

## ***Interactive comment on “Constraining a global ecosystem model with multi-site eddy-covariance data” by S. Kuppel et al.***

**S. Kuppel et al.**

sylvain.kuppel@lsce.ipsl.fr

Received and published: 9 August 2012

The authors would like to thank the Reviewer for the very helpful comments and suggestions. The comments have been taken into consideration in the revised manuscript. We answer them individually in the following.

**Kuppel and colleagues advance from previous approaches to optimize process-oriented terrestrial ecosystem models (TEMs) against eddy covariance data by calibrating sets of common parameters simultaneously for several sites. Most TEMs rely on the plant functional type (PFT) classes to prescribe parameter vectors ( $x$ ) that control functional responses of carbon and water fluxes to environmental drivers. However, once multiple sites of the same PFT are used indepen-**

C3123

**dently to estimate  $x$  it is common that these are often different. By performing multi-site (MS) optimizations, the current study circumvents the need to aggregate  $x$  or to regionalize  $x$  according to any factors other than PFT classifications and shows that the model performance does not decrement significantly to single-site (SS) optimizations. In some cases it even improves. It is then an approach that has significant advantages regarding practical applications in parameter regionalization and parametric uncertainty propagation. It also holds promise for deeper analysis of the PFT concept in TEMs as the most important parameter vector ‘covariate’.**

**The current manuscript is a robust piece of work that in my view only fails in clarifying and corroborating certain methodological options:**

**– By defining a set of common parameters the authors implicitly assert them as linked to the plant functional type. Since these range between vegetation to soil and energy balance parameters, the authors should clarify if this was purely a practical decision (because in TEMs parameters are prescribed per PFT). If so, discuss limitations given some soil water availability parameters would also vary between sites or the soil decomposition parameters could be widely considered constant.**

The choice of keeping the largest common set of parameters follows from the philosophy of a global ecosystem model, with common parameters for a given Plant Functional Type: the so-called ‘practical decision’ by the reviewer. Note that 12 PFTs are used to describe all ecosystems. Our overall objective is to improve the global model performances using the information from scarce in situ measurements. We thus choose to keep most parameters as generic and common to all sites, except the one that scales the initial soil carbon pool sizes. Such parameter directly reflects the history of land cover changes at each site which are very different between the selected sites, and

C3124

thus need to be optimized independently. We acknowledge that these assumptions are strong when using small-scale in situ observation, especially for soil water availability parameters ( $Dpu_{cste}$  and  $Hum_{cste}$ ), as they strongly vary across ecosystems and soil conditions. However, we kept the main 'global philosophy' in order to assess the potential of such approach for global simulation. In the next version of the model, we will reconsider this choice and render the soil water holding capacity as a function of soil type and soil classifications. The text in section 2.4 has been revised in that spirit:

*'The initial state of the model is optimized with the only parameter chosen as site-specific: a multiplier of the different soil carbon pool contents, which are closely related to the local land-use history (Carvalhais et al., 2008). The rest of the parameters are considered as generic across sites in this study. We acknowledge that this assumption brings some limitations given the potential inter-site variability of some parameters (e.g. soil water availability), which will be kept in mind in the results analysis.'*

**– The 'performed optimizations' are more than just MS vs SS, since the authors also explore the role of individual data streams (LE and NEE jointly and separately) and do heterotrophic respiration experiments – which are seen first only here. Some introduction/motivation should be also given previously about these experiments.**

The title of this section is indeed somewhat misleading. We changed it to *'Performed optimizations: single-site and multi-site'* and a full presentation of the different experiment has been integrated in the text in section 2.4:

*'Besides, two other data assimilation experiments are conducted for the purpose of this study. First, a series of multi-site optimizations are performed in order to evaluate the individual impact of the parameters related to heterotrophic respiration, notably regarding the initial soil carbon content (see section 3.2.1.). Second, we use the multi-*

C3125

*site procedure to separately assimilate each one of the data stream (NEE and LE), so as to evaluate their respective information content. In the case of LE, non-sensitive parameters were left out the optimization.*

*In the end, we performed the following optimizations, including sensitivity tests:*

- 12 reference SS optimizations (21 parameters for each)
- 1 reference MS optimization (20 generic parameters, 12 site-specific)
- 6 MS optimizations with different parameters related to heterotrophic respiration left out
- 1 MS optimization with only NEE data (20 generic parameters, 12 site-specific)
- 1 MS optimization with only LE data (14 generic parameters)'

