ISIMIP2b Simulation Protocol
Published on 30 January 2017

The simulation protocol describes the simulation scenarios, input data sets and output variables necessary to participate in the ISIMIP2b simulation round. The scientific rationale and more detailed information about the pre-processing of input data can be found in the accompanying description paper Frieler et al. *Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b),* Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-229, in review, 2016.

Contents

1 Scenario design

The simulation scenarios are divided into three groups, depicted in **Figure 1** and **Figure 2**, directed at addressing distinct scientific questions:

- Quantification of pure climate-change effects of the historical warming compared to pre-industrial reference levels (Group 1).
- Future impact projections accounting for low (RCP2.6) and high (RCP6.0) greenhouse gas emissions assuming present day socio-economic conditions (Group 2).
- Future impact projections accounting for low (RCP2.6) and high (RCP6.0) greenhouse gas emissions assuming dynamic future socio-economic conditions (Group 3).

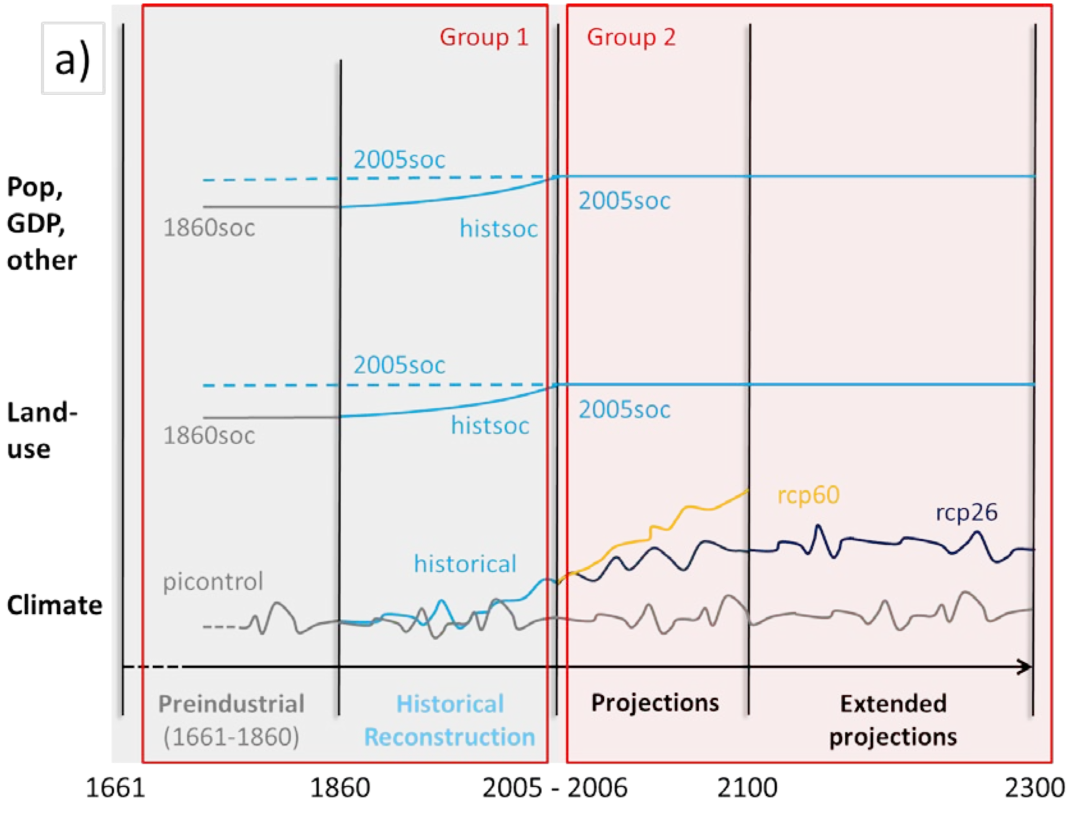

Figure 1 Schematic representation of the scenario design for ISIMIP2b. “Land use” also includes irrigation. “Other” includes other non-climatic anthropogenic forcing factors and management, such as fertilizer input, selection of crop varieties, flood protection levels, dams and reservoirs, water abstraction for human use, fishing effort, atmospheric nitrogen deposition, etc.. Panel a) shows the Group 1 and Group 2 runs. **Group 1** consists of model runs to separate the pure effect of the historical climate change from other human influences. Models that cannot account for changes in a particular forcing factor are asked to hold that forcing factor at 2005 levels (2005soc, dashed lines). **Group 2** consists of model runs to estimate the pure effect of the future climate change assuming fixed year 2005 levels of population, economic development, land use and management (2005soc).

