Calibrated Land Model Intercomparison Project 1st Stage Draft Protocol

Background

Motivation

CalLMIP is a new initiative led by members of the Analysis and Integration of Modeling Earth Systems (AIMES) Land Data Assimilation Working Group and the International Land Modeling Forum (ILMF) Parameter Estimation Working Group.

This initiative is targeted at global-scale land models - i.e., prognostic, process-based land surface or terrestrial biosphere models that form part of Earth system models, or intermediate complexity models and vegetation demographic models that are routinely run at global scale (for example for the global carbon budget).

The overall goals envisaged for CalLMIP are:

- To quantify the reduction in inter-model spread as a result of parameter calibration using NEE, Qle and Qh for a wide range of variables (e.g., carbon and water stores and fluxes, aboveground biomass, soil carbon).
- To encourage the use of uncertainty quantification methods for constraining parameter uncertainty in global land models.
- To better understand and exchange knowledge on the variety of methods being used in global land model parameter calibration in order to identify best practices.

Guiding principles for CalLMIP

1. **First phase goal:** A community-led initiative to evaluate how well models' existing ("in-house") parameter calibration efforts reduce model-data misfit and the spread across models for key variables.

2. **‘Come as you are’ approach:** To leverage current model parameter estimation systems, rather than requiring development of new systems, maintaining flexibility to accommodate existing methods.

3. **Keep it as simple as possible:** Phase 1 is a starting point from which additional analysis/experiments can develop e.g., no global future projections, no tests of

different calibration set-up.

4. **All outputs available in public repository**: for access to the model outputs and additional analyses by the community, or access to the data repository provided, e.g. through user account on their HPC system

5. **Site optimization focus:** Key variables **to opt into** for optimization include carbon fluxes, evapotranspiration/latent heat flux, sensible heat flux (any number of variables to output).

Proposed timeline for the CalLMIP Phase 1

- March/April 2025: Three community sessions to decide on protocol, data, sites etc.
- May-August 2025: Preparation of input and observation files and data harmonization by the MacBean lab for all modeling groups.
- November 2025: Deliver inputs and finalised protocol to modeling groups. Draft of protocol paper circulated (if we decide to go ahead with this).
- November 2025: Webinar to present the first stage protocol and gather feedback on future CalLMIP phases and goals to the wider community.
- **November to February 2026: Phase 1a "test" calibration at 1 site – model prior and posterior simulations post calibration uploaded to modelevaluation.org by February 27th 2026.**
    - *Note – during this period we will be redoing the site selection process based on the FLUXNET 2025 release in December 2025 (see Section 2 for further details).*
- March to July 2026: Full multi-site Phase 1 parameter calibration experiments carried out by individual modeling groups – model prior and posterior simulations uploaded to modelevaluation.org by July 31st 2026.
- August to October 2026:
    - Global scale simulations with calibrated parameters. Calibrated model simulations uploaded to modelevaluation.org by October 30th 2026.
    - Initial analysis and draft paper for site-level calibration intercomparison with all participating model groups.

- Virtual meetings with participating groups to discuss initial analyses and intercomparison.

- Sept 2026: Presentation of initial site level calibration results at the 2nd Land Surface Modeling Summit.

- December 2026: Presentation of initial results from global post-optimization simulations at AGU 2026.

- January to April 2027:

    - Finalise site level intercomparison paper and submit.

    - Draft paper for global post-optimization intercomparison

    - Planning for possible Phase 2.

Protocol Plan

1. Calibration experiment set-up

1.1 Calibration Experiments

1. NEE, LE (Qle) and H (Qh)
2. NEE only
3. LE only
4. H only
5. NEE and LE
6. LE and H

1.2 Elements that will be constant across all experiments and modeling groups

To ensure comparability across modeling groups, certain key aspects of the calibration experiment will be standardized:

- Flux tower data: a common climate forcing dataset will be used for all experiments. This will be prepared by the members of Natasha MacBean's lab group and distributed by the CalLMIP planning committee to participating modeling groups. The dataset will include data for a fixed number of sites (see Section 2) that represent the diversity of PFTs worldwide. For each site, the following will be provided: i) gap-filled *in situ* meteorological forcing data (variables typically used in land surface models – see Section 3) from the flux tower for the entire calibration and validation period; ii) measured (i.e., "raw" or "uncorrected") daily observations of NEE, Qle, and Qh, with measurement errors, for calibration only (see Section 4); and iii) site meta-data including soil characteristics and vegetation type, canopy height?, forest age?.
- Calibration experiments (required): groups must run calibration Experiment 1 (i.e., using NEE, Qle and Qh together in the parameter calibration – see Section 1.1). (Note groups can also opt into additional Experiments 2 to 6) – see Section 1.1). However, the method for combining all 3 datastreams into the same calibration experiment (e.g., one at a time/stepwise or all-together/simultaneous) is left to the group's discretion.
- Spin-up protocol: each group must ensure that models start from a stable, physically realistic state, preventing biases from arbitrary initial conditions; therefore, a consistent spin-up protocol will be described in this protocol (see Section 7) (unless the group has a set functionality to prescribe initial conditions/states).
- Posterior uncertainties on parameters and variables optimized will be required.
- Output file format will be standardized across models (see Section 6).
- Calibration/validation setup: a set number of sites will be set aside for spatial validation and 1 year for each site for temporal validation (Section 8).
- Measurement errors will be specified for the flux data.
- All methods and choices must be thoroughly documented by all groups (see Section 10 for information required).

1.3 Elements that will be defined by each modeling group (flexible)

While the core structure of the CalLMIP experiments will be standardized (see Section 1.2), modeling groups will have flexibility in defining certain methodological aspects:

- Model version: each group will select which version of their model they want to submit to CalLMIP. This includes choices such as whether LAI is prescribed or treated as a prognostic variable, and how to translate vegetation and soil characteristic information from each site to fit their model. If a group runs a variant of a land surface model that contributed to CMIP6, and can run a CMIP6-like configuration, and/or if they have the version for CMIP7-FT, we ask that one or both of these versions be prioritized.

- Parameter selection: each group will decide which and how many parameters it wants to optimize. Parameters may vary spatially (e.g., by PFT) but should be time-invariant. The modeling groups can choose their method for parameter selection (e.g., expert elicitation, sensitivity analysis). Furthermore, each group will determine what prior parameter bounds and uncertainties to use for their parameter estimation experiments.

- Calibration experiments (optional): groups can opt into running the additional, option Experiments 2 to 6 (see Section 1.1).

- The method for multi-datastream calibration (e.g., one at a time/stepwise or all-together/simultaneous) is left to the group's discretion.

- Parameter estimation method: each group will select its own method for calibration (e.g., gradient-based optimization, MCMC approaches, or machine learning techniques).

- Number of PFTs to be calibrated for each site: each group will decide whether they calibrate the parameters for all PFTs at a given site or only for the most dominant PFT, depending on model configuration and parameter calibration system.

- Cross/multi-site parameter approach: parameter optimization must result in a set of "operational" parameters (e.g., for running the model globally), rather than site-specific parameter sets, although PFT-specific and soil-specific parameters, etc, are allowed. The specific method for this implementation is left to each group's discretion (e.g., hierarchical Bayesian models, individual site calibrations followed by averaging, multi-site optimizations for each PFT, parameter ensembles etc).

- Model errors are specific to each model and need to be estimated by each group and combined with measurement errors provided, as appropriate for the parameter calibration method.

2. Site selection

We have selected 22 sites from the 42 sites originally selected from the PLUMBER2 list of sites. Below is a description of the site selection process and the sites selected.

However, **PLEASE NOTE that we are planning to re-do the site selection process to add more sites with higher global coverage once FLUXNET 2025 is available (AGU, December 2025).** We will follow the same site selection criteria as outlined below, but we are expecting a much wider range of sites will be available: a) because of increased record lengths; and b) because all sites will have uncertainties through the harmonized flux data processing that all networks with FLUXNET are working on for FLUXNET 2025. Therefore, we are proceeding with Phase1a, which is a test calibration at 1 site selected from the 22 PLUMBER 2 sites selected below. We will then update the sites selected in January/February 2026 before finalising the list of sites for the full Phase 1b multisite experiment, which will start in March 2026.

The initial 42 sites from PLUMBER2 were chosen by Gab Abramowitz and colleagues for their own model development pipeline. The main criteria for selecting these sites were: (a) length of record (meaning model spin-up issues were less likely to interfere with evaluation), and (b) data quality (less gap-filling of key driving variables - precipitation, temp, radiation, humidity, wind) (G. Abramowitz, personal communication, March 26th, 2025). We added the following additional criteria for site selection within CalLMIP:

1. Minimum a four year record length. Within these four years there had to be at least three years worth of observation (specifically, NEE) half hourly (or hourly) data after taking into account missing and gap-filled data. Only gap-filled data from NEE, Qle, and Qh was assessed to determine if there was an appropriate amount of data.
2. Netcdf files needed to include uncertainties for NEE, Qle, and Qh. In the PLUMBER2 archive, only netcdf files using FluxNet 2015 contained uncertainties. Files processed using OzFlux provided no uncertainties while LaThuille only provided uncertainties for NEE. Thus, for the 22 selected sites there are no [Note: OzFlux uncertainties are provided in FLUXNET2015. This would give us 2-4 extra sites (Au-Tum, Au-How and maybe Au-ASM and AU-Stp if there's enough data). If we include crop and wetland sites that would give us an additional 6 crop sites and 2 additional wetland sites].

