The Inter-Sectoral Impact Model Intercomparison Project

ISIMIP2a Simulation protocol (extended version)

Last updated on 15 September 2020

Table of Contents

1 Introduction

1.1 ISIMIP: General concept

ISIMIP provides a framework for the collation of a consistent set of climate impact data across sectors and scales. This framework will serve as a basis for model evaluation and improvement, allowing for improved estimates of the biophysical and socio-economic impacts of climate change at different levels of global warming. It also provides a unique opportunity for considering interactions between climate change impacts across sectors through consistent scenarios.

ISIMIP is intended to be structured in successive rounds, each having its own focus topics and focus regions that inform the scenario design. The main components of the ISIMIP framework are:

- A common set of climate and other input data which will be distributed via a central database;
- A common modelling protocol to ensure consistency across sectors and scales in terms of data, format and scenario set-up;
- A central archive where the output data will be collected and made available to the scientific community.

1.2 General remark regarding adaptation

As in the ISIMIP Fast Track, simulations should not be designed to describe the effects of different adaptation measures. In contrast, it is the aim to describe the impacts of climate change on different sectors under "present-day" management assumptions. There are individual exceptions to this general rule (such as "naturalized" runs within the water sector). These exceptions are clearly specified in the sector-specific scenario set-up of the simulations. Wherever such an exception is not mentioned please choose the "present day" option regarding management.
In particular, the historical validation runs should be "as close to the real historic conditions as possible" (i.e. to the extent to which this can be achieved without major model improvement).

2 General Design of ISIMIP2a – Focus topic

"Extreme events and variability" was chosen as the focus topic of ISIMIP2a, reflecting the interest of the community as well as stakeholders in investigating and improving the representation of variability and extreme events, in particular in impact models and along the entire modelling chain. Therefore, the model evaluation and validation task specified in this protocol is explicitly designed to evaluate the models' ability to reproduce observed historical variability, responses to extreme climatic events such as heat waves, droughts, floods, heavy rains and storms, and representation of extreme impact events. Based on these evaluation exercises, modelling teams will have the opportunity to adjust model parameters and implement necessary model improvements. Moreover, the ISIMIP2a simulations serve to validate the impact models that are used for future projections in ISIMIP2b.

Note that the emphasis on this focus topic does not exclude some other work not directly related to extreme events and variability. For example, in particular for regional models and in new sectors, it may be important to first calibrate and investigate performance for average condition.

3 Motivation of experiment design

This chapter provides a short description of the scientific rationale behind the design of each of the experiments in ISIMIP2a. The details of the experiments are further described in the remainder of the protocol.

The overarching objective of the historical validation experiment is to gain insight into the ability of current impact models to reproduce observed features of simulated variables, with an emphasis on (but not limited to) variability and extreme events. Simulations are designed such as to match historical conditions as closely as possible, within the limitations of e.g. availability of historical forcing data, variety of model formulations, and model development resources. In addition, ISIMIP2a serves to evaluate the models used for future projections in ISIMIP2b. Therefore, it is **important to apply the same model version for ISIMIP2a and ISIMIP2b**.

Four different observations-based historical climate datasets will be used to force impact models, to allow for a comparison of the different historical simulations. Each data set has its own strengths and weaknesses (e.g. regarding temporal extent, quality of specific variables, previous application within the modeling community) and generally represents a plausible reconstruction of the terrestrial climate of the past ~100 years. The different historical simulations will allow a systematic quantification of the effect of the choice of forcing data on impact model results; allow comparison to previous studies using either of these datasets; and provide an extensive data base for model evaluation and impact assessment, in particular with regard to the focus topic (e.g. a certain climatic extreme event could be better reproduced in one data set than in the other).

Addendum August, 2019: ISIMIP2a now includes a fifth and sixth observations-based climate dataset (GSWP3-W5E5 and GSWP3-EWEMBI). The W5E5 and EWEMBI data contained in these datasets span the period 1979-2016, and were backward-extended to 1901 using GSWP3, hence the names. See details below.

4 Common input data and settings for all sectors

This chapter describes climate forcing data and other input data that should be used by modelling groups in all sectors. Note that several different experiments with differences in input data and other settings are requested; see the sectoral chapters for a list of the requested experiments. In this chapter, we only describe the common input data sets.
If you require additional input data that is not specified in this chapter, please use your default data source. In case anything remains unclear please contact the coordination team or sectoral coordinators.

4.1 Atmospheric data

Please use the historical climate data listed in **Table** 1 for the historical calibration and validation runs. The runs should start in 1901, or earlier if spin-up is needed (see below and Section 5.1). All data will be available through the ISIMIP website, www.isimip.org. Separate historical simulations should be conducted with each of four different datasets, **in the order indicated** in the last column of **Table** 1. This is because each of the datasets has its own advantages and shortcomings, and thus by using several of them, it will be possible to assess the influence of the choice of forcing data on the overall results. Moreover, this procedure serves the needs of the different participating sectors (e.g. data over ocean is needed for the fisheries sector) and facilitates consistency with other model intercomparison exercises (e.g. ISIMIP Fast Track; GSWP3). Modelling groups that cannot run all datasets before the submission deadline should nonetheless begin in the order indicated and inform the ISIMIP coordination team.

Table 1: Historical (atmospheric) climate data sets to be used in calibration and validation runs. All data sets contain the variables near-surface air temperature (tas), precipitation (pr), near-surface relative humidity (rhs), surface downwelling longwave radiation (rlds), surface downwelling shortwave radiation (rsds), surface pressure (ps), near-surface wind speed (wind), and partly also daily minimum and maximum near-surface air temperature (tasmin and tasmax, resp.). Note that simulations should be conducted with each of these datasets.

Dataset	Reanalysis	Years	Resolution, coverage	Bias target	Priority; comments
WATCH-WFDEI (WATCH a.k.a. WFD: Weedon, et al., 2011; WFDEI: Weedon et al., 2014)	ERA-40, ERA-Interim	1901- 2016	0.5° Land	CRU, GPCC	1 Combined forcing file (WFD 1901-1978, WFDEI 1979-2016) will be provided by ISIMIP. WFDEI precipitation and snowfall data are those corrected with GPCC. **NOTE** a discontinuity in the data exists at the transition from WATCH to WFDEI, and results must be interpreted with caution.
GSWP3-W5E5 (GSWP3: Dirmeyer et al., 2006; W5E5: Lange S. , 2019a; Cucchi et al., 2020)	ERA5	1901-2016	0.5° Land + Ocean	CRU, GPCC, GPCP	2 Combined forcing file (homogenized GSWP3 1901-1978, W5E5 1979-2016) will be provided by ISIMIP. To minimize discontinuities at the 1978/1979 transition, GSWP3 data were homogenized with W5E5 data using the ISIMIP3BASD v2.4.1 bias adjustment method (Lange, 2019c; Lange, 2020).
GSWP3-EWEMBI (GSWP3: Dirmeyer et al., 2006; EWEMBI: Lange S., 2019b)	ERA-Interim	1901-2016	0.5° Land + Ocean	GPCC, GPCP, CRU, SRB	3 Combined forcing file (homogenized GSWP3 1901-1978, EWEMBI 1979-2016) will be provided by ISIMIP. To minimize discontinuities at the 1978/1979 transition, GSWP3 data were homogenized with EWEMBI data using the ISIMIP3BASD v2.4.1 bias adjustment method (Lange, 2019c; Lange, 2020).

GSWP3 (Dirmeyer et al., 2006)	20CR	1901-2010	0.5° Land + Ocean	GPCC, GPCP, CPC-Unified, CRU, SRB	4 Based on dynamical downscaling. Further details on the global dynamical downscaling method are given in (Yoshimura & Kanamitsu, 2008) (Yoshimura & Kanamitsu, 2013).
PGMFD v2.1 (Princeton) (Sheffield et al., 2006)	NCEP/NCAR Reanalysis 1	1901-2012	0.5° Land + Ocean	CRU, SRB, TRMM, GPCP, WMO validated against GSWP2	5
WATCH (WFD) (Weedon, et al., 2011)	ERA-40	1901-2001	0.5° Land	CRU, GPCC	6

Historical **CO2 concentrations** are also provided in the input data archive (/ISIMIP/ISIMIP2a/InputData/climate_co2/co2 /historical_CO2_annual_1765_2018.txt). They are based on time series of global atmospheric CO2-concentrations from (Meinshausen, Raper, & Wigley, 2011) for 1765-2005 and (Dlugokencky & Tans, 2019) from 2006-2018.

Note that simulation results only need to be submitted for the reporting periods specified in Section 5.1. The parts of the climate forcing data prior to the reporting period may be used for spin-up purposes and/or to facilitate further analyses. Simulation results for years outside the reporting period may still be submitted to the ISIMIP repository on a voluntary basis.

4.2 Oceanic data

See Section 12.

4.3 Land-use/land-cover

We provide a time-varying historical land-use (LU) data set that should be used for the historical validation runs. The time series starts in 1861 and ends in 2018 (files under /ISIMIP/ISIMIP2a/InputData/landuse_humaninfluences/landuse/ and /ISIMIP/ISIMIP2a/InputData /landuse_humaninfluences/n-fertilizer/[1]) and should be applied for the spin-up as well as for the historical runs, as described above. This landuse data is ultimately based on the HYDE 3.2 data set (Klein Goldewijk, 2016). This data was in turn harmonized by the land use group of George Hurtt at the University of Maryland College Park, which provides the “Land-Use Harmonization" (LUH2 v2h) data set (Hurtt, Chini, Sahajpal, Frolking, & et al, In prep.) [see also https://luh.umd.edu/]. This data set provides land use categories, pastures and rangeland, 5 crop types (C3 annual, C3 perennial, C4 annual, C4 perennial and C3 nitrogen fixing crops). Furthermore, the LUH2 v2h data set gives information on management, i.e. it provides information of irrigated vs rainfed areas and on fertilization rates. We interpolated this data set onto the ISIMIP standard grid in order to generate the “landuse-totals”, the “landuse-pastures”, the “landuse-urbanareas” and the “n-fertilizer-5crops” ISIMIP input files. In order to downscale the 5 crop files to the 15 crops (maize, groundnut, rapeseed, soybeans, sunflower, rice, sugarcane, pulses, temperate cereals [incl. wheat], temperate roots, tropical cereals, tropical roots, others annual, others perennial, and others N-fixing), the Monfreda data set (Monfreda, Ramankutty, & Foley, 2008) has been used. In this step the 5 crop types are split into 15 crop types according to the ratios given by the Monfreda data. Models that simulate their own natural vegetation should report that. All these grid cell shares don’t necessarily add up to 1, since we ignore some landuse categories such as natural vegetation, ponds, highways and so on. If you need this (in previous protocol versions called “others”), please calculate 1-all other categories.

[1] Previous versions of these data in ‘deprecated’ folders within these same locations.

Table 2: Land-cover and soil data to be used in historical validation runs.

Dataset	Description	More info	Scale	Variables included
Mandatory				
Historical land use patterns	Combination of HYDE3.2/LUH2h v2h and Monfreda land use data. **Note:** Data covers from 1861 to 2018.	HYDE 3.2: (Klein Goldewijk, 2016) LUH v2h: (Hurtt, Chini, Sahajpal, Frolking, & et al, In prep.) Monfreda: (Monfreda, Ramankutty, & Foley, 2008)	Global 0.5°Annual	Irrigated and rainfed crop areas for the following crop classes: c3per, c4per, maize, oil_crops_groundnut, oil_crops_rapeseed, oil_crops_soybean, oil_crops_sunflower, others_c3ann, others_c3nfx, pulses, rice, temperate_cereals, temperate_roots, tropical_cereals, tropical_roots Furthermore, there are pastures, rangeland, managed_pastures and urbanareas,
Historical area covered by natural vegetation	Derived from "HYDE3/ MIRCA" as 1-(all agricultural area). Note that forestry is counted as natural vegetation because of lack of historical forestry data.		See above	Fraction of grid cell covered by natural vegetation.
Optional				
HWSD or GSWP3 (upscaled version of HWSD)	soil map	See http://hydro.iis.u-tokyo.ac.jp/~sujan/research/gswp3/soil-texture-map.html, upscaling method A. Each model has the option to use their own soil datasets if preferred.	Global 30 arc sec (HWSD) *or* 0.5° (GSWP3) Fixed	Soil types: Sand, Loamy Sand, Sandy Loam, Loam, Silt Loam, Silt, Sandy Clay Loam, Clay Loam, Silty Clay Loam, Sandy Clay, Silty Clay, Clay, Ice.

Table 3: Agricultural land-use categories.

Land-use type	Historical reconstruction	Disaggregation into functional crop types (LUH2)	Individual crops or crop groups
Irrigated crops	LUH2	Total cropland disaggregated into: C_3 annual, C_3 nitrogen-fixing, C_3 perennial, C_4 annual, C_4 perennial (contains only sugarcane)	C_3 annual disaggregated into: rapeseed, rice, temperate cereals, temperate roots, tropical roots, sunflower, others C_3 annual C_3 perennial: (no further disaggregation) C_3 nitrogen-fixing disaggregated into: groundnut, pulses, soybean, others C_3 nitrogen-fixing C_4 annual disaggregated into: maize, tropical cereals C_4 perennial: sugarcane
Rainfed crops	LUH2	Total cropland disaggregated into: C_3 annual, C_3 nitrogen-fixing, C_3 perennial, C_4 annual, C_4 perennial (contains only sugarcane)	C_3 annual disaggregated into: rapeseed, rice, temperate cereals, temperate roots, tropical roots, sunflower, others C_3 annual C_3 perennial: (no further disaggregation) C_3 nitrogen-fixing disaggregated into: groundnut, pulses, soybean, others C_3 nitrogen-fixing C_4 annual disaggregated into: maize, tropical cereals C_4 perennial: sugarcane
Pastures	LUH2	Total pastures are provided.	In addition, pastures are split into managed pastures and (natural) rangelands
Bioenergy production (rainfed grass)	-		e.g., miscanthus (if you use a different bioenergy crop, please indicate this in the model description)
Bioenergy production (rainfed trees)	-		e.g., poplar (temperate), eucalyptus (tropical) (if you use a different bioenergy crop, please indicate this in the model description)
Urban	LUH2		Variable is called urban areas

4.4 Socio-economic input

Table 4**:** Socio-economic data provided for ISIMIP2a.

Variable	Name	Unit	Time coverage, frequency	Comments
Population, country level and 0.5°x0.5° grid level[2]	pop	number of people	1860-2017, annual time steps	National and gridded historic population data are taken from the History Database of the Global Environment (HYDE) version 3.2.1 (http://themasites.pbl.nl/tridion/en/themasites/hyde/download/index-2.html). Estimates for total as well as urban and rural population counts are available. This does not only extend previously provided population data but might also outdate previous data to a series of bug fixes since HYDE version 3.2.000.
Gross Domestic Product (GDP), country level and 0.5°x0.5° grid level	gdp	GDP PPP 2005 USD	1861-2016, annual time steps	Historic country-level GDP data are an extension of the data provided by Geiger, 2018 (https://www.earth-syst-sci-data.net/10/847/2018/essd-10-847-2018.html), and are derived mainly from the Maddison Project database. Gridded GDP data is based on the downscaling algorithm developed by Murakami & Yamagata, 2019 (https://www.mdpi.com/2071-1050/11/7/2106, http://dataservices.gfz-potsdam.de/pik/showshort.php?id=escidoc:2740907). The years 2000-2016 have been rescaled to provide a better match to the updated national GDP time series. All datasets are also available via the ISIMIP website, www.isimip.org.

[2] Various other resolutions (0.1° x 0.1°, 0.125° x 0.125°, 0.0833° x 0.0833° (5arcmin), 0.0416° x 0.4166° (2.5arcmin)) are also available.

4.5 Other human influences

For all these input variables, we describe reconstructions to be used for the historical (**hist**) simulations (see **Table** 5).

Table 5: Data sets representing "other human influences" for the historical simulations (**hist**).

Driver	Historical reconstruction
Reservoirs & dams (1800-2025) • **location** • **upstream area** • **capacity** • **construction/commissioning year**	Global data on 0.5° grid, mapped to the DDm30 routing network. Including some dams with planned completion in the future. Sources: - 6862 dams of capacity > 100 mcm (0.1 km^3) from GranD database (http://wp.geog.mcgill.ca/hydrolab/grand/). - 50 dams of capacity > 1000 mcm (1 Gt) with construction/commissioning year after 2010 from Kansas State University (KSU, Dr. Jida Wang). A file is provided with information on dams/reservoirs from both sources, with locations adapted to DDM30 to best match reported upstream areas. **Note on mapping**: Dam locations are originally given as point-information and simple mapping to the nearest grid cell center in the DDM30 stream network can lead to inconsistencies. For dams from GRanD, the comparison of GRanD upstream area and DDM30 upstream area was used as criterion to check the agreement of the location. A correction was applied to all dams with an upstream area in the GRanD data bigger than 10000 km^2 (844 dams, making up 94% of the total upstream area) and more than 50% deviation from the upstream area reported in GranD: They were shifted to the 8 possible neighboring cell centers and the dam was moved to the grid cell centre resulting in the smallest deviation in the upstream area. In total, 144 dams have been moved. For dams from KSU, no upstream area was provided and thus no comparison was possible; dams locations were mapped to DDM30 manually. **Note on usage**: Because the data from KSU is yet unpublished, modeling teams using it are asked to offer co-authorship to the team at KSU on any resulting publications. Please contact info@isimip.org in case of questions.
N fertilizer use (kg per ha of cropland)	Annual crop-specific input per ha of crop land for C_3 and C_4 annual, C_3 and C_4 perennial and C_3 Nitrogen fixing. This data set is part of the LUH2 dataset developed for CMIP6 (Hurtt, Chini, Sahajpal, Frolking, & et al, In prep.) based on HYDE3.2. The 2016 value is assumed to be identical to the 2015 value.
Nitrogen (NH_X and NO_Y) deposition	Monthly, 0.5°gridded data for global nitrogen deposition (1860-2016) NH_X and NO_Y deposition (g N m-2 yr-1) from the NCAR Chemistry-Climate Model Initiative (CCMI). Nitrogen deposition data was interpolated to 0.5° by 0.5° by the nearest grid. Data in 2015 and 2016 is assumed to be same as that in 2014. More information is

	available in (Tian, et al., 2018).
Fishing intensity	Depending on model construction, one of: Fishing effort from the Sea Around Us Project (SAUP); catch data from the Regional Fisheries Management Organizations (RFMOs) local fisheries agencies; exponential fishing technology increase and SAUP economic reconstructions. This data can currently not be hosted on ISIMIP servers; modelers are asked to contact the sectoral coordinators of the marine ecosystems and fisheries sector to gain access to this data.
Forest management	Based on observed stem numbers.
Water abstraction for domestic and industrial uses	For modelling groups that do not have their own representation, we provide files containing the multi-model mean domestic and industrial water withdrawal and consumption generated from ISIMIP2a varsoc runs of WaterGAP, PCR-GLOBWB, and H08. This data is available in the ISIMIP2b archive, from 1901 until 2005 (ISIMIP2b/InputData/water_abstraction/histsoc). Remaining years after 2005 should be filled in with the analogous data from RCP6.0 (ISIMIP2b/InputData/water_abstraction/rcp60soc/), which is based on multi-model projections from the Water Futures and Solutions project (Wada et al., 2016, http://www.geosci-model-dev.net/9/175/2016/).

