Interactive comment on "Development of $PM_{2.5}$ source impact spatial fields using a hybrid source apportionment air quality model" by C. E. Ivey et al.

Anonymous Referee #1
Received and published: 25 March 2015

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.

We would first like to thank the referee for the constructive review of our manuscript. As noted, the aim of this paper is to develop a modeling method that addresses the need for creating spatially and temporally complete source impact fields, particularly to serve as exposure surrogates for health studies. We believe this work will positively benefit those conducting health studies as well as policy assessments regarding ambient $PM_{2.5}$ concentrations.

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.

We thank the reviewer for catching those and they have been fixed. We renumbered the figures and tables according to their appearance in the text. We also eliminated figures (previously Figs. S8-10) in the supplementary information that were not mentioned in the text.

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).

We added the following performance metrics in the abstract: % error relative to the observed quantity for trace metals (e.g., V, Mg, P) and correlation coefficients for ions (e.g., NO_3^{-2}, SO_4^{-2}, NH_4^{+}).

"Correlations improved for concentrations of major ions, including nitrate (CMAQ-DDM: 0.404, SH: 0.449), ammonium (CMAQ-DDM: 0.454, SH: 0.492), and sulfate (CMAQ-DDM: 0.706, SH: 0.730). Errors in simulated concentrations of trace metals were reduced considerably: 295% (CMAQ-DDM) to 139% (SH) for vanadium; and 1340% (CMAQ-DDM) to 326% (SH) for manganese. Errors in simulated concentrations of very trace components are expected to remain given the uncertainties in source profiles."

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

Average Rj values were presented for withheld CSN observation locations in the supplementary information (Table S1). We added an additional figure to display the 9-day averaged Rj spatial fields for four sources (dust, on-road diesel, on-road gasoline, woodstoves) similar to Figure 2 (Fig. S1).

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?

The referee brings up an important point that we clarify in the revised manuscript (conclusion). The purpose of the CTM-RM hybrid method is not to create a new chemical transport or atmospheric physics model, but to use statistical methods and measurements of PM species concentrations to improve source impact estimates. This improvement is two fold, 1) using the observed species concentrations to improve the predicted species concentrations and 2) incorporating emissions and atmospheric transformation processes from the model improves the secondary PM impact estimates beyond what is achievable using only a receptor model. We are not adjusting the source strengths, but adjusting the source impacts, and this distinction is important. We agree with the reviewer that other processes (e.g., wind, rain, or model parameter errors) can also impact the simulated impact of a source at a specific receptor, though errors in emissions (particularly highly variable ones such as dust) on a specific day can be large. Thus, the adjustment, as implemented here, is not on the emissions, but on the source-specific impacts. That is why we interpolate the receptor-specific adjustment spatially as opposed to using the adjustment to estimate how much the source strength should be adjusted and re-running the model.

The CTM-RM adjustment takes into account several sources of uncertainty: measurement uncertainty, modeled concentration uncertainty, as well as uncertainties in source impact estimates. These are represented as $\sigma_{i,obs}^2$, $\sigma_{i,SP}^2$ (now named $\sigma_{i,CTM}^2$, see question 3 below), and $\sigma_{\ln(R_j)^2}$, respectively. Additionally, using measured concentrations as the main point of reference for adjusting source impacts directly addresses the uncertainty of modeled processes (e.g.. wind-blown dust, SOA yield) by increasing or decreasing source impacts in light of an under- or overestimation in modeled concentrations. The 2nd term in the objective function prevents a non-real adjustment by further weighting the term with R_j uncertainty.

Since biomass burning and SOA formation have different compositions and source impact profiles, distinguishing between the two discrepancies can be determined by examining source impact adjustments. Further, collinearity is not an issue with this method because each source impact profile is different.

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.

The RM portion of the hybrid approach is based on minimizing the weighted difference between observed concentrations and simulated concentrations, while also accounting for uncertainties in source strengths, source profiles and observations, in essence an extended chemical mass balance (CMB) approach. In section 2.2 (pg. 650, line 17), we note that the CTM-RM "uses an effective variance approach to balance model outputs." The effective variance approach is also utilized by versions of the CMB approach (Watson et al., 1984). As suggested by the referee, we extended this section to add additional details and make this clearer in the description of the CTM-RM hybrid method in section 2.2.