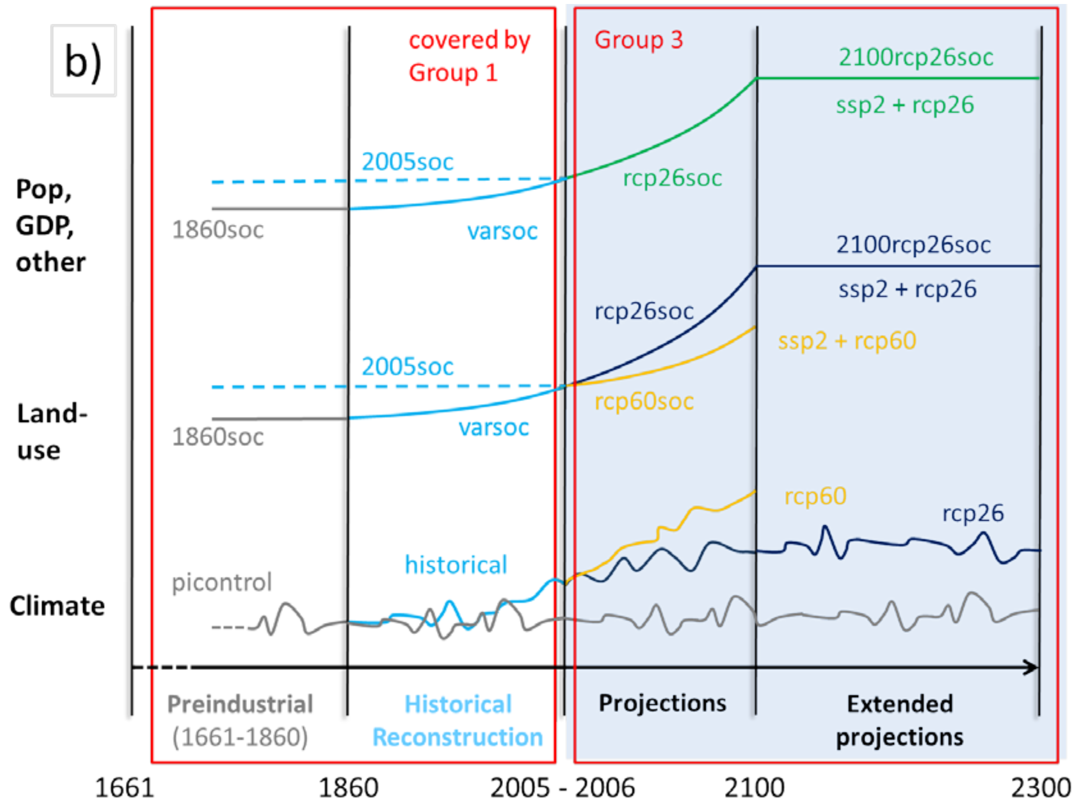

Figure 2 Schematic representation of the scenario design for **Group 3** runs. Group 3 consists of model runs to quantify the effects of the land use (and irrigation) changes, and changes in population, GDP, and management from 2005 onwards associated with RCP6.0 (no mitigation scenario under SSP2) and RCP2.6 (strong mitigation scenario under SSP2). Forcing factors for which no future scenarios exist (e.g. dams/reservoirs) are held constant after 2005.