**Regarding the evaluation of the different optimization exercises:**

**– A table on model performance for the different optimizations (MS versus SS) would be very helpful in synthesizing the current results. In this regard the current exercise is solely based on the RMS metric to evaluate the model performance at site level. Other metrics like correlation or model efficiency which translate changes in model behavior could also help understand if the improvements would also be paralleled by changes in model sensitivities, especially since for some sites the parameter differences between SS and MS is very significant.**

Before focusing on the RMS, we have tested several types of metrics in order to evaluate our optimizations. Among these, we think that the correlation did not add much to

C3126

the message, as we found that the results support what can be said using the RMS in terms of improvement from the prior and similarities/differences between SS and MS optimizations. Also, we preferred to emphasize the analysis of the model performance by separating the different time scales, while discussing the model efficiency would have lengthened the text without adding significantly new elements to the analysis.

**– Given the importance of KsoilC in the MS and SS optimizations it could be relevant to compare modeled estimates of soil C pools against site level observations at the sites [e.g. Schrumpp et al., 2011] as an independent data source that would corroborate the different approaches and the heterotrophic respiration experiments.**

This independent comparison would be indeed very useful. Unfortunately the data was not available at all the sites for a systematic comparison. Such in-depth analysis, although highly relevant, would require a substantial lengthening of the paper and is clearly beyond the scope of the current paper objectives. As mentioned in the conclusion we would like to achieve it in subsequent studies.

**– Since MODIS products also include LAI and FPAR products, along with NDVI and EVI, it would be good to justify why no direct comparisons between the same quantities (LAI or FPAR) were performed.**

LAI and FPAR products require an intermediate model, usually an empirical algorithm or an inverse radiative transfer scheme, to be derived from surface reflectances. While this intermediate step potentially adds uncertainties, it has been reported that significant divergences remain between LAI (and consequently FPAR) products coming from different instruments and reference ground-based measurements (Garrigues et al., 2008). The NDVI product, on the other hand, is directly calculated from the surface reflectances and would then remain truer to the measured signal. Moreover the LAI product is usually more subject to saturation problem for high values than NDVI.

C3127

Overall, we choose to compare only the seasonal variations of the satellite and model signal, using the correlation coefficient as a metric, and not to compare the raw fAPAR values. We added one sentence in the revised version to justify our choice :

*'Note that NDVI was preferred to other satellite products (such as LAI or FAPAR) as it is directly calculated from the surface reflectances, contrary to LAI and FAPAR which require intermediate models to be generate, thus possibly adding significant uncertainty to the retrieved data (Garrigues et al., 2008).'*

**Particular comments:**

**P3319, L14-17: Liu and Gupta [2007] refer the initial states as another source of uncertainties / mismatch between model and observations.**

Indeed, the initial state should be listed a distinct source of error. While we can reasonably consider that the water variables are annually reinitialized in temperate deciduous broadleaf forests, we acknowledge that the initial state of carbon-cycle variables plays a major role in the carbon dynamics of the modeled ecosystems. In the optimization procedure, we only take into account the initial soil carbon stocks via the scaling parameter  $K_{soilC}$  (Table 2 in the manuscript) applied to both slow and passive soil carbon pool in order to account for the local land-use history. The initial state of the biomass is currently taken from the spinup of the model, which assumes a mature forest (contrasting with the young stand age at some sites considered in this study) and does not take into account potential disruptive events (thinning, fire. . .). The optimization of the initial state of the biomass will be integrated in future studies (in preparation), as it is beyond the scope of this paper. We have mentioned the importance of initial state in the revised manuscript:

- in the introduction,

C3128

*'The discrepancies between the fluxes simulated by terrestrial biosphere models and the observations have five main origins: errors in flux measurements, errors in meteorological forcings, error in structural equations of the model (including missing processes), inadequate calibration of the model parameters, and erroneous initial state of the model. While the first two types of error are foreign to the biosphere model itself, the last three items are crucial to improve model simulations.'*

- in section 2.4,

*'The initial state of the model is optimized with the only parameter chosen as site-specific: a multiplier of the different soil carbon pool contents, which are closely related to the local land-use history (Carvalhais et al., 2008).'*

**P3323, L4-5: how much data was it actually used for all sites and individually per site?**

Over a total of 43 site-years used in this study, 196 days of NEE data and 2 of LE data were left out. Table 1 gives the detail for each site. We added a sentence in the revised version to summarize this information:

