Geosci. Model Dev. Discuss., 4, C854–C860, 2011
www.geosci-model-dev-discuss.net/4/C854/2011/

**Geoscientific
Model Development
Discussions**

Interactive comment on "Use of agricultural statistics to verify the internannual variability in land surface models: a case study over France with ISBA-A-gs" *by* J.-C. Calvet et al.

Anonymous Referee #2

Received and published: 14 October 2011

This paper utilizes agricultural statistics to evaluate the ability of the ISBA-A-gs model to capture interannual variability in biomass and crop yields. The authors find that the parameter governing maximum water holding capacity (MaxAWC) is a strong control on the amplitude of interannual variability in biomass yield. Despite not explicitly representing specific aspects of annual crop production that differ from natural vegetation, the model is able to capture a significant level of interannual variability for crops in a handful of sites, particularly when the MaxAWC parameter is optimized for each site. However for several crop sites the model cannot produce a significant correlation for any parameter values indicating that the model lacks the structure necessary to describe the phenomena contributing to yield variability at those sites. Grassland

variability is more easily captured by the model but the R^2 value sill benefits from optimized MaxAWC. Lower optimal values of MaxAWC for grasslands correspond to an independent estimation of actual water holding capacity indicating that the tuned parameter values correspond to real differences in the characteristics of crop and grassland systems. This paper indicates the importance of soil water capacity in governing annual time scale variability in ecosystem-climate interactions. It also establishes a set of criteria for evaluating models' ability to represent interannual variability. I recommend the paper be accepted for publication subject to the concerns and calls for clarification listed below.

General Comments

1) Limited Sensitivity Analysis – The authors conducted a preliminary sensitivity analysis to identify which parameters to tune in order to optimize R^2 to yield data at each site. However they chose to do this only for a single region with a single crop – Rye – which yielded the best R^2 using default parameter values. They do not justify why this particular crop and region were chosen and do not discuss the possibility that parameter sensitivity would vary as a function of vegetation type, crop type and/or region. Furthermore they conducted a one-at-a-time sensitivity analysis, which fails to capture interactions among model parameters. While a more sophisticated sensitivity analysis examining interactions may be beyond the scope of the study, the limitations of the approach used should be discussed more thoroughly and the specific region and crop chosen should be justified.

2) Tuning one feature of the model at the possible expense of other features – In addition to altering the character of interannual aboveground biomass yield variability, the two parameters tuned affect the mean level of aboveground biomass as well. For instance, it can be seen from Figs 7 and 8 that the reducing MaxAWC from 200mm to 50mm leads to an overall reduction of mean aboveground biomass on the order of 35% (from $\sim$1.5 kg/m^2 to $\sim$1 kg/m^2). Given that the authors find variation in the optimal MaxAWC value across sites for the same vegetation and crop types, an ob-

vious question is whether the optimized parameters improve the model's correlation to the geographic pattern of mean yield statistics. If this metric also improves, then a stronger case could be made that the model is capturing a meaningful feature of agricultural yield variability (explaining both temporal and spatial yield variability).

3) Use of modeled benchmarking data for grasses but not crops calls into question conclusions regarding differences in the model's ability to capture crop vs. grass inter-annual variability – Very different sorts of data are used to benchmark crops vs. grass yields. It is not clear to what extent the better model fit to grass data is a function of the nature of the data itself. In particular shared model biases could contribute to improved fit for grasses. A better case should be made for why comparison to ISOP data is useful and why crop and grassland Rˆ2 values are not directly comparable.

Specific Comments

Page 1479, line 15: The case is made that due to uncertainty in remotely sensed LAI data, in-situ biomass measurements are needed. However in line 6, three studies are sited that assimilate satellite LAI data. Do these studies suffer from the deficiency in LAI measurements mentioned here? In order to motivate the use of in-situ data in this study, more clarification is needed of when LAI data are useful and when they are not.

Page 1481, line 1: It would be useful to state more specifically how the photosynthesis parameterization differs from the standard Farquhar model?

Section 2.3.3: It is not clear from this section whether the Agreste data, the ISOP data or both will be used for benchmarking. It would be helpful to discuss the advantages or disadvantages of using one vs. the other and explain why the authors chose to use both.

Page 1485, lines 6,7: Are the values chosen for MaxAWC and gm realistic? A discussion is needed of how these ranges were chosen and why.

Page 1485, line 8: A discussion of the method used to find the optimal parameter

values at each site should be included here. It is not clear from the present description why the simulation was repeated 48 times at each site with the various parameter values described. Since the goal is to optimize Rˆ2 at each site, that goal should be stated in the methods.

