REQUEST FOR AN EMI R&D PROJECT 2027–2029

Country/Organisation: Spain

Principal Investigator[1]: Etienne Tourigny

Affiliation: Barcelona Supercomputing Center

Address: Plaça Eusebi Güell, 1-3
08034 Barcelona (Spain)

Other researchers: BSC : Roberto Bilbao, Raffaele Bernardello, Vladimir Lapin, Valentina Sicardi, Stefano Materia, Lina Teckentrup, Pablo Ortega, Markus Donat, Francisco Doblas-Reyes, Amirpasha Mozafarri, Iria Ayan
MPG : Alexander J. Winkler

Project Title:

Development of new seasonal-to-decadal climate and carbon cycle prediction system using EC-Earth4

To make changes to an existing project please submit an amended version of the original form

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SPESICCF**	
Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)	2027	
Would you accept support for 1 year only, if necessary?	YES x	NO ☐

Computer resources required for project year:		**2027**	**2028**	**2029**
High Performance Computing Facility	[SBU]	140	300	200
Graphics Processing Unit Cluster-A	[GBU]	x	x	x
Graphics Processing Unit Cluster-B	[GBU]	x	x	x
Accumulated data storage (total archive volume)[2]	[GB]	0	0	0

---

[1] The Principal Investigator will act as contact person for this EMI R&D Project and, in particular, will be asked to register the project, provide annual progress reports of the project's activities, etc.
[2] These figures refer to data archived in ECFS and MARS. If e.g. you archive x GB in year one and y GB in year two and don't delete anything you need to request x + y GB for the second project year etc.
[3]The number of vGPU is referred to the equivalent number of virtualized vGPUs with 8GB memory.

EWC resources required for project year:		**2027**	**2028**	**2029**
Number of vCPUs	[#]			
Total memory	[GB]			
Storage	[GB]			
Number of vGPUs[3]	[#]			

Continue overleaf.

Principal Investigator: Etienne Tourigny

Project Title: Development of new seasonal-to-decadal climate and carbon cycle prediction system using EC-Earth4

Extended abstract

All EMI R&D Project requests should provide an abstract/project description including a scientific plan, a justification of the computer resources requested and the technical characteristics of the code to be used. The completed form should be submitted/uploaded at https://www.ecmwf.int/en/research/special-projects/special-project-application/special-project-request-submission.

Following submission by the relevant Member State the EMI R&D Project requests will be published on the ECMWF website and evaluated by ECMWF and its Scientific Advisory Committee. The requests are evaluated based on their scientific and technical quality, and the justification of the resources requested. Previous EMI R&D Project reports and the use of ECMWF software and data infrastructure will also be considered in the evaluation process.

Requests exceeding 5,000,000 SBU should be more detailed (3-5 pages).

Summary

In this EMI R&D Project request, we aim to develop and use a new seasonal-to-decadal climate and carbon cycle prediction system based on EC-Earth4. We will incorporate several ongoing developments which are part of a long-term strategy to use ML/AI techniques to provide time-varying vegetation and land carbon fluxes to overcome the limitations of the current system in place, and other methods to tackle longstanding climate biases and initialization shocks. Specifically, we will perform seasonal forecast demonstrators with ML/AI generated time-varying vegetation fields, perform a decadal prediction hindcast using the new system and finally integrate a novel ML/AI land carbon flux module to perform operational multi-annual predictions of the climate and global carbon cycle in support of the Global Carbon Budget and WMO Global Annual-to-Decadal Climate Update.

We build upon previous ECMWF special projects led by the PI and colleagues, as well as past and ongoing large European projects and service contracts.

