Geoscientific

Model Development

Discussions

GMDD

Interactive

comment

Interactive comment on "The impacts of data constraints on the predictive performance of a general process-based crop model (PeakN-crop v1.0)" *by* Silvia Caldararu et al.

Silvia Caldararu et al.
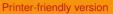
Matthew.Smith@Microsoft.com

Received and published: 30 January 2017

Response to Reviewer 1

First of all we would like to thank the reviewer for the very thorough analysis and constructive comments. We believe that addressing these comments will add value to our work. We would also like to thank him for signing the review, which is an important part of the open review process.

Below are our responses to the specific comments made, including how we intend to address them with specific changes in the manuscript.

Comment from reviewer 1: "A first criticism is that there seems to be some ambiguity

about the exact objective of the predictions. In the introduction the authors speak of application to a "generic farm location", whereas evaluation is based on comparison with regional yields. Prediction for a farm, with uniform management, is quite different than prediction for a region. The paper seems more oriented toward regional prediction, since county yields are one of the data sets used as input and are also used for evaluation. On the other hand, the landcover input data was aggregated to 3km by 3 km pixels, which is generally intermediate between farm and county scale. In any case, it is essential to clarify the spatial scale of interest."

Response from Authors 1: The model is intended, in an ideal situation, to be used at the local scale. However, given the lack of availability of the data, particularly the yield information, this is difficult. The model as applied here is indeed at an intermediate scale, as the eddy flux data is always field scale, while the crop yield is regional. We agree that this discrepancy between the different spatial scales in the data has not been discussed sufficiency and we have now added this o the discussion.

Proposed changes to Manuscript 1: We will add to the end of the introduction a clarifying statement about the scale of application, such as

"While our study is part of a boarder scientific objective to enable more accurate field scale predictions, the lack of availability of field scale datasets to train and validate our model means that the scale of model evaluation for our study here is a mix of field (flux tower) and regional scales (county and country level for yield estimates and 3 by 3 km scale for photosynthetic activity)."

And we will add a dedicated paragraph to the discussion about this issue

"One additional complication is the different spatial scales of the three datasets - while the eddy covariance data is at the scale of the flux tower footprint, which can be seen as equivalent to the individual field scale, the fAPAR and yield data correspond to larger scales (county and country level for the yield data and a 3 by 3 km scale for the fAPAR data). The assumption behind our analysis is that the conditions at field

scale are representative of the regional scale, so that there would be no discrepancy between model predictions at these different scales. This is obviously a source of error, especially at the wheat sites in Europe, which will be located over a much more heterogeneous landscape. Further sources of data at the field scale would be required to identify the model error caused by the discrepancy in spatial scales."

Comment from Reviewer 2: "Much of the evaluation is based on comparing the model using data constrained parameters to the model with prior parameters. This is not a very interesting comparison. The prior parameters were chosen quite arbitrarily by the authors to represent essentially a total lack of information about the parameter values. The fact that adding some information improves the situation is hardly surprising."

Response from Authors 2: This is indeed a not very surprising result. However, we have chosen the model with prior parameters as a benchmark as this offers a worst case scenario of parameter values and can be used to show relative improvement in model error and uncertainty when using the different datasets.

Proposed changes to Manuscript 2: We will include a statement in the introduction to reflect on the fact that we fully expect data constraining the model to improve predictive accuracy but that we're more interested in how much the predictions are improved. We propose to insert just after we state the paper aims:

"We expect the qualitative answer to the first question to be that utilising empirical data does enable the model to make better predictions because that's a typical outcome of our parameter estimation approach. However we are more interested in the quantitative answer; i.e. how much? For example, the generation of a model that could make extremely precise and accurate predictions would suggest that data-constraining general models with the datasets we identify could provide an extremely useful tools for agricultural predictions and forecasts. Alternatively, the generation of a model that makes very imprecise predictions would suggest that more data collection and model improvement is needed for the model to have practical applications."

