Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-13-AC2, 2016

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "A new step-wise Carbon Cycle Data Assimilation System using multiple data streams to constrain the simulated land surface carbon cycle" *by* P. Peylin et al.

P. Peylin et al.

peylin@lsce.ipsl.fr

Received and published: 12 June 2016

Peylin et al. contribute an interesting study on the effect of using a stepwise optimisation rather than merging all data streams in a single cost function. This is a procedure that we have used in e.g. catchment scale water quality modelling where one first calibrates the water cycle before calibrating parameters relevant to nutrient diffusion (e.g. Exbrayat et al., 2011). However, this approach has not been investigated in details in the frame of the (global) carbon cycle. Therefore, I agree with reviewer #1 that this paper is highly relevant to the community. I particularly like the several steps used by the authors to reconcile site-scale calibration with global atmospheric concentrations.

I have some very minor comments on the paper that should be straightforward to ad-

Printer-friendly version

Discussion paper

dress:

We thank the reviewer for his positive evaluation of the paper.

p.5 l.1: another alternative to stepwise optimisation and simultaneous procedure is a simultaneous, multi-objective approach based on the Pareto ranking of several cost functions to account for trade-offs (e.g. Yapo et al., 1998). Would it be realistic to use such an approach in this system to avoid the increase in RMSE against MODIS NDVI from step1 to step2 (fig. 8, TeBD)?

The suggestion of using a simultaneous multi-objective approach based on the Pareto ranking of several cost functions to account for trade-offs as in Yapo et al. (1998) could indeed appear as an alternative approach. However, it is based on a random generation of parameter sets in order to find the so-called "Pareto parameter space" where all parameter sets lead to equal overall objective function but with trade-offs between individual cost functions. With the global carbon cycle optimization problem and the ORCHIDEE model the approach in Yapo et al. (1998) is from a computational point of view not feasible at all. For a specific hydrology problem with 13 parameters to be optimized and two cost functions the proposed algorithm required more than 68000 function evaluations to converge to a solution. Given that our total number of parameters is on the order of 100 and that we can not afford a large number of global simulation for the evaluation of the cost function linked to atmospheric CO_2 data (few tens to a hundred), such method is thus clearly not suitable.

Finally we should mention that the increase in RMSE against MODIS NDVI from step1 to step2 is relatively small compared to the initial improvement of the RMSE during step1 optimization.

p.5 l.12-13: Using a restricted number of parameters is a valid point but it needs to be mentioned here that one must proceed to some sort of sensitivity analysis to accurately select these parameters.

Printer-friendly version

Discussion paper

We agree that the selection of parameters should follow a rigorous sensitivity analysis, using for example the MORRIS algorithm (global sensitivity analysis; Morris, 1991). In our case, we have done such sensitivity analysis in previous studies with the ORCHIDEE model; we have thus kept the same set of parameters that was identified in these studies. We have improved the text adding: "(following a global sensitivity analysis)".

p.8 l.1: why not using only days with data?

We agree with the reviewer that we could have used only the days with data. However, for practical implementation it was easier to interpolate the time series that have sparse data. We have checked however that using this interpolation does not change the results of the optimization.

p.9 l.26: please mention the resolution of the model here

We added the resolution of the LMDz model.

p. 10 l.1: is it robust to assume that carbon pools are at equilibrium in 1990? Could this system use a prior from soil and biomass maps instead (like e.g. Bloom et al., 2016)?

The choice to bring the model at equilibrium for the carbon pools in 1990 is a compromise. We do not think that our optimization system could easily use soil map (such as the HWSD data set) or forest biomass estimates (such as the Saachi et al (2011) map for the Tropics).

For the soil carbon content, it is not straightforward to use the HWSD map to force a global model like ORCHIDEE. First, with the "CENTURY" soil carbon model used in ORCHIDEE, the turn over time of the soil organic matter (for each reservoirs) together with the rate of organic matter input to the soil (litter and root turnover) determine a total soil carbon content that is in balance with all components of the model. It is thus difficult to optimize the soil carbon content with global estimates such as HWSD, while keeping the internal model consistency. One way would be to optimize the parameters controlling the turnover times as well as the soil carbon input over a long spin up period (several 1000 years); this is currently not feasible with the optimization of several parameters, especially without the adjoint of the model. Second the model does not represent yet high soil carbon content such as peat land or permafrost, while these soil type are usually taken into account in the observations. Finally, the HWSD soil C map corresponds primarily to the carbon content from 0 to 1 meter of soil; we would need first to adjust the observation so that it matches the total soil carbon content that is modeled. Overall, it is a rather complex process to optimize the soil carbon content in the case of ORCHIDEE if we want to keep the internal coherence of the model to improve its predictive skill. Ongoing works are occurring at LSCE to determine the best approach to assimilate soil carbon content observations or potentially turnover time of soil organic matter (derived from observations) into ORCHIDEE.

Similar issues also pertain to the assimilation of forest biomass maps (such as Saatchi et al. 2011). We cannot change easily the above ground biomass in ORCHIDEE without violating the overall carbon allocation scheme and the internal consistency of the model (i.e., the ratio of carbon content between the different reservoirs). We thus would need to perform an optimization of the parameters controlling carbon allocation as well as the input of carbon over a period corresponding at least to the age of the forest. This is a difficult task that was investigated in Thum et al. (in revision) for site scale observations.

The study of Bloom et al. (2016) is slightly different in the sense that the model is less complex than ORCHIDEE and that they could do an optimization over long time period, as needed for carbon stock observations. However, we acknowledge that using carbon stock observations is the next challenge as already mentioned it in the discussion section. For the soil carbon content we thus choose a compromise which is to bring the model to equilibrium in 1990, perform a transient simulation over 10 years and then optimize an ensemble of coefficients that scale (either at site level or globally for several regions) the initial soil carbon content (the slow and passive pools) in order to account for the all past effect that led to deviation from the “equilibrium assumption”.

p. 18 l.6: How is fire simulated during spin-up?

During spin-up the ORCHIDEE model is run with fires provided by a “generic” fire emission module that is thus not based on the satellite observation of fire occurrence and fire extension. Such simplification is necessary, as we do know have the past history of fire occurrence to guide the spin-up procedure. We have précised in the text that the inclusion of GFEDv3 fire emission is not used for the model spin-up.

p.26 l.29: perhaps "ecosystem data streams" is more correct (LE is not carbon sensu stricto)

Indeed LE is not a carbon cycle data stream but it is related to the carbon cycle through the stomatal conductance at the leaf level. In order to keep the word “carbon” we have change the expression to “carbon cycle related data streams”

p. 29 l.17: see also Bloom et al. (2016)

We added the reference “Bloom et al. (2016)”

Figure 1: lower box, "carbon"

Corrected

Figure 5: please add scale on x-axis

We added the scale on x-axis to mention the month of the year.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-13, 2016.