2 Input data

2.1 Climate input data

- Bias-corrected to the EWEMBI data set at daily temporal and 0.5° horizontal resolution using updated versions of Fast-Track methods (see bias-correction Fact Sheet at www.isimip.org for methods description and further references).
- Daily time step, 0.5° horizontal resolution
- Pre-industrial (1661-1860), historical (1861-2005) and future (RCP2.6 and RCP6.0) conditions provided based on CMIP5 output of GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR and MIROC5. Output from two GCMs (GFDL-ESM2M and IPSL-CM5A-LR) includes the physical and biogeochemical ocean data required by the marine ecosystem sector of ISIMIP (see FISH-MIP, www.isimip.org/gettingstarted/marine-ecosystems-fisheries/).
- Priorities:
  1. IPSL-CM5A-LR
  2. GFDL-ESM2M
  3. MIROC5
  4. HadGEM2-ES

Table 1 Bias-corrected climate variables, including data sources of individual EWEMBI variables.

Variable	Short name	Unit
Near-Surface Relative Humidity	hurs	%
Near-Surface Specific Humidity	huss	kg kg^{-1}
Precipitation (rainfall + snowfall)	pr	kg m^{-2} s^{-1}
Snowfall Flux	prsn	kg m^{-2} s^{-1}
Surface Air Pressure	ps	Pa
Surface Downwelling Longwave Radiation	rlds	W m^{-2}
Surface Downwelling Shortwave Radiation	rsds	W m^{-2}
Near-Surface Wind Speed	sfcWind	m s^{-1}
Near-Surface Air Temperature	tas	K
Daily Maximum Near-Surface Air Temperature	tasmax	K
Daily Minimum Near-Surface Air Temperature	tasmin	K

Table 2 Variables provided without bias correction

Variable	Short name	**Unit**
Eastward Near-Surface Wind	uas	m s^{-1}
Northward Near-Surface Wind	vas	m s^{-1}
Eastward Wind at 250 and 850 hPa levels	ua	m s^{-1}
Northward Wind at 250 and 850 hPa levels	va	m s^{-1}

Table 3 Variables provided without bias correction at monthly resolution.

Variable	Short name	**Unit**
Ocean variables (for marine ecosystems & fisheries sector)		
Sea Water X Velocity	uo	m s^{-1}
Sea Water Y Velocity	vo	m s^{-1}
Sea Water Z Velocity	wo	m s^{-1}
Sea Water Temperature	to	K
Dissolved Oxygen Concentration	o2	mol m^{-3}
Total Primary Organic Carbon Production (by all types of phytoplankton) [calculated as sum of lpp + spp (IPSL) or sum of lpp + spp + dpp (GFDL)]	intpp	mol C m^{-2} s^{-1}
Small Phytoplankton Productivity	spp	mol C m^{-3} s^{-1}
Large Phytoplankton Productivity	lpp	mol C m^{-3} s^{-1}
Diazotroph Primary Productivity	dpp	mol C m^{-3} s^{-1}
Total Phytoplankton Carbon Concentration [sum of lphy + sphy (IPSL) or lphy + sphy + dphy (GFDL)]	phy	mol C m^{-3}
Small Phytoplankton Carbon Concentration	sphy	mol C m^{-3}

Large Phytoplankton Carbon Concentration	lphy	mol C m^{-3}
Diazotroph Carbon Concentration	dphy [diaz]	mol C m^{-3}
Total Zooplankton Carbon Concentration [sum of lzoo + szoo]	zooc	mol C m^{-3}
Small Zooplankton Carbon Concentration	szoo	mol C m^{-3}
Large Zooplankton Carbon Concentration	lzoo	mol C m^{-3}
pH	ph	1
Sea Water Salinity	so	psu
Sea Ice Fraction	sic	%
Large size-class particulate organic carbon pool	goc	mmol C m^{-3}
Photosynthetically-active radiation	Par	Einstein m^{-2} day^{-1}
Ocean variables (for tropical cyclones)		
Depth-resolved monthly mean Sea Water Potential Temperature	thetao	K
Sea Surface Temperature	tos	K
Atmospheric variables (for tropical cyclones)		
Air Temperature at all atmospheric model levels	ta	K
Specific Humidity at all atmospheric model levels	hus	kg kg^{-1}

2.2 Land-use patterns

The following land-use data are provided and described in detail in **Table 4**:

