9 Biomes

Table 19 provides an overview of all experiments to be run in the biomes sector in ISIMIP2a. This table is for your reference only; please read chapters 1-5 and this section carefully before beginning with the experiments.

Table 19 Experiment summary for Biomes models

	Climate Data	Scenario	Population/GDP	Land use (LU)	Other settings (sens-scenario)	# runs
Historical runs	PGMFD v.2 (Princeton)	hist	varsoc (see **Table 18**; if varsoc not possible, please submit the presoc run)	Hyde3 + MIRCA (no LU specifier)	historical CO2 (co2)	1
	GSWP3	hist	varsoc	Hyde3 + MIRCA (no LU specifier)	historical CO2 (co2)	1
	WATCH (WFD)	hist	varsoc	Hyde3 + MIRCA (no LU specifier)	historical CO2 (co2)	1
	WATCH+WFDEI.GPCC	hist	varsoc	Hyde3 + MIRCA (no LU specifier)	historical CO2 (co2)	1
	Optional run: PGMFD v.2 (Princeton)	hist	nat	reference run, natural vegetation only, no land-use	historical CO2 (co2)	1
	Optional run: PGMFD v.2 (Princeton)	Hist	varsoc	Hyde3 + MIRCA	fix at pre-industrial levels (pico2) = 280ppm	1

Future runs	GCM1 (HadGEM2-ES)	hist+ 2.6 (rcp2p6) 6.0 (rcp6p0)	pressoc + SSP2	Hyde3 + MIRCA Constant from 2000 onwards (lufix)	historical CO2 + RCP2.6 RCP6.0 (co2)	2
	GCM1 (HadGEM2-ES)	hist+ 2.6 (rcp2p6) 6.0 (rcp6p0)	pressoc + SSP2	Hyde3 + MIRCA Constant from 2000 onwards (lufix)	historical CO2 + fixed CO2 from 2000 onwards (noco2)	2
	GCM1 (HadGEM2-ES)	hist+ 2.6 (rcp2p6) 6.0 (rcp6p0)	pressoc + SSP2	Hyde3 + MIRCA MAgPIE from 2000 onwards (luvar)	historical CO2 + RCP2.6 RCP6.0 (co2)	2
	GCM2 (IPSL-CM5A-LR)	hist+ 2.6 (rcp2p6) 6.0 (rcp6p0)	pressoc + SSP2	Hyde3 + MIRCA Constant from 2000 onwards (lufix)	historical CO2 + RCP2.6 RCP6.0 (co2)	2
	GCM2 (IPSL-CM5A-LR)	hist+ 2.6 (rcp2p6) 6.0 (rcp6p0)	pressoc + SSP2	Hyde3 + MIRCA MAgPIE from 2000 onwards (luvar)	historical CO2 + RCP2.6 RCP6.0 (co2)	2
						13

9.1 Sector-specific input

Table 20 Biomes-specific input data

Dataset	Description	More info	Dates	Scale	Variables included
Optional (does not have to be harmonized):					
GSWP3	soil data	each model does have the option to use their own vegetation and soil datasets if they prefer		global, 30 arc sec (HWSD) or 0.5° (GSWP3), fixed	soil type
CRU elevation data					

9.2 Output data

IMPORTANT: The output variables reported for the biomes sector are also appropriate for use in the permafrost sector described in Section 7.6.

Table 21 Variables to be reported by biomes models

long name	units		output variable name	frequency	comment
Essential outputs					
Pools					
Carbon Mass in Vegetation biomass	kg m-2	per pft and gridcell total	**cveg_<pft>**	year	Gridcell total VegC is essential. Per PFT information is desirable.
Carbon Mass in Litter Pool	kg m-2	per pft and gridcell total	**clitter_<pft>**	year	Info for each individual pool.
Carbon Mass in Soil Pool	kg m-2	per pft and gridcell	**csoil_<pft>**	year	Info for each individual pool.

		total			
Fluxes					
Carbon Mass Flux out of atmosphere due to Gross Primary Production on Land	kg m-2 s-1	per pft and gridcell total	**gpp_<pft>**	mon (day)	
Carbon Mass Flux into atmosphere due to Autotrophic (Plant) Respiration on Land	kg m-2 s-1	per pft and gridcell total	**ra_<pft>**	mon (day)	
Carbon Mass Flux out of atmosphere due to Net Primary Production on Land	kg m-2 s-1	per pft and gridcell total	**npp_<pft>**	mon (day)	
Net Primary Production on Land allocated to leaf biomass	kg m-2 s-1	per pft and per gridcell	**npp_landleaf_<pft>**	mon (day)	
Net Primary Production on Land allocated to fine root biomass	kg m-2 s-1	per pft and per gridcell	**npp_landroot_<pft>**	mon (day)	
Net Primary Production on Land allocated to above ground wood biomass	kg m-2 s-1	per pft and per gridcell	**npp_abovegroundwood_<pft>**	mon (day)	
Net Primary Production on Land allocated to below ground wood biomass	kg m-2 s-1	per pft and per gridcell	**npp_belowgroundwood_<pft>**	mon (day)	
Carbon Mass Flux into atmosphere due to Heterotrophic	kg m-2 s-1	per pft and gridcell	**rh_<pft>**	mon (day)	