Current seasonal-to-decadal forecast systems with EC-Earth3

The Climate Variability and Change Group of BSC's Earth Sciences Department has developed over the years a state-of-the-art decadal prediction system based on the EC-Earth 3 Global Climate Model (GCM) (Döscher et al, 2022). The original system (Bilbao et al., 2021) is used operationally to inform the World Meteorological Organisation Global Annual-to-Decadal Climate Update (Hermanson, L., et.al, 2022); the BSC is currently one of the four Global Producing Centers. This system was used to produce the CMIP6 Decadal Climate Prediction Project (DCPP; Boer et al., 2016) simulations which have been used in many studies exploring seasonal-to-decadal predictability and climate variability (e.g. Smith et al. 2020). As part of the Climate-Carbon Interactions in the Coming Century (CCiCC) H2020 project, the system was upgraded (Li et al, 2026; Bernardello et al, 2024) to perform multi-annual initialized predictions of the full carbon cycle with EC-Earth3-CC, using TM5 for atmospheric CO_2 transport, LPJ-GUESS vegetation model for land carbon stocks and fluxes and PISCES for ocean biogeochemistry and CO_2 fluxes. The carbon-cycle prediction system has been used to inform the Global Carbon Budget (GCB) reports since 2023 (Friedlingstein et al. 2023, 2025) with multi-year prediction of land and ocean CO_2 fluxes and atmospheric CO_2 concentrations. Extended 5-year hindcasts and predictions will soon be used to inform the WMO Lead Centre for Annual-to-Decadal Climate Prediction (Li et al, 2026). However, managing the two (GCM and CC) prediction systems in parallel and the aging infrastructure has been difficult to maintain. Also, the lack of native CO_2 tracer support with proper mass conservation in the IFS version of EC-Earth3, technical difficulties in initializing the LPJ-GUESS model on November 1st (required for operational forecasts), systematic vegetation shock when being initialized from ERA5 and persistent initialization difficulties in the Northern Atlantic (leading to the collapse of the AMOC) are strong incentives to develop a more modern system.

New seasonal-to-decadal forecast system with EC-Earth4

Additionally, BSC is a partner in the COPERNICUS PROJECT C3S2_375 - Development of C3S Decadal Prediction Service. In this contract, operational forecasts are performed every year, with the last forecast to be performed for 2028. As part of this service contract, BSC is improving the operationalization of its decadal prediction system and plans to update it to the newest EC-Earth model version — EC-Earth4 – which is being finalized and will participate in the CMIP7 AFT (Assessment Fast Track).

The EC-Earth4 GCM contains updates to NEMO (from version 3.6 to 4.2.2) and IFS (from IFS cycle 36 to OpenIFS cycle 48R1 and the Antarctic Freshwater Emulator (AIME) and PISM 2.2 Ice Sheet model (both improvements included in EC-Earth3-ESM-1 for the OptimESM project). The upgrade to OpenIFS includes an option to run the model in Single Precision (SP) which improves greatly computational performance and cost, as well as 3 options for aerosols: MACv2-SP aerosols are was used in the EC-Earth3, M7 with simplified chemistry and M7 with full chemistry. Support for atmospheric CO_2 and appropriate mass fixers is possible after work done in previous ECMWF special projects – OpenIFS Modeling of the Atmospheric Carbon Cycle (spnlpete-2022, spnlpete-2023). The complete version of EC-Earth4 (EC-Earth4-ESM) also contains LPJ-GUESS (for time-varying vegetation and CO_2 fluxes) and PISCES (for ocean CO_2 fluxes), allowing EC-Earth4-ESM to be used for emission-driven runs of the AFT.

However, LPJ-GUESS is not an ideal candidate to be used in the seasonal-to-decadal prediction system due to the limitations described earlier. We propose two alternative approaches using hybrid ML/AI based techniques for having time-varying vegetation and land CO_2 fluxes in our prediction system, in order to have a single unified system for predicting the global climate and the carbon cycle. The hope is that with a hybrid ML/AI approach we can correct the vegetation shocks when initializing the model from observations. The approach is illustrated in Figure 1.

