Reviewer #3:

Independent evaluation: We find many statements of MMM status by HTAP2 within this manuscript; however, the independent evaluation supported by other measurements can increase the persuasion of this MMF result. I agree that MMF will be better in theory but we do not lead the proof that "MMF does give better deposition estimates by incorporating in-situ measurements" (P14, L262-263) without independent validations.

Response: There is no way to do a spatially overlapping comparison, Figures 4, S4, and S5 do clearly demonstrate the impact of MMF on the model estimates. By definition, MMF values will be closer to the measured values because they explicitly incorporate the measured values along with the model estimates.

Not machine learning, but show with test/training site

Response: We thank the reviewer for the comment. As you state, we are not using a machine learning methodology. The common machine learning methods require a large dataset for training and testing. Training and testing sites are not applicable in this case because each measurement is influencing the model to force it towards the measured value. A testing site could be chosen that is influenced by one or several training sites, but taking out the testing site from the full dataset will influence the results. IDW does not build a linear regression or other relationship between inputs and outputs the way machine learning does; it is simply adjusting the values nearby observations.

Already in use in EPA

Response: Yes, exactly, that is why we are applying this method globally. It is already commonly used within the US.

The largest change over the ocean: I can follow that one of the reasons for coarse grid resolution will lead to the largest changes over oceans as listed in Table 1. However, because this was not helped by the observational fact (e.g., ship-borne measurements), how can we interpret this MMF result? Is it possible to only focus on the grid where the observation was available within the 1 by 1 grid in Table 1?

Response: We thank the reviewer for the comment. We have redone the calculations with a 1, degree, 2.5 degree, and 5 degree interpolation distance. As such the results have changed and we expect there to be an increase over the ocean with the larger interpolation distances. This is because some observations are either in grid cells classified as ocean or coastal or are influencing ocean or coastal grid cells.

Specific comments

P2, L18-19 (Abstract): Why sulfur trends were not stated? Moreover, according to my major concerns, please rewrite this abstract. It should be clarified the validation of this MMF result.

Response: Sulfur trends are now stated more explicitly in the abstract and the results sections. We cannot have an independent evaluation; this is not a machine learning approach.

P2, L45: It is ambiguous what "it" indicates. Is it ambient concentration or dry deposition?

Response: "It" refers to dry deposition and the sentence has been updated to reflect that.

Line 45: "Dry deposition is inferred from continuous measurements combined with modeled dry deposition velocities…"

P3, L55: No need for the repetition of EANET.

Response: Thanks for the comment. The reference has been changed.

P4, L78-79: I noticed Tan et al.'s paper is updated recently (http://dx.doi.org/10.1016/j.scitotenv.2022.158007). What are the differences between this update and this study?

Response: Tan et al.'s newest paper focuses only on China and only on wet nitrogen deposition.

P4, L86-88: This sentence is the result and is not appropriate to be stated in this introduction section.

Response: The sentence has been removed.

P4, L89: How about preparing table summarization for these available datasets? It will be kind to wide readers.

Response: We have added a new table as Table 1 to summarize the datasets.

Table 1: Sources of deposition observations.

Name	**Source**	**Number of Observations**	**Region**	**Value**
NTN, AIRMoN	NADP	247	USA	wet deposition
CASTNET	NADP	84	USA	dry deposition
CAPMoN	NAtChem	27	Canada	wet and dry deposition

EMEP	EMEP	86	Europe	wet deposition
Li et al. Study	Li et al. 2019	407	China	wet deposition
EANET	EANET	47	East Asia	wet and dry deposition
IDAF	INDAAF	1	Niger	wet deposition

P6, L137: It is one of an approach to use wet deposition itself, but their elements (concentration in precipitation and precipitation amount) could be the target of MMF. I can see some relevant discussion in Section 5, but for example, the project of MICS-Asia used the fusion for monthly-accumulated precipitation (https://doi.org/10.5194/acp-21-8709-2021). It will be better for readers why wet deposition is targeted as MMF in this study.

Response: We thank the reviewer for the comment. While we agree with the reviewer that it is common to fuse the concentration in precipitation, we are using an ensemble mean wet deposition grid created from 11 HTAP-II models and therefore we do not have independent precipitation and concentration grids. It is not reasonable to follow the method the reviewer describes given the constraints of our datasets.

P8, L183 (Table 1): It is kind to provide the region map for this analysis as a supplemental figure.

The region map is provided as a supplemental figure (SF1), adapted from Tan et al. 2018.

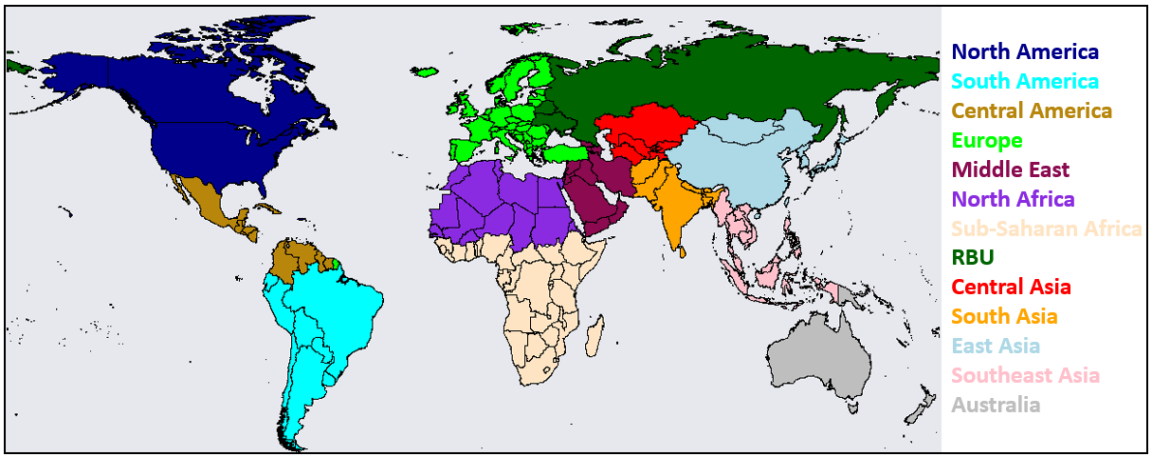

Figure S1. A world map showing the regions used to calculate the totals presented in Table 1. This figure is adapted from Tan et al., 2018, based on their region divisions.

