

The authors would like to thank the Reviewer for the very helpful comments and suggestions. We will take them into consideration in the revised manuscript and we will insert the flow-chart as suggested. Moreover, we will modify some figures according to the Reviewer's suggestion. Finally, we would like to respond to some individual points:

>What is about the uncertainty of different model-data fusion approaches

Model-data fusion (MDF) relies on the combination of models with observational constrains through an optimization approach. In this way, model parameters, model states and their respective uncertainties can be estimated, conditional on the data.

MDF requires assumptions about the data and how to define the "cost function" (e.g. likelihood function, etc.) that is the basis of optimization. Different assumptions lead to different cost functions and therefore to different optima.

Here, we used the least square cost function (i.e. sum of squared error) assuming that data are homoscedastic and with gaussian errors. Different cost functions can be used with different MDF approaches that can be classified as follow (Williams et al., 2009): 1) "Global search" algorithms (e.g. Metropolis or Metropolis Hastings, genetic algorithms); 2. "Gradient-descent" algorithms (e.g. Quasi-Newton methods, Levenberg-Marquardt), "Sequential" algorithms (e.g. Kalman filter). An intercomparison of MDF algorithms is presented in Fox et al., (2009).

In this work we chose a "Global search" method (i.e. Metropolis algorithm) although the same cost function could be used with other approaches (down-gradient methods). Getting stuck in a local minimum is an important potential source of uncertainty and there is no algorithm yet developed that is guaranteed to find the global optimum. Here, we ran the optimizations several times, with different starting points, to verify that we were finding the global optimum. Metropolis is known to be good at avoiding being trapped in local minima because allows for changes in the searching direction. For these reasons, for phenological research Metropolis is a widely used method (e.g. Chuine et al, 1998, Schaber and Badeck 2003 among the others).

Page 888 line 12 and 13 authors wrote" around 2.0" and in line 23 "around 6.0" whereas in the table 3 you put <2.0 and <6.0. Please explain those numbers and the use of those numbers in more detail.

We have included in the text a description of the $\Delta AICc$ values as follow: if the difference, $\Delta AICc$, is zero or very small, both models are essentially equally likely to be the best model. If $\Delta AICc$ is ~ 2.0 , then the model with the lower $AICc$ is almost three times more likely to be best. If, however, $\Delta AICc$ is ~ 6.0 , then the model with the lower $AICc$ is about 20 times more likely to be best. Then, in Tables 3 we classified each model according to the difference between its $AICc$ and the one of the best model. Models were classified as models with $\Delta AICc < 2$ (models considered as equal or up to 3 times worse than the best), models with $2 < \Delta AICc < 6$ (essentially models with very low probability to be considered good as the best) and model with $\Delta AICc > 6$ (poor models compared to the best). For more detail on the use of the AIC we refer to Anderson et al., (2000).

Page 889 line 22 explain shortly the Sen's slope estimation

The Sen's slope estimator (Sen, 1968) is a non-parametric procedure for estimation of trend magnitude for univariate time-series with monotonous trends and no seasonal or other cycles in the data (no autocorrelation in the time series). However, Yue et al., (2002) proposed a method to minimize the influence of autocorrelation on the ability to detect trends with non parametric methods.

Non-parametric methods are preferred to parametric methods (i.e. regression analysis: regression slope) because no assumption is made regarding the probability distribution of data while parametric tests assume normally distributed variables (homogeneous variance) and homoscedastic.

The second advantage is its robustness to outliers or to abrupt breaks due to inhomogeneous time series (Hirsch et al. 1982).

The Sen's Slope estimates the slope of a time series computing slopes for all the pairs of ordinal time points and then uses the median of these slopes as an estimate of the overall slope. Here, we used the Sen's slope estimator to assess the slope (or magnitude) of trends in bud-burst dates.

We will add this clarification and a short description of the method in the Materials and Methods section providing to the readers more references.

Fig. 2a? Explain more why the bud burst date overlap sometime e.g. in the 2070s. Where the warmer scenario show a delayed onset!?

This can be explained by the convergence of temperatures simulated by the two different scenarios and the interaction between photoperiod limitation and temperature for the warmer scenario (A1fi). Figure 2a shows that around 2070 the temperature of the Scenario A1fi decreases while the scenario B1 report particularly high mean temperatures in that period. The convergence of temperatures modeled by the two scenarios is very likely the reason of the overlapping between bud-burst dates simulated for the two scenarios.

Fig. 3 legend is not understandable. I do not know what represent a and what are the bars for?

We will improve the legend of this figure. Grey bars represent the average budburst dates across species (Fig 3a) and the average budburst for each species (Fig 3b).

References:

Anderson, D.R., Burnham, K.P., Thompson, W.L., 2000. Null hypothesis testing: problems, prevalence and an alternative. *J. Wildlife Manage.* 64, 912–923.

Chuine, I., Cour, P., and Rousseau, D. D. (1998). Fitting models predicting dates of flowering of temperate-zone trees using simulated annealing, *Plant Cell and Environment*, 21, 455-466.

Fox, A., Williams, M., Richardson, A.D., Cameron, D., Gove, J.H., Quaife, T., Ricciuto, D., Reichstein, M., Tomelleri, E., Trudinger, C.M., Van Wijk, M.T. (2009). The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data, *Agricultural and Forest Meteorology*, 149, 10, 1, 1597-1615.

Hirsch, R.M.; Slack, J.R.; Smith, R.A. (1982). Techniques of trend analysis for monthly water quality data. *Water Resources Research* **18** (1): 107–121, doi:10.1029/WR018i001p00107. 58201.

Schaber, J., and Badeck, F. W. (2003). Physiology-based phenology models for forest tree species in Germany, *International Journal of Biometeorology*, 47, 193-201.

Sen, P.K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association* **63**: 1379–1389, JSTOR 2285891, MR02

Yue, S, Pilon, P, Phinney, B, Cavadias, G (2002). The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes* 16:1807–1829.