We appreciate the generally favorable nature of the peer reviews and the opportunity to enhance the paper by responding to specific comments. Reviewer comments are in italics, with author's response provided below each corresponding comment. Note that line numbers in responses refer to the revised manuscript.

Response to RC1: Anonymous Referee #1:

1. *The first paragraph of Sect. 1 introduces the importance of N fertilizer on agricultural land and its implication on N emissions, but neglected those from non-cultivated land. In addition, it is not clear why only NO and* NO_2 *emissions are mentioned in this paragraph, and their relationships. Merging this part with the third paragraph might improve the logic here.*

> We address the importance of soil nitrogen emissions from non-cultivated (non-agricultural) land in the second paragraph of Section 1 (Lines 50-53):
>
> "Recent studies have shown higher soil NO_x even in non-agricultural areas like forests to significantly impact summertime ozone in CONUS (Hickman et al., 2010; Travis et al., 2016). Consequently, it is increasingly important to estimate both N fertilizer-induced and non-agricultural NH_3 and NO_x emissions in air quality models."
>
> We adopt the reviewer's suggestion to improve the logic of the flow in the third paragraph (Lines 54-58):
>
> "Soil NO emissions tend to peak in the summertime, when they can contribute from 15-40% of total tropospheric NO_2 column in the continental U.S. (CONUS) (Williams et al., 1992; Hudman et al., 2012; Rasool et al., 2016). Summer is also the peak season for ozone concentrations (Cooper et al., 2014; Strode et al., 2015) and the time when photochemistry is most sensitive to NO_x (Simon et al., 2014)."

2. *L79: the impacts of* N_2O *emissions are not introduced as that for* NO_x*,* NH_3*, and HONO.*

> We add the following statement in Line 71:

“Soils and agriculture are the leading emitters of N_2O, a potent greenhouse gas (IPCC, 2013).”

3. *L223: it is a little confusing on the different versions of CMAQ and the schemes of NO emission in these versions. For example, is YL or BEIS used in CMAQ? In which version. This confusing issue can also be found in later of the manuscript due to too many schemes, methods, interaction systems, datasets are introduced here. A clarification of the abbreviations and the purpose of them could be useful to readers.*

To clarify ‘CMAQ-YL’ (in Line 224) refers to the original CMAQ v5.1 which has the Yienger and Levy (1995) (abbreviated as YL) scheme for soil NO estimation used in the Biogenic Emission Inventory System (BEIS) for in-line biogenic emissions calculation in CMAQ. The term ‘CMAQ-YL’ term was used to highlight that CMAQ’s default soil scheme differs slightly from the original scheme presented by Yienger and Levy (1995) (refer to section 2.1, Lines 216-241). To further avoid confusion, we added in Line 239:

“However, for sake of simplicity we refer to ‘CMAQ-YL’ merely as ‘YL’.”

The only other two variations to this original CMAQ v5.1 code are the replacement of YL in the in-line BEIS with:

a) ‘BDSNP’ (earlier implemented in previous version of CMAQ i.e. v5.0.2 presented in Rasool et al. (2016) and updated for v5.1 for this paper), and

b) The new ‘Mechanistic’ (or ‘Mech.’) scheme implemented in CMAQ v5.1 and presented in this paper for the first time.

Fig. 1, Section 2.1, Tables 1 and 2 clearly described and distinguished these three variations for soil N estimation in CMAQ v5.1 (CMAQ-YL or ‘YL’ actually being the original implementation of CMAQ v5.1 available in CMAQ’s official distribution from U.S. EPA). In addition, results presented in the paper are compared between these three different schemes.

4. *Similar to points 3, Sect. 2.2 is a little hard to follow given different land covers are used and converted in different model.*

> Table A2 gives the mapping of NLCD 40 land cover types to MODIS 24. Also, Table A1 gives the different climate zones in which the respective MODIS 24 land cover types fall. Our mechanistic scheme only uses NLCD 40 as it is the default land cover definition used in CMAQ. NLCD 40 to MODIS 24 conversion was needed only in BDSNP as it used constant soil NO emission factors related to non-agricultural land covers (classified as per MODIS 24 nomenclature), which have also been described in Rasool et al. (2016).

5. *Sect. 2.6, the mechanisms are very well organized and presented. But it could be better whenever the factors impacting concentrations or fluxes can be referenced (e.g., fXXX in those equations).*

> We actually do reference factors affecting soil nitrogen fluxes as ‘fXXX’ (generic form of function) in Equations 2-7 in Section 2.1 (Overview of Soil N schemes) for ‘YL’, ‘BDSNP’ and ‘Mech.’ schemes. These factors or functions (fXXX) are expanded in detail in different equations throughout Sections 2.5 and 2.6.

6. *The model comparison and evaluation are only conducted for two months in one year (May and July of 2011). It is crucial to explain the reasons in more detail. Readers may very curious about why. For example, why not using multi-month (e.g., for a whole year) and multi-year (e.g., 5-10 years) for evaluation? Is that due to the availability of observations? If so, it would be necessary to list the available observations. Unless using two months of a single year is well justified, it could be good to use more observations for seasonality, or even interannual variability, given that the purpose of a model (and the evaluation) is to be able to simulate spatio-temporal changes.*

> The period between 1 May to 1 August has been established to exhibit ~ 2/3 of the total annual soil NO_x budget (Hudman et al., 2010). Hudman et al. (2012) also exhibited the soil NO_x to be maximal in the months of May (onset of growing season) and July (offset of growing season). Rasool et al. (2016) also established by running a standalone BDSNP soil NO parametrization for the whole year that May and July have the highest soil NO fluxes.

The observational studies giving maximum soil NO emission rates in Table 3 happen during May and July as well. Hence, for a computationally intensive, regional-scale (Continental US i.e., CONUS) simulation like ours that involved both EPIC and CMAQ, focusing on May and July makes sense based on the above justifications. For inter- and intra-annual variability, EPIC derived soil Nitrogen pools, other relevant soil properties, emission inventories and meteorology for different years are required, which is beyond the scope of this study.

7. *Sect. 2.8, what about the validation of* N_2O *emissions?*

We stated in Lines 647-649 in section 3.1, that:

"However, unlike NO_x emissions, for N_2O no background conditions or emission inventory is in place in CMAQ's chemical transport model, so comparisons with ambient observations are not yet possible."

This highlights the need for further work within CMAQ to include greenhouse gases like N_2O, which are not accounted in its chemical transport model currently. The absence of background/initial conditions and an emissions inventory for other sources of N_2O resulted in our choice to keep N_2O as a separate diagnostic output of our emissions model rather than an input to chemical transport modeling.

8. *Sect. 3.1, it is not clear what is the anthropogenic emissions. Please define it? Whether emissions caused by fertilizer application are anthropogenic?*

We clarify that we were referring to anthropogenic fossil fuel emissions (Lines 621-623):

"However, the aggregated budget of soil NO is much less than anthropogenic NO_x from non-soil related sources, because fossil fuel use is concentrated in a limited number of urbanized and industrial locations."

9. *L630: when exactly the peak emissions happened in site observation? Are they also in May and July?*

> Yes, as mentioned in response to RC1 comment #6: The observational studies giving maximal soil NO emission rates across different sites in Table 3 happen during May and July (onset and offset of growing season respectively). That is also a reason justifying the simulations during these months specifically, besides May and July also being the peak soil NO_x months in the year.

10. *L638: differences are obvious also in Canada. It may be good to explain this too.*

> BDSNP estimates higher soil NO emissions than the other models in forested regions of northeastern Canada, like due to the higher emission factor that it assigns to forest biomes (Rasool et al., 2016). The mechanistic scheme estimates lower emissions there because it tracks the actual N transformation processes.

11. *Sect.3.2, Why not directly compare it with observations like in Fig. S2-b. It should be mentioned that negative bias in difference means less bias compared to observation. Statistics on the mean biases from different schemes are important, and should be presented. For example, the 1:1 scatter plot compared to observations, which may quantify the improvements and disadvantages.*

> Our aim is to show the difference that results from using the 'BDSNP' and 'Mech.' schemes relative to original CMAQ (i.e. 'YL').
>
> We added the following in Lines 673-674 as suggested by the reviewer:
>
> "In addition, negative bias in difference means less bias compared to observation (Figures 6-10)."
>
> We assert that spatial plots of statistics like Mean Bias are preferable to scatter plots because they represent spatial patterns in model performance.

12. *Fig.10: mechanistic scheme is worse compare to that of YL in northeast US. Can it be explained?*

Mechanistic scheme estimates for total NO_x are lower than those from YL in northeast US, as evident in Figure 5. That explains the higher positive bias in $PM_{2.5}$ NO_3 in Mechanistic scheme compared to YL with respect to the observation in Figure 10. This underestimation may be attributed to lack of excess manure N that is applied to agricultural filed in vicinity of animal feedlots while estimating soil N in EPIC (also described in Lines 719-727). Additionally, EPIC optimizes the fertilizer application rate to account for the modeled plant nutrient demand. This is often an underestimate of real world practices as discussed in the last paragraph of section 3.3. We are currently working on how to best address this discrepancy within the EPIC-CMAQ modeling system.

13. *L717: please explain the exact regions and locations.*

Specified the regions in Lines 719-722 as:

"Underestimates of soil N in some regions with an abundance of animal farms, such as parts of Colorado, New Mexico, north Texas, California, the Northeast U.S., and the Midwest, may be attributed to the lack of representation of farm-level manure N management practices, in which manure application can exceed the EPIC estimate of optimal crop demand."

14. *L752-753: it could be helpful to show the general performance on the dry and wet conditions used (simulated by other models).*

Fig. S7 in supplementary material shows estimated low soil moisture to also exhibit very dry conditions in Texas for May and July 2011, while relatively moist conditions with highest soil moisture in the Northeast and Pacific Northwest primarily in May 2011. Hence, the WRF meteorological model simulation for soil moisture for both dry and wet conditions in this paper performs reasonably well in comparison to the actual reported wet and dry conditions in 2011 as reported by NOAA's Palmer indices for wet and dry conditions across CONUS in 2011, as cited in Line 753.

15. *L760: it may be good to indicate from literature the importance of manure management (e.g., compared to N fertilizer) in these regions.*

We do address the detrimental impact of land application as part of manure management in Lines 722-727:

"Farms in the vicinity of concentrated animal units often apply N in excess of the crop N requirements as part of the manure management strategy, typically increasing the N emissions (Montes et al., 2013). USDA has reported that confined animal units/livestock production correlates with increasing amounts of farm-level excess N (Kellogg et al., 2000; Ribaudo and Sneeringer, 2016). Model representations of these practices are needed to better estimate the impact of nitrogen in the environment."

To clarify the importance of manure management compared to N fertilizer in the U.S., we present the further explanation in Lines 764-773:

"In the U.S., 60 percent of Nitrogen from manure produced on animal feedlot operations cannot be applied to their own land because they are in 'excess' of USDA advised agronomic rates. Most U.S. counties with animal farms have adequate crop acres not associated with animal operations, but within the county, on which it is feasible to spread the excess manure at agronomic rates at certain additional cost. However, 20 percent of the total U.S. on-farm excess manure nitrogen is produced in counties with insufficient cropland for its application at agronomic rates (Gollehon et al., 2001). For areas without adequate land, alternatives to local land application such as energy production (for example, biofuel) are needed. In absence of such a mitigation strategy, excess manure N applied on soil contributes is susceptible to reactive N emissions and leaching (Ribaudo et al., 2003; Ribaudo et al., 2012)."

The following citations are added in 'reference' section:

Gollehon, N. R., Caswell, M., Ribaudo, M., Kellogg, R. L., Lander, C., and Letson, D.: Confined animal production and manure nutrients, United States Department of Agriculture, Economic Research Service, 2001.

Ribaudo, M., Livingston, M., and Williamson, J.: Nitrogen management on us corn acres, 2001-10, United States Department of Agriculture, Economic Research Service, 2012.

Ribaudo, M., Gollehon, N., and Agapoff, J.: Land application of manure by animal feeding operations: Is more land needed?, Journal of Soil and Water Conservation, 58, 30-38, 2003.

16. *It is the first process-based scheme in a photochemical model. But authors may need to mention where this kind of mechanisms have been used before (e.g., crop models, terrestrial vegetation models, etc.), and the advantages.*

We already have addressed the advantages of using mechanistic model like DayCENT and listed similar process-based models in Lines 369-381:

"One of the advantages of using DayCENT is its ability to simulate all types of terrestrial ecosystems. DayCENT is one of the only biogeochemical models which not only provides a process-based representation of soil N emissions, but has also been calibrated and validated across an array of conditions for crop productivity, soil C, soil temperature and water content, N_2O, and soil NO_3^- (Necpálová et al., 2015). Hence, mechanistic models like DayCENT yield more reliable results by applying validated controls of soil properties like soil temperature and moisture, which are the key process controls to nitrification and denitrification. More recent mechanistic models like DNDC, MicNit, ECOSYS, and COUPMODEL are quite similar to DayCENT in their representation of nitrification and denitrification process. However, these models have not been as widely evaluated and impose greater computational costs (Butterbach-Bahl et al., 2013). DayCENT also enhances consistency in our mechanistic model by utilizing the same C-N mineralization scheme (taken from the CENTURY model (Parton et al., 2001)) that is used in EPIC."

Minor remarks:

L346: Wang et al.: please provide the year of this publication.

Wang et al. (1998), edited in Line 347

L457: NH4+?

NH_4 changed to NH_4^+ in Line 458

Response to RC2: Anonymous Referee #2:

Comment 1: *Figure 3. The authors explain the results due to "likely" causes. Figure 3c does not convey clearly the results intended by the authors. This part should be clarified.*

To clarify, we referred to 'likely' causes in Lines 610-615 as the differences between BDSNP implemented in GEOS-Chem (Hudman et al., 2012) and in CMAQ (Rasool et al., 2016) to be the finer land use definition and daily scale and finer resolution EPIC soil N data, which has been illustrated in greater detail in Rasool et al. (2016). Fig. 3c on the other hand is the nitrogen oxide flux from the mechanistic scheme, which has a dynamic representation of C-N mineralization, absent in both YL and BDSNP. We further edited Lines 612-617 as:

"Hudman et al. (2012) found nearly twice as large of a gap between BDSNP and YL in GEOS-Chem; the narrower gap here likely results from our use of sub-grid biome classification and EPIC fertilizer data (Rasool et al., 2016). The mechanistic scheme (Figure 3c) generates emission estimates that are closer to the YL scheme but with greater spatial and temporal heterogeneity, reflecting its use of a more dynamic soil N and C pools."

Comment 2: *It also appears that the process-based methods introduced in the CMAQ framework cannot be rigorously tested due to lack or old data, which detracts somehow from the considerable efforts made to improve the accuracy in soil N emission predictions. Presentation quality is fine.*

This work highlights the scarcity and need of observation of soil nitrogen fluxes (especially NO_x, HONO and NH_3 that affect air quality) on a frequent basis and in more locations.

Firstly, agricultural study sites such as the Kellogg Biological Station (https://lter.kbs.msu.edu/datatables/177) are quite rare and not well aligned with ambient air quality observation networks. Secondly, the N_2O measured at agricultural sites is unaccounted for in most chemical transport models like CMAQ. In addition, these chamber studies are designed more with the aim of looking at difference between various management practices on a field scale, which would require running different simulations of biogeochemical models (EPIC or DAYCENT), which is computationally expensive for a regional scale (CONUS) implementation like this, but ideally extend to future research plans.

However, improvements in modeled estimates in comparison to observed OMI NO_2 column, measured concentrations of NO_x, O_3, $PM_{2.5}$ NO_3 and some available soil NO emission rates, with 'Mechanistic' scheme does provide an indication that we are moving towards the right direction.

Comment 3: *Whenever possible, authors should include estimates of estimation or observational errors (e.g. Table 3).*

Table 3 gives the comparison of maximum soil NO emission rates observed for various sites with those corresponding to the three modeling approaches presented ('YL', BDSNP' and 'Mech.').

Comment 4: *Abbreviations used in tables and figures should be explained in the table titles or figure captions. Tables and figures should stand on their own.*

Edits have been made to define abbreviations at first use in both tables and figures as well.

Comment 5: *Since CMAQ already uses EPIC to simulate NH3 bi-directional exchange, the authors should acknowledge recent documentation of process-based denitrification approaches used in EPIC: Izaurralde et al. (2017). Ecol. Modelling 359:349-362 doi:10.1016/j.ecolmodel.2017.06.007. (see line 481).*

Izaurralde et al. (2017) added to line 482, with full citation in ‘reference’ section as:

Izaurralde, R. C., McGill, W. B., Williams, J. R., Jones, C. D., Link, R. P., Manowitz, D. H., Schwab, D. E., Zhang, X., Robertson, G. P., and Millar, N.: Simulating microbial denitrification with EPIC: Model description and evaluation, Ecological modelling, 359, 349-362, 2017.

Comment 6: *The methodology and Figure 2 do not describe well the treatment of soil layer processes. EPIC simulates soil C and N transformation layer by layer up to 15. Is it the same for DayCent? How are the results from one model past to the other? Are these calculations done for the surface layer?*

EPIC is coupled with CMAQ through the FEST-C interphase to be compatible with the regional scale (CONUS) implementation in CMAQ. All EPIC output variables provided to CMAQ as input for calculating soil N emissions are for the soil depth from 0 to 1 cm and from 1 cm to 10 cm (prefixed as L1 and L2 in FEST-C), respectively. Bash et al. (2013) also modeled Ammonia evasion from soil and NH_4^+ nitrification losses for CMAQ, utilizing FEST-C interphase soil layers with depths of 1 cm and 10 cm, keeping things consistent in treatment of soil layers when it comes to treatment of different soil N cycling processes.

To clarify more, DAYCENT’s soil N gas sub-module was not run separately, but was ported and coded in the new ‘Mech.’ scheme in CMAQ and calculations in terms of soil layers were always consistent with the above-described approach for EPIC-CMAQ (i.e. top 10 cm soil layer, where the soil N cycling mostly occurs).

Briefly, the CONUS regional-scale implementation of EPIC and DAYCENT in CMAQ do not use all the soil layers except for topsoil (top 10 cm) used in the original plot-scale implementations of EPIC and DAYCENT. This is justified, as total N-cycling microbial biomass (N and C) in topsoil are about one to two orders of magnitude higher than that in subsoils (> 10 cm). This suggests that N cycling mainly occurred in topsoil, given that exponential declines in soil C and N resources occur in subsoils (Tang et al., 2018). Non-agricultural soil nutrient and properties data used in the new ‘Mechanistic’ scheme were available for the top 30 cm soil layer from the most recent global compilation of such data across different biomes (Xu et al., 2015), but are still consistent with the topsoil (i.e., top 10 cm L1 + L2) configuration for N cycling as used in this work. This is supported by the fact that studies have shown topsoil depth (even 0-5 cm) mineralizable N to be representative of the 0–30 cm depth, as 0-15 cm N-cycling biomass drops considerably as it reaches 10 cm depth and is significantly higher than N-cycling biomass available at soil depths > 15 cm (Dessureault-Rompré et al., 2016).

Dessureault-Rompré, J., Zebarth, B.J., Burton, D.L. and Grant, C.A.: Depth distribution of mineralizable nitrogen pools in contrasting soils in a semi-arid climate. *Canadian Journal of Soil Science*, *96*(1), pp.1-11, 2016.

Tang, Y., Yu, G., Zhang, X., Wang, Q., Ge, J., & Liu, S.: Changes in nitrogen-cycling microbial communities with depth in temperate and subtropical forest soils. *Applied Soil Ecology*, *124*, 218-228, 2018.

Xu et al. (2015) is in ‘reference’ section in main manuscript

Comment 7: *The authors should mention what impact could have an increase in the spatial resolution of the simulation in order to better capture the soil / management heterogeneity.*

Spatial scale-dependent variation in soil/management heterogeneity can substantially influence how an analysis has to be approached; i.e. whether to opt for regional scale or more of plot-scale (<10m). Implications of various spatial resolution in soil ecology are manifold one of which is pertaining to microbial-plant community diversity. However, how heterogeneity in soil bacterial communities influences biogeochemical soil N cycling between local (< 10 m) and landscape (e.g., CONUS 12 km x 12 km in our case) scales still needs further research (O'Brien et al., 2016).

O'brien, S.L., Gibbons, S.M., Owens, S.M., Hampton-Marcell, J., Johnston, E.R., Jastrow, J.D., Gilbert, J.A., Meyer, F. and Antonopoulos, D.A.: Spatial scale drives patterns in soil bacterial diversity. *Environmental microbiology*, *18*(6), pp.2039-2051, 2016.