- Historical land-use (LU) changes from the HYDE3.2 data (Klein Goldewijk, 2016) (see **Figure 3**). Three, consistently generated disaggregation levels are provided:
    - Rainfed crop land, irrigated crop land, pastures and total crop land (the sum of rainfed and irrigated) – filename includes “landuse-totals”;
    - As above, with crop land divided into 5 functional crop types (LUH2) – filename includes “landuse-5crops”;

- As above, with crop land divided into 15 individual crops or crop groups (based on (Monfreda et al., 2008)) – filename includes “landuse-15crops”;

- Transient, future LU patterns generated by the LU model MAgPIE (Popp et al., 2014; Stevanović et al., 2016), assuming population growth and economic development as described in SSP2, for climate-change scenarios using RCP2.6 and RCP6.0 (see **Figure 3**).

The transition from historical to future LU patterns requires a harmonisation between the land-use classes and areas between the different data sets. A full description of how this will be done will appear here shortly.

Table 4 Agricultural land-use categories

Land-use type	Historical reconstruction	Future projections	Disaggregation into functional crop types (LUH2)	Individual crops or crop groups
Irrigated crops	HYDE	MAgPIE	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	HYDE	MAgPIE	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) C3 nitrogen-fixing disaggregated into: groundnut, pulses, soybean, others C_3 nitrogen-fixing C_4 annual disaggregated into: maize, tropical cereals C_4 perennial: sugarcane

Managed grassland (pastures)	HYDE	MAgPIE		
bioenergy production (rainfed grass)	-	MAgPIE		
bioenergy production (rainfed trees)	-	MAgPIE		
Urban	HYDE	constant (HYDE)		
Other (natural vegetation etc.)	1 - everything else	1 - everything else	The LUH2 data set includes additional natural land classes, which are consistent with the historical LU data provided here, and could be provided upon request.	(to be specified)

