MS number: gmd-2013-93

Reviewer #1

We appreciate this referee for encouraging comments. Our responses are as follows.

1. First of all, the authors did not show the way to change the parameter numbers in iterative optimization process. Do you deterministically estimate the next step parameter numbers by tangent linear algorithm like TAF for all the parameters at once over multidimensional parameter field? or Do you move the parameter number with prescribed interval for each parameter and move its number toward the direction where you get smaller misfit function (calculate S(m)/p x 15 PFTs x 13 parameters for each iteration)?

   In the minimization of cost function, modeled values $\mathbf{G}(\mathbf{m})$ for CO_2 is determined by first derivatives of net ecosystem CO_2 exchange rate from the biosphere model $\mathbf{B}$ for finite differences of parameters Δ $\mathbf{m}$,

$$\mathbf{G} = \mathbf{A}\left[\mathbf{B}(\mathbf{m}_\mathrm{p}) + \frac{\mathbf{B}(\mathbf{m}_\mathrm{p} + \Delta\mathbf{m}) - \mathbf{B}(\mathbf{m}_\mathrm{p})}{\Delta\mathbf{m}}\right] \tag{A1}$$

   where $\mathbf{A}$ is atmospheric transport model described as a linear function and $\mathbf{m}_\mathrm{p}$ is biospheric model parameters. Regarding NPP and AGB, $\mathbf{G}$ is determined by Equation A1 without multiplication of $\mathbf{A}$. Optimum model parameters $\mathbf{m}$ in an iteration are determined using singular value decomposition. In the calculation, derivatives from model parameters $\mathbf{m}_\mathrm{p}$ are output as $\mathbf{m}$ to minimize the cost function. $\mathbf{m}_\mathrm{p}$ are updated in the next iteration by adding these derivatives (i.e. $\mathbf{m}_\mathrm{p}$+$\mathbf{m}$), and CO_2, NPP and AGB are newly estimated, then cost function is minimized with Equation A1. Since the biosphere model is nonlinear, optimum parameters are estimated by repeating this process with Δ $\mathbf{m}$ little by little.

2. How you calculate the Chi2 for three different data streams is poorly explained. I understand as you calculate Chi2 differently for each three different data streams with parameter, then sum them up. Also what is the difference between Chi2 and Eqn 4? Or do you calculate the Chi2 for CO2+Chi2 for AGB+Chi2 for NPP+Chi2 for params at once? Which is correct? You should also make the list of Chi2 value for each term in both prior and posterior status. Moreover, you should mention why you can use the standard deviations in the calculation of the 7.5° grid mean values for AGB and NPP as the uncertainties for AGB and NPP (denoted as CD). How to set them is quite important information determining the relative influence by each term on reducing the misfit function.

   We calculated χ^2 for CO_2, NPP and AGB simultaneously as,

$$\chi^2 = \sum_{i=0}^{N_\mathrm{obs}} (G_i - d_{\mathrm{obs},i})^2 / C_{D,i} + \sum_{j=0}^{N_p} (m_j - m_{p,j})^2 / C_{M,j} \tag{A2}$$

where N_{obs} is number of observations, G is modeled values for CO_2, NPP and AGB, d_{obs} is observations, m is optimum model parameters, m_p is model parameters, and C_D and C_M are uncertainties for d_{obs} and m_p, respectively.

I am revising the manuscript on $\mathbf{C}_{\mathrm{D}}$.

3. Relative influence by each parameter and by each data steam on reducing the mismatch is unclear. You compared the fractional shift in parameter value (1 - P_post/P_prior) in Fig 2. It does not allow us to understand on which parameter has reduced the mismatch between modeling and observation the most itself. I first recommend you to show the relative change in parameter numbers based on parameter uncertainty ((P_prior-P_post)/P_prior_uncertainty) instead of current fractional shift, 1 - P_post/P_prior. Also if possible, youd better show us the relative reduction in parameter uncertainty, 1-sigma_post/sigma_prior, where sigma is a standard deviation of the respective parameter uncertainty before or after assimilation, which comes from 2nd derivative of mistfit function, and which may not be able to be calculated by your optimization method, I guess. I am also very interested in the case if you assimilate either of atmospheric CO2 or of biometric data: AGB and NPP. Will single data assimilation improve other data streams? Which data stream is more influential separately on annual NEP, CO2 seasonality, and physiological parameter shift?

   I agree with your suggestion. I am revising the manuscript following your advice.

4. Reliability of three data streams should be discussed. Atmos CO2 concentration seems to have relatively minor uncertainty due to normalized sampling methods. But, other AGB and NPP may have large uncertainty on their accuracies. Both are from satellite data, which potentially suffer from changeable sampling accuracy depending on satellite angle (too low angle in high latitudes) and surface optical condition (Cloudy condition in Tropical regions). GPPDI AGB data also must have suffered from ununiformed field sampling methods. More than that, I like to know if it was good idea to incorporate three different items for improving the simulation. I guess that only assimilation of atmos CO2 conc is enough to simulate well against AGB and NPP simulation. Because number of data points is much larger in atmos CO2 than in other two, usually large Chi2 for atmos CO2 will prevent the optimization of VISIT to match with AGB and NPP observations. Of course, it really depends on how you set the data uncertainty in misfit function. At least, anyway, for seasonality, the annual mean values of AGB and NPP will not affect that much.

   As you mentioned, I applied NPP and AGB data with large uncertainties for VISIT optimization. This is to prevent simulating unrealistic values of carbon pools when optimizing VISIT using only CO_2. Because atmospheric CO_2 is explained by only net ecosystem CO_2 exchange rate that is difference between ecosystem respiration and gross primary productivity, which does not ensure accuracy of other variables such as carbon pools in the model simulation. But it is right that NPP and AGB with large uncertainties do not constrain the optimization scheme in this study. I will add more analysis on this point in the manuscript.