4.6 Focus regions

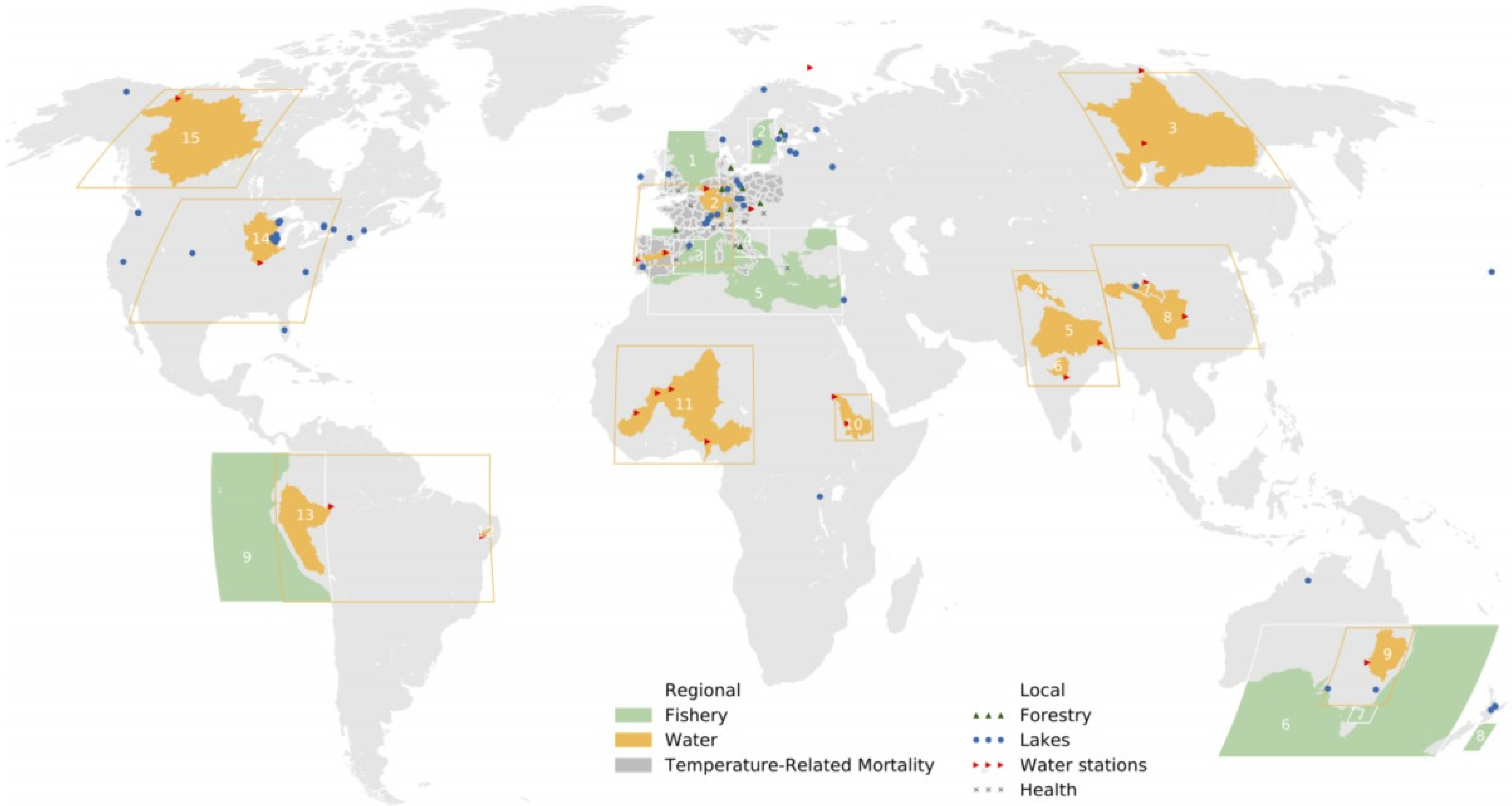

Figure 1: ISIMIP2a extended focus regions. See **Section** 6-14 for sector-specific details.

4.7 Scenario design

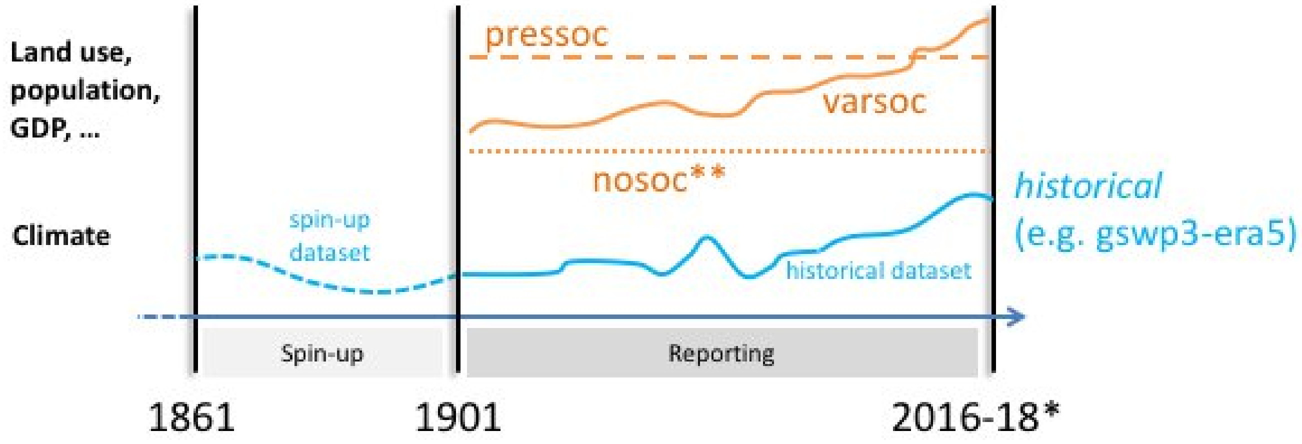

Figure 2: Schematic representation of the scenario design for ISIMIP2a. In addition to land use, population, and GDP, other non-climatic anthropogenic forcing factors and management (such as irrigation, fertilizer input, selection of crop varieties, flood protection levels, dams and reservoirs, water abstraction for human use, fishing effort, atmospheric nitrogen deposition, etc.) should follow the same scenario. *End year is different for each dataset. **no human influences except year-2000 land-use.

Table 6: General definition of socio-economic, CO_2 sensitivity, and agricultural management scenarios. See sector-specific sections for further specifications.

	Scenario name	Description
Socio-economic scenarios	nosoc	No human influences except for year-2000 constant land-use patterns. No anthropogenic water abstraction (e.g. irrigation), no reservoirs/dams. No population and GDP data prescribed. For forest models, this means running the model without any management (but nitrogen deposition should be included). If your model includes natural regeneration, please only regeneration those species previously present on the plot.
	pressoc	Present-day human impact runs: only climate varies; keep all other settings (population, GDP, land-use, technological progress, etc.) constant at year 2000 values. This run will be used to quantify adaptation pressure under current socio-economic conditions. For water models, pressoc includes present-day irrigation and other water uses / reservoirs. See further details for health models in **Section** 14.
	varsoc	Not only climate but also population, GDP, land-use, technological progress, etc. varies over the historical period. For forest simulations, this means: Manage forests according to historical management guidelines without species change. See further details for health models in **Section** 14.
	nat	A natural vegetation only run without any land-use pattern (optional for biomes models). This is a reference run to separate fluxes from natural vegetation and agriculture in runs with historic land-use. It is like the nosoc run but without land-use. If your model does not distinguish between natural and managed land, the “nat” run will be identical to the “nosoc” run.
CO_2 sensitivity scenarios	co2	Transient CO2 concentration (taken historical data) for CO2 fertilization effects. If your model does not implement CO2 fertilization using transient CO2 concentrations at all, please use your own implementation and include that in the reporting. For simulations without any CO2 implementation also use the “co2” tag to indicate that this driver is given implicitly in the driving climate forcing.

	noco2	Sensitivity experiment: only applies to models that take CO2 into account. CO2 concentration fixed at present-day value, i.e. run with transient historical CO2 up to the year 2000 and keep CO2 fixed at 369ppm thereafter.
	co2const	Keep CO2 concentration constant at 1971 level. For spin-up, use time-varying, historical CO2 concentration until 1970.
Agricultural management scenarios	firr (for crop models)	Full irrigation, i.e. until the soil is saturated.
	noirr (for crop models and hydrological models)	No irrigation. Both firr and noirr are required for the agro-economic models.
	cirr (for hydrological models)	Constrained irrigation.
	harmnon (for crop models)	Full fertilizer run, i.e. ensuring no N constraints, harmonized sowing and harvesting dates.
	fullharm (for crop models)	Full harmonization with regard to fertilizer application rates, sowing, and harvesting dates.
	default (for crop models)	“Best guess” representation of historical conditions regarding fertilizer application rates, sowing and variety settings that best reproduce given harvesting dates.

4.8 Spin-up

Simulation results should be reported from 1901 onwards. For models requiring spin-up, we provide 120 years of data produced by first removing a linear trend from the 1901-1930 portion of each forcing dataset, and then resampling these years at random (files <variable>_<climate_forcing>_1781_1900_spinup.nc4 in detrended/ subfolders). If more than 120 years of spin-up are needed, these data can be repeated as often as needed. For combined datasets (e.g. GSWP3-W5E5) use the spin-up dataset corresponding to the *first* part of the dataset.

After running the necessary length of spin-up data, continue with the respective climate forcing data set, starting in 1901. See Section 4.1 for more information on the historical climate forcing data sets.

Use historical CO_2 concentration, as provided in the input data archive (filename: /ISIMIP/ISIMIP2a/InputData/climate_co2/co2/historical_CO2_annual_1765_2018.txt), for the part of the spin-up directly preceding the reporting period. When using a longer spin-up period that (nominally) extends back further than 1765, please keep CO_2 concentration constant at 1765 level until reaching the year corresponding to 1765.

5 Reporting model results

5.1 Formatting, naming conventions and publication process

5.1.1 File names

Any model output files submitted to ISIMIP should follow the naming convention below. This naming convention is designed to be applicable across all sectors in ISIMIP. The file name is supposed to reflect the input data and scenario choices that went into the simulation; NOT to reflect specificities of individual models (these should be documented in the model database on the ISIMIP website). Please keep this in mind when preparing your files, and feel free to ask the coordination team in case of doubt.

File names consist of a series of identifiers, separated by underscores; see examples below. Things to note:

- Report **one** variable per file
- In filenames, use **lowercase** letters only
- Use underscore ("_") to separate identifiers
- Variable names consist of a single word without hyphens or underscores
- Use hyphens ("-") to separate strings within an identifier, e.g. in a model name
- NetCDF file extension is .nc4

In case of any questions, please contact the coordination team (info@isimip.org) before submitting files.
The file names should follow this convention for the **historical calibration and validation runs**:

<model-name>_<obs>_<bias-correction>_<clim-scenario>_<socio-econ-scenario>_<sens-scenarios>_<variable>_<region>_<timestep>_<start-year>_<end-year>.nc4

where the parts in brackets should be replaced with the appropriate identifier as in **Table** 7, and the order of identifiers should be respected. Identifiers may be dependent on the sector. The associated sectors are given in brackets in **Table** 7.
For example:

swim_watch_nobc_hist_nosoc_co2_dis_blue-nile-khartoum_daily_1971_2001.nc4 for the regional water sector and

lpjml_gswp3_nobc_hist_2005soc_co2_yield-mai-firr-default_global_annual_1971_2012.nc4 for a maize crop model run.

Table 7: File name specifiers for output data.

Item	Possible specifiers (use lowercase letters only!)	Explanation
<model-name>	*(your model name as registered with ISIMIP)*	Name of the impacts model. **IMPORTANT:** If you previously submitted data to ISIMIP2a and your model has been modified in the meantime, then • either run any new ISIMIP2a-extended simulations using the **same model version** as for the initial ISIMIP2a simulations, • or **assign a new name** to your updated model. E.g. if your model was called BESTModel1, you may call it BESTModel1b for the new uploads
<obs>	gswp3, princeton, watch, wfdei, gswp3-ewembi, gswp3-w5e5 (all sectors) localclim (Forests)	Name of the observations-based dataset providing the climate forcing data. For locally observed weather data from a weather station or similar which is in direct proximity of the forest stand and has been used for detailed model evaluation runs.
<bias-correction>	nobc	Indicates that no bias correction was performed on the climate data (e.g. observational data or ocean data).
<clim-scenario>	hist	Historical climate information.
<socio-econ-scenario>	nosoc pressoc varsoc nat	See **Table 6** in Section 4.7

<sens-scenario>	co2 noco2 co2const	See **Table 6** in Section 4.7
<variable>	*(variable names listed in the sector specific output tables)*	Output variable of the impact model. The identifier can also be used for information about the plant functional type (pft) in the biomes and permafrost sectors; the pft/species naming is model-specific and hence has to be reported in the impact-model database (www.isimip.org/impactmodels). In the forest sector, the identifier might contain information about the tree species; the species names codes are listed in Table 23. In the health sector, the identifier might contain information about the realization (se Section 14.2); the naming is model-specific and hence has to be reported in the impact model database (www.isimip.org/impactmodels).
	Plus a combination of the following settings for crop models (separated by dashes):	
	firr (for crop models) noirr (for crop models)	See **Table 6** in Section 4.7
	harmnon (for crop models) fullharm (for crop models) default (for crop models)	See **Table 6** in Section 4.7
<region>	global common name of river basin and gauge station as listed in Table 13. **Table** 13 forest site name as defined in Table 20**.** city/country/region name	For global simulations. For simulations covering one of several basins or a single location within a focus region (in the form "[river basin]-[station name]", e.g., "rhine-lobith"). For simulations of the regional forest models. For Health/Temperature-related mortality
<timestep>	daily, monthly, annual, decadal	
<start-year>_<end-year>		For the forest simulations, the full simulation period can be covered in one file; e.g. 1952-2012.

5.1.2 Format for gridded data

Gridded data should be returned in NetCDF4 format with a compression level of at least 5. It is important that you comply precisely with the formatting specified below, in order to facilitate the analysis of your simulation results in the ISIMIP framework. Incorrect formatting can seriously delay the analysis or lead to errors. For questions or clarifications, please contact the ISIMIP coordination team. Further information and instructions follow in this section and at the ISIMIP website (https://www.isimip.org/protocol/preparing-simulation-files/).

Global data are to be submitted for the ranges **-89.75** to **89.75** degrees latitude, and **-179.75** to **179.75** degrees longitude, i.e. 360 rows and 720 columns, or 259200 grid cells total. Please report the output data row-wise starting at 89.75 and -179.75, and ending at -89.75 and 179.75. The standard horizontal resolution is 0.5x0.5 degrees, corresponding to the resolution of the climate input data; with reporting intervals of 0.5 degrees_east for longitude, and of -0.5 degrees_north for latitude. Submitting data at lower resolution than 0.5x0.5 degrees is only encouraged in exceptional cases; in those cases, the above numbers will change accordingly (e.g., a 1x1 degree grid would have 180 rows, from 89.5 to -89.5 degrees latitude).

Antarctica and Greenland do not have to be simulated. If you are limited by data (e.g. soil data) you can also reduce the latitude range of your simulations, however, the **minimal latitude range** to be simulated is -60 to + 67 degrees. **Important:** When *reporting* results the whole grid as specified above should be reported – pixels you did not simulate should be filled with the missing value marker (1.e+20f).

Regional model teams should interpolate their output data to the same, common 0.5x0.5 degree grid as described above, and submit in the same NetCDF file format. Each file should **cover the entire globe** (even though the filename should contain the name of the region), with any grid cells outside the simulated region to be filled with missing values (1.e+20f). This will not lead to significantly larger files as long as NetCDF compression is used. **Exception:** Single (one-point) timeseries do not have to be embedded into the 0.5x0.5 degree grid, but should be reported as NetCDF4 files with the coordinates of the point included in the header information (e.g., lon : 172.84 degrees_east; lat : 47.08 degrees_north).

Note that submitting results in this format is important to facilitate comparison among different models and between global and regional scale. The **ISIMIP coordination team will be glad to assist** with the preparation of these files if necessary. **In addition** to the global file, regional model teams may submit a second file containing the output data in their default format. This may be e.g. on a finer resolution than 0.5°, on a non-regular grid, etc.

Please note the corresponding file naming conventions in Section 5.1.1.

5.1.3 NetCDF files: general conventions

Latitude, longitude and time should be included as individual variables in each file, following the conventions of **Table** 8. When defining the variable in a file, time should be the first dimension; e.g., albedo(time,lat,lon).

Table 8: Naming and format conventions for NetCDF files.