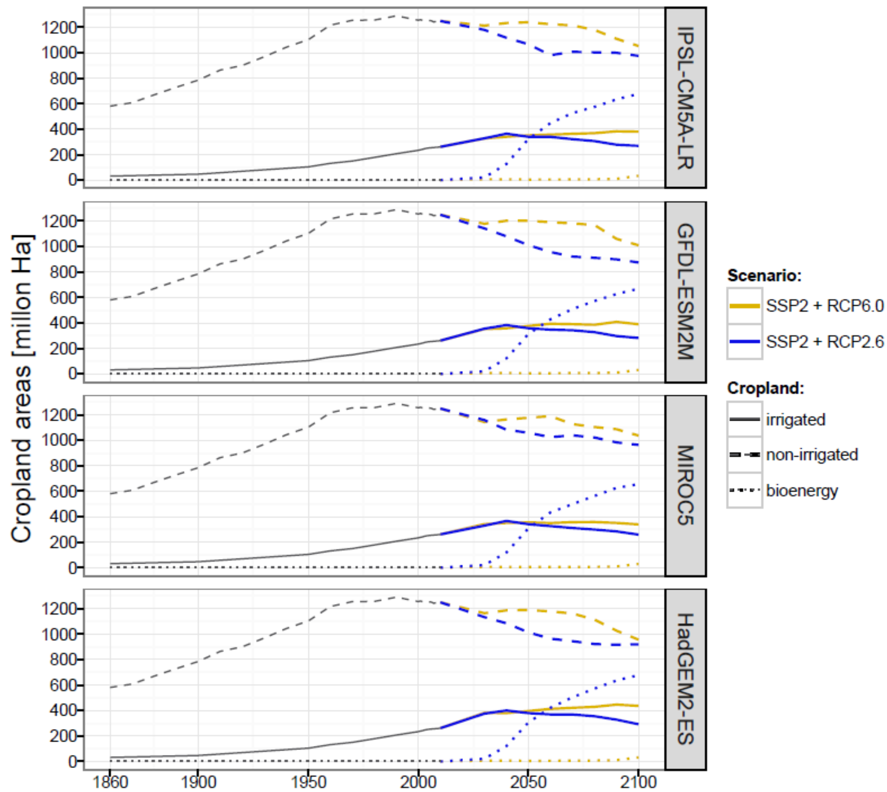

Figure 3 Time series of total crop land (irrigated (solid lines) and non-irrigated (dashed lines)) as reconstructed for the historical period (1860 - 2010) based on HYDE3.2 (Klein Goldewijk, 2016) and projected under SSP2 (2030-2100) assuming no explicit mitigation of greenhouse gas emissions (RCP6.0, yellow line) and strong mitigation (RCP2.6, dark blue line) as suggested by MAgPIE. Future projections also include land areas for second generation bioenergy production (not included in “total crop land”) for the demand generated from the Integrated Assessment Modelling Framework REMIND/MAgPIE, as implemented in the SSP exercise (dotted lines). Global data were linearly interpolated between the historical data set and the projections.

2.3 Sea-level rise patterns

Table 5 Information on sea-level-rise data.

Driver	Historical reconstruction	Future projections	Long-term projections
Sea-level rise	Observed time series up to 2000	From 2000 onwards, spatial patterns derived from GCMs. Regional variation of sea-level rise from glaciers and the large ice sheets are scaled from their respective gravitational patterns.	Constrained extrapolations have been extended to 2299.

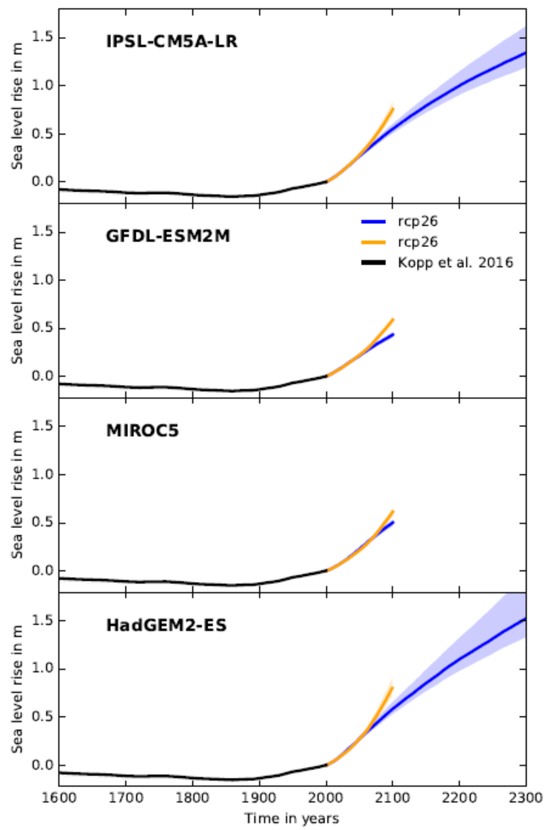

Figure 4 Time series of global total sea-level rise based on observations (Kopp et al., 2016, black line) until year 2000 and global-mean-temperature change from IPSL-CM5A-LR (panel 1), GFDL-ESM2M (panel 2), MIROC5 (panel 3) and HadGEM2-ES (panel 4) after year 2000: solid lines: Median projections, shaded areas: uncertainty range between the 5th and 95th percentile of the uncertainty distribution associated with the ice components. Blue: RCP2.6, yellow: RCP6.0. All time series relative to year 2000. Non-climate-driven contribution from glaciers and land water storage are added to the projections.

