**Interactive comment on** “Evaluation of forest fire models on a large observation database” **by J.-B. Filippi et al.**

J.-B. Filippi et al.
filippi@univ-corse.fr

Received and published: 19 May 2014

Dear Reviewer,

thanks for your comments that will help to greatly enhance the paper. Please find these answers and propositions for changes. We will wait for other reviews to start implementing those.

*In terms of presentation quality, I rate current presentation quality as Fair (3), and I explain why below. To put my rating in context, I believe that Excellent (1) should be reserved for papers that make exceptionally significant contributions.*

Thanks for the comment, we agree that the excellent should only be for exceptional
papers. We hope that with your helpful comments we can reach Good (2).

Other authors have attempted to determine the forecast quality of a single fire spread rate formulation for immediate operational use. This study is the first, to my knowledge, to use fire spread perimeters from a large number (80) of relatively small (simple fuel, Corsican) wildfires to compare the performance of the most current common European (Balbi) against the performance of the most current common American (Rothermel) fire-spread rate models, all run in fully-automated operational modes.

We are happy that readers understood the point of testing in this fully automated/AKO operational procedure. The motivation was searching how to evaluate, not the actual evaluation results.

Page 3225, lines 18 to 20. The authors point out that later versions of the Rothermel model are available. One criticism of the study might be the older version of the Rothermel model used. Please explain why you used the older version as opposed to a more current version.

The older version is well documented, the most cited and with parameters that can be matched to most common fuel characterization data. If we were to use another formulation, it would require to select sub-models (acceleration, spotting....) based on choices that would be subjective. We will further explain our choice in the text (section 2.2).

Page 3228, lines 6 to 28. Rewrite. As written it is not clear how the measured wind and fuel data are automatically preprocessed and readied for the fire propagation solver. On page 3235, lines 20 to 23, the authors write that "The meteorological values were taken at the closest observation station, even though the actual wind direction at the exact fire location may be significantly different." I understand that the authors were attempting to evaluate these models in a purely operational context, but the authors should explain, STEP-BY-STEP, how a user
goes from the raw station wind observations to the wind value input into the fire propagation solver.

We will include a step-by-step procedure description in section 3.2, with a diagram of the data flow.

The reason for this request is the authors use of WindNinja to output wind at the same resolution as the elevation field. This implies that the nearest (raw) station observation was NOT the wind observation used by ForeFire, but rather a WindNinja interpolated value. Please clarify.

Yes, you understood correctly. The station data is used to generate the WindNinja field. The step-by-step description will clarify that.

Page 3228, lines 21 to 23. When used operationally, the nearest upstream station wind is used to drive the Rothermel spread rate model. Please explain why winds from the nearest station, not the nearest upstream station, were used.

This would indeed be more relevant (upstream), not only for Rothermel, but for all models. Unfortunately, there is not one station per stream, or even for a few streams, topography of Corsica is rather chaotic as valleys are not more than a few kilometres wide. This is why we opted for a wind pre-processor to somehow "rectify", but we are aware (and already commented) that it is an important limitation, without clear alternative choice. We will clarify that no obvious upstream station is available.

Page 3231, lines 13 to 24. If some of the simulated results are based on data of poor quality, should they not be dropped from the scores? Or at least should scores based on some data quality selection also be reported? This would show just how much a poor quality wind forecast, for example, impacts fire spread forecasting.

We do not think that it is necessary because the scores are already ranked, so that one may just look at the good part of the graph. It is impossible to make an a priori
evaluation of the data quality. We will write a comment on the fact that "as data quality cannot be checked before evaluation, all scores are reported".

Page 3236, lines 21 to 24 (Appendix A). One reason I am not satisfied with presentation quality is that the authors present, but do not discuss, figures 5 to 7. These figures examine a few individual fires tabled in Table 3, and so they should be discussed either under Section 4.2 or the Appendix A. Please include a small discussion for each of these figures.

You are right, these are barely discussed in the legends. We will add a paragraph for the table and a paragraph for the figures in the appendix and refer to the appendix in the text.

The other reason is I think that Fig 2 needs a small change. I request that the grey used to delimit the upper and lower bounds of the lines be less opaque (and maybe a different colour), so that the lines showing scores become easier to compare/see. We will use a lighter grey. I don’t believe that the grey colour used elsewhere to show burned area needs to be changed, as long as the caption in figures 3 to 7 tells the reader that the gray area is the burned area.

We will update the captions.

I am not going to correct language/grammar/typos in the article, except for one thing. In science, datum is singular and data are plural. So change sentences that use, for example, "the data is ..." to "the data are ..."

Thanks, we will correct that and will proofread to find the other typos. Regards

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 2, 3219, 2014.