Respiration on Land		total			
Carbon Mass Flux into atmosphere due to total Carbon emissions from Fire	kg m-2 s-1	per pft and gridcell total	**fireint_<pft>**	mon (day)	
Fraction of cell burnt by fire	Fractional	Per pft and gridcell total	**firefrac_<pft>**		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)	kg m-2 s-1	per pft and gridcell total	**ecoatmflux_c_<pft>**	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.
Root autotrophic respiration	kg m-2 s-1	per pft and gridcell total	**rr_<pft>**	mon (day)	
Structure					
Fraction of absorbed photosynthetically active radiation	%	per pft and gridcell average	**fapar_<pft>**	mon (day)	
Leaf Area Index	1	per pft and gridcell average	**lai_<pft>**	mon (day)	
Plant Functional Type Grid Fraction	%	per gridcell	**pft_<pft>**	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	kg m-2 s-1	per pft and gridcell	**evap_<pft>**	mon (day)	
Evaporation from Canopy (interception)	kg m-2 s-1	per pft and gridcell	**intercep_<pft>**	mon (day)	the canopy evaporation+sublimation (if present in model).
Water Evaporation from Soil	kg m-2 s-1	per gridcell	**esoil**	mon (day)	includes sublimation.
Transpiration	kg m-2 s-1	per pft and gridcell	**trans_<pft>**	mon (day)	
Total Runoff	kg m-2 s-1	per gridcell	**qtot**	mon (day)	the total runoff (including "drainage" through the base of the soil model) leaving the land portion of the grid cell.
Soil Moisture	kg m-2	per gridcell	**soilmoist**	mon (day)	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	kg m-2 s-1	per gridcell	**qs**	mon (day)	the total surface runoff leaving the land portion of the grid cell.

Frozen soil moisture for each layer	kg m-2	per gridcell	**soilmoistfroz**	mon	Please provide soil moisture for all depth levels and indicate depth in m. Please provide for purposes of permafrost sector.
Snow depth	m	per gridcell	**snd**	mon	Grid cell mean depth of snowpack. Please provide for purposes of permafrost sector.
Snow water equivalent	Kg m-2	per gridcell	**swe**	mon	snow depth x snow density
Annual maximum thaw depth	m	per gridcell	**thawdepth**	year	calculated from daily thaw depths
Optional outputs					
Carbon Mass in Leaves	kg m-2	per pft and gridcell	**cleaf_<pft>**	year	
Carbon Mass in Wood	kg m-2	per pft and gridcell	**cwood_<pft>**	year	including sapwood and hardwood
Carbon Mass in Roots	kg m-2	per pft and gridcell	**croot_<pft>**	year	including fine and coarse roots
Others					
Temperature of Soil	K	per gridcell	**tsl**	mon (day)	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	%	per gridcell	**burntarea**	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 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.

9.3 Experiments

9.3.1 Historic runs and validation exercise

Table 22: Validation datasets for biomes models. Please note the data use restrictions indicated below the table.

Dataset	Source and further information	Variables included	Period	Scale	comment
SeaWiFS	Derived from SeaWiFS remotely sensed fAPAR product http://oceancolor.gsfc.nasa.gov/SeaWiFS/ Gobron et al., 2006	fAPAR (fraction of incident Photosynthetically Active Radiation that is absorbed by green vegetation (also called 'green vegetation cover')	1998-2005, monthly resolution	0.5 x 0.5 degrees spatial resolution	Reliable fAPAR values cannot be obtained when solar incidence is > 50°; cells where fAPAR could not be obtained for any month were excluded from the provided data set.
EVI	http://modis.gsfc.nasa.gov/data/dataprod/dataproducts.php?MOD_NUMBER=13	fAPAR	monthly resolution	0.5 x 0.5 degrees spatial resolution	The derivation of the FAPAR data is based on Eq. 11 in Xiao et al., 2005 (Ecological Applications, vol. 15, no. 3, pp. 954--969), which equates MODIS Enhanced Vegetation Index (EVI) to FAPARpav. For upscaling the MODIS monthly EVI data (either MYD13C2 or MOD13C2) from its native 0.05 degree resolution to 0.5 degree, we use a simple averaging method.