*'Note that individual days with more than 20% of missing half-hourly observations were not included in the assimilation; over a total of 43 site-years used in this study, this missing data represents 196 and 2 days in the NEE and LE data streams, respectively.'*

**P3323, L6-18: from the construction of  $R$  described in this section it is not clear how different are the observational and prior model error and what are the properties and distributions of the model errors used in  $R$  (e.g. is the error heteroscedastic? What the mean error is in NEE and LE to the cost function? Does it also vary by site?). It is also not clear the role of the factor  $k\sigma$  in the construction of  $R$ .**

C3129

**Table 1.** Length of the selected gap-filled data at each site, with the number of assimilated NEE and LE days.

Site	Period	Number of days	NEE days used	LE days used
DE-Hai	2000-2006	2555	2528	2555
DK-Sor	2004-2006	1095	1095	1095
FR-Fon	2006	365	365	365
FR-Hes	2001-2003	1095	1095	1095
JP-Tak	1999-2004	2190	2151	2190
UK-Ham	2004-2005	730	682	728
US-Bar	2004-2005	730	715	730
US-Ha1	2003-2006	1460	1452	1460
US-LPH	2003-2004	730	730	730
US-MOz	2005-2006	730	730	730
US-UMB	1999-2003	1825	1825	1825
US-WCr	1999-2004	2190	2131	2190

In this study, we assume Gaussian errors distributions both for parameters and observations (3323/23-24), thus for model errors as well. In addition, the Bayesian framework only deals with centered errors, thus our cost function does not include any systematic error. Also, for simplicity we assume a homoscedastic error both NEE and LE, and we define one daily uncertainty for each one of the two data streams.

Regarding NEE, although the model error is yet to be well characterized in TEM, it is commonly assumed that it could be on the order of 1-2 gC/m<sup>2</sup>/d, much larger (in terms of variance) than the random flux measurement error (0.4 gC/m<sup>2</sup>/d, Hollinger and Richardson, 2005). Further support in that direction was recently found by Kuppel et al. (submitted).

Finally, we chose to keep  $R$  diagonal because the model error structure is poorly known

C3130

at present. For want of a better calibration, for each type of data (NEE and LE) we base our estimation of the variance on the prior model-data mean square difference (MSD). Based on the 'inflated variances' method recommended by Chevallier (2007), the variances in  $\mathbf{R}$  are calculated by multiplying the MSD by the 'inflation factor'  $k_{\sigma,igma}$ , which represents our estimation of the characteristic autocorrelation time length of the model error (here, 30 days).

Also, for simplicity the following revised text has been moved to the section 2.3, which described the data assimilation system :

*'Regarding the observational error statistics, the error covariance matrix  $\mathbf{R}$  should include both the error on the measurements and the error on the model process representation. On the one hand, the random measurement error on the observed fluxes is known not to be constant and can be estimated as the residual of the gap-filling algorithm (Richardson et al., 2008). On the other hand, model errors are rather difficult to assess and may be much larger than the measurement error itself. Therefore, we chose to focus on the model error whose correlations cannot be neglected (Chevallier et al., 2006). The difficulty of evaluating the structure of the model error leads us to keep  $R$  diagonal and, as compensation, artificially inflate the variances (Chevallier, 2007). First, the variances in  $\mathbf{R}$  are defined as constant in time for each type of observation (NEE and LE). Second, their values are chosen based on the mean squared difference between the prior model and the observations, multiplied by the inflation factor  $k_{\sigma,igma}$ , which represents our estimation of the characteristic autocorrelation time length (in days) of the model error. The value of  $k_{\sigma,igma}$  is fixed to 30, which for the error propagation is equivalent to assimilating one observation every 30 days.'*

**P3325, L9-11: analogous to Carvalhais et al. [2008]**

C3131

This has been added to the text.

**P3327, L5-6: but performance statistics are computed on daily data?**

We smoothed the signal only for display purposes, performances are indeed computed with the daily averaged data. We are now more specific in the text:

*'Figure 2 shows the seasonal cycles of NEE and LE at two of the twelve sites used in this study, where only two years of data are shown. Plots with all years at all the sites can be found in Appendix A. Note that for the sake of clarity, data have been smoothed with a 15-day moving average window in all the figures showing seasonal cycles, but not in any of the optimizations.'*

**P3332, L2-8: a small table showing the reductions in RMS according to the experimental setup would be very helpful in grasping these more objectively.**

Table 2 has been added (Table 3 in the revised manuscript), while the text has modified:

*'We have further investigated such hypothesis with 6 MS experiments where, as compared to the reference MS optimization, one or more of the four parameters related to the heterotrophic respiration ( $K_{soilC}$ ,  $Q_{10}$ ,  $HR_{H,b}$  and  $HR_{H,c}$ ) are in turn left out of the assimilation procedure. Table 3 shows that without  $K_{soilC}$  there is a 38.5% reduction of the model-data RMS at yearly time scale, a value significantly lower than in the standard case (49.3%). A similar degraded performance is found when the three other parameters are simultaneously left out of the optimization, whereas we observe a slight improvement (as compared to the standard MS case) when each one of them is individually excluded. When the four parameters are not considered, the optimization procedure becomes significantly less efficient with a 16.6% RMS reduction on yearly average.'*

**P3332, L20-22: could also occur because of the correlation in the drivers.**

C3132

**Table 2.** All-site yearly reduction of the model-data NEE RMS from the prior model, for various combinations of  $R_h$  parameters left out the optimization.

$R_h$ parameter(s) left out	None	$K_{soilC}$	$Q_{10}$	$HR_{H,b}$	$HR_{H,c}$	$Q_{10}+HR_{H,b}+HR_{H,c}$	all
Yearly RMS reduction (%)	49.3	38.5	49.3	54.3	55.2	38.6	16.6

If the Referee's remark is about the error correlation in the drivers, we must point out that we are using the *observed* meteorology, so that we can reasonably assume that the error correlation between the different drivers is not significant. The error of one captor is probably independent of the others.

But if the Referee means the correlation between the drivers themselves, the impact on the posterior parameters errors is not clear to us: although we suspect that it would have an influence, it is something we cannot directly test. In this context, we did not change the text as we do not foresee any specific need to mention this point.

**P3333, L15-16: known as the equifinality problem [see for example Franks et al., 1997; Medlyn et al., 2005].**

We have added it to the revised manuscript:

*'This indicates that we might face an equifinality problem: a range of different parameter sets may yield similar model performance, and similar model predictions.'*

**P3336, L1-7: the  $R_a$  parameters also include a scalar that implicitly scales  $R_a$  to mismatches in vegetation biomass  $GR_{frac}$ , which could explain the summer mismatch in NEE if  $R_a$  is in general overestimated during this period. We should see that this deduction is very linked to the model structure since  $R_g$  does not seem to depend on instantaneous productivity, as is usually assumed in other models. Also here we see  $c_{0,i}$  – how does this coefficient vary?**

C3133

We added to the text that our deductions in this paragraph are specific to the ORCHIDEE model. As for  $c_{0,i}$  (g/g/day), this coefficient does not vary, it is prescribed per PFT and its value also depends on the biomass compartment considered ( $2.62 \times 10^{-3}$  (leaf),  $1.19 \times 10^{-4}$  (sapwood, fruits, and carbohydrate reserve),  $1.67 \times 10^{-3}$  (roots) and 0 (heartwood)). We added it to the text :

*'...where  $T_i$ ,  $B_i$  and  $LAI$  are respectively the soil or surface temperature, the biomass content and the leaf area index, while  $c_{0,i}$  ( $g \cdot g^{-1} \cdot day^{-1}$ ) is the maintenance respiration coefficient at  $0^\circ C$ , which is prescribed depending on the PFT and the biomass pool  $i$ : 0 (heartwood),  $1.19 \times 10^{-4}$  (sapwood, fruits, and carbohydrate reserve),  $1.67 \times 10^{-3}$  (roots) and  $2.62 \times 10^{-3}$  (leaf).'*

**P3336: Shouldn't the energy balance parameters be site dependent?**

In the current effort of spatialization of the information, we chose to test a generic set of energy balance parameters. Given the good agreement between the MS- and SS-optimized LE fluxes, we chose to consider this genericity assumption as reasonable.

**P3338, L14: "green" should be grey.**

Corrected.

**P3338, L16: "purple" should be black.**

Corrected.

