



Copernicus Climate Change Service



Updated document describing the design of the production chain

C3S_432_Lot3_UFZ - Global Multi-model hydrological Seasonal predictions

Issued by: UFZ / Luis Samaniego

Date: 08/06/2020

Ref: C3S_432_Lot3_UFZ_D1.2.1_v2

Official reference number service contract: 2019/C3S_432_Lot3_UFZ/SC1

This document has been produced in the context of the Copernicus Climate Change Service (C3S). The activities leading to these results have been contracted by the European Centre for Medium-Range Weather Forecasts, operator of C3S on behalf of the European Union (Delegation Agreement signed on 11/11/2014). All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission and the European Centre for Medium-Range Weather Forecasts has no liability in respect of this document, which is merely representing the authors' view.



Contributors

UFZ

Luis Samaniego
Stephan Thober
Oldrich Rakovec

UU

Niko Wanderds
Edwin Sutanudjaja

UKCEH

Alberto Martinez de la Torre
Eleanor Blyth



Table of Contents

1 The ULYSSES production chain	6
1.1 Features	6
1.2 Objectives	6
2 Overview of the ULYSSES modeling chain	7
2.1 SF Workflow	7
2.2 Data formats and meta information	15
2.3 The work flow manager	15
2.4 Quality control	15
3 Description of the subcomponents	17
3.1 Bias correction	17
3.2 General procedure	17
3.3 Detailed procedure	19
3.4 Downscaling	21
3.5 Land surface models	25
4 Calibration and cross-validation	29
4.1 Protocol	29



Executive Summary

This report presents an updated version of the technical proposal that describes the design of the production chain of ULYSSES . It includes details of the several algorithms, workflows, and output data streams. This document also include responses to all request raised by the by the ECMWF described in C3S_432_Lot3_UFZ_1.1.1.



1 The ULYSSES production chain

1.1 Features

The ULYSSES production chain will be operated directly in the ECMWF IT infrastructure (i.e. supercomputers). This will make data transfer redundant, significantly reducing the delay between data processing and delivery. The production chain will be operated with the ECflow suite.

There are five major features of the ULYSSES production chain:

1. We will deliver hindcasts (1993–2019) and operational forecasts (2020–2021) for all key ECVs for a multi-model ensemble containing four state-of-the-art hydrological models for lead times up to six months at a spatial resolution of 0.1° globally.
2. We will provide a flexible, scalable and extendable production chain. This will be achieved by implementing the entire production chain in an ecFlow suite. The ecFlow environment will be highly modularised (i.e., separation of initialisation, creation of initial hydrologic conditions, and operational forecast, downscaling, bias-correction, CDS upload, among others). This allows the inclusion of new datasets, meteorological seasonal forecasts and hydrological models with minimal effort. The ecFlow suite provides an overview of the entire production chain and clear traceability of provenance of data and workflows. ECflow enables a seamless integration into continuous operationalization.
3. We will provide a comprehensive documentation and user support. Our documentation provides various documents such as factsheets, videos and infographics to fully support the multi-tier approach of the CDS helpdesk. This will maximize the user uptake of the products and thus the benefit of the service for society.
4. We will use four diverse hydrological models (HMs) with a consistent setup employing identical geophysical information to reduce uncertainty that might stem from the use of different ancillary datasets (Samaniego et al., 2017). In our system, differences between models are solely due to differences in process descriptions. This allows us to attribute any difference in the outcomes directly to the difference in process description of the models within the multi-model ensemble. We will use existing parametrisations and our experience of delivery of global modeling from previous projects (i.e. earthH2Observe and EDgE) to guarantee timely delivery of ECVs to CDS.
5. All HMs share a common streamflow routing algorithm (mRM) to minimize uncertainties derived from the river network characterisation and its parameterization.

1.2 Objectives

These features will provide an unique wealth of information with respect to hydrological multi-model ensemble seasonal forecasting. ULYSSES will allow C3S to achieve the following objectives:

- Deliver good model performance that enables to provide locally relevant information (Bierkens, 2015).
- Provide hindcasts and operational seasonal forecasts of hydrologic variables listed in Table 3 with lead time of 6 months.

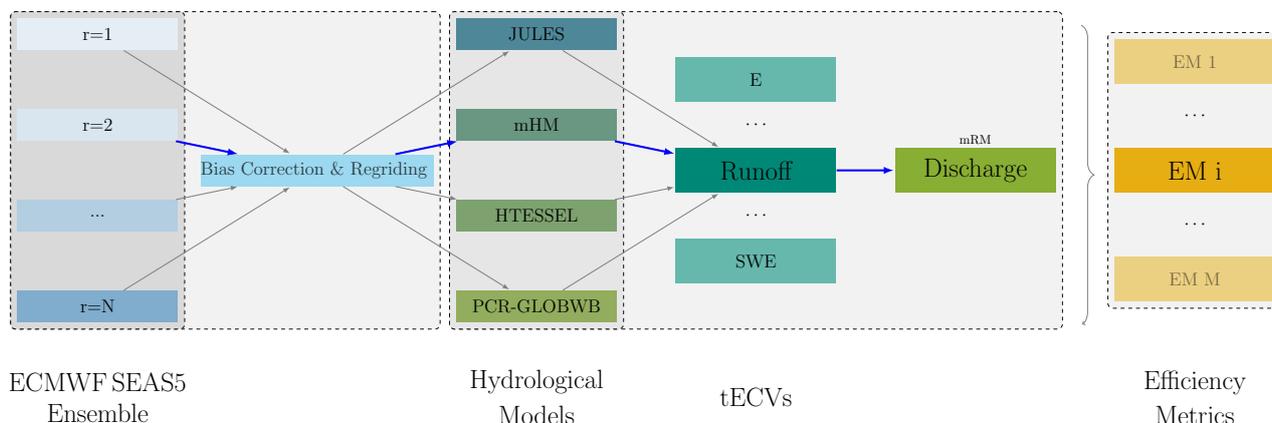


Figure 1: General scheme proposed in ULYSSES to generate an ensemble of hydrologic variables at global scale.

- Generate a consistent set of ECVs based on hydrological simulations driven by ECMWF S5 ensemble forcing data at 0.1° global grid and a list of 50 selected stations.
- Provide detailed user documentation (e.g. product fact sheets, web stories) on the system: from the establishment of the modelling chain, to case studies on how the data may be used.
- Provide the setups for four hydrological models, the routing model and tools to update these setups if new datasets from the other Lots of this ITT become available.
- Provide ecFlow suite for the production chain.

2 Overview of the ULYSSES modeling chain

The EDgE PoC model chain (Samaniego et al., 2019) forms the basis for the ULYSSES production chain (Fig. 1).

This production chain provides a state-of-the-art hydrological multi-model seasonal forecasting system, capturing uncertainty of hydrologic variables. The hydrological models used comprise a wide range of process representations covering a gradient from complex land surface models (HTESSEL, JULES) to global hydrological models (PCR-GLOBWB, mHM), providing a good range of different model types. Operational forecasts of key hydrologic variables based on this multi-model ensemble at a spatial resolution of 0.1° globally are unprecedented. ULYSSES will push the boundary of knowledge and data supplied to end-users and provide a step forward in terms of understanding how uncertain process representations impact hydrological forecasts.

2.1 SF Workflow

The detailed processing steps of the ULYSSES production chain are shown in Fig. 2, which is a continuation of the modeling chain created within the EDgE project. The workflow will be run within an ECflow environment to include all processing steps in one tool allowing an overview of ongoing tasks and provide an efficient error handling. The workflow comprises one-time operations and monthly operations. The one-time operations include:

- the setup of hydrological models and,
- the creation of historical initial states using ERA5-Land* as meteorological forcing.



* NOTE: Following the KO meeting and subsequent discussion, a forcing similar to ERA5-Land should be used within the operational chain, but derived by the contractors from ERA5T — which “provides preliminary data for ERA5 (successor of ERA-Interim) on a daily basis, with a 5-day delay from real time” <https://confluence.ecmwf.int/display/CUSF/Release+of+ERA5T>. — using similar downscaling/re-gridding/correction algorithms. ECMWF will provide support regarding the processing steps/algorithms to be used. This is necessary because ERA5-Land is currently produced operationally with a latency of 2-3months, which is not appropriate to provide initial conditions for the operational hydrological seasonal forecasting system proposed in ULYSSES . However, historical ERA5-Land could be used for the calibration/ testing of hydrological models if work has already started to avoid unnecessary delays in the project as it is expected that the ERA5-land-post-processed data to be identical or very similar to ERA5-land.

The monthly operations are composed of:

- the updating of initial hydrologic states,
- downscaling and bias correction of meteorological seasonal forecasting data,
- the generation of the ensemble of hydrologic variables,
- quality checking and the uploading of the data to the climate data store (CDS).

Details of the monthly operations are shown in Fig. 1.