2.4 Population patterns and economic output (Gross Domestic Product, GDP)

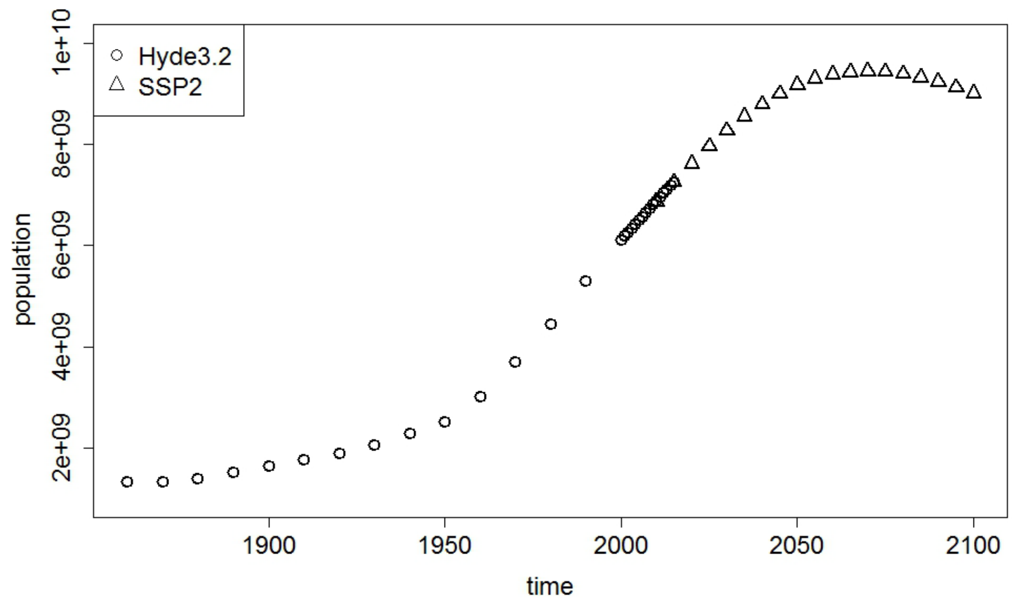

Figure 5 Time series of global population for the historical period (dots) and future projections following the SSP2 storyline (triangles).

Table 6 Socio-economic input data

Driver	Historical reconstruction	Future projections
GDP	• Annual country-level data from the Maddison project (Bolt and van Zanden, 2014, www.ggdc.net/maddison/maddison-project/home.htm)	• Annual and 10-year country-level data based on OECD projections from the SSP database (Dellink et al., 2015, https://secure.iiasa.ac.at/web-apps/ene/SspDb/)
Population	• Annual data on a 0.5° grid based on the HYDE3.2 database (Klein Goldewijk et al., 2010, 2011).	• Annual data on a 0.5° grid based on the national SSP2 population projections as described in Samir and Lutz, (2014). • Country-level age-specific data in 5-year age groups and all-age mortality rates in 5-year time.

2.5 Other human influences

For all of these input variables, we describe reconstructions to be used for the historical **histsoc** simulations (see **Table 7**). For models that do not allow for time-varying human influences across the historical period, human influences should be fixed at present-day (**2005soc**) levels (see dashed line in **Figure 1**, Group 1). Beyond 2005 all human influences should be held constant (Group 2) or varied according to SSP2 if associated projections are available (**Figure 2**, Group 3). Within ISIMIP2b projections are provided for future irrigation-water extraction, fertilizer application rates and nitrogen deposition (see **Table 7**).

Table 7 Data sets representing "other human influences" for the historical simulations (**histsoc**, Group 1) and the future projections accounting for changes in socio-economic drivers (**rcp26soc**/**rcp60soc**, Group 2).

Driver	Historical reconstruction	Future projections
Reservoirs & dams • **location** • **upstream area** • **capacity** • **construction/commissioning year**	Global data on 0.5° grid based on GranD database and the DDm30 routing network. Documentation: http://www.gwsp.org/products/grand-database.html **Note**: Simple interpolation can result in inconsistencies between the GranD database and the DDM30 routing network (wrong upstream area due to misaligned dam/reservoir location). A file is provided with locations of all larger dams/reservoirs adapted to DDM30 so as to best match reported upstream areas.	No future data sets are provided. Held fixed at year 2005 levels in all simulations.

Water abstraction for domestic and industrial uses	Generated by each modelling group individually (e.g. following the **varsoc** scenario in ISIMIP2a). Modelling groups that do not have their own representation could use an average of the ISIMIP2a data generated by the other models (available upon request). Before 1901 water abstraction for domestic and industrial uses is fixed at 1901 values.	Generated by each modelling group individually. For modelling groups that do not have their own representation, we provide files containing the multi-model mean (from WaterGAP, PCR-GLOBWB and H08) for domestic and industrial uses under SSP2 from the Water Futures and Solutions (WFaS) (Wada et al., 2016) project. Since this data is only available until 2050, the values should be kept constant from 2050 onwards. Also, the data provided for rcp26soc and rcp60soc are identical and both taken from simulations based on RCP6.0. RCP2.6 was not considered by WFaS. The difference is expected to be small compared to the influence of socio-economic conditions.
Irrigation water extraction (km^3)	Individually derived from the land-use and irrigation patterns provided. Water directly used for livestock (e.g. animal husbandry and drinking), except for indirect uses by irrigation of feed crops, is expected to be very low (Müller Schmied et al., 2016) and could be set to zero if not directly represented in the individual models.	Derived from future land-use and irrigation patterns provided based on output from the MAgPIE model (see section 0). Land-use projections are provided for: • SSP2+RCP6.0 • SSP2+RCP2.6; Direct water use for livestock should be ignored (i.e. can be set to zero).

