso-scale application (see T1.4.2-4). HYPERstreamHS will also be modified in order to ease the coupling with PF-CLM outputs. This SubTask will contribute Deliverable D1.2 due at M18 of the project.
T1.4.2-2	**Set-up of the model in the Adige river basin**. In this SubTask HYPERstreamHS will be tested in terms of capabilities to reproduce historical streamflow data and water management in large storage reservoirs. In particular, detailed information available in the Adige river basin (12000 km^2, south-eastern Alpine region) on existing hydropower plants including their location, installed capacity, type of power generation, level-storage relationships and spillways geometry, production data, and reservoir levels will be used to simulate hydropower generation through the use of plant-specific models. Historical hydropower generation data will be also used to validate the hydropower generation model in the Adige.
T1.4.2-3	**Hydrological benchmarking exercise in the Adige river basin**. In collaboration with theFZJ-IBG-3 partner, we plan to perform an hydrological benchmarking exercise in the Adige river basin to identify the most suitable modelling framework to be adopted for the simulations in the meso-scale watershed. In particular we plan to evaluate suitable modifications of the CLM3.5 by introducing a simple module dealing with deep infiltration to aquifers and return flow, and to replace the grid based routing scheme with the multi-scale grid independent scheme embedded in the HYPERstreamHS hydrological model. The performance of a calibrated HYPERstreamHS stand-alone model in reproducing observed streamflow will be also evaluated to provide a benchmark for assessing the effectiveness of the introduced parameterizations.
T1.4.2-4	**Set-up of the model in the Italian Alpine region**. The modeling framework identified in SubTask 1.4.2-3 will then be extended to the entire investigation region, by including the representation of storage reservoirs and hydropower plants located therein. Historical hydropower generation data available in the meso-scale watershed will be used to validate the hydropower generation model at such spatial scale.

Timeline

Refactoring of HYPERstreamHS (T1.4.2-1) is currently ongoing and such activity is expected to contribute to MS1. First results concerning the set-up of the hydropower model in the Adige river basin (T4.1.1-2) will contribute to Deliverable D1.2. Results for T.1.4.2-3 and T.1.4.2-3 will be presented on the final Deliverable D1.3.

WATER	Project Months											
T1.4.2		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.4.2-1												
T1.4.2-2												
T1.4.2-3												
T1.4.2-4												

Risk and mitigation

T1.4.2-4: Simulations of hydropower production in the entire Italian Alpine region is critically connected to the collection of reliable information on existing large storage reservoirs and connected hydropower plants (i.e. location of penstock and water diversion channels, installed capacity, type of power generation, level-storage relationships and spillways geometry, production data, reservoir levels, etc.). This collection strictly depends on the willingness to sharing data from hydropower companies managing the plants in the investigated area, or from the possibility to retrieve such information from public available repositories or hydrological plans from river basin authorities. To mitigate such risk, a reduction of the investigated area can be decided in case of lack of data in one or more regions.

6.3 Task T1.4.3

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

T1.4.3-1	**Set up of hydrodynamic surrogate model over Gironde estuary or Garonne river in France**. The modeling framework proposed in this subtask will be carried out on a real test case for the Gironde estuary where Blayais nuclear power plant is located and where security is at stake in case of flooding. In case the sources of uncertainties linked to atmospheric forcing on the estuarine area are too difficult to take into account (large input space with multivariate temporal and spatial correlations), the test case catchment will be switched and the study will be carried out on the Garonne river (Tonneins-La Réole) where no atmospheric forcing is required. The developments will be achieved with the hydrodynamics solver TELEMAC (OpenSource software developed by EDF) using the OpenTURNS UQ dedicated library (Open Source software developed by EDF, AirbusGroup, Phimeca).
T1.4.3-2	**Sensitivity analysis using the surrogate model**. The surrogate model implemented in the T1.4.3-1 will be first used to carry out a global sensitivity analysis to classifying sources of uncertainty to explain water level variance. The use of a surrogate model based on Galerkin projection (Polynomial Chaos Expansion) will allow us to identify and quantify uncertainties in hydrologic forcing, friction coefficient and river geometry for stationary flows. The non-stationary (time varying) flow will also be investigated though some discontinuities in the surface response are to be expected. Such an occurrence may complicate the surrogate construction and could require more advanced solutions that are currently under investigation as part of a PhD project (S. El Garroussi), funded by CERFACS and Région Midi-Pyrénées.

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.

Timeline

The construction of the hydrodynamics model for the Garonne catchment will be achieved for MS1 and a sensitivity study has been carried out with respect to friction coefficient and upstream discharge perturbed with a Gaussian process realization for stationnary flow. Work on non-stationnary flow is on going and mixture of expert surrogate is currently been investigated in T1.4.3-2 and will contribute to the D1.2.

