Protocol for initial mortality benchmarking simulations
V3.0, 01.11.22

Aim

To 1) evaluate an ensemble of demography models against observations of forest regrowth and stand structure. 2) assess the impact on disturbance rate increases on forest ecosystem carbon density.

Story plan
- Afforestation and carbon credit plantations. How can we assess resilience in the context of increasing disturbance frequencies but also increasing resources (CO2)?
- In 3 different biomes, what level of increase in disturbance rate would lead to the replacement of a forest? And how would higher [CO2] ameliorate disturbance competition?
- This is the first time an ensemble of demography models has been evaluated against observations. They allow us to explicitly capture the process of growth after planting and of mortality.
- We show how key climate-related forest variables change over the course of regrowth (Biomass, roughness, albedo)
  (for diagnostic purposes also woody productivity and mortality?)
- We evaluate biomass against regrowth data and old-growth data for 3 regions (sites).
- We examine whether the models capture well known aspects of stand size structure (self-thinning)
- We test whether the change in disturbance rates alter the balance / coexistence of PFTs (= lead to another kind of forest).
- Implications…

Simulation setup

If your runs contain any diversions from the protocol (variables, dimensions, units), please make a record and share with your outputs.

Outline: Single grid cell simulations, one for each major forest type. Co-located with reference data for comparison. We simulate, starting from bare ground, forest regrowth to the point of an old growth forest (calibrate your own model to give the regrowth curve).
Baseline run (P0) - use for calibration. We will provide regrowth curves for each ecosystem.
Then use the mortality/disturbance rates you have obtained there as baseline rate for changes in the sensitivity run (PS).

Locations:
Based on centre-referenced 0.5° grid cell.
- Boreal: Finland (Lon 23.25°, Lat 62.25°) 143 m.
- Temperate: Poland, Bialowieza (Lon 23.75°, Lat 52.75°)
- Tropical: Barro Colorado Island (Lon -79.75°, Lat 9.25°)

Suggested site locations are based on a combination of data availability (plot observations of forest dynamics growth and mortality rates, stand structure, DBH), biome coverage and experience of this group.

Simulation setup:

Baseline run (P0 spinup + 30 + clear cut + 420)**:**
- Spinup model according to your own procedure using the ==provided climate, CO₂ and N deposition are held constant at recent historical levels throughout the whole simulation at 2020 values== (as specified below).
  - Note: For the entire simulation (all steps below:) ==Keep repeating the 30 years of the climate dataset as forcing==.
- then run ==30 years== (to characterise the forest post-spinup)
- then reset to bare ground (year 31) and simulate the next ==420 years== of regrowth – to get the full regrowth trajectory displayed in the ==figure above==
- (this makes ==450 years simulation== in total after spin-up).

If logging occurs in your model, turn it off (we want natural forests).
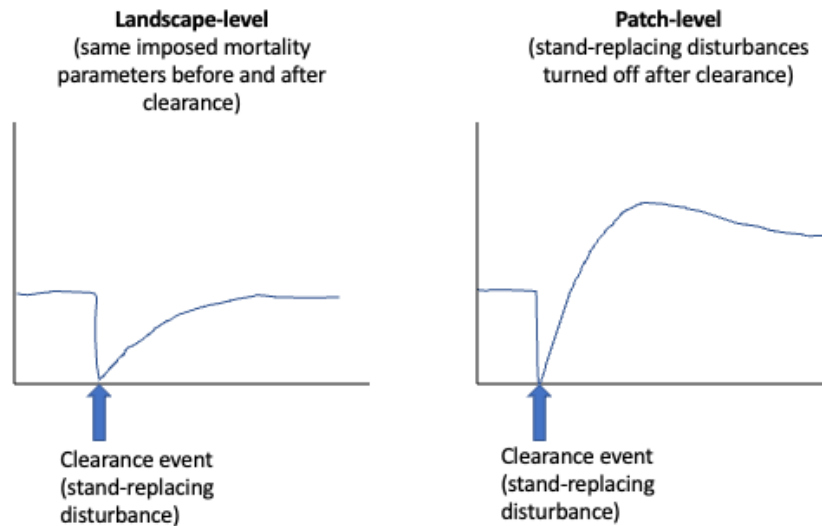==Submit all outputs from the post-spin-up phase==.

