Response to Referee #1 (Assistant Professor, Matthias Morzfeld)

General comments: The paper presents a description of the application of a particle filter to assimilate leaf are index observations into a global vegetation model. The paper is concise, and presents only the necessary information, no long background/review section is provided. I find the numerical experiments and results convincing (pending specific comments below).

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

My main concern is that this paper might not be suitable for Nonlinear Processes in Geophysics. The reason is that the journal emphasizes new methods, applied to realistic problems. The paper simply presents an "old" method applied to a new problem. I find it interesting to read that a particle filter can solve an important and "real" data assimilation, however the general NPG readership might get bored. The authors should decide wether NPG is the best journal to reach the audience they want to reach. This is also reflected by the references, only few of which are to articles in journals similar to NPG. I suspect that this paper would also make a fine contribution in a journal that is more focused on, e.g., Earth system modeling. The authors may want to consider going that route.

Response: We would really appreciate the kind suggestion. After carefully considering other potential options, we reached to the conclusion that we would like to have this paper published in NPG because the main focus of this paper is methodology. It would not be trivial if well-known, “old” methods work with a new problem. As we mentioned in the discussion paper (P6. Line 26-27), “To the best of the authors’ knowledge, this is the first study to assimilate the fine time-scale satellite data with an *individual-based* DGVM”. We also explored sensitivities to the filter parameters such as the initial and resampling perturbation sizes and particle size. In the revised manuscript, Section 4 was added about the sensitivity experiments. We believe these methodological explorations would be useful for the readers of NPG. In addition, this study performs simple experiments at a single location with only a couple of plant functional types (PFTs). Future studies will include spatial distributions with more PFTs, and will be more relevant to domain-specific journals. We added related descriptions in the revised manuscript (P.2 Line 8-11, P9.

Line 14 - 16).

Specific comments:

1. I wonder if there is any sensitivity to how repeated particles are perturbed after resampling. The authors chose a random perturbation, but miss to motivate their choice. I think the paper should contain numerical experiments where it is shown that either the method is robust to (small) changes in how repeated particles are perturbed, or it should be reported how the perturbations influence the results.

Response: Following the suggestion, we performed additional experiments with different random perturbation settings, and found that the filter collapsed for biomass with smaller perturbation settings, though estimated model parameters and other state variables were estimated accurately without collapse. We added a new section in the revised manuscript to show these results and to discuss the sensitivity to the perturbation settings (P7. Line 12 -21).

2. The number of particle used is typically important for the results one obtains with a particle filter. Indeed, much of the meteorological literature says that the number of particles required is excessive. To address this issue, I would suggest to run more numerical experiments with a varying number of particles. One can then compute, e.g., means and variances, and check that the method has converged when, e.g., 8000 particles are used. Specifically, I suggest experiments with 4000, 8000 and perhaps 16000 particles (if possible).

Response: Following the suggestion, we performed additional experiments with different particle sizes ranging from 500 to 16000. The results showed that the filter collapsed for biomass with 4000 particles or less when we keep the same random perturbation setting. Also, the model parameter and other state variable were not estimated accurately with 500 particles with a smaller random perturbation setting. We included the sensitivity to the particle sizes along with the sensitivity to the random perturbation settings (cf. previous comment) in the revised manuscript (P7. Line 12 -21).

3. I wonder what happens when the data assimilation is initialized with a "smaller" initial uncertainty. The authors define intervals for the parameters, but do not mention how they came up with these intervals. It would be interesting to see what happens when these intervals are shortened or widened. In particular, the particle filter has no mechanism to bring the parameters to values that are not contained within the initial set. This could make things difficult for the

"real life" application. Again, I suggest to investigate this issue with more numerical experiments.

Response: We would appreciate the suggestion to perform additional sensitivity experiments on the initial parameter uncertainties. We selected the intervals of the initial parameters 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). We added the specific references in the revised manuscript (P5. Line 8-9).

Following the comment, we performed additional experiments with different initial parameter uncertainties. The results showed that the parameters were not estimated accurately with wider initial intervals when 2000 particles or less are used, although they were estimated accurately with narrower initial intervals. With 8000 particles, the parameters were estimated accurately even with wider initial intervals. Sampling a wider interval with a smaller particle size generally reduces the particle density, or the effective number of particles, so that the results seem to be reasonable. We included these sensitivity results in the revised manuscript (P6. Line 29 - P7. Line 10).

4. In figs.4 (right column), 5d (right column), 7 (right column), and 8d (right column): it seems that the data assimilation only impacts the parameter estimates for parts of the year, however data are assimilated every 4 days. The authors miss to provide a clear explanation of why that is the case.

Response: We assimilated the LAI only when 0.5 or larger, therefore, data assimilation has impacts only in the summer when the leaves appear (i.e., LAI $\geq$ 0.5). We added this discussion in the revised manuscript (P5. Line 33 - P6. Line 1).

5. I would remove all NODA figures, as they do not really carry information. It is clear that when no data assimilation is used, no parameter is changed.

Response: We believe that it would be important to show NODA figures to highlight the impact of data assimilation. Since this is the first time to apply DA to SEIB-DGVM, it is unclear what impact DA would bring. Therefore, we would really like to keep the side-by-side comparison of the experiments with and without DA.