Some Fluxnet files were further excluded because Qg (ground heat flux) was not measured. Though Qg is not a mandatory measurement within the FluxNet processing pipelines Qg is essential for calculations of energy balance closure. EBC allowed for the calculation of an energy balance correction factor that was then used in an equation to calculate the joint uncertainty.
After these two rounds of filtering there were 22 sites passed to the third level of screening.

Once sites were confirmed to have minimum 3 years worth of data + uncertainties, they were examined for information regarding plant functional types.
The results of examining sites PFT information are located within the PLUMBER2 Site information tab.

PFT information within the netcdf's were spotty and to get a percentage of plant fractional types consulting external sources was often required.
Sources included the PLUMBER2 netcdf, Shi et al. 2024 dataset, collaboration with Vincent Tartaglione and Cédric Bacour, and reliable internet sources
Shi et al. (2024):
https://essd.copernicus.org/articles/17/117/2025/essd-17-117-2025.html and https://zenodo.org/records/12596218
Vincent Tartaglione and Cedric Bacour provided an excel document listing the ORCHIDEE PFT information for each site based on an extensive analysis they have conducted on PFT information for fluxnet sites (as yet unpublished).
Internet sources were either papers or flux tower information sites
To ensure that a given source had correct PFT information I used a ⅔ rule for my sources, where at minimum different 2 sources had to provide similar PFT information. Note that for some of the SW US sites (US-SRG, US-SRM, US-Whs and US-Wkg) NM thinks that all the PFT descriptions are slightly off! This is probably true for all temperate or tropical mixed shrub grass ecosystems (and the C3 vs C4 split is also likely to be wrong in the SW US). NM will update the PFT table to reflect numbers from her paper in the coming days.

Information on soil composition was collected from netcdf's when possible and compiled into a list within the PLUMBER2 site information tab. Availability of soil data was not a considering factor whether a site was considered eligible.

The sites are:

AU-Tum (EBF cool, 16 yrs); AU-How (WSA tropical, 15 yrs); AU-Cum (EBF, 5 yrs); AU-ASM (ENF, hot, 7 yrs); AU-GWW (SAV, 5 yrs); AU-Ctr (EBF tropical, 8 yrs); AU-Stp (GRA, 8 yrs); BR-Sa3 (EBF tropical, 3 yrs); CA-Qfo (ENF cold/snow, 7 yrs); CH-Dav (ENF, 18 yrs); CN-Din (EBF, 3 yrs); CN-Cha (MF, 3 yrs); DE-Geb (CRO, 14 yrs); DE-Gri (GRA, 11 yrs); DE-Hai (DBF, 13 yrs); DE-Tha (ENF, 17 yrs); DK-Sor (DBF, 18 yrs); FI-Hyy (ENF cold/snow, 19 yrs); FR-Gri (CVM, 9 yrs); FR-Pue (EBF, 15 yrs); GF-Guy (EBF tropical, 11 yrs); IT-Lav (ENF, 10 yrs); IT-MBo (GRA, 10 yrs); IT-Noe (CSH, 11 yrs); NL-Loo (ENF, 17 yrs); RU-Fyo (ENF, 12 yrs); US-Blo (ENF, 7 yrs); US-GLE (ENF - cold/snow, 6 yrs); US-Ha1 (MF, 15 yrs); US-Me2 (ENF, 13 yrs); US-MMS (DBF, 16 yrs); US-Myb (WET, 4 yrs); US-NR1 (ENF, 16 yrs); US-PFa (MF, 20 yrs); US-FPe (GRA, 7 yrs); US-SRM (WSA, 11 yrs); US-SRG (GRA, 6 yrs); US-Ton (WSA, 14 yrs); US-UMB (DBF, 15 yrs); US-Var (GRA, 14 yrs); US-Whs (OSH hot, 7 yrs); US-Wkg (GRA, 10 yrs).