Page 1485, paragraph beginning at line 20: The analysis presented in this paragraph seems arbitrary and doesn't flow well with the rest of the paper, partly because results are being presented in the methods section. Figure 4 is mentioned but not discussed. Its relevance to the analysis presented elsewhere is not clear. It is not clear why the values of MaxAWC are set at their specific levels and why they differ for crops and grassland? It is also not clear why non-default values are chosen for the gm parameter for grasslands? Why does the gm parameter differ between managed and unmanaged grasslands? Why is this single province chosen? I would recommend relocating and revamping this whole paragraph so that the methodological choices are explained better and results are tied into the broader objectives of the paper or dropping it along with Figures 4 and 5.

Page 1486, lines15-20: The beginning of this paragraph describes the methods for the preliminary sensitivity analysis. It would make more sense to describe these methods in the methods section before describing the method used to find the optimal values for the two parameters chosen.

Page 1487, line 27: Figures 9 and 10 add little information beyond table 2 except for showing the spatial pattern of Rˆ2 significance. However the spatial pattern is not discussed. If the spatial pattern is not important to the objectives of the paper these figures could be dropped. Otherwise the significance of the pattern should be discussed

Section 4.1: The change in the number of sites with significant Rˆ2 values is used as a metric to judge the sensitivity of model fit to fixing versus optimizing the two model parameters MaxAWC and gm. For croplands, a high sensitivity is found but for grasslands a large number of sites are significantly correlated to the model regardless of whether

MaxAWC is held constant or set at its site-specific optimal value. The authors conclude that croplands are more sensitive than grasslands to the value of MaxAWC and go on to draw conclusions about the differences between cereal versus forage pea crops and between managed versus unmanaged grasslands. However, the metric based on the number of significantly correlated sites ignores changes to the Rˆ2 value obtained by fixing versus holding the MaxAWC value constant. As can be seen from table 2, the improvements in Rˆ2 among the significantly correlated sites going from fixed gm and fixed MAXAWC to fixed gm with optimal MaxAWC (from the 2nd last to 3rd last row of table 2) is similar for cereals and unmanaged grasslands ($\sim$.1 improvement in Rˆ2). Examining the change in Rˆ2 tells a different story than the number of significant sites. Thus, while the model describes some amount of the interannual variance in grassland yields regardless of whether MaxAWC is fixed or varied, a seemingly significant additional amount of variance is described with the optimal MaxAWC values. Through this lens, the model is just as sensitive to MaxAWC for unmanaged grasslands as cereals and the conclusions drawn from this section are incorrect.

Section 4.2 – It is also likely that features of crop production not explicitly represented by the model are changing over time and this contributes to the poor Rˆ2 values for crop sites. Some of these limitations are listed in the intro on page1480, line 7 and the poor fit of the model to crop sites is not surprising given these issues. The discussion in this section would benefit from mentioning these limitations. As it is currently written, this section seems to attribute the poor fit for some crop sites solely to geogeaphic variability within departments.

Page 1490, line 6: Presumably soil type varies among grasslands as well. Do the authors mean to say that there is more variability of soild type among crops? If so, this needs to be clarified. If not, variation in soil type would not explain why crops are more heterogeneous than grasslands.

Page 1492, line 10: Do the authors mean to say that the siting of croplands on better soils explains 1) why MaxAWC is lower for grasslands within the INRA data or 2) why

the optimized model estimate of MaxAWC falls at the lower end of the INRA range for grasslands. 1) makes sense but 2) does not. From the language it is unclear which is meant.

Fig 7 and 8: The vertical axes are not equal on all figures which masks the effect of changing MaxAWC on the mean aboveground biomass yield.

Technical Comments:

Page 1479, line 3 “permits to simulate” is unclear grammatically, perhaps “permits the simulatio of”

Page 1479, lines 16-19: This sentence is unclear.

Page 1482, lines 5,7: “growing at springtime”, “growing at summertime” is grammatically unusual, perhaps “spring crops” “summer crops” is better

Page 1482, lines 14-18: It is not clear whether the dots in figure 1 are departments or SAFRAN gridcells

Page 1483, line 21: by “all these crops” do the authors mean to say that each crop considered alone covers a significant fraction of each of the 45 departments or that total crop area covers a significant fraction in those departments with some crops not present in all departments? Please clarify

Page 1484, line14: It is unclear what the ISOP regions were used for in this sentence

Page 1490, line 1: change “that” to “so”

Page 1491, line 12: Please explain what the 3-layer force-restore soil model is

Page 1492, lines 14-15: Please explain this analysis more clearly or drop it. Table 2 : some vertical lines separating different groups of columns would help

Fig 6: gc and thetac are listed in reverse order