General comments

The paper introduces and discusses a method that utilizes kriging to spatially interpolate source-specific impact adjustment factors to generate revised CTM source impact fields from the CTM-RM method results. The method is then applied to January 2004 over the continental United States.

The paper addresses a relevant issue concerning the growing need to produce detailed and sound estimations of the contribution of the different emission sources to the PM concentration.

To this aim the paper introduces a novel approach, partially based on a previous work, combining features of both source and receptor oriented modelling techniques.

For this reason this work certainly fits the scope of GMD.

The paper is well written, with concise and clear statements, however there a few general and specific questions that should be addressed before publications.

General and specific questions are detailed in the following.

First of all there are some inconsistencies in figures and table citations; particularly the order of figures and tables does not exactly reflect the corresponding order of citation in the text. Moreover there are some figures and tables included in the paper but never cited or commented.

The abstract is concise and complete, but I suggest to introduce also a few quantitative evaluation of the improvement in model performance (e.g some measure of error and its corresponding reduction).

I would also suggest to introduce one or two figures describing the average spatial field of R's (likewise figure 5 does for concentrations).

There are also a few general questions that should be addressed in the discussion, namely:

1. The objective function described in (1) introduces a set of "adjustment" factors that allows filling the gap between observed and CMAQ concentrations modulating the influence of each source to the total concentration. But we know that the discrepancy between modelled and measured concentrations can rely not only on the emission strengths but also on other inputs (e.g. meteorology) as well as model formulation. Supposing for example that the overestimation of dust is related to an overestimation of wind, either that the overestimation of biomass burning is yielded by a too efficient SOA formation, how can the CTM-RM hybrid method improve the model performance for the right reason?

2. The paper introduces the concept of CTM-RM hybrid approach, also specifying that the RM model is CMB. But where does the RM actually contribute to the hybrid modelling approach? It seems that the objective function takes into account the source profiles, but not the results of a concentration apportionment. Probably more details are needed.

3. Related to the previous point, there is also another issue concerning the proper definition of "source profiles". Do they represent an "emission speciation profile", thus describing the source fingerprint at the emission point? Either do they represent the source fingerprint at the receptor? In case they correspond to the first definition, have they been somehow compared to the emission speciation profiles adopted by SMOKE/CMAQ to define the 33 emission categories? In the second case (source profile at the receptor) I suppose they cannot be considered totally invariant, because the fingerprint of a source can change according to the travel time (e.g. different deposition rates for primary compounds; chemical transformation for secondary compounds). Please, briefly discuss this issue, if the authors consider that is relevant for the optimization process.

4. The authors correctly point out that the "SAs" terms cannot strictly be considered as Source Contribution Estimates (SCEs), while they should be seen as sensitivity terms. This discrepancy can be relevant for sources like livestock that strongly contributes to ammonia emissions but not to NO_X. As a consequence, in terms of SCEs they contribute just to ammonia, but in terms of sensitivity they influence both ammonia and nitrate. Therefore, in case of livestock, considering the sensitivity analysis as a source apportionment, may imply an overestimation of the contribution of this source category. Do the authors consider that this aspect would help to explain the increased ranking of livestock after the adjustment? Do they consider this increase reliable?

5. The key aspect of novelty of the paper concerns the development of gridded CTM-RM source apportionment results, based on findings at the receptor sites. Did the authors investigate the issues related to the spatial representativeness of the measurement sites?

Specific comments

P650 – (1). Though the objective function is properly referred (Hu et al., 2014), I suggest to add a few details to make it more readable.
- Are all daily quantities ?
- Add a few details about: Uncertainties in observation measurement, source profiles and source strength
- Are Rj expressed as function of receptor and time?

P650 R19 – As already mentioned authors should briefly discuss the definition of "source profiles" (general question #3)

P652 R12 – if there are 189 CSN stations with 9 days, why N= 75 instead of about 170 ?

P652 R24 – Authors say that 41 species were used for CTM-RM/SH optimization, though some of them were seldom above the detection limit. Can these species introduce too much uncertainty in the optimization phase?

P653 R5-8 – The maps show several "hot spots" with strong spatial gradients. Could this effect be related to the spatial representativeness of measurement sites? (see also General question #5).
In some cases the observed value does not correspond to the surrounding gridded value. Are they withheld data?
P653 R9 – How many data are considered in table S2?

P653 R10 – Figure S3 shows that almost all Rs are < 1.0 suggesting that CMAQ-DDM estimations are always overestimated at all sites, for all sources. Any comment?

P654 R9-13 – Are the overestimations concerning Fig S6 and S7 expressed in terms of "factors", likewise fig. S5? They seem too high.

P654 R22-25 – See General question #4

P655 (4) – What do "i" and "N" account for?

P656 R5-10 – Authors discuss categories showing high absolute values of RMSE. Maybe some comments could be added also for categories showing a relevant RMSE with respect to the corresponding average and median (e.g. dust)

P657 R1 and R4 – What does N represent?

P657 R24 – Could it be useful adding also some information about emissions of the main precursors (NOX, NH3, SO2 and VOC)?

P658 R25-P659 R14 – The concept of source profile and its role should be better clarified (see also general question #3)

P659 R14-16 – Authors state that just through changes in emissions they can improve the model results and performance. But, how can they deal with discrepancies not directly related to emissions? See also general question #1

P670-671 figure 5 – Authors may also include a pair of total PM2.5 gridded fields, also overlapping observed data. This would give an idea of the actual improvement before and after the implementation of the correction factors.

Tables S3-S5 – Some error metrics (e.g. RMSE) could be added for each species to quantify the changes in model performance between CMAQ-DDM and the hybrid approaches

Tables S6-S8 – They should be commented because some results are not very clear. For example Beta coefficient for PM2.5 in table S6 decreases from 0.43 to 0.27 and 0.24. I would expect an increase of beta coefficient toward 1.0, in case of improved model performance

Technical corrections
P654 R8 – Figure 2?

P657 R24 – Why Tables S1 is placed before Table S2?

P668 Figure 3 – Is it cited in the text?

Tables S4 and S5 – HYB should be SH (spatial Hybrid)?

Tables S6-S8 – they are cited but not commented

Figures S4 and S8 – they seem not cited