WATER	Project Months											
T1.4.3		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.4.3-1												
T1.4.3-2												
T1.4.3-3												

Risk and mitigation

T1.4.3-1: The computation of the surrogate for the Gironde estuary may be complex due to the uncertainty in atmospheric fields due to multivariate temporal and spatial correlation. Such correlations are complex to represent with Gaussian Process that we plan to use to build an ensemble of perturbations for the forcing. A simplified working case would be the Garonne river that is also 2D with active flood plain, where atmospheric forcings are not needed. In this case, only univariate temporal forcing fields are required in the Design of Experiment (sampling of the inputs space), thus limiting the difficulty in the generation of the learning database. The merits of working on the Garonne stands in the simplicity of the forcing fields compared to those of the Gironde estuary, plus the computational cost of the TELEMAC model is significantly smaller on the Garonne. This should not cause any delay in the project. The inconvenient of Garonne test case is that it does not include any electricity production site (such as Blayais on the Gironde), the demonstration would thus be less applicative to the industrial context. Yet the methodology remains appropriate: the merits of running sensitivity analysis would thus be demonstrated.

T1.4.3-2: Building a surrogate model for active flood plain evolves discontinuity in the response surface where there may be no water (for several sets of input variables), then some water for other input sets. The discontinuity is triggering and may be left aside for the time being, limiting the study to stationary flow. If that choice has to be made, the delay in the project should be minimal as the stationary surrogate model has already been validated in the context of S. El Garroussi's PhD.

T1.4.3-3: Chaining PF-CLM discharge with the hydrodynamics model requires that discharge is computed at neighboring points to the inflow. This has not been discussed yet between CERFACS and FZJ. If surrogate model for non-stationary flow are not successfully built, we could limit the Parflow-TELEMAC chain to a full direct model simulation, leaving aside the surrogate part, and focus on the impact of using simulated discharge versus observed discharge as input to TELEMAC.

This would be an informative study as well, even though it may be expensive and thus should remain within the scope of a research project.

6.4 Task T1.4.4

The main task T1.4.4 is subdivided in 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. Furthermore, literature research on the state of the art of OED for reservoir engineering and other geoscientific applications will be done.
T1.4.4-2	**Software access**. A basic OED workflow for geothermal reservoir models has been developed as part of the research project MeProRiskII by partners at FSU Jena [10] [14].The software framework EFCOSS (Environment for Combining Optimization and Simulation Software) was used for linking our in-house geothermal simulation code SHEMAT-Suite to mathematical optimization software. The first step is to get access to the EFCOSS code in order to be able to use the existing OED workflow for geothermal reservoir models. Training and a practical introduction to the usage of the OED workflow will be pursued.
T1.4.4-3	**Learning and understanding the EFCOSS software environment and its coupling with SHEMAT-Suite**. This subtask we aim at understanding the OED workflow, which involves the software framework EFCOSS and the geothermal reservoir simulator SHEMAT-Suite, by getting acquainted with the source code. This step is essential for adding new functions to the source code or adapt existing functions. The existing workflow comprises the functionality to compute the optimal surface position for a borehole in order to measure temperature for estimating a rock property (e.g., permeability) with least uncertainty based on an OED algorithm. Understanding the source code will help us understand if and how we can change experimental conditions such as predicting not only the optimal borehole surface location but also the optimal borehole depth. Moreover, it is essential to understand how and at which points in the source code SHEMAT-Suite is coupled to the EFCOSS framework for identifying possibilities of performance optimization.
T1.4.4-4	**Reproducing OED simulations and developing a post-processing workflow**. In order to learn how to apply the OED workflow, we reproduce two existing OED studies by FSU Jena group: (1) a synthetic two-dimensional model [13] and (2) a three-dimensional geothermal reservoir model [14]. 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**. Develop a more detailed concept for the improvement of the existing OED workflow than the general description of the task in the Grant Agreement. This subtask involves identifying relevant and feasible extensions of the existing OED functionalities for geothermal reservoir engineering. So far, it has been shown that it is possible to predict the optimal borehole location at the surface for estimating permeability of one or more subsurface model units with least uncertainty from temperature logs measured within this borehole [13], [14]. This can be extended for example by predicting the optimal depth of temperature logs or by finding optimal locations for estimating other petrophysical parameters such as thermal conductivity or porosity. Moreover, this subtask includes identifying and defining interesting research questions such as studying the influence of apriori measurement locations or of data quality on the OED outcome. Additionally, performance bottlenecks of the workflow will be identified and feasible optimization strategies will be defined. The results of this subtask will define the work to be done in T1.4.4-6 and T1.4.4-7.
T1.4.4-6	**Setting-up a testmodel suite and extending OED functionality**. A suite of synthetic testmodels will be used for studying aspects defined in T1.4.4-5 and for optimizing the OED workflow. The test suite should comprise synthetic two- and three-dimensional geothermal reservoir models of different size and complexity. In addition and according to the outcomes of T1.4.4-5, new functionalities will be added to the existing OED workflow and defined research aspects will be studied using the testmodel suite which was set-up in this task.
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. For example, for determining the optimal borehole location, the Fisher Matrix needs to be computed repeatedly for every location within the model domain. Thus, the forward model (i.e., SHEMAT-Suite) is called and computed repeatedly. This process could be parallelized over computing nodes by using an MPI backend. This will increase the performance of the OED workflow, what in turn will enable us to apply the OED workflow to highly resolved reservoir models or larger model domains which were not computationally feasible so far.