"The effective variance approach is also utilized by versions of CMB (Watson et al., 1984), and the optimization method used here is, in essence, an extended CMB approach. Uncertainties in the first term of the objective function serve as effective variances of the numerator and are specified for each species *i*."

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.

This is an important point covered in Hu et al. (2014) and is related to the prior response (point 1). In this manuscript, "source profiles" describe the source fingerprint at the receptor, i.e., that the source profile can be altered, e.g., by the formation of secondary species. However, for many of the species, there is no secondary formation. We assume that within the accumulation mode (which contains most of the fine PM mass in CMAQ) that the composition of the primary portion of the PM2.5 from any source is the same, but secondary species can be formed, and primary components lost to deposition, altering the source profile at the receptor. The specific steps taken in applying source profiles to CMAQ-generated data, and those steps are described as follows. In a publication by Adam Reff et al. (2009), source profiles for 84 source categories were presented, which were aggregated from roughly 300 $PM_{2.5}$ SPECIATE v4.0 profiles and contain estimates of trace metal contributions. The 84 PM2.5 profiles were further aggregated into 33 categories, consistent with the sources of interest in this study. Then the trace metals contributions in the 33 profiles were used to speciate the "unidentified" portion of PM2.5 (species name: A25) as output by CMAQ (v4.7.1). The contributions of 35 trace metal species were normalized to one and then used to split the unidentified PM2.5, and results for these species are used. At the receptor, both the primary and secondary PM2.5 contribution at the receptor is used to determine the new, receptor-oriented, source profile. This same approach was used to generate trace metal species concentrations in the preceding publication by Hu et al., and more details about this method may be found there (Hu et al., 2014, *ACP*).

We do agree that source fingerprints may change from source to receptor due to physical and chemical changes. But CMAQ, in principle, captures those changes for emissions from each source from their emitting point to the receptor point and represents those changes in the calculated source impact values. We included a more detailed discussion of the source profiles in section 2.2 in the revision.

"Source impact profiles are derived from the information provided by Reff et al (2009). In this manuscript, "source impact profiles" are different than "source profiles" in that they describe the source fingerprint at the receptor. In other words, the source profile can be altered, for example by the formation of secondary species. However, for many of the species, there is no secondary formation. It is assumed that within the accumulation mode, which contains most of the fine PM mass in CMAQ, the composition of the primary portion of the PM2.5 from any source is the same, but secondary species can be formed, altering the source profile at the receptor. The specific steps taken in applying source profiles to CMAQ-generated data, and those steps are described as follows. Source profiles for 84 source categories were presented in Reff et al. (2009), which were aggregated from roughly 300 PM2.5 SPECIATE v4.0 profiles and contain estimates of trace metal contributions. The 84 PM2.5 profiles were further aggregated into 33 categories, consistent with the sources of interest in this study. Then the contributions in the 33 profiles were used to speciate the "other" (sometimes called unidentified) portion of PM2.5 (species name: A25) as output by CMAQ. The contributions of the 35 trace species were then used to split the "other" PM2.5 in to individual species, and results for these species, along with the other primary and secondary species are used. At the receptor, both the primary and secondary PM2.5 contribution at the receptor are used to

determine the new, receptor-oriented, source profiles. This same approach was used to generate receptor-oriented profiles in the preceding publication by Hu et al. (2014)."

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 NOX. 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?