==**Important:** We will provide data for regrowth curves and old growth biomass for each of the 3 locations. You should adjust your P0 simulation as necessary to fit well to these datasets.==

Sensitivity runs (PS)
- Repeat the P0 ==run 6 times==, each with an increase in ==disturbance rate== from baseline to most extreme: The levels you run are up to you but should be set such that at the highest rate Forest transitions into a non-forest ecosystem after disturbance-recovery (representing a grassland-ecosystem). The definition of a forest is taken from the FAO (page 4 of https://www.fao.org/3/I8661EN/i8661en.pdf): " [...] ==trees higher than 5 meters and a canopy cover of more than 10 percent== [...]". It is up to you to choose ==the way to increase your disturbance rate== (e.g., increasing patch destroying disturbances, ==background mortality, everything==… it should be motivated by a choice of mechanism(s) which is consistent with ==climate-induced disturbance increases==). Note that for several of the models this will require some trial and error.
- ==Repeat the above 7 simulations (P0 + 6 sensitivities) with CO2 level of 562 ppm== (2020 level + 150 ppm, as used in FACE sites)

Patch vs landscape dynamics

Landscape-level
(same imposed mortality parameters before and after clearance)

Patch-level
(stand-replacing disturbances turned off after clearance)

Clearance event
(stand-replacing disturbance)

Clearance event
(stand-replacing disturbance)

We want to create patch-level regrowth curves for evaluation, but landscape-level regrowth curves for the figures in the impact section. Depending on your model structure, you may create this separation between patch and landscape by carrying out an additional simulation, or providing information from your standard simulation that allows us to recreate the patch-level regrowth curve. Patch-level dynamics are only needed for P0, not for the sensitivity simulations.

Vegetation composition: Allow the below species / types of trees to be the ones to grow there. If you normally simulate grasses within a forest grid cell, simulate them. 'convert' to your corresponding PFT- all modellers have to ensure that appropriate PFTs /species can grow.

Lon 23.25°, Lat 62.25° Boreal: Finland (FIN): Pinus sylvestris - shade intolerant needleleaf (PFT1), Picea abies - shade tolerant needleleaf (PFT2), Betula pendula - shade intolerant broadleaf deciduous (PFT3), Grasses combined (PFT8).

Lon 23.75°, Lat 52.75° Temperate; Bialowieza (BIA): Picea abies - shade tolerant needleleaf (PFT2), *Betula spp.*- Shade intolerant broadleaf deciduous (PFT3), *Carpinus betulus* or *Tilia cordata* - (intermediate) shade tolerant broadleaf deciduous (PFT4), Grasses (PFT8).

Lon -79.75°, Lat 9.25° Tropical; Barro Colorado Island (BCI): Tropical broadleaf evergreen shade intolerant (PFT5) + Tropical broadleaf evergreen shade tolerant (PFT6) Tropical broadleaf deciduous (PFT7), Grasses (PFT8).

(PFT1-8 reflect dimensions in netcdf, see section outputs)

Forcing data:

P0/PS:
- <mark>TRENDY (cru_jra)</mark> (1991-2020) ,<mark>6 hourly data randomly reordered by year</mark>, repeated) – with further temporal disaggregation done locally by each group for now.
    - Order of years:
    - 2014 1996 2010 1998 1991 2002 2012 2016 2017 1999 2008 2001 2015 2019 2006 1993 1994 2020 2000 2004 1992 1995 2018 2011 2005 2007 2013 2009 1997 2003
    - Note: this sequence of years within the 30 year cycle should always follow the order as indicated above. it should be repeated, not generated anew for each cycle.
- <mark>Fixed CO_2 at 2020 levels: 412 ppm</mark>
- <mark>Other forcing (e.g. N): KgN/ha/yr</mark>
    - FIN1: 5.01
    - BIA: 10.07
    - BCI   3.57
- Soil information is every model's default.