The EC-Earth4 model is mostly ready to use and is currently being tuned for participation in the CMIP7 AFT. A few requirements to perform initialized predictions have already been implemented : framework for producing in-house ocean and ocean BGC reconstructions, full-field initializing of the atmosphere from ERA5 (based on Bernardello et al, 2024), support for running several start dates and members, integration into the Autosubmit workflow manager (Manubens-Gil et. al, 2016). The remaining technical developments will be implemented before the end of 2026 : proper support for saving openifs and nemo restarts, to be used as initial conditions ; use of the OpenIFS Data Hub to download initial conditions for all start dates, generation of proper land initial conditions using ecLand (as part of ongoing Special Project spesiccf-2026).

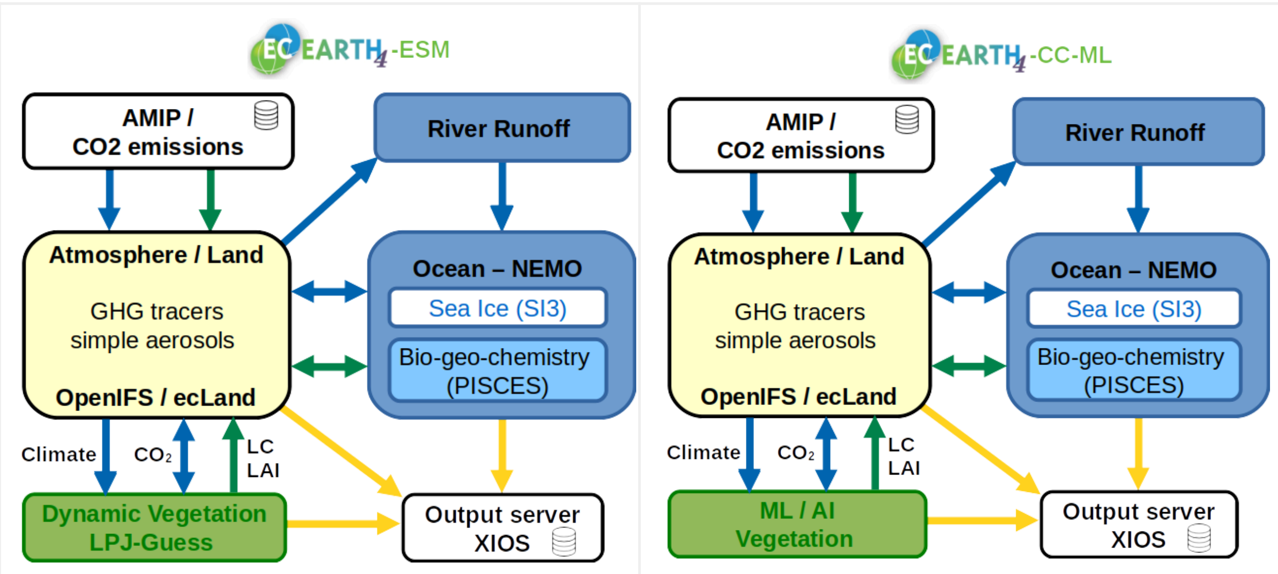

Figure 1 - Schematic of EC-Earth4 ESM version (AIME and PISM components omitted) with LPJ-GUESS (left) ; CC-ML version (right) with ML/AI methods time-varying vegetation fields and land CO_2 fluxes.

Time-varying vegetation fields

BSC has been involved in many research projects to provide time-varying vegetation fields for ecLand, the land surface component of IFS. Work started in the CERISE Horizon Europe research project, in order to provide time-varying vegetation fields for ERA6-land prototype (CERISE, 2025) and has evolved into more ambitious work for the CONCERTO and TerraDT Horizon Europe projects. The overall goal is to produce time-varying Land Use (LU), LC and LAI boundary conditions using ML/AI observationally-driven approaches and spanning the historical period of 1850 to future scenarios. The modelling framework is used to produce the high-resolution (1km), long-temporal-coverage land product family named TerraNostra. It includes the Land Use product TerraNostra Land use

(TerraNostra LU), and will be followed by TerraNostra Land Cover (TerraNostra LC) and TerraNostra Leaf Area Index (TerraNostra LAI).