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 et al.) based on HYDE3.2.	Inorganic N fertilizer use per area of crop land provided by MAgPIE, different for SSP2+RCP2.6 and SSP2+RCP6.0
Nitrogen (NH_X and NO_Y) deposition	Annual, 0.5° gridded data for 1850-2005 derived by taking the average of three atmospheric chemistry models (GISS-E2-R, CCSM-CAM3.5, and GFDL-AM3) in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) (0.5° x 0.5°) (Lamarque et al., 2013a, 2013b). GISS-E2-R provided monthly data; CCSM-CAM3.5 provided monthly data in each decade from 1850s to the 2000s; and GFDL-AM3 provided monthly data for 1850-1860, 1871-1950, 1961-1980, 1991-2000 and 2001-2010. Annual deposition rates calculated by aggregating the monthly data, and deposition rates in years without model output were calculated according to spline interpolation (CCSM-CAM3.5) or linear interpolation (for GFDL). The original deposition data was downscaled to spatial resolution of half degree (90° N to 90° S, 180° W to 180° E) by applying the nearest interpolation.	As per historical reconstruction for 2006-2099 following RCP2.6 and RCP6.0.
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. Given that the SAUP historical reconstruction starts in 1950, fishing effort should be held at a constant 1950 value from 1860-1950.	Held constant after 2005 (**2005soc**)

2.6 Focus Regions

Simulation data are welcome for all world regions, with a preference for those shown in dark orange in **Figure 6** and defined in **Table 8**, if feasible with your model. For regions not defined in the protocol, please contact the ISIMIP Team to agree on appropriate naming and define the location of the region in the metadata of your output files.

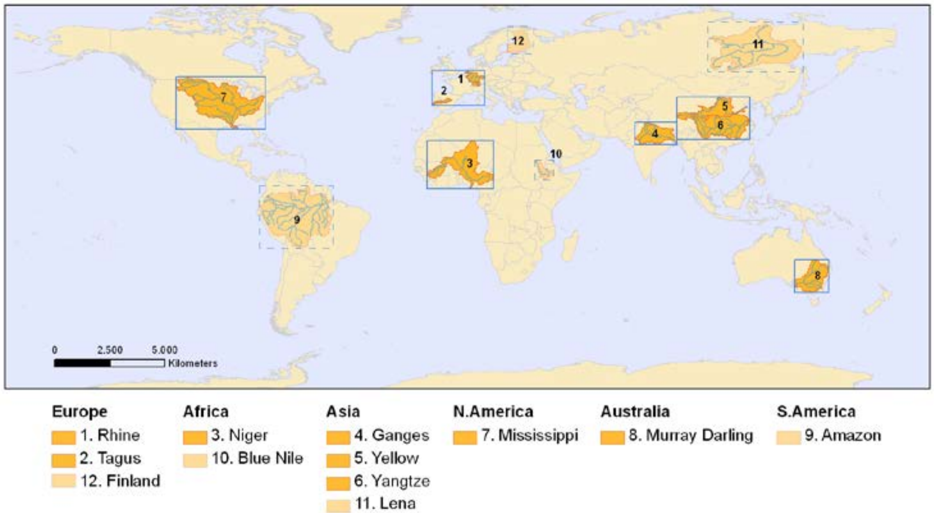

Figure 6 ISIMIP focus regions. Solid boxes (centered on river basins marked in dark orange) indicate cross-sectoral focus regions. Dashed boxes and river basins/countries marked in light orange indicate possible sector-specific extensions (e.g. forests in Finland and the Amazon region, water in the Upper Amazon, Lena, and Blue Nile river basins).