For outputs see next page…

Outputs

The aim of the outputs below is to be able to intercompare and evaluate:
   a. the baseline stand structure,
   b. mortality rates by both stem and biomass metrics, thereby together constraining variation in mortality rate across the size distribution,
   c. the mechanisms which are driving mortality at different stages in the age of the stand.
If you feel that there is a key output that is needed to interpret your model's behaviour which is not included here, please feel free to submit that as well (with a short explanation).

Please supply outputs as netcdf files. Use the example netcdf metadata and dimension format provided for each run for your own output, that will make the analysis later much easier. The example files are in the shared folder and can also be accessed from here: https://lu.box.com/s/mt9s1j72axfef8d9oaj91k9tp2kdvvt3 Please provide separate files for each site.

The key outputs are those related to mortality and stand structure

Where outputs are requested by size class (sizeclass), the following classes are requested, since they reflect available observations. If your model does not simulate a size class, please fill it with zero. If a tree is already included in a smaller size class it should not be included in a bigger one (e.g. <5 means 1.0-4.999):
<1 cm, <5, <10, <15, <20, <30, <40, <50, <60, <70, <80, <90, <100, <150, <200, >=200

Additional information:
To keep things transparent and more interpretable, we should provide some PFT-specific parameters/ behaviours: For each of your PFTs, please submit a description of the PFT, along with parameter values for (if not applicable, mark with X):

PFT name and short description	Vcmax $(\mu mol\ m^{-2}\ s^{-1})$	Stomatal slope (Unitless / $kPa^{0.5}$)* *prob need people to indicate the model for this to make sense	Leaf / wood / root lifespan (years)

Also, please report the changes you have made to mortality rates for each sensitivity runs:

PFT	PS	Mortality rate changed	Values
	1 ("baseline- old growth calibrated,P0)		

	…				
	7 ("transition to non-forest")				

File naming conventions:
If possible, create per-variable files, using the following naming conventions, this makes my life a lot easier. If you diverge from this format, please provide the necessary information to post-process them into this format later:
Label files as follows,
<model name>_<variable name>_<simulation>_<location>[_<mortality_rate_id>]
Where "simulation" is P0 and PS and "location" is FIN, BIA or BCI,
[ ] = additional naming convention relevant for sensitivity runs, with <mortality_rate_id> being a number between 1 and 8, where
1 = Calibrated morality rate to reproduce old-growth biomass content
…
7 = mortality rate that turns the forest into a non-forest (i.e., the above FAO forest definition is no longer valid)
Example: FATES_cwood_P0_Bia and FATES_cwood_PS_Bia_1 (should be equivalent)

Note:
- *Dimensions in italics* are optional dimensions, e.g., pft dimension
- The pft dimension is to be filled up the following way: place output from Pinus sylvestris/PFT1 (see above) into the pft-dimension called 1(where possible). place output from Picea abies/PFT2 into pft-dimesion called 2, and so on... Pfts that are absent at a particular location should have 0 in the output variable.

Long name	Output variable name	Units	Timestep	Dimensions	Notes
Pools and structure					
Carbon mass in vegetation by PFT	cveg	kg C m^{-2}	annual	pft, time	
Aboveground biomass	AGB	kg C m^{-2}	annual	time	Necessary to evaluate against regrowth curves (add, if possible, no need to delay simulations for it, we will otherwise do a Simple transformation of Cveg *0.75 - to

					remove roots)
Carbon mass in wood by PFT	cwood	kg C m^{-2}	annual	pft, time	
Carbon mass in wood by size class	cwood_size	kg C m^{-2}	annual	sizeclass, time	
Stem number by size class	nstem_size	count ha^{-1}	annual	sizeclass, time	
Leaf area index	lai	m^2 m^{-2}	annual	pft, time	
Crown area	CA	m^2 ha^{-1}	annual	pft, time	I
Basal Area	BA	m^2 ha^{-1}	annual	pft,time	
95 th percentile of tree height	height	m	annual	pft,time	
Roughness length (momentum)	z0	m	annual	time	Optional.
Albedo	albedo	n/a	annual	time	Optional.
Fluxes					
Woody biomass growth	WBgrowth	kg C m^{-2}yr^{-1}	annual	pft,time	gross growth
Basal area growth	BAgrowth	m^{2}ha^{-1} yr^{-1}	annual	pft,time	