P9, L189 (Figure 2): How about presenting the difference between MMF and MMM to clarify the difference driven by data fusion in this study? This result will clarify the impact of MMF compared to MMM and can help to understand the result listed in Table 1.

Response: We have changed Figure 4 and added figures S4 and S5 to present the different between MMM and MMF to demonstrate the differences driven by the data fusion method.

P10, L202 (Figure 3): But MMF used EMEP dataset itself, so this kind of comparison seems to be meaningless.

Response: Yes, it is true that MMF uses the EMEP dataset; Figure 3 shows to what extent IDW can "correct" the model to match the EMEP results. We cannot have a testing/training split dataset because we are not doing machine learning and there is no model to apply to testing data; therefore, we are limited to looking at the distribution and characteristics of the EMEP and the MMF data points.

P10, L202-203: I do not follow this sentence for East Asia. From this figure, MMF still underestimated the observational values.

Response: Yes, MMF still underestimates the observational values because it is effectively a weighted "average" between observations and model estimates. Therefore, it can correct the model where there are measurements, but the rest of the region's results are based solely on the model. So even "nudging" the model estimates in some grid cells toward the observations is not enough to fully correct the model at those places or over the entire region.

P12, L226-229: Within this context, TDep is regarded as truth?

Response: Yes, that is correct. TDep is widely in use and is created using CMAQ estimates and PRISM precipitation reanalysis and observations. There is a whole team working just on TDep and it has been endorsed by the scientific community.

P12, L230: Why NH4 is only presented? In addition, because MMF uses NADP dataset itself, what is the meaning of this kind of evaluation?

Response: We thank the reviewer for the comment. We now include all species in either the main text or the supplemental figures.

Technical corrections

P7, L169: In this Figure 1, "concentration in precipitation" multiplied by "precipitation" should be "wet deposition"? Please confirm this illustration.

Response: Correct. The Figure has been updated to reflect the change of phrase.

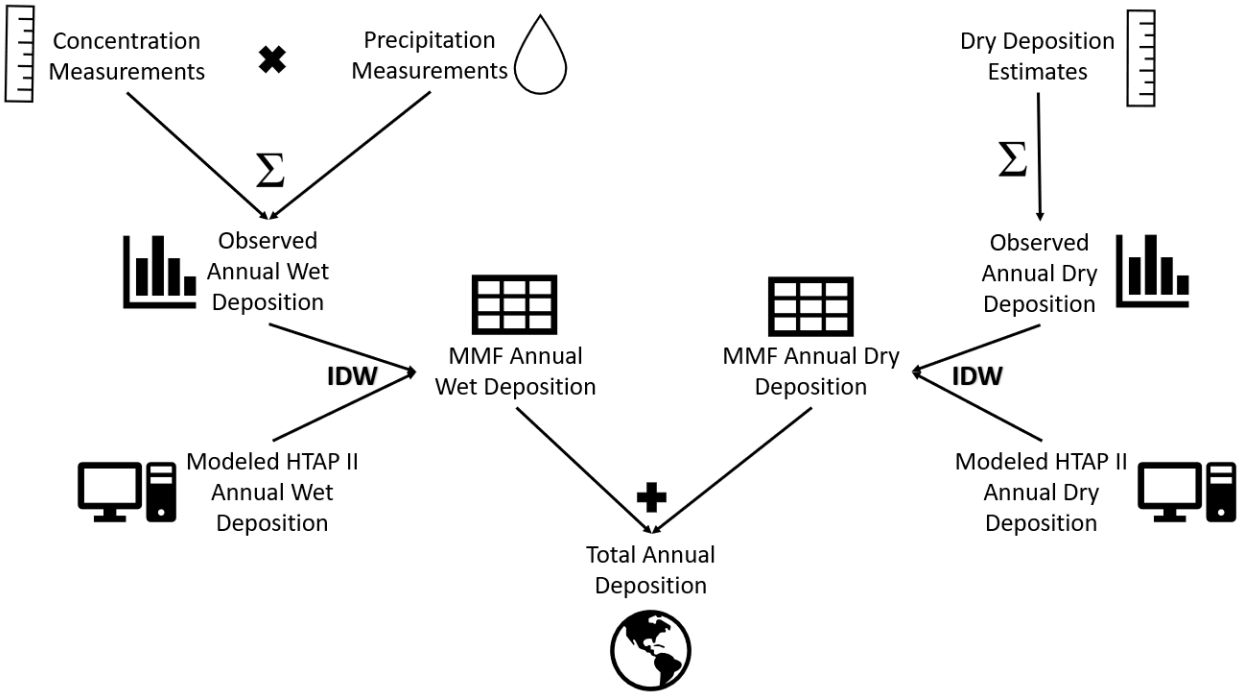

Figure 1. A flowchart describes the MMF methodology implemented in this paper.