Dimension	Name	Unit
X	Lon	degrees east
Y	Lat	degrees north
T	Time	<time steps> since 1901-01-01 00:00:00 (where <time steps> is replaced by days, months, etc., according to the time step the data is reported on) Note: crop models use a different time step; see Section 10.3.
missing value	1.e+20f	
fill value	1.e+20f	

Once you have created your NetCDF file, you can check the metadata by running the command ncdump -h; an example output is given on the ISIMIP website, www.isimip.org (Protocol -> Important information about preparing your simulation files).

5.1.4 Format for non-gridded data

Data that is not defined on a grid, such as point-based data (e.g. for particular gauges or data for world regions), should nonetheless be reported in NetCDF format (e.g. as a separate file for each gauge or region), each file containing a single time series. The ISIMIP coordination team will assist with the preparation of these files where necessary. **It is important that all ISIMIP results exist in NetCDF format**, in order to be compatible with the structure and functionalities of the ESGF repository.

5.1.5 Variables with layers

For variables that can be simulated on multiple **fixed layers** (e.g. variables with DBH classes in Forest Models, or with fixed depth layers in the Lakes sector), we require the following:

- The level dimension has a specific name per sector; i.e. 'levlak' for the lakes sector, 'lev' for the water sector or 'dbh_class' for forest models
- For variables where a level cover a range, like soil depth or dbh class, specify the lower and upper boundaries of every layer, with data corresponding to the midpoint of each layer (e.g. dbh class or depth layer)
    - The boundaries of the top/bottom layers will end up in a variable called 'depth_bnds', and the midpoint will be in 'depth'
    - 'depth' and 'depth_bnds' are double

- For variables where a level has the function of an index, it is not necessary to indicate boundaries
- For variables where it is possible to have layers or not (e.g. variable “harv” in Forest Models), add global attribute ’dbhclass_profile’ and use label ‘true’ if the file contains layers (e.g. multiple dbh classes) or ‘false’ depending on the case

For variables simulated on **depth layers varying** over time and/or over space (e.g. soil moisture on layers that can get deeper or shallower over time, or have different depths at different locations), we distinguish between variables where the levels vary per grid cell, and variables where the levels vary over time. For such variables, we additionally require the following:

- For variables where depth of layers varies over time, add global attribute ’time_varying_soil_layer_depth’ and use label ‘true’ or ‘false’ depending on the case
- For variables where depth of layers varies per grid cell, add global attribute ‘location_varying_soil_layer_depth’ and use label ‘true’ or ‘false’ depending on the case

More information and example outputs are available on the ISIMIP website, www.isimip.org (Protocol -> Important information about preparing your simulation files).

5.1.6 Time intervals

Please submit your output data in chunks according to the **Table** 9 depending on the time step of the output variable you are reporting (the requested time step for each variable is listed in the sector-specific tables).

Table 9: Definition of time intervals

Daily time step	For files holding *global* daily data, files should cover 10 years starting in the second year of a decade and end in the first year of the next decade (e.g. 1991-2000). If the time period starts after the second year of the decade or ends before the first year of the new decade, the start or end year of the time period should be used as the start or end year of the file respectively. *Non-global* daily data should be submitted for the entire simulation period in single files per variable.
monthly, annual, or decadal time step	Output should be reported in one single time series file per experiment.
30-year averages (biodiversity	Output should be reported in one single file per period.

only)	
5-year period (health)	Output should be reported in one single file per period.
growing-season (agriculture)	Output should be reported in one single file per period. Time dimension is replaced by a unitless coordinate variable with integer values, or counter, named 'growing-season', indicating the number of growing season since starting year of the period.

5.1.7 Submission of simulations

Data should be submitted to a dedicated file system on a central server located at DKRZ Hamburg. More information on how to access this server and on formatting of your files is available on the ISIMIP website at https://www.isimip.org/protocol/preparing-simulation-files/. The ISIMIP coordination team will gladly provide assistance upon request.

5.1.8 Format checks

The **ISIMIP data management team will check** the formatting of the files submitted to the server at DKRZ Hamburg. File naming, variable and dimensions definitions and units, grid formatting, time axis, coverage of simulation period and global metadata are reviewed and corrected when possible. Files with severe errors will be reported and a following submission will be requested. Files passing the checks will be published in the OutputArea of DKRZ. Further information on the checks performed can be found at https://www.isimip.org/protocol/preparing-simulation-files/#quality-check-of-your-simulation-data.

5.2 Documentation of models and experiments

In addition to adhering to the common settings described in this protocol, it is essential to keep track of further details regarding the configuration of each individual model. This ensures that the simulation results can appropriately be interpreted by authors of multi-model studies, and that these can remain transparent and usable for a long time into the future. For this purpose, the ISIMIP coordination team provides a questionnaire that all modelling teams are asked to answer. The questionnaire is accessible online through the ISIMIP website; for assistance please write to Info@isimip.org.

6 Water

6.1 Experiments

Table 10 provides an overview of all experiments to be run in the water sector in ISIMIP2a.

Table 10: Summary of experiments for water models.

Climate Data	Scenario	Human Impacts	Other settings (sens-scenario)	# runs
WATCH-WFDEI	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
GSWP3-W5E5	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
GSWP3-EWEMBI	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
GSWP3	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
PGMFD v.2 (Princeton)	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
WATCH (WFD)	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
Additional sector-specific run: PGMFD v.2 (Princeton)	Hist	varsoc	constant CO2 at 1971 levels (co2const)	1

6.2 Sector-specific input data

In ISIMIP2a, hydrological modelling teams are asked to take into account the historical evolution of irrigated areas, dams and reservoirs, in order to obtain a more realistic estimate of the historical evolution of runoff and discharge. The data sources to be used are listed in **Table** 11, along with a soil and vegetation dataset that may be used optionally.

Table 11: Input data to be used for the historical runs (ISIMIP2a), in addition to the common data listed in Section 4.

Dataset	Description	More info	Scale	Variables included, comments
Mandatory (if feasible)				
Dams/Reservoirs		See **Table** 5 (Other human influences) http://www.gwsp.org/products/grand-database.html		
DDM30 routing network, mapped to the CRU land mask	flow directions, slope, and basin numbers	Note: The routing network includes large lakes that are not included in the provided land mask. These cells should not be included when results are submitted and there should be no runoff added to the river network from these cells. I.e. these cells are included only for transportation purposes (streamflow).	global, 0.5°	**for global models only**[3]
Optional (does not have to be harmonized)				
HWSD, or GSWP3 (upscaled version of HWSD)	soil map	See http://hydro.iis.u-tokyo.ac.jp/~sujan/research/gswp3/soil-texture-map.html, upscaling method A. Each model does have the option to use their own soil datasets if they prefer.	global, 30 arc sec (HWSD) or 0.5° (GSWP3), fixed	soil type
GLIMS (Global Land Ice Measurements from Space)	Glacier distribution	See http://www.glims.org/About/		
HydroSHEDS	Topography/routing network	Hydrographically corrected SRTM data. Available in 3 resolutions, includes accumulated upstream area. Also,		**for regional models only**[3]

[3] To allow a direct intercomparison of river flows between global and regional models on a gridded basis, the runoff produced by the global models could be collected and routed through the HydroSHEDS network as a post-processing step, using a single routing model. Volunteers for this task are welcome.

		HydroSHEDS is not available north of 60 degrees, due to limitations in the SRTM data at high latitudes.		

6.3 Output data

Note that variable names are chosen to comply, where feasible, with the ALMA convention[4] and the names used in WATCH/WaterMIP. Although variable names are mixed-case here, make sure to use **only lower-case** letters in the output filenames (see Section 5.1.1).
All variables are to be reported as time-averages with the indicated resolution; do not report instantaneous values ('snapshots'). An exception is **maxdis**, which is the maximum daily-average discharge in a given month, to be reported on a monthly basis (see below).
Water balance equation in terms of requested output variables:

rainf + snowf = evap + qtot,

where **Evap** is the sum of interception, transpiration, sublimation, and evaporation from the surface. This equation only holds on timescales long enough for changes in water storage (e.g. in soil and groundwater) to average out.

IMPORTANT: Some output variables reported for the water sector are also appropriate for use in the permafrost sector described in Section 11*; these are marked with an *. If you plan to submit simulations for the permafrost sector, note that additional variables are also required for the permafrost sector (see* **Table** 28*).*

Table 12: Output variables to be reported by water sector models. Variables highlighted in orange are requested from both global and regional models, if computed; variables highlighted in purple are requested only from regional models; others only from global models.

[4] http://www.lmd.jussieu.fr/~polcher/ALMA/convention_output_3.html

Variable (long name)	Variable name	Unit (NetCDF format)	Resolution	Comments
Hydrological Variables				
*Runoff	**qtot**	kg m-2 s-1	daily* (0.5°x0.5°)	Total (surface + subsurface) runoff (qtot = qs + qsb). *if daily resolution not possible, please provide monthly[5]. Please also deliver for the permafrost sector.
Surface runoff	**qs**	kg m-2 s-1	monthly (0.5°x0.5°)	Water that leaves the surface layer (top soil layer) e.g. as overland flow / fast runoff.
Subsurface runoff	**qsb**	kg m-2 s-1	monthly (0.5°x0.5°)	Sum of water that flows out from subsurface layer(s) including the groundwater layer (if present). Equals qg in case of a groundwater layer below only one soil layer.
Groundwater recharge	**qr**	kg m-2 s-1	monthly (0.5°x0.5°)	Water that percolates through the soil layer(s) into the groundwater layer. In case seepage is simulated but no groundwater layer is present, report seepage as qr and qg.
Groundwater recharge	**qr**	kg m-2 s-1	monthly (average for basin until gauge location)	Water that percolates through the soil layer(s) into the groundwater layer. In case seepage is simulated but no groundwater layer is present, report seepage as qr and qg.
Groundwater runoff	**qg**	kg m-2 s-1	monthly (0.5°x0.5°)	Water that leaves the groundwater layer. In case seepage is simulated but no groundwater layer is present, report seepage as qr and qg.
Discharge (gridded)	**dis**	m3 s-1	daily* (0.5°x0.5°)	*if daily resolution not possible, please provide monthly.
Discharge (gauge level)	**dis**	m3 s-1	daily* (at gauge locations; see **Table** 13)	*if daily resolution not possible, please provide monthly.
Monthly maximum of daily discharge	**maxdis**	m3 s-1	monthly (0.5°x0.5°)	Optional variable – please report if daily discharge data is not reported.
Monthly minimum of daily discharge	**mindis**	m3 s-1	monthly (0.5°x0.5°)	Optional variable – please report if daily discharge data is not reported.

[5] If storage issues keep you from reporting daily data, please contact the ISIMIP team to discuss potential solutions.

Evapotranspiration	**evap**	kg m-2 s-1	monthly (0.5°x0.5°)	Sum of transpiration, evaporation, interception and sublimation.
Evapotranspiration	**evap**	kg m-2 s-1	monthly (average for basin until gauge location)	Sum of transpiration, evaporation, interception losses, and sublimation.
Potential Evapotranspiration	**potevap**	kg m-2 s-1	monthly (0.5°x0.5°)	As evap, but with all resistances set to zero, except the aerodynamic resistance.
Potential Evapotranspiration	**potevap**	kg m-2 s-1	monthly (average for basin until gauge location)	As evap, but with all resistances set to zero, except the aerodynamic resistance.
*Soil moisture	**soilmoist**	kg m-2	monthly (0.5°x0.5°)	Please provide soil moisture for all depth layers (i.e. 3D-field), and indicate depth in m. If depth varies over time or space, see instructions in Section 5.1.5. Please also deliver for the permafrost sector.
Soil moisture	**soilmoist**	kg m-2	monthly (average for basin until gauge location)	Please provide soil moisture for all depth layers (i.e. 3D-field), and indicate depth in m. If depth varies over time or space, see instructions in Section 5.1.5. Please also deliver for the permafrost sector.
Soil moisture, root zone	**rootmoist**	kg m-2	monthly (0.5°x0.5°)	Total simulated soil moisture available for evapotranspiration. If simulated by the model. Please indicate the depth of the root zone for each vegetation type in your model. If depth varies over time or space, see instructions in Section 5.1.5.
Frozen soil moisture for each layer	**soilmoistfroz**	kg m-2	monthly (0.5°x0.5°)	Soil_frozen_water_content This variable only for the purposes of the permafrost sector.
*Temperature of Soil	**tsl**	K	daily* (0.5°x0.5°)	Temperature of each soil layer. Reported as "missing" for grid cells occupied entirely by "sea". **THIS IS THE MOST IMPORTANT VARIABLE FOR THE PERMAFROST SECTOR.**

				Also need depths in meters. Daily would be great, but otherwise monthly would work. If depth varies over time or space, see instructions in Section 5.1.5. This variable only for the purposes of the permafrost sector. *if daily resolution not possible, please provide monthly.
*Snow depth	**snd**	m	monthly (0.5°x0.5°)	Grid cell mean depth of snowpack. This variable only for the purposes of the permafrost sector.
*Snow water equivalent (= snow water storage)	**swe**	kg m-2	monthly (0.5°x0.5°)	Total water mass of the snowpack (liquid or frozen), averaged over a grid cell. Please also deliver for the permafrost sector.
Total water storage	**tws**	kg m-2	monthly (0.5°x0.5°)	Mean monthly water storage in all compartments. Please indicate in the netcdf metadata which storage compartments are considered.
Canopy water storage	**canopystor**	kg m-2	monthly (0.5°x0.5°)	Mean monthly water storage in the canopy.
Glacier storage	**glacierstor**	kg m-2	monthly (0.5°x0.5°)	Mean monthly water storage in glaciers.
Groundwater storage	**groundwstor**	kg m-2	monthly (0.5°x0.5°)	Mean monthly water storage in groundwater layer.
Lake storage	**lakestor**	kg m-2	monthly (0.5°x0.5°)	Mean monthly water storage in lakes (except reservoirs).
Wetland storage	**wetlandstor**	kg m-2	monthly (0.5°x0.5°)	Mean monthly water storage in wetlands.
Reservoir storage	**reservoirstor**	kg m-2	monthly (0.5°x0.5°)	Mean monthly water storage in reservoirs.

River storage	**riverstor**	kg m-2	monthly (0.5°x0.5°)	Mean monthly water storage in rivers.
*Annual maximum thaw depth	**thawdepth**	m	monthly (0.5°x0.5°)	Calculated from daily thaw depths.
River temperature	**triver**	K	monthly (0.5°x0.5°)	Mean monthly water temperature in river (representative of the average temperature across the channel volume).
Rainfall	**rainf**	kg m-2 s-1	monthly (0.5°x0.5°)	These variables are required for test purposes only. If you need to reduce output data volumes, please provide these variables only once, with the first (test) data set you submit, e.g. for the first decade of each experiment. NOTE: rainf + snowf = total precipitation
Snowfall	**snowf**	kg m-2 s-1	monthly (0.5°x0.5°)	
Water management variables (for models that consider water management/human impacts)				
Irrigation water demand (=potential irrigation water withdrawal)	**pirrww**	kg m-2 s-1	monthly (0.5°x0.5°)	Irrigation water withdrawal, assuming unlimited water supply.
Actual irrigation water withdrawal	**airrww**	kg m-2 s-1	monthly (0.5°x0.5°)	Irrigation water withdrawal, taking water availability into account; please provide if computed.
Potential irrigation water consumption	**pirruse**	kg m-2 s-1	monthly (0.5°x0.5°)	Portion of withdrawal that is evapo-transpired, assuming unlimited water supply.
Actual irrigation water consumption	**airruse**	kg m-2 s-1	monthly (0.5°x0.5°)	Portion of withdrawal that is evapotranspired, taking water availability into account; if computed.
Actual green water consumption on irrigated cropland	**airrusegreen**	kg m-2 s-1	monthly (0.5°x0.5°)	Actual evapotranspiration from rainwater over irrigated cropland; if computed.
Potential green water consumption on irrigated cropland	**pirrusegreen**	kg m-2 s-1	monthly (0.5°x0.5°)	Potential evapotranspiration from rainwater over irrigated cropland; if computed and different from AIrrUseGreen.
Actual green water consumption on	**arainfusegreen**	kg m-2 s-1	monthly (0.5°x0.5°)	Actual evapotranspiration from rainwater over rainfed cropland; if computed.

rainfed cropland				
Actual domestic water withdrawal	**adomww**	kg m-2 s-1	monthly (0.5°x0.5°)	If computed.
Actual domestic water consumption	**adomuse**	kg m-2 s-1	monthly (0.5°x0.5°)	If computed.
Actual manufacturing water withdrawal	**amanww**	kg m-2 s-1	monthly (0.5°x0.5°)	If computed.
Actual Manufacturing water consumption	**amanuse**	kg m-2 s-1	monthly (0.5°x0.5°)	If computed.
Actual electricity water withdrawal	**aelecww**	kg m-2 s-1	monthly (0.5°x0.5°)	If computed.
Actual electricity water consumption	**aelecuse**	kg m-2 s-1	monthly (0.5°x0.5°)	If computed.
Actual livestock water withdrawal	**aliveww**	kg m-2 s-1	monthly (0.5°x0.5°)	If computed.
Actual livestock water consumption	**aliveuse**	kg m-2 s-1	monthly (0.5°x0.5°)	If computed.
Total (all sectors) actual water consumption	**atotuse**	kg m-2 s-1	monthly (0.5°x0.5°)	Sum of actual water consumption from all sectors. Please indicate in metadata which sectors are included.
Total (all sectors) actual water withdrawal	**atotww**	kg m-2 s-1	monthly (0.5°x0.5°)	Sum of actual water withdrawal from all sectors. Please indicate in metadata which sectors are included.
Total (all sectors) potential water withdrawal	**ptotww**	kg m-2 s-1	monthly (0.5°x0.5°)	Sum of potential (i.e. assuming unlimited water supply) water withdrawal from all sectors. Please indicate in metadata which sectors are included.