Technical corrections:

I find the use of "newly" in the first sentence of the abstract a bit unusual. I would suggest to re-formulate this sentence. The sentence also appears again later on (p.2. line 6, p.6 line 27), and there it should also be changed.

Response: We corrected these sentences in the revised manuscript (P1. Line 11, P2. Line 7).

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 revised 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 made these assumptions clearer in the revised manuscript (P9. Line 11-13).

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 included these in the revised manuscript (P2. Line 25-31).

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 included 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 (P4. Line 22-23).

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 in the discussion paper, 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 revised the manuscript to describe more explicitly about the missing data from the original quality-controlled MODIS data (P7. Line 31 - P8. Line 2).

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 added a new section to show the sensitivity to the particle size in the revised manuscript (P6. Line 29 - P7. Line 21).

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 in the discussion paper, 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 explicitly described about the differences of the observation error standard deviations between the simulated and real observations in the revised manuscript (P8. Line 3-6).

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 added the references and explicit descriptions about the perturbation settings in the revised manuscript (P5. Line 8-9, P5. Line18-19).

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 added a new section to show these results and to discuss the sensitivity to the perturbation settings in the revised manuscript (P6. Line29 - P7. Line 21).

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 500. We described about the collapse in the new section on the sensitivity to the particle size and random perturbation size (P6. Line29 - P7. Line 21).

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 in the discussion paper, “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 0.5 or larger, the LAI observation exists only in the summer season. Model state and parameters are estimated together at DA, and the model systematic errors associated with the parameters are corrected by DA with parameter estimation. No explicit bias correction is applied. We added these descriptions (P5. Line 4-5, P5. Line 14-17)

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 discussed this point in the revised manuscript (P6. Line 6-8).

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 less than 0.5. 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 added these discussions in the revised manuscript (P5. Line 3-5).

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 in the discussion paper, "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. However, in another OSSE, the nature run used a small Pmax value; the results showed that the parameters for grass showed significantly larger uncertainties, while the parameters for forest were estimated well. 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 included these results and discussions in the revised manuscript (P6. Line 14-27).

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 added the descriptions in the revised manuscript (P7. Line 24-27).

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 in the discussion paper 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 revised this sentence about the observation error standard deviations to avoid potential misunderstanding (P8. Line 3-6).

Technical corrections

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

Response: We corrected it accordingly (P1. Line 11, P2. Line 7).

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 replaced “assuming” with “simulating” in the revised manuscript (P1. Line 13, P2. Line 14, P5. Line 2).

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 revised it accordingly (P2. Line11).

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

Response: We revised it accordingly (P3. Lines 10, 17).

Non-Gaussian data assimilation of satellite-based Leaf Area Index observations with an individual-based dynamic global vegetation model

Hazuki Arakida[1], Takemasa Miyoshi[1, 2, 3], Takeshi Ise[4], Shin-ichiro Shima[1, 5], and Shunji Kotsuki[1]

[1]RIKEN Advanced Institute for Computational Science, Kobe, 650-0047, Japan
[2]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD 20742, USA
[3]Application Laboratory, Japan Agency for Marine-Earth Science and Technology, Yokohama, 236-0001, Japan
[4]Field Science Education and Research Center, Kyoto University, Kyoto, 606-8502, Japan
[5]Graduate School of Simulation Studies, University of Hyogo, Kobe, 650-0047, Japan

Correspondence to: Hazuki Arakida (hazuki.arakida@riken.jp) and Takemasa Miyoshi (takemasa.miyoshi@riken.jp)

Abstract. We ~~newly~~ developed a data assimilation system based on a particle filter approach with the Spatially Explicit Individual-Based Dynamic Global Vegetation Model (SEIB-DGVM). We first performed an idealized observing system simulation experiment to evaluate the impact of assimilating the leaf area index (LAI) data every 4 days, ~~assuming~~simulating the satellite-based LAI. Although we assimilated only LAI as a whole, the forest and grass LAIs were estimated separately with high accuracy. Uncertain model parameters and other state variables were also estimated accurately. Therefore, we extended the experiment to the real world using the real Moderate Resolution Imaging Spectroradiometer (MODIS) LAI data, and obtained promising results.

1 Introduction

The terrestrial biosphere is an important part of the Earth System Model (ESM) to simulate the carbon and water cycles. However, terrestrial biosphere models tend to have large uncertainties, for example, in phenology (Richardson et al., 2012; Murray-Tortarolo et al., 2013) and in spatial distributions of plant species (Cheaib et al., 2012). Recently, data assimilation (DA) methods which incorporate observation data into models have been applied to terrestrial biosphere models to reduce the uncertainties in the state variables and model parameters (Luo et al., 2011; Peng et al., 2011). Previous studies have successfully applied the ensemble Kalman filter (e.g., Evensen, 2003; Williams et al., 2005; Quaife et al., 2008; Stöckli et al., 2011) or adjoint method (e.g., Kaminski et al., 2013; Kato et al., 2013) to the *static* vegetation models, but studies with the *dynamic* global vegetation models (DGVMs) are still limited (Luo et al., 2011; Peng et al., 2011), although Hartig et al. (2012) pointed out the importance.

