General comments

Rabin et al. study the effects of different land use practices on altering the fire regime. Following previous studies they find that land use may have a distinct effect on both the total area burnt as the timing of fire activity. Moreover, they separate fires occurring in natural vegetation from pasture burning and provide a first global estimate of associated burned area and carbon emissions. The manuscript is well written and presents interesting original work of interest for publication in Biogeosciences. However, their methods are a little confusing and some of the model results are not well presented. Moreover, the limitations of their statistical model are not fully discussed. I would recommend publication in Biogeosciences only after sufficiently considering the comments below.

Specific comments

Methods

Please add numbering to the equations.

In equation 2, (line 4 on page 10822) you optimize $\hat{F}_{k,i}$ where k stands for c, p or o. However, you minimize the sum of squared errors for each analysis region (or group of grid cells), let's say r, and not for each grid cell i. It may thus be better to use something like $\hat{F}_{k,r}$ where r is the analysis region (or cluster of grid cells), and k may be c, p or o.

To continue, in your first equation on page 10821 $F_{k,i}$ stands for the "fraction of that land-use type that burned in that grid cell". However, when optimizing $\hat{F}_{k,i}$ in the next equation (first equation on p. 10822) $\hat{F}_{k,i}$ values represent a different thing: the slopes of multiple linear regression between the spatial distribution of BA and the three land use types. This means that $\hat{F}_{k,i}$ cannot be interpreted exactly the same as $F_{k,i}$ and is not "the fraction of land-use type that burns across the region". It would help the reader if you would redefine and clearly state what $\hat{F}_{k,i}$ stands for, and then if you want to explain it in a less technical way you could state "this is related to the fraction of land-use type that burns across the region".

The second definition (P10823 L12) "the net effect of land use k on fire in the region, expressed as a fraction of the area of land use k in the region" is confusing. The negative slopes are interesting and I agree that they may well represent a real aspect of the system. But with this definition you undermine your more realistic interpretation that a certain land use may affect fire activity in different land use classes in the vicinity (often within the same 0.25° grid cell). The fact that land use only explains a fraction of the variation of BA and your model often converges to regional mean values is a different issue, discussed below.

Data

The Methods of GFED are explained by van der Werf et al., 2010, and adjusted by Randerson et al., 2012.

Results

Figures 2 and 3. It would be easier for the reader if you would just present the annual burned area and carbon emissions split up by the different land use types (k) here. Then you can remove the

“cropland and crop+ categories” which are confusing and just use one “total”. It is not clear what applying the “model” adds to the results presented in these two figures opposed to a simple estimate of burned area per land use type “crop, pasture, natural, total”.

The authors provide little insights in the model performance and background data-sets. How much of the spatial variation in burned area can be explained by the distribution of the three land use classes for each analysis region (e.g., r-squared)? What are the actual $\hat{F}_{k,i}$ values per analysis region? And what does the land use distribution look like? It would be interesting to see some of these figures either in the main body of text or in the Annex material. If you think the paper is becoming too long you could merge figures 2 and 3, or remove all the current annex figures and make your point about the current interpretation of negative $\hat{F}_{k,i}$ a little stronger by citing more literature.

Fig. 4, this is an interesting figure. You state that “Numbers can be interpreted as ..”, but for the reader it would be easier if you first state what the numbers actually are, something like “the maps show $\hat{F}_{k,i}$ times the area of k”. Then in the next phrase you can say, “this can be interpreted as..”. Many people have knowledge about these type of models, providing such information makes it easier for them to interpret your results.

Figures 5 and 8, it seems that many of the regions are resolved as “mean” values, and that the different land use types provide only limited information on the spatial distribution of annual burned area within the analysis regions. This may be a consequence of:

1. Burned area in many analysis regions is dominated by a single land use.
2. In many cases little of the spatial distribution of BA can be explained by land use.

Both will make the values converge to the mean. This should be more clearly discussed.

Figure 7a, please explain the meaning of the three colors in the caption of the figure. Figure 7b, It may look nicer if you would delete the white space on the x-axis before “August”.

Figure 8b. This figure is a little counter intuitive now. First you say “(b) each grid cell”, but then the fit and the equations are presented for binned-mean values. First, depending on your bin-size the slope and r-squared will vary, which makes the results subjective. Second, you have already presented the “binned” results in Fig. 8a (using analysis regions for bins). It would be interesting to read here how much of the spatial variation in burned area is actually captured at the grid cell level by your model.

Discussion

For the reader it may be easier if you better separate the actual results and the discussion. Some of the results section reads more like a discussion while the discussion is sometimes very technical, how do your results relate to other literature?

4.3 Impacts of regional analysis

This is an interesting discussion. Poor performance for Europe seems to be mostly a matter of Europe having many fires in all three land use classes while their spatial distribution may provide little information on the distribution of these fires. It may help to better discuss the differences in land use management between different areas. For example in Eastern Europe and Russia agricultural fires

might be common practice but similar fires will not be found in Spain or Italy. In a similar way, it will be hard to compare pastures across the world. Some of the grazed savannas will appear so close to natural vegetation that the a measure of “livestock density” may be more useful than “grazed or not grazed”. On top of that, what about the naturally occurring herds of herbivores, especially in Africa. The FAO has published an interesting map of global livestock density.

The good performance of the model for Boreal Asia and for analysis regions where nearly all fires occur in a single land use type is obvious because the model just represents the mean values of the observations for that land use in the analysis region. The high number of analysis regions where this is the case is partly a consequence of the way the authors have defined the analysis regions in the first place. The authors should better acknowledge/discuss this. Now you state “Another, more general consequence .. in the results”. Here it would really help if you would have presented how much of the spatial variation in burned area could actually be explained by land use. And then just state something like “On average, only xx% of the spatial variation in burned area could be explained by land use, hence for many of the regions the $\hat{F}_{k,i}$ values simply represent the mean burned area for the given land use in the analysis region.” Finally, a short discussion of the alternative sources of spatial variation in burned area might be helpful.