Timeline

The timeline table illustrates the planned timeline for finishing the seven subtasks described above. Since this report is delivered in month six, T1.4.4-1 to T1.4.4-3 have been already completed successfully and T1.4.4-4 is ongoing. We have access to the EFCOSS software environment and were able to reproduce existing OED simulations using EFCOSS and SHEMAT-Suite. A post-processing workflow for the OED results is currently been developed. In the following months, we are going to define a detailed research and development concept for OED for geothermal modelling and compile a suitable set of test models. The outcome will be presented in the first Milestone MS1. Subsequently, the second and third deliverables will document the improvements of the OED-workflow functionality and performance.

WATER	Project Months											
T1.4.4		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.4.4-1												
T1.4.4-2												
T1.4.4-3												
T1.4.4-4												
T1.4.4-5												
T1.4.4-6												
T1.4.4-7												

Risk and mitigation

T1.4.4-2: No access to the existing OED workflow because parts of the code were developed by a third partner (FSU Jena group). In this case we would have to develop our own implementation of optimization algorithms to SHEMAT-Suite. The main focus of T1.4.4 would then change from extending and optimizing an existing implementation to developing a new OED implementation, which would probably delay the subsequent subtasks for few months. However, in that case T1.4.4-3 would be obsolete so that some extra months become available. We will report on the necessity of adapting the proposed Timeline in the Milestone report MS1.
T14.4.-3 & -4: The presence of code errors in OED code or its link to SHEMAT-Suite, possibly found during T1.4.4-3 and T1.4.4-4, might delay the schedule. This delay is expected to be in the range of maximum two months. If necessary, it could be compensated by reducing the extent of and time spent on SubTask 1.4.4-6. We will report on the necessity of adapting the SubTasks in the Milestone report MS1.
T1.4.4-7: The developments necessary for optimizing the OED workflow might be numerically or computationally more challenging and thus might be more time consuming than expected. In that case, we would seek for help from HPC experts within the EoCoE-II team in order to reduce the delay of T1.4.4. This is the most severe risk of T1.4.4. It will either delay the completion of this task by up to several months or result in a code that is less optimized than planned originally. Both, a delay or a less optimize code will affect T1.4.5 because we would have to coarsen model discretization or reduce size of the model domain for completing T1.4.5 on time. We will report on the status of T1.4.4-7 and possible consequences for T1.4.5 in the second deliverable (D1.2).

6.5 Task T1.4.5

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

T1.4.5-1	**Evaluation of the study area**. We intend to use a geothermal reservoir in Tuscany (Italy) as a study area for applying the OED workflow developed in T1.4.4. We will evaluate the feasibility of this study area for T1.4.5 from practical, physical and scientific points of view. In addition, we will evaluate other optional study areas (worldwide) which are – as the Tuscany reservoir – digitally recorded in their geological structures already. Finally, we will decide on a suitable study area and show case for applying the OED workflow.

T1.4.5-2	**Data collection**. Once the study area will be defined, we will gather the available data on the geological structure (i.e., the 3D geological model) and collect additional relevant data such as borehole logs. Depending on the study area, this task might include literature research, online database research, as well as communication with respective companies and authorities. The accessibility of data will have been clarified as part of the decision making process in T1.4.5-1.
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 workflow.
T1.4.5-5	**OED study**. Finally, we will apply the OED workflow developed in T1.4.4 to the geothermal reservoir model. An overview of the workflow is described in the section on Task 1.4.4. The aim is to predict optimal borehole locations or logging depths for new temperature measurements in order to deterministically invert for physical rock properties (e.g., permeability, thermal conductivity) with least uncertainty.