Table 8 List of ISIMIP focus regions as shown in **Figure 6**.

Focus region (shortname)	Zonal extent	Meridional extent	River basin(s) or Region (shortname)
North America (nam)	114°0’W– 77°30’W	28°30’N–50°0’N	Mississippi (Mississippi)
Western Europe (weu)	9°30’W–12°0’E	38°30’N–52°30’N	Rhine and Tagus (rhine)
West Africa (waf)	12°0’W–16°0’E	4°0’N–24°30’N	Niger (niger)
South Asia (sas)	73°0’E–90°30’E	22°0’N–31°30’N	Ganges (ganges)
China (chi)	90°30’E–120°30’E	24°0’N–42°0’N	Yellow and Yangtze (yellow, yangtze)
Australia (aus)	138°30’E–152°30’E	38°0’S –24°30’S	Murray Darling (murrydarling)
Amazon (ama)	80°0’W –50°0’W	20°0’S –5°30’N	Amazon (amazon)
Finland (fin)	21°0’E–32°0’E	59°30’N–79°30’N	-
Blue Nile (blu)	32°30’E - 40°0’E	8°0’N - 16°0’N	Blue Nile (bluenile)
Lena (len)	103°0’E - 141°30’E	52°0’N - 72°0’N	Lena (lena)

3 Conventions for File Names and Formats

3.1 General Notes

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. The ISIMIP Team will be glad to assist with the preparation of these files if necessary.

For questions or clarifications, please contact info@isimip.org or the data manager directly (buechner@pik-potsdam.de) before submitting files.

3.1.1 File names

Things to note:

- Report **one** variable per file
- Use **lowercase** letters in file names only
- Separate only specifiers with underscore "_"
- Use hyphens for specifier internal string separation, e.g. in model name
- NetCDF file extension is .nc4

The file name format is:

<modelname>_<gcm>_<climate_scenario>_<soc-scenario>_<co2sens-scenarios>_<variable>_<region>_<timestep>_<start-year>_<end-year>.nc4

The identifiers in brackets should be replaced with the appropriate identifiers from **Table 9**. Specifiers may be dependent on the sector. The identifiers <variable> might also contain information about the plant functional type (in the biomes and permafrost sectors). The pft naming is model-specific and hence has to be reported in the impact-model database entries for each model (www.isimip.org/impactmodels). In the forestry sector the identifier <variable> might contain information about the tree species. The species names codes are listed in **Table 16**.

Examples:

lpjml_ipsl-cm5a-lr_historical_histsoc_co2_qtot_global_annual_1861_2005.nc4

lpjml_ipsl-cm5a-lr_rcp26_rcp26soc_2005co2_yield_mai_global_annual_2006_2099.nc4

Table 9 Identifiers for file naming convention.

Item	Possible specifiers	Description
<modelname>		Model name
<gcm>	hadgem2-es, ipsl-cm5a-lr, miroc5, gfdl-esm2m	Name of the General Circulation Model from which climate-forcing data was used. Where point data has been used, include the name of the position, e.g. hadgem2-esForestBilyKriz

<climate_scenario>	picontrol, historical rcp26, rcp60	Climate & CO2 concentration scenario (RCP)
<soc -scenario>	nosoc, 1860soc, histsoc, 2005soc, rcp26soc, rcp60soc, 2100rcp26soc	Scenario describing other human influences, such as land use and land management.
<co2sens-scenario>	co2, 2005co2	'co2' for all experiments other than the sensitivity experiments for which 2005co2 is explicitly written. Note: even models in which CO2 has no effect should use the co2 identifier relevant to the experiment.
<variable>		Output variable names – see sector-specific tables.
<region>	global, [region]	Regions names given in Section 2.6.
<timestep>	3hr, daily, monthly, annual	The temporal resolution of your output data files.
<start-year>_<end-year>	e.g. 1861_1870	Files should be uploaded in 10-year pieces. For the transition from the historical to the future period (2005-2006), files should be separated, i.e. the identifiers would be 2001_2005 and 2006_2010.