Carbon Mass Flux lost from live wood due to mortality or other turnover process	cmort	kg C m^{-2}s^{-1}	annual	sizeclass, time	**IMPORTANT:** If possible, please split these outputs by mortality/turnover mechanism. Each mortality/turnover mechanism in your model (e.g. bioclimatic limits, disturbance, fire) should be presented separately. Please do not include any litter loss in these fluxes (for instance during fire). The sum of all these mechanisms must equal the total vegetation carbon turnover for that vegtype. Please only include wood in carbon fluxes.
Stem number Flux lost from vegetation due to mortality or other turnover process	stemmort	Count ha^{-1}yr^{-1}	annual	sizeclass, time, *pft*	**Note:** If time constraints do not allow this division of outputs, then total values for cmort and stemmort summed across all PFTs and size classes will provide substantial information already.
Carbon Mass	gpp	kg C	annual	pft, time	Positive flux is into land

Flux out of Atmosphere due to Gross Primary Production on Land		$m^{-2}s^{-1}$			
Carbon Mass Flux out of Atmosphere due to Net Primary Production on Land	npp	kg C $m^{-2}s^{-1}$	annual	pft, time	Positive flux is into land
Carbon Mass Flux out of Atmosphere due to Net Biospheric Production on Land	nbp	kg C $m^{-2}s^{-1}$	annual	time	This is the net mass flux of carbon between land and atmosphere calculated as photosynthesis MINUS the sum of plant and soil respiration, and carbon fluxes from fire. If your model includes other relevant fluxes such as reproduction, please include in the calculation as appropriate. Positive flux is into the land.
Note (esp. for global runs): Any other variables related to drought stress (MDK comment: "For the LPJ type assumption, I would also collect the nutrient variable, i.e. the one that is used with water stress to determine root/leaf allocation"					

Compare/calibrate output against
1) Biome-specific regrowth curves (REGROWTH)
2) biomass measurements, at single site old growth forest from that biome (EQ. BIOMASS DYNAMCIS)
3) stand structure from the same single old growth forest site (STAND_STRUCTURE)
   a. NSTEM_SIZE
   b. CWOOD_SIZE

Pick as many or the four calibration targets as possible, as specified below:

REGROWTH:
Data: 20-year average value.
Model output: cveg*0.75 or AGB should fall within that range for the year bins (mido´points are given in the benchmak files for plotting: 10,30,50,90,110,130,150,…Your model will have all years since disturbance, so you can eyeball the goodness, by simply plotting the output into the graph, like this: red is the data and black/blue model output (you will probably only have one black line):
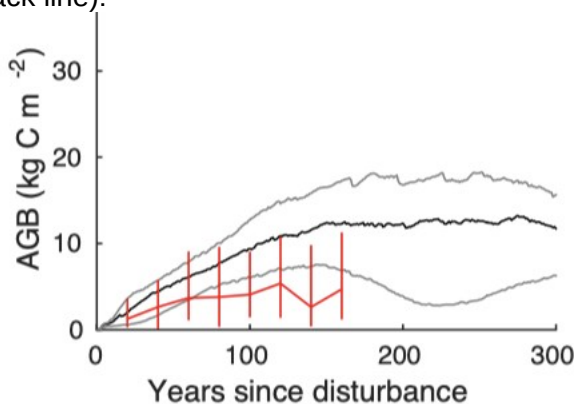
Please note that the target for the Boreal Forest is going to change, but you can use the one provided one as a starting point for now.