Timeline

The table below illustrates the planned timeline for finishing the five subtasks described above. Since T1.4.4 and T1.4.5 are consecutive, the beginning of the main work on T1.4.5 is in month 22 when T1.4.4 will be completed. Only Subtask 1.4.5-1 - the evaluation of the study area - is planned to be completed before the Milestone MS1 because it may result in reorienting of the scientific objectives. We will report the results of T1.4.5 in the third deliverable D1.3 at the end of the project.

WATER	Project Months											
T1.4.5		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.4.5-1												
T1.4.5-2												
T1.4.5-3												
T1.4.5-4												
T1.4.5-5												

Risk and mitigation

T1.4.4-2: The originally proposed show case of a geothermal reservoir in Tuscany might not be suitable or feasible for this task. The area is mostly under concession of Enel Green Power, Italy, and was the target for a previous project where OED developments firstly had been applied. However, currently we have no cooperation agreement with Enel Green Power, who owns most of the relevant data. Additionally, the very hot Tuscany reservoir involves particular complications due to water/steam phase transition and possible two-phase flow which burden the elaboration of the OED problem unnecessarily. For mitigating this risk we plan to evaluate the study of this area at an early time in the course of the project (T1.4.5-1). Thus, we will have the possibility to decide on another study area if necessary, without delaying the progress of the task. The T1.4.5-1 is scheduled before the Milestone in MS1 so that we can report on the decision about a suitable study area in the Milestone report.
T1.4.5-3 & -4: Model set-up or model calibration could be more complex than expected. The modelling process might take longer than expected or the calibration result might not be satisfactorily. If we encounter respective difficulties that may result in a delay of more than two months, we would use a quasi-synthetic version of the model for performing the OED simulations instead of the realistic reservoir model. Thus, we can prevent a delay of T1.4.5 beyond the final deliverable.
T1.4.5-5: The code performance may not be high enough for large scale OED simulation. If we were not able to optimize the OED workflow well enough or in time for large scale OED simulations, we could reduce the model size accordingly, for instance by using a coarser discretization or only a subset of the reservoir. Thus we would still be able to process T1.4.5-5 in time for the final deliverable at the end of the project.

6.6 Partners and resources

T1.4.1: Bibi S. Naz, is working full-time position for WP1 since January 2019. Bibi Naz is working on all subtasks. provides support with parflow technical implementation and visualization of the results and report writing.
T1.4.2: UNITN is the main contributor of this Task. FZJ will contribute to the Task activities by providing during the second half of the project hyper-resolved gridded runoff time series in the Italian alpine region by means of ParFlow model. Two permanent staff members of the UNITN are involved in the task: prof. Bruno Majone and prof. Alberto Bellin. Two PhD students, Andrea Galletti and Alessandro Todaro, previously enrolled with other funding, will contribute to the task during the first 18 months of the project. Patrick Zuliani has been enrolled as PhD student specifically for this project and he is already active on the task activities starting from January 1st 2019. Andrea Galletti and Alessandro Todaro will work on subtask T1.4.2-1 and T1.4.2-2, while Patrick Zulian, after adequate technical formation, will contribute to subtasks T1.4.2-1, T1.4.2-3 and T1.4.2-4. Bruno Majone is PI of UNITN group and will supervise the activities of all subtasks. Alberto Bellin provides support with HYPERstreamHS technical implementation and interpretation of the results.
T1.4.3: CERFACS is the only partner involved in this main Task. Sophie Ricci and Nicole Goutal works on all tasks of T1.4.3 and will supervise the activities of all subtasks. Andrea Piacentini has a consulting role supporting with the TELEMAC software. Siham El Garrousi is a PhD student working on surrogate models for TELEMAC in non-stationary flow. Nicole Goutal is a permanent staff from EDF and will collaborate with CERFACS on hydrodynamics and surrogate modeling. She is the co-advisor ofSiham El Garroussi.
T1.4.4 & T1.4.5: The partners involved in this Main Task are the Institute for Applied Geophysics and Geothermal Energy (GGE), E.ON Energy Research Center, RWTH Aachen University· The financial resources dedicated to this tasks are employed to hire one researcher for 16 person-months (PMs) deployed to the scientific tasks 1.4.4 and 1.4.5 within WP1. Johanna Bruckmann (M.Sc.

Applied Geosciences) will work these 16 PMs for WP1 with a half-time position (50 %, 20 hours per week) spread over 32 months (January 2019 - August 2021). Johanna Bruckmann works on all subtasks of Task 1.4.4. The other team members have a supervising function. Gabriele Marquart supervises the work as temporary head of the institute and the reservoir modelling group at GGE until September 2019. Her succession is still unclear. Christoph Clauser has a supervising function for the scientific work being Emeritus Professor at GGE and PhD supervisor for Johanna Bruckmann.