GIMMS g3 NDVI	http://www.mdpi.com/2072-4292/5/2/927	fAPAR	from 1981 15-days resolution	0.5 x 0.5 degrees spatial resolution	Zhu et al. *Remote Sens.* **2013**, *5*(2), 927-948; doi:10.3390/rs5020927
NDVI3g		fAPAR	15 days	???	
fapar3g	Upon request to U Boston Contact : Ranga B Myneni <rmyneni@bu.edu>	fAPAR	15 days	Global 0.05°	
Geoland-2 LAI	Fusion of SPOT4-VGT & AVHRR http://www.geoland2.eu/core-mapping-services/biopar.html	LAI	15 days ? TBC ?	Global 0.05° resolution for AVHRR and 1 km resolution for SPOT4-VGT	From 1981 to 1999, LAI, FAPAR and FCover are derived from NOAA/AVHRR Long Term Data Record (LTRD) dataset provided by NASA and the University of Maryland. They cover the globe at 0.05° resolution. From 2000 to the present, LAI, FAPAR and FCover are derived from SPOT/VGT data at 1km resolution.
Processed FLUXNET data [10]		GPP (Gross Primary Production)	monthly values, for the available period at each site	Different sites	Processing means partitioning of net carbon fluxes into GPP and respiration, and screening for outliers. In addition, gap-filling has been applied to shortwave solar radiation, followed by conversion to photosynthetic flux density (PPFD). Monthly GPP has been calculated by applying fitted relationships between GPP and PPFD (based on half-hourly data) throughout each month. Months with inadequate data to fit such relationships have been discarded.

10

					As long as the data are not published they are only available for the sole purpose of model evaluation in the context of ISIMIP. They are provided by Colin Prentice (colin.prentice@mq.edu.au) and Tyler Davis (tyler.davis@imperial.ac.uk) who would welcome any feedback.
Luyssaert forest site productivity dataset and extended version	matteo.campioli@uantwerpen.be Luyssaert et al. forest site data can be found at http://daac.ornl.gov/VEGETATION/guides/forest_carbon_flux.html	GPP (Gross Primary Production) + NPP (Net Primary production) + NEP for various sites where these variables are directly measured	Annual site data. The extended data have not been published yet and will be made available after the first publication		These data concern specific forest ecosystems that are not in equilibrium from previous disturbance. A specific site simulation protocol will be needed for comparison of NEP and carbon stocks. NPP and GPP data can be compared directly with model output, given the ISIMIP simulation protocol
CDIAC atmospheric CO2 concentrations	Carbon Dioxide Information Analysis Center (CDIAC, cdiac.ornl.gov)	Atmospheric CO2 concentration (seasonal phase and concentration)	1998-2005	26 sites	
De-trended CO2 inversion	Keeling, 2008; Bousquet et al., 2000, Rödenbeck et al., 2003; Baker et al., 2006; Chevalier et al., 2010	Atmospheric CO2 concentration (inter-annual comparison)	1980-2006	Different sites	

Data use Restrictions for data sets in Table 26**:**

GPP data derived from FLUXNET data

The data set is provided by Colin Prentice (colin.prentice@mq.edu.au) and Tyler Davis (tyler.davis@imperial.ac.uk) who would welcome any feedback. As long as the data are not published they are only available for the sole purpose of model evaluation in the context of ISIMIP. The underlying FLUXNET measurements and have to be acknowledged in any publication in the following way:

This work used eddy covariance data acquired by the FLUXNET community and in particular - the following networks: AmeriFlux (U.S. Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DE-FG02-04ER63917 and DE-FG02-04ER63911)), AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada (supported - CFCAS, NSERC, BIOCAP, Environment Canada, and NRCan), GreenGrass, KoFlux, LBA, NECC, OzFlux, TCOS-Siberia, USCCC. We acknowledge the financial support to the eddy covariance data harmonization provided - CarboEuropeIP, FAO-GTOS-TCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Université Laval and Environment Canada and US Department of Energy and the database development and technical support from Bekeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California - Berkeley, University of Virginia.