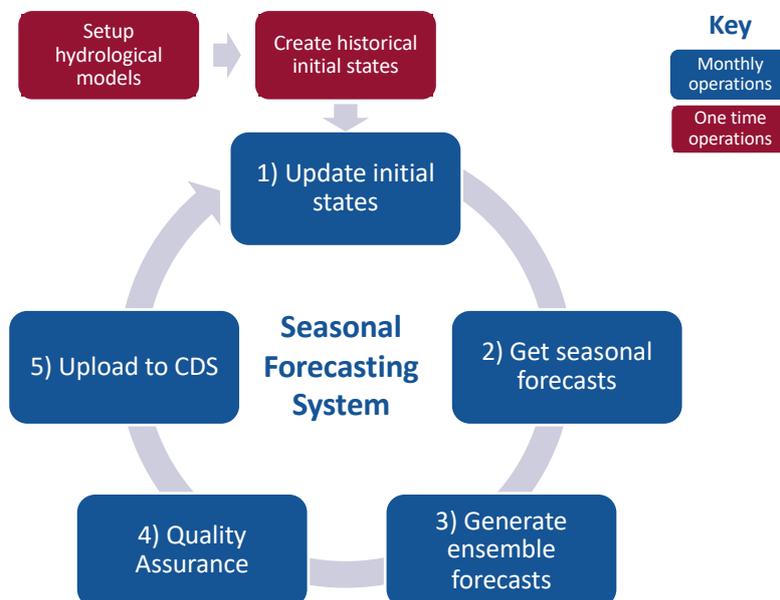


Figure 2: General workflow scheme to be followed in the ULYSSES and its five stages, shown in blue. Red boxes show the operations that are only carried out once during the project. The outputs of the modelling chain (ECVs) are described in (Table 3). This figure is based on Samaniego et al. (2019).

We will run the ULYSSES production chain for each of the four hydrological models individually (HMs: HTESSEL, JULES, mHM, and PCR-GLOBWB; see Section below for detailed model description). The modelling protocol is defined as follows:



1. The first step is a one-time operation and contains the setup of the HMs (Fig. 2). We will use the geophysical datasets listed in Table 1 to create model ancillary files at a global 0.1° spatial resolution compatible with the domain of the ERA5-Land dataset. These are freely available datasets that are widely used in the hydrologic community. For each HM, where model parameters have to be derived from these datasets (e.g., porosity from SoilGrids), standard transfer functions or look-up tables will be shared and applied to these datasets. The datasets listed in Table 1 are only used initially and we will provide software tools to easily update the model ancillary files to the new dataset produced within Lot 2 of this ITT once they become available. The historical forcing data could be either extracted from the CDS or directly from ECMWF/MARS archiving system. Note the earlier comment that the ERA5-Land dataset will not be available directly from CDS/MARS for the operational running of the system, and should be re-created by the contractors as an additional processing loop of downscaling/regridding/correction.

Table 1: List of third-party datasets used as inputs for the HMs in ULYSSES

Dataset Name	Description	Source
GlobCOVER	Global land cover data v2.2	http://due.esrin.esa.int/page_globcover.php
SoilGrids250m	Global 3D Soil Information System at 250 m resolution	http://www.isric.org/index.php/explore/soilgrids
GMTED2010	Global 7.5 Arc-Second Elevation	https://topotools.cr.usgs.gov/gmted_viewer/gmted2010_global_grids.php
GLiM	Global hydrogeology map	http://dx.doi.org/10.1594/PANGAEA.788537

2. The second step of our production chain is a one-time operation devoted to creating the historic initial conditions (Fig. 2). All meteorological forcing variables (Table 2) from the ERA5-Land dataset will be generated “on-the-fly” during operation as indicated above. All models will be run with ERA5-Land from 1993 to 2019 for model spin-up. It should be noted that, the operational chain will be built using a dataset derived from ERA5T (i.e., timely updates of the ERA5 climate reanalysis) and equivalent to ERA5-Land as ERA5-LandT might not be available in due time for the operational activities of ULYSSES.

Restart files will be stored for the end of this simulation period. These restart files will be used to perform a second run from 1993 to 2019. Within this second run, restart files will be stored at the end of each month and will follow the naming convention:

`<model_name>_restart_<yyyy>_<mm>_<dd>.<file_type>`

The file type will be model dependent and cannot be further harmonised within ULYSSES (the file type is netCDF for all HMs but PCR-GLOBWB). These restart files will be used to generate the hindcast data. The last restart file at 31.12.2019 will be used for the operational forecasts.

For each model, the multiscale Routing Model (mRM, [Thober et al., 2019](#)) will be used as post-processor to transform runoff to discharge. mRM will be spun-up for each HM in the same way as described above and all restart files of the HMs will be complemented by restart files of mRM using the following naming conventions:

`mRM_<model_name>_restart_<yyyy>_<mm>_<dd>.nc.`



In the following, we refer to a model run as a simulation of a HM in conjunction with mRM as postprocessor for simplicity.

For the second simulation and all following simulation, we will store all required output variables for all models. The list of all output variables is contained in Table 3.

Table 2: List of forcing variables provided in the CDS for ERA5-Land at 0.1° spatial resolution and hourly temporal resolution and ECMWF S5 at 1° spatial resolution and 6-hourly temporal resolution. Variables marked by an asterisk are derived variables from those available in the CDS.

Variable	Unit	Standard long name
pr	$\text{kg m}^{-2} \text{s}^{-1}$	precipitation
tas	K	near-surface air temperature
tasmax*	K	daily maximum near-surface air temperature
tasmin*	K	daily minimum near-surface air temperature
ps	Pa	surface air pressure
sfcWind*	m s^{-1}	near-surface wind speed
huss*	1	near-surface specific humidity
rsds	W m^{-2}	downward shortwave radiation
rlds	W m^{-2}	downward thermal radiation

The output files will follow the following name convention:

`<model_name>_<forcing_dataset>_<realisation>_<start_date>_<end_date>.nc`

For the output generated by ERA5-Land, annual files will be stored. The *realisation* refers to the ECMWF S5 ensemble member and will be omitted for the ERA5-Land simulations. Two modules of the mesoscale Hydrologic Model mHM will be used within ULYSSES as stand-alone tools. A simple energy balance module was developed to compute skin temperatures for mHM (Zink et al., 2018) that will be used as post-processor for mHM and PCR-GLOBWB. Both of these models typically do not compute surface temperature internally (Tsurf in Table 3). The second one is a module to calculate potential evapotranspiration (PET). PET can be calculated at different levels of complexity. Within mHM, the simple temperature-based Hargreaves-Samani scheme (Hargreaves and Samani, 1985), the radiation-based Priestley-Taylor scheme (Priestley and Taylor, 2003), and the classical Penman-Monteith scheme (Monteith, 1981), are available. Within hydrological modelling, the Hargreaves-Samani equation is a well-accepted choice. Its advantages in comparison to the other schemes is that the input data is less uncertain. It uses daily average temperature and the diurnal temperature range, whilst the other two schemes rely on radiation fluxes. The latter are highly dependent on cloud cover that is more challenging to observe and simulate than temperature. Since ULYSSES is targeted to provide a hydrological forecasting system, the Hargreaves-Samani scheme will be used to calculate PET as output.

3. The first monthly operation step of our modelling chain is the update of the initial hydrologic states (Fig. 2). For this purpose, the meteorological variables starting from the date of the latest restart file to the beginning of the operational seasonal forecast are going to be derived within the ULYSSES processing chain from ERA5T dataset (see details above and the ECMWF recommendations in this respect provided in D1.1.1). Subsequently all LSMs/HMs are run for this period. Restart files will be stored at the end of the simulation period for the initialisation of the forecast



Table 3: Essential Climate Variables (ECVs) produced within the ULYSSES production chain. All variables are produced on a global scale at a resolution of 0.1°.

Symbol	Long name	Standard name	Unit (CF v1.7)	Definition	Positive direction
Tsurf	Average surface temperature	surface_temperature	K	Average of all vegetation, bare soil and snow skin temperatures	
P	total precipitation	precipitation_flux	kg m ⁻² s ⁻¹	Average of total precipitation (Rainf+Snowf)	downwards
ET	Total evapotranspiration	water_evaporation_flux	kg m ⁻² s ⁻¹	Sum of all evaporation sources, averaged over a grid cell	downwards
PET	potential evapotranspiration	water_potential_evaporation_flux	kg m ⁻² s ⁻¹	The flux as computed for evapotranspiration but with all resistances set to zero, except the aerodynamic resistance	downwards
SWE	Snow water equivalent	snow_water_equivalent	kg m ⁻²	The amount of liquid water contained within the snow pack	Into grid cell
Qsm	snowmelt	surface_snow_melt_flux	kg m ⁻² s ⁻¹	Average liquid water generated from solid to liquid phase change in the snow	solid to liquid
SM	Percentage of water wrt the available volume	Total volumetric soil moisture	%SM or m ³ m ⁻³	Volumetric soil moisture content in the soil layers at the end of each model time step	Inside grid cell
Qr	Total runoff	runoff_flux	kg m ⁻² s ⁻¹	Average total liquid water draining from land	into grid cell
Qp	Discharge (gauge level)	N.A	m ³ s ⁻¹	Water volume leaving the cell	downstream
Q	Gridded river discharge	N.A	m ³ s ⁻¹	Water volume leaving the cell	downstream

starting in the next month. The model output files will be appended to that of the given year. ECMWF will provide support regarding the processing/algorithms to produce a dataset equivalent to ERA5-Land. The ERA5T dataset will be extracted from ECMWF/MARS storage as this is more efficient than downloading it from CDS/MARS.

