Top-down Constraints on Regional Nitrous Oxide and Methane Emissions within the U.S. Corn Belt

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Zichong Chen

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Timothy J. Griffis
Dylan B. Millet

12 April 2018

Acknowledgements

First and foremost, I am extremely grateful to my advisor and co-advisor, Dr. Tim Griffis and Dr. Dylan Millet, for the endless support both academically and personally. The experience, big-picture vision, and enthusiasm have been a tremendous treasure to me and my research. I appreciate all the contributions of time, ideas, and funding to make my Ph.D. experience stimulating. The joy and enthusiasm they have for their research was contagious and motivational for me. I could not have imagined having a better advisor and mentor for my Ph.D. study.

Besides my advisor and co-advisor, I would like to thank the rest of my thesis committee: Drs. John Baker and John Zobitz, for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

I am thankful for the helpful discussions with Drs. Xuhui Lee, Rod Venterea, Jeff Wood, Peter Turner, Kelley Wells, Lu Hu, Malte J. Deventer, and Ming Chen. Further, all of the following projects were made possible with the technical and logistical support provided by Matt Erickson. I thank my fellow members in the Griffis group and Millet group- including Xiang Li, Xiaosong Zhao, Ke Xiao, Kendall King, Cody Winker, Cheng Hu, Peter Ganzlin, Meng Xing, Lee Miller, Hariprasad Alwe, Sreelekha Chaliyakunnel, Xin Chen and Xueying Yu for the collaboration, support and helpful discussions.
Last but not the least, I would like to thank my parents and younger sister for supporting me spiritually throughout writing this thesis and my life in general.

Abstract

Nitrous oxide (N_2O) and methane (CH_4) are important long-lived greenhouse gases. Their concentrations are increasing ~0.85 ppb and ~6 ppb per year, respectively. Agricultural sources account for about 60% and 42% of the total global anthropogenic emissions, respectively. Nitrous oxide emissions within the US Corn Belt have been previously estimated to be 200-900% larger than predictions from emission inventories, implying that one or more source categories in bottom-up approaches are underestimated. Furthermore, the CH_4 budget of this region, and its source partitioning, are poorly constrained, with uncertainties ranging from ~50% to 100%. To address these concerns, my dissertation research sought to: 1) quantify the relative importance of direct versus indirect N_2O emissions within the region and to explore its seasonality and inter-annual variability; 2) assess retrospective and future N_2O emissions at fine spatiotemporal scales to help identify hotspots, hot moments, and mitigation priorities for the US Corn Belt; and 3) partition CH_4 emissions into natural (e.g. wetlands) and anthropogenic (e.g. livestock, waste, and natural gas) sources and explore its temporal variability.

Here I interpreted hourly N_2O concentrations measured during 2010 and 2011 at a tall tower located in the US Corn Belt using a time-inverted transport model and a scale factor Bayesian inverse (SFBI) method to simultaneously constrain direct and indirect agricultural emissions. Data and analyses suggested an indirect emission factor for runoff and leaching ranging from 0.014–0.035, which represents an upward adjustment of 1.9–4.6 times relative to bottom-up inventories such as the Intergovernmental Panel on Climate Change (IPCC).

To better understand the temporal and spatial dynamics of N_2O emissions within the region, I advanced a measurement-modeling framework for identifying potential regional hotspots. I used an Eulerian approach based on the Weather Research and Forecasting Chemistry (WRF-Chem) model and implemented the land surface scheme CLM45-BGC-CROP. The model results were strongly correlated ($r=0.69 - 0.82$, $p<0.01$) with multiple tall tower observations across the Corn Belt. Model simulations over a 40-year time period, under the RCP8.5 business-as-usual scenario, indicated that indirect N_2O emissions were associated with a relatively large increasing trend of 0.36 nmol m^{-2} s^{-1} per year, highlighting that southern Minnesota and the lower Midwest regions will likely represent regional hotspots and mitigation priorities in the future.

To estimate the total CH_4 budget and to partition emissions into natural and anthropogenic sources, I combined tall tower (185 m) aerodynamic flux measurements and the SFBI technique for the period June 2016 to September 2017. Flux partitioning revealed that the regional natural, anthropogenic, and total CH_4 emissions were 4.0 ± 1.2, 7.8 ± 1.6, and 11.8 ± 2.0 Tg CH_4 yr^{-1}. The total regional CH_4 emissions for the Corn Belt accounted for ~23% of the overall US budget.