Total (all sectors) potential water consumption	**ptotuse**	kg m-2 s-1	monthly (0.5°x0.5°)	Sum of potential (i.e. assuming unlimited water supply) water consumption from all sectors. Please indicate in metadata which sectors are included.
Static output (Note: data that cannot be submitted in NetCDF format may be submitted in another suitable format directly via email to Info@isimip.org)				
Natural vegetation types	**Names to be coordinated with biomes/ecosystem sector**	N/A	static (0.5°x0.5°)	Map of natural vegetation / land surface types as used by the model. Please include a description of the parameters and their values associated with these vegetation types (parameter values could be supplied as spatial fields where appropriate). In your description please also provide details of the evapotranspiration scheme used by your model.
Soil types	**soil**		static (0.5°x0.5°)	Soil types or texture classes as used by your model. Please include a description of each type or class, especially if these are different from the standard HSWD and GSWP3 soil types. Please also include a description of the parameters and values associated with these soil types (parameter values could be submitted as spatial fields where appropriate).
Leaf Area Index	**lai (to be coordinated with other sectors)**		static (0.5°x0.5°) or monthly (0.5°x0.5°) where appropriate	If used by or computed by the model.
Agricultural variables (optional output for all water models that also simulate crop yields)				
Crop yields (dry matter)	**yield-<crop>-<irrigation setting>**	dry matter (t ha-1 per growing season)	per growing season (0.5°x0.5°)	Irrigation setting = "cirr" for "constrained irrigation" or "noirr" for rainfed.
Actual planting dates	**plantday-<crop>-<irrigation setting>**	day of year	per growing season (0.5°x0.5°)	Julian dates.
Actual planting year	**plantyear-<crop>-<irrigation setting>**	year of planting	per growing season (0.5°x0.5°)	This allows for clear identification of planting that is also easy to follow for potential users from outside the project.

Anthesis dates	**anthday-<crop>-<irrigation setting>**	day of year of anthesis	per growing season (0.5°x0.5°)	Together with the year of anthesis added to the list of outputs (see below) it allows for clear identification of anthesis that is also easy to follow for potential users from outside the project.
Year of anthesis	**anthyear-<crop>-<irrigation setting>**	year of anthesis	per growing season (0.5°x0.5°)	It allows for clear identification of anthesis that is also easy to follow for potential users from outside the project.
Maturity dates	**matyday-<crop>-<irrigation setting>**	day of year of maturity	per growing season (0.5°x0.5°)	Together with the year of maturity added to the list of outputs (see below) it allows for clear identification of maturity that is also easy to follow for potential users from outside the project.
Year of maturity	**matyyear-<crop>-<irrigation setting>**	year of maturity	per growing season (0.5°x0.5°)	It allows for clear identification of maturity that is also easy to follow for potential users from outside the project.
Nitrogen application rate	**initr-<crop>-<irrigation setting>**	kg ha-1 per growing season	per growing season (0.5°x0.5°)	Total nitrogen application rate. If organic and inorganic amendments are applied, rate should be reported as inorganic nitrogen equivalent (ignoring residues).
Above-ground biomass (dry matter)	**biom-<crop>-<irrigation setting>**	Dry matter (t ha-1 per growing season)	per growing season (0.5°x0.5°)	The whole plant biomass above ground.
Soil carbon emissions	**sco2-<crop>-<irrigation setting>**	kg C ha-1	per growing season (0.5°x0.5°)	Ideally should be modeled with realistic land-use history and initial carbon pools. Subject to extra study.
Nitrous oxide emissions	**sn2o-<crop>-<irrigation setting>**	kg N2O-N ha-1	per growing season (0.5°x0.5°)	Ideally should be modeled with realistic land-use history and initial carbon pools. Subject to extra study.

* If storage issues keep you from reporting daily data, please contact the ISIMIP team to discuss potential solutions.

Comments related to the optional agricultural outputs

The reporting of the crop yield-related outputs differs from the reporting of other variables in the water sector, as it is not done according to calendar years but according to **growing seasons** to resolve potential multiple harvests. See the agriculture section (section 10) for details. Simulations should be provided for the four major **crops** (wheat, maize, soy, and rice) but output for other crops and also bioenergy crops is highly welcome, too; see Section 10 for crop naming.

Yields simulations provided in the water sector should account for **irrigation water constraints**. For each crop, yields should be reported separately for irrigated land (cirr for "constrained irrigation") and rainfed land (noirr). This complements the full irrigation (firr) pure crop runs requested in the agriculture part of the protocol (Section 10).
Those models that cannot simulate time varying management/human impacts/fertilizer input should keep these fixed at year 2000 levels throughout the simulations.

6.4 Additional information for regional hydrological models

CALIBRATION: Please use WFDEI (from 1979 to 2016) for calibration, for all simulations.

Table 13**:** Catchment gauging stations for reporting regional hydrological model results.

River Basin (short name for filenames)	Station for calibration and validation (short name for filenames)	Coordinates Lat/Lon	GRDC Station Code	Data availability (monthly discharge)	Data availability (daily discharge)	Area upstream of gauge (km^2) according to GRDC or GIS
Amazon (amazon)	São Paulo de Olivenca (sao-paulo-de-olivenca)	-3.45/-68.75	3623100	1979-1993	1973-2010	990781
Blue Nile (blue-nile)	El-Deim, Sudan Border (el-diem) Khartoum (khartoum)	11/35 15.62/32.55	n.a.* 1663100	1961-2002 1900-1982	n.a. n.a.	160000 325000
Danube (danube)	Wien-Nußdorf (wien-nussdorf)	48.25/16.3	6242500	1828-1899	1900-to date	101700
Ganges (ganges)	Farakka (farakka)	25/87.92	2846800	1949-1973	n.a.	835000
Godavari (godavari)	Tekra (tekra)	19/80	n.a.	1964-2017	1964-2017	119781
Indus	Tarbela Reservoir (tarbela)	72.86/ 34.33	n.a.	2000-2016	2000-2016	173345
Lena (lena)	Krestovski (krestovski) Stolb (stolb)	59.73/113.17	2903427	1936-2002	1936-1999	440000

		72.37/126.8	2903430	1978-1994	1951-2002	2460000
Mackenzie (mackenzie)	Artic Red River (artic-red-river)	67.4583/-133.745	4208025	1972-1996	1972-2015	1660000
Mississippi	Alton (alton)	38.885/-90.1809	4119800	1928-1984	1933-1987	444185
Murray Darling (darling)	Louth (louth)	-30.5318/ 145.1144	5204250	1954-2000	1954-2008	489300
Niger (niger)	Dire (dire) Koulikoro (koulikoro) Lokoja (lokoja) Tossaye (tossaye)	16.2667/-3.3833 12,8667/-7,55 7,8/6,7667 16.9416/ -0.579166	1134700 1134100 1834101 1134850	1924-2012 1907-2012 2007-2012 1954-1992	1924-2003 1907-2006 1970-2006 1954-1992	340000 120000 2074171 348000
Pajeú (pajeu)	Floresta (floresta)	-8,6089,-38,5767	n.a. (National system for information on water resources, Brasil)	n.a.	n.a.	12266
Rhine (rhine)	Lobith (lobith)	51.84/6.11	6435060	1901-1996	1901-2010	160800
Tagus (tagus)	Almourol (almourol) Trillo (trillo)	39.47/-8.37 40.7/-2.58	6113050 6213800	1973-1990 1977-1984	1982-1990 1977-1984	61490 3253
Yangtze	Cuntan (cuntan)	29,616667/106,6	n.a.	1987-2006	1987-2006	804859
Yellow, Huang He (yellow)	Tangnaihai (tangnaihai)	35.5/100.15	n.a.	1971-2002	1971-2002	121000

Note: If GRDC station is not available, the data availability is indicated for data from other sources; *GRDC data reported as poor

7 Lakes

7.1 Experiments

Simulations of climate-change effects on lakes will be made using coupled lake-hydrodynamic and water-quality models. Models can operate on the global scale (uncalibrated) or on a number of case-study lakes (calibrated). Both global and local models will conduct the same set of simulations.

Table 14: Summary of experiments for lake models.

Climate Data	Scenario	Human Impacts	Other settings (sens-scenario)	# runs
WATCH-WFDEI	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
GSWP3-W5E5	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
GSWP3-EWEMBI	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
GSWP3	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
PGMFD v.2 (Princeton)	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3
WATCH (WFD)	Hist	nosoc pressoc varsoc	historical CO2 (co2)	3

See **Table** 6 and **Table** 7 for an explanation of the nosoc, pressoc, and varsoc experiments. Depending on whether and how human influences are included, a given model may not be able to run all three experiments.

7.2 Sector-specific input data

Global lake models

Global-scale simulations should be performed either assuming a lake present in every pixel or using grid-scale lake fraction based on the Global Lake and Wetland Database (GLWD) (Lehner & Döll, 2004) and available on the DKRZ input data repository at /work/bb0820/ISIMIP/ISIMIP2a/InputData/lakes/pctlake.nc4 (Subin, Riley, & Mironov, 2012). Since a 0.5°x0.5° pixel potentially contains multiple lakes with different characteristics (e.g. in terms of bathymetry, transparency, fetch), it is not possible to fully represent this subgrid-scale heterogeneity. Instead, the global-scale lake simulations should represent a 'representative lake' for a given pixel. Consequently, no stringent requirement is imposed with respect to lake depth, light extinction coefficient or initial conditions.

For lake depth, modellers are encouraged to use the data from the Global Lake Data Base (GLDB). A regridded lake depth field based on GLDBv1 (Kourzeneva, 2010) is available at 0.5°x0.5° resolution on the DKRZ input data repository at /work/bb0820/ISIMIP/ISIMIP2a/InputData/lakes/lakedepth.nc4; this field was aggregated from 30 arc sec to 1.9°x2.5° and then interpolated again to 0.5°x0.5° (Subin, Riley, & Mironov, 2012), but modellers may choose to use the more recent GLDBv2 available at 30 arc sec (http://www.flake.igb-berlin.de/ep-data.shtml) (Choulga, Kourzeneva, Zakharova, & Doganovsky, 2014). Modellers are requested to document their approach regarding lake depth, light extinction coefficient and initial conditions in the ISIMIP Impact Model Database (www.isimip.org/impactmodels). In case the lake model has no built-in calculation of the light extinction coefficient, modellers may consider using the parameterisation proposed by (Shatwell, Thiery, & Kirillin, 2019): extcoeff = 5.681 * max(depth,1) ^(-0.795), derived from a collection of 1258 lakes, or the parameterisation proposed by (Håkanson, 1995): extcoeff = 1.1925 * max(lakedepth,1)^(-0.424), derived from 88 Swedish glacial lakes. Yet it should be noted that modellers are free to decide how to represent extinction coefficient.

Local lake models

Simulations will be made for case-study lakes selected based on the availability of high-quality meteorological and limnological observations, thereby aiming for a good spread across climates and lake types. Model inputs consist of the meteorological variables given in **Table** 1, water inputs from hydrological model simulations, and nutrient loads estimated using simple loading function (Haith & Shoemaker., 1987) (Schneiderman, Pierson, Lounsbury, & Zion, 2002) or statistical estimation procedures. In addition, site-specific data will be needed such as lake bathymetry data. Direct climate effects on lakes that influence factors such as water temperature stratification period, mixing depth etc. will be simulated using climate scenarios shown in **Table** 14, and water inflows from hydrologic model simulations based on the experiments described in Section 6. Lake water quality simulations, which affect factors such as phytoplankton and nutrient levels, will also need to include simple nutrient loading inputs linked to the hydrologic model simulations.

Reporting
All variables are to be reported as time-averages with the indicated resolution.
For depth-varying variables, data should be provided either as fully resolved vertical profiles, or, if that is not possible, as a mean of the epilimnion or mixed layer (“mean epi”) and mean of the hypolimnion (“mean hypo”). When the lake is simulated as completely mixed or isothermal, the mean of the entire water column is assigned to the epilimnion, and the hyolimnion concentration is set to a missing value.
See section 5.1.5 for further information on file formatting.

Diagnostic for lake stratification
As density is a non-linear function of temperature and a global analysis requires examination of a wide range of lake temperatures it is preferable to use a density-derived definition of stratification to a purely temperature-related definition, as follows:
Calculate density (ρ) from temperature using the formula (Millero & Poisson, 1981):
ρ = 999.842594 + (6.793952 x 10-2 t) – (9.095290 x 10-3 t2) + (1.001685 x 10-4 t3) – (1.120083 x 10-6 t4) + (6.536336 x 10-9 t5),
where t is water temperature of the lake layer in °C.
Define the lake to be stratified whenever the density difference between the surface and the bottom of the lake is greater than 0.1 kg m-3. Note this definition does not distinguish between ‘normal’ and ‘reverse’ stratification. Reverse stratification means that the surface is colder than the bottom, but the surface water density is less than the maximum density of water, found particularly under ice. While a separate step can be used to distinguish these events by assessing whether the surface temperature is greater than or less than 3.98 °C, this separation is not requested by the protocol.
Note that the range of model outputs will vary from model to model. Below are generic outputs that capture the basic information provided by most lake-eutrophication models. Modelling groups whose models do not provide all information listed here are invited to report on the reduced set of variables implemented in their models.

7.3 Output Data

Table 15: Output variables to be reported by lake models.

Variable (long name)	Variable name	Unit (NetCDF format)	Spatial Resolution	Temporal Resolution	Depth Resolution	Comments
Hydrothermal Variables						
Thermal stratification	**strat**	*unitless*	Representative lake associated with grid	Daily	None	1 if lake grid cell is thermally stratified 0 if lake grid cell is not thermally

			cell			stratified
Depth of Thermocline	**thermodepth**	m	Representative lake associated with grid cell	Daily	None	Depth corresponding the maximum water density gradient
Water temperature	**watertemp**	K	Representative lake associated with grid cell	Daily	Full Profile	Simulated water temperature. Layer averages and full profiles. See Section 5.1.5 for details on reporting
Surface temperature	**surftemp**	K	Representative lake associated with grid cell	Daily (monthly)	None	Average of the upper layer in case not simulated directly
Bottom temperature	**bottemp**	K	Representative lake associated with grid cell	Daily (monthly)	None	Average of the lowest layer in case not simulated directly
Lake ice cover	**ice**	*unitless*	Representative lake associated with grid cell	Daily	None	1 if ice cover is present in lake grid cell 0 if no ice cover is present in lake grid cell
Lake layer ice mass fraction	**lakeicefrac**	*unitless*	Representative lake associated with grid cell	Daily (monthly)	Mean Epi	Fraction of mass of a given layer taken up by ice
Ice thickness	**icethick**	m	Representative lake associated with grid cell	Daily (monthly)	None	
Snow thickness	**snowthick**	m	Representative lake associated with grid cell	Daily (monthly)	None	
Temperature at the ice upper surface	**icetemp**	K	Representative lake associated with grid cell	Monthly	None	
Temperature at the snow upper surface	**snowtemp**	K	Representative lake associated with grid	Monthly	None	

			cell			
Sensible heat flux at the lake-atmosphere interface	**sensheatf**	W m-2	Representative lake associated with grid cell	Daily (monthly)	None	At the surface of snow, ice or water depending on the layer in contact with the atmosphere. **Positive if upwards.**
Latent heat flux at the lake-atmosphere interface	**latentheatf**	W m-2	Representative lake associated with grid cell	Daily (monthly)	None	**See sensible heat flux**
Momentum flux at the lake-atmosphere interface	**momf**	kg m-1 s-2	Representative lake associated with grid cell	Daily (monthly)	None	**See sensible heat flux**
Upward shortwave radiation flux at the lake-atmosphere interface	**swup**	W m-2	Representative lake associated with grid cell	Daily (monthly)	None	**See sensible heat flux.** Not to be confused with net shortwave radiation
Upward longwave radiation flux at the lake-atmosphere interface	**lwup**	W m-2	Representative lake associated with grid cell	Daily (monthly)	None	**See sensible heat flux.** Not to be confused with net longwave radiation
Downward heat flux at the lake-atmosphere interface	**lakeheatf**	W m-2	Representative lake associated with grid cell	Daily (monthly)	None	**See sensible heat flux** the residual term of the surface energy balance, i.e. the net amount of energy that enters the lake on daily time scale: lakeheatf = swdown - swup + lwdown - lwup - sensheatf - latenheatf (terms defined positive when directed upwards)
Turbulent diffusivity of heat	**turbdiffheat**	m2 s-1	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	Only if computed by the model. See Section 5.1.5 for details on reporting

Surface albedo	**albedo**	*unitless*	Representative lake associated with grid cell	Daily (monthly)	None	Albedo of the surface interacting with the atmosphere (water, ice or snow)
Light extinction coefficient	**extcoeff**	m-1	Representative lake associated with grid cell	Constant	None	only to be reported for global models, local models should use extcoeff as input
Sediment upward heat flux at the lake-sediment interface	**sedheatf**	W m-2	Representative lake associated with grid cell	Daily (monthly)	None	Positive if upwards. Only if computed by the model
Water Quality Variables						
Chlorophyll Concentration	**chl**	g-3 m-3	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	Total water chlorophyll concentration – indicator of phytoplankton. See Section 5.1.5 for details on reporting
Phytoplankton Functional group biomass	**phytobio**	mole m-3 as carbon	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	Different models will have different numbers of functional groups so that the reporting of these will vary by model. See Section 5.1.5 for details on reporting
Zoo plankton biomass	**zoobio**	mole m-3 as carbon	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	Total simulated Zooplankton biomass. See Section 5.1.5 for details on reporting
Total Phosphorus	**tp**	mole m-3	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	See Section 5.1.5 for details on reporting
Particulate Phosphorus	**pp**	mole m-3	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	See Section 5.1.5 for details on reporting
Total Dissolved	**tpd**	mole m-3	Representative lake	Daily (monthly)	Either full	Some models may also output data for

Phosphorus			associated with grid cell		profile, or mean epi and mean hypo	soluble reactive phosphorus (SRP). See Section 5.1.5 for details on reporting
Total Nitrogen	**tn**	mole m-3	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	See Section 5.1.5 for details on reporting
Particulate Nitrogen	**pn**	mole m-3	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	See Section 5.1.5 for details on reporting
Total Dissolved Nitrogen	**tdn**	mole m-3	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	Some models may also output data for Nitrate (N02) nitrite (NO3) and ammonium (NH4). See Section 5.1.5 for details on reporting
Dissolved Oxygen	**do**	mole m-3	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	See Section 5.1.5 for details on reporting
Dissolved Organic Carbon	**doc**	mole m-3	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	Not always available. See Section 5.1.5 for details on reporting
Dissolved Silica	**si**	mole m-3	Representative lake associated with grid cell	Daily (monthly)	Either full profile, or mean epi and mean hypo	Not always available. See Section 5.1.5 for details on reporting

7.4 Additional information for local lake models

7.4.1 Lake sites

Table 16: Lake site specifications for local lake models. A document with additional information is maintained by the sector coordinators and provided at https://docs.google.com/spreadsheets/d/1UY_KSR02o7LtmNoOs6jOgOxdcFEKrf7MmhR2BYDlm-Q/edit#gid=555498854.