4. The second monthly operation step of our modelling chain is the preprocessing of the seasonal forecast data (Fig. 2), that is divided into three subtasks:

(a) Retrieve of all meteorological variables (Table 2) for ECMWF S5 for a lead time of 180 days



starting at the initialization date from ECMWF/MARS storage as soon it is produced. It should be noted that S5 data is produced around the day 5th of each month but not made publicly available at the CDS/MARS before day 10th. Retrieving from within ECMWF IT infrastructure will be much more efficient and faster than if retrieved from cloud-based CDS/MARS infrastructure. This procedure is beneficial for ULYSSES because it provides a time buffer for the processing of the hydrological seasonal forecasts before they are to be published on the CDS/MARS, hence giving more time for bug-fixes if delays/issues arise.

- (b) Downscaling of the ECMWF S5 from the 1° to 0.1° using the method described in Section 3. Alternatively, the same downscaling/regridding algorithm used to process ERA5T data could be used to keep consistency across all processing steps. The contractor will evaluate both methods and take a decision once the hindcast operations start.
- (c) Bias-correcting the downscaled ECMWF S5 data with ERA5-Land using an expanded method of that used within ISI-MIP ([Hempel et al., 2013](#)). See Section 3 for the details.

The forcing files will follow the name convention:

<model_name>_<realisation>_<resolution>_<start_date>_<end_date>.nc.

This step is carried out for the operational forecasts immediately after they become available. We will also do this for the hindcast data over the life time of the project.

- 5. The third monthly operation step of our modelling chain is the generation of the ensemble forecast (Fig. 2). The forcing data created within the previous step will be used to force the HM and store all output variables following in files with the following name convention:

Output_<model_name>_<realisation>_<resolution>_<start_date>_<end_date>.nc.

In total, 51 output files containing data for 180 days forecast will be created.

- 6. The fourth step of the monthly operations is the evaluation and quality control (EQC) of the created data (Fig. 2). We will use different tools to assess the quality of the reference dataset and the seasonal hindcasts and forecasts. Details of the EQC are provided below.
- 7. The fifth and final step of our modelling chain is the upload of the generated forecast data to the CDS. For this purpose, files containing abstracts, detailed descriptions of dataset, variables, etc., in formats (e.g. .yaml) required by the CDS will be adapted for each HM.

A detailed overview of the processing steps in the production chain and associated time line is provided in Table 4 and 5 for the monthly update of the reference run and the operational forecast, respectively. The production of the monthly update of the reference run needs to be carried out before the operational seasonal forecasts of all ECVs can be produced because the monthly update of the reference run creates the initialisation files for the operational forecasts. We assume that the latency of the operational reanalysis product is five days and is available for a given month m before the operational seasonal forecast is available for that month. The production time is scheduled to not exceed 10 hours with the production time of the model outputs to take place within the first 8 hours. This will provide us enough time to react to errors in the production chain.



Table 4: Production time line of the monthly updates of the **historical and operational** reference (i.e. restart) runs based on the ERA5-Land forcing data with associated time requirements in quarters of an hour. The provided time refers to the processing step when it is finished.

Time	Processing Step	Description
T_r	Real-time reanalysis product for the end of month m is available.	
$T_r + 0.25$ h	Copy files containing forcing variables (Table 2) to working directory.	
$T_r + 0.50$ h	Quality check 1 (QC1)	QC1 checks for missing values, not a number values and whether all forcing variables are within a physically plausible range (e.g., non-negative precipitation).
$T_r + 0.75$ h	Downscale ERA5T meteorological forcing to 0.1° spatial resolution.	
$T_r + 1.00$ h	Create meteorological forcings (ERA5-LandT) for predefined subdomains.	
$T_r + 1.25$ h	Bias-correct meteorological forcings (ERA5-LandT) on predefined subdomains.	
$T_r + 1.50$ h	QC2	QC2 performs the same checks as QC1, but for the bias-corrected and downscaled meteorological forcing data.
$T_r + 2.25$ h	Run all hydrological models on subdomains.	We will request as many compute cores as are needed to run all models in at most one hour. We expect to require approximately 100 compute cores. Please note that the number of compute cores will not be equally distributed among the hydrologic models.
$T_r + 2.5$ h	Run the multiscale Routing Model (mRM) on subdomains.	This step and the previous one generate the initial conditions for the next forecasting step.
$T_r + 2.75$ h	QC3	QC3 checks whether all HMs and mRM finished and that the output of all models is within reasonable physical bounds (e.g. no negative snow water equivalent).
$T_r + 3.25$ h	Join model outputs that has been created for subdomains into one file per ensemble member.	
$T_r + 3.50$ h	Create meta-data files and pseudo-manifest files for data upload to CDS.	
$T_r + 3.75$ h	QC4	QC4 checks whether all subdomain files have been correctly joined and all meta-data files and pseudo-manifest files have been created.
$T_r + 4.00$ h	Upload data to CDS.	
$T_r + 4.25$ h	Copy data to archiving system.	
$T_r + 4.50$ h	QC5	QC5 validates the succesful upload to CDS and copy to the archiving system via checksums.



Table 5: Production time line of monthly **seasonal operational forecasts** with associated time requirements in quarters of an hour. The provided time notes when the processing step is finished.

Time	Processing Step	Description
T_o	Operational ECMWF S5 forecast is available.	It is assumed that all required initialization files are available.
$T_o + 0.25$ h	Copy files containing forcing variables (Table 2) to working directory.	
$T_o + 0.50$ h	Quality check 1 (QC1)	QC1 checks for missing values, not a number values and whether all forcing variables are within a physically plausible range (e.g., non-negative precipitation).
$T_o + 0.75$ h	Downscale ECMWF S5 meteorological forcing to 0.1° spatial resolution.	
$T_o + 1.00$ h	Create ECMWF S5 meteorological forcings for predefined subdomains.	
$T_o + 1.25$ h	Bias-correct ECMWF S5 meteorological forcings on predefined subdomains.	
$T_o + 1.50$ h	QC2	QC2 performs the same checks as QC1, but for the bias-corrected and downscaled meteorological forcing data.
$T_o + 5.50$ h	Run all hydrological models on subdomains using the initial conditions derived in Table 4, see $T_r + 2.5$ h.	We will request as many compute cores as are needed to run all models in at most four hours. Given the estimate of SBUs in this document, we expect to require approximately 3 200 compute cores. Please note that the number of compute cores will not be equally distributed among the hydrologic models.
$T_o + 6.5$ h	Run the multiscale Routing Model (mRM) on subdomains.	
$T_o + 6.75$ h	QC3	QC3 checks whether all HMs and mRM finished and that the output of all models is within reasonable physical bounds (e.g. no negative snow water equivalent).
$T_o + 7.25$ h	Join model outputs that has been created for subdomains into one file per ensemble member.	
$T_o + 7.50$ h	Create meta-data files and pseudo-manifest files for data upload to CDS.	
$T_o + 7.75$ h	QC4	QC4 checks whether all subdomain files have been correctly joined and all meta-data files and pseudo-manifest files have been created.
$T_o + 8.75$ h	Upload data to CDS.	
$T_o + 9.75$ h	Copy data to archiving system.	
$T_o + 10.00$ h	QC5	QC5 validates the succesful upload to CDS and copy to the archiving system via checksums.



2.2 Data formats and meta information

The metadata, the file naming convention, and all datasets generated in ULYSSES will be in NetCDF Climate and Forecast (CFv1.7) standard (CF, 2007) and will be compliant with the Common Data Model of C3S. The metadata of the hydrological models will follow the Assistance for Land-surface Modelling activities (ALMA) convention (Bowling and Polcher, 2001) implemented successfully in previous European projects (e.g., FP7-WATCH, ISI-MIP, Earth2Observe). Metadata associated with geospatial data text and image will conform to the Dublin Core Convention (Initiativ, 2020), using a vocabulary of fifteen properties for use in resource description. Metadata will be contained in the header of the netCDF output files. Alongside this, full documentation will be prepared in the prototyping phase and provided to the CDS in advance of data production, containing abstracts, detailed descriptions of dataset, variables, etc., in formats (e.g. .yaml) required by the CDS, together with manifest files. This will provide assurance that the data within the production phase can be incorporated seamlessly into the CDS, providing full information to end users.

2.3 The work flow manager

ecFlow is a work flow manager that facilitates the execution of a large number of programs (with dependencies on each other and on time) in a controlled environment (<https://confluence.ecmwf.int/display/ECFLOW>). It provides good restart capabilities with a reasonable tolerance for hardware and software failures. ECMWF uses ECflow to run all their operational suites across a range of platforms.