We would first like to clarify the meaning of SAs in this manuscript. SAs denote the sensitivity of ambient particulate matter concentrations to emissions. In agreement with the referee, SAs are not total source contribution estimates (SCEs). However, we do not agree that our methods lead necessarily to an overestimation of the contribution. We do a mass balance on each individual speies as part of the method, and given that we have the inventoried sources (we use a comprehensive emissions inventory), so the result should not be biased high for any specific source. Our methods are novel in that, although some sources may not emit a certain pollutant, there still be some interactions with emissions from other sources, and they capture those interactions. For example, in the case of agricultural emissions, although little NOx is directly emitted, the influence on nitrate concentrations is precisely what we hope to estimate, as quantifying inter-source interactions (traditionally not quantified in source apportionment methods) are important in determining the primary and secondary impacts of sources on air quality. Our hybrid source- and receptor-oriented approach takes this into account, and can help elucidate impacts from source interactions. In this case, source impacts are the concentrations of pollutants that arise as a result of direct emissions and secondary interactions (both formation and destruction processes), and there is no truly unique impact of a source on secondary species when source interactions are involved. We added this explanation in the discussion section in the revised manuscript.

"The spatial hybrid method is also novel in that, although some sources may not emit a certain pollutant, there still may be some interactions with emissions from other sources leading to those species being part of the source impact. For example, in the case of agricultural fertilizer emissions, although NOx is not directly emitted, the influence on nitrate concentrations is calculated, taking account of inter-source interactions (traditionally not quantified in receptor-oriented source apportionment methods) that are important in determining the primary and secondary impacts of sources on air quality. This hybrid source- and receptor-oriented approach takes this into account and can determine impacts from complex source interactions. However, this also shows that the formation of secondary species is often dependent upon multiple sources, and the impact of one source is dependent upon other sources, leading to ambiguity in source attribution. The approach here uses the sensitivities at current conditions."

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?

First, we thank reviewer for noting spatial novel aspects of the manuscript. The spatial representativeness of our hybrid results were tested/evaluated by performing cross-validation through data withholding and by comparing to an independent dataset. Data withholding was used to evaluate the initial model development for spatial interpolation, where 10% of the monitors were randomly removed from the spatial dataset and the remaining 90% of points were kriged. The Rj value at the monitor and the corresponding kriged value from the spatial grid were compared. Evaluation metrics are presented in the

supplementary information (Table S2 and Figure S3) and indicate that the observation data available for this study can adequately serve as inputs for spatial interpolation.

Following the model evaluation using data withholding an independent dataset was used to evaluate the model. Gridded hybrid concentrations were evaluated using observation data from the IMPROVE monitoring network. Data from IMPROVE was not used for model development. Results of this evaluation are found in the supplementary information (Tables S7 and S8). Correlation coefficients comparing observed and gridded concentrations are highest for more abundantly available species concentrations (higher frequency of measurement being above detection limit).

We addressed this comment in section 2.4 in the revised manuscript:

"This study uses available speciated CSN data over the entire U.S., thereby providing a very spatially heterogeneous dataset that is representative of key emissions and meteorology in each region. The lack of rural data available may present uncertainties in the spatial representativeness of R_j values outside of urban regions."

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?

In revision, we added more details about the objective function, its inputs, and the spatiotemporal characteristics of the R outputs (Section 2.2). R values are specific to one site and one day, as the method is applied at monitors when speciated $PM_{2.5}$ data is available on observation days.

"The initial R_j values are specific to one site and one day, as the method is applied at monitors when speciated PM2.5 data is available on observation days, and are then kriged and interpolated. The terms c_i^{obs} and c_i^{sim} represent the observed and CMAQ-simulated concentrations, respectively; Γ weights the amount of change in source impact. Uncertainties in observation measurement ($\sigma_{i,obs}$), modeled concentrations ($\sigma_{i,CTM}$), and source strength ($\sigma_{ln(R_j)}$) are also included in the model. Specifically, $\sigma_{i,obs}$ is reported with measurements for each day from the CSN network; $\sigma_{i,CTM}$ is modeled error, which is proportional to observed concentrations and remains constant for all sites and days; and $\sigma_{ln(R_j)}$ is uncertainty in source contribution expressed as the log of the factor of uncertainty, which also remains constant for each site and day."

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

Please see our response to general question #3. We elaborated further on the source profiles in section 2.2 of the revised manuscript.