Lake name	Lake name in file name (reporting)	Reservoir or lake?	Country	Latitude (dec deg)	Longitude (dec deg)
Allequash Lake	allequash	lake	USA	46,04	-89,62
Alqueva Reservoir	alqueva	reservoir	Portugal	38,20	-7,49
Lake Annecy	annecy	lake	France	45,87	6,17
Lake Annie	annie	lake	USA	27,21	-81,35
Lake Argyle	argyle	reservoir	Australia	-16,31	128,68
Lake Biel	biel	lake	Switzerland	47,08	7,16
Big Muskellunge Lake	big-muskellunge	lake	USA	46,02	-89,61
Black Oak Lake	black-oak	lake	USA	46,16	-89,32
Lake Bourget	bourget	lake	France	45,76	5,86
Lake Burley Griffin	burley-griffin	reservoir	Australia	-35,30	149,07

Crystal Lake	crystal-lake	lake	USA	46,00	-89,61
Crystal Bog	crystal-bog	lake	USA	46,01	-89,61
Delavan Lake	delavan	lake	USA	42,61	-88,60
Dickie Lake	dickie	lake	Canada	45,15	-79,09
Eagle Lake	eagle	lake	Canada	44,68	-76,70
Ekoln basin of Mälaren	ekoln	lake	Sweden	59,75	17,62
Lake Erken	erken	lake	Sweden	59,84	18,63
Esthwaite Water	esthwaite-water	lake	United Kingdom	54,37	-2,99
Falling Creek Reservoir	falling-creek	reservoir	USA	37,31	-79,84
Lake Feeagh	feeagh	lake	Ireland	53,90	-9,50
Fish Lake	fish	lake	USA	43,29	-89,65
Lake Geneva	geneva	lake	France/Switzerland	46,45	6,59
Great Pond	great	lake	USA	44,53	-69,89
Green Lake	green	lake	USA	43,81	-89,00
Harp Lake	harp	lake	Canada	45,38	-79,13

Kilpisjärvi	kilpisjarvi	lake	Finland	69,03	20,77
Lake Kinneret	kinneret	lake	Israel	32,49	35,35
Lake Kivu	kivu	lake	Rwanda/DR Congo	-1,73	29,24
Klicava Reservoir	klicava	reservoir	Czechia	50,07	13,93
Lake Kuivajarvi	kuivajarvi	lake	Finland	60,47	23,51
Lake Langtjern	langtjern	lake	Norway	60,37	9,73
Laramie Lake	laramie	lake	USA	40,62	-105,84
Lower Lake Zurich	lower-zurich	lake	Switzerland	47,28	8,58
Lake Mendota	mendota	lake	USA	43,10	-89,41
Lake Monona	monona	lake	USA	43,06	-89,36
Mozhaysk reservoir	mozhaysk	reservoir	Russia	55,59	35,82
Mt Bold	mt-bold	reservoir	Australia	-35,12	138,71
Lake Müggelsee	mueggelsee	lake	Germany	52,43	13,65
Lake Neuchâtel	neuchatel	lake	Switzerland	46.54	6.52

Ngoring	ngoring	lake	China	34,90	97,70
Lake Nohipalo Mustjärv	nohipalo-mustjaerv	lake	Estonia	57,93	27,34
Lake Nohipalo Valgejärv	nohipalo-valgejaerv	lake	Estonia	57,94	27,35
Okauchee Lake	okauchee	lake	USA	43,13	-88,43
Lake Pääjärvi	paajarvi	lake	Finland	61,07	25,13
Rappbode Reservoir	rappbode	reservoir	Germany	51,74	10,89
Rimov Reservoir	rimov	reservoir	Czechia	48,85	14,49
Lake Rotorua	rotorua	lake	New Zealand	-38.08	176.28
Lake Sammamish	sammamish	lake	USA	47,59	-122,10
Sau Reservoir	sau	reservoir	Spain	41,97	2,40
Sparkling Lake	sparkling	lake	USA	46,01	-89,70
Lake Stechlin	stechlin	lake	Germany	53,17	13,03
Lake Sunapee	sunapee	lake	USA	43,23	-72,50
Lake Tahoe	tahoe	reservoir	USA	39,09	-120,03
Lake Tarawera	tarawera	lake	New Zealand	-38,21	176,43

Lake Taupo	taupo	lake	New Zealand	-38,80	175,89
Toolik Lake	toolik	lake	USA	68,63	-149,60
Trout Lake	trout-lake	lake	USA	46,03	-89,67
Trout Bog	trout-bog	lake	USA	46,04	-89,69
Two Sisters Lake	two-sisters	lake	USA	45,77	-89,53
Lake Vendyurskoe	vendyurskoe	lake	Russia	62,10	33,10
lake Võrtsjärv	vortsjaerv	lake	Estonia	58,31	26,01
Lake Waahi	waahi	lake	New Zealand	37,33	175,07
Lake Washington	washington	lake	USA	47,64	-122,27
Windermere	windermere	lake	United Kingdom	54,31	-2,95
Lake Wingra	wingra	lake	USA	43,05	-89,43
Zlutice Reservoir	zlutice	reservoir	Czechia	50,09	13,11

8 Biomes

8.1 Experiments

Table 17 provides an overview of all experiments to be run in the biomes sector in ISIMIP2a.

Table 17: Experiment summary for Biomes models. For an explanation of the varsoc, pressoc and nat settings see **Table** 7 (and **Table** 10 if your model is also a water model). If varsoc is not possible, please submit the pressoc run.

Climate Data	Scenario	Human impacts (see Table 7)	Other settings (sens-scenario)	# runs
WATCH-WFDEI	hist	varsoc	historical CO2 (co2)	1
GSWP3-W5E5	hist	varsoc	historical CO2 (co2)	1
GSWP3-EWEMBI	hist	varsoc	historical CO2 (co2)	1
GSWP3	hist	varsoc	historical CO2 (co2)	1
PGMFD v.2 (Princeton)	hist	varsoc (see **Table** 10)	historical CO2 (co2)	1
WATCH (WFD)	hist	varsoc	historical CO2 (co2)	1
Additional sector-specific run: PGMFD v.2 (Princeton)	hist	nat	historical CO2 (co2)	1
Additional sector-specific run: PGMFD v.2 (Princeton)	hist	varsoc	fix at pre-industrial levels (pico2) = 280ppm	1
				8

Please note: these tables do not include all necessary information and should be used as a reference only once the sector-specific and cross-sectoral protocol has been read in full.

8.2 Output data

IMPORTANT: The output variables reported for the biomes sector are also appropriate for use in the permafrost sector described in Section 11.

Table 18: Variables to be reported by biomes models.

Variable (long name)	**Variable name**	**Unit (NetCDF format)**		**Resolution**	**Comment**
			Essential outputs		
Pools					
Carbon Mass in Vegetation biomass	**cveg-<pft>**	kg m-2	per pft and gridcell total	year	Gridcell total is essential. Per PFT information is desirable.
*Carbon Mass in aboveground vegetation biomass	**cvegag-<pft>**	kg m-2	per pft and gridcell total	year	Gridcell total cvegag is essential. Per PFT information is desirable.
*Carbon Mass in belowground vegetation biomass	**cvegbg-<pft>**	kg m-2	per pft and gridcell total	year	Gridcell total cvegbg is essential. Per PFT information is desirable.
Carbon Mass in Litter Pool	**clitter-<pft>**	kg m-2	per pft and gridcell total	year	Info for each individual pool.
Carbon Mass in Soil Pool	**csoil-<pft>**	kg m-2	per pft and gridcell total	year	Info for each individual pool.
Fluxes					
Carbon Mass Flux out of atmosphere due to Gross Primary Production on Land	**gpp-<pft>**	kg m-2 s-1	per pft and gridcell total	day (mon)	
Carbon Mass Flux into atmosphere due to Autotrophic (Plant) Respiration on Land	**ra-<pft>**	kg m-2 s-1	per pft and gridcell total	day (mon)	
Carbon Mass Flux out of atmosphere due to Net	**npp-<pft>**	kg m-2 s-1	per pft and gridcell total	day (mon)	

Primary Production on Land					
Net Primary Production on Land allocated to leaf biomass	**npplandleaf-<pft>**	kg m-2 s-1	per pft and per gridcell	day (mon)	
Net Primary Production on Land allocated to fine root biomass	**npplandroot-<pft>**	kg m-2 s-1	per pft and per gridcell	day (mon)	
Net Primary Production on Land allocated to above ground wood biomass	**nppabovegroundwood-<pft>**	kg m-2 s-1	per pft and per gridcell	day (mon)	
Net Primary Production on Land allocated to below ground wood biomass	**nppbelowgroundwood-<pft>**	kg m-2 s-1	per pft and per gridcell	day (mon)	
Carbon Mass Flux into atmosphere due to Heterotrophic Respiration on Land	**rh-<pft>**	kg m-2 s-1	per pft and gridcell total	day (mon)	
Carbon Mass Flux into atmosphere due to total Carbon emissions from Fire	**fireint-<pft>**	kg m-2 s-1	per pft and gridcell total	day (mon)	
Fraction of cell burnt by fire	**firefrac-<pft>**	Fractional	Per pft and gridcell total		Burnt area fraction: single value for each scenario corresponding to year 2100.
Carbon Mass Flux out of Atmosphere due to Net biome Production on Land (NBP) (please specify if NBP≠NPP+HR+Fires in your model)	**ecoatmfluxc-<pft>**	kg m-2 s-1	per pft and gridcell total	day (mon)	This is the net mass flux of carbon between land and atmosphere. calculated as photosynthesis MINUS the sum of plant and soil respiration, carbonfluxes from fire, harvest, grazing and land use change. Positive flux is into the land.
Root autotrophic respiration	**rr-<pft>**	kg m-2 s-1	per pft and gridcell total	day (mon)	
Structure					

Fraction of absorbed photosynthetically active radiation	**fapar-<pft>**	%	per pft and gridcell average	day (mon)	
Leaf Area Index	**lai-<pft>**	1	per pft and gridcell average	day (mon)	
Plant Functional Type Grid Fraction	**pft-<pft>**	%	per gridcell	year (or once if static)	The categories may differ from model to model, depending on their PFT definitions. This may include natural PFTs, anthropogenic PFTs, bare soil, lakes, urban areas, etc. Sum of all should equal the fraction of the grid-cell that is land.
Hydrological variables					
Total Evapo-Transpiration	**evap-<pft>**	kg m-2 s-1	per pft and gridcell	day (mon)	
Evaporation from Canopy (interception)	**intercep-<pft>**	kg m-2 s-1	per pft and gridcell	day (mon)	The canopy evaporation+sublimation (if present in model).
Water Evaporation from Soil	**esoil**	kg m-2 s-1	per gridcell	day (mon)	Includes sublimation.
Transpiration	**trans-<pft>**	kg m-2 s-1	per pft and gridcell	day (mon)	
Total Runoff	**qtot**	kg m-2 s-1	per gridcell	day (mon)	The total runoff (including "drainage" through the base of the soil model) leaving the land portion of the grid cell.
Soil Moisture	**soilmoist**	kg m-2	per gridcell	day (mon)	If possible, please provide soil moisture for all depth layers (i.e. 3D-field), and indicate depth in m. Otherwise, provide soil moisture of entire column.
Surface Runoff	**qs**	kg m-2 s-1	per gridcell	day (mon)	The total surface runoff leaving the land portion of the grid cell.
Frozen soil moisture for each layer	**soilmoistfroz**	kg m-2	per gridcell	mon	Please provide soil moisture for all depth levels and indicate depth in m.

					Please provide for purposes of permafrost sector.
Snow depth	**snd**	m	per gridcell	mon	Grid cell mean depth of snowpack. Please provide for purposes of permafrost sector.
Snow water equivalent	**swe**	kg m-2	per gridcell	mon	Snow depth x snow density.
Annual maximum thaw depth	**thawdepth**	m	per gridcell	year	Calculated from daily thaw depths.
Optional outputs					
Carbon Mass in Leaves	**cleaf-<pft>**	kg m-2	per pft and gridcell	year	
Carbon Mass in Wood	**cwood -<pft>**	kg m-2	per pft and gridcell	year	Including sapwood and hardwood.
Carbon Mass in Roots	**croot-<pft>**	kg m-2	per pft and gridcell	year	Including fine and coarse roots.
Others					
Temperature of Soil	**tsl**	K	per gridcell	day (mon)	Temperature of each soil layer. Reported as "missing" for grid cells occupied entirely by "sea". **THIS IS THE MOST IMPORTANT VARIABLE FOR THE PERMAFROST SECTOR.** Also need depths in meters. Daily would be great, but otherwise monthly would work.
Burnt Area Fraction	**burntarea**	%	per gridcell	day (mon)	Fraction of entire grid cell that is covered by burnt vegetation.

Note: If you cannot provide the data at the temporal or spatial resolution specified, please provide it the highest possible resolution of your model. Please contact the coordination team (Info@isimip.org) to for any further clarification, or to discuss the equivalent variable in your model.

9 Forest Models (Regional, Forest stand-level models)

PROFOUND Contributors: Christopher Reyer, Susana Barreiro, Harald Bugmann, Alessio Collalti, Klara Dolos, Louis Francois, Venceslas Goudiaby, Carlos Gracia, Thomas Hickler, Mathieu Jonard, Chris Kollas, Koen Kramer, Petra Lasch-Born, Denis Loustau, Annikki Mäkelä, Simon Martel, Daniel Nadal I Sala, Delphine Picart, David Price, Santiago Sabaté, Monia Santini, Rupert Seidl, Felicitas Suckow, Margarida Tomé, Giorgio Vacchiano

9.1 Introduction to multi-model simulations in ISIMIP2a and PROFOUND

This is an overview document to support multi-model simulations of forest stand models for both model evaluation with observed data. A number of sites has been selected in the COST Action PROFOUND (http://cost-profound.eu/site/) for which a) a wide range of forest models can be rather easily initialized, b) observational data is available for model evaluation and b) additional local driving datasets are available such as N-deposition or locally observed climate (**Table** 20). To get access to this PROFOUND Database, please contact reyer@pik-potsdam.de. A few important particularities for the forest simulations are listed below.

1) **Management:** The modeling experiments mostly encompass managed forests. The standard management ("varsoc") during the historical period is the observed management as defined by the data available for each site (please only use the reduction in stem numbers to design the management).

2) **Calibration:** Some of the models may require some kind of calibration or model development before they can contribute to ISIMIP. Such alterations of the model can influence the results of a model comparison and "model calibration" is understood differently by different modelers. All alterations to the model in the framework of this exercise should be reported in the model experiment documentation provided together with the upload of the simulations. Whenever the model calibration or development is driven by an improvement of the model after a comparison to data that were originally made available in ISIMIP for model evaluation, a part of those data should be kept aside for model evaluation and not used for calibration.
    a. Model development needed to run a model at specific sites is welcomed and needs to be transparent/ properly documented (e.g. adjustment of phenology model to include chilling effects). This is also applicable for more general calibration (i.e. fixing parameters once but not changing afterwards) for example to include a new tree species in a model.

    b. Manual or automatic site-specific "tuning" of species-specific and process-specific parameters should be avoided. The same "model" (i.e. also with the same parameter values) should be used in all simulations. If needed, any tuning needs to be documented in a transparent way and should be backed up by existing data (e.g. from TRY-database). If your model contains genetic processes where the change in parameters is part of the model processes, this is naturally part of “your model approach” and should be clearly spelled out as part of the documentation of your model. In this specific case, please contact the sectoral coordinators to discuss if it makes sense to include a “genetic adaptation” and a “parameter-fixed, control” run.

3) **Reporting Period**: Each phase of ISIMIP has its own reporting period (e.g. 1971-2000 for ISIMIP2A) but since we have sometimes data for model initialization and validation going back even further in time, you should always start your reporting period for the first time step for which stand data is available (e.g. 1948 for the Peitz stand) and run your model until the last point in time where climate data is available. Similarly, if the model runs only start later than, e.g. 1971, the reporting period is shorter. If the data for model initialization is only available very late (e.g. KROOF starts in 1998 only, you do not need to run your model for those climatic datasets which end early (e.g. Watch ending in 2001 already).

9.2 Experiments

Table 19 provides an overview of all experiments to be run with regional forest models in ISIMIP. This table is for your reference only; please read chapters 1-5 of the general ISIMIP protocol and this whole section carefully before beginning with the experiments. In case of any questions please contact info@isimip.org. Please note that aside from harmonized climate, stand, management and soil input, the default settings of your model should be used. Also note that for output data files the **file name is all lower case**!

Table 19: Experiment summary for regional forest models. Each experiment is to be carried out for each site named in **Table** 20**.** For management scenarios see **Table** 21 - 23.