We will create an ECflow suite to handle all tasks related to the modeling chain (Section 2.1). The ECflow environment will be highly modularised and small tasks will be created for the individual processing steps that make up the modelling chain (i.e., separation of initialisation, creation of initial hydrologic conditions, and operational forecast, downscaling, bias-correction, HM execution, CDS upload, among others). The ECflow suite thus provides an overview over the entire production chain and a clear traceability of provenance of data and workflows. This suite will also provide a comprehensive error handling. If for any reason, a part of the production chain fails, an error will be raised to inform the modeler. ECflow allows the modeller to easily restart the production chain from any given subtask without the requirement to rerun the entire production chain. This will help to keep the delay between data acquisition and delivery as short as possible, even if errors occur.

The modular nature of the ECflow suite enables the use of new datasets, meteorological seasonal forecasts, and HMs with minimal effort and facilitates a seamless integration of the ULYSSES system into continuous operationalization.

The production chain will allow portability of input data to other models and users, once the required data is uploaded in the Climate Data Store (CDS) as requested by ECMWF in C3S_432_Lot3_UFZ_-1.1.1. Details on how this tool is built and its user manual will be presented in the final deliverable of ULYSSES and will be available in Confluence.

2.4 Quality control

The quality control comprises two kinds of quality, technical quality and skill-based quality. The technical quality control checks that all metadata has been created correctly within the files and that the upload of the data to the CDS was successful. The skill-based quality computes skill measures for the created forecasts and issues warning if low skill is detected compared to climatology. At month $t+1$,

the following efficiency measure (Fig. 1) is calculated for forecasts at month $t+1, \dots, t+5$. Let k be the number of ensemble members

$$n_j(t) = \frac{\sum_{i=1}^k 1[v_i(t) < P^{0.5}(v, t)] + \sum_{i=1}^k 1[v_i(t) > P^{99.5}(v, t)]}{k} 100, \quad (1)$$

where n_j denotes for a specific grid cell j the percentage of ensemble members that are outside of the 99% confidence interval of the climatology. Within equation 1, k denotes the number of ensemble members (255 for the entire ensemble), v the analysed ECV, $P^x(v, t)$ the percentile x for variable v and month t . For each ECV listed in Table 3, we create maps showing the percentage of outliers for each grid cell. During 99% of the time, this efficiency measure should be less than one, indicating that all forecasts have been created and no extreme values are forecasted. If this number is higher than one, then a warning is issued because an unusual situation is occurring. In areas with particularly high numbers, we will investigate whether an extreme event is forecasted or whether a mistake in the production chain happened. For example, if one HM failed for some reason, then n will be at least 20%.

Additionally, skill measures will be calculated for the forecasts created within previous months. At month $t+1$, the forecasts of month t for this month will be compared against the simulations based on forcing data equivalent to ERA5-Land. In this case, ERA5-Land based forcings (see note above on how this data will be derived in ULYSSES) are used as a reference. This skill metrics comprise for each ECV individually the bias, and spread. This procedure is applied to all forecasts of that month for lead times up to six months. Let, l the lead time and m the month, then the bias is defined for a given model as:

$$\epsilon_{ilm} = \frac{1}{n} \sum_{j=1}^n (SF_{ilm}^j - ERA5L_{im}) \quad (2)$$

Given the above notation, the spread is defined as

$$s_{ilm} = P^{75}(SF_{ilm}) - P^{25}(SF_{ilm}), \quad (3)$$

where P^x denotes the x -th percentile of the ensemble. All the measures are used to quantify the skill of the ensemble forecast. The correct interpretation of this skill can only be conducted if the models are also compared against observations. For example, hydrological models with a high auto-correlation in their state variables tend to have high skill, but do not necessarily have the lowest error when compared against observations (Wanders et al., 2019). Hence, we will additionally compare simulated fluxes and state variables obtained from the ERA5-Land reference run against observations with the iLAMB tool used within earth2Observe (Luo et al., 2012). By using this international standard comparison with data, we will enable model development to be carried out externally to ULYSSES, which will eventually benefit the C3S system.

We will guarantee full reproducibility of the ULYSSES production chain by making all codes, including the ECflow suite, available via a open access GIT repository for future applications. Individual codes are written in Fortran (mHM (www.ufz.de/mhm), JULES, HTESEL, bias-correction, downscaling, mRM), or Python (PCR-GLOBWB (https://github.com/UU-Hydro/PCR-GLOBWB_model)). Feedback from the



Evaluation and Quality Control (EQC) framework can be integrated into the ULYSSES quality control outlined in this Section.

3 Description of the subcomponents

3.1 Bias correction

It is recognised that climate models are an “approximation of the real climate system and have different simplifications resulting in biases of the simulated climate when compared to the observed one”. The same is true for dynamical seasonal forecasting models. Consequently, they cannot “always” be used directly to drive HMs, although this is the ultimate aim of Ulysses.

To overcome this problem, we propose the use of a “bias correction” (BC) of the modelled outputs towards the observed climatology. There are a number of statistical approaches to undertake this correction (e.g., quantile-mapping technique, cumulative distribution function-transform, among others). It should be noted, however, that bias adjustments to raw climate outputs introduce uncertainties, and hence should be used with care. In Ulysses, we will expand the technique used in the ISI-MIP project for precipitation and temperature to all variables listed because it is a well-established method and preserves the trend of the original data. This method is separated into two steps. First, long-term mean values are corrected. Second, the daily variabilities are corrected. An additive error model is assumed for temperature. In this case, the normally distributed residual errors are corrected using a quantile-mapping approach. A multiplicative error model is used for precipitation, where a gamma distribution is assumed to adjust precipitation amounts during wet months with an additional adjustment for the dry/wet month frequency. We will extend the additive error model to correct the forcing variables other than precipitation and temperature. The bias-correction method has been developed to adjust daily values. The hydrological models HTESSEL, and JULES require sub-daily values as forcing. A multiplicative scaling factor will be used to adjust these sub-daily values with respect to the corrected daily value. ERA5-Land will be used as a reference dataset to correct against. It is important to note that the bias correction will be applied after the downscaling to make full use of the variability entailed in the ERA5-Land reference dataset. Moreover, the downscaling of ECMWF S5 will use the same algorithm as the one applied to downscale ERA5 to the 0.1°, to keep the consistency of both data sets.

The bias-correction method applied in the Ulysses project closely follows the method proposed by Lange (2019). This method has been developed for climate change projections and is modified here for seasonal forecasts.

3.2 General procedure

The bias-correction consists of eight steps, that are used to correct the variables listed in Table 6 at the daily time scale.



Table 6: List of forcing variables to be Bias Corrected

Standard long name	Short Variable	Unit
Daily mean near-surface relative humidity	rh	Pa Pa ⁻¹ or %rh
Daily mean precipitation	rainfall	kg m ⁻² s ⁻¹
Daily mean snowfall flux	snowfall	kg m ⁻² s ⁻¹
Daily mean sea-level pressure	sp	Pa
Daily mean surface downwelling longwave radiation	strd	W m ⁻²
Daily mean surface downwelling shortwave radiation	ssrd	W m ⁻²
Daily mean near-surface wind speed	wind	m s ⁻¹
Daily mean near-surface air temperature	t2m	K
Daily maximum near-surface air temperature	t2mmax	K
Daily minimum near-surface air temperature	t2mmin	K

We follow the nomenclature of Lange (2019) and define the historic observations as x_{hist}^{obs} and historic simulations as x_{hist}^{sim} . Lange (2019) also considered simulations during future periods and focussed on correcting future values, but this is not the case for seasonal forecasts.

The outline of the bias-correction steps are as follows.

1. Rescale x_{hist}^{obs} and x_{hist}^{sim} for ssrd to the interval (0, 1).
2. Replace missing values in x_{hist}^{obs} and x_{hist}^{sim} for snowfall ratio (i.e., the ratio between snowfall and rainfall, that is not defined in case of dry days) by random sampling from available values.
3. Detrend x_{hist}^{obs} and x_{hist}^{sim} for sp, strd and t2m.
4. Randomize values beyond thresholds for bounded variables (e.g., rainfall, snowfall).
5. Use parametric quantile mapping to adjust the distribution of values in x_{hist}^{sim} to that of x_{hist}^{obs} for all variables. For bounded values, also adjust the frequency of values beyond thresholds.
6. Add trend for sp, strd and t2m.
7. Rescale ssrd.

A summary of all variables and the procedure used is provided in this table.

Table 7: Parameters used for each variable

Variable Short name	Lower bound	Lower threshold	Upper bound	Upper threshold	Distribution	Detrend
rh	0	0.01	100	99.99	beta	no
rainfall	0	0.1/86400	-	-	gamma	no
snowfallratio	0	0.0001	1	0.9999	beta	no
psl	-	-	-	-	normal	yes
strd	-	-	-	-	normal	yes
ssrd	0	0.0001	1	0.9999	beta	no
wind	0	0.01	-	-	weibull	no
t2m	-	-	-	-	normal	yes
tasrange	0	0.01	-	-	Rice	no



3.3 Detailed procedure

Indexes for location, month and lead time are dropped. All statistics (e.g., distribution function, trends, etc.) are estimated for each location, month, and lead time separately.