Partner	People	Role	Task
FZJ	Stefan Kollet	Leader of SC Water for energy	
FZJ	Bibi S. Naz	Co-Leader of SC Water for energy and main researcher	All tasks of T1.4.1
UNITN	Bruno Maione	Permanent staff. Supervisor of all activities of T.1.4.2	All tasks of T1.4.2
UNITN	Alberto Bellin	Permanent staff	All tasks of T1.4.2
UNITN	Andrea Galletti	PhD student	T1.4.2-1, T1.4.2-2
UNITN	Alessandro Todaro	PhD student	T1.4.2-1, T1.4.2-2
UNITN	Patrick Zuliani	PhD student	T1.4.2-1, T1.4.2-3, T1.4.2-4
CERFACS	Sophie Ricci	Permanent staff	T1.4.3, T1.4.3-3
CERFACS	Andrea Piacentini	Consultant on TELEMAC	T1.4.3-1
CERFACS	Siham El Garrousi	PhD student	T1.4.3-2
EDF	Nicole Goutal	Permanent staff	T1.4.3, T1.4.3-1
RWTH	Johanna Bruckmann	PhD student, main researcher	All tasks of T1.4.4 & T1.4.5
RWTH	Gabriele Marquart	Temporary supervisor	All tasks of T1.4.4 & T1.4.5
RWTH	Christoph Clauser	PhD supervisor	All tasks of T1.4.4 & T1.4.5

7 Fusion for Energy (T1.5)

GYSELA is a 5D full-f and flux-driven gyrokinetic Fortran parallel code that solves Vlasov and Poisson equations to simulate electrostatic plasma turbulence and transport in the core of Tokamak devices. Parallelization efficiency is presently limited by network capability and memory performance. Within EoCoE-II, the new release GYSELA-X will target exascale supercomputers. This is particularly required by the need to address electromagnetic turbulence from the core to the far edge region in ITER-relevant magnetic geometries. These constitute groundbreaking physics upgrades. Task 1.5 aims at paving the way towards the preliminary exploitation of GYSELA-X. T1.5 is divided in thre main tasks:

T1.5.1	Prototype of GYSELA-X: arbitrary magnetic equilibrium in limiter configuration
T1.5.2	Advanced GYSELA-X: X-point configuration & alternative/complementary methods
T1.5.3	Core-edge-SOL physics: GYSELA-X & Tokam3X

T1.5.1 focuses on the basics with the help of prototypes, namely the derivation/implementation of the new set of equations and the identification of optimal choices for handling non-circular geometries.
T1.5.2 deploys the solutions in the advanced version, and extends the problem to critical large temperature variation and X-point issues.
T1.5.3 takes advantage of all the developments to address first physics issues.

7.1 Task T1.5.1

The main task T1.5.1 is subdivided in two subtasks:

T1.5.1-1	**Identify minimal set of equations for ion turbulence**. This subtask is split in two activities. *Core physics.* – So far, GYSELA works in the electrostatic limit. The gyrokinetic equations including electromagnetic effects and dealing with generalized magnetic geometries will be derived for core plasmas, i.e. for closed magnetic field lines. The strategies which have revealed to be most numerically efficient in other codes (so-called "magnetic cancellation" issue) will be adapted to the semi-Lagrangian numerical scheme of GYSELA-X. *Edge & SOL physics.* – The treatment of the boundary layer, characterized by a transition from closed (edge region) to open ("Scrape-Off Layer" or SOL) magnetic surfaces will be worked out. The foreseen method is to use immersed boundary conditions (via the penalization technique) in the SOL to account for the presence of material surfaces. The issue of the X-point (location where one component of the magnetic field vanishes) will be addressed from a theoretical point of view.
T1.5.1-2	**Comparative efficiencies of flux-surface aligned and Cartesian mesh grids in poloidal plane, and of patches with regular or irregular meshes**. This subtask is split in two activities, each dealing with one of the two main solvers. *Poisson (& Ampére) solver.* – The approach using flux-aligned grid for D-shape geometry will be pursued and its efficiency evaluated for Poisson eq. (inversion of a generalized 2D Laplacian). We will additionally work out then details a Cartesian grid for D-shape geometry, implement it and optimize it in GYSELA-X. Multigrid solvers are likely to be required. The case of circular geometry with multi-patches will be addressed within the WP3 (collab. with CERFACS). Eventually, a multigrid solver for a D-shape magnetic geometry—i.e. including the mapping—will be developed and implemented in GYSELA-X. *Vlasov solver.* – We will develop an hybrid method for the Vlasov solver which will able to resolve D-shape geometries and then implement it in GYSELA-X within the frame of a PhD thesis.

Timeline

FUSION T1.5.1	Project Months											
		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.5.1-1												
T1.5.1-2												

Progress and final reports will be written on each of these subtasks for Milestone MS1 and deliverable D1.2, respectively, including numerical and physics advances.