EQ. BIOMASS DYNAMICS:
Data: depending on site, single year census or average census year.
Model output: cveg*0.75 or AGB should fall within the range of the data most of the time, when at equilibrium (e.g., probably starting from simulation year 300+, depending on the model).

STAND STRUCTURE – NSTEM_SIZE:
Data: depending on the site: BIA: max, min and median value, derived from 5 sites; BCI; 10th, and 90th percentile, median. Values from bootstrapping using 100ha subsets of single site.
Model output: nstem_size of the last simulation year (420) should fall within the range specific to the site (see data).
For some sites/DBH-classes, data is marked with NA, there are no observations for this, but still report your outputs for this.
Please note that the target for BCI may change a bit, but you can use this one as a starting point for now.

STAND STRUCTURE – CWOOD_SIZE:

Data: depending on the site: BIA: max, min and median value, derived from 5 sites; BCI; 10th, and 90ᵗʰ percentile, median. Values from bootstrapping using 100ha subsets of single site.
Model output: cwood_size of the last simulation year, should roughly fall within the range specific to the site (see data).
For some sites/DBH-classes, data is marked with NA, there are no observations for this, but still report your outputs for this.
Please note that the target for BCI may change a bit, but you can use this one as a starting point for now.
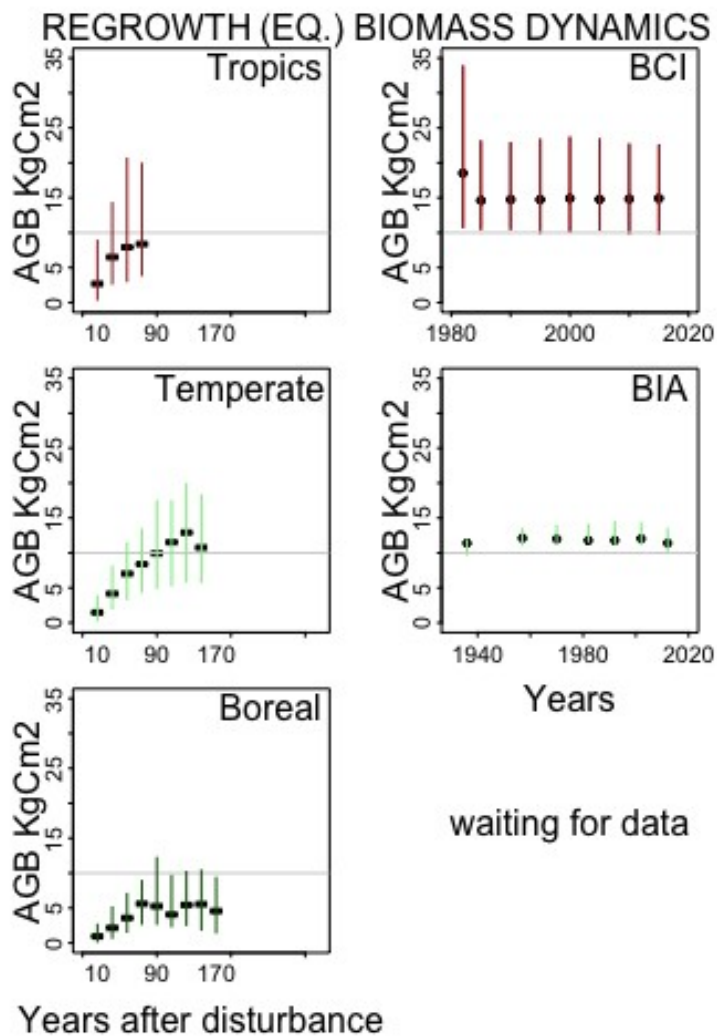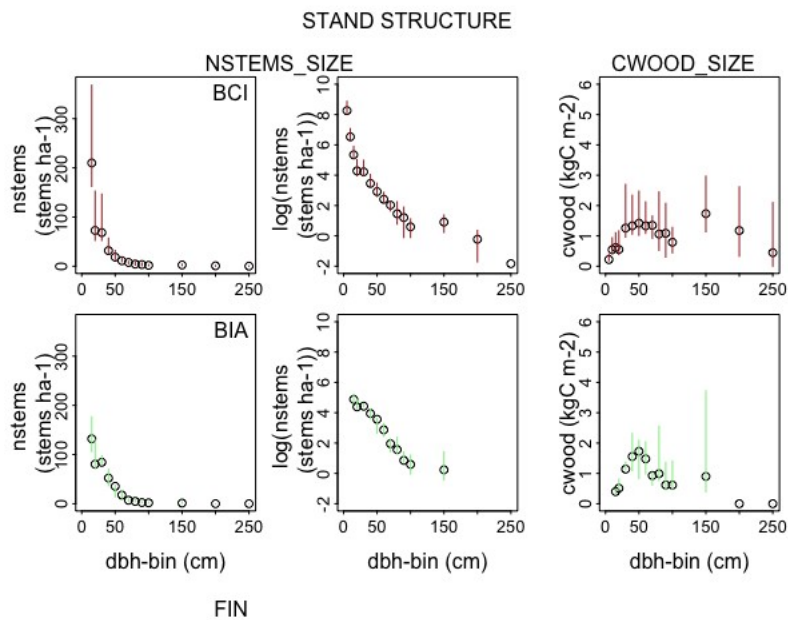

