Response to Referee #2 (Dr. Malaquias Pena Mendez)

General comments:

1. This paper applies a particle filter data assimilation scheme to assimilate MODIS Leaf Area Index (LAI) data into an explicit individual-plants dynamical global vegetation model. Results indicate that the scheme reduces the uncertainty of the LAI analyses as compared to random initialization. Furthermore, the technique appears to successfully estimate the model parameters that control separately LAI for the forest and for the grass types, out of whole LAI observations.
2. The content of the paper is relevant for Earth Systems and non-linear modeling. It addresses one important aspect to increase models realism through the use of information contained in fine time-scale observed data.
3. The problem addressed in this study is challenging considering the nonlinearities in the dynamics of the vegetation, the multitude of interactive physical and biogeochemical processes taking place at the local and regional scales, and the fact that not all state variables are observed or retrieved by satellite.
4. The study is well executed to proof the concept with all the needed elements (calibrated model, quality controlled data, an optimized data assimilation scheme) and reduced (only a few geographical points) scope to make it successful.

Response: Thank you very much for the useful, constructive comments to improve the paper. We will revise the manuscript accordingly. Our point-by-point responses are shown in blue.

Specific comments and questions:

1. While the description is succinct and easy to follow, it needs to make explicit major assumptions made and the problems one may encounter if they were to be relaxed.

Response: Thank you very much for the suggestion. We made strong assumptions in the OSSE, and relaxed some in the real-world experiment. We still made strong assumptions in the limited area application. We will make these assumptions clear in the revised manuscript.

2. Below is a list of questions that arose while reading the manuscript:

2.1: What modifications to the original vegetation model were made to adapt it to the DA scheme? Were these only the changes in parameter values we see in the appendix?

Response: The model simulates daily states, but the original model outputs were only once per year. Daily outputs are needed for data assimilation every once in four days. Therefore, we modified the model code to output the daily states every 4 days. In addition, the original model code assumed running for many years continuously, and the initial seed for the random number generator was fixed. Since in this study we stopped the model every 4 days, and the same seed was repeated every time we started the model. Therefore, we modified the model code to randomly generate the seed for the random number generator every time when we initiate the model. These are the only modifications not shown in the paper, because we thought these were only minor technical modifications. To explicitly describe all necessary changes to the existing model code, we will include these in the appendix.

2.2: Are the field observations in the Siberia Yakutsk Larch forest site independent or were they included to create the climate forcing data 2001-2007 in the vegetation model?

Response: The climate forcing data were created using the NCEP/NCAR reanalysis data and CRU observation based data. The observed climate data at this site were not directly used in our experiments, but these data might be included in the NCEP/NCAR reanalysis data. It is not simple to find if the site observation data were reported through GTS and included in the NCEP/NCAR reanalysis. Therefore, we will include a sentence in the revised manuscript describing the possibility that the observations at the site may be used in the NCEP/NCAR reanalysis through GTS.

2.3: Was the 2004-2007 period of MODIS 4-days frequency data continuous on the study site? Were there missing data? How was the missing data handled?

Response: There are a number of missing data in the quality-controlled MODIS data. Therefore, as we have described in P5. Line 30-31, if the number of the quality controlled MODIS data in the 10-km radius contains less than 300 grid points, we set these data as the missing data. We will revise the manuscript to describe more explicitly about the missing data from the original quality-controlled MODIS data.

2.4: Was the 8000 particles generated decided by computer capacity, or any other criteria?

Response: In response to the other reviewer's comment #2, we performed additional experiments with different particle sizes. We will add a new section to show the sensitivity to the particle size in the revised manuscript.

2.5: Simulated observations (in the OSSE) versus real observations: How do they compare? Were the real observations also normally distributed? Were standard deviations of real observations about 10% as in the OSSE experiment?

Response: We assumed the normal distribution for the real observation error. The error standard deviation is included in the MODIS dataset that we used (Knyazikhin et al., 1999). As already described in P.6 Lines 1-2, we used "the median of the error standard deviations" in the 10km radius. The standard deviations of the real observations are different from those of the OSSE, as indicated by Figs. 3-a and 6-a. We will explicitly describe about the differences of the observation error standard deviations between the simulated and real observations in the revised manuscript.

2.6: Did you follow any particular rule to determine the perturbation size of Pmax and Dor? In the study you allow larger amplitude perturbations for forest than for grass types. The amplitude of Dor is relatively very small.

Response: There are two perturbation settings for the model parameters: the initial perturbation sizes and the random perturbation sizes when resampling. We selected the initial perturbation sizes based on the ecological knowledge from the previous studies (Kolari et al., 2006; Zeng et al., 2011; Zhao et al., 2015; Takagi et al., 2015). The initial Pmax perturbation size for grass is 4 times smaller than that of forest. The initial Dor perturbation sizes for grass and forest are the same. The random perturbation sizes when resampling follow the initial perturbation sizes. We will add the references and explicit descriptions about the perturbation settings in the revised manuscript.

Following the comment of the other referee, we performed additional experiments with different random perturbation settings for the initial perturbation sizes and the random perturbation sizes when resampling. We will add a new section to show these results and to discuss the sensitivity to the perturbation settings in the revised manuscript.

2.7: The manuscript indicates that perturbations of parameters are applied only to duplicated particles. Since the particle DA scheme eliminates particles far away from observations (Fig. 1), that would mean that the range of the distribution of all the particles decreases after several cycles at least compared to the initial (uniform) distribution. Is this correct? Still, you do not report any issue with collapsing of the DA scheme when observations are outside the range of the distribution of particles. Can you please, elaborate more on this issue?

Response: Yes, it is correct that the range of the distribution of all the particles decreases after several cycles. If we apply proper random perturbations to the duplicated particles, we can avoid filter collapse. However, our additional experiments showed filter collapse when the particle size is 1000. We will describe about the collapse in the new section on the sensitivity to the particle size and random perturbation size.

2.8: The NODA and the TEST experiments; Figure 3. How the 8000 particles are inserted at the initial conditions? Is this done every 4 days with a uniform distribution each time? The TEST experiment appears to reduce a big systematic error that appear during the growing months. Traditional DA schemes apply a bias-correction strategy of the First Guess prior to performing the analysis. Does this mean that particle DA also removes systematic errors?

Response: As already described in P.4 Line 28, "The 8000 particles at the end of the 103-year spin-up runs are used as the initial conditions for DA", and the NODA and TEST experiments start from the same initial 8000 particles. The 8000 particles continue to be the same until the first observation of LAI is assimilated. Since the LAI is observed only when greater than 0.5, the LAI observation exists only in the summer season. The systematic errors in NODA comes from the uncertain parameter settings. TEST can estimate parameters using observed LAI, and therefore, can reduce the systematic errors. This is different from the bias-correction strategy of the first guess. There is no explicit bias correction applied to the TEST experiment. So, we understand that this particular particle DA can reduce systematic errors by estimating the uncertain model parameters. We will discuss this point in the revised manuscript.

2.9: TEST experiment; Figure 3a (forecast+grass). Please, explain the problem at the end of the fall months (circled in blue in the attached figure). Can this be attributed to neglecting observations when LAI<0.5 ? Will this be removed if observations are added there?

Response: Yes, because we assume observation to be available only when LAI>0.5, it is difficult to estimate the LAI when observed LAI is 0.5 or smaller. We did not perform experiments with small LAI observations, because the MODIS data for the real-world experiment did not include LAI<=0.5 (Fig. 6). There are too few data with real MODIS LAI<=0.5, and our preprocessing assigns the missing value. We will add these discussions in the revised manuscript.

2.10: It is obscure to me how come the individual LAI of Forest and Grass are accurately estimated out of the whole LAI. Even when the whole LAI estimation is incorrect as in the periods in the blue circle in the attached figure. What mechanism or statistical assumption within the DA process makes the partitioning of LAI correct? Is this pure chance?

Response: Thank you for the comment, which initiated further analysis of the results that were not shown in the manuscript. As already described in P.5, Lines 21-22, "To investigate the sensitivity to the choice of the nature run, we performed similar OSSEs by replacing the nature run with other randomly-chosen parameter sets". We investigated these different OSSEs more carefully and found that the parameters for grass were estimated well when the nature run used a larger Pmax value for grass (Figure 1 left, as shown in the manuscript). However, in another OSSE (Figure 1 right), the nature run used a small Pmax value; the results showed that the parameters for grass showed significantly larger uncertainties (Fig. 1 c, e), while the parameters for forest were estimated well (Fig. 1 b, d). Larger Pmax values for grass produce more grass LAI, which can be observed with the observing threshold of LAI > 0.5 near the growing and falling periods (shown by the blue circles provided by the reviewer). With smaller Pmax values for grass, the small grass LAI cannot be observed directly, but the large LAI observations in the summer season predominantly suggest forest LAI. This would allow to estimate the forest parameters well, although the grass parameters showed larger uncertainties. We will include these results and discussions in the revised manuscript.

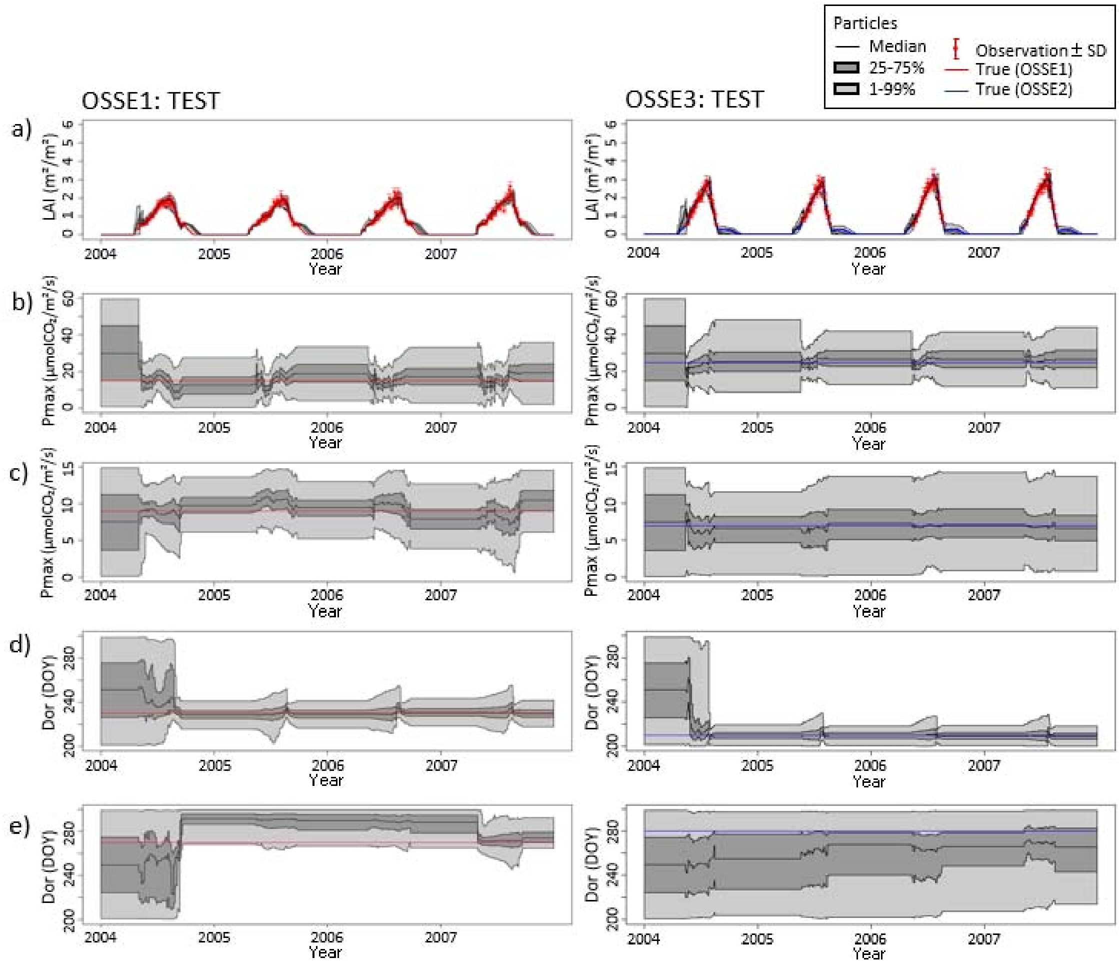

Figure 1. The results of the OSSE1 and OSSE3 with data assimilation. The particle sizes and the perturbation sizes are the same as the discussion paper. a) Time series of LAI for forest + grass, b) Pmax for forest, c) Pmax for grass, d) Dor for forest, e) Dor for grass. Dark and light gray areas indicate the quartiles and 1-99 % quantiles of the particles as shown in legend. Thick black curves indicate the medians. Red dots with error bars indicate the observations and their error standard deviations, and red and blue lines indicate the true values used for the nature runs.

2.10: In the Real-World Experiment; there is no detail on the perturbation strategy, so I suppose it is the same as in the OSSE experiment.

Response: We did not provide the details of the experimental settings. Yes, the perturbation strategy of the real-world experiment is same as that of the OSSE. We will add the descriptions in the revised manuscript.

2.11: The observation error standard deviation in the real case needs more explanation. What is the truth from which the error is estimated? Is this the in situ observation? Is this error an input in the DA scheme?

Response: As described in our previous response to the comment #2.5, P6 lines 1-2 reads "The observation error standard deviations are assigned to each LAI datum from the original source, and we took the median of the error standard deviations." We rely on the original MODIS data source about the estimate of the observation error standard deviation. We will revise this sentence about the observation error standard deviations to avoid potential misunderstanding.

Technical corrections

Page 1. Abstract. ".. newly developed" should be ".. developed". You repeated that later on in the text.

Response: We will correct it accordingly.

Page 1. Abstract. ".., assuming the satellite-based LAI." This is an incomplete statement. You repeated this statement problem in the introduction, page 2, row 11. Maybe you meant to say "using" instead of "assuming".

Response: We do assume the satellite-based LAI data for the OSSE, but more precisely, we "simulate" the satellite-based LAI in the OSSE. Therefore, we will replace "assuming" with "simulating" in the revised manuscript.

Page 2. row 8. "straightforward" may be replaced by "numerically straightforward". In this context, it is not simple to go from local to global because spatial covariances become relevant.

Response: We agree, and will revise it accordingly.

Page 2. (last) row 31. "phase space stays the same" may be "phase space dimension stays the same"

Response: We will revise it accordingly.