List of Tables

List of Figures

Figure 4.2. The *a priori* annual average emissions from (a) livestock; (b) natural gas +oil; (c) coal mining; (d) waste; (e) natural wetlands and (f) total (unit is $\log_{10}$(nmol m^{-2} s^{-1})). Anthropogenic source categories including livestock, oil/gas, coal mining, and waste were from *Maasakkers et al*. [2016], and the natural wetland emissions were derived from *Bloom et al*. [2017]. The KCMP tall tower is indicated by the black crosses in each panel.

Figure 4.3. Source footprint functions (units: $\log_{10}$(ppm μmol^{-1} m^2 s)) for measurements at the KCMP tall tower (indicated by crosses). (a) December, 2016; (b) March. 2017; (c) August, 2017. (d) Normalized cumulative contribution as a function of source footprint scale (ppm μmol^{-1} m^2 s) and the corresponding distance from the tall tower (km).

Figure 4.4. Hourly mean CH_4 mixing ratios measured at 3 m (panel a) and 185 m (panel b) as a function of wind direction. The color bar represents the friction velocity value (m s^{-1}) Panel c shows the mean methane mixing ratios (+/- 1 standard deviation) as a function of wind direction binned using 15 degree intervals. Panel d shows the relation between CH_4 mixing ratios measured at 3 m versus 185 m. The color bar indicates the friction velocity value. Each data point represents an hourly average value. Note that the 1:1 line is obscured by the best-fit linear regression line.

Figure 4.5. Median monthly vertical CH_4 mixing ratio gradient (Panel a), and median monthly flux-gradient estimate (Panel b). Error bars represent the standard error. Note that August, 2016 is not reported due to a large fraction of missing hourly observations.

Figure 4.6 a). Monthly linear regression between the observed and simulated CH_4 mixing ratios for *a priori* and *a posteriori* simulations between June 2016 and September 2017

Table of Contents

Chapter 1
Introduction

Two important greenhouse gases (GHGs) contributing to global warming include nitrous oxide (N_2O) and methane (CH_4) [*Tian et al*., 2012; *Kirschke et al*., 2013; *Jacob et al*., 2016; *Griffis et al*., 2017]. Nitrous oxide has a global warming potential 298 times that of CO_2 on a 100-yr time horizon [*Myhre et al*., 2013] and is the dominant stratospheric ozone depleting source [*Ravishankara et al*., 2009]. Atmospheric N_2O has increased by 19% compared to the preindustrial level [*Myhre et al*., 2013, Figure 1.1]. This increase has largely been attributed to agricultural practices, which account for ~60% of the anthropogenic emission budget [*Davidson*, 2009; *Davidson and Kanter*, 2014].

Methane (CH_4) has a global warming potential 28 times CO_2 on a 100-yr time horizon. The global CH_4 concentration has increased from 720 ppb to 1800 ppb since preindustrial times [*Hartmann et al*., 2013]. The dominant sources of CH_4 include natural wetlands, livestock, oil/gas systems, landfills, coal mines, wastewater management, and rice cultivation [*Wecht et al*., 2014; *Jacob et al*., 2016]. The present-day global emission of CH_4 is well constrained and estimated to be 550 ± 60 Tg yr^{-1}. However, the magnitude and seasonality of CH_4 emissions from different source categories and source regions are highly uncertain [*Hartmann et al*., 2013; *Kirschke et al*., 2013].