These new LC and LAI datasets will be used for seasonal forecasts demonstrators in the CONCERTO project, as part of Task 6.2: Seasonal Forecast Demonstrators. The participating organizations to these demonstrators are ECMWF (IFS/ecLand), CMCC (CLM) and BSC (EC-Earth4/ecLand).

The goals of these demonstrators are to:

- Demonstrate and quantify the added value of the improved land-surface and carbon-cycle developments in a seasonal prediction framework, particularly for extreme events.
- Test improved land initialisation and model developments in seasonal forecasts
- Focus on well-documented extreme events (e.g. 2003, 2018, 2022 European summer extremes; 2010/11 Austral summer extremes), where land feedbacks (soil moisture – temperature coupling, vegetation stress) are expected to matter most.
- Evaluate impact on forecast skill, means, trends and land-atmosphere fluxes relevant to the carbon-cycle.

Hybrid modelling of the land carbon cycle

The BSC is involved in the PostPeak ERC Starting Grant (PI Alexander J. Winkler, more details at https://cordis.europa.eu/project/id/101222996) – How Land Carbon Dynamics Shape the Rise and Fall of Atmospheric CO2. In this project, a hybrid process-based and ML/AI data-driven model system CHOPINN (Carbon Hybrid-modelling using Observations and Process-Informed Neural Networks) will be developed, illustrated in Figure 2.

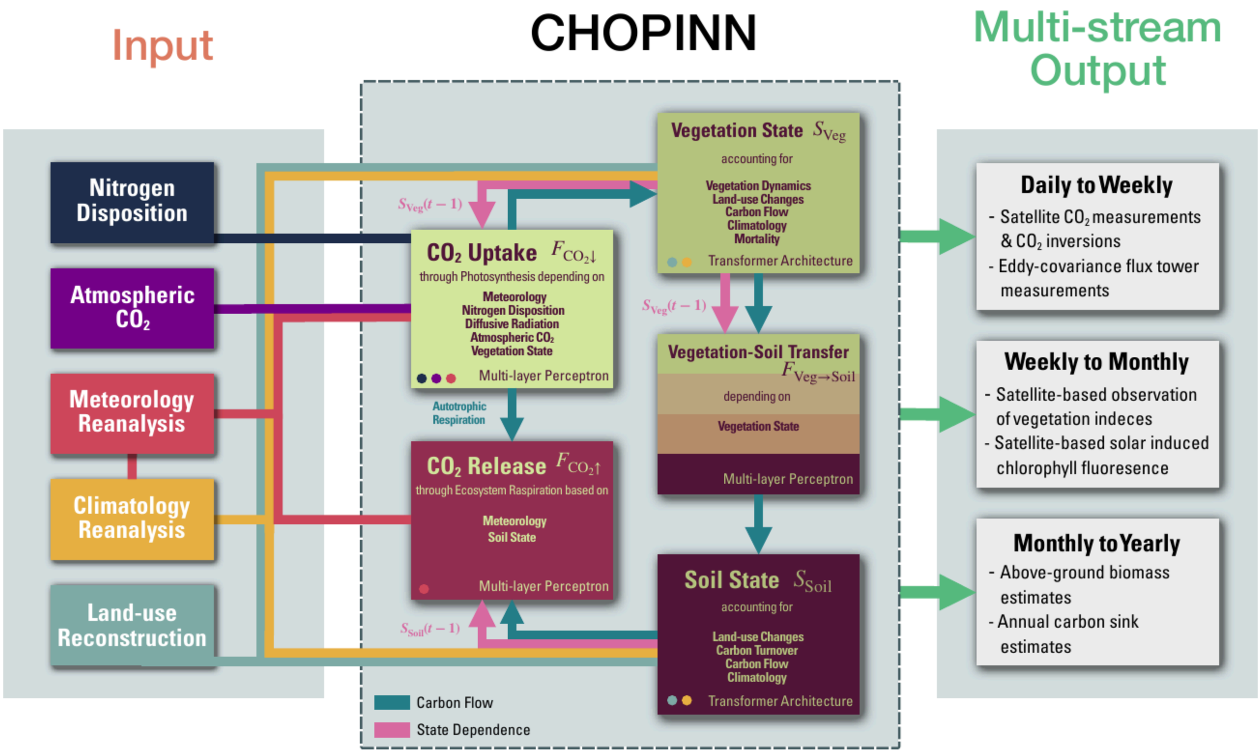

Figure 2 - Schematic of the CHOPINN modelling framework developed in the PostPeak project.