Breakdown of IGBP vegetation types:

ENF: 11, EBF: 7, GRA: 7, DBF: 4, WSA: 3, MF: 3, SAV: 1, CRO: 1, CVM: 1, CSH: 1, WET: 1, OSH: 1 (see here for IGBP classification: https://fluxnet.org/data/badm-data-templates/igbp-classification/).

Breakdown of countries:
Australia (7), Brazil (1), Canada (1), Switzerland (1), China (2), Germany (4), Denmark (1), Finland (1), France (2), French Guiana (1), Italy (3), Netherlands (1), Russia (1), USA (16)

3. Forcing Datasets

Standardized in situ flux tower meteorological forcing data will be used to ensure consistency across sites. The forcing datasets will include LWdown, SWdown, Tair, Qair, Rainf, Wind and PSurf for a number of flux tower sites. We will also provide the NCO code needed to convert Rainf into Rain and Snow in the modelevaluation.org information for those modelers who can use/need both variables. Atmospheric CO2 forcing data will also be provided.

The sites used in the first phase of CalLMIP will be based on the PLUMBER2 dataset. PLUMBER2 provides pre-selected sites from FLUXNET2015 (CC-BY-4.0 licensed sites), FLUXNET La Thuile Free-Fair-Use subset and OzFlux data, that meet quality control criteria and have minimal data-sharing restrictions. PLUMBER2 quality control steps ensure that all sites have complete metadata, including canopy height and IGBP vegetation type, and contain full years of data with minimal missing values for key variables. Furthermore, standardized gap-filling methods (e.g., LWdown from Abramowitz et al. (2012), PSurf from elevation and temperature) and FLUXNET2015-aligned energy balance corrections have been applied to these sites as part of the PLUMBER2 processing steps.

From the 170 PLUMBER2 sites, a subset (~40 sites) will be chosen by the CalLMIP planning committee for use in CalLMIP experiments. These will be chosen to maximize diversity of PFTs worldwide and ensure at least X full years of data.

The same forcing files in NetCDF format will be provided to all groups for all selected sites. Individual modeling groups will then need to modify the files to fit their model (for example, changing the variable names, adjusting or renaming the units and other attribute information, separating snow and rain etc).

4. Optimization Datasets

Optimization data will include standard *in situ* flux tower data for optimization: NEE, Qle and Qh. Datasets will be provided at daily frequency for the selected sites for the calibration period only. One year of data will be retained for temporal validation.

Only measured flux data (i.e. non gap-filled, no energy balance closure corrected data) will be provided. Days that have required more than 25% gap-filling will be set to NaNs.

Daily measurement errors will be provided to all groups based on FLUXNET processing (Pastorello et al., 2017). *These need to be combined with modelling errors for the observation error covariance matrix in certain DA methods (e.g. variational)*. Each modeling group will decide on their own method for estimating model error.

The same observation flux files in NetCDF format will be provided to all groups for all selected sites. Individual modeling groups will then need to modify the files to work with their model/calibration system (for example, changing the variable names, adjusting or renaming the units and other attribute information).

5. Required Outputs

i. Each group will conduct two simulations per site for each experiment they perform (note: 1 experiment is required – see Section 1.2 and 1.3) to ensure a consistent evaluation of model performance before and after calibration:

   a. **prior simulation**, representing an out-of-the-box simulation using default model parameters in the model version being used at each site (i.e. no prior parameter tuning or updating of the model to fit the data). This will serve as a baseline to assess the impact of calibration.
   b. **posterior simulation**, using the optimized parameters and reflecting the improvements achieved through the calibration process.

ii. All groups must output the following variables (per grid cell, not per PFT, age cohort or otherwise): NEE, Qle, H, GPP, Reco, Transpiration, Bare Soil Evaporation, Ground Heat Flux, Total Column Soil Moisture, LAI, Aboveground Biomass, Total Soil Carbon, land surface temperature, ???

iii. All groups must provide posterior uncertainties (or ensemble simulations) on variables optimised within the model output file.

iv. Model groups do not need to provide posterior parameter values or their uncertainties. However, we ask that modeling groups retain this data in case we (as a community) decide that this information could be useful for further analysis or discussion.

v. All groups must output files according to output file specifications listed in Section 6.

vi. All groups must follow the spin-up and transient protocol prior to performing the parameter calibrations (see Section 7).

vii. Each group must also provide detailed metadata alongside their results. This information will be input directly into a draft paper documenting the 1st phase protocol. Information required listed in Section 10.