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

On page 652, line 12, N = 75 refers to the number of withheld CSN observations, which were used for model cross-validation. These 75 observations (space-time pairs) were randomly selected by removing 10% of the available observations with speciated $PM_{2.5}$ data on each observation day. Speciation is conducted every three or six days. We clarified this in Section 2.4 of the revision:

Response to Referee #1

"Performance of the spatial extension was evaluated using a data withholding approach. To evaluate the method, we removed 10% of the available observation (75 sets of observations at the monitors with speciated PM2.5 data) and re-ran the spatial hybrid model. This led to a total of 75 observation sets being used in the model evaluation. All references to "withheld CSN data" refer to these 75 sets of withheld data. "

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?

The referee raises an important note regarding the introduction of added uncertainty due to some species concentrations measured at CSN sites being below detection limit. We also considered the possibility of added uncertainty due to detection limit issues. We tested the optimization with the absence of species with limited availability, and we found no significant differences in model performance. Further, knowing that the observations are below the detection limit is useful information. The uncertainty value used in weighting the observations reduces the impact of those observations, but the information is still captured. We also added a summary of this explanation in Section 2.4 of the revised manuscript:

"Also note that 41 species, including total PM, were used for spatial field construction, but only results for 20 species are presented for comparison of CSN results and 15 species for SEARCH and IMPROVE results, as measurements for some trace metals are seldom above measurement detection limit. The possibility of added uncertainty in the optimization step due to detection limit issues was considered. The optimization was tested with the absence of species with limited availability, and no significant differences in model performance were found. The use of the measurement uncertainty in the objective function minimizes the role of those measurements on days when they are below the detection limit, but still accounts for the levels being low. Using all available measurements in the optimization model is the preferred approach."

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?

The referee points out that in Figure 2 (cited as Figure 3 in text) the Rjs at the measurement sites do not correspond to the surrounding kriged results. We found that the spatial grid is shifted slightly to the east, hence this issue is not related to spatial representativeness. The observations align well with the kriged fields when re-plotted. Figure 2, now Figure 3 in the revision, was replotted.

P653 R9 – How many data are considered in table S2?

In Table S2, 75 data points, corresponding to 10% of monitors with available speciated PM2.5 data from each observation day. All references to metrics at "withheld CSN data" refer to these 75 monitoring locations throughout the manuscript. We made designation clearer in the method evaluation section (2.4), as well as in the table headers. Please see the response above.

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?

We observed that most R values are less than 1.0, which indeed indicates that the hybrid-adjustment is reducing the initial CMAQ-DDM estimated source impact. However, for some sites and days, R values are greater than one (see Figure S2). We replotted Figure S2 with the y-axis on a log scale in order to

better see the values greater than 1.0. Further, in Figure S3, the cumulative distribution plots exceed 1.0 (x-axis) for dust, lawn waste burning, prescribed burning, and woodstoves. These sources are highly variable day-to-day, and underestimations are possible in cases where the original emissions missed an actual burn or dust event. We addressed this comment in section 3.1 of the revision.

"The cumulative distribution plots exceed 1.0 (x-axis) for dust, lawn waste burning, prescribed burning, and woodstoves. These sources are highly variable day-to-day, and underestimations are possible in cases where the original emissions missed an actual burn or dust event."

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

The referee addresses that the factors presented in reference to the differences between the gridded spatial fields of CMAQ-DDM and spatial hybrid concentrations. The factors indeed do not appear to be reflected in the spatial field plots (Figs. S5-S7). However, the factors were calculated based on the average ratio of CMAQ-DDM to spatial hybrid grids over the entire domain. For instance, on Jan. 4, on average, the CMAQ-DDM grid values were a factor of 3 times higher than spatial hybrid grid values for biomass burning (Fig. S5). Over a large portion of the domain (boundaries, central US), impacts are near zero, and ratios may be influenced by numerical noise. We should note that there was a small error in the plots (Rjs spatial hybrid impacts fields were oriented in reverse), which produced plots that would lead the referee to believe that our estimates were too high. We have replotted Figures S5-7, and the new plots reflect the factors explained above.

P654 R22-25 – See General question #4

We addressed this comment in general question #4.