	Climate Data	Scenario	Management	Other settings (sens-scenario)	# runs
Historical runs without disturbances (Experiment 1a)	Observations from local meteorological station or likewise	hist	1. Observed management (varsoc) 2. Natural reference run (nosoc)	historical CO_2 without disturbances (co2), EMEP-N-deposition	2
	WATCH-WFDEI	hist	1. Observed management (varsoc) 2. Natural reference run (nosoc)	historical CO_2 without disturbances (co2), EMEP-N-deposition	2
	GSWP3-W5E5	hist	1. Observed management (varsoc) 2. Natural reference run (nosoc)	historical CO_2 without disturbances (co2), EMEP-N-deposition	2
	GSWP3-EWEMBI	hist	1. Observed management (varsoc)	historical CO_2 without disturbances	2

			2. Natural reference run (nosoc)	(co2), EMEP-N-deposition	
	GSWP3	hist	1. Observed management (varsoc) 2. Natural reference run (nosoc)	historical CO_2 without disturbances (co2), EMEP-N-deposition	2
	PGMFD v.2 (Princeton)	hist	1. Observed management (varsoc) 2. Natural reference run (nosoc)	historical CO_2 without disturbances (co2), EMEP-N-deposition	2
	WATCH (WFD)	hist	1. Observed management (varsoc) 2. Natural reference run (nosoc)	historical CO_2 without disturbances (co2), EMEP-N-deposition	2

Please note: these tables do not include all necessary information and should be used as a reference only once the sector-specific and cross-sectoral protocol has been read in full.

9.3 Sector-specific input

The input and evaluation data is provided through the PROFOUND database including an R-package to explore the database. Until the database is officially released, please get in touch with Christopher Reyer (reyer@pik-potsdam.de) to access it.

Table 20: Overview of the forest stands to be simulated in ISIMIP/PROFOUND.

Site name (for filenames)	Lat	Lon	Country	Forest type	Species	Thinning during historical time period	Comments
hyytiala	61.8475	24.295	FI	Even-aged conifer	pisy, piab with some deciduous mix	below	Note that an experimental plot of pine contains a lot of data while footprint of flux tower is larger. Please note that the deciduous admixtures only appear in the data at a later stage and hence do not need to be simulated. Only simulate pine and spruce (no hard-woods) and regenerate as pure pine stand
peitz	51.9166	14.35	DE	Even-aged conifer	pisy	below	Managed using a weak thinning from below.
solling-beech	51.77	9.57	DE	Even-aged deciduous	fasy	above	

solling-spruce	51.77	9.57	DE	Even-aged conifer	piab	below	
soro	55.485844	11.644616	DK	Even-aged deciduous	fasy	above	
kroof	48.25	11.4	DE	Mixed deciduous and conifers	fasy, piab, acpl, lade, pisy, quro	below	Unmanaged/ thinning from below in past 20 years for all species.
le-bray	44.71711	-0.7693	FR	Even-aged conifer	pipi	below	
collelongo	41.8494	13.5881	IT	Even-aged deciduous	fasy	above	
bily-kriz	49.3	18.32	CZ	Even-aged conifer	piab	below	

Table 21: Planting information for the sites included in the simulation experiments. DBH is defined as diameter at breast height of 1.30m. The numbers in brackets indicate plausible ranges.

Name	Density (ha^{-1})	Age (years)	Height (m)	DBH (cm)	Age when DBH is reached (years)	Remarks
bily-kriz	4500	4	0.5	na	9	Historical planting density was 5000/ha but current practices are 4500/ha only.
collelongo	10000	4	1.3	0.1	4	Only a rough approximation, usually natural regeneration is the regeneration method.
hyytala	2250 (2000-2500)	2	0.25 (0.2-0.3)	na	6 (5-7)	Regenerate as pure pine stand
kroof (beech)	6000 (5000-7000)	2	0.6 (0.5-0.7)	na	5	The planting density is for single-species stands, hence when regenerating the 2-species-stand KROOF, the planting density of each species should be halved
kroof (spruce)	2250 (2000-2500)	2	0.35 (0.3-0.4)	na	7	See above
le-bray	1250 (1000-14000)	1	0.2 (0.1-0.25)	na	3 (2-5)	These are the current practices (De Lary, October, 2015) and should be used for future regeneration. Historically, the site was seeded with 3000-5000 seedlings per ha and then cleared once or twice to reach a density of 1250/ha at 7-year old when seedlings reach the size for DBH recruitment. Modelers could mimic this by "planting" trees with DBH of 7.5cm and 6m height in 1978 with a density of 1250 trees/ha
peitz	9000 (8000-10000)	2	0.175 (0.1-0.25)	na	5	The “age when DBH is reached = 5” is an estimate
solling-beech	6000 (5000-7000)	2	0.6 (0.5-0.7)	na	5	The actual stand was established in 1847 from natural regeneration. Until begin of measurements in 1966, the stand was regularly thinned. All figures in table are estimates. Natural regeneration is the recommended regeneration method of stand establishment; stem count in 2014: 130
solling-spruce	2250 (2000-2500)	2	0.35 (0.3-0.4)	na	7	The actual stand was planted in 1891 on a former meadow. Until begin of measurements in 1966, the

						stand was regularly thinned. All figures in table are estimates; stem count in 2014: 290
soro	6000	4	0.82	na	6	Planted in 1921, stem count in 288 ha-1 in 2010 (Wu, et al., 2013)

9.4 Output data

Table 22: Variables to be reported by forest models. Abbreviations are provided in **Table** 23. Variables should be reported as documented in Section 5.

Variable (long name)	Variable name	Units (NetCDF format)		Resolution	DBH class resolution	Comment
Essential outputs						
Mean DBH	**dbh-<species/total>**	cm	per species and stand total	annual	None	
Mean DBH of 100 highest trees	**dbhdomhei**	cm	stand total	annual	None	100 highest trees per hectare.
Stand Height	**hei-<species/total>**	m	per species and stand total	annual	None	For models including natural regeneration this variable may not make sense, please report domhei.
Dominant Height	**domhei**	m	stand total	annual	None	Mean height of the 100 highest trees per hectare.
Stand Density	**density-<species/total>**	ha-1	per species and stand total	annual	None	
Basal Area	**ba-<species/total>**	m2 ha-1	per species and stand total	annual	None	
Volume of Dead Trees	**mort-<species/total>**	m3 ha-1	per species and stand total	annual	None	

Harvest by dbh-class	**harv-<species/total>**	m3 ha-1	per species and stand total and dbh-class	annual	Either dbh classes or total	See Section 5.1.5
Remaining stem number after disturbance and management by dbh class	**stemno-<species/total**	ha-1	per species and stand total and dbh-class	annual	Either dbh classes or total	See Section 5.1.5
Stand Volume	**vol-<species/total>**	m3 ha-1	per species and stand total	annual	None	
Carbon Mass in Vegetation biomass	**cveg-<species/total>**	kg m-2	per species and stand total	annual	None	As kg carbon*m^{-2}
*Carbon Mass in aboveground vegetation biomass	**cvegag-<species/total>**	kg m-2	per species and stand total	annual	None	As kg carbon*m^{-2}
*Carbon Mass in belowground vegetation biomass	**cvegbg-<species/total>**	kg m-2	per species and stand total	annual	None	As kg carbon*m^{-2}
Carbon Mass in Litter Pool	**clitter-<species/total>**	kg m-2	per species and stand total	annual	None	As kg carbon*m^{-2}, Info for each individual pool.
Carbon Mass in Soil Pool	**csoil-<species/total>**	kg m-2	per species and stand total	annual	None	As kg carbon*m^{-2}, Info for each individual soil layer
Tree age by dbh class	**age-<species/total**	yr	per species and stand total and dbh-class	annual	Either dbh classes or total	See Section 5.1.5
Gross Primary Production	**gpp-<species/total>**	kg m-2 s-1	per species and stand total	daily	None	As kg carbon*m^{-2}*s^{-1}
Net Primary Production	**npp-<species/total>**	kg m-2 s-1	per species and stand total	daily	None	As kg carbon*m^{-2}*s^{-1}
Autotrophic (Plant) Respiration	**ra-<species/total>**	kg m-2 s-1	per species and stand total	daily	None	As kg carbon*m^{-2}*s^{-1}
Heterotrophic Respiration	**rh-< total>**	kg m-2 s-1	stand total	daily	None	As kg carbon*m^{-2}*s^{-1}
Net Ecosystem Exchange	**nee-<total>**	kg m-2 s-1	per stand	daily	None	As kg carbon*m^{-2}*s^{-1}
Mean Annual Increment	**mai-<species/total>**	m3 ha-1	per species and stand	annual	None	

			total			
Fraction of absorbed photosynthetically active radiation	**fapar-<species/total>**	%	per species and stand total	daily	None	Value between 0 and 100.
Leaf Area Index	**lai-<species/total>**	m2 m-2	per species and stand total	monthly	None	
Species composition	**species-<species>**	%	per ha	annual (or once if static)	None	As % of basal area; the categories may differ from model to model, depending on their species and stand definitions.
Total Evapotranspiration	**evap**	kg m-2 s-1	stand total	daily	None	sum of transpiration, evaporation, interception and sublimation. (=intercept + esoil + trans)
Evaporation from Canopy (interception)	**intercep-<species/total>**	kg m-2 s-1	per species and stand total	daily	None	the canopy evaporation + sublimation (if present in model).
Water Evaporation from Soil	**esoil**	kg m-2 s-1	per stand	daily	None	includes sublimation.
Transpiration	**trans-<species/total>**	kg m-2 s-1	per species and stand total	daily	None	
Soil Moisture	**soilmoist**	kg m-2	per stand	daily	None	If possible, please provide soil moisture for all depth layers (i.e. 3D-field), and indicate depth in m. Otherwise, provide soil moisture of entire column.

Optional outputs						
Removed stem numbers by size class by natural mortality	**mortstemno-<species/total>**	ha-1	per species and stand total and dbh-class	annual	Either dbh classes or total	As trees per hectare. See Section 5.1.5
Removed stem numbers by size class by management	**harvstemno-<species/total>**	ha-1	per species and stand total and dbh-class	annual	Either dbh classes or total	As trees per hectare. See Section 5.1.5
Volume of disturbance damage	**dist-<dist-name>**	m3 ha-1	per species and stand total	annual	None	
Nitrogen of annual Litter	**nlit-<species/total>**	g m-2 a-1	per species and stand total	annual	None	As g Nitrogen m^{-2} a^{-1}
Nitrogen in Soil	**nsoil-<total>**	g m-2 a-1	stand total	annual	None	As g Nitrogen m^{-2} a^{-1}
Net Primary Production allocated to leaf biomass	**nppleaf-<species>**	kg m-2 s-1	per species and stand total	daily	None	As kg carbon*m^{-2}*s^{-1}
Net Primary Production allocated to fine root biomass	**npproot-<species>**	kg m-2 s-1	per species and stand total	daily	None	As kg carbon*m^{-2}*s^{-1}
Net Primary Production allocated to above ground wood biomass	**nppagwood-<species>**	kg m-2 s-1	per species and stand total	daily	None	As kg carbon*m^{-2}*s^{-1}
Net Primary Production allocated to below ground wood biomass	**nppbgwood-<species>**	kg m-2 s-1	per species and stand total	daily	None	As kg carbon*m^{-2}*s^{-1}
Root autotrophic respiration	**rr-<species/total>**	kg m-2 s-1	per species and stand total	daily	None	As kg carbon*m^{-2}*s^{-1}
Carbon Mass in Leaves	**cleaf-<species>**	kg m-2	per species and stand total	annual	None	
Carbon Mass in Wood	**cwood-<species>**	kg m-2	per species and stand total	annual	None	including sapwood and hardwood
Carbon Mass in Roots	**croot-<species>**	kg m-2	per species and stand total	annual	None	including fine and coarse roots
Temperature of Soil	**tsl**	K	per stand	daily	None	Temperature of each soil layer

Note: If you cannot provide the data at the temporal or spatial resolution specified, please provide it the highest possible resolution of your model. Please contact the coordination team (info@isimip.org) to for any further clarification, or to discuss the equivalent variable in your model.

Table 23: Codes for species, disturbance names and dbh classes as used in protocol (species, dist_name, dbhclass).

Long name	Short name
Fagus sylvatica	fasy
Quercus robur	quro
Quercus petraea	qupe
Pinus sylvestris	pisy
Picea abies	piab
Pinus pinaster	pipi
Larix decidua	lade
Acer platanoides	acpl
Eucalyptus globulus	eugl
Betula pendula	bepe
Betula pubescens	bepu
Robinia pseudoacacia	rops
Fraxinus excelsior	frex
Populus nigra	poni
Sorbus aucuparia	soau
C3 grass	c3gr
hard woods	hawo
fire	fi
wind	wi
Insects	ins
Drought	dr
Grazing	graz
Diseases	dis
DBH_class_<X>-<X+5>*	dbh_c<X>
DBH_class_>140*	dbh_c140

*the boundaries of the dbh classes should interpreted as follows: dbh_class_0-5 = 0 to<5 cm; dbh_class_5-10 =5 to<10 cm, etc…. the dbh class dbh_c140 includes all trees of 140cm dbh and larger.

10 Agriculture (crop modelling)

This section lays out the global output protocol for the agricultural sector's contribution to ISIMIP. For further details, please contact AgMIP (ag-grid@agmip.org) and ISIMIP (info@isimip.org).
Note that the variable names are chosen to comply with AgMIP conventions or are harmonized with the conventions used in the ISIMIP water sector (for irrigation water). They are given in lower-case letters only in order to prevent the use of mixed-case names in the file names (see Section 5.1.1). **Table** 6 provides an overview of all experiments to be run in the agriculture (crop modelling) sector in ISIMIP2a.

10.1 Experiments

Table 24: Experiment summary for crop models.

	Climate Data	Scenario	Management settings	Land use (LU)	Other settings (sens-scenario)	Irrigation	# Runs
Historical runs	WFD+WFDEI	hist	default (present day) (default) fully harmonized (fullharm) harmonized season, no N constraints (harmnon)	pure crop run (no LU specifier)	historical CO2 (no co2 specifier)	firr noirr	6
	GSWP3-W5E5	hist	default (present day) (default)	pure crop run (no LU specifier)	historical CO2 (no co2 specifier)	firr noirr	2
	GSWP3-EWEMBI	hist	default (present day) (default)	pure crop run (no LU specifier)	historical CO2 (no co2 specifier)	firr noirr	2
	GSWP3	hist	default (present day) (default)	pure crop run (no LU specifier)	historical CO2 (no co2 specifier)	firr noirr	2
	PGMFD v.2 (Princeton)	hist	default (present day) (default)	pure crop run (no LU specifier)	historical CO2 (no co2 specifier)	firr noirr	2
	WATCH (WFD)	hist	default (present day) (default)	pure crop run (no LU specifier)	historical CO2 (no co2 specifier)	firr noirr	2
							12 (per crop)

10.2 Sector-specific input

Some GGCMs require inputs on planting dates, crop variety parameters, fertilizer use and possibly other management specifics. While the agreement for the fast-track was to use each model's setting that best represents current management patterns, we'll have specific inputs on planting dates and maturity dates (to allow for spatially-explicit variety parameterization) as well as fertilizer use (N, P, K). Some experiments will be run with harmonized input data (validation and attribution studies), and some with default model settings.

Table 25: Crop-model-specific input data.

Variable	Source*	Units	Notes
Planting **dates**	(Sacks, Deryng, Foley, & Ramankutty, 2010), (Portmann, Siebert, & Döll, 2010), supplemented with a rule-based approach as implemented in LPJmL in regions without observational data (see Elliott et al. 2015).	Julian days (Jan 1st= 1,…)	Planting dates for primary seasons per crop and grod cell.
Approximate maturity	(Sacks, Deryng, Foley, & Ramankutty, 2010), (Portmann, Siebert, & Döll, 2010) , supplemented with a rule-based approach as implemented in LPJmL in regions without observational data (see Elliott et al. 2015).	days from planting	Growing season length in days.
Fertilizers and manure	(Mueller, et al., 2012), (Potter, Ramankutty,	kg ha-1 yr-1	Average nitrogen, phosphorus, and potassium application rates in each grid cell, with organic and inorganic amendments aggregated

	Bennett, & Donner, 2011), (Liu, et al., 2010), (Foley, et al., 2011)		and converted to an “effective inorganic application rate”.
Historical [CO2]	Mauna Loa/RCP historical	ppm	Annual [CO2] values from 1900-2013.

10.3 Output data and definitions

Crop Priority and naming list:

1. Wheat[6], maize, soy, rice [whe, mai, soy, ric]
2. All others: Sugarcane, sorghum, millet, rapeseed, sugar beet, barley, rye, oat [sug, sor, mil, rap, sgb, bar, rye, and oat] + managed grass [mgr], field peas [pea], cassava [cas], sunflower [sun], groundnuts [nut], bean [ben], potato [pot], bioenergy crops such as poplar [pop], eucalyptus [euc], miscanthus [mis] … **Note:** planting and maturity dates for bioenergy crops shall only be reported if meaningful (i.e. not for perennials).

Reporting per growing seasons:
To resolve potential double harvests within one year, crop yields should be reported per growing season and not per calendar year. Thus, in the NetCDF output files, do not use a time dimension but instead a unitless coordinate variable with integer values; more information on how to construct these files in **Section** 5.1.6 and in our ISIMIP website (https://www.isimip.org/protocol/preparing-simulation-files/). Cumulative growing season variables such as, e.g., actual evapotranspiration or precipitation are to be accumulated over the growing season. The first season in the file (growing-season=0) is then the first complete growing season of the time period provided by the input data without any assumed spin-up data, which equates to the growing season with the first planting after this date. To ensure that data can be matched to individual years in post-processing, it is essential to also provide the actual planting dates (as day of the year), actual planting years (year),

[6] There will be no distinction between winter and spring wheat.

anthesis dates (as day of the year), year of anthesis (year), maturity dates (day of the year), and year of maturity (year). This procedure is identical to the GGCMI convention (Elliott, et al., 2015).

Table 26: Output variables for crop models.