To meet the unprecedented need for food, fiber and biofuels, the global demand for synthetic nitrogen (N) has been forecast to reach 119 Tg N yr^{-1} in the year 2018 [FAO, 2018], and about 4% of them is emitted as N_2O [*Crutzen et al*., 2008; *Smith et al*., 2012]. In recent decades bottom-up and top-down approaches have been widely used to estimate regional to global scale N_2O emissions. Bottom-up approaches rely on activity data and emission factors (EFs) to estimate the emissions [*De Klein et al*., 2006], while the top-down estimates are derived from atmospheric observations [*Smith et al*., 2012; *Jacob et*

al., 2016]. At global scales, top-down and bottom-up emission estimates have shown reasonable agreement [*Crutzen et al*., 2008]. However, a large discrepancy in regional N_2O emission estimates reveals the existence of compensating biases over different regions of the world [*Griffis et al*., 2013]. For example, the US Corn Belt, one of the most intensively managed landscapes, produces 40% of the world's grain [*NRCS*, 2012], and consumes 6.2 ± 0.9 Tg N yr-1 synthetic fertilizer with a remarkable growth rate of 0.08 Tg N yr^{-1} in recent decades [*Griffis et al*., 2017]. Within this region, top-down emission estimates [*Kort et al*., 2008; *Miller et al*., 2012; *Griffis et al*., 2013; *Zhang et al*., 2014] have shown to be 2- to 9-fold larger than the bottom-up estimates, implying significant underestimate in the EFs for the source(s). However, a severe lack of observations and high spatial variability in N_2O emissions have posed a major challenge in constraining the sources and in reconciling the difference between top-down and bottom-up emission estimates.

A number of studies [*Stehfest et al*., 2006; *Shchebak et al.,* 2014] have indicated that the EF for direct emissions (i.e., directly from farmlands) is well constrained. However, there are large uncertainties, up to 50-fold, in the indirect emissions associated with leaching and runoff, attributed to a severe lack of observations at the appropriate spatial and temporal scales [*Outram and Hiscock*, 2012; *Turner et al*., 2015]. Top-down and bottom-up estimates are challenging to reconcile, largely due to the fact that the former often only provides aggregate emission totals and the latter relies on tallying up detailed EFs for individual sources.

Within the US Corn Belt, livestock production is also significant with a population of 728 million animals, including cattle, swine, sheep, layers, broilers, turkey and horses [*USDA*

NASS, 2014]. Wetlands are largely distributed in central and northern Minnesota as well as Wisconsin. Thus, this region is an important source of both natural and anthropogenic CH_4 emissions. However, the CH_4 budget of this region, and its source partitioning, are poorly constrained, with uncertainties ranging from ~50% to 100% [*National Research Council*, 2010; *Dlugokencky et al*., 2011].

To advance the scientific understanding of N_2O and CH_4 emissions from the US Corn Belt the goals of this research are three-fold: 1) quantify the relative importance of direct versus indirect N_2O emissions within the region and to explore its seasonality and inter-annual variability; 2) assess retrospective and future N_2O emissions at fine spatiotemporal scales to help identify regional hotspots, hot moments, and mitigation priorities for the US Corn Belt; and 3) partition CH_4 emissions into natural (e.g. wetlands) and anthropogenic (e.g. livestock, waste, and natural gas) sources and explore its temporal variability.

In chapter 2, we employed a novel inverse modeling technique paired with high-precision tall tower observation data, to simultaneously constrain the direct and indirect N_2O sources within the Corn Belt. This chapter explores the relative importance of direct and indirect sources and provides strong evidence that the discrepancy between top-down and bottom-up emission estimates can be reconciled via upward adjustment of the indirect emission factors.

Findings from Chapter 2 have provided further motivation determining the accurate quantification of N_2O emission at higher spatial and temporal resolution. This is essential in facilitating effective mitigation efforts with knowledge of high emissions in terms of locations (hot spots) and time (hot moments). Chapter 3 develops an Eulerian modeling

framework to extract the most information possible from multiple tall tower concentration observations within the Corn Belt. Efforts in this chapter focus on the retrospective estimate of fine-scale spatiotemporal N_2O variability and also predict the future trends *via* incorporating a process-based land surface scheme into the Eulerian modeling framework. This chapter includes an assessment of mitigation priorities in guiding effective GHG reductions for the region.

Furthermore, it's of critical importance to delve into the relative importance of natural and anthropogenic CH_4 sources within this region. In Chapter 4, tall tower (185 m) aerodynamic flux measurements and a scale factor Bayesian inversion were used to constrain the monthly budget and to partition the CH_4 emissions into natural (e.g. wetlands) and anthropogenic (e.g. livestock, waste, and natural gas) sources within this region. This chapter concludes with an evaluation of the contribution from the Corn Belt to the overall US emission budgets.