The processing of the data makes use of the daily shortwave radiation provided by WATCH forcing data. The use of this dataset should also be acknowledged with a citation similar to:

Weedon, G. P., Gomes, S., Balsamo, G., Best, M. J., Bellouin, N. & Viterbo, P. (2012) WATCH forcing databased on ERA-INTERIM. Retrieved 10 September 2013, from ftp://rfdata:forceDATA@ftp.iiasa.ac.at

9.3.2 Basic Metrics to measure the agreement between observations and simulations

1. Spatial agreement (calculated at each point in time)

Step 1

$$NME^{space} = \sum_i \left|x_i^{sim} - x_i^{obs}\right| \,/\, \sum_i \left|x_i^{obs} - \bar{x}^{obs}\right|$$

$$NMSE^{space} = \sum_i (x_i^{sim} - x_i^{obs})^2 \,/\, \sum_i (x_i^{obs} - \bar{x}^{obs})^2$$

Step 2 (removing the influence of the mean)

$$NME^{space} = \sum_i \left|(x_i^{sim} - \bar{x}^{sim}) - (x_i^{obs} - \bar{x}^{obs})\right| \,/\, \sum_i \left|x_i^{obs} - \bar{x}^{obs}\right|$$

$$NMSE^{space} = \sum_i ((x_i^{sim} - \bar{x}^{sim}) - (x_i^{obs} - \bar{x}^{obs}))^2 \,/\, \sum_i (x_i^{obs} - \bar{x}^{obs})^2$$

Step 3 (removing the influence of the variability)

$$NME^{space} = \sum_i \left|\frac{x_i^{sim} - \bar{x}^{sim}}{\sum_i \left|x_i^{sim} - \bar{x}^{sim}\right|/n} - \frac{x_i^{obs} - \bar{x}^{obs}}{\sum_i \left|x_i^{obs} - \bar{x}^{obs}\right|/n}\right| \,/\, \sum_i \left|x_i^{obs} - \bar{x}^{obs}\right|$$

$$NMSE^{space} = \sum_i \left(\frac{x_i^{sim} - \bar{x}^{sim}}{\sqrt{\sum_i (x_i^{sim} - \bar{x}^{sim})^2/n}} - \frac{x_i^{obe} - \bar{x}^{obs}}{\sqrt{\sum_i (x_i^{obs} - \bar{x}^{obs})^2/n}}\right)^2 \,/\, \sum_i (x_i^{obs} - \bar{x}^{obs})^2$$

Where x_i^{obs} and $x_i^{sim}$are the observed and simulated values of variable x in grid cell or at site I, respectively. $\bar{x}^{obs}$ is the mean observed values across all site or grid cells.

2. Temporal agreement (on global or regional level)

$NME^{time}$and $NMSE^{time}$ are calculated analogously to the above specifications where x_i^{obs} and $x_i^{sim}$are the observed and simulated global (or regional) mean values of variable x in year of month i, respectively. $\bar{x}^{obs}$ and $\bar{x}^{sim}$ is the mean observed and simulated values across all years and months, respectively.

3. Agreement with regard to seasonality

To compare the observed and simulated seasonality each simulated or observed month is represented by a vector in the complex plane, where the length of the vector corresponds to the magnitude of the variable for the specific month and the direction of the vector corresponds to the time of the year represented by the angle

$$\theta_t = 2\pi \frac{t-1}{12}$$

with month 1 (January) arbitrarily set to an angle of 0. A mean vector L is calculated by averaging the real and the imaginary parts of the 12 vectors x_t:

$L_x = \sum_t x_t \cos\theta_t$ and $L_y = \sum_t x_t \sin\theta_t$

The length of the mean vector divided by the annual value stands for seasonal concentration, $C = \sqrt{L_x^2 + L_y^2}/\sum_t x_t$ and $P = \arctan(L_x/L_y)$ stands for its phase. Thus if the variable is concentrated all in one month, seasonal concentration is equal to 1 and the phase corresponds to that month. If the variable is evenly spread over all months then the concentration is equal to zero and the phase is undefined. If either modeled or observed values have zero values for all months in a given cell or site then that cell/site is not included in the comparison. Modelled and observed phase are compared using mean phase difference

$$MPD = \frac{1}{\pi} \arccos[\frac{\cos(\omega_i - \vartheta_i)}{n}],$$

where ω_i is the modeled phase and ϑ_i is the observed phase. The measure can be interpreted as the average timing error as a proportion of the maximum error (6 months). For seasonal CO2 concentrations, where the data are monthly deviations from the mean CO2, we compare the seasonal amplitude instead of the seasonal concentration by comparing the simulated and observational sum of the absolute CO2 deviations for each month using the NME or NMSE from step 1 above.

9.3.3 Fast track runs for new models

Please consult the fast track protocol Section 7 for those runs and related information. It is available at www.isimip.org > Getting Started > ISIMIP Fast Track Protocol. In case of any questions please contact info@isimip.org. Please note that aside from harmonized climate and socio-economic input the default settings of your model should be used. Also note that for output data files the file name (as specified in Section 5.2 of the fast track protocol) is all lower case!