Variable (long name)	Variable name	Unit	Resolution	Comments
Key model output				
Crop yields	**yield-<crop>-<irrigation setting>**	dry matter (t ha-1 per growing season)	per growing season (0.5°x0.5°)	Crop-specific Yield may be identical to above-ground biomass (biom) if the entire plant is harvested, e.g. for bioenergy production.
Irrigation water withdrawal (assuming unlimited water supply)	**pirrww-<crop>-<irrigation setting>**	mm per growing season	per growing season (0.5°x0.5°)	*Irrigation water withdrawn in case of optimal irrigation (in addition to rainfall), assuming no losses in conveyance and application.*
Key diagnostic variables				
Actual evapotranspiration	**aet-<crop>-<irrigation setting>**	mm per growing season	per growing season (0.5°x0.5°)	portion of all water (including rain) that is evapo-transpired, the water amount should be accumulated over the entire growing period (**not** the calendar year)
Nitrogen application rate	**initr-<crop>-<irrigation setting>**	kg ha-1 per growing season	per growing season (0.5°x0.5°)	Total nitrogen application rate. If organic and inorganic amendments are applied, rate should be reported as effective inorganic nitrogen input (ignoring residues).
Actual planting dates	**plantday-<crop>-<irrigation setting>**	Day of year	per growing season (0.5°x0.5°)	

Anthesis dates	**anthday-<crop>-<irrigation setting>**	Days from planting date	per growing season (0.5°x0.5°)	
Maturity dates	**matyday-<crop>-<irrigation setting>**	Days from planting date	per growing season (0.5°x0.5°)	
Additional output variables (optional)				
Above ground biomass (dry matter)	**biom-<crop>-<irrigation setting>**	t ha-1 per growing season	per growing season (0.5°x0.5°)	The whole plant biomass above ground
Soil carbon emissions	**sco2-<crop>-<irrigation setting>**	kg C ha-1	per growing season (0.5°x0.5°)	Ideally should be modeled with realistic land-use history and initial carbon pools. Subject to extra study.
Nitrous oxide emissions	**sn2o-<crop>-<irrigation setting>**	kg N2O-N ha-1	per growing season (0.5°x0.5°)	Ideally should be modeled with realistic land-use history and initial carbon pools. Subject to extra study.
Total N uptake (total growing season sum)	**tnup-<crop>-<irrigation setting>**	kg ha -1 yr -1	monthly (0.5°x0.5°)	Nitrogen balance: uptake
Total N inputs (total growing season sum)	**tnin-<crop>-<irrigation setting>**	kg ha -1 yr -1	monthly (0.5°x0.5°)	Nitrogen balance: inputs
Total N losses (total growing season sum)	**tnloss-<crop>-<irrigation setting>**	kg ha -1 yr -1	monthly (0.5°x0.5°)	Nitrogen balance: losses
Growing season temperature sum	**sumt_<crop>**	deg c-days yr-1	per growing season (0.5°x0.5°)	Sum of daily mean temperature over growing season
Growing season radiation	**gsrsds_<crop>**	w m-2 yr-1	per growing season (0.5°x0.5°)	Average growing season shortwave solar radiation
Growing season precipitation	**gsprcp_<crop>**	mm ha-1 yr-1	per growing season	Total growing season precipitation per crop

			(0.5°x0.5°)	

Note: The reporting periods for some output variables were changed from "yearly" to "per growing season" in April 2019. Please be aware that model outputs submitted before this date, may still contain yearly data. Some models (e.g., LPJmL) report outputs for additional crops ("cas" cassava, "mil" millet, "nut" groundnut, "pea" peas, "rap" rapeseed, "sgb" sugar beet, "sug" sugarcane, "sun" sunflower, "mgr" managed grass). The model EPIC-BOKU provides outputs for alternative PET equations (Hargreaves (hg), Penman-Monteith (pe), Priestley Taylor (pt), Baier-Robertson (br)).

10.4 Experiments

10.4.1 Historic runs and validation experiment

Specification of the historical run

Simulations for the historical period should be provided as pure crop runs (i.e. assuming the crop growing all over the world), based on the climate input described in Section 4. For each crop, there should be a full irrigation run (firr) and a no-irrigation run (noirr). Within ISIMIP2a we also ask for historical runs with three different degrees of harmonization as given in **Table** 27.

Table 27: Scenario settings for crop model simulations

Simulation	Comments
Default	Model should use their individual "best representation" of the historical period with regard to sowing dates, harvesting dates, fertilizer application rates and crop varieties.
Fully harmonized	Simulations based on prescribed "present day" fertilization rates (available for download) and fixed planting and harvesting dates (also available for download). Modelers should have planting as closely as possible to these dates, but it may be admissible to use these dates as indicators for planting windows (depending on model specifics).
Harmonized seasons with no N constraints	For models with an explicit description of the nitrogen cycle: harmnon simulations should be run with nitrogen stress turned off completely or (if that's not possible) with very high N application rates to make model results comparable between those GGCMs that have explicit N dynamics and those that do not. For models without the nitrogen cycle: harmnon and fullharm simulations are the same and do not need to be duplicated.

Each of these three variants should be combined with a no-irrigation and full irrigation assumption, resulting (for the models with an explicit representation of the nitrogen cycle) in 6 runs for the respective climate input data set (cf. **Table** 6).

Specification of PET equation

Running simulations with different PET equations implicate submitting different version of your model, with a consequent different model name; i.e. if you create a second set of simulations using Priestley Taylor PET equation, you shall use your <model-name> in the initial version, and <model-name>-pt in the second run. We recommend you these abbreviations: 'hg' for Hargreaves, 'pe' for Penman-Monteith, 'pt' for Priestley Taylor, and 'br' for Baier-Robertson.

Specification of the validation procedure

For the validation task the pure crop simulations should

1) be masked by the following LU patterns: "Dynamic MIRCA" (reconstruction of historical LU based on HYDE and MIRCA2000, see Section 4.3.

2) averaging and aggregation will be performed in the post-processing and depending on what data we compare to. It could include de-trending (to compare with possibly de-trended observations).

11 Permafrost

11.1 Experiments

The permafrost sector in ISIMIP2a will not require any additional runs. The runs developed for the biomes sector and the water sector can also be assessed by the permafrost sector (see Section 4.7 Scenario design for the scenario setup). Finnland (region 12) and the Lena catchment (region 11) are the two regions affected by permafrost. Therefore, any runs over these regions can be assessed for permafrost. Permafrost will require additional output data. Models which do not include a carbon cycle should still submit the requested hydrological variables as these can be used to assess permafrost extent and thaw.

11.2 Sector-specific input

None

11.3 Output data

Table 28 below is very similar to **Table** 18 in the Biomes sector, but with some hydrological variables added. **Soil temperature at each model level is the most important variable – if that is all you can deliver then please do so, it will be useful.**

Table 28: Variables to be reported for the permafrost sector.

Variable (long name)	Variable name	Unit (NetCDF format)		Resolution	Comment
			Essential outputs		
Temperature of Soil	**tsl**	K	per gridcell	Day (mon)	Temperature of each soil layer. Reported as "missing" for grid cells occupied entirely by "sea". **THIS IS THE MOST IMPORTANT VARIABLE.** Also need depths in meters. Daily would be great, but otherwise monthly would work.

Pools (as Biomes output Table)					
Carbon Mass in Vegetation	**cveg-<pft>**	kg m-2	per pft and gridcell total	year	Gridcell total cveg is essential. Per PFT information is desirable.
*Carbon Mass in aboveground vegetation biomass	**cvegag-<pft>**	kg m-2	per pft and gridcell total	year	Gridcell total cvegag is essential. Per PFT information is desirable.
*Carbon Mass in belowground vegetation biomass	**cvegbg-<pft>**	kg m-2	per pft and gridcell total	year	Gridcell total cvegbg is essential. Per PFT information is desirable.
Carbon Mass in Litter Pool	**clitter**	kg m-2	per gridcell	year	Total of all pools. Info for each individual pool is desirable.
Carbon Mass in Soil Pool	**csoil**	kg m-2	per gridcell	year	Total of all pools. Info for each individual pool is desirable.
Fluxes (as Biomes output Table)					
Carbon Mass Flux out of Atmosphere due to Gross Primary Production on Land	**gpp**	kg m-2 s-1	per gridcell	mon (day)	
Carbon Mass Flux into Atmosphere due to Autotrophic (Plant) Respiration on Land	**ra**	kg m-2 s-1	per gridcell	mon (day)	
Carbon Mass Flux out of Atmosphere due to Net Primary Production on Land	**npp**	kg m-2 s-1	per gridcell	mon (day)	
Carbon Mass Flux into Atmosphere due to Heterotrophic Respiration on Land	**rh**	kg m-2 s-1	per gridcell	mon (day)	
Carbon Mass Flux into Atmosphere due to	**fireint**	kg m-2 s-1	per gridcell	mon (day)	

CO2 Emission from Fire					
Fraction of cell burnt by fire	**firefrac**	Fractional	per gridcell		Burnt area fraction: single value for each scenario corresponding to year 2100
Carbon Mass Flux out of Atmosphere due to Net Biospheric Production on Land	**ecoatmfluxc**	kg m-2 s-1	per gridcell	mon (day)	This is the net mass flux of carbon between land and atmosphere calculated as photosynthesis MINUS the sum of plant and soil respiration, carbonfluxes from fire, harvest, grazing and land use change. Positive flux is into the land.
Structure (as Biomes output Table)					
Fraction of absorbed photosynthetically active radiation	**fapar-<pft>**	%	per pft and gridcell average	mon (day)	
Leaf Area Index	**lai-<pft>**	1	per pft and gridcell average	mon (day)	
Plant Functional Type Grid Fraction	**pft-<pft>**	%	per gridcell	year (or once if static)	The categories may differ from model to model, depending on their PFT definitions. This may include natural PFTs, anthropogenic PFTs, bare soil, lakes, urban areas, etc. Sum of all should equal the fraction of the grid-cell that is land.
Soil moisture for each layer	**soilmoist**	kg m-2	per gridcell	mon	Please provide soil moisture for all depth levels and indicate depth in m. (As for Water sector)
Frozen soil moisture for each layer	**soilmoistfroz**	kg m-2	per gridcell	mon	Please provide soil moisture for all depth levels and indicate depth in m. **This is a new variable.**
Snow depth	**snd**	m	per gridcell	Day	Grid cell mean depth of snowpack. **This is a new variable.**
Annual maximum thaw depth	**thawdepth**	m	per gridcell	year	Calculated from daily thaw depths
Snow water equivalent	**swe**	kg m-2	per gridcell	mon	Total water mass of the snowpack (liquid or frozen) averaged over grid cell (As for Water sector)
Runoff	**qtot**	kg m-2 s-1	per gridcell	mon (day)	Total runoff leaving the land portion of the grid cell (this

					is in both Biomes and Water Tables)
Optional outputs					
Burnt Area Fraction	**burntarea**	%	per gridcell	mon (day)	fraction of entire grid cell that is covered by burnt vegetation

Note: If you cannot provide the data at the temporal or spatial resolution specified, please provide the highest possible resolution of your model. Please contact the coordination team (Info@isimip.org) to for any further clarification, or to discuss the equivalent variable in your model.

12 Marine Ecosystems & Fisheries

12.1 Experiments

Table 29: Summary of historical runs for global and regional marine ecosystem & fisheries models. Priority should be given to the fishing scenario (time-varying fishing effort). Any other impacts not mentioned here should be held constant at year-2000 levels.

Climate data GCM	Scenario	Fishing effort	Ocean acidification	# runs
GFDL ESM2 (re-analysis)	hist	fishing (time-varying effort/mortality) no-fishing (zero effort/mortality)	time-varying pH	2

12.2 Sector-specific input

Climate-related forcing for historical simulations

Table 30: Historical and future forcing datasets for global and regional models.

Dataset description	Time period	Comments
GFDL reanalysis product CORE-forced MOM-SIS-TOPAZ	1959-2004	observation/re-analysis based time-series (1.0° x 1.0° degree) as described in (Stock, Dunne, & John, 2014) => includes observed climate variability

Table 31: Forcing variables provided as input for global and regional marine fisheries models.

Variable (long name)	Variable name	Unit (NetCDF format)	Resolution	Comments
Sea water X velocity	*uo*	m s-1	Monthly	surface
Sear water Y velocity	*vo*	m s-1	Monthly	surface
Sea water temperature	*to*	K	Monthly	surface and bottom

Sea ice concentration	*sic*	%	Monthly	
Dissolved oxygen concentration	o2	mol m-3	Monthly	surface and bottom
Total primary organic carbon production (by all types of phytoplankton)	*intpp*	mol C m-3 s-1	Monthly	depth-integrated To be calculated as *intpp* = *intpp_lphy* + *intpp_sphy* + *intpp_diaz*
Small phytoplankton productivity	*intpp_sphy*	mol C m-3 s-1	Monthly	depth-integrated
Large phytoplankton productivity	*intpp_lphy*	mol C m-3 s-1	Monthly	depth-integrated
pH	*ph*	1	Monthly	surface and bottom
Salinity	*so*	psu	Monthly	surface and bottom

12.2.1 Historical fishing effort

For this round, modelers will use their own default fishing effort and catch data. In most cases this will be Sea-Around-Us-Project (SAUP) data (http://www.seaaroundus.org/data/#/eez) obtained through a memorandum of understanding (MOU) or data from Regional Fisheries Management Organizations (RFMOs) or local fisheries agencies. Modelers that do not have access to these data are asked to contact the ISIMIP sectoral coordinators.

12.2.2 Spin-up and initialization

Input data is provided from 1950 to 2004. Years until 1970 can be replicated as needed and used for spin-up. Historical reporting is from 1971-2005, but if your model starts later, start when your model normally starts!

12.3 Output data

- **Provide temporally (monthly) and spatially (1 x 1 degree grid) explicit column-integrated time series (1971-2004)** (All files should be saved with .nc4 file extension; a conversion script for .csv files can be found at: http://vre1.dkrz.de).

- **Use variable names as specified in Table** 32**, and check the overall ISIMIP simulation protocol for how to name your files**
- **If there is no data value for outputs, use the value: 1.e+20f**
- **Mandatory output**: this is the priority for first round of model comparisons (provide as many as possible!)
- **Optional output**: if you can, please store or upload all output you receive from your model, we may eventually use it

Table 32: Common output variables to be provided by global and regional marine fisheries models.

Variable (long name)	Variable name	Unit (NetCDF format)	Resolution	Comments
Mandatory output from global and regional models (provide as many as possible)				
TOTAL system biomass density	**tsb**	g C m-2	monthly	all primary producers and consumers
TOTAL consumer biomass density	**tcb**	g C m-2	monthly	all consumers (trophic level >1, vertebrates and invertebrates)
Biomass density of consumers >10cm	**b10cm**	g C m-2	monthly	if L infinity is >10 cm, include in >10 cm class
Biomass density of consumers >30cm	**b30cm**	g C m-2	monthly	if L infinity is >30 cm, include in >30 cm class
TOTAL Catch (all commercial functional groups / size classes)	**tc**	g m-2	monthly	catch at sea (commercial landings plus discards, fish and invertebrates)
TOTAL Landings (all commercial functional groups / size classes)	**tla**	g m-2	monthly	commercial landings (catch without discards, fish and invertebrates)
Optional output from global and regional models				
Biomass density of commercial species	**bcom**	g C m-2	monthly	Discarded species not included (Fish and invertebrates)
Biomass density of large consumers >90cm and <100kg	**blarge**	g C m-2	monthly	

Biomass density of medium consumers >30cm and <90cm	**bmed**	g C m-2	monthly	
Biomass density of small consumers <30cm	**bsmall**	g C m-2	monthly	
Biomass density (by functional group / size class)	**b-<class>-<group>**	g C m-2	monthly	Provide name of each size class (<class>) and functional group (<group>) used, and provide a definition of each class/group
Catch (by functional group / size class)	**c-<class>-<group>**	g m-2	monthly	Provide name of each size class (<class>) and functional group (<group>) used, and provide a definition of each class/group
Catch of large consumers >90cm and <100kg	**clarge**	g m-2	monthly	
Catch of medium consumers >30cm and <90cm	**cmed**	g m-2	monthly	
Catch of small consumers <30cm	**csmall**	g m-2	monthly	
TOTAL Catch of consumers >10cm	**tc10cm**	g m-2	monthly	
TOTAL Catch of consumers >30cm	**tc30cm**	g m-2	monthly	

12.4 Additional information for regional marine ecosystem & fisheries models

12.4.1 Ocean regions

Table 33: Ocean regions

Ocean regions (short name for use in file names)		
North Sea (north-sea)	4°30’W–9°30’E	50°30’N–62°30’N
Baltic Sea (baltic-sea)	15°30’E–23°30’E	55°30’N–64°30’N

North-west Meditteranean (nw-med-sea)	1°30’W–6°30’E	36°30’N–43°30’N
Adriatic Sea (adriatic-sea)	11°30’E–20°30’E	39°30’N–45°30’N
Mediterranean Sea (med-glob)	6°30’W–35°30’E	29°30’N–45°30’N
South-East Australia (se-australia)	120°30’E–170°30’E	47°30’S–23°30’S
Eastern Bass Strait (east-bass-strait)	145°30’E–151°30’E	41°30’S–37°30’S
Cook Strait (cook-strait)	174°30’E–179°30’E	46°30’S–40°30’S
North Humboldt Sea (humboldt-n)	93°30’W–69°30’W	20°30’S–6°30’N

13 Terrestrial Biodiversity

13.1 Experiments

Table 34: provides an overview of all experiments to be run in the terrestrial-biodiversity sector in ISIMIP2a.

Table 34: Experiment summary for terrestrial-biodiversity models.

	Climate Data	Scenario	Human influences, land use (LU)	Other settings (sens-scenario)	# runs
Historical runs	EWEMBI	hist	nat	no CO2	1

13.2 Sector-specific input

Table 35: Biodiversity-specific input data used for building our models.

Dataset	Description	More info	Dates	Scale	Variables included
EWEMBI	Bioclimatic variables	30-year monthly means of minimum temperature (tasmin), maximum temperature (tasmax) and total precipitation (pr) were calculated and used to derive 19 bioclimatic variables; see (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005)	**30-yr averages of 1980 – 2009 (1995)**	global, 0.5° (EWEMBI)	Bio4 (*temperature seasonality*), *Bio5 (max temperature of warmest month)*, *Bio12 (annual precipitation) and Bio15 (precipitation seasonality)*, *Bio18 (precipitation of warmest quarter) and Bio19 (precipitation of coldest quarter)*

13.3 Output data

Table 36: Output variables to be reported by terrestrial-biodiversity sector models.

