We thank the reviewer for his/her comments. In the following please find our point-by-point response to these comments, marked with an "R"

General comments:

This study attempts to optimize emissions of N_2O in the Corn Belt region of the US using observations of atmospheric N_2O mixing ratios from a single tall tower site, KCMP. However, the methods used are not appropriate and involve gross assumptions that influence the results.

More specifically, the method employed to optimize the emissions is over simplified: only one scalar is estimated and used to scale all the emissions in the inner domain (corresponding to the whole Corn Belt region) with no accounting for errors in the spatial distribution of the prior emissions.

R: Thank you for this comment. As a point of clarification, we did not scale all the emissions in the inner domain. We actually only scaled the agricultural emissions -- the sum of 'manure management', 'agricultural soil', 'indirect emission from agriculture', and 'agricultural waster burn' emissions -- in EDGAR42 in the grids that belong to the Corn Belt, which comprise 43% of the grids in the inner domain. In the revised manuscript, we have revised the description of the methodology:

"The third run was a scaled simulation whereby the surface flux in the grid cells belonging to the Corn Belt in the inner domain was the sum of natural soil emissions, EDGAR42 non-agricultural emissions, and a multiple of EDGAR42 agricultural emissions. The multiplier values are 3, 6, 12, and 25, the exact choice depending on the modeling period (Table 2). For the grid cells not belonging to the Corn Belt in the inner domain and the grid cells in the outer domain, the multiplier values are set to one". (lines 1-5, page 5 in the revised manuscript)

About the spatial distribution of the prior emissions, our reasons for using one scale factor are as follows:

(a) All inversion methods have their own advantages and disadvantages. We acknowledge that our inversion method has limitations associated with the assumption of prior spatial distribution, but the comparison between the modeled and observed spatial patterns in the atmospheric N_2O concentration (Fig. 6b in the revised manuscript) illustrates that the EDGAR42 as the prior distribution is acceptable for the Corn Belt. The Emissions Database for Global Atmospheric Research version v4.2FT2010 -- EDGARv4.2 uses a bottom-up approach and the default emission factors from the IPCC Guidelines (IPCC, 2006) to estimate the N_2O emissions except for road transport. So the spatial distribution of the N_2O emissions in EDGAR42 is based on the intensity of agricultural activities. Because the optimized spatial distribution of N_2O emission for

the Central U. S. is strongly similar to that of fertilization (Miller et al., 2012; discussed below), the spatial distribution in EDGAR42 is a good prior distribution.

(b) To further address the issue raised by the reviewer, we have performed additional modeling experiments to investigate the uncertainty introduced by the prior spatial distribution. In one modeling experiment, we enlarged the agricultural emissions for the whole modeling area (outer domain). In another experiment, we used the spatial distribution of the N_2O emission strictly following the spatial distribution of the fertilization rate. The constrained multiplier did not change much: the constrained multiplier for June 1-20 changes by less than 14% and less than 10% in these two modeling experiments, respectively. These results are summarized in Section 4.3 "Other sources of uncertainty".

(c) Miller et al. (2012) used a geostatistical approach to determine which of the existing emission databases, i.e., DLEM versus EDGAR32FT2000, is a more reasonable priori inventory for the central U.S. They conclude that EDGAR32FT2000 is the best priori spatial distribution for N_2O flux for their Bayesian inversion. They also show that a posteriori N_2O flux distribution from their geostatistical and Bayesian inversions are strongly similar to the spatial distribution of nitrogen fertilizer application, as shown in Figure R1.
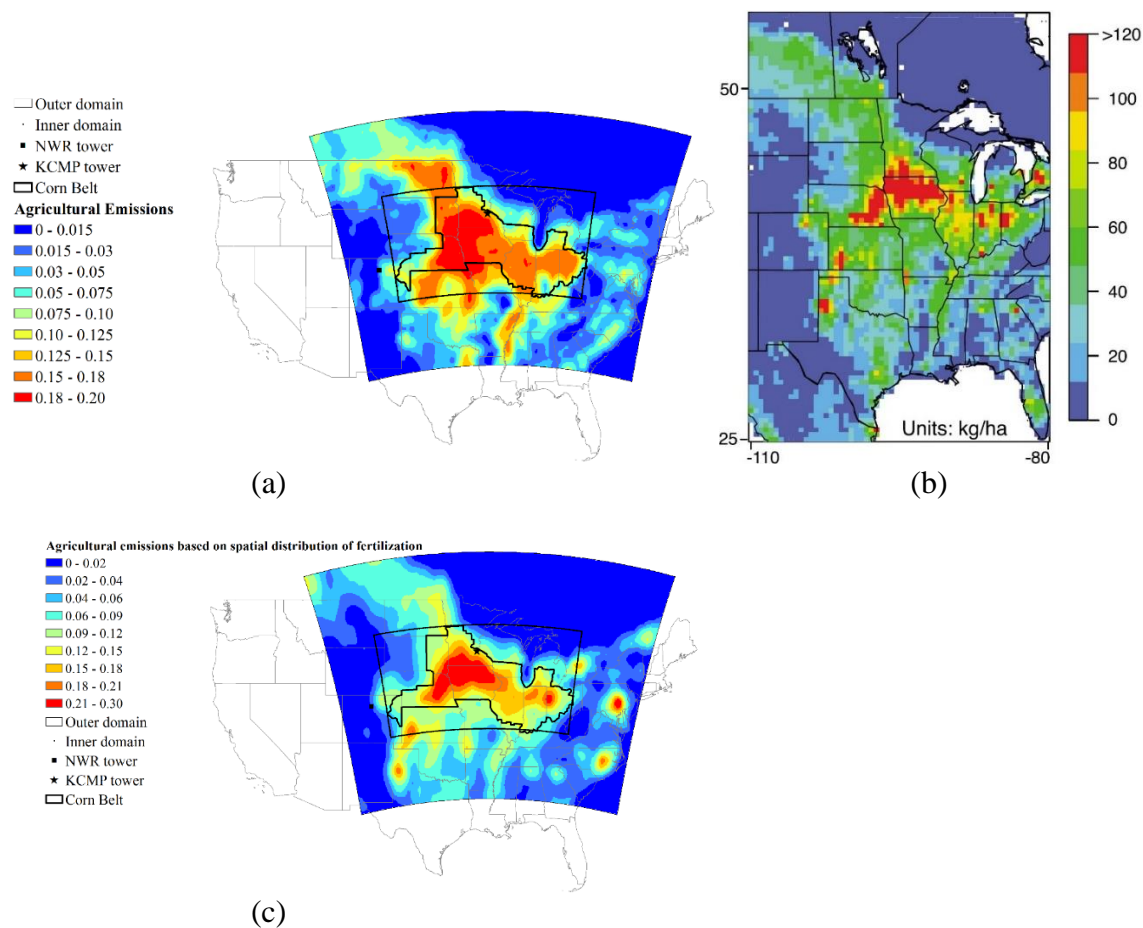

(a)

(b)

(c)

Figure R1. a: the agricultural emissions in EDGAR42 (unit: nmol m^{-2} s^{-1}). b: Figure 9 in Miller et al. (2012). c: agricultural emissions (nmol m^2 s^{-1}) based on the spatial distribution of N fertilization provided by Potter et al. (2010).

Like EDGAR32FT2000, the agricultural N$_2$O emissions in EDGAR42 are also dominated by emissions associated with nitrogen fertilizer application, and therefore have strongly similar distributions to the nitrogen fertilizer use pattern, as shown in the Figure R1 (a) and (b).