Risk and mitigation

The CEA-IRFM team has just lost one of the pillar developers of the code, with an expertise on computational science and high-level parallelism which cannot be replaced by non-permanent staff. The whole project has been restructured so as to preserve its main goals. Yet, a few more prospective activities had to be abandoned (cf. WP2, Task 2.6).
T1.5.1-1: No specific risk identified regarding this subtask.
T1.5.1-2: If we face difficulty/delay to find adequate postdocs and HLST support, then the work will still be done by permanent staff but some delays in achieving the intended results my be possible. Also, some developments will be delayed, some others may be abandoned. If it occurs, the choice will retain the best option. In any case, the main goals of the project will be preserved.

7.2 Task T1.5.2

The advanced version of GYSELA-X will be developed by extending the prototype up to the wall of the reactor. The specific treatment of the presence of an X-point (locus where one component of the magnetic field vanishes; there, the magnetic field line reduces to a circle) will be addressed by means of dedicated work on reduced models and/or simplified cases. The main task T1.5.2 is subdivided in two subtasks:

T1.5.2-1	**Handle the X-point of ITER-relevant magnetic configurations with multi-patch or flux coordinate independent schemes**. The "flux coordinate independent" numerical scheme, initially developed for fluid codes, will be implemented and its efficiency to treat the dynamics across the X-point will be tested in the GENE Eulerian code. Immersed boundary conditions via the penalization technique will be implemented for the Vlasov solver in GYSELA-X. These conditions require to add extra linear operators to the Vlasov equation at the wall locations. These operartors account for plasma-wall interactions, which are mandatory for properly recovering the structure and magnitude of equilibrium flows. The details of a Poisson solver capable of handling the expected large gradients of the electric potential in the vicinity of the X-point will be worked out. To this end, the efficiency of those methods developed for the general D-shape geometry will be evaluated.

T1.5.2-2	**Multi-patch treatment of the Vlasov equation to handle large variations of temperature**. Large temperature variations (by up to 2 orders of magnitude) develop from core to peripheral fusion plasmas, the so-called Scrape-Off Layer. Also, temperature gradients are large at the edge, with gradient lengths reaching up to a few turbulence correlation lengths. Handling these requires the development of a multi-patch treatment, at least in the (radial) direction of the gradients, so as to avoid wasting memory resources. Such a multi-patch, which is also required for D-shape geometry, will be developed and implemented in GYSELA-X.

Timeline

There is little interplay between these different developments, so that they can mostly been done in parallel. The way to proceed is either to use different prototypes, or to create branches in the git version of the main code. Both approaches will be used. The full integration and tests of the successful developments will be done before the end of the project.

FUSION	Project Months											
T1.5.2		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.5.2-1					■	■	■	■	■	■	■	■
T1.5.2-2				■	■	■	■	■	■	■	■	■

Progress and final reports will be written on each of these subtasks for Milestone MS1 and deliverable D1.2, respectively, including numerical and physics advances.

Risk and mitigation

The same risks as the ones identified for task 1.5.1-2 also apply here. In addition, we have identified a few others more specific to these tasks:
T1.5.2-1: Numerical instabilities may occur when implementing immersed boundary conditions in the open field line region, especially in the regime where both ions and electrons are treated kinetically. Indeed, in this case, turbulent particle transport adds up to the problem, and the system has to ensure a bipolarity despite the different mobilities of the two species. Possible solutions include the use of numerical schemes with enhanced damping.
T1.5.2-2: Discontinuities may appear at the interface of the patches that may be difficult to handle. Resolving these possible issues will be considered as a priority task, and might delay other activities.

7.3 Task T1.5.3

The objective is to address the physics of edge-core interplay in tokamak plasmas by means of axisymmetric and turbulent simulations with the newly developed GYSELA-X code. Special attention will be paid to the structure and dynamics of large scale axisymmetric flows in the vicinity of the transition layer between closed (core) and open (Scrape-Off Layer or SOL) field line regions. These flows, which are expected to be large in this region, are known to play a critical role in turbulence self-organisation and saturation. The obtained flows will be compared to simulation results obtained with the 3D fluid code Tokam3X. The main task T1.5.3 is subdivided in two subtasks:

T1.5.3-1	**Recover large scale equilibrium flows expected in the SOL**. Simulations covering both core and edge plasma with immersed boundary conditions will be run in the axisymmetric regime (hence without turbulence). Once the steady-state is reached, the structure and magnitude of equilibrium flows will be compared to results of the fluid code Tokam3X in similar regimes.
T1.5.3-2	**Highlight the critical interplay between core, edge and SOL turbulence and transport in (some of) the ITER relevant parameter range**. The interplay between core, edge and SOL plasmas (including immersed boundaries) will be addressed in the turbulent regime. Special attention will be paid to the role of large scale flows and their interaction with turbulence spreading. Preliminary simulations will run at modest temperature variations from core to edge, followed by more ambitious ones at ITER relevant large temperature variations.