The *static* vegetation models are time-independent and do not include the vegetation succession process (Peng, 2000). Alternatively, DGVMs include the vegetation succession process and can simulate carbon and water cycle changes linking to the vegetation shift under the changing climate. Especially, *individual-based* DGVMs simulate local interactions among individual plants such as competitions for light and water, so that the model can simulate the vegetation succession more explicitly (Smith et al., 2001; Sato et al. 2007). Garetta et al. (2010) pioneered to apply DA to an *individual-based* DGVM for paleoclimate, but no study has been published thus far to assimilate fine time-scale data from satellites and ground stations using an *individual-based* DGVM. If the initial vegetation structure and the model parameters of an *individual-based* DGVM are estimated more accurately by assimilating the fine time-scale data, the uncertainties of the simulated future vegetation would be greatly reduced.

This study explores to assimilate frequent satellite-based Leaf Area Index (LAI) data with an *individual-based* DGVM known as the SEIB-DGVM, standing for Spatially Explicit Individual-Based DGVM (Sato et al., 2007). We~~ newly~~ developed a non-Gaussian ensemble DA system with the SEIB-DGVM based on a particle filter approach. ~~As the first step~~Although the particle filter is an existing, well-known approach, this is the first attempt to apply it to an *individual-based* DGVM with frequent LAI data. Therefore, we focus on the methodological development in this study ~~performs~~and perform a series of numerical experiments at a ~~local scale.~~single location with only a couple of plant functional types (PFTs) as the first step. It would be numerically straightforward to extend it to the global scale in the future studies, since the local-scale experiments can be performed in parallel for different locations. In ~~this~~the present study, we first perform ~~an~~ idealized ~~experiment~~simulation experiments to investigate how well we can estimate the model parameters associated with phenology by assimilating the LAI data every 4 days, ~~assuming~~simulating the satellite-based LAI product from the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Terra and Aqua spacecraft. We also investigate to what extent assimilating the LAI data could improve the estimates of the state variables such as GPP (Gross Primary Production), RE (Ecosystem Respiration), NEE (Net Ecosystem Exchange), and biomass, the most fundamental variables for carbon cycle and vegetation states. Sensitivities to the filter settings such as the random perturbation sizes and particle sizes are also investigated. Following the idealized ~~experiment~~experiments, we perform an experiment using the real MODIS LAI observation data to see how well the proposed approach ~~works~~performs in the real world.

2 Methods

2.1 SEIB-DGVM

The SEIB-DGVM simulates establishment, growth, and decay of the individuals of prescribed ~~plant functional types (~~PFTs~~)~~ within a spatially explicit virtual forest (Sato et al., 2007), forced by climate conditions such as air temperature, soil temperature, cloudiness, precipitation, humidity, and winds. We used version 2.71 (Sato and Ise, 2012) but with minimal modifications for DA~~ (cf. Appendix).~~. The model simulates daily states, but the original model outputs were only once per year. Outputs are needed for DA every once in 4 days, so that we modified the model code to output the model 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 stop the model every 4 days, and the same seed is repeated every time when we start 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. Other modifications are summarized in Appendix.

Among the various model outputs ranging from individual tree height to soil water content (Sato et al., 2007, with updated information available from the package of version 2.71), we focus on LAI because it is a key to the vegetation model, and because previous studies show a promise in assimilating satellite-based LAI data with a *static* vegetation model (Stöckli et al., 2011) and a *non-individual-based* DGVM (Demarty et al., 2007). We extend the previous studies to assimilate the LAI data with the *individual-based* DGVM.

2.2 Particle filter-based DA

Individual-based DGVMs include highly nonlinear processes such as occasional establishment and death of individual plants. These processes produce and eliminate state variables, and the phase space changes time to time. DA methods that have been used in geophysical applications usually assume that the state variables are defined uniquely for the given dynamical system and that the phase space dimension stays the same. The widely-used ensemble Kalman filter, for example, finds the best linear combination of the ensemble with optimal fit to the observations, but it is not trivial to define a linear combination or even the ensemble mean for the variables missing in some ensemble members. Therefore, it would not be trivial to apply the widely-used DA methods to *individual-based* DGVMs.

Alternatively, particle filters run independent parallel simulations or particles and represent the probability density function (PDF) explicitly by assigning probability to each particle. Therefore, particle filters can handle non-Gaussianity and nonlinearity explicitly, and can be applied to the *individual-based* DGVMs in a straightforward manner (e.g., Garetta et al., 2010) even though the phase space dimension is different for each particle.