The figure for agricultural emissions in EDGAR42 (Figure R1 (a)) has been added in the revised manuscript as supplemental Figure S1.

Figure R2 shows the agricultural emission strength and the fertilization rate as functions of distance from the KCMP tower in the south wind sector (90° – 270°). The two quantities show strongly similar overall decreasing trends as the distance from the KCMP tower increases (upper panel in Figure R2), thus confirming that the dominant driver of spatial variations in N$_2$O agricultural emission is N fertilizer use. Within the radius of about 500 km from the tower, the emission strength and the fertilizer application rate are approximately constant with distance (lower panel in Figure R2), noting that 500 km is the distance from the KCMP tower to the south boundary of the Corn Belt. The lack of sensitivity to distance in the Corn Belt is the main reason why we have adopted a single scale factor for the whole Corn Belt.

The figure below has been added in the revised manuscript as supplemental Figure S4.
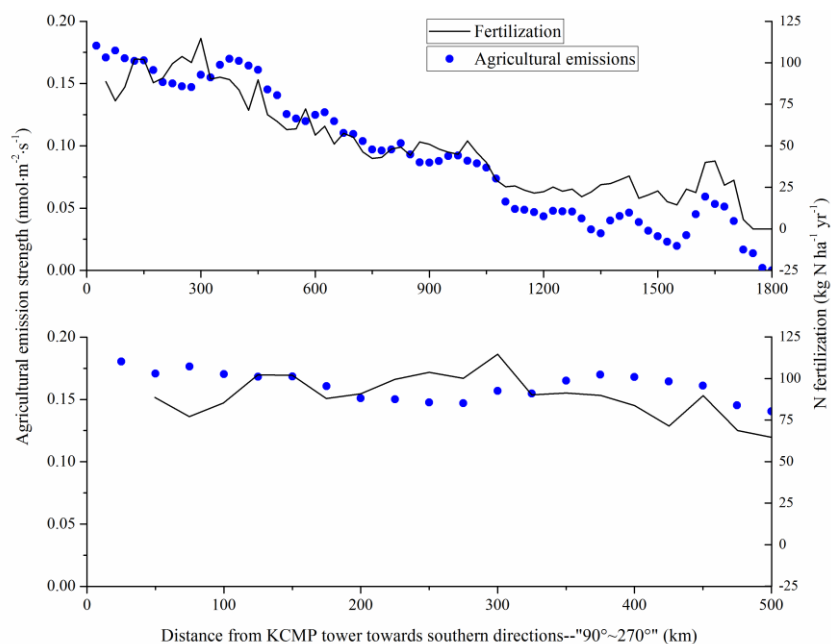

Figure R2. Changes of agricultural emission strength and the fertilization rate with distance from the KCMP tower in the south wind sector (90° – 270°).

To analyze the uncertainty related to the prior spatial distribution, we conducted two modeling experiments for June 1-20 whereby the prior emissions are based on the spatial distribution of fertilization using the default (multiplier: 1) and a scale (multiplier: 25.0) factor (Figure R3). Specifically,

The emission strength at a specific grid = mean areal emission strength for the outer (or inner) domain in EDGAR42 × fertilization at the grid ÷ mean areal fertilization for the outer (or inner) domain

The agricultural emissions based on the spatial distribution of fertilization for the default simulation are shown in Figure R1 (c). The constrained multiplier for June 1-20 based on these experiments is 17.7 – 25.3, which is close to the constrained multiplier of 19.0 – 28.1 in our manuscript (Table 2).
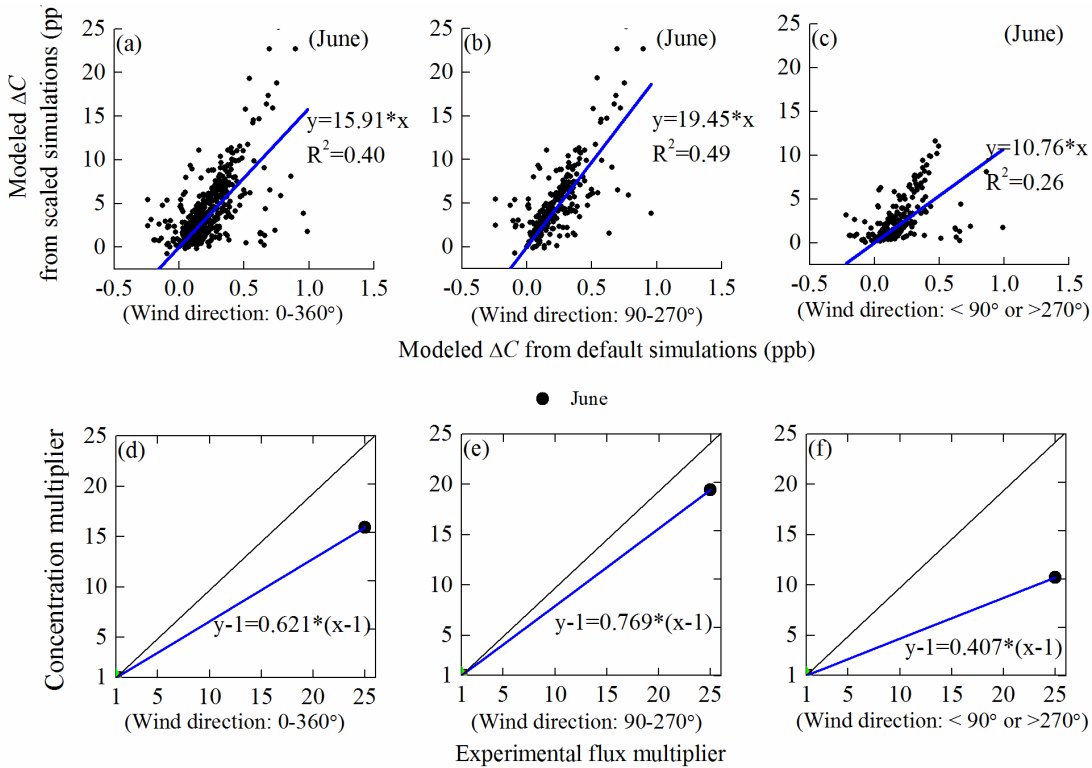

Figure R3. Relationships for different model runs between concentration multiplier and experimental flux multiplier. The modeled N_2O mixing ratio enhancement ΔC was obtained from default and scaled simulations for 185 m at the KCMP tower. The scale simulation shown in panels a – c uses a multiplier of 25.0. The regression slope in panels a – c is represented by the black circle in panels d – f. This figure is similar to Fig. 3 in the manuscript, but the prior distribution is based on the spatial distribution of the fertilization.

(d) In their inverse analysis, Xiang et al. (2013) reported a clear difference in the spatial distribution between the prior emission based on EDGAR4 and the posterior emission for

4

California. But their case is very different from ours: the N_2O emissions in the Corn Belt are comparatively more uniform spatially, whereas huge spatial heterogeneity exists in the N_2O emissions in California due to the heterogeneous land use. For example, large cities in California are emission hot spots: EDGAR 2000 indicate strong urban emissions from the San Francisco Bay area, and the coastal part of Los Angeles is strong source region in EDGAR 4 (Figs. 3 and 6 in Xiang et al., 2013). Their inverse results are improved after incorporating land use maps in their inversion.

Inversion studies for European N_2O emissions, where agriculture is also the dominant source of the N_2O flux, illustrate that the posteriori N_2O fluxes largely preserve the spatial distribution of a priori inventory in EDGAR4 (Corazza et al., 2011) or EDGAR4.1 (Bergamaschi et al., 2015).

(e) Consistent with Miller et al. (2012), Chen et al. (2016) also used the EDGAR42 emission as the priori emission for the Corn Belt. Chen et al. (2016) employed the STILT model and a Bayesian method to optimize each type of EDGAR42 emissions (e.g., direct emissions, indirect emission, and nature soil emission) for the Corn Belt, and obtained an optimum scale factor for each emission type, but spatial analysis was not included in their study. The present study includes an evaluation of the influence of N_2O emissions from Corn Belt on the spatial patterns of atmospheric N_2O concentration, as shown in Fig. 6.

The low correlation between the modelled and observed mixing ratios, even after optimization, indicates that the spatial distribution of the prior emissions is likely wrong.

R: We admit that the correlation is not as high as we would have liked, but the R^2 value between observation and modeling results after optimization is similar to those reported by other related studies. For example, the optimization in Kim et al. (2013) improved their R^2 for CO at the KCMP tower from 0.29 to 0.48, with an improvement of 0.19. The improvement of R^2 for N_2O in the present study is 0.33 (from 0.02 to 0.35). The R^2 value reported by Xiang et al. between simulation and observation for the atmospheric N_2O concentration in California is 0.29 – 0.33, and is similar to ours.

We have added the following text in the Discussion section:

"The improvement brought by the optimization to the correlation between the modeled and observed concentrations is similar to that reported by Kim et al. (2013) for CO for the same measurement tower. Their optimization improved their R^2 from 0.29 to 0.48, with an improvement of 0.19. The improvement of R^2 for N_2O in the present study is 0.33 (from 0.02 to 0.35). The R^2 value reported by Xiang et al. (2013) between optimized simulation results and the observation for the atmospheric N_2O concentration in California is 0.29 – 0.33, and is similar to ours. The remaining variations, not explained by the model, may be caused by the model's

inherent limitation in terms of simulating boundary layer transport processes and by errors in the spatial distribution of the prior emissions. "

"although the substantial noises (abrupt increases and decreases) in the observation have contributed to the low correlation coefficients between the simulated and the observed time series" (lines 13 and 14, page 6 in original manuscript) has been deleted.

Strong emissions close to the tower would have a very strong influence on the mixing ratio, if these are missing in the prior, then scaling all emissions equally will not correct for the error in the prior.

R: We agree with the reviewer on this point. However, there is no evidence that emissions near the tower are stronger than those near the edge of the Corn Belt (Figure R2). Since the prior emission is based on agricultural activity data reported at a fine spatial resolution, it is unlikely that large emission hot spots would have been missed by the inventory. (According to USDA Agricultural Resource Management Surveys -- https://www.ers.usda.gov/data-products/chart-gallery, the Corn Belt is one of the agricultural regions in the world that have the most accurate fertilizer use statistics.)

Please also refer to our response to Major concern 1) below.

There are also concerns about the representation of the background mixing ratio, which was based on observations at Niwot Ridge, a high altitude site (>3000 m above sea level).

R: we select Niwot Ridge as the background site for four reasons: (a) it is located at the upwind side of the Corn Belt; (b) it is basically outside of the zone of influence of the Corn Belt emissions (Fig. 6); (c) observations at Niwot Ridge are continuous in time, whereas observations at other sites are made only once or twice a day; (d) we want to be consistent with former studies, i.e., Griffis et al. (2013) and Chen et al. (2016), who also use Niwot Ridge as the background site.

In the revised manuscript, we added the following text:

'Another advantage of using NWR as opposed other NOAA monitoring sites is that the observation at NWR is continuous in time whereas measurements at other sites are made only once or twice per day. Griffis et al. (2013) and Chen et al. (2016) also used the observation at NWR as the background concentration'. (lines 22-24 in page 5)

To analyze the uncertainty caused by representation of the background mixing ratio, we have done additional inverse analysis using the AMT (Argyle, located in Maine) as the background site. The two background sites show slightly different concentrations (Figure R4). The absolute change of the constrained multiplier is $1.4 - 1.9$, $1.9 - 3.3$, $1.1 - 1.8$, and $0 - 0.1$ for $1^{st} - 20^{th}$ in

June, August, October, and December, respectively; corresponding relative change in the optimized flux is 6.8 – 8.0 %, 20.4 – 25.4 %, 32.4 – 38.3 %, and 0 – 2.8 % (Table R1). As a result, the inverse uncertainty caused by background selection is larger in August and October than that in June and December.

These new results are included in the online supplement (Figure S5 and Table S4).
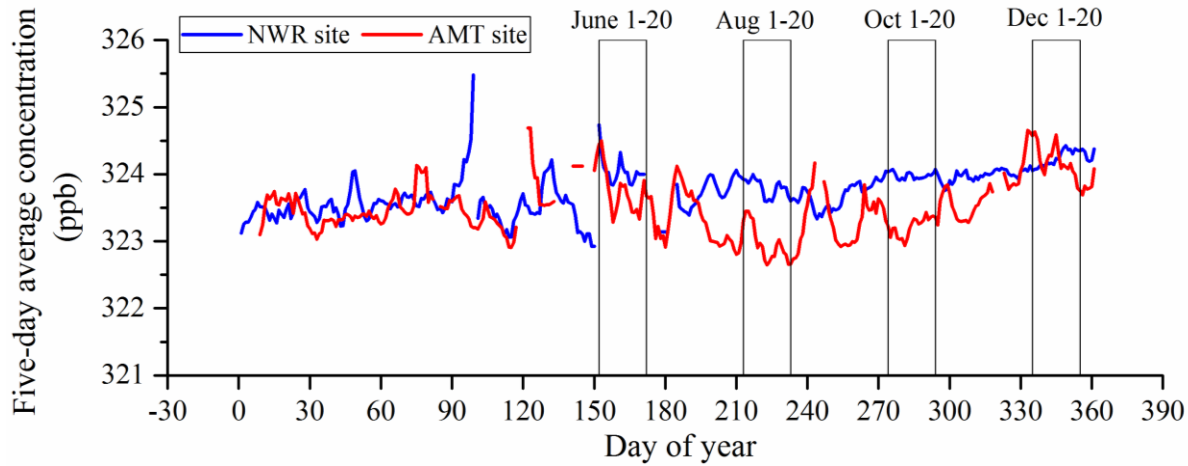

Figure R4. Observed N_2O mixing ratio at the Niwot Ridge and "Argyle, Maine" (AMT) sites.

Table R1. Experimental and optimized flux multiplier M_F **using AMT as background site**. Values in brackets are constrained agricultural emission flux in units of nmol m^{-2} s^{-1}.

Time	June 1 – 20	August 1 – 20	October 1 – 20	December 1 – 20
Experimental	0, 1, 25	0, 1, 12	0, 1, 3	0, 1, 6
Optimized [a]	20.4 (3.12)	11.2 (1.71)	4.5 (0.68)	3.0 (0.47)
Optimized [b]	24.3 (3.71)	14.0 (2.15)	5.2 (0.80)	3.7 (0.57)
Optimized [c]	30.0 (4.59)	16.3 (2.49)	6.5 (0.99)	4.4 (0.67)

Notes: a, b, c: using observation data at heights of 32, 100, and 185 m, respectively.

Table 2 in the present study. Experimental and optimized flux multiplier M_F. Values in brackets are constrained agricultural emission flux in units of nmol m^{-2} s^{-1}.

Time	June 1 – 20	August 1 – 20	October 1 – 20	December 1 – 20
Experimental	0, 1, 25	0, 1, 12	0, 1, 3	0, 1, 6
Optimized [a]	19.0 (2.91)	9.3 (1.43)	3.4 (0.52)	3.0 (0.47)
Optimized [b]	22.5 (3.44)	11.6 (1.77)	3.8 (0.59)	3.6 (0.55)
Optimized [c]	28.1 (4.29)	13.0 (1.99)	4.7 (0.72)	4.3 (0.66)

Notes: a, b, c: using observation data at heights of 32, 100, and 185 m, respectively.

Furthermore, there are concerns about the accuracy of the modelled atmospheric transport (see specific comments below). Additionally, in places, the references do not support the statements made (some examples are given under the specific comments below).

R: these concerns will be taken care of in the following responses.

Major concerns:

1) Account for errors in the spatial distribution of the prior emissions. This would mean solving scalars for the emissions (or for the emissions themselves) for each grid cell or subregions of the inner domain.

R: In principle, we agree with the reviewer that we should solve the scalar for each grid cell. But in practice, this would be extremely difficult, if not mathematically impossible, given that we only have observation at one location but there are 7991 grid cells in the Corn Belt. The problem would become wildly unconstrained. This is a limitation of our study, and is also a limitation of other N_2O inversion studies, e.g., Mikaloff Fletcher et al., 2004; Carouge et al., 2010; Bousquet et al., 2011; Chen et al., 2016. For example, Miller et al. (2012) used data at four tower sites to do inverse analysis, and pointed out that 'with limited atmospheric data, the Bayesian inversion cannot fully correct discrepancies in spatial distribution among inventories. In this case, the geostatistical inversion can inform our choice of a priori inventory'. In the Bayesian inversion analysis reported by Chen et al. (2016), they were only able to solve a single scalar for the whole Corn Belt.

In response, we have added the following text:

"In the EDGAR inventory, the N_2O emission is determined with the IPCC-type methodology using agricultural activity data and standardized emission factors. Our study makes an implicit assumption these emission factors are biased similarly in all model grids in the Corn Belt."

---

2) Improve the estimate of the background concentration to account for the origin of the air masses when entering the inner domain

R: Unlike some other modeling studies (e.g., Kort et al. 2008), here we compare the modeled N_2O concentration enhancement, instead of the absolute concentration itself, with the observation. The concentration enhancement is calculated as the difference in N_2O mole fraction between the default or scaled simulation and the background simulation. The main purpose of doing this is actually to limit the effect of air mass origin. This effect is further reduced by using the initial and boundary conditions produced by a global model.

8

We have added the following text:

'The initial and boundary conditions for the meteorological field were obtained from the weather forecast model Global Forecast System (ftp://nomads.ncdc.noaa.gov/GFS/analysis_only), and the initial and the lateral boundary conditions for the N_2O concentration was provided by Model for Ozone and Related Chemical Tracers (MOZART) version 4 (http://www.acom.ucar.edu/wrf-chem/mozart.shtml)'. (lines 15-19, page 4)

---

3) Account for uncertainties in the observed and modeled concentrations

R: One source of model uncertainty is related to the spatial extent to which the scaling is applied. In all of our previous scaled experiments, the emission scaling was restricted to the grid cells in the Corn Belt. We have done additional experiments by applying the scaling to the whole modeling area (outer domain). The results are summarized in Table R2. Compared with the original results (Table 2), the change to the optimized flux is less than 14%.

This table is added to the online supplement (Table S2).

Table R2. Experimental and optimized flux multiplier M_F **using enlarged agricultural emissions for the whole modeling domain**. Values in brackets are constrained agricultural emission flux in units of nmol m^{-2} s^{-1}.

Time	June 1 – 20	August 1 – 20	October 1 – 20	December 1 – 20
Experimental	0, 1, 25			
Optimized [a]	16.8 (2.58)	8.4 (1.28)	3.1 (0.47)	2.8 (0.43)
Optimized [b]	19.4 (2.97)	10.1 (1.54)	3.4 (0.52)	3.2 (0.49)
Optimized [c]	24.1 (3.68)	11.3 (1.72)	4.2 (0.64)	4.3 (0.66)

Notes: a, b, c: using observation data at heights of 32, 100, and 185 m, respectively.

In another experiment, we used a spatial distribution of the N_2O emission that strictly follows the spatial distribution of the fertilization rate. The results are summarized in Table R3. Compared with the original results (Table 2), the change to the optimized flux is less than 10%.

This table is added to the online supplement (Table S3).

Table R3. Experimental and optimized flux multiplier M_F **using the spatial distribution of a prior emission in proportion to fertilization**. Values in brackets are constrained agricultural emission flux in units of nmol m^{-2} s^{-1}.

Time	June 1 – 20	August 1 – 20	October 1 – 20	December 1 – 20
Experimental	0, 1, 25			
Optimized [a]	17.7 (2.70)	9.2 (1.41)	3.3 (0.51)	3.0 (0.46)
Optimized [b]	20.9 (3.19)	11.2 (1.71)	3.7 (0.57)	3.5 (0.53)
Optimized [c]	25.3 (3. 86)	12.5 (1.92)	4.6 (0.70)	4.2 (0.65)

Notes: a, b, c: using observation data at heights of 32, 100, and 185 m, respectively.

In the revised manuscript, we have added an additional section to describe the uncertainties in the observed and modeled results, namely,

'4.3 Other sources of uncertainty

Errors in the inverse analysis can arise from uncertainties in the model simulations and in the observed concentration. A large source of modeling uncertainties is related to the uniform scaling factor applied to the agricultural emissions in all the grid cells in the Corn Belt. Since sources closer to the observation tower have a stronger influence on the observed concentration than those farther away, a uniform scaling may bring some uncertainty into the inverse analysis. Miller et al. (2012) reported that the spatial patterns of the N_2O fluxes from both geostatistical and Bayesian inversions are strongly similar to that of nitrogen fertilizer application rate. We have analyzed the agricultural emission strength and the fertilization data and presented the results as functions of distance from the KCMP tower in the south wind sector (90° – 270°; Figure S4). The two quantities show strongly similar overall decreasing trends as the distance from the KCMP tower increases, thus confirming that the dominant driver of spatial variations in N_2O agricultural emission is N fertilizer use. Within the radius of about 500 km from the tower, the emission strength and the fertilizer application rate are approximately constant with distance, noting that 500 km is the distance from the KCMP tower to the south boundary of the Corn Belt. Because of the lack of sensitivity to distance within the Corn Belt, the uncertainty caused by the uniform scaling of the prior is probably not too large.

To further investigate the inverse uncertainty, we have done additional modeling simulations by applying the scaling to the agricultural emissions in the whole modeling domain (including both the inner and the outer domain), instead of the Corn Belt only. The results, summarized in Supplementary Table S2, show that changes to the optimized flux is less than 14%. We have also done simulations by using the spatial distribution of the N_2O emission that strictly follows the spatial distribution of the fertilization rate. The results illustrate that changes to the optimized flux is less than 10% (Table S3).

Another uncertainty in the modeling and the subsequent inverse analysis is the model background. Unlike some other modeling studies (e.g., Kort et al. 2008), here we compare the modeled N_2O concentration enhancement, instead of the absolute concentration itself, with the observation. The concentration enhancement is calculated as the difference in the N_2O mole fraction between the default or scaled simulation and the background simulation. The main purpose of doing this is to limit the effect of air mass origin. This effect is further reduced by using the initial and boundary conditions produced by a global model (MOZART4, http://www.acom.ucar.edu/wrf-chem/mozart.shtml).

The observation background is also a source of inverse uncertainty. In the present study, we used NWR as the background site for reasons stated in Section 2.3. We have also used AMT located downwind of the simulation domain (Fig. 1) as the background site to investigate the uncertainty caused by the background mixing ratio. The N_2O mixing ratio is nearly the same between the two background sites during June and December, but the mixing ratio at AMT is 0.6 – 0.8 ppb lower than at NWR in August and October (Figure S5). The relative change in the optimized flux is 7 – 8% and 0 – 3% in June and December, respectively, and is 20 – 25% and 32 – 38% in August and October, respectively (Table S4).

Uncertainties also exist in the monitoring data obtained at the KCMP tower. The abrupt increases and decreases by as much as 31 ppb in less than 2 hours) in the observed N_2O mixing ratio (e.g., at day of year 160 and 164, Fig. 3a) are clearly measurement noise related to the sampling and calibration procedures. Such large measurement noises may be one reason for why even after optimization, the correlation between the modeled and observed concentrations is not very strong'.

---

4) Address concerns about the model representation of concentrations at the KCMP site, e.g., the vertical layer selected to represent the observations

R: we have added the following text in the revised manuscript,

'The heights of layers 1, 2, and 3 are around 30, 100, and 190 m, respectively, so we compared the modeling results of these layers with observations at height of 32, 100, and 185 m, respectively'. (lines 10-12, page 4)

5) Include observations also from the NOAA sites to help constrain the emissions

R: Thank you for this suggestion. Our strategy is to use these data as independent evaluation of the modeled concentration (revised Section 3.5). In response to this comment, we have tried to constrain the emission using observations made at WBI. WBI is located in the center of the Corn Belt. The other NOAA sites located outside the Corn Belt are no sensitive to the emissions in the

Corn Belt according to our model simulations (Figure 6). The results are summarized in the following Figure R5 and Table R4.
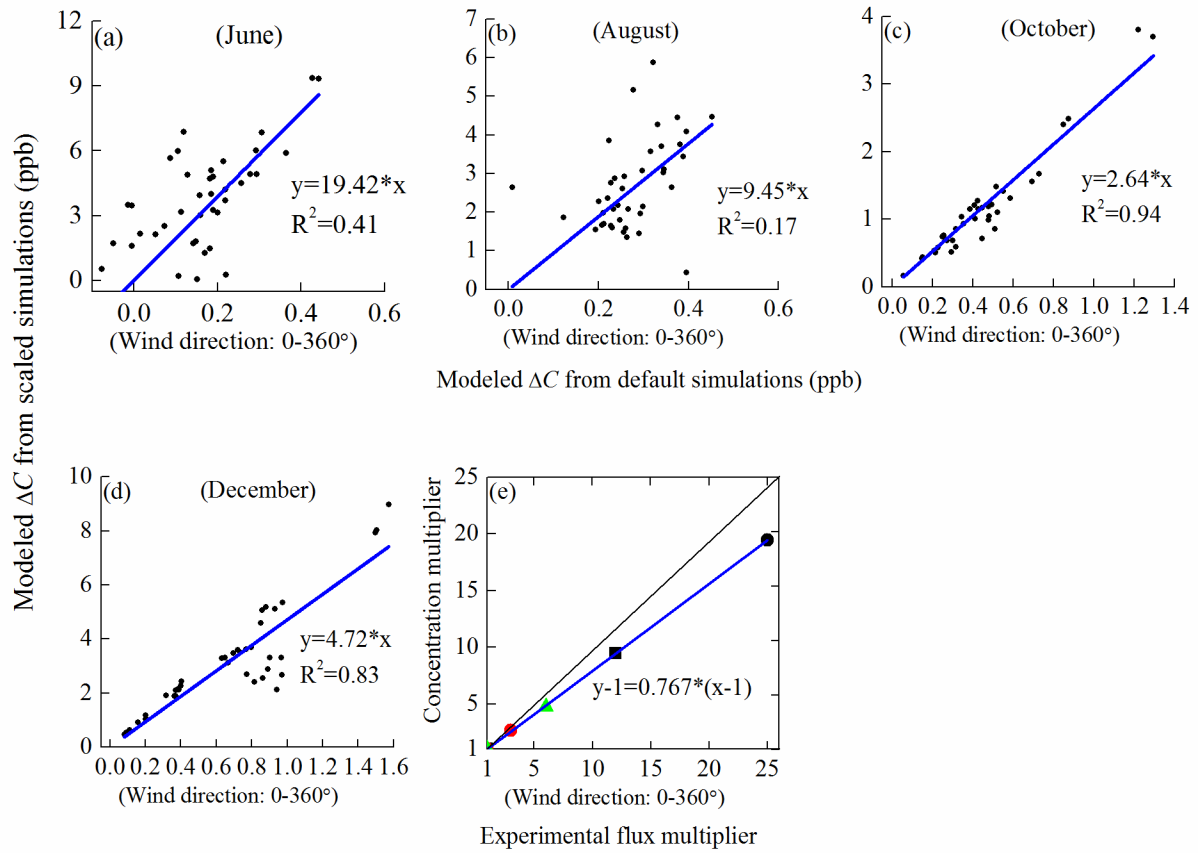

Figure R5. Inverse equations derived from the NOAA discrete monitoring at WBI (UTC hours 19 and 20 only).

Table R4. Experimental and optimized flux multiplier M_F **using NOAA discrete monitoring at WBI**. Values in brackets are constrained agricultural emission flux in units of nmol m^{-2} s^{-1}.

Time	June 1 – 20	August 1 – 20	October 1 – 20	December 1 – 20
Experimental	0, 1, 25	0, 1, 12	0, 1, 3	0, 1, 6
Optimized	28.8 (3.28)	2.2 (0.69)	–	0.38 (0.99)

The constrained multiplier for June using the data at WBI site (28.8) is close to the inverse results using the data at the KCMP site (28.1, measurement height 185 m), indicating large emissions in June. For August and December, the constrained multipliers are smaller than those obtained from KCMP, possibly because of measurement bias errors in these low flux periods due to the limited number of observations (one or two discrete samples per day). This bias problem is most obvious in October: The observed N_2O mixing ratio at WBI near the center of the Corn Belt is even smaller than that at the background, as indicated by the negative ΔC (Table 3).

12

Producing a negative ΔC would require a negative flux in the model domain, which is unreasonable.

Specific comments:

P1, L18-19: EDGAR agricultural emissions (direct and indirect) have no seasonality, therefore, it is expectable that during the peak emission season, the inventory emissions will be much lower. However, what is important is how the inventory compares to the annual mean emissions.

R: This is a good point. In the revised manuscript, we have added the following discussion:

'The annual N_2O emission flux from Corn Belt calculated using the IPCC inventory methodology is 0.19 nmol m^{-2} s^{-1} (Griffis et al., 2013), slightly smaller than the annual flux from Corn Belt in EDGAR42 of 0.21 nmol m^{-2} s^{-1}. The constrained annual N_2O emission flux over the central U. S. in Miller et al. (2012) is around 0.40 nmol m^{-2} s^{-1}, namely, around two times that from the IPCC inventory methodology. The constrained N_2O emission fluxes during the four study periods in the present study are all larger than those for the same periods in Miller et al. (2012), so the constrained annual N_2O emission fluxes from the Corn Belt should be larger than the values in EDGAR42 and those calculated using the IPCC inventory methodology'. (lines 27-32, page 10 in the revised manuscript)

P1, L17: How were the emissions optimized? This should be stated in the abstract.

R: The optimization method is now described the abstract:

'We derived a simple equation to relate the emission strengths to atmospheric N_2O mixing ratios, and used the equation and hourly atmospheric N_2O measurements at the KCMP tall tower in Minnesota to constrain agricultural N_2O emissions'. (lines 14-16, page 1)

P1, L19: The authors state the "total emissions" which would imply the integrated emission (i.e., in units of mass) but give the emission in units of flux. Rather they should say that this is the "mean" emission.

R: it has been corrected in the revised manuscript. (line 20, page 1)

P1, L21: Do the authors refer here to their optimized fluxes? It is not clear. If they are referring to the optimized fluxes (which are higher in June than EDGAR) this does not indicate that the IPCC emissions are underestimated. Firstly, EDGAR is an independent inventory and is not the

same as what is reported to the IPCC (also note the IPCC does not estimate emissions). Secondly, only annual mean emissions are reported to the IPCC, so it is not possible to compare the optimized emissions (based on about 80 days) with the annual mean. Lastly, the fact that concentrations simulated using the optimized emissions (i.e., based on the comparison with observations at KCMP tower) compare well with NOAA observations does not necessarily mean that the fluxes are an improvement. This depends on the footprint of the NOAA observations and if there is a notable improvement compared to using the prior, i.e., EDGAR.

R: we added 'after optimization' in the revised manuscript. (line 22, page 1)

About the comparison between emissions calculated using the IPCC methodology and those from the EDGAR database, please refer to our responses above.

P1, L27: ODP needs to be specified the first time

R: It has been specified in the revised manuscript. (line 28, page 1)

P1, L28: May refer to Prather et al 2015, lifetime of 116 years

R: This citation has been added. (line 29, page 1)

P1, L36: Both references given are only about European emissions (not about North American or the Corn Belt). Please either change this sentence use references about the Corn Belt.

R: Three citations have been added for the Corn Belt, namely, Griffis et al., 2013; Zhang et al., 2014; and Chen et al., 2016. (lines 1-2, page 2)

P2, L2: There is nothing empirical about the equations used in inverse methods (i.e., top-down methods). The optimization can be done in various ways but always based on statistical theorems, e.g., Baye's Theorem.

R: This sentence has been changed to 'Top-down estimates of the emissions are usually determined from observed atmospheric N_2O mixing ratios, a transport model, and model optimization'. (lines 4-5, page 2)

P2, L22: This may be true of STILT but it is not generally true of Lagrangian models. Lagrangian models can also be used to simulate fields of mixing ratios in space and time. Please correct this statement.

14

R: 'Lagrangian models' has been changed to 'STILT' in the revised manuscript. (line 26, page 2)

P2, L24-25: This is not true. The reference Corrazza et al. presents a Eulerian model based inversion study for Europe and there are multiple other studies, e.g., Bergamaschi et al., 2015.

R: We have changed the sentence 'no modeling studies have been published on the relationship between the spatial characteristics of surface emissions and the atmospheric N_2O mixing ratio at the regional scale' to 'no modeling studies have been published on the relationship between the spatial characteristics of surface emissions and the atmospheric N_2O mixing ratio for the Corn Belt'. (lines 28-29, page 2)

P2, L31: "Empirical" is not correct in this context. Empirical must be observation/experiment based.

R: 'Empirical' here has been deleted in the revised manuscript.

P4, L17: Changes in tropospheric N_2O mixing ratio depend on the surface fluxes but also the atmospheric transport. Please add "atmospheric transport".

R: It has been added. (line 34, page 4)

P4, L19: Vertical diffusion does not contribute much to the vertical mixing of N_2O, but rather turbulence in the BL and convection. Please correct this.

R: 'vertical diffusion' has been changed to 'turbulent diffusion and convection' in the revised manuscript. (line 35, page 4)

Fig 1: By "default emission" do the authors mean "prior emission". If the prior is meant this should be stated (and not "default").

R: 'default emission' has been changed to 'prior emission'.

Eq. 1: Since N_2O is inert in the troposphere, the transport is linear (and can be defined as a linear operator) and since the authors are considering the change in mixing ratio with respect to a background, why is "a" not equal to one? In other words, why should Mc not be equal to Mf?

R: The primary reason for this may be because the KCMP tower is close to the northern boundary of the Corn Belt. Consider an extreme condition – all winds are northern wind: increase of the surface emission in the Corn Belt would have very limited influence on the atmospheric N_2O concentration enhancement, and 'a' would be less than one.

P4. L35-37: The background defined for the observations at KCMP tower should be equivalent to the background in your model. The modeled background mixing ratio is that modeled using natural soil and non-agricultural emission over the whole domain plus some initial boundary mixing ratio (which is not discussed). However, the mixing ratio observed at NWR could be very different from the model background. NWR is a high altitude station (>3000 m above sea level). The authors should re-do the calculations using a background comparable to their model background. Also, they should state what was used for the initial mixing ratio at the model outer boundary. Was this based on global model simulations or other?

R: In response to this comment, we have added the following text to Section 2.3:

'Next, we used the observed enhancement, ΔC, defined as the actual concentration observed at the KCMP tall tower minus a background concentration and adjusted for a small gradient in the modeled N_2O mixing ratio between KCMP and the background concentration site from the background simulation, to constrain the flux multiplier (and the surface emission flux). A numerical example is given in the Supplementary Information on how this is done.'

The reviewer is correct in pointing out that the observed background location does not match the background location used by the model. In our inverse analysis, we made adjustment for the spatial mismatch. A numerical example is given in the online supplement to illustrate how this is done. For the reviewer's convenience, the online information is duplicated here:

'The concentration observed at NWR in south winds during June 1-20 is 323.981 ppb and that at the height of 185 m on the KCMP tower is 329.424 for the same period, giving an (unadjusted) concentration enhancement of

Term 1 = 329.424 - 323.981 = 5.443 ppb.

The concentration given by the background simulation for the height of 185 m at the KCMP grid point is 321.352 ppb. In this simulation, the agricultural emission is set to zero in the Corn Belt. The concentration given by the default simulation for the same height in this grid point is 321.606 ppb. In the default simulation, the emission intensity is the prior (that is, the flux given by EDGAR42). The modeled concentration enhancement is

Term 2 = 321.606 - 321.352 = 0.254 ppb.

If EDGAR42 were perfect (and the model had no errors), Term 2 would match Term 1 perfectly. However, the fact that Term 2 is much smaller than Term 1 indicates that the prior emission is biased too low.

We use the difference between the observed and the model default concentration enhancement to find an optimal flux multiplier. But before doing so, we should adjust Term 1 slightly to account for the fact the observed background site is NWR but the modeled background is at the KCMP grid. This adjustment for this spatial mismatch is made using the background simulation results. It is worth mentioning that the influence of emissions from the Corn Belt on the atmospheric N_2O concentration at the NWR site is very limited. For example, the modeled N_2O concentrations at the NWR site from the background and scaled simulations are quite close (Table S1). According to the background simulation, the mean concentration at the NWR and the KCMP model grids is 321.257 and 321.352 ppb, respectively. So the amount of adjustment is

Term 3 = 321.352 - 321.257 = 0.095 ppb.

The concentration multiplier constrained by the observation is

M_C = (Term 1 – Term 3) / Term 2 = 21.055

Substituting this value in the inversion Equation (1) and noting that a = 0.740 for south winds, we obtain a constrained flux multiplier M_F = 28.102'.

The initial and boundary conditions have been described in section '2.2 Model setup', as

'The initial and boundary conditions for the meteorological field were obtained from the weather forecast model Global Forecast System (ftp://nomads.ncdc.noaa.gov/GFS/analysis_only). The initial and boundary conditions for the N_2O mixing ratio for each modeling period were obtained from Model for Ozone and Related Chemical Tracers (MOZART) version 4 (http://www.acom.ucar.edu/wrf-chem/mozart.shtml)'. (lines 15-19, page 4)

P5. L1-2: The authors mean that the observed enhancement should match the difference between the default and background simulations?

R: Yes, 'the observed enhancement should match the difference between the default and background simulations', if the prior emission strength is accurate. The text has been clarified.

17

P5. L2-3: The disagreement could also be due to errors in the assumed background based on NWR (see comment above).

R: please see the responses above.

P5. L3-5: The model errors are critical to the results obtained and need to be accounted for in the method. It is not acceptable to leave the model errors to the discussion.

R: Thank you for this suggestion. We have moved this part to the Results section. Specifically, we have moved the PBL height analysis to Section 3.1 "Mixing height" and the spatial comparison to Section 3.5 "Spatial variations of modeled N_2O concentration".

P5. L5-12: The method presented assumes that the adjustment needed to the prior fluxes is constant throughout the inner domain. This requires a further assumption that the change in mixing ratio at KCMP is equally sensitive to all fluxes in the inner domain. This assumption needs to be stated as it is a very important assumption and has large implications on the results obtained. (This is known as "aggregation error" in inverse problems). A more rigorous approach would be to estimate an emission-scaling factor for sub regions within the inner domain, e.g. based on the grid cells of WRF-Chem, which would reduce the aggregation error.

R: These assumptions are stated. (lines 7-8, page 5)

Fig. 2a-c: I am surprised that the correlations are so low for the comparison of scaled and default simulated mixing ratios. Can the authors please explain why the correlations are not much higher? Also, it should be stated in the caption that figures a-c are using an emission-scaling factor of 25.

R: Please refer to our response to the general comments.

'The scale simulation shown in panels a – c uses a multiplier of 25.0' has been added to the caption of Fig. 3.

Fig. 3. and P6, L25-35: The correlations even with the optimized fluxes are quite low. Did the authors examine different model layers to see which one best represents KCMP observations at 185 m above ground level? Are other species measured at KCMP, e.g., CH4 or CO? If so, how well does WRF-Chem model capture the variability in these? This could tell you about the general performance of your transport model.

R: Please refer to our response to the general comments above.

According to Kim et al. (2013), the WRF-Chem model explains 29 % and 48 % of the variability in CO observed at the KCMP tower before and after optimization. This point is now noted in the text. No long-term measurement of CH_4 is available at KCMP.

P7, L14-18: Turbulent mixing is the main process how tracers are mixed in the boundary layer, diffusion plays only a minor role, if any.

R: 'diffusion' has been changed to 'transport' or 'vertical mixing' in the revised manuscript. (lines 6-9, page 6)

P7, L38: The authors should also state the local time for the data selected since this is what is relevant in the selection.

R: In the notes of Table3, we have added the information on conversion of local time to UTC time for each site.

P8, L9-11: This is rather an indication that the background (based on NWR) is far too simplified. The background calculation does not take into account where the air masses are actually coming from into the domain, neither the altitude nor the horizontal direction.

R: If the low bias was only for one moment (or one day), one possible reason may be related to the background. But because the bias was for the average of twenty days, we prefer to believe that it is caused by uncertainty in the observations.

P8, L11-15: The variability in the observations should not be considered as a problem, but rather as information. The fact that the model cannot capture the variability shows that there is a problem in the model, in the fluxes and/or the transport.

P8, L18-20: Low variability of the observations (or small enhancements above background) do not limit their usefulness in inversions. Rather, it is also important information, which indicates that the emissions are low during these periods. It does require, however, an accurate background estimate (which is just as important in periods of high emissions). Furthermore, the variability of the observations gives important information about the sources, as the atmospheric transport is not constant in time, i.e., changes in atmospheric transport mean that the sensitivity of the observations to different areas changes in time. Lastly, why were the NOAA observations not used in the optimization as well?

R: We agree that the variability of the observations gives important information. That is why we modeled the multiple-day and sub-day dynamics for the KCMP site, where hourly data is available, as shown in Fig. 4. There is no hourly data from the NOAA sites.

We also wish to point out that some of the variability in the observations are clearly measurement errors. For example, the unreasonably negative ΔC at WBI, a site located at the center of the Corn Belt, suggests large measurement uncertainties associated with flask sampling (Table 3). When the measured variabilities are dominated by measurement errors, the resulting inversion results will also be subject to large errors.

We have improved the clarity of the text here:

"For August, October, and December with much weaker emissions than in June, the agreement between the observed and modeled ΔC is not as good. The observed N_2O mixing ratio at WBI near the center of the Corn Belt is even smaller than that at the background during October 1-20, as indicated by the negative ΔC at WBI, a site located at the center of the Corn Belt (Table 3), which is unreasonable and suggests large uncertainties in the concentration measurements. The largest disagreement occurred in the October period: the modeled mean ΔC with $M_F = 3.0$ is 0.81 ppb for the four sites in the model domain but the observed mean ΔC is actually 0.01 ppb.

By measuring spatial variations in atmospheric N_2O, an observational network consisting of multiple sites has the potential to help constrain inverse analysis using Eulerian tracer transport models, if the measurement reflects true natural variations and is unaffected by measurement uncertainties. Given the flask measurement uncertainties noted above, such N_2O inversion would be difficult for low emission periods. The large measurement uncertainties may have explained why Miller et al. (2012) limited their geostatistical inversion to an early summer period."

Section 4.2: An important analysis that is missing here is which model layers best represents the observations at the three heights on the tower, especially 185 m height from which observations were used for the optimization. Since the WRF-Chem model overestimates the vertical gradient between 32 and 185 m both during daytime (factor of 8) and nighttime (factor of 4) there appears to be an issue with vertical mixing between the layers used to represent the 32 and 185 m heights. It seems that the vertical mixing is too weak and that fluxes at the surface are not adequately mixed in the BL up to, e.g., 185 m, such that the model underestimates the concentrations at this height.

R: The modeled vertical mixing, represented by the N_2O concentration gradient between 32 m and 185 m, is slightly smaller than observation (modeled gradient – observed gradient = 0.22 ppb) during the daytime. The larger modeled gradient than observation at night may be related to the underestimation of vertical mixing in the modeling during night periods, as pointed out by the reviewer.

Figure R6 shows the actual comparison of the observed and modeled vertical gradients between the 32-m and the 185-m heights. The model is in better agreement with the observation on day 165, 167, and 172 than on others. Meanwhile, the modeled mixing heights (shown in the figure below), reflecting the strength of vertical mixing, are strongly similar to NCEP-NARR data.

It is possible that the N_2O concentration at the height of 185 m was underestimated and the concentration at the height of 32 m was overestimated. The true emission strength should lie somewhere between the two fluxes constrained by using the data from the 185 m and 32 m height. Even though for brevity we only presented the data graphs for the 185-m height, the inversion results using measurements at the other two heights (32 m and 100 m) are summarized in the data tables (Table 2). In the abstract and the conclusion section, the optimized flux is given in a range of values for each period (e.g., lines 20 - 21, page 1; lines 20 - 23, page 12).

We have added the two figures below into the revised manuscript as supplemental figures (Figures S2 and S3).
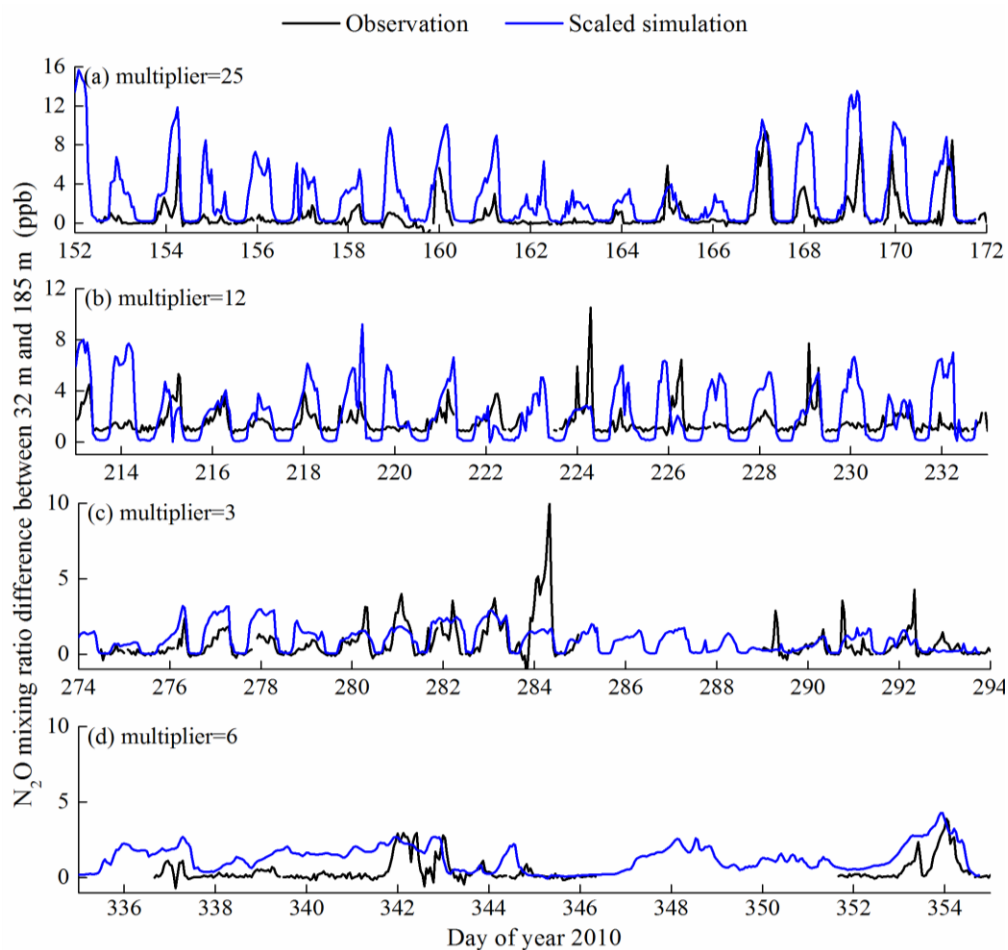

Figure R6. Observed and modeled vertical N_2O mixing ratio gradients (32 m minus 185 m) at the KCMP tower site.
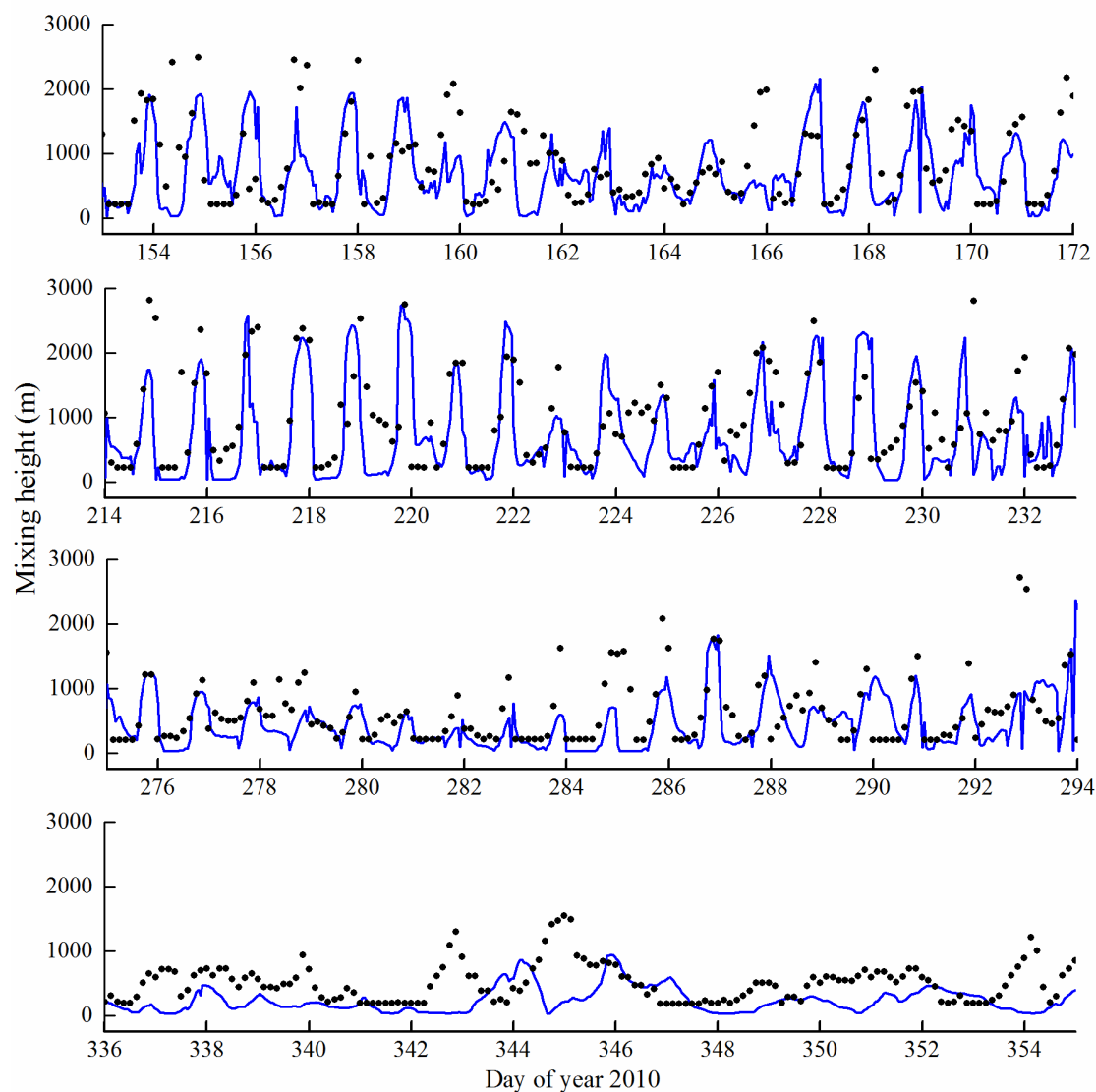
Figure R7. Simulated mixing heights and the NCEP-NARR data (dots) for different periods.

P9, L21-25: The statement about the Miller et al. (2012) study is not correct. Miller et al. use both a geostatistical and a Bayesian inversion for the optimization with a spatial resolution of 1x1 degrees. They do not fit a single scaling factor for all of the central US.

R: 'a single emission scaling' has been changed to 'the emission scaling' in the revised manuscript. (line 23, page 10)

If we understood correctly, Miller et al. (2012) used the geostatistical method to help inform their choice of a priori inventory (i.e., DLEM versus EDGAR32FT2000) for their Bayesian inversion, and eventually selected EDGAR database as the prior inventory. Four sites (LEF, WBI, BAO, and WKT) were used for the Bayesian inversion: there is a footprint for each tower site,

and a scale factor is applied the whole footprint area. Their final estimate is the average of four scaled values.

Technical comments:

Fig. 2 caption: delete "of" after "relationships.

R: it has been corrected.

P6, L25: "noise" is always singular.

R: it has been corrected.