Fig1. Benchmark /calibration targets REGROWTH AND BIOMASS DYNAMICS for Tropics and Temperate biomes, alongside equilibrium biomass dynamics at respresentative old-growth

forest sites (BIA = Bialowieza, BCI = Barro Colorado Island. **Note that AGB is in kgCm2! And that this is woody carbon biomass only.**

Fig2. Stand structure benchmarks/calibration targets. First row is BCI data and second row is BIA data. NSTEMS_SIZE is visualised twice here, once on a log axis. **Note that CWOOD_SIZE for BCI may be subject to change, but get started with this for now**

Dataset description:

benchmark_regrowth_and_eq_dynamics.csv
contains benchmark targets REGROWTH and EQ. BIOMASS DYNAMICS

File header names	Description
"bin_ranges"	upper and lower range of 20 year data bins "(" = up to, and "[" = includes. E.g. "(20,40]" includes all values >20, and <= 40.
"bin_num"	Mid point of 20 year data bins. This is the point to compare your model against
"AGB_kgCm2_med"	Median value of 20-year binned observations of aboveground woody biomass biomass. Equivalent model output variable is cveg*0.75 or AGB.
"AGB_kgCm2_10"	10th percentile of of 20-year binned observations of aboveground woody biomass biomass. Equivalent model output variable is cveg*0.75 or AGB
"AGB_kgCm2_90"	90th percentile of of 20-year binned observations of aboveground woody biomass biomass. Equivalent model output variable is cveg*0.75 or AGB
"Biome"	Data originates from either Tropical, Temperate or Boreal biomes.
"final"	'Y' = final version of data. 'N' = potentially not final version, so use for now (already reasonable best guess), but be prepared to change.

benchmark_stand_structure.csv
contains benachmark targets STAND_STRUCTURE:NSTEMS_SIZE and STAND_STRUCTURE:CWOOD_SIZE.

File header names	Description
"dbh_classes_num"	Numeric type of dbh classes for quick plotting.
"dbh_classes"	Dbh classes as in the protocol
"nstem_size_ha.1"	Median nstem_size (Stem number by size clas)s in count /ha
"nstem_size_upper_ha.1"	Upper observed limit for nstem_size
"nstem_size_lower_ha.1"	Lower observed limit for nstem_size
"cwood_size_kgCm.2"	Median carbon mass in aboveground wood, by sizeclass in kgC m-2
"cwood_size_upper_kgCm.2 "	Upper observed limit
"cwood_size_lower_kgCm.2"	Lower observed limit
"site"	BCI or BIA, later also FIN
"final	Y' = final version of data. 'N' = potentially not final version, so use for now (already a very good guess), but be prepared to change.