**P3340, 3.5.1.: Different methods to decompose NEE into GPP and  $R_{eco}$  have been compared before [Desai et al., 2008]. These would probably stand as a better benchmark for the modeled fluxes here. The comparison shown here is very oriented to a site where the actual NEE from both approaches is very different**

C3134

**(Granier:  $-289 \text{ gC.m}^{-2}.\text{yr}^{-1}$  against  $-562 \text{ gC.m}^{-2}.\text{yr}^{-1}$  from the 'Lathuile' dataset). Another point to consider would be to see if the uncertainties from both inCux partitioning and modeled inCuxes intersect.**

Indeed we realize that this section lacks spatial representativity in order to make a robust point out of the data of Granier et al. (2008) at Hesse site. As another referee suggested, we will omit this paragraph in the revised manuscript and only mentioned that such in-depth comparison with site-specific studies is beyond the scope of the paper but would need further attention in a more systematic way at all sites:

*'In-depth comparisons with site-specific gross flux estimates at each site (e.g. Granier et al., 2008) is beyond the scope of this paper but would deserve further attention for a more precise evaluation of the optimization procedure at all sites.'*

**P3342, L6-9: this could be something very specific to the structure of ORCHIDEE (see above) and see Keenan et al. [2012].**

Our statement is indeed somewhat too general; we have tempered it in the revised manuscript with:

*'Overall, our optimization scheme is able to provide a set of parameters which fairly improves the simulation of assimilation and respiration processes in the ORCHIDEE model, although we have chosen to assimilate daily NEE and not to separate between nighttime and daytime values.'*

**P3343, L2-4: since the (mis)matches in NDVI/FPAR time series stemming from the land cover component are already considered (P3343, L4-10) the ignored cells for comparison should only target those where no seasonal cycle is seen in the observational data, and not in the observational and model data. The comparisons would be biased optimistic by excluding cells with no seasonal cycle from model outputs that might have it in the observations.**

C3135

This criterion was originally formulated to work with several PFTs, however in the cases of deciduous broadleaf forests the model always simulates a significant seasonality. As a result, we checked that removing this filter on the models output did not affect the results as the DBF PFT is dominant (50% threshold). Therefore, we have modified the sentence accordingly :

*'The mesh cells where no clear NDVI annual cycle is visible are ignored in the calculation, i.e., when observed time series have a standard deviation lower than 0.04.'*

**P3345, L18-20: Could the link between the current work and the work of Santaren et al. [2012, which is not published yet] be more concrete?**

The paper is still not yet submitted (delays in the completion), and referring to it does not add much to the discussion. Further to a suggestion made by the other referee, we have omitted this sentence in the revised manuscript.

**P3345, L24-27: would it be expected that the bias in snow sublimation propagate to the current parameter sets?**

The ORCHIDEE model tends to overestimate the snow sublimation during at the end of winter and at spring time. During this period, we expect the optimization to produce larger albedo values (parameter  $K_{albedo,veg}$ ) and reduced rugosity length (parameter  $Z_{overheight}$ ). However, given the relative shortness of this period, we assume little effect on the other parameters' optimization. We conducted a sensitivity test at JP-Tak site (one of the most affected by this feature in our list of sites), artificially reducing the snow sublimation during spring time and concluded that the impact was indeed negligible. We added in the text that such effect has a minor impact on the parameter values :

*'The sublimation-related misfit of LE is not corrected, as we did not include specific parameters of snow build up and sublimation. Note, however, that sensitivity tests*

C3136

*have shown no significant impact of this discrepancy of LE upon the optimization of the other energy balance parameters.'*

**P3346, L15-26:** It seems also that the DBFs in the Northern Hemisphere are in much colder regions than in the Southern Hemisphere. The sites considered are between 36° and 56° N, and only 2 sites below 42. There is no indication if the mean annual temperature (or any other climate diagnostic) would significantly differ this optimization convergence would occur. Could this be given as a reason behind the different conclusions in this study and Groenendijk et al. [2011]?

At the global scale our analysis is only based on the phenology, which is solely governed by a temperature criterion for the considered PFT, temperate deciduous broadleaf forest (DBF). We acknowledge that a significant difference in mean annual temperature in some of the regions considered in the southern hemisphere would certainly affect the effectiveness of the optimized parameters regarding leaf onset ( $K_{pheno,crit}$ ) and senescence ( $T_{senes}$ ). However, we think that the major factor of discrepancy is the phenology scheme itself, as ecosystems such as those in arid Australian forests are likely to be primarily controlled by the soil water availability. Thus our results might call for a refinement of the phenology scheme for such PFT with a slightly different scheme for the drought-controlled DBF leaf onset and leaf senescence. Such refinement is beyond the scope of this paper. The conclusion has been rephrased in that direction:

*'From this starting point, the MS optimization brings a slight improvement in the northern hemisphere, and contrasting results in the southern hemisphere, where none of the sites used in the optimization are located: significant improvement in South Africa but degradation in Australia. At the global scale, the correlation median shifts from 0.83 to 0.88. The degradation in Australia might reflect the limits of the phenological scheme of deciduous forests in the model, solely based on a temperature criterion.'*

C3137

*The tree species in the arid Australian forests, although classified as DBF, are likely to have a phenology strongly controlled by the available soil moisture, a feature much less prevalent at the sites used in this study. At present, we can only suggest the need for further investigations regarding the formulation of the DBF phenology in the model, towards a refinement of the PFT classification.'*

**P3356, Table 2:** the addition of a column with the posterior parameter values and uncertainties would be very useful.

Given the number of different optimization set-ups, with different number of parameters and uncertainty components (MS, SS...), summarizing it in a single table, especially in Table 2, did not seem very convenient. We think that the Fig. 4 already provides a useful summary.

**P3359, Figure 2 (and beyond):** no uncertainties in the data? In this case (Fig. 2b NEE) we also see that the MS optimizations perform better than the SS. It seems to happen in some cases. The reason this is happening could be related to the uncertainties included in the cost function (R), which vary between SS and MS settings.

We took the data as a reference, quantifying the uncertainty for the modeled fluxes. Indeed MS optimization sometimes performs better than the SS, but in our view this might also be linked to local minima in the convergence of the SS cost function, where the MS cost function benefits from a larger  $n_{obs}/n_{para}$  ratio which would somewhat 'smooth' the solution space for the cost function. The uncertainties included in the cost function, based on the prior model-data mismatch, are however independent from the optimization considered (MS or SS).

**P3361, Figure 4:** Does KsoilC correlate with NEE?

C3138



This is a very good point. We looked at the correlation between the annual NEE ( $\text{gC/m}^2/\text{yr}$ , averaged over all the years available at each site) and the optimized values of  $K_{\text{soilC}}$ . In both multi-site and single-site optimizations, we had observed no significant correlation (lower than 0.3), which suggests that there are important drivers of the annual carbon budget other than the initial carbon stocks, although the optimization of the latter significantly reduces the annual mismatch between model and observation.

#### References :

- Carvalhais, N., et al. (2008), Implications of the carbon cycle steady state assumption for biogeochemical modeling performance and inverse parameter retrieval, *Global Biogeochem Cy*, 22(2), doi:10.1029/2007GB003033.
  - Chevallier, F. (2007), Impact of correlated observation errors on inverted CO<sub>2</sub>(2) surface fluxes from OCO measurements, *Geophys. Res. Lett.*, 34, L13802, doi:10.1029/2007gl030463.
  - Desai, A. R., et al. (2008), Cross-site evaluation of eddy covariance GPP and RE decomposition techniques, *Agr Forest Meteorol*, 148(6-7), 821-838.
  - Franks, S. W., et al. (1997), On the sensitivity of soil-vegetation-atmosphere transfer (SVAT) schemes: Equifinality and the problem of robust calibration, *Agr Forest Meteorol*, 86(1-2), 63-75.
  - Garrigues, S., et al. (2008), Validation and intercomparison of global Leaf Area Index products derived from remote sensing data, *J. Geophys. Res.*, 113, G02028, doi:10.1029/2007JG000635.
  - Groenendijk, M., et al. (2011), Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data, *Agr Forest Meteorol*, 151(1), 22-38.
- C3139
- Hollinger, D. Y., and Richardson, A. D. (2005), Uncertainty in eddy covariance measurements and its application to physiological models, *Tree Physiology*, 25, 873-885.
  - Keenan, T. F., et al. (2012), Terrestrial biosphere model performance for inter-annual variability of land-atmosphere CO<sub>2</sub> exchange, *Global Change Biol*, 18(6), 1971-1987.
  - Kuppel, S., et al. (submitted), Quantifying the model structural error in Carbon Cycle Data Assimilation Systems.
  - Liu, Y. Q., and H. V. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resour Res*, 43(7), doi:10.1029/2006WR005756.
  - Medlyn, B. E., et al. (2005), On the validation of models of forest CO<sub>2</sub> exchange using eddy covariance data: some perils and pitfalls, *Tree Physiol*, 25(7), 839-857.
  - Schrumpf, M., et al. (2011), How accurately can soil organic carbon stocks and stock changes be quantified by soil inventories?, *Biogeosciences*, 8(5), 1193-1212.

---

Interactive comment on *Biogeosciences Discuss.*, 9, 3317, 2012.