3.3.1 Step 0: Preparation

1. Calculate daily values as arithmetic mean for all variables.
2. Calculate snowfallratio from rainfall and snowfall.
3. Relative humidity is calculated as Q_{air}/Q_{sat} . Q_{sat} is calculated following the procedure used in the statistical downscaling (see Section on Specific humidity).
4. Calculate temperature range as: $tasrange = \max(t2m) - \min(t2m)$
5. psl is calculated from sp using the same lapse-rate correction used in the statistical downscaling.

3.3.2 Step 1: rescaling ssrd

Estimate annual cycle of upper bounds of surface solar radiation at daily resolution as

$$m_d = \max_y(ssrd_d^y), \quad (4)$$

$$b_d = \frac{1}{31} \sum_{d-15, d+15} (m_d), \quad (5)$$

where d is the day of the year and y is the year index. Please note that January 1st is the day after December 31st to calculate continuous annual cycle. In other words, b_d is the running 31 day window mean of multiyear daily maximum values.

All values are then scaled according to

$$x_d^y = x_d^y / b_d \quad (6)$$

In step 7, the inverse scaling is conducted.

3.3.3 Step 2: Fill missing values in snowfallratio

Snowfall ratio is not defined on days with no precipitation. For these days a random value from the available numbers is chosen. In detail, p is drawn from the interval $[0, 100] \in \mathbf{R}$ with uniform probability and the p th percentile of the available values is chosen. This preserves the distribution of values.

3.3.4 Step 3: Detrending

Linear trends in sp , $strd$ and $t2m$ are removed from both x^{obs} and x^{sim} . Those trends are restored in step 6. Trend lines t^{obs} and t^{sim} are estimated as linear regression of annual values. Let x_d^y be the value at day d and year y and let t^d be the value for the corresponding year y . The trends line are shifted such that $\sum t^y = 0$. Values are then detrended by subtracting the trend line.



3.3.5 Step 4: Bounded values

This step only applies to variables that have lower or upper bounds. All values below the lower threshold α are replaced by random numbers randomly drawn from the open interval (a, α) . Conversely, all values above the upper threshold β are replaced by random numbers in the open interval (β, b) .

3.3.6 Step 5: quantile mapping

This step is the core of the bias correction method. For unbounded variables, it consists of a parametric quantile mapping of x^{sim} to x^{obs} . For variables with at least one bound, it consists of a bias adjustment of the frequency of values beyond the threshold and a parametric quantile mapping of all other values.

Frequency of values beyond a threshold are adjusted as follows. For variables with a lower bound a and a lower threshold α , let P^{obs} and P^{sim} denote the relative frequencies of values smaller than α in x^{obs} and x^{sim} , respectively. If x^{sim} is of length n , then the lowest nP^{obs} values of x^{sim} are set to a . Conversely, for variables with an upper bound b and upper threshold β , the relative frequency of values greater than β is corrected by setting the highest nP^{obs} values to b . All other values are bias-corrected using a parametric quantile mapping.

Following Lange (2019), distributions used for parametric quantile mapping are the beta distribution for bounded climate variables (rh, snowfallratio, ssrd), the gamma distribution for rainfall, the normal distribution for unbounded climate variables (sp, strd, t2m), the Weibull distribution for wind, and the Rice distribution for tasrange. For unbounded variables, these distributions are fitted to all values. For bounded variables, these distributions are fitted to all values that are not beyond the thresholds. We use a simple gradient descent optimizer and the RMSE objective function to fit these distributions. Let \hat{F}^{obs} and \hat{F}^{sim} be the fitted cumulative distribution functions.

The parametric quantile mapping is very simplified in comparison to that used in the ISIMIP3 presented by Lange (2019) because no future period exist. The bias corrected value is calculated as

$$x_{bc}^{sim} = \hat{F}_{obs}^{-1}(\hat{F}^{sim}(x^{sim})). \quad (7)$$

3.3.7 Step 6: Apply trends

Apply the observed trends to the bias-corrected values for the variables sp, strd and t2m.

3.3.8 Step 7 : Inverse rescaling of ssrd

Calculate actual surface short-wave downward radiation values by

$$x_d = x_d * b_d, \quad (8)$$

with the annual cycle of daily maximum values estimated in step 1 for the observed values.

3.3.9 Postprocessing

1. Using the bias-corrected t2m and sp, calculate saturated specific humidity Q_{sat} using the equations outlined in the downscaling algorithm. Specific humidity is then calculated as:



$$Qa = Qsat * \frac{rh}{100} \quad (9)$$

2. Calculate rainfall and snowfall from rainfall and snowfall ratio.
3. Calculate surface pressure from psl using the lapse-rate correction applied in the statistical downscaling.
4. Calculate sub-daily values by preserving the ratio between the original sub-daily value and the original daily mean value. In detail,

$$x_h = x_d \frac{x_h^o}{x_d^o}, \quad (10)$$

where x_h, x_d, x_h^o, x_d^o is the hourly value, the bias-corrected value, the original hourly value and the mean value calculated in step 0, respectively.

3.4 Downscaling

The methods described in the following use methods described in the WATCH project. For a full-length description with references, see the technical description available under:

<http://www.eu-watch.org/media/default.aspx/emma/org/10376311/WATCH+Technical+Report+Number+22+The+WATCH+forcing+data+1958-2001+A+meteorological+forcing+dataset+for+land+surface-+and+hydrological-models.pdf>

All indexes for time and space are dropped because all operations will be executed for each time step, may it be 1-hourly or 6-hourly and grid cell.

As suggested in the Terms of Reference, the bilinear interpolation method will be used for the disaggregating all forcing data sets described in Table 2. It should be noted, however, that the ULYSSES Team may be forced to propose another disaggregation method if the suggested approach is found, for instance, to be not-fit-for-purpose, or that the resulting fields are too smooth and poorly describing the orographic effects observed in variables such as precipitation and temperature. It is also plausible that the bilinear interpolation method induce bias, which in turn, lead to poor forecasting skills. If any of this case is true, the external drift krigging method used (Samaniego et al., 2019) in the EDgE project will be used. See ECMWF recommendation described in C3S_432_Lot3_UFZ_1.1.1.

To ensure consistency in the modelling chain, the downscaling technique used for ERA5T and ECMWF-S5 datasets will be the same. Either the method proposed here or the one used to generate ERA5-Land will be used. The decision on which one should be taken will be decided by the contractor in due time when the first hindcast experiments are performed.

3.4.1 Nomenclature for the bilinear interpolation

Bilinear interpolation within a square of Lat/Lon values is given on Wikipedia (https://en.wikipedia.org/wiki/Bilinear_interpolation). We follow this nomenclature here:



Let $Q_{11} = f(x_1, y_1)$, $Q_{12} = f(x_1, y_2)$, $Q_{21} = f(x_2, y_1)$, $Q_{22} = f(x_2, y_2)$ the values of the variable Q at two different locations (x_1, y_1) and (x_2, y_2) respectively. The value of Q at (x, y) can then be calculated as follows:

$$Q_{xy} = \frac{1}{(x_2 - x_1)(y_2 - y_1)} [x_2 - x \quad x - x_1] \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} y_2 - y \\ y - y_1 \end{bmatrix} \quad (11)$$

This interpolator function will be applied to all forcing variables and is denoted as the $\mathcal{I}\langle Q \rangle = Q_{xy}$ operator.

3.4.2 Wind

Let W_o be the wind field at the source resolution to be interpolated. This can be calculated from longitudinal and meridional wind speed using the Pythagorean theorem. The interpolated wind speed is then calculated as follow:

$$W_i = \mathcal{I}\langle W_o \rangle \quad (12)$$

3.4.3 Temperature

Let T_o be the source temperature field to be interpolated. The interpolated temperature field T_i is then calculated as follows:

$$T_s = T_o + 0.0065Z_o, \quad (13)$$

$$T_i = \mathcal{I}\langle T_s \rangle - 0.0065Z_i, \quad (14)$$

where Z_o and Z_i are the elevation fields of the source and interpolated resolution in [m], respectively.

3.4.4 Surface Pressure

Let S_o be the source surface pressure field to be interpolated. The interpolated surface pressure field is the calculated as follows:

$$S_s = \frac{S_o}{(T_o/T_s)^{g/-\gamma R}}, \quad (15)$$

$$S_i = \mathcal{I}\langle S_s \rangle \left(\frac{T_i}{T_i + 0.0065Z_o} \right)^{g/-\gamma R}, \quad (16)$$

where, $g = 9.81 \text{ m s}^{-2}$, $\gamma = -0.0065$ (is the negative value of the environmental lapse rate), R is the gas constant of air $287 \text{ J kg}^{-1} \text{ K}^{-1}$.