Interactive
comment

Comment from Reviewer 3: "A much more relevant comparison would be between the data constrained model and long term average county yields. Does the model do better than simply assuming that the future is like the aver-age of the past? This is analogous to comparing climate forecasts with climatology.

Response from Authors 3: This is a very good idea, however, we consider that such a comparison with long terms yields is premature. We have not performed this analysis for two reasons: inadequate data and the mix of spatial scales. As mentioned our ultimate objective is to enable better field scale predictions, however the only field scale data we have, the flux datasets, typically only cover one year and different sites cover different years. This makes comparing model predictions to a data average for one site almost meaningless. Our other datasets do have longer, more complete time series. The fAPAR time series reflects vegetation dynamics over a wider spatial area of vegetation than an individual field, including non-agricultural vegetation and so itself is only a coarse indicator of field level dynamics - we are therefore not interested in how well the model captures the complete seasonal dynamics of that dataset. It would be interesting to understand whether a model such as that developed here can out-perform historical data at predicting large scale yield variations. This should be dealt with in a dedicated future piece of work because a number of other analyses should be conducted alongside such a comparison to make it insightful. For example we currently do not account for variations in environmental drivers and agricultural practices at smaller spatial scales across a region that would also help to explain yield variations observed at the larger spatial scales. These could be accounted for by simulating our model for every relevant field within the region and aggregating the predictions. A dedicated comparison of the predictions of our model with those of statistical averages over historical data is a natural avenue for future research.

Proposed changes to Manuscript 3: We have thought hard about how to implement this request but, as explained above, evaluation for each dataset raises further issues. We have therefore decided to defer this comparison to another study, ideally one in which

Interactive
comment

Discussion paper

our model can be trained and evaluated against a richer dataset of location specific data. We propose to include a brief statement in the discussion (under future data needs) stating

"If this model, or any other similar process-based data constrained crop model, is to be used for scientific purposes to understand the response of crops to climate across the globe, the ideal data would be a global data set, such as space-based vegetation observations, combined with high quality field level data that would ideally include growth timeseries, final grain yield and information about agricultural practices. However, if the model is to be used for agricultural purposes, to aid decision making at the local level, high quality field level data would be sufficient. A valuable evaluation in such studies, not conducted here for brevity and due to a lack of location-specific data, would be to compare the predictive accuracy of the model against the predictive accuracy of a statistical average over the data. Such an analysis would reveal whether and how much benefit is gained by using a data constrained model for predictions"

Comment from Reviewer 4: "Another aspect of the evaluation is that uncertainty intervals are given for each prediction. This is extremely informative and pertinent, and is a very valuable addition to the comparison between the mean prediction and observed values. However, the uncertainty results need to be discussed more thoroughly. For example, it seems that the 95% intervals for yield cover all historic yields at most sites (Fig. 3a). Surely this uncertainty is so large as to render the results useless. More discussion is required here."

Response from Authors 4: We agree that the uncertainty in yield predictions is very large, making it largely unsuitable for predicting interannual variation in crop yields. However, as the reviewer also points out, the model parametrisation method presented in our paper has the advantage that is does provide an uncertainty estimate. Hence, any future improvements in model performance, especially through added sources of data are easily quantifiable.

Proposed changes to Manuscript 4: We will expand an existing paragraph in the discussion to read

"Model uncertainty is difficult to compare with previous crop modelling studies, as models with fixed parameter values do not often provide uncertainty estimates. In fact, providing uncertainty values for all model variables and parameters is one of the advantages of a data constrained model. In the current model, uncertainty is highest at the start of the season for all variables but decreases rapidly and final yield uncertainty is much lower. This is due to thresholds: abrupt changes from one growing stage to another when small differences in parameters can lead to large differences in resulting variables. It is, however, important to note that the uncertainty in our yield predictions remains high and the model in its current form is unlikely to provide accurate predictions for practical applications without the addition of new data (Section 7.4). We have however shown that the use of three different data types does reduce prediction uncertainty - pointing to an avenue for future model improvement."