Optional:

- Groups using ensemble methods can submit prior and posterior ensemble simulations; however, this is not required for groups not using ensemble methods. The choice of how many ensembles to include, and the method for selecting those ensembles, is left to each model group.

- Groups may perform the experiments with different model versions. In that case, we ask model groups to submit all required outputs for each model version and follow the file naming specifications listed in Section 6.

6. Output file specifications

Model results must be output at daily frequency and stored in NetCDF files, using standard **CMIP** variable names, and CF-compliant metadata, ensuring consistency in variable naming, units, and documentation. A list of standard names is provided here (note, for CMIP variable names see column 2, not column 1): https://modelevaluation.org/variableStandards.

All output variables (optimized and additional) will be stored in the same output file, including uncertainties on optimized variables.

A standardized naming system for output file names will be required: [*model_name*].[*model_version*]_Expt[*CalLMIP_expt_no.*]_[*SiteName*]_[*Cal/Val*]_[*Prior/Posterior for calibration or Temporal/Spatial for validation*].nc

E.g. the prior simulations for the DK-Sor site for the required experiment would be

7. Spin-up and transient protocol for the parameter optimizations

Each modeling group must ensure that a full spin-up is performed that is appropriate for the variables they are optimizing. We will follow the PLUMBER2 spin-up protocol. Models will cycle over the flux tower meteorological forcing files until equilibrium has been reached in the model states of interest (e.g., soil C or soil moisture). For models focusing on carbon fluxes, a full spin-up to equilibrium carbon stocks (and N and P) will be required (likely 100s to 1000s of years depending on model structure). For models that are only focused on hydrological or energy balance variables and have prescribed LAI, a shorter spin-up will be sufficient (likely at least 10 years of simulation).

For carbon stocks we will spin up to equilibrium representing the year 1850 using atmospheric CO_2 concentration and N deposition levels of 285 ppm and 0.79 kg N ha-1 yr-1, respectively. The transient simulation will cover the year 1851 to the first year of the forcing data. Climate forcing during this transient simulation will continue to cycle over the available flux tower site meteorological forcing files, but models should use historical changes in atmospheric CO_2 concentration and N deposition (files will be provided).

Note 1: If groups can and want to initialize their model runs by prescribing initial conditions and states they can do that.

Note 2: We will endeavour to find forest age/disturbance that may be required for example by vegetation demography/cohort models but this needs to be further discussed with relevant modeling groups.

8. Optimization validation at site-scale

Site-level temporal validation will be conducted for the optimized variables using the last year for each site. Spatial validation will be conducted by excluding X sites from the calibration. Meteorological forcing files will be provided for the validation period and additional sites, but no flux data will be provided. Model outputs and output file specifications must be the same for the validation runs as for the prior and posterior simulations.

Note 1: Additional evaluation of variables not provided for optimization (as opposed to an assessment of reduction in inter-model spread) will depend on data availability at the selected sites. If CalLMIP participants know of in situ data on model states (e.g., soil *moisture, land surface temperature, aboveground biomass or soil C stocks) that is readily available, please let us know.*

9. Post-calibration global simulations.

Two other potential sets of runs (with additional analyses) have been proposed for the CalLMIP 1st stage. An outline of what these additional runs would require and the protocol for doing so is provided below. During the 3rd session of the workshop we will collectively discuss whether to do either or both of these additional runs as well as the specific details.

Post-optimization global runs over the last two decades

Global scale runs (for both prior default parameters and posterior optimized parameters) would enable us to evaluate large-scale model performance. Similarly to the site-level experiments, the input data for the global simulations, such as climate forcing, atmospheric CO2, land use/cover change, N deposition etc will be the same across models and will follow the TRENDY protocol – i.e. we will use the TRENDY input data and the same spin-up and transient protocol unless the full TRENDY spin-up is not required). These input data would also be prepared by the MacBean lab and distributed by the CalLMIP planning committee. The outputs will be standardized across models including the variables required (likely will include model states and not just fluxes), time period (likely at least the last 20 years), frequency of output (likely monthly and/or annual), and the spatial resolution (0.5 degrees, 1 degree or 2 degrees). These will be decided upon by CalLMIP participants during the third session to enable as many groups to participate as possible.

The datasets and evaluation/benchmarking tools that will be used for evaluation/benchmarking of global scale simulations will be decided at a later date. Options include satellite based products and machine learning based upscaled products. Options for benchmarking and analysis include iLAMB, Ecological Forecasting Initiative (EFI) computational infrastructure, ESMValTool.