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

In Equation 4, *i* represents monitors and *N* represents the total number of monitors withheld for evaluation. We will clarify this in the next revision. Error of the modeled concentration of each species is calculated as the average of the errors over all withheld observations. The notation of Equation 4 will also be modified for clarity (see below).

$$Error = \frac{1}{N}\sum_{i=1}^{N}\frac{|obs_i - sim_i|}{obs_i} \quad (4)$$

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)

We included the discussion of sources with similar mean and median values and correspondingly low RMSEs, such as livestock, Mexican combustion, and nonroad natural gas combustion (Table S1 in the revision).

"Sources such as diesel, liquid petroleum gas, non-road natural gas, and Mexican combustion all had very low RMSEs, mean R values near 1, and median R values near 1. This indicates that there is little to no adjustment to these source impacts and that kriging captures the R values calculated by the CTM-RM application."

Response to Referee #1

P657 R1 and R4 – What does N represent?

In these references, N represents the number of monitors used for evaluation. We changed the notation to "N = ## monitors."

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

We added domain totals of the emissions of precursors and discussed their role in the formation of secondary $PM_{2.5}$ in Section 3.1.

"Coal combustion, which includes the secondary formation of sulfate, remains in the top three sources for average hybrid PM2.5 source contributions at withheld observation locations, as its emissions uncertainties are low due to the availability of continuous emission monitoring data. SO_2 emissions are large (Jan. 2004 domain totals: 72924.7 metric tons per day), as are NOx (74619.7 metric tons per day) (Table S9). During the study period, coal combustion had the highest contribution to SO2 emissions (35080.3 metric tons/day) and the second highest contribution to NOx emissions (14250.1 metric tons per day) behind mobile sources. The source impacts found here account for the transformation of these gaseous emissions from coal combustion."

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

Please see response to general question #3. We clarified the role of source profiles in the revision.

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

Please see the response to general question #1.

P670-671 figure 5 – Authors may also include a pair of total $PM_{2.5}$ gridded fields, also overlapping observed data. This would give an idea of the actual improvement before and after the implementation of the correction factors.

We added plots of CMAQ-DDM and spatial hybrid total $PM_{2.5}$ fields to Fig. 5, as well as a discussion of the performance in estimating $PM_{2.5}$. We plotted overlapped observations for one observation day in January, which gives a better idea of improvement after implementing the adjustment factors (Fig 5).

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

We added columns for RMSE in Tables S3, S5, and S6.

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.

The referee raises an interesting point. Indeed some performance indicators for some species indicate poorer correlation, such as the beta values for calcium for CMAQ-DDM (beta = 1.22) and spatial hybrid (beta = 0.16) comparison (Table S6). However, all metrics presented must be taken into account and

evaluated holistically. The alpha values for calcium indicate an improvement in performance, as the spatial hybrid value (alpha = 0.044) is closer to 0.0 than the CMAQ-DDM value (alpha = 0.13). Further, mean concentrations at withheld observation locations also indicate better performance of the spatial hybrid model, where mean calcium concentrations were 0.0407 (observed), 0.182 (CMAQ-DDM), and 0.0501 (spatial hybrid) (Table S3). According to the mean concentrations, the spatial hybrid method performs best. Throughout the analysis, CMAQ-DDM estimates of trace metal concentrations were orders of magnitude too high, while spatial hybrid results were closer to observations. While some individual metrics indicate better performance of the base CMAQ-DDM, overall performance of the spatial hybrid method is most favorable. An important point is that the species where performance is less good are typically those species that have a smaller role in determining source impacts, e.g., they are very trace species and/or have high uncertainties (relative to their observed concentrations) in the measurements or source profiles. We added this discussion to Section 3.1.

Technical corrections P654 R8 – Figure 2?

This part of the manuscript is in reference to Figs. S5-7.

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

As addressed earlier, tables and figures will be renumbered according to their appearance in the text.

P668 Figure 3 – Is it cited in the text?

Tables and figures will be renumbered according to their appearance in the text.

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

Yes, HYB should be SH, as this was an oversight. Throughout the course of the production of this manuscript, we changed the notation of the results to spatial hybrid.

Tables S6-S8 – they are cited but not commented

We added relevant comments about these tables in the next revision.

Figures S4 and S8 – they seem not cited

These figures were removed in the revision.