Here we adopt an efficient particle filter approach known as the Sequential Importance Resampling (SIR~~;~~, Fig. 1) (Gordon et al., 1993). First, n parallel simulations are performed, and each simulation is considered as a particle representing the true state of the system with equal probability. Next, likelihood $l_t^{(i)}$ is calculated for each particle using the Gaussian likelihood function:

$$l_t^{(i)} = \mathrm{p}(y_t | x_{t|t-1}{}^{(i)}) = \frac{1}{\sqrt{2\pi \cdot \sigma^2}} \exp\left\{-\frac{(y_t - x_{t|t-1}{}^{(i)})^2}{2 \cdot \sigma^2}\right\} \qquad for\ i = 1, \ldots, n. \qquad (1)$$

Here, $x_{t|t-1}{}^{(i)}$ denotes the simulated LAI of the ith particle at time t from the previous time step t-1, y_t the observed LAI at time t, and σ the observation error standard deviation. Since the prior probability is uniform, Bayes' rule gives that the posterior probability of the ith particle is proportional to $l_t^{(i)}$, i.e., the particles closer to the observation have more probability. Next, we resample the particles, so that each particle has equal probability. The particles with more probability (larger $l_t^{(i)}$) are duplicated, and the particles with less probability (smaller $l_t^{(i)}$) are removed. If n is sufficiently large, we can evaluate the posterior PDF accurately. Each resampled particle represents the true state of the system with equal probability and acts as the initial particle for the next time step. This Bayesian framework is repeated.

2.3 OSSE and the real-world experiment

We first perform ~~an~~a series of idealized Observing System Simulation ~~Experiment (OSSE~~Experiments (OSSEs). The OSSE (e.g., Atlas, 1997) is a widely-used approach in meteorological DA to test the general performance of a DA system and to evaluate the impact of specific observing systems. OSSE has the nature run, which is usually generated by running a simulation for a certain period. Observation data are simulated from the nature run by applying the observation operator, i.e., converting the model variables to the observed variables. Here, we add artificial random noise to simulate the observation error. DA experiments are initiated from the state independent of the nature run, and the simulated observations are assimilated. The resulting analyses and subsequent forecasts are compared with the nature run to evaluate the performance of DA. Once an OSSE is done, it is straightforward to extend the OSSE to the real world by simply replacing the simulated observations with the real-world observations.

3 OSSE

3.1 Experimental design

~~—~~To generate the nature run, the SEIB-DGVM was initialized with the bare ground (i.e., no plant at the beginning) and was run for 107 years using the climate forcing data from year 2001 to 2010 available at the SEIB-DGVM webpage (http://seib-dgvm.com/). Here, the 10-year forcing data are repeated for the 107-year simulation, and the last 7 years from year 101 to 107 use the actual climate forcing of 2001 to 2007, so that we call year 101 to 107 to be 2001 to 2007. The daily climate data were generated by the procedure of Sato and Ise (2012) with updated information available at the SEIB-DGVM webpage, based on the monthly Climate Research Unit observation-based data (CRU-TS3.22 0.5-degree monthly climate time series) (Harris et al., 2014) and the daily data from the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) reanalysis (Kalnay et al., 1996). We chose the study area at one of the AsiaFlux sites, the Siberia Yakutsk Larch forest site at Spasskaya Pad, the middle basin of River Lena (62º 15' 18" N, 129º 14' 29" E). The observed climate data at this site were not directly used in this study, but these data may have been included in the NCEP/NCAR reanalysis. Field ~~observation~~observed carbon flux data are available as the ground truth to verify the DA results at this site. Forced by the climate data, the SEIB-DGVM simulates the vegetation shifts from the bare ground to a grassland, and then to a forest. The two PFTs, the boreal deciduous needle leaved trees and C3 grass, are the dominant PFTs in this study area. Therefore, we do not consider the other PFTs in this study~~.~~ following Sato et al. (2010). We call these two PFTs simply "forest" and "grass".

The nature run (Fig. 2 a) was performed with the "true" parameter values Pmax = 15 $\mu molCO_2\ m^{-2}\ s^{-1}$ and Dor = 230 DOY (day of year) for forest and Pmax = 9 $\mu molCO_2\ m^{-2}\ s^{-1}$ and Dor = 270 DOY for grass, where Pmax and Dor stand for the maximum photosynthesis rate and the start date of the dormancy, respectively (Fig. 2 b). Hereafter we omit the units for Pmax ($\mu molCO_2\ m^{-2}\ s^{-1}$) and Dor (DOY) for simplicity. The LAI observations for the last 4 years from year 2004 to 2007 were created by adding independent Gaussian random noise to the LAI values from the nature run (Fig. 2 a) every 4 days,

~~assuming~~simulating the MODIS LAI product. Here, the observation error standard deviation was given by 10 % of the nature run LAI value. The observed LAI < 0.5 ~~are discarded and~~ were not used for DA because the MODIS data for the real-world experiment did not include LAI < 0.5. There are too few data with real MODIS LAI < 0.5, and we assign the missing value in preprocessing. Since the LAI is observed only when 0.5 or larger, the LAI observation exists only in the summer season.

Next, 8000 particles (parallel simulations) were generated with uniformly perturbed parameters: Pmax = [0, 60] for forest, Pmax = [0, 15] for grass, Dor = [200, 300] for both. Here, [*a*, *b*] denotes random draws from the uniform distribution between *a* and *b*. These initial perturbation sizes are based on the previous studies (Kolari et., 2006; Zeng et al., 2011; Zhao et al., 2015; Takagi et al., 2015). We ran 8000 parallel simulations for 103 years for spin-up from the bare ground using the same climate forcing data as the nature run. In the course of the vegetation succession, these randomly perturbed parameter sets result in a variety of LAI simulations (Fig. 2 b).