Comment from Reviewer 5: "The uncertainty calculations are based on propagating uncertainty in the parameter values through the model. It is not clear if residual error is included when calculating uncertainty intervals or not. It should be."

Response from Authors 5: The model uncertainty is calculated by sampling parameters from the posterior distributions and then computing model predictions with the sampled parameters, which results in a distribution of model predictions from which we can calculate a predicted mean and confidence intervals. We realise that this was insufficiently well explained clearly in the paper.

Proposed changes to Manuscript 5: We will include a fuller explanation of how we propagate uncertainty in the parameter values of the model:

"To calculate uncertainty for the model predictions we sample parameter values from their respective posterior distribution and compute predictions with each parameter combination, which results in a corresponding distribution of model predictions. We

report this prediction distribution uncertainty using 95th percent confidence intervals. This predicted distribution does not include the prescribed or inferred uncertainty about observations, $\sigma\_(x,D)$, our predicted distributions correspond to the state being predicted and not the observations of that state."

Comment from Reviewer 6: "Also, there are other sources of uncertainty than the parameters which might be quite important, in particular uncertainty about management practices. This should at least be discussed."

Response from Authors 6: We agree and will make a change to the manuscript to reflect this.

Proposed changes to Manuscript 6: We will include a discussion of model uncertainty related to management practices, in particular sowing and harvest dates as well as fertilizer input, such as

"In addition to the three datasets used for parametrisation, the model also requires input data in the form of sowing and harvest dates and fertiliser inputs. Additional uncertainty is associated with these datasets which is not available nor accounted for in our analyses. For example, the crop calendar (Sacks et al., 2010) and Nitrogen Fertilizer Application (Potter et al., 2010) datasets are global data collections that will imperfectly represent the value for any given location. Alternatives to these global datasets would be to use location-specific data, or to infer the values. Location specific data has the advantage of more accurately reflecting the situation at a given site and would therefore be useful when the model is applied at the field scale, but such data is unlikely to be available for all sites. Successful inference of the values would depend on if there is enough information in the datasets used to infer the model parameters. If there is inadequate data then there would be excessive degrees of freedom for inference, leading to the wrong parameter values begin inferred and the model performing poorly in novel situations. Therefore, the decision whether to obtain more data or infer unknown quantities in future applications of our model and inference framework depends on the

Interactive comment

[Printer-friendly version](#)

[Discussion paper](#)

data availability and the intended scales of application."

Comment from Reviewer 7: "It seems that the likelihood used here for the Bayesian estimation assumes that all data are independent. This is of course almost certainly false for time series data. Taking non independence into account by dividing by the number of measurements is only a very crude approximation."

Response from Authors 7: The division by the number of measurements is not meant to account for non independence, rather it accounts for the different number of data points in each time series so that each dataset is given equal importance for fitting purposes. We agree that the timeseries data is most likely not independent but the independence assumption is often used when fitting models to eddy covariance flux data in order to simplify the formulation of the likelihood function.

Proposed changes to Manuscript 7: We will include the point made by the reviewer after we describe our Likelihood function.

"Note that with this definition of the likelihood we are treating every data point as independent, that is the likelihood of a value at time ,t, is treated independently from the likelihoods at preceding times. This is only an approximation but is commonly used in parameter estimation studies because the additional mathematical and computational complexity of accounting for non-independent data."

Comment from Reviewer 8: "More detail about the model would be helpful. How exactly is the date of flowering calculated? According to the text, the switch from vegetative to reproductive growth occurs when increased vegetative fractions would not result in an overall increase in growth rate. Is this calculated day by day or is there some averaging over environmental conditions to ensure that the plant doesn't respond to conditions on one specific day?"