Post-optimization site-based future simulations up to 2100

Future simulations for both prior and posterior parameter sets would enable us to have an assessment of how parameter optimization is affecting our future predictions, albeit not at global scale. For these runs we would provide offline climate forcing from the recent CMIP6 ESM outputs (to be decided) from the pre-industrial to 2100, allowing a full transient run (with changing climate, CO_2, LUC) to be performed. Land use change, future CO_2 concentrations and N deposition data would be provided (to be decided).

Note 1: We note that both sets of additional simulations require different forcing data to what will be used for the optimizations. We therefore will need to be careful not to discuss model "improvement" per se as biases/uncertainties due to parameters calibrated to different forcing may propagate in unknown ways to global scale or future simulations. What we can assess is the difference between prior and posterior and how that compares to benchmark data (for global runs) or how that changes future predictions.

Note 2: We caution that global runs and new future simulations when model states have not been optimized is problematic and needs to be addressed in the optimization and/or considered when performing the additional runs/analyses.

Note 3: There is potential to compare the calibrated outputs with empirical models, similar to PLUMBER2. This will be discussed further with Gab Abramowitz.

10. Model/experiment information table

The following information will be required from each modeling group. This information will be input directly into a CalLMIP 1st Phase protocol paper that we aim to publish in GMD. Information required includes:

To be input from draft paper…

11. Additional Information

11.1 Participants

Model	Who	Institution	Email	Which variables do you plan to optimize?
CABLE-POP	Matthias Cuntz	INRAE, France	matthias.cuntz @inrae.fr	NEE, LE
ORCHIDEE	S. Beylat V. Tartaglione V. Bastrikov C. Bacour P. Peylin L. Olivera	LSCE CEA/IPSL, France	orchidas-callmi p@lsce.ipsl.fr	
GAIAFLUX	Gerbrand Koren	Utrecht University, Netherlands	g.b.koren@uu. nl	NEE, GPP

11.2 Potential Additional Experiments

Additional experiments	Participant(s)
Comparing parameter optimization with site forcing versus climate reanalyses	??
Compare the same optimisation method across different models	ORCHIDEE (2-3 different versions of the model)
Compare different optimization methods across the same model	CLM, ORCHIDEE (to be determined for ORCHIDEE depending on the differences obtained)

Include state variables in optimization (e.g., soil C, soilm)	CLM

12. Open Data Policy

All outputs from CalLMIP parameter calibration experiments will be deposited in a publicly available repository (such as …). These model outputs will be freely available for public use with an authorship and acknowledgement policy. Users of these model outputs will be required to reach out to the individuals and modeling groups who performed the experiments to offer co-authorship within X years of the model outputs being deposited in the repository. Users will also be required to cite the CalLMIP 1st Phase paper (citation will be provided).

13. CalLMIP co-authorship policy

Any individual who contributes to performing parameter calibration experiments, and/or who contributes in an intellectual capacity to the design of the experiments and analyses, and/or who advises individuals carrying out parameter calibration experiments, and/or who documents the experiment information will be included as a co-author on the paper resulting from the CalLMIP 1st Phase experiments. Co-authors will also be expected to provide feedback on the structure and framing of the paper as well as on draft(s) of the paper.

14. Future vision for CalLMIP

Tiered Experiments (Nancy Kiang)

1. Canopy structure: given observed canopy height(s), LAI, and demography, calibrate allometry parameters for biomass. Non-demographic DGVMs would simplify the canopy accordingly, aim to calculate correct biomass.

2. Canopy biophysics: given Canopy Structure, seasonal LAI, initial soil carbon, and meteorological forcings (including soil moisture), calibrate leaf specific area, leaf biophysics, autotrophic respiration, and soil respiration parameters for NEE and transpiration.
3. Other optional things to calibrate with biophysics: leaf optical properties to get APAR, albedo.
4. Land surface hydrology: Given (1) and (2) and precipitation, calibrate soil physics parameters for soil moisture.
5. Phenology: calibrate parameters for seasonal LAI timing.
    a. For some models, this may be only a function of meteorology and soil moisture, independent of biophysics simulation.
    b. For others, stress (cold, water) and the carbon balance might also influence the phenology.   So calibration would be, given (1)-(4), for both timing of LAI and fluxes.
6. Carbon allocation: given (1)-(5) calibrate allocation to stem (height and dbh) growth, storage, leaf growth, senescence, retranslocation to get LAI, stem growth, litterfall/soil carbon.
7. I guess we’re not doing disturbance or other ecological dynamics.