CHOPINN is designed to operate either as a standalone offline tool forced by meteorological boundary conditions, or as a coupled component within the Earth system models ICON and EC-Earth4. It will predict terrestrial carbon fluxes across timescales ranging from seasonal to multi-decadal, enabling near-real-time monitoring (roughly one-month delay from the real-time given update of atmospheric forcing files), near-term land carbon flux prediction (1–10 years), and long-range climate projection.

As an offline tool, it will provide low-latency carbon flux estimates driven by observed or reanalysis meteorology. As a coupled component within ICON and EC-Earth4, it will replace existing land carbon cycle representations with a machine-learning-based, observation- and process-guided framework that captures ecosystem dynamics with greater fidelity. This supports the study of land–atmosphere carbon feedbacks, the trajectory of the terrestrial carbon sink, near-term prediction of the carbon cycle, and the exploration of scenarios beyond peak atmospheric CO_2. Together, these developments address several long-standing research questions in land carbon–climate interactions and global carbon cycle–climate feedbacks, spanning timescales from daily to centennial dynamics, as outlined in the PostPeak ERC Starting Grant.

CHOPINN will be integrated into ICON-ESM and EC-Earth4 using a flexible coupling approach (using oasis or YAC couplers) and replace or enhance the vegetation components (JSBACH for ICON and LPJ-GUESS for EC-Earth4). For EC-Earth4, CHOPINN will communicate daily with OpenIFS, receiving meteorological fields from OpenIFS and sending CO_2 fluxes and optionally LAI. The EC-Earth4 + CHOPINN modelling system (hereby known as EC-Earth4-CC-ML) will be used for centennial-scale studies of the global carbon cycle as well as near-term prediction of the carbon cycle.

Proposed computational activities

A description of the model configurations used in this project is provided in Table 1, as well as an estimate of the speed in SYPD (Simulated Years Per Day) and cost in CHSY (Computing Hours per Simulated Year). Values for the GCM configurations were obtained on hpc2020 using Single Precision for OpenIFS and the standard TL255L91-eORCA1 grid. Values for the CC-ML configuration (with PISCES and CHOPINN for the carbon cycle) are estimated based on previous experiments with the full ESM configuration. The addition of PISCES requires more computational resources but does not slow down the coupled system. In addition, we will use two aerosol implementations, one is the simple MACv2-SP prescribed aerosols and the other more detailed M7 (SimChem) interactive aerosols scheme, which is ~4 times more computationally demanding and ~4 times slower than the MACv2-SP version.

Table 1 - Model configurations used in this project with components, aerosol scheme and estimated computational cost in CHSY

Configuration	main components	aerosols	CHSY
EC-Earth4-GCM-SP	oifs, nemo	MACv2-SP	1100
EC-Earth4-GCM-M7	oifs, nemo	M7 SimChem	4000
EC-Earth4-CC-ML-SP	oifs, nemo, pisces, chopinn	MACv2-SP	1800
EC-Earth4-CC-ML-M7	oifs, nemo, pisces, chopinn	M7 SimChem	4980

The first computational activity to be performed with the computing hours of this EMI R&D Project will be the seasonal forecast demonstrator for the CONCERTO project.