Response from Authors 8: The start of the reproductive growth is calculated using an average of environmental conditions for the peak carbon criteria. Due to the continuous

Interactive comment

decrease in soil N driven only by plant uptake and the relative simplicity of the model which does not take temperature and moisture into account for nutrient uptake, this averaging is not necessary for the peak N criteria. We will clarify this in the text

Proposed changes to Manuscript 8: We will include a statement in the methods about this:

"The peak nitrogen condition is achieved when an increase in root mass does not result in an increase in nitrogen uptake. This condition is achieved in nitrogen limited environments where the nitrogen available in the soil is depleted through the period of vegetative growth. This assumption can be considered valid in agricultural systems where the major nitrogen input into the system during the growing period comes solely from agricultural fertilisers. Soil nitrogen decays monotonically through the season in our model due to the simplicity with which we model nitrogen uptake and so detecting the peak nitrogen condition is straightforward. Similarly, the peak carbon flowering condition is triggered when the addition of aboveground biomass would not lead to an increase in net carbon gain, due to self-shading in the canopy. To calculate the peak carbon trigger we use the environmental variables averaged over p days, to avoid flowering being triggered by short-term environmental fluctuations. We infer p alongside the other parameters in our model."

Comment from Reviewer 9: "What exactly are the management inputs required for the model? The authors mention sowing and harvest date, but aren't sowing density and fertilizer inputs also required? The required management information should be made clear, as well as the sources of this information."

Response from Authors 9: This is correct, in addition to sowing and harvest dates, fertilizer input and planting density are also required, as well as irrigation regimes for future versions of the model which would take into account water limitation. We will include more information on this, including sources of the values used on the study.

Proposed changes to Manuscript 9: We propose to include these details in the methods

"Fertilizer input data were obtained from the published site descriptions (see Table 1 for references) or from the Nitrogen Fertilizer Application database (Potter et al. 2010). The model implemented in this study does not require any additional information on irrigation or soil properties."

Comment from Reviewer 10: "There is also no information on soils. Apparently this information is not needed here thanks to the assumptions that there is no water limitation, and that initial soil N is negligible compared to fertilizer N. In general, however, it will be necessary to have soil information."

Response from Authors 10: Additional soil information would be essential for versions of the model that include water limitation and very important if the N uptake model was made more complex, for example if we included information about rooting depth or different forms of available nitrogen.

Proposed changes to Manuscript 10: We will include a paragraph on this point in the discussion

"The model in the version presented in this paper does not include any water limitation to growth due mainly to a lack of data constraint on any water related parameters, as we found that latent heat data from EC towers is not sufficient. Below-ground measurements of not only root growth but also soil water properties would again provide some of the necessary information. Such belowground data, especially if supplemented by nutrient concentrations can also help constrain a more complex version of the nitrogen uptake scheme, which could be improved to include more explicit soil-plant interactions and additional processes such as biological nitrogen fixation for legumes."

Comment from Reviewer 11: "The authors suggest that the model could be tested by comparing different model structures. Perhaps more useful would be to test the model proposed here with much more detailed input data, in order to reduce the data as a source of error and thereby isolate the amount of error due to the model."

Response from Authors 11: Comparing different model structures and including better quality data are two different, but valid ways of identifying sources of error in our analysis and the two combined would give the most accurate analysis.

Proposed changes to Manuscript 11: We believe this point is already covered by one of the final statements in our Discussion

"If this model, or any other similar process-based data constrained crop model, is to be used for scientific purposes to understand the response of crops to climate across the globe, the ideal data would be a global data set, such as space-based vegetation observations, combined with high quality field level data that would ideally include growth timeseries, final grain yield and information about agricultural practices. However, if the model is to be used for agricultural purposes, to aid decision making at the local level, high quality field level data would be sufficient."

Comment from Reviewer 12: "P7 L23-24. "given the model" needs to be omitted"

Proposed changes to Manuscript 12: This will be corrected.

―――――――――――――――――――――――