The 8000 particles at the end of the 103-year spin-up runs are used as the initial conditions for DA. The simulated LAI observations are assimilated every 4 days. The nature run and particle filter use the same climate forcing data, so that the difference comes from the model parameter values. ~~To avoid the exact duplications and convergence~~The particles continue to be the free runs until the first LAI observation is assimilated in the summer season. The state variables and model parameters are estimated together at DA, and the model systematic errors associated with the model parameters are corrected by DA with parameter estimation. No explicit bias correction is applied. To avoid the exact duplications after resampling, the model parameters Pmax and Dor are randomly perturbed for the duplicated particles. Here, random draws [-4, 4] are added to Pmax for forest and to Dor for both forest and grass, and [-1, 1] ~~is~~are added to Pmax for grass because the initial Pmax perturbation size for grass is a quarter of that of forest. In case that these perturbed parameters exceed the corresponding initial parameter range, the excess value was bounced back from the limits. To assess the impact of DA, we also perform an experiment without DA ("NODA" hereafter), and compare to the experiment with DA ("TEST" hereafter).

3.2 Results

Figure 3 shows the time series of LAI for NODA (left) and TEST (right). The observations (Fig. 3 a, ~~red~~blue dots with error bars) cannot distinguish the forest and grass, but the model simulates LAIs for forest and grass separately (Fig. 3 b, c). Although the particles without DA are widely spread (left, gray areas), DA makes the particles much narrower (right) and consistent with the nature run (red curves). With DA, the median of the particles for forest is almost identical to the nature run for the entire 4 years (Fig. 3 b, right). As for grass, the median of the particles is also very close to the nature run with DA, but in the first three years the dormancy period is delayed (Fig. 3 c, right).

The model parameters are estimated accurately (Fig. 4). There is no direct observation of these parameters, so that the estimations are purely due to DA of the LAI observations. Although the particles of the NODA experiment are uniformly distributed (Fig. 4, left), DA makes the particles close to the true parameters (Fig. 4, right). Since we assimilated the LAI only when 0.5 or larger, DA has an impact only in the summer season when the leaves grow. It takes 1-4 years until the true values fall within the quartiles of the particles. The Pmax estimates for both forest and grass show occasional jumps, but tend to stay around the true values (Fig. 4 a, b). Dor for forest seems the most accurate and stable after the dormancy period of the first year (Fig. 4 c). Dor for grass takes the longest; the estimation is not accurate until the dormancy of the fourth year (Fig. 4 d). This may be related to the previous results showing the erroneous estimates of the grass LAI near the dormancy period in the first 3 years (Fig. 3 c). The systematic errors in NODA come from the uncertain parameter settings. TEST can estimate the parameters through DA, and can reduce the systematic errors. This is different from the bias-correction strategy of the first guess.

Other model variables such as GPP, RE, NEE and biomass show large improvements (Fig. 5). Although the particles of the NODA experiment are widely spread, DA with only LAI observations greatly reduces the uncertainties for the four variables, and the estimations are generally reasonable.

4 Sensitivity experiments for OSSE

4.1 Sensitivity to the nature run

To investigate the sensitivity to the choice of the nature run, we performed ~~similar~~ two additional OSSEs, which we call “OSSE2” and “OSSE3”, by ~~replacing the~~generating different nature ~~run~~runs with ~~other randomly-chosen~~ different parameter sets~~. Also, we generated different~~ (Table 1). The random numbers for the observation errors~~. The~~ are also different. The other settings follow the TEST experiment.

The results ~~showed~~ show that both OSSE2 and OSSE3 perform well in general. Namely, the LAI and parameters are estimated generally well (Fig. 6). We find the main difference between OSSE2 and OSSE3 in the parameters for grass (Fig. 6 c, e). OSSE3 shows significantly larger uncertainties for the parameters for grass. In OSSE2, the Pmax value for grass is larger and produces more grass LAI. Since grass starts to grow earlier and stays longer than forest, it is critical to have LAI observations near the emerging and falling periods for estimating the grass parameters. Due to the larger Pmax value for grass in OSSE2, LAI can be observed with the observing threshold of LAI = 0.5 near the emerging and falling periods. By contrast, in OSSE3, the Pmax value for grass is smaller, and the small grass LAI < 0.5 cannot be observed. We can see this in the LAI time series (Fig. 6 a, right) near the tails in the spring and fall seasons every year. The uncertainties of LAI are not reduced year by year, corresponding to the large uncertainties of the grass parameters. In the summer, LAI becomes larger mostly due to forest, so that the forest parameters can be estimated well.

4.2 Sensitivity to the initial perturbation size

Here we investigate the sensitivity to the initial perturbation sizes with particle sizes ranging from 1000 to 16000. Table 2 shows the three initial perturbation settings: small, moderate and large. For the TEST experiment, the moderate initial perturbation sizes were used. We perform additional sensitivity experiments with the small and large initial perturbation sizes. Except for the initial perturbation sizes and the particle size, the experiments follow the TEST experiment.