Timeline

FUSION	Project Months											
T1.5.3		D1.1		MS1		D1.2						D1.3
Subtasks	3	6	9	12	15	18	21	24	27	30	33	36
T1.5.3-1												
T1.5.3-2												

Subtask 1.5.3-1 requires the successful implementation of immersed boundary conditions in GYSELA-X (task 1.5.2-1). Yet, the axisymmetric simulations considered here are expected to be less numerically challenging than turbulent ones, hence more likely not to exhibit possible instabilities. Subtask 1.5.3-2 is aimed at capitalizing all numerical upgrades implemented in GYSELA-X so as to address new ITER-relevant physical issues, namely the edge-core interplay in complex geometry. The results might well be simply preliminary and proof-of-principle, depending on the various delays encountered along the road. Yet, GYSELA-X aims at being the reference code for the 5-10 years after the end of EoCoE-II, so that new physics will surely be unraveled, either in the time frame of EoCoE-II or in the following years.

Risk and mitigation

General risk: this task capitalizes on many of the upgrades and groundbreaking developments performed on GYSELA-X within the whole EoCoE-II project. It may happen, if critical delays occur, that either less ambitious simulations from the physical point of view will be carried out, or vice-versa that they will be performed at reduced parallel efficiency and exascale-relevant performance. Risks regarding subtask 2) and mitigations: If we face difficulty/delay to find adequate postdocs and HLST support, then the work will still be done by permanent personnel but at lower speed.

7.4 Partners and resources

We plan to hire:
M10: an expert for the implementation and optimization in GYSELA of a Cartesian grid for D-shape geometry. The project of CEA-IRFM+MPG has been selected by the European HLST (High Level Support Team) program on fusion.
M10: a postdoc for implementing FCI approach in GENE from the EoCoE-II budget. The hired resource will work in MPG.

M10: Emily Bourne (EB), PhD for implementing an hybrid method for Vlasov solver on D-shape geometry in GYSELA-X as part of the European PhD programme “NUMERICS”. The hired resource will work in CEA-IRFM & AMU
M12: a postdoc EoCoE-II for implementing the multigrid solver for D-shape in GYSELA-X. The hired resource will work in MPG.
M12: Hire a second postdoc from EoCoE-II budget for multipatch treatment to handle large temperature variation from core to Scrape-Off Layer (this resource is shared with WP2, T2.6). This resource will work in CEA-IRFM.
M18: a PhD student for running core-edge simulations with the newly developed GYSELA-X code. The hired resource will work at CEA-IRFM. As part of a technical challenge a postdoc will be hired for developing multigrid solver for circular geometry with multi-patches EoCoE-II budget. This resource will be located in CERFACS and work on tasks described in WP3.

Partner	People	Role	Task
CEA-IRFM	Guilhem Dif-Pradalier	Permanent Staff (Physics)	T1.5.1-1, T1.5.2-1, T1.5.3-1, T1.5.3-2
CEA-IRFM	Xavier Garbet	Permanent Staff (Physics)	T1.5.1-1
CEA-IRFM	Philippe Ghendrih	Permanent Staff (Physics)	T1.5.1-1, T1.5.2-1, T1.5.3-1, T1.5.3-2
CEA-IRFM	Yanick Sarazin	Permanent Staff (Physics)	T1.5.2-1, T1.5.3-2
CEA-IRFM	Patrick Tamain	Permanent Staff (Physics)	T1.5.3-1
CEA-IRFM	Virginie Grandgirard	Permanent Staff (Numerical Analysis)	T1.5.1-2, T1.5.2-1, T1.5.2-2
CEA-IRFM	Elisabetta Caschera	PhD student (Physics)	T1.5.1-1, T1.5.2-1
CEA-IRFM	Camille Gillot	PhD student (Physics)	T1.5.1-1
MPG	Yaman Güçlü	Expert (Applied Math)	T1.5.1-2
MPG	Ahmed Ratnani	Expert (Applied Math)	T1.5.1-2
MPG	Kab Seok Kang	Expert (Applied Math)	T1.5.1-2,
MPG	Frank Jenko	Permanent Staff (Physics)	T1.5.2-1
MPG	Eric Sonnendrücker	Expert (Applied Math)	T1.5.1-1, T1.5.1-2, T1.5.2-1
INRIA/AMU	Michel Mehrenberger	Expert (Applied Math)	T1.5.1-2, T1.5.2-1, T1.5.2-2
CEA-IRFM/AMU	Emily Bourne	PhD student	T1.5.1-2, T1.5.2-1, T1.5.2-2

References

[1] “For wind turbines in complex terrain, the devil is in the detail”, Julia Lange, Jakob Mann, Jacob Berg, Dan Parvu, Ryan Kilpatrick, Adrian Costache, Jubayer Chowdhury, Kamran Siddiqui and Horia Hangan, Environmental Research Letters, Volume 12, Number 9, 2017. doi 10.1088/1748-9326/aa81db.