Variable (long name)	Variable name	Units (NetCDF format)	Frequency	Comment
Essential outputs				
Species probability of occurrence				
Amphibian species probability of occurrence	**amphibianprob**	Probability of occurrence per cell[1]	30-year period centered around 1995 (1980 – 2009)	Results from individual SDMs assuming no dispersal.[2]
Terrestrial bird species probability of occurrence	**birdprob**			
Terrestrial mammal species probability of occurrence	**mammalprob**			
Summed probability of occurrence				
Amphibian summed probability of occurrence	**amphibiansumprob**	Summed probability of occurrence per cell[1]	30-year period centered around 1995 (1980 – 2009)	Aggregated results from individual SDMs assuming no dispersal.[2]
Terrestrial bird summed probability of occurrence	**birdsumprob**			
Terrestrial mammal summed probability of occurrence	**mammalsumprob**			
Endemic summed probability of occurrence				
Summed probability of endemic amphibian species[3]	**endamphibiansumprob**	Summed probability of occurrence per cell[1]	30-year period centered around 1995 (1980 – 2009)	Aggregated results from individual SDMs assuming no dispersal.[2]
Summed probability of endemic terrestrial bird species[3]	**endbirdsumprob**			
Summed probability of endemic terrestrial mammal species[3]	**endmammalsumprob**			
Threatened summed probability of occurrence				
Summed probability of threatened amphibian species[4]	**thramphibiansumprob**	Summed probability of occurrence per cell[1]	30-year period centered around	Aggregated results from individual SDMs assuming

Summed probability of threatened terrestrial bird species[4]	**thrbirdsumprob**		1995 (1980 – 2009)	no dispersal.[2]
Summed probability of threatened terrestrial mammal species[4]	**thrmammalsumprob**			
Species richness				
Amphibian species richness	**amphibiansr**	Estimated number of species (species richness) per cell	30-year period centered around 1995 (1980 – 2009)	Results from macroecological richness models
Terrestrial bird species richness	**birdsr**			
Terrestrial mammal species richness	**mammalsr**			

[1] For the Maximum Entropy (MaxEnt) model algorithm the output is not probability, but habitat suitability/relative occurrence probability. Values also range between 0 and 1.
[2] No dispersal assumes that species can only be present where they are actually present according to the IUCN and BirdLife range maps.
[3] Endemic (range-restricted) species are the smallest ranging 15% of all species.
[4] Threatened species are all species that are either (i) critically endangered, (ii) endangered or (iii) vulnerable according to their IUCN red list status.

14 Health: Temperature-related mortality (TRM)

There are no restrictions regarding the type of (empirical) models (GAMs, DLNMs, linear threshold model, etc.) to be used as long as the methodology has been documented in previous peer-reviewed publications. It also does not matter at which spatial scale the model operates (city-scale, regional, national, global), with the possible restrictions stemming from the input data provided.

14.1 Experiments

See **Table** 6 for a general explanation of the pressoc, and varsoc experiments. Specification for TRM models are

pressoc: no long-term trend in the relationship between temperature and mortality. E.g., constant exposure-response function (ERF), if possible, estimated from observational data in period centred on the year 2000

varsoc: reflecting historical trend in the relationship between temperature and mortality. E.g., varying ERF, as estimated from observations in adjacent subperiods; if possible, extrapolate to reporting years outside of observational period using external factors (such as climatic factors, etc.).

Table 37: Summary of experiments for TRM models.

Climate Data	Scenario	Human Impacts	# runs
WATCH-WFDEI	Hist	pressoc varsoc	2
GSWP3-W5E5	Hist	pressoc varsoc	2
GSWP3-EWEMBI	Hist	pressoc varsoc	2
GSWP3	Hist	pressoc varsoc	2
PGMFD v.2 (Princeton)	Hist	pressoc varsoc	2

14.2 Output data

Table 38: Output variables to be reported by TRM models

Variable (long name)	Variable name	Unit	Temporal resolution	Comments
Number of deaths attributable to cold	**ancold-<r>**	1	daily	For ERF models, this occurs when temperature is below threshold (e.g., minimum mortality temperature (MMT)). Report 0 if temperature above threshold. Can have gender, age, etc. dimensions; see below.
Number of deaths attributable to heat	**anheat-<r>**	1	daily	Temperature above threshold (ERFs). Report 0 if temperature below threshold. Can have gender, age, etc. dimensions; see below.
Baseline total mortality	**btm**	1	daily	To be reported as annual series of mean daily total mortality, or as a single number of mean daily mortality; to be used for computations of attributable fractions. Can have gender, age, etc. dimensions; see below.
Population	**pop**	1	annual or 5-year intervals	Baseline population data should be provided for computations of mortality rates (i.e. deaths per total population). Can have gender, age, etc. dimensions; see below.

Instructions on reporting results:

- If different **realizations** of the model are applied, then these should be indicated by the specifier **<r>**. E.g. to reflect a central, upper, and lower estimate of the ERF:
  <r> = lower, central, upper
  Please explain the meaning of these realizations in the online model documentation; contact the ISIMIP coordination team in case of questions.
- If data are disaggregated e.g. by **age group, gender**, etc., they should be reported along an additional dimension, described by an auxiliary coordinate variable, in the NetCDF files. See the example provided at https://www.isimip.org/protocol/preparing-simulation-files/.
- For local (non-gridded) data, **locations (cities/regions/countries)** should be reported along an additional dimension called *location*, with the location name given as string in an auxiliary coordinate variable called *location_name*, in the NetCDF files. In addition, coordinates of

the location should be reported in auxiliary variables called *location_lat* and *location_lon*. See the example provided at https://www.isimip.org/protocol/preparing-simulation-files/. The <region> specifier in the file name should be set to "local".

- For gridded data, the <region> specifier in the file name should be "global" or indicate a region or country.

15 References

Arnell, N. (1999). A simple water balance model for the simulation of streamflow over a large geographic domain. *Journal of Hydrology*, *217*(3–4), 314–335.

Cescatti, A., & Piutti, E. (1998). Silvicultural alternatives, competition regime and sensitivity to climate in a European beech forest. *Forest Ecology and Management*, *102*(2), 213-223.

Choulga, M., Kourzeneva, E., Zakharova, E., & Doganovsky, A. (2014). Estimation of the mean depth of boreal lakes for use in numerical weather prediction and climate modelling, Tellus A. *Dyn. Meteorol. Oceanogr.*, *66*(1), 21295.

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H. and Buontempo, C. (2020) WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data*, 12, 2097–2120.

Davie, J. C., Falloon, P. D., Kahana, R., Dankers, R., Betts, R., Portmann, F. T., . . . Arnell, N. (2013). Comparing projections of future changes in runoff and water resources from hydrological and ecosystem models in ISI-MIP. *Earth System Dynamics Discussions*, *4*(1), 279–315.

De Lary, R. (October, 2015). *Massif des Landes de Gascogne. II – ETAT DES CONNAISSANCES TECHNIQUES*. Bordeaux: CRPF Aquitaine.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T. and Hanasaki, N. (2006) GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface. *Bulletin of the American Meteorological Society*, 87(10), 1381–98.

Dlugokencky, E., & Tans, P. (2019). *Trends in atmospheric carbon dioxide*. Retrieved November 2, 2019, from National Oceanic & Atmospheric Administration, Earth System Research Laboratory (NOAA/ESRL): https://www.esrl.noaa.gov/gmd/ccgg/trends/gl_data.html

Döll, P., & Schmied, H. M. (2012). How is the impact of climate change on river flow regimes related to the impact on mean annual runoff? A global-scale analysis. *Environmental Research Letters*, *7*(1), 14037.

Döll, P., Kaspar, F., & Lehner, B. (2003). A global hydrological model for deriving water availability indicators: Model tuning and validation. *Journal of Hydrology*, *270*(1–2), 105–134.

Duncker, P. S., Barreiro, S. M., Hengeveld, G. M., Lind, T., Mason, W. L., Ambrozy, S., & Spiecker, H. (2012). Classification of Forest Management Approaches: A New Conceptual Framework and Its Applicability to European Forestry. *Ecology and Society*, *17*(4).

Elliott, J., Müller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K. J., Büchner, M., . . . Ruane, A. C. (2015). The Global Gridded Crop Model Intercomparison: Data and modeling protocols for Phase 1 (v1.0). *Geosci. Model Dev.*, *8*, 261-277.

Fekete, B. M., Vörösmarty, C. J., & Grabs, W. (2000). Global Composite Runoff Fields on Observed River Discharge and Simulated Water Balances. *GRDC Reports*, *22*(115).

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., . . . Hill. (2011). Solutions for a cultivated planet. *Nature*, *478*(7369), 337-342.

Fürstenau, C., Badeck, F. W., Lasch, P., Lexer, M. J., Lindner, M., Mohr, P., & Suckow, F. (2007). Multiple-use forest management in consideration of climate change and the interests of stakeholder groups. *Eur J Forest Res*, *126*, 225-239.

González, J. R., & Palahí, M. (2005). Optimising the management of Pinus sylvestris L. stand under risk of fire in Catalonia (north-east of Spain). *Ann. For. Sci. 62*, *62*, 493-501.

Gosling, S. N., & Arnell, N. W. (2011). Simulating current global river runoff with a global hydrological model: Model revisions, validation, and sensitivity analysis. *Hydrological Processes*, *25*(7), 1129–1145.

Gosling, S. N., Warren, R., Arnell, N. W., Good, P., Caesar, J., Bernie, D., . . . Smith, S. M. (2011). A review of recent developments in climate change science. Part II: The global-scale impacts of climate change. *Progress in Physical Geography*, *35*(4), 443–464.

Gutsch, M., Lasch, P., Suckow, F., & Reyer, C. (2011). Management of mixed oak-pine forests under climate scenario uncertainty. *Forest Systems*, *20*(3), 453-463.

Haddeland, I. C. (2011). Multimodel estimate of the global terrestrial water balance: setup and first results. *Journal of Hydrometeorology*, *110531121709055*.

Haith, D. A., & Shoemaker., L. L. (1987). Generalized Watershed Loading Functions for stream flow nutrients. *Water Resour. Bull.*, *23*, 471-478.

Håkanson, L. (1995). Models to predict Secchi depth in small glacial lakes. *Aquatic Science*, *57*(1), 31–53.

Hanewinkela, M., & Pretzsch, H. (2000). Modelling the conversion from even-aged to uneven-aged stands of Norway spruce (Picea abies L. Karst.) with a distance-dependent growth simulator. *Forest Ecology and Management*, *134*, 55-70.

Hein, S., & Dhôte, J.-F. (2006). Effect of species composition, stand density and site index on the basal area increment of oak trees (Quercus sp.) in mixed stands with beech (Fagus sylvatica L.) in northern France. *Ann. For. Sci.*, *63*, 457-467.

Hijmans, R., Cameron, S., Parra, J., Jones, P., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965-1978.

Hurtt, G., Chini, L., Sahajpal, R., Frolking, S., & et al, .. (In prep.). Harmonization of global land-use change and management for the period 850-2100. *Geoscientific Model Development*.

Kerr, G. (1996). The effect of heavy or 'free growth' thinning on oak (Quercus petraea and Q. robur). *Forestry: An International Journal of Forest Research*, *69*(4), 303-317.

Kim, H. (. (n.d.). *Global Soil Wetness Project Phase 3*. Retrieved from Global Soil Wetness Project Phase 3: http://hydro.iis.u-tokyo.ac.jp/GSWP3/

Klein Goldewijk, D. i. (2016). *A historical land use data set for the Holocene; HYDE 3.2 (replaced)*. Utrecht University. DANS.

Koster, R. D., Fekete, B. M., Huffman, G. J., & Stackhouse, P. W. (2006). Revisiting a hydrological analysis framework with International Satellite Land Surface Climatology Project Initiative 2 rainfall, net radiation, and runoff fields. *Journal of Geophysical Research*, *111*(D22), D22S05.

Kourzeneva, E. (2010). External data for lake parameterization in Numerical Weather Prediction and climate modeling. *Boreal Environ. Res.*, *15*(2), 165–177.

Lähde, E., Laiho, O., & Lin, J. C. (2010). Silvicultural alternatives in an uneven-sized forest dominated by Picea abies. *Journal of Forest Research*, *15*(1), 14-20.

Lange, S. (2019a). WFDE5 over land merged with ERA5 over the ocean (W5E5). V. 1.0. doi:10.5880/pik.2019.023

Lange, S. (2019b). Earth2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI) v1.1. *GFZ Data Services*. doi:10.5880/pik.2019.004

Lange, S. (2019c). Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development*, 12, 3055–3070.

Lange, S. (2020). ISIMIP3BASD v2.4.1. *Zenodo*, doi:10.5281/zenodo.3898426.

Lascha, P., Badecka, F.-W., Suckowa, F., Lindnera, M., & Mohr, P. (2005). Model-based analysis of management alternatives at stand and regional level in Brandenburg. *Forest Ecology and Management*, *207*, 59-74.

Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.*, *296*(1-4), 1-22.

Liu, J., You, L., Amini, M., Obersteiner, M., Herrero, M., Zehnder, A. J., & Yang, H. (2010). A high-resolution assessment on global nitrogen flows in cropland. *National Academy of Sciences*, *107*(17), 8035-8040.

Loustau, D., Bosc, A., Colin, A., Ogée, J., Davi, H., Francois, C., . . . Delage, F. (2005). Modeling climate change effects on the potential production of French plains forests at the sub-regional level. *Tree physiology*, *25*, 813-23.

Meinshausen, M., Raper, S. C., & Wigley, T. M. (2011). Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 – Part 1: Model description and calibration. *Atmospheric Chemistry and Physics*, *11*(4), 1417–1456.

Millero, F., & Poisson, A. (1981). International one-atmosphere equation of state of seawater. *Deep-Sea Research*, *28*, 625-629.

Monfreda, C., Ramankutty, N., & Foley, J. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, *22*(GB1022).

Mueller, N., Gerber, J., Johnston, M., Ray, D., Ramankutty, N., & Foley, J. (2012). Closing yield gaps through nutrient and water management. *Nature*, *490*, 254-257.

Mund, M. (2004). *Carbon pools of European beech forests (Fagus sylvatica) under different silvicultural management*. Göttingen: Forschungszentrum Waldökosysteme.

Oleson, K. W., Niu, G.-Y., Yang, Z.-L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., . . . Qian, T. (2008). Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research*, *113*(G1), G01021.

Pape, R. (1999). Effects of Thinning Regime on the Wood Properties and Stem Quality of Picea abies. *Scandinavian Journal of Forest Research*, *14*(1), 38-50.

Portmann, F., Siebert, S., & Döll, P. (2010). MIRCA2000 – global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, *24*(1).

Potter, P., Ramankutty, N., Bennett, E. M., & Donner, S. D. (2011). Global fertilizer and manure, version 1: nitrogen fertilizer application. *NASA Socioeconomic Data and Applications Center*.

Pukkala, T., Miina, J., Kurttila, M., & Kolström, T. (1998). A spatial yield model for optimizing the thinning regime of mixed stands of Pinus sylvestris and Picea abies. *Scandinavian Journal of Forest Research*, *13*(1-4), 31-42.

Sacks, W. J., Deryng, D., Foley, J. A., & Ramankutty, N. (2010). Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*, *19*(5), 607-620.

Schneiderman, E. M., Pierson, D. C., Lounsbury, D. G., & Zion, M. S. (2002). Modeling the hydrochemistry of the Cannonsville watershed with Generalized Watershed Loading Functions (GWLF). *J. Am. Water Resour. Assoc.*, *38*, 1323-1347.

Schütz, J.-P., Götz, M., Schmid, W., & Mandallaz, D. (2006). Vulnerability of spruce (Picea abies) and beech (Fagus sylvatica) forest stands to storms and consequences for silviculture. *Eur J Forest Res*, *125*, 291-302.

Shatwell, T., Thiery, W., & Kirillin, G. (2019). Future projections of temperature and mixing regime of European temperate lakes. *Hydrology and Earth System Sciences*, *23*(3), 1533-1551.

Sheffield, J., Goteti, G., & Wood, E. F. (2006). Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. *Journal of Climate*, *19*(13), 3088–3111.

Štefančík, I. (2012). Growth characteristics of oak (Quercus petraea [Mattusch.] Liebl.) stand under different thinning regimes. *Journal of Forest Science*, *58*(2), 67-78.

Sterba, H. (1987). Estimating Potential Density from Thinning Experiments and Inventory Data. *Forest Science*, *33*(4), 1022-1034.

Stock, C. A., Dunne, J. P., & John, J. G. (2014). Global-scale carbon and energy flows through the marine planktonic food web: An analysis with a coupled physical-biological model. *Progress in Oceanography*, *120*, 1-28.

Subin, Z. M., Riley, W. J., & Mironov, D. (2012). An improved lake model for climate simulations: Model structure, evaluation, and sensitivity analyses in CESM1. *J. Adv. Model. Earth Syst.*, *4*(1), M02001.

Thivolle-Cazat, A. (2013). *Disponibilité en bois en Aquitaine de 2012 à 2025*. Bordeaux: FCBA, IGN, INRA, CRPF Aquitaine.

Tian, H., Yang, J., Lu, C., Xu, R., Canadell, J. G., Jackson, R., . . . Wini. (2018). The global N2O Model Intercomparison Project (NMIP): Objectives, Simu- lation Protocol and Expected Products. *B. Am. Meteorol. Soc.*

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, *50*, 7505–7514.

Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., . . . Best, M. (2011). Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century. *Journal of Hydrometeorology*, *12*(5), 823–848.

Wu, B., Yu, B., Yue, W., Shu, S., Tan, W., Hu, C., . . . Liu, H. (2013). A Voxel-Based Method for Automated Identification and Morphological Parameters Estimation of Individual Street Trees from Mobile Laser Scanning Data. *Remote Sensing*, *5*(2), 584-611.

Yoshimura, K., & Kanamitsu, M. (2008). Dynamical Global Downscaling of Global Reanalysis. *Monthly Weather Review*, *136*(8), 2983–2998.

Yoshimura, K., & Kanamitsu, M. (2013). Incremental Correction for the Dynamical Downscaling of Ensemble Mean Atmospheric Fields. *Monthly Weather Review*, *141*(9), 3087–3101.