Tables 3 show the mean absolute errors (MAE) and the widths of the 1-99% quantiles, respectively, averaged over a year in 2007. We consider that the filter diverges when the MAE is larger than the half width of the 1-99 % quantiles, as shown by gray shades in the tables. The results show that the filter diverges for biomass in 10 out of 15 experiments. The 5 experiments that do not diverge are (4000; small), (8000; small), (16000; small), (8000; moderate) = TEST, and (16000; moderate), where (;) denotes (particle size; initial perturbation sizes). (1000; large) causes filter divergence for most variables and parameters. (2000; large) shows filter divergence for Dor for grass in addition to biomass. Sampling a wider interval with a smaller particle size generally reduces the particle density, or the effective number of the particles, so that the results seem to be reasonable.

4.3 Sensitivity to the resampling perturbation size

Here we investigate the sensitivity to the resampling perturbation sizes with particle sizes ranging from 500 to 16000, in a similar way as the previous subsection. Table 4 shows the three resampling perturbation settings: small, moderate, and large. For the TEST experiment, the moderate resampling perturbation sizes were used.

Tables 5 show similar ~~results with the different nature runs and random numbers~~tables as Tables 3 but for the sensitivity to the resampling perturbation sizes. We use the similar notation of (;) denoting (particle size; resampling perturbation setting). The results show that the filter diverges for biomass in 13 out of 18 experiments. The 5 experiments that do not diverge are (4000; large), (8000; moderate) = TEST, (8000; large), (16000; moderate), and (16000; large). (500; small) is most unstable, with more variable and parameter showing filter divergence. Resampling perturbations act as variance inflation in the ensemble filters (e.g., Anderson and Anderson 1999). It is known that variance inflation generally stabilizes the filter, and the results obtained here seem to be consistent.

45 Real-world experiment

45.1 Experimental settings

~~The~~Here the OSSE ~~was~~is extended to the real world by replacing the simulated observations with the real observations. The sensitivity results in the previous section showed that the settings used for the TEST experiment provided stable filter performance; therefore, we follow the TEST experiment here with the moderate initial and resampling perturbation sizes and with 8000 particles.

Since the OSSE used the actual climate forcing in 2004 to 2007, we used the quality~~-~~-controlled MODIS LAI product of MCD15A3 for those years with flagged as “good quality”, “Terra or Aqua”, “detectors apparently fine for up to 50 %”, “significant clouds not present”, and “main method used with or without saturation”. We took the median of the LAI observations in the 10-km radius from the study site (62º 15' 18'' N, 129º 14' 29'' E). ~~The gridded MODIS LAI data have the~~

~~grid resolution of 1 km, so that~~There are a number of missing data in the quality-controlled MODIS data. Therefore, if the number of the data in the 10-km radius ~~area contains 314 grid points. If more~~is less than ~~14 grids are missing~~300, we ~~apply~~ set these data as the missing data for DA. Since the MODIS data resolution is 1 km, the 10-km radius area contains about 314 data. The observation error standard deviations are assigned to each LAI datum ~~from~~in the original ~~source~~MODIS product (Knyazikhin et al., 1999). We rely on the estimate of the observation error standard deviations, and ~~we~~ took the median of the error standard deviations~~.~~ in the same way as getting the LAI data. The observation error standard deviation is used in the particle filter when computing the likelihood function (Eq. 1).

The model-simulated NEE was validated with the field observation data at this AsiaFlux site (Ohta et al., 2001; 2008; 2014). The data was quality controlled by the steady-state test as indicated by the quality flag 0. Although the model simulates daily-average NEE, the field observation data represent instantaneous NEE every 30 minutes. The observation data are missing frequently, and it is not trivial to derive daily averages. Therefore, the raw data are compared with the DA results directly. This allows only a rough verification about whether or not the simulated NEE is in a reasonable range, but this is the only possible verification with an independent source.

~~4~~5.2 Results

Figures ~~6,~~ 7, 8, and ~~8~~9 show similar time series to Figs. 3, 4, and 5, respectively, but with the real MODIS LAI observations. Although the particles of the NODA experiments are widely spread, DA makes the particles much narrower (right) for all variables and parameters. With DA, the median of LAI is very close to the observations, within the range of the observation error standard deviations (Fig. ~~6~~7 a). The grass and forest LAIs are estimated separately (Fig. ~~6~~7 b, c), but there is no direct observation or other verification truth to compare with. This is similar to the model parameters (Fig. ~~7~~8) and other model variables (Fig. ~~8~~9) except for NEE, for which direct field observation data are available. As in the OSSE results, the range of uncertainties for NEE is reduced significantly by DA (Fig. ~~8~~9 c). Since the field observations are made instantaneously every 30 minutes, the observation values (red) appear to have a wider range. However, the SEIB-DGVM simulates only daily-average NEE, and it is not straightforward to compare the outputs from SEIB-DGVM with the field observations. We still find that the median of NEE becomes closer to the observations, particularly near the dormancy period. The simulated NEE generally stays within the reasonable range compared with the field observations. In general, the particle filter shows promising results with the real MODIS LAI data.

~~5~~6 Conclusion

We assimilated the satellite-based MODIS LAI data using a non-Gaussian ensemble DA system with the SEIB-DGVM based on the SIR particle filter approach. To the best of the authors' knowledge, this is the first study to assimilate the fine time-scale satellite data with an *individual-based* DGVM. We found that DA performed generally well both for OSSE and real-world experiment, so that the ~~newly developed DA system greatly reduced the~~ uncertainties of the state variables and model parameters~~.~~ were greatly reduced. Additional sensitivity experiments for OSSE revealed general robustness but some sensitivities to the nature run, initial and resampling perturbation sizes, and particle size, particularly for biomass.