The established protocol is :

- 4-month-long initialized seasonal hindcasts in atmosphere-only mode
- Land surface initialization based on offline land surface model simulations forced by ERA5
- Hindcast Period: At least 20 years for robust skill assessment
- 20 Ensemble members
- Start date: 01-May
- Forcings:
    - SST/sea-ice - all models use ESA-CCI
    - Standard forcings (GHG, aerosols, solar etc) according to each modelling centre, but to be identical across CONTROL and sensitivity experiments

And the planned experiments for EC-Earth4 are shown in Table 2

Table 2 - Planned seasonal forecast demonstrator experiments for EC-Earth4 in CONCERTO Task 6.2

Experiment	Description	Purpose	Initial conditions	Forcings
CONTROL	Control experiment	Baseline	Offline Land-only run	SST/Sea Ice or fully coupled

CONCERTO LULC-LAI	Implementation of WP2 developments	Impact of varying LULC/LAI	Offline Land-only run	SST/Sea Ice or fully coupled
PROG_LAI	Prognostic LAI developed in WP5	Impact of enabling prognostic LAI	Offline Land-only run	SST/Sea Ice or fully coupled

The baseline seasonal forecast experiments will be conducted as soon as the basic prediction system is ready for use (planned for early 2027). An evaluation of the prediction skill as well as the necessary modifications to EC-Earth4 (incorporation of time-varying vegetation and prognostic LAI from ECMWF contribution) will be done before the sensitivity experiments will be performed. These experiments will be done in two sets, one using the simple MACv2-SP prescribed aerosols and with the more detailed M7 (SimChem) interactive aerosols, in order to evaluate the added benefit of using interactive aerosols.

The second set of simulations are the initialized decadal predictions, which are initialized on November 1st and are done for 11 years (as to obtain a full 10 years of prediction). We will start with multi-year (3) hindcasts of 5 members as a proof of concept and to perform an initial skill assessment and provide any needed corrective measures). As in the seasonal hindcast demonstrators we will compare hindcasts done using MACv2-SP and M7-SimChem aerosols to quantify the added skill from interactive aerosols. We plan to conduct these simulations during 2027, after the evaluation of the baseline seasonal hindcasts.

We will use the outputs of the multi-year (3) hindcast to run CHOPINN in offline mode and evaluate the skill in prediction of land CO_2 fluxes. This will allow to fine-tune the EC-Earth4-CC-ML system, which will be used in another set of multi-year hindcast tests in the second half of 2028. We will also perform the full hindcast and forecast using the EC-Earth4-GCM-SP system, in time to produce the 2029 initialized forecast for the C3S2_375 contract.

Finally, in the first half of 2029 we will perform an entire hindcast with the carbon cycle and more members (15), in time for a contribution to the Global Carbon Budget (2029) using a single model version for prediction of climate and atmospheric CO_2, and use this system for the following year's WMO official predictions.

Long decadal hindcasts using M7 (which is around 3 times more costly than the MACv2-SP version) are not requested in thisEMI R&D Project request and could be performed in another HPC or as part of another EMI R&D Project request, should initial tests show a substantial improvement over the version with MACv2-SP aerosols.

In Table 3 we summarize these planned experiments and total required computing time. We estimate for years 2027, 2028 and 2029 a total of 111, 238 and 157 M SBU, respectively. In order to account for a buffer of 25% including post-processing, errors and re-runs we request a total of 140, 300 and 200 M SBU. respectively. Should there be insufficient resources available for allocating this demand, we would accept a reduction of this total amount and trim down the member size accordingly. We do not require any significant data storage at this time, as the results will be uploaded to BSC storage facilities as part of our Autosubmit workflow.

Table 3 - Summary of planned experiments and total required computing time.

Experiment	period	model configuration	start dates	years	members	# exps	years	CHSY	M CPUh	M SBU
seasonal hindcast demo. (SP)	Q1 2027	EC-Earth4-GCM-SP	21	0.33	25	3	525	1100	0.58	11
seasonal hindcast demo. (M7)	Q2 2027	EC-Earth4-GCM-M7	21	0.33	25	3	525	4000	2.10	40
multi-year hindcast test (SP)	Q3 2027	EC-Earth4-GCM-SP	42	3	5	1	630	1100	0.69	13
multi-year hindcast test (M7)	Q4 2027	EC-Earth4-GCM-M7	42	3	5	1	630	4000	2.52	48
multi-year hindcast test (CC-SP)	Q3 2028	EC-Earth4-CC-ML-SP	42	3	5	1	630	1800	1.13	21
multi-year hindcast test (CC-M7)	Q4 2028	EC-Earth4-CC-ML-M7	42	3	5	1	630	4980	3.14	59
decadal hindcast (SP)	Q2 2028	EC-Earth4-GCM-SP	42	11	10	1	4620	1800	8.32	157
~~decadal hindcast (M7)~~	~~2028~~	~~EC-Earth4-GCM-M7~~	~~42~~	~~11~~	~~10~~	~~1~~	~~4620~~	~~4980~~	~~23.01~~	~~435~~
decadal hindcast (CC-SP)	2029	EC-Earth4-CC-ML-SP	42	11	10	1	4620	1800	8.32	157
~~decadal hindcast (CC-M7)~~	~~2029~~	~~EC-Earth4-CC-ML-M7~~	~~42~~	~~11~~	~~10~~	~~1~~	~~4620~~	~~4980~~	~~23.01~~	~~435~~
2027 total									5.89	**111**
2028 total									12.59	**238**
2029 total									8	**157**

References

Bernardello, R., Sicardi, V., Lapin, V., Ortega, P., Ruprich-Robert, Y., Tourigny, E., and Ferrer, E.: Ocean biogeochemical reconstructions to estimate historical ocean CO2 uptake, Earth Syst. Dynam., 15, 1255–1275, https://doi.org/10.5194/esd-15-1255-2024, 2024.
Bilbao, R., and Coauthors, 2021: Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth. Earth System Dynamics, 12, 173–196, https://doi.org/10.5194/esd-12-173-2021.
Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., Kushnir, Y., Kimoto, M., Meehl, G. A., Msadek, R., Mueller, W. A., Taylor, K. E., Zwiers, F., Rixen, M., Ruprich-Robert, Y., and Eade, R.: The Decadal
CERISE, 2025. D7.4 Time varying Lake cover, Land cover and LAI and extension of CONFESS vegetation data back to 1925 datasets. https://www.cerise-project.eu/sites/default/files/2025-05/CERISE-D7-4-V1.0.pdf
Climate Prediction Project (DCPP) contribution to CMIP6, Geosci. Model Dev., 9, 3751–3777, https://doi.org/10.5194/gmd-9-3751-2016, 2016.
Döscher, R., and Coauthors, 2022: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6. Geoscientific Model Development, 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022.
Friedlingstein, P., and Coauthors, 2023: Global carbon budget 2023. Earth System Science Data Discussions, 15, 5301–5369.
Friedlingstein, P., and Coauthors, 2025: Global carbon budget 2024. Earth System Science Data, 17 (3), 965–1039.
Hermanson, L., and Coauthors, 2022: WMO Global Annual to Decadal Climate Update: A Prediction for 2021–25. Bulletin of the American Meteorological Society, 103, E1117–E1129, https://doi.org/10.1175/BAMS-D-20-0311.1.
Li, H., and Coauthors, 2026: Multimodel Reconstructions and Predictions of the CO2 Fluxes and Atmospheric CO2 Variations. Bulletin of the American Meteorological Society, 107, E1352–E1369, https://doi.org/10.1175/BAMS-D-25-0139.1.
D. Manubens-Gil, C. P., J. Vegas-Regidor. (2016). Seamless management of ensemble climate prediction experiments on HPC platforms. 2016 International Conference on High Performance Computing & Simulation (HPCS), Innsbruck. https://doi.org/10.1109/HPCSim.
2016.7568429
Smith, D. M., and Coauthors, 2020: North Atlantic climate far more predictable than models imply. Nature, 583, 796–800, https://doi.org/10.1038/s41586-020-2525-0.