This manuscript presents a top-down estimate of ammonia emissions in China using IASI satellite observations and WRF-Chem simulations. The study finds a notable contribution from non-agricultural sources and highlights their implications for $PM_{2.5}$-related health burdens. The topic is of broad interest to the community. However, the manuscript requires major revision before it can be considered for publication. I have several concerns about the MLR-based top-down estimation approach, and the associated uncertainties need to be more clearly explained and discussed to evaluate the robustness of the results. Additionally, the interpretation of the top-down emission estimates should be strengthened by providing deeper insights and more comprehensive comparisons with existing literature.

Major comments:

1. Although the authors do not mention, it seems to me that the MLR-based approach assumes a linear relationship between NH_3 emissions and column concentrations and attributes all discrepancies between satellite observed and simulated NH_3 columns to local NH_3 emissions. It is important for the authors to clarify these assumptions and discuss the associated uncertainties. Specifically, (1) How might the nonlinear response of NH3 concentrations to emissions affect the results? For example, do the summed contributions from individual sources (SA_agriculture + SA_non-agriculture + SA_transport) approximate the simulated total NH_3 column from the prior? (2) To what extent could transport of emissions from nearby grid cells influence the posterior results and cause spatial misattribution of emissions? (3) How might uncertainties in emissions of other pollutants (e.g., SO_2, NO_x) or in the model's representation of inorganic aerosol formation, impact the posterior estimates of NH_3 emissions? Could the large discrepancies between prior simulated and observed NH3 column concentrations due to these factors? Some quantitative discussion or sensitivity analysis on these points would help strengthen the credibility of the posterior estimates.

2. It is not very clear to me how the current MLR framework can separate AGR and non-AGR NH_3 emissions. Some clarification would be helpful, as noted in the minor comments below. Generally, in each grid cell j, I would expect the temporal variations in SA_agriculture and SA_non-agriculture to be perfectly correlated and differ only in magnitude, thus they could not be separated in the regression. While WRF-Chem may simulate different day-to-day variations for SA_agriculture and SA_non-agriculture, that reflects the effects of transport from surrounding grid cells, which contradicts the assumption that transport effects are negligible. Also, it would be useful to explain why the emission corrections primarily affect the non-AGR sector. Given that non-AGR emissions are relatively small in the prior, one would expect SA_non-agriculture to be much smaller than SA_agriculture in Eqn (4). Is the regression

coefficient b significantly larger than a, and if so, what is the reason for that? Specifically, in northern Henan, the posterior results show decreased AGR but increased non-AGR emissions compared to the prior, which seems hard to understand and needs further explanation.

3. The finding of substantially higher non-agricultural (non-AGR) NH_3 emissions compared to prior estimates is certainly interesting and important. However, the discussion of the posterior results in Section 4 currently focuses mainly on reporting emission magnitudes and sectoral contributions, with limited interpretation of the underlying causes or contextualization within existing literature. I would encourage the authors to expand this discussion by addressing the following points:

(1) What are the potential reasons for the apparent underestimation of non-AGR NH_3 emissions in current bottom-up inventories? (2) What types of non-agricultural sources (e.g., industrial processes, transportation) are most likely responsible, based on current understanding? (3) How do your findings about non-AGR emissions compare with previous top-down estimates? The discussion in Lines 259–281 is helpful, but it could be further strengthened by emphasizing on observation-based or model-based studies that have investigated non-AGR NH_3 sources. It would also be valuable to highlight how your results build upon or differ from those studies, and what novel insights your analysis contributes to this topic.

Minor comments:

1. Line 26/28 and elsewhere: please remove the "·" between Tg and yr-1. Also, replace Gg mon-1 with either Gg month-1 or Gg mo-1 to follow standard unit concentions.

2. Line 37-40. The summary statement is too general. It would be more informative to highlight the insights into non-agricultural ammonia emissions and their implications.

3. Line 127: Is biomass burning emission also treated online? Just checking, as this is not commonly the case.

4. Line 134: The "last accessed" date should reflect the actual date when the data were downloaded.

5. In section 2.2, what's the overpass time of IASI data? Do you use level-2 satellite data?

6. Line 140-141: The description is unclear. How is the neural network applied to improve the data quality, was it developed by the authors or sourced elsewhere?

7. Line 125/156: Which version of MEIC is used?

8. Line 162-163 and Section 4.1: Please clarify how the model and observations are sampled for comparison.

9. Line 179: The term “errors” is vague. Consider using clearer language such as “underestimated by 30%” or “biased low by 30%.”

10. Line 181-189: Figure 6 can be described with the text here.

11. Section 3: Please be consistent in the use of statistical metrics. RMSE is used for IASI comparisons, while IOA and MFB are used for surface observations. A brief explanation of why different metrics are applied, and what each evaluates, would be helpful.

12. Line 195: You mention deriving posterior emissions for four months—how are prior/posterior simulations compared with observations across the seasons? Are the same scale factors applied to all three months in each season? Please clarify. Given that WRF-Chem simulations are available for the full year, it would be more consistent to derive monthly emissions for all 12 months, which should follow the same procedure and would not require much additional effort.

13. Line 200: In Line 138, you mentioned that IASI data were regridded to the model resolution, but here you refer to single-pixel comparisons, which is somewhat confusing. Please clarify how the satellite data were matched to model outputs.

14. Line 202: What does “regression factor” refer to? Is it the same as the emission scale factor?

15. Line 206: Is TA_satellite the monthly average or the daily average of NH3 concentrations?

16. Line 208: Should be “outside transportation, AGR emissions, non-AGR emissions, respectively’

17. Line 211: The term “control emissions” is unclear. Do you mean emissions were zeroed out? Also, please replace “cycle” with “experiment.”

18. Line 214: What is A_blank used for?

19. Line 203/216: Earlier you use k for month and j for region, but later this is reversed. Please ensure consistency throughout. Also, using “grid cell j” is clearer than “region j” or “area j.”

20. Line 217: It is unclear why does the regression is derived mathematically imply it needs to be corrected? Please clarify the motivation for adjusting the regression coefficients.

21. Line 218-231: The description of the correction process is not very clear. It is unclear what is meant by “goodness of fit,” how the “invalid” regression coefficients are defined, and what fraction of them are removed. The phrase “make a trade-off” in Line 225 is vague and would benefit from clarification. Additionally, it is not explained how the adjustment factors a_n and b_n are derived or what their physical meaning is. The choice of a 30% threshold in Line 229 also seems arbitrary—particularly in high-NH_3 regions, where it could allow larger discrepancies between observations and simulations, but the physical basis for this threshold is not clearly explained.