5. This time, the authors did optimize only the physiological parameters. But, I guess that phenology and soil water physics-related parameters are also very important for ecosystem modeling. So, the remaining discrepancy would be improved by further parameter optimization. Discuss them.

   I will add discussion on phenology and soil water physics.

6. Totally, the paper is to be improved its explanation on parameter optimization scheme and misfit estimation scheme, and the expression on the improvement of parameter uncertainty and misfit between modeling and observation.

   I agree with your suggestion. I will revise the documents especially on optimization scheme and its results.

7. Text: This paper should apply the present tense on every sentence because this is the modeling research not the field measurement.

   All sentences are changed to present tense.

8. Page 9, Line 23: First of all, I like to know if the authors assimilated three observations simultaneously or separately. If you did simultaneously, the formulation should have two more terms for NPP and AGB.

   All parameters were optimized simultaneously for CO_2, NPP and AGB. Please have a look at reply to comment #2.

9. Page 10, Line 6-7: If you assimilate for three variables, you have to have three terms for variables + one term for parameter. Why dont you have them? i.e. S(m) = 1/2(Chi2 for CO2 + Chi2 for AGB + Chi2 for NPP + Chi2 for parameters)

   As commented above, parameters were optimized simultaneously for three variables, so that cost function used is written by Equation 13 in the manuscript. In this calculation, for example, $\mathbf{d}_{\mathrm{obs}}$ is comprised of observations for CO_2, NPP and AGB.

10. Page 10, Line 10-11: How did you determine the criteria for CM? Is there any proper observation or literature for them (This study fixed CM at 10% around each mp, and at 2 oC for Topt and Tmin.)?

    Values in CM are key to determine optimum parameters, but unfortunately there are no reliable observations on parameters being optimized in this study. I recognize your suggestion and will reconstruct $\mathbf{C}_{\mathrm{M}}$ on revised manuscript.

11. Page 10, Line 20-21: You have to tell us how to move the parameter numbers in iteration. You move the number by the prescribed small interval to make the slope of S(m)/parameter vector, and do that for each parameter again and again to reach

the sufficiently small Cai2. Am I right? Also you have to tell how many iterations were done, and what is the criteria to stop the iteration.

Parameters were moved little by little by updating parameters using the results of minimization, see reply to comment #1. I am writing details of iteration.

12. Page 12, Line 12-14: Why did you use them as uncertainties? Its the standard deviation in spatial distribution but the uncertainty on estimation accuracy.

    There is no information on uncertainty of NPP and AGB observations. Standard deviation was used as substitute for uncertainty. But I will replace this with other way following previous study.

13. Page 12, Line 17, X2 = 9.80 for AGB and NPP: Are Chi2s same for both AGB and NPP, respectively? Or 9.8 is the sum of Chi2 for AGB and NPP?

    9.8 is χ^2 for both AGB and NPP. When calculated χ^2 for each variable, values of χ^2 differ.

14. Page 12, Line 17-19: Show us the prior Chi2 values. I know that prior Chi2 does not have much meanings cos it can be any number as long as you put arbitral numbers for prior parameters. But, like to know how the combination of AGB and NPP made change in Chi2.

    I will show it on revised manuscript.

15. Page 12, Line 20-21: Cant you estimate the relative contribution on reducing the Chi2 by each parameter? Current relative shift in parameter number actually shows how the prior and posterior parameters are different in terms of absolute number. But, we do not know yet the actual influence by each parameter for total simulation accuracy. I think that at least you have to show the relative change of parameters: (param_post- param_prior)/param_uncertainty_prior

    I've never tried it, but it is possible to estimate the relative contribution of each parameter to reduction of χ^2. This suggestion could improve our optimization scheme, and I will consider how to adopt this.

16. Page 18, Line 1: AGB.

    Fixed.

17. Page 19, Line 7-10: You can express the simultaneous shift in two parameters by calculating the covariance of relative change in parameter number, then show them up in 13x13 matrix table. These parameter co-shifts are very important.

    I appreciate this comment and will include this analysis in the revision.

18. Page 19, Line 18, NEP of 2.0 PgCyr1: I have a big concern about this number. How could the NEP be positive? Cos, the VISIT was spun-up for 2000 yrs with present cli- mate to get the C storage equilibrium. If the initial state of storage was fixed to produce neutral C flux, NEP cannot be shifted positively even the physiological parameters were modified. Another thing is that the size of NEP, which corresponds to residual net terrestrial C uptake of 2.6 PgCyr-1 in IPCC (2007), proves that the VISIT underestimates the NEP. You have to consider the harvest for Forest and Cropland, and other carbon flows as you mentioned about forest fire.

    This suggestion is right. This uptake may be caused by insufficient spin-up period. I will check the state of carbon fluxes during spin-up and revise the results if modification needs.

    I am not going to include the influence of harvest in this paper. Harvest or disturbance have significant impact on carbon pools and fluxes, and subsequent parameter optimization. It leads to completely different results for optimization of interval of harvest and disturbance and its intensity, not optimization of physiological parameters.

19. Page 20, Line 10: I admit that the authors succeeded to incorporate three data streams at once for optimizing the VISIT parameters. But, you never explained us how better you could simulate the ecosystem flux and storage and atmospheric CO2 concentration compared to the prior simulations. And you should tell us which data stream and which parameter is more influential to reduce the misfit between modeling and observation.

    I will revise the manuscript following this comment. But prior parameters are artificially given as an initial setting of model as described in the manuscript, and comparison of posterior with prior does not make sense.