In Chapter 5, the overall conclusions of this dissertation research are presented along with an outline of future research directions.

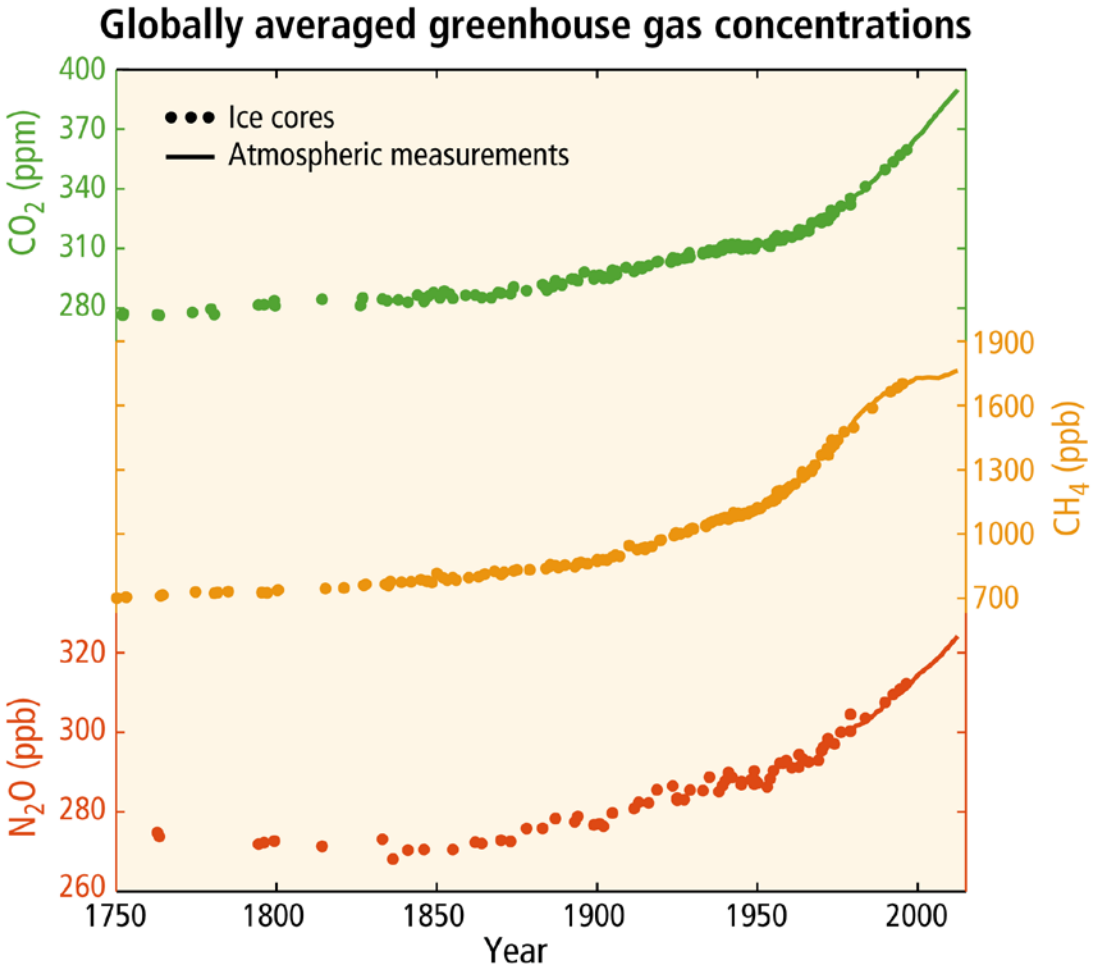


Figure 1.1. Observed changes in atmospheric greenhouse gas concentrations. Atmospheric concentrations of carbon dioxide (CO_2, green), methane (CH_4, orange), and nitrous oxide (N_2O, red), *IPCC Fifth Assessment Report*, 2013.

Chapter 2
Partitioning N_2O Emissions within the US Corn Belt using an Inverse Modeling Approach

Chen, Z. et al. (2016), Partitioning N_2O emissions within the US Corn Belt using an inverse modeling approach, *Global Biogeochem. Cycles*, *30*, 1192–1205, doi:10.1002/2015GB005313.