3.4.5 Specific humidity

The strategy for calculating specific humidity is to keep the relative humidity fixed and apply it to saturated vapour pressure at the high-resolution grid.

Let T_o be the source temperature field. First, a transformation to degree Celsius is applied

$$T_c = T_o - 273.15 \quad (17)$$

Next, the saturated vapour pressure is calculated using the following equation:

$$E_o = A \exp(B - T_c/D)T_c/(T_c + C), \quad (18)$$

where the parameters A, B, C, D are chosen according to T_c . When T_c is above 0°C , then $A = 6.1121, B = 18.729, C = 257.87, D = 227.3$, else $A = 6.1115, B = 23.036, C = 279.82, D = 333.7$.

Next, an enhancement factor is calculated:

$$F_e = 1. + X + (S_o/100.)(Y + (ZT_c^2)), \quad (19)$$

where the parameters X, Y, Z are chosen according to T_c . When T_c is above 0°C , then $X = 0.00072, Y = 3.210^{-6}, Z = 6.910^{-10}$, else $X = 0.00022, Y = 3.8310^{-6}, Z = 6.410^{-10}$.

Next, an enhanced saturated vapour pressure is calculated.

$$E_e = E_o F_e \quad (20)$$

The saturated specific humidity at the source resolution Q_o is then calculated as:

$$Q_{s_o} = \frac{0.62198E_e}{S_o/100. - 0.37802E_e} \quad (21)$$

The relative humidity at the source resolution is then calculated as

$$R_o = 100 \frac{Q_{a_o}}{Q_{s_o}}, \quad (22)$$

where Q_{a_o} is the specific humidity at the source resolution.

Next, the relative humidity is interpolated bilinearly to the target resolution.



$$R_i = \mathcal{I}\langle(R_o) \rangle \quad (23)$$

Next, the process that has been applied at the source resolution is reversed. In detail, saturation vapour pressure deficit is calculated at the target resolution using the same equation that have been used at the source resolution (in particular, using the same parameters).

$$T_{ci} = T_i - 273.15, \quad (24)$$

$$E_i = A \exp(B - T_{ci}/D) T_{ci} / (T_{ci} + C), \quad (25)$$

$$F_i = 1. + X + (S_i/100.) (Y + (Z T_{ci}^2)), \quad (26)$$

$$E_{ei} = E_i F_i, \quad (27)$$

$$Qs_i = \frac{0.62198 E_{ei}}{S_i/100. - 0.37802 E_{ei}}. \quad (28)$$

Finally, specific humidity is calculated at the high resolution using saturated vapour pressure and interpolated relative humidity.

$$Qa_i = Qs_i \frac{R_i}{100}. \quad (29)$$

3.4.6 Downward long-wave radiation

The downward long-wave radiation at the source grid might contain slightly negative values. Those are replaced by local linear interpolation of adjacent values in time. Let L_o be the corrected non-negative field that is interpolated bilinearly using

$$L_t = \mathcal{I}\langle L_o \rangle. \quad (30)$$

These values L_t are then elevation corrected using temperature, surface pressure, and specific humidity. First, vapour pressure and emissivity is calculated at the source resolution using

$$E_o = \frac{S_o/100. Qa_o}{0.62198} \quad (31)$$

$$\epsilon_o = 1.08(1. - \exp[-E_o^{T_o/2016}]) \quad (32)$$

The same calculation are carried out at the high resolution

$$E_i = \frac{S_i/100. Qa_i}{0.62198} \quad (33)$$

$$\epsilon_i = 1.08(1. - \exp[-E_i^{T_i/2016}]) \quad (34)$$



Next, the interpolated values are corrected using the Stefan-Boltzmann law:

$$L_i = \frac{\epsilon_i \sigma}{\epsilon_o \sigma} \left(\frac{T_i}{T_o} \right)^4 L_t \quad (35)$$

where σ is the Stefan-Boltzmann constant $5.6704 \text{ W/m}^2/\text{K}^4$.

3.4.7 Downward short-wave radiation

Let S_o be the downward short-wave radiation at the source resolution. The interpolated downward short-wave radiation is then calculated as

$$S_i = \mathcal{I}\langle S_o \rangle \quad (36)$$

No further cloud cover correction or aerosol correction is applied. It is assumed that the bias correction can cope with these.

3.4.8 Precipitation

Within the WATCH project, a simple bilinear interpolation with P_o as the source field:

$$P_i = \mathcal{I}\langle P_o \rangle \quad (37)$$

It is known that reanalysis data has a bias in number of wet days. Within WATCH, a simple wet day correction was carried out with removing days with smallest positive precipitation until the number of wet days was matching the observed one. CRU data set was used as observations.

Within Ulysses, an alternative approach will be investigated in case the bilinearly interpolated field will not lead to satisfactory results. The alternative approach is to use external drift kriging with elevation as external drift. This approach was used in the EDgE project and allows the representation of the relationship between precipitation amounts and elevation.

3.5 Land surface models

All HMs used in ULYSSES will be run over the global domain at a spatial resolution of 0.1° grid spacing. In this project we will invest in harmonizing the setup among models by using only one source of geophysical information. In this way, we will impose physical consistency within the multi-model ensemble.

Each of the four HMs used within ULYSSES brings value to the multi-model ensemble. This limited number of four models span a considerable range of process representation in hydrologic modelling. The mesoscale Hydrologic Model (mHM) and PCR-GLOBWB (PGB) both represent hydrologic land-surface processes explicitly and use a potential evapotranspiration approach instead of a full representation of the energy cycle. mHM uses a state-of-the-art parameter regionalization, i.e., the Multi-



scale Parameter Regionalization (Samaniego et al., 2010), that allows to apply mHM at multiple spatial resolutions. PGB is unique on its representation of the human-water interactions and groundwater withdrawal rates, but these features will not be used in ULYSSES to ease comparison with the other LSMs/HM. Both models have shown to be among the best performing models for predicting stream-flow at the global scale. In contrast to mHM and PGB, the two hydrologic models HTESSEL and JULES explicitly resolve the energy cycle and provide estimates for all water and energy fluxes and state variables at the land-surface. These two models describe the vegetation in detail using 9 land cover tiles. As a result of using these four models, a relatively large uncertainty within the forecasts can be expected at a reasonable computational cost.

The inclusion of the human-water interactions, including groundwater abstractions, irrigation and reservoirs are not included for ULYSSES, but are a promising expansion as the service evolves.

3.5.1 mHM

The mesoscale Hydrologic Model (mHM, (Samaniego et al., 2010, Kumar et al., 2013)) developed by UFZ is a spatially explicit distributed hydrological model that uses grid cells as a primary hydrologic unit, and accounts for the following processes: canopy interception, snow accumulation and melt, soil moisture dynamics, infiltration and surface runoff, evapotranspiration, subsurface storage and discharge generation, deep percolation and baseflow and discharge attenuation and flood routing.

The model is driven by hourly or daily meteorological forcing data (e.g., precipitation, temperature), and it utilizes observable physical characteristics (e.g., soil textural, vegetation, and geological properties) to infer the spatial variability of the required model parameters. To date, the model has been successfully applied and tested in more than 300 Pan EU basins, as well as India, and USA, ranging in size from 4 to 550,000 km² at spatial resolutions (or grid size) varied between 1 km and 100 km. mHM has been evaluated as one of the few models ready to be implemented in an operational setting (Kauffeldt et al., 2016).

mHM source code (www.ufz.de/mhm) is highly modular and is written in Fortran and freely available under <https://git.ufz.de/mhm>. The model parameter set of mHM will be derived from a global optimization using discharge, ET, and TWS anomaly. This parameter set will be available by 31st October 2019. If it is not available at the beginning of the project, we will use the parameter set that was derived within the EDgE project. This parameter set provides a compromise solution for the hydroclimatic regimes that are present in Europe (Marx et al., 2018, Samaniego et al., 2018, Thober et al., 2018, Wanders et al., 2019).

3.5.2 PCR-GLOBWB

PCR-GLOBWB is a grid-based global hydrology and water resources model developed at Utrecht University (Sutanudjaja et al., 2018). It is a modular model coded in Python and PCRaster-Python routines. Currently two published model versions are available: one at a spatial resolution of half degree (50×50 km at the equator) and one with cells sized 5 arcminutes (10×10 km). Time steps for hydrology and water use are one-day while internal time stepping for hydrodynamic river routing is variable. PCR-GLOBWB differs from all other hydrological model included in this tender, in that it includes dynamic human-water use.

For each grid cell and each time step, PCR-GLOBWB simulates moisture storage in three vertically stacked upper soil layers, as well as the water exchange between the soil, the atmosphere and the un-



derlying groundwater reservoir. The exchange with the atmosphere comprises of precipitation, evaporation from soils, open water, snow and soils and plant transpiration, while the model also simulates snow accumulation, snowmelt and glacier melt. Subgrid-variability of land use, soils and topography is included. PCR-GLOBWB includes improved subgrid schemes for runoff-infiltration partitioning, interflow, groundwater recharge and baseflow, as well as routing of water over the terrain.

Compared to the other models used in this project, PCR-GLOBWB has a stronger emphasis on groundwater flows and soil water storage. These properties have a positive impact when using the model for seasonal forecasts of droughts, as they are largely dominated by the initial hydrological conditions (Wanders et al., 2019). In addition, PCR-GLOBWB can be partially or fully coupled to a two-layer global groundwater model (available at half degree or 5 arcminutes) based on MODFLOW. Recent work also includes coupling PCR-GLOBWB to DFLOW-FM a flexible mesh version of DELFT3D that can be used to solve the 2D shallow water equations for detailed inundation studies.

PCR-GLOBWB resolves the global water balance at a daily temporal resolution and has provided necessary input to a range of applications: energy modelling, biodiversity assessment, climate change studies, seasonal forecasting of floods and droughts, water scarcity assessments and projections of future conflicts. PCR-GLOBWB is widely used within the global hydrological community, and environmental science community and used in 750+ scientific publications.

3.5.3 HTESSEL

HTESSEL is the land-surface model of the coupled ECMWF system 5 seasonal forecasting models (Johnson et al., 2019). Like other land-surface schemes (e.g., Noah-MP, CLM, Terra-ML), it is used to calculate water and energy exchange fluxes between land and atmosphere. Process descriptions have been also improved over the past years with respect to its predecessor Tessel (Balsamo et al., 2011). For example, calculating surface runoff following the variable infiltration capacity approach and a spatially varying soil map were introduced. Similar to JULES, there are only two runoff components (surface and subsurface runoff) and no further delay of these. HTESSEL simulates vegetation on eight different tiles within a grid, including bare soil, low vegetation to high vegetation. Within each of the tiles, the information of different PFTs is lumped. Using HTESSEL allows us to further understand the effect of the downscaling and bias-correction of the ULYSSES production chain because HTESSEL is the land-surface model used within SEAS5. HTESSEL is owned by ECMWF and will be used under the Apache 2.0 license.

3.5.4 JULES

JULES — The JULES (Joint UK Land Environment Simulator) model is a core component of both the Met Office's modelling infrastructure and NERC's Earth System Modelling Strategy. It simulates the land response to climate and weather with options to study the water and carbon cycle, either interlinked or separately (Best et al., 2011). It can be run at any scale, depending on the meteorological driving data and the soils and land cover data to inform the model. The system that the code needs to run is essentially the JULES modelling system.

The hydrology in JULES is relatively sophisticated and includes rainfall intensity modelling, saturation and infiltration excess surface runoff generation as well as drainage, soil freezing, multi-layer soil moisture and snow models and river routing as well as separate presentation of transpiration, interception and soil surface evaporation and a full energy balance (heat and radiation).



JULES has been developed by a wide community of UK researchers, coordinated by the Met Office and CEH. JULES is available to anyone for non-commercial use, free of charge. This has led to a large and diverse community from across the globe using JULES to study land surface processes on a wide variety of temporal and spatial scales.

This model has been used in a variety of water-related global applications, in particular FP7-WATCH, ISI-MIP and earthH2Observe where the model was used within an ensemble of models to assess the global water cycle. It has been used to study the water budgets of the UK ([Martínez-De La Torre et al., 2019](#)) and as part of several studies to study the global water cycle ([Schellekens et al., 2017](#)).

JULES is similar to HTESSEL ([Johnson et al., 2019](#), [Balsamo et al., 2011](#)) (The ECMWF LSM currently used in GloFAS) in several aspects: it represents the land cover with tiles of different vegetation and surface types (e.g. deciduous and evergreen trees, grasses, urban, ice and water). The sub-surface has a layered soil-physics routine which solves the Darcy-Richards equations, giving vertical profiles of soil moisture and heat which mean that soil freezing and its impact on hydrology can be modelled. The hydrology in JULES is also similar, with saturation excess runoff generation dependent on the topography, soil-drainage and a snow model as well as separate presentation of transpiration, interception and soil surface evaporation. However, there is a key module in JULES which ensures that the hydrology is sensitive to rainfall intensity. By including a representation of the distribution of rainfall intensity, and how that varies with rainfall type, JULES is able to distinguish between intense thunderstorms and large-scale rain and the impact that it has on water-loss through interception and the impact on infiltration excess runoff. With increasing rainfall intensity expected to rise, this factor may have a significant role on the global hydrological cycle.

3.5.5 The multiscale Routing Model mRM

The derivation of the streamflow for all HMs will be carried out with the multiscale routing model (mRM) ([Thober et al., 2019](#)). The sum of all runoff components of each hydrological model is used as input for mRM. We developed mRM, a Fortran-written software tool, to provide a simple solution to route water through a river network by representing the river reaches connecting two rectangular grid cells. mRM solves the kinematic wave equation using a finite difference scheme. An adaptive time stepping scheme fulfilling a numerical stability criterion is used to guarantee applicability at a range of spatial and temporal scales. The time-step of the employed Muskingum algorithm (ranging from hours to days) in the routing scheme is dependent on the spatial resolution such that the hydrological model outputs become quasi-scale invariant. It is worth mentioning that mRM can be easily run at various spatial resolution by only changing one number in a configuration file. No further adjustment of the input data is required ([Thober et al., 2019](#)). mRM has the ability to simultaneously route cell-generated runoff to multiple outlets, allowing streamflow to be generated over the entire domain within a single integration time step. ULYSSES will make full use of the key feature of mRM which is that the river network can be represented at a different resolution than the hydrological models are applied on. It can also handle daily input and disaggregates this input in time if this is required by the numerical integration time step. The river network will be stored in a restart file to minimise execution time. mRM outputs gridded discharge and time series at pre-defined locations which is exactly what is required in this Lot. We have used mRM successfully in EDgE and since then, have introduced parallel compute paradigms in the Fortran source code to make full use of high performance compute clusters. We use both openMP and MPI paradigms within mRM. Overall, mRM is ideal within a seasonal hydrologic ensemble forecasting system. By using one routing model, we reduce the uncertainty to the hydrologic modelling system and in addition, we allow for further developments of this aspect of



the system without affecting the multiple runoff-generation models. mRM has small execution times by using parallel code, and provides the output at a grid and pre-defined locations.

4 Calibration and cross-validation

Within the ULYSSES project, all four models are set up and, if possibly, automatically calibrated) by every modeling group. This will lead to one (i.e., the best possible) parameter set applicable for the global-scale seamless simulation. The overall procedure should follow, if possible, the protocol proposed by (Samaniego et al., 2017, Rakovec et al., 2016a).

4.1 Protocol

- All four models will be setup on the common geomorphologic/land- cover characteristics provided in Table 1.
- Historical calibration runs conducted with the ERA5-Land data set (or equivalent, as chosen by the contractor) at 0.1°. The variables are listed in Table 2.
- All model will be evaluated against three main water balance components:
 - Daily/monthly streamflow (i.e., routed runoff with mRM) using a preselected GRDC gauges having daily records from 1985-2019, with at least of 5years of record. The period from 1980 to 1984 will be used for model warm-up.
 - Monthly simulated grids of evapotranspiration (ET) against FLUXNET from 1981 to at least 2011 at 0.5°grid.
 - Monthly simulated grids of terrestrial water storage anomaly (TWSA) against GRACE (NASA-GFZ) from 2004 to 2016 at 1°grid.
- All models will report the following metrics:
 - The KGE metric (Gupta et al., 2009) including its components for each GRCD station on calibrated and validated sites. A site used for calibration should not be used for cross-validation. No separation between calibration/validation periods.
 - Rank correlation for actual ET and TWSA fields.
 - Pearson correlation for standardized anomalies (i.e., removed seasonality) ET and TWSA fields.
 - Taylor diagrams for ET and TWS as shown in Rakovec et al. (2016b).
- The selection of GRDC data will follow strict hydro-climatological testing, test of jumps and should not exhibit anthropogenic interferences during the record period. Selection criteria are describes in Rakovec et al. (2016a). The number of selected sites should contain at least 20 basin in each updated Köppen–Geiger hydro-climatic zones. The selected basins should pass a Budiko-test (i.e., should not deviate more than 10% from the ideal theoretical relationships proposed by Schreiber, 1904; Ol'dekop, 1911 and Budyko, 1974 citepZink:2017hk.
- The drainage area of the calibration/cross-validation basins should vary from 10 km² to 10000 km².



- CDF of model performance should follow the schema proposed in [Rakovec et al. \(2019\)](#). Here KGR and CDFs of its components should be depicted.
- If the LSM/HMs can perform tests at multiple scales, then fluxmatching test between 0.1° and 0.5° and 1.0°, respectively should be performed as proposed in [Samaniego et al. \(2017\)](#).



References

- G Balsamo, F Pappenberger, E Dutra, P Viterbo, and B Van den Hurk. A revised land hydrology in the ECMWF model: a step towards daily water flux prediction in a fully-closed water cycle. *Hydrological Processes*, 25(7): 1046–1054, 2011.
- M J Best, M Pryor, D B Clark, G G Rooney, R L H Essery, C B Ménard, J M Edwards, M A Hendry, A Porson, N Gedney, L M Mercado, S Sitch, E Blyth, O Boucher, P M Cox, C S B Grimmond, and R J Harding. *The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes*. *Geoscientific Model Development*, 4(3):677–699, 2011.
- Marc F P Bierkens. *Global hydrology 2015: State, trends, and directions*. *Water Resources Research*, 51: 4923–4947, July 2015.
- L. Bowling and J. Polcher. *The ALMA data exchange convention*. Online, 2001. <http://www.lmd.jussieu.fr/polcher/ALMA/>.
- MetaData CF. *NetCDF CF Metadata Conventions, Version 1.7*. Online, 2007. <http://cfconventions.org>.
- Hoshin V Gupta, Harald Kling, Koray K Yilmaz, and Guillermo F Martinez. *Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling*. *Journal of Hydrology*, 377(1-2):80–91, October 2009.
- George H Hargreaves and Zohrab A Samani. *Reference Crop Evapotranspiration from Temperature*. *Applied engineering in agriculture*, 1(2):96–99, 1985.
- S Hempel, K Frieler, L Warszawski, J Schewe, and F Piontek. *A trend-preserving bias correction – the ISI-MIP approach*. *Earth System Dynamics*, 4(2):219–236, July 2013.
- Dublin Core Metadata Initiativ. *DCMI, Dublin core metadata element set, Version 1.1*. Online, 2020. <http://dublincore.org/documents/dces/>.
- Stephanie J Johnson, Timothy N Stockdale, Laura Ferranti, Magdalena A Balmaseda, Franco Molteni, Linus Magnusson, Steffen Tietsche, Damien Decremet, Antje Weisheimer, Gianpaolo Balsamo, Sarah P E Keeley, Kristian Mogensen, Hao Zuo, and Beatriz M Monge-Sanz. *SEAS5: the new ECMWF seasonal forecast system*. *Geoscientific Model Development*, 12(3):1087–1117, 2019.
- A Kauffeldt, F Wetterhall, F Pappenberger, P Salamon, and J Thielen. *Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level*. *Environmental Modelling & Software*, 75:68–76, January 2016.
- Rohini Kumar, Luis Samaniego, and Sabine Attinger. *Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations*. *Water Resources Research*, 49:360–379, January 2013.
- Stefan Lange. *Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0)*. *Geoscientific Model Development*, 12(7):3055–3070, 2019.
- Y Q Luo, J T Randerson, G Abramowitz, C Bacour, E Blyth, N Carvalhais, P Ciais, D Dalmonch, J B Fisher, R Fisher, P Friedlingstein, K Hibbard, F Hoffman, D Huntzinger, C D Jones, C Koven, D Lawrence, D J Li, M Mahecha, S L Niu, R Norby, S L Piao, X Qi, P Peylin, I C Prentice, W Riley, M Reichstein, C Schwalm, Y P Wang, J Y Xia, S ZAEHLE, and X H Zhou. *A framework for benchmarking land models*. *Biogeosciences*, 9(10):3857–3874, October 2012.
- Alberto Martínez-De La Torre, Eleanor M Blyth, and Graham P Weedon. *Using observed river flow data to improve the hydrological functioning of the JULES land surface model (vn4.3) used for regional coupled modelling in Great Britain (UKC2)*. *Geoscientific Model Development*, 12(2):765–784, 2019.
- Andreas Marx, Rohini Kumar, Stephan Thober, Oldrich Rakovec, Niko Wanders, Matthias Zink, Eric F Wood, Ming Pan, Justin Sheffield, and Luis Samaniego. *Climate change alters low flows in Europe under global warming of 1.5, 2, and 3 °C*. *Hydrology and Earth System Sciences*, 22(2):1017–1032, 2018.
- J L Monteith. *Evaporation and surface temperature*. *Quarterly Journal of the Royal Meteorological Society*, 107 (451):1–27, January 1981.
- C H B Priestley and R J Taylor. *On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters*. *Monthly Weather Review*, pages 1–12, November 2003.



- O Rakovec, R Kumar, S Attinger, and L Samaniego. Improving the realism of hydrologic model functioning through multivariate parameter estimation. *Water Resources Research*, 52(10):7779–7792, 2016a.
- Oldrich Rakovec, Rohini Kumar, Juliane Mai, Matthias Cuntz, Stephan Thober, Matthias Zink, Sabine Attinger, David Schäfer, Martin Schrön, and Luis Samaniego. Multiscale and Multivariate Evaluation of Water Fluxes and States over European River Basins. *Journal of Hydrometeorology*, 17(1):287–307, 2016b.
- Oldrich Rakovec, Naoki Mizukami, Rohini Kumar, Andrew J Newman, Stephan Thober, Andrew W Wood, Martyn P Clark, and Luis Samaniego. Diagnostic Evaluation of Large-Domain Hydrologic Models Calibrated Across the Contiguous United States. *Journal of Geophysical Research-Atmospheres*, 124(24):13991–14007, December 2019.
- L Samaniego, S Thober, R Kumar, N Wanders, O Rakovec, M Pan, M Zink, J Sheffield, E F Wood, and A Marx. Anthropogenic warming exacerbates European soil moisture droughts. *Nature Climate Change*, 5:1117–21, April 2018.
- Luis Samaniego, Rohini Kumar, and Sabine Attinger. Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, 46(5), 2010.
- Luis Samaniego, Rohini Kumar, Stephan Thober, Oldrich Rakovec, Matthias Zink, Niko Wanders, Stephanie Eisner, Hannes Müller Schmied, Edwin H Sutanudjaja, Kirsten Warrach-Sagi, and Sabine Attinger. Toward seamless hydrologic predictions across spatial scales. *Hydrology and Earth System Sciences*, 21(9):4323–4346, 2017.
- Luis Samaniego, Stephan Thober, Niko Wanders, Ming Pan, Oldrich Rakovec, Justin Sheffield, Eric F Wood, Christel Prudhomme, G Rees, Helen Houghton-Carr, Matthew Fry, Katie Smith, Glenn Watts, Hege Hisdal, Teodoro Estrela, Carlo Buontempo, Andreas Marx, and Rohini Kumar. Hydrological Forecasts and Projections for Improved Decision-Making in the Water Sector in Europe. *Bulletin of the American Meteorological Society*, 100(12):2451–2472, December 2019.
- Jaap Schellekens, Emanuel Dutra, Alberto Martínez-De La Torre, Gianpaolo Balsamo, Albert Van Dijk, Frederiek Sperna Weiland, Marie Minvielle, Jean Christophe Calvet, Bertrand Decharme, Stephanie Eisner, Gabriel Fink, Martina Flörke, Stefanie Peßenteiner, Rens Van Beek, Jan Polcher, Hylke Beck, Rene Orth, Ben Calton, Sophia Burke, Wouter Dorigo, and Graham P Weedon. A global water resources ensemble of hydrological models: The earthH2Observe Tier-1 dataset. *Earth System Science Data*, 9(2):389–413, July 2017.
- Edwin H Sutanudjaja, Rens Van Beek, Niko Wanders, Yoshihide Wada, Joyce H C Bosmans, Niels Drost, Ruud J Van Der Ent, Inge E M De Graaf, Jannis M Hoch, Kor De Jong, Derek Karssenbergh, Patricia López López, Stefanie Peßenteiner, Oliver Schmitz, Menno W Straatsma, Ekkamol Vannamettee, Dominik Wisser, and Marc F P Bierkens. PCR-GLOBWB 2: A 5 arcmin global hydrological and water resources model. *Geoscientific Model Development*, 11(6):2429–2453, June 2018.
- Stephan Thober, Rohini Kumar, Niko Wanders, Andreas Marx, Ming Pan, Oldrich Rakovec, Luis Samaniego, Justin Sheffield, Eric F Wood, and Matthias Zink. Multi-model ensemble projections of European river floods and high flows at 1.5, 2, and 3 degrees global warming. *Environmental Research Letters*, 13(1):014003–11, January 2018.
- Stephan Thober, Matthias Cuntz, Matthias Kelbling, Rohini Kumar, Juliane Mai, and Luis Samaniego. The multiscale routing model mRM v1.0: simple river routing at resolutions from 1 to 50 km. *Geoscientific Model Development*, 12(6):2501–2521, 2019.
- Niko Wanders, Stephan Thober, Rohini Kumar, Ming Pan, Justin Sheffield, Luis Samaniego, and Eric F Wood. Development and Evaluation of a Pan-European Multimodel Seasonal Hydrological Forecasting System. *Journal of Hydrometeorology*, 20(1):99–115, January 2019.
- Matthias Zink, Juliane Mai, Matthias Cuntz, and Luis Samaniego. Conditioning a Hydrologic Model Using Patterns of Remotely Sensed Land Surface Temperature. *Water Resources Research*, 54(4):2976–2998, April 2018.



This page is intentionally left blank



ECMWF – Shinfield Park, Reading, RG2 9AX, UK
Contact: info@copernicus-climate.eu