Although we assimilated only LAI as a whole, the forest and grass LAIs were estimated separately. This suggests that the satellite-based DA reduce the uncertainties in the initial vegetation structure of the *individual-based* DGVM toward the simulation of the future vegetation change. Another notable results include that the model parameters of the *individual-based* DGVM were estimated successfully, and that the uncertainties in the unobserved model variables relevant to carbon cycle and vegetation states were also reduced significantly. Similarly to the previous studies with a *static* vegetation model (Stöckli et al., 2011) and a *non-individual-based* DGVM (Demarty et al., 2007), the results in the present study also suggest that LAI be the key to DA for phenology and carbon dynamics.

As a potential limitation, it is important to note that we have made strong assumptions in OSSE. For example, the only source of model imperfections was the model parameter uncertainties of the four parameters. It was also assumed that the observation error statistics were perfectly known. These conditions would have never been met in the real-world experiment.

As the first step, this study focused on the methodological development of the data assimilation system with SEIB-DGVM and estimated only four parameters of two PFTs using LAI observations at a single location. As a next step, more parameters and distributions of more diverse PFTs should be considered at different locations. Local-scale experiments can be performed in parallel for different locations since the satellite-based LAI observations are available globally. The simulation with the initial states and parameter sets obtained from the SEIB-DGVM-based DA system would be expected to improve the estimates of the carbon cycle changes ~~at~~over the ~~global domain~~globe.

Acknowledgements

The authors thank Hisashi Sato, the main developer of SEIB-DGVM, for useful discussions. The source code of SEIB-DGVM and the climate forcing data are available at http://seib-dgvm.com/. MODIS LAI product of MCD15A3 was retrieved from the online Data Pool, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (https://lpdaac.usgs.gov/data_access/data_pool). Carbon flux data was retrieved from AsiaFlux database (https://db.cger.nies.go.jp/asiafluxdb/).

References

Anderson, J. L., and Anderson, S. L.: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, Mon. Wea. Rev., 127, 2741–2758, 1999.

Atlas, R.: Atmospheric observations and experiments to assess their usefulness in data assimilation, J. Meteorol. Soc. Jpn., 75, 111–130, 1997.

Cheaib, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrêne, E., François, C., Gritti, E. S., Legay, M., Pagé, C., Thuiller, W., Viovy, N., and Leadley, P.: Climate change impacts on tree ranges: model intercomparison facilitates understanding and quantification of uncertainty, Ecol. Lett., 15, 533–544, 2012.

Demarty, J., Chevallier, F., Friend, A. D., Viovy, N., Piao, S., and Ciais, P.: Assimilation of global MODIS leaf area index retrievals within a terrestrial biosphere model, Geophys. Res. Lett., 34, L15402, 2007.

Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean. Dynam., 53, 343-367, 2003.

Garreta, V., Miller, P. A., Guiot, J., Hély, C., Brewer, S., Sykes, M. T., and Litt, T.: A method for climate and vegetation reconstruction through the inversion of a dynamic vegetation model, Clim. Dynam., 35, 371-389, 2010.

Gordon, N. J., Salmond, D. J., and Smith, A. F. M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation, IEE Proc. F, 140, 107-113, 1993.

Harris, I., Jones, P.D., Osborn, T.J., and Lister, D.H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, Int. J. Climatol., 34, 623-642, 2014.

Hartig, F., Dyke, J., Hickler, T., Higgins, S. I., O'Hara, R. B., Scheiter, S., and Huth, A.: Connecting dynamic vegetation models to data – an inverse perspective, J. Biogeogr., 39, 2240–2252, 2012.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, Bull. Amer, Meteor. Soc., 77, 437-471, 1996.

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S., and Schimel, D. S.: Ecological forecasting and data assimilation in a data-rich era, Ecol. Appl., 21, 1429–1442, 2011.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J. P., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E. N., Köstler, C., Mathieu, P. -P., Pinty, B., Reick, C. H., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., Terwisscha van Scheltinga, A., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, J. Geophys. Res. Biogeosci., 118: 1414-1426, 2013.

Kato, T., Knorr, W., Scholze, M., Veenendaal, E., Kaminski, T., Kattge, J., and Gobron, N.: Simultaneous assimilation of satellite and eddy covariance data for improving terrestrial water and carbon simulations at a semi-arid woodland site in Botswana, Biogeosciences, 10, 789-802, 2013.

Knyazikhin, Y., Glassy, J., Privette, J. L., Tian, Y., Lotsch, A., Zhang, Y., Wang, Y., Morisette, J. T., Votava, P., Myneni, R.B., Nemani, R. R., Running, S.W.: MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) product (MOD15) algorithm theoretical basis document, http://eospso.gsfc.nasa.gov/atbd/modistables.html, 1999.

Kolari, P., Pumpanen, J., Kulmala, L., Ilvesniemi, H., Nikinmaa, E., Grönholm, T., Hari, P.: Forest floor vegetation plays an important role in photosynthetic production of boreal forests, Forest Ecol. Manag., 211, 241-248, 2006.