Synopsis

Nitrous oxide (N_2O) emissions within the US Corn Belt have been previously estimated to be 200-900% larger than predictions from emission inventories, implying that one or more source categories in bottom-up approaches are underestimated. Here we interpret hourly N_2O concentrations measured during 2010 and 2011 at a tall tower using a time-inverted transport model and a scale factor Bayesian inverse method to simultaneously constrain direct and indirect agricultural emissions. The optimization revealed that both agricultural source categories were underestimated by the Intergovernmental Panel on Climate Change (IPCC) inventory approach. However, the magnitude of the discrepancies differed substantially, ranging from 42–58% and 200–525% for direct and indirect components, respectively. Optimized agricultural N_2O budgets for the Corn Belt were 319±184 (total), 188±66 (direct), and 131±118 Gg-N yr^{-1} (indirect) in 2010, versus 471±326, 198±80, and 273±246 Gg-N yr^{-1} in 2011. We aqstttribute the inter-annual differences to varying moisture conditions, with increased precipitation in 2011 amplifying emissions. We found that indirect emissions represented 41–58% of the total agricultural budget, a considerably larger portion than the 25–30% predicted in bottom-up inventories, further highlighting the need for improved constraints on this source category. These findings further support the hypothesis that indirect emissions are presently underestimated in bottom-up inventories. Based on our results, we suggest an indirect emission factor for runoff and leaching ranging from 0.014–0.035 for the Corn Belt, which represents an upward adjustment of 1.9–4.6 times relative to the IPCC and is in agreement with recent bottom-up field studies.

Introduction

As one of the most intensively managed agricultural areas in the world, the US Corn Belt plays an important role in meeting global demands for corn, soybean and biofuel production. To sustain this production, ~5.0 Tg of nitrogen (N) is applied as synthetic fertilizer to fields in the Corn Belt each year [*Millar et al*., 2010; *Griffis et al*., 2013]. Large inputs of synthetic N fertilizer and biological N fixation associated with legume cropping are directly related to increasing concentrations of atmospheric nitrous oxide (N_2O), a greenhouse gas with a 100-yr average global warming potential (GWP) 298 times larger than an equal mass of CO_2 [*Myhre et al.,* 2013].

Constraints on regional to continental scale N_2O budgets are needed to develop baseline emission estimates that can be used as a reference in order to inform and assess policy and mitigation strategies. Investigations using top-down or bottom-up methodologies have been used to estimate N_2O emissions [*Kort et al*., 2008; *Miller et al*., 2012; *Griffis et al*., 2013].Using a top-down constraint and short-term air flask observations, *Kort et al.,* [2008] estimated that emissions over North America were up to 3-fold larger than from bottom-up inventories. With regional tall tower measurements and daily flask data, *Miller et al.,* [2012] estimated that regional budgets were at least 2-fold larger than bottom-up inventories when performing geostatistical and Bayesian inverse analyses. Using two-years of hourly tall tower observations, *Griffis et al.,* [2013] applied atmospheric boundary layer approaches to estimate the N_2O budget within the US Corn Belt. Their estimates agreed well with other top-down estimates based on inverse analyses [*Kort et*

al., 2008; *Miller et al*., 2012], and were 2- to 9-fold greater than bottom-up approaches. Based on a global analysis of N entering agricultural systems, *Crutzen et al*. [2008] and *Smith et al*. [2012] concluded that N_2O emissions would need to be 2- to 3.4-fold larger than current bottom-up inventories in order to match the observed changes in atmospheric N_2O concentration over the period 1860 to 2000. The large difference between top-down and bottom-up N_2O budgets at regional to continental scales implies that the emission inventories are not adequately accounting for N_2O sources. However, there is a good agreement between top-down and bottom-up methodologies at the global scale [*Del Grosso et al*., 2008; *Thompson et al.* 2014].