[2] “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”, G. Psofogiannakis and A. C. T. van Duin, Surface Science 646 (2016) 253–260

[3] “Firedrake: Automating the Finite Element Method by Composing Abstractions”, F. Rathgeber, D. A. Ham, L. Mitchell, M. Lange, F. Luporini, A. T. T. Mcrae, G-T. Bercea, G. R. Markall, and P. H. J. Kelly, ACM Transactions on Mathematical Software, **43** Article 24 (2016), DOI:10.1145/2998441

[4] “A fast and robust numerical scheme for solving models of charge carrier transport and ion vacancy motion in perovskite solar cells”, N.E. Courtier, G. Richardson and J.M. Foster, Applied Mathematical Modelling **63** 329 (2018), DOI:10.1016/j.apm.2018.06.051

[5] “A domain specific language for performance portable molecular dynamics algorithms”, W. R. Saunders, J. Grant, E. H. Müller, Computer Physics Communications **224** 119 (2018), DOI:10.1016/j.cpc.2017.11.006

[6] “Fast electrostatic solvers for kinetic Monte Carlo simulations”, W. R. Saunders, J. Grant, E. H. Müller, I. R. Thompson, (2019), arXiv:1905.04065

[7] “Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model”, S. J. Kollet and R. M. Maxwell, Advances in Water Resources, Volume 29, Number 7, 945 (2006), DOI:10.1016/j.advwatres.2005.08.006

[8] “Improvements to the Community Land Model and their impact on the hydrological cycle,” Oleson, K. W. and Niu, G.-Y. and Yang, Z.-L. and Lawrence, D. M. and Thornton, P. E. and Lawrence, P. J. and Stöckli, R. and Dickinson, R. E. and Bonan, G. B. and Levis, S. and Dai, A. and Qian, T. Journal of Geophysical Research: Biogeosciences, Volume 113, Number G1, (2008), DOI:10.1029/2007JG000563.

[9] Atkinson, A.C., Donev, A.N., 1992. “Optimum experimental designs”, Oxford Science Publications, Clarendon Press

[10] Marquart, G. (ed), 2015. Verbundprojekt MeProRisk-II: Optimierungsstrategien und Risikoanalyse für tiefe geothermische Reservoire – Eine Machbarkeitsstudie. Teilprojekt A: umerische Modellbildung und Optimierungsverfahren. Endbericht BMWi-Projekt 0325389A

[11] “Joint three-dimensional inversion of coupled groundwater flow and heat transfer based on automatic differentiation: sensiivity calculation, verification and synthetic examples”, Rath, V., Wolf, A., Bücker, H.M., Geophys. J. Int., 167, 453 (2006).

[12] EFCOSS: “An Interactive Environment Facilitating Optimal Experimental Design”, Rasch, A., Bücker, H.M., ACM Transactions on Mathematical Software, 37(2) (2010).

[13] “Model identification for flow simulations in geothermal reservoirs: towards optimally drilling boreholes”, Padalkina, K., Bücker, H.M., Herty, M., Rath, V., Seidler, R., Proc. Appl. Math. Mech. 13(1), 345 (2013).

[14] "Optimal experimental design for reservoir property estimates in geothermal exploration", Seidler, R., Padalkina, K., Bücker, H.M., Ebigbo, A., Herty, M., Marquart, G., Niederau, J., Comput. Geoscience, **20**, 375 (2016).

[15] "A Scale-Consistent Terrestrial Systems Modeling Platform Based on COSMO, CLM, and ParFlow,"Shrestha, P. and Sulis, M. and Masbou, M. and Kollet, S. and Simmer, C., Volume 142, Number 9, 3466, (2014). DOI:10.1175/MWR-D-14-00029.1.

[16] " Improving soil moisture and runoff simulations at 3 km over Europe using land surface data assimilation" Naz, B. S., Kurtz, W., Montzka, C., Sharples, W., Goergen, K., Keune, J., Gao, H., Springer, A., Hendricks Franssen, H.-J., and Kollet, S., Hydrol. Earth Syst. Sci., Volume 23, 277, (2019) DOI:10.5194/hess-23-277-2019.