Murray-Tortarolo, G., Anav, A., Friedlingstein, P., Sitch, S., Piao, S., Zhu, Z., Poulter, B., Zaehle, S., Ahlström, A., Lomas, M., Levis, S., Viovy, N., and Zeng, N.: Evaluation of land surface models in reproducing satellite-derived LAI over the high-latitude Northern Hemisphere. Part I: Uncoupled DGVMs, Remote Sens., 5, 4819-4838, 2013.

Ohta, T., Hiyama, T., Tanaka, H., Kuwada, T., Maximov, T.C., Ohata, T., and Fukushima, Y.: Seasonal variation in the energy and water exchanges above and below a larch forest in eastern Siberia, Hydrol. Process., 15, 1459-1476, 2001.

Ohta, T., Maximov, T. C., Dolman, A. J., Nakai, T., van der Molen, M. K., Kononov, A. V., Maximov, A. P., Hiyama, T., Iijima, Y., Moors, E. J., Tanaka, H., Toba, T., and Yabuki, H.: Interannual variation of water balance and summer evapotranspiration in an eastern Siberian larch forest over a 7-year period (1998-2006), Agr. Forest Meteorol., 148, 1941-1953, 2008.

Ohta, T., Kotani, A., Iijima, Y., Maximov, T.C., Ito, S., Hanamura, M., Kononov, A.V., and Maximov, A.P.: Effects of waterlogging on water and carbon dioxide fluxes and environmental variables in a Siberian larch forest, 1998–2011, Agr. Forest Meteorol., 188, 64-75, 2014.

Peng, C.: From static biogeographical model to dynamic global vegetation model: a global perspective on modelling vegetation dynamics, Ecol. Model., 135, 33-54, 2000.

Peng, C., Guiot, J., Wu, H., Jiang, H., and Luo, Y.: Integrating models with data in ecology and palaeoecology: advances towards a model–data fusion approach, Ecol. Lett., 14, 522–536, 2011.

Quaife, T., Lewis, P., De Kauwe, M., Williams, M., Law, B. E., Disney, M., and Bowyer, P.: Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter, Remote Sens. Environ., 112, 1347-1364, 2008.

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., Dietze, M. C., Dragoni, D., Garrity, S. R., Gough, C. M., Grant, R., Hollinger, D. Y., Margolis, H. A., McCaughey, H., Migliavacca, M., Monson, R. K., Munger, J. W., Poulter, B., Raczka, B. M., Ricciuto, D. M., Sahoo, A. K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J., and Xue, Y.: Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis, Glob. Change Biol., 18, 566–584, 2012.

Sato, H., Itoh, A., and Kohyama, T.: SEIB–DGVM: A new Dynamic Global Vegetation Model using a spatially explicit individual-based approach, Ecol. Model., 200, 279-307, 2007.

Sato, H., Kobayashi, H., and ~~Ise, T.: Effect~~Delbart, N.: Simulation study of ~~plant dynamic processes on African~~ the vegetation ~~responses to climate change: Analysis~~structure and function in eastern Siberian larch forests using the ~~spatially explicit~~ individual-based ~~dynamic global~~ vegetation model ~~(~~SEIB-DGVM~~), J. Geophys. Res., 117, G03017, 2012~~, Forest Ecol. Manag., 259, 301-311, 2010.

Sato, H., and Ise, T.: Effect of plant dynamic processes on African vegetation responses to climate change: Analysis using the spatially explicit individual-based dynamic global vegetation model (SEIB-DGVM), J. Geophys. Res., 117, G03017, 2012.

Sato, H., Kobayashi, H., Iwahana, G., and Ohta, T.: Endurance of larch forest ecosystems in eastern Siberia under warming trends, Ecol. Evol., 6, 5690–5704, 2016.

Smith, B., Prentice, I. C., and Sykes, M. T.: Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space, Global Ecol. Biogeogr., 10, 621–637, 2001.

Stöckli, R., Rutishauser, T., Baker, I., Liniger, M. A., and Denning, A. S.: A global reanalysis of vegetation phenology, J. Geophys. Res., 116, G03020, 2011.

Takagi, K., Hirata, R., Ide, R., Ueyama, M., Ichii, K., Saigusa, N., Hirano, T., Asanuma, J., Li, S-G., Machimura, T., Nakai, Y., Ohta, T., Takahashi, Y.: Spatial and seasonal variations of CO2 flux and photosynthetic and respiratory parameters of larch forests in East Asia, Soil Sci. Plant Nutr., 61, 61-75, 2015.

Williams, M., Schwarz, P. A., Law, B. E., Irvine, J., and Kurpius, M. R.: An improved analysis of forest carbon dynamics using data assimilation, Glob. Change Biol., 11, 89–105, 2005.

Zeng, H., Jia, G., Epstein, H.: Recent changes in phenology over the northern high latitudes detected from multi-satellite data, Environ. Res. Lett., 6, 1-11, 2011.

Zhao, J., Zhang, H., Zhang, Z., Guo, X., Li, X., Chen, C.: Spatial and temporal changes in vegetation phenology at middle and high latitudes of the Northern Hemisphere over the past three decades, Remote Sens., 7, 10973-10995, 2015.