Agricultural N_2O emissions arise from direct emissions from fertilized soils, and through two indirect pathways: (i) from the deposition of NH_3 and NO_x volatilized from synthetic fertilizer and manure; and (ii) from the leaching and runoff of fertilizer and manure N, mainly as nitrate (NO_3^-). Previous studies from various agricultural fields within Minnesota indicate that direct N_2O emission is about 1.3% of applied synthetic N and in excellent agreement with the IPCC direct emission factor [*Fassbinder et al*., 2013; *Griffis et al*., 2013]. A recent meta-analysis has shown that this emission factor is well constrained but increases nonlinearly as N addition exceeds crop demand [*Shcherbak et al*., 2014]. One of the largest sources of uncertainty in the bottom-up estimates is related to the indirect sources from agricultural ecosystems. For example, indirect emissions from leaching and runoff have an emission factor range of 0.0005 to 0.025 (a 50-fold range) due to a lack of observations at the appropriate spatial and temporal scales [*Nevison et al*., 2000; *Outram and Hiscock*, 2012; *Turner et al*., 2015].

Agricultural drainage systems (i.e., tile lines and ditches, and surface waters), necessary to support crop production in the US Corn Belt, have shown the potential for large episodic emissions. The indirect emissions from these streams remain poorly quantified at the appropriate spatial scales and are thought to contribute substantially to regional emissions [*Beaulieu et al*., 2011, *Outram and Hiscock*, 2012, *Turner et al*., 2015]. Approximately 2.3 million hectares of land have been drained for agriculture in Minnesota. These fine-scale drainage features represent a key hydrological conduit for the transport of N [*Alexander et al*., 2000]. *Turner et al*., [2015] analyzed N_2O emissions as a function of Strahler stream order in southern Minnesota, and found that headwater streams (i.e., stream order of 1) were the strongest sources, emitting 60% of the riverine budget. Their scaling suggested that by accounting for emissions from zero-order streams that the regional N_2O budget would more than double. Further, these indirect emissions may become more important because drainage and stream flow have both increased within the region over the last 50 years [*Baker et al*., 2012].

In this paper we employed an inverse analysis to simultaneously constrain the direct and indirect N_2O sources within the US Corn Belt. Tall tower measurements were used to provide high-resolution information with a large-scale footprint that is highly sensitive to emission from the US Corn Belt. A Bayesian inverse analysis was adopted to identify and constrain the N_2O budget, sources, and sensitivity to environmental drivers. Here we address the following questions:

(1) Can Bayesian inverse modelling be used to objectively constrain the direct and indirect sources contributing to N_2O emissions in the US Corn Belt?

(2) To what sources are observed tall tower N_2O concentrations most sensitive?

(3) What are the seasonal patterns of the direct and indirect N_2O emissions?

(4) How do the environmental factors (e.g., precipitation, air temperature, soil moisture, and surface runoff) influence N_2O emissions from direct versus indirect emissions?

Methods

Approach Overview

The inverse approach employed in this study is outlined in Figure 2.1. The WRF3.5 model [*Zhao et al*., 2009; *Nehrkorn et al*., 2010; *Jeong et al*., 2012] provides the conditions of wind, atmospheric stability, and planetary boundary layer (PBL) height to drive the STILT transport model [*Gerbig et al*., 2003; *Lin et al*., 2003; *Lin et al*., 2004], which was used to estimate the tall tower concentration source footprint. The source footprint was multiplied by *a priori* emission estimates derived from a variety of sources, which is added to the background concentrations, to obtain an initial guess of N_2O concentrations at the tall tower receptor. With these initial estimates and the tall tower concentration observations, a Bayesian inverse model was used to optimize the *a priori* emissions along with the relative contributions of direct and indirect sources.

Study Domain

Our study domain is focused on the US Corn Belt and includes the major corn/soybean production systems in Minnesota, Illinois, Indiana, Iowa, Missouri, Ohio, South Dakota, and Wisconsin. The N_2O concentration measurements were made at the University of Minnesota tall tower Trace Gas Observatory (KCMP tall tower, 44.689°N, 93.073°W; 244 m height) over a two-year period (2010 and 2011). The KCMP tower is located approximately 25 km south of Minneapolis-St. Paul, MN. In the vicinity of the tower, agriculture represents approximately 46% of the land use and is representative of the larger US Corn Belt region [*Zhang et al*., 2014].

Tall Tower N_2O Measurements

During 2010-2011, air samples were analyzed at the tall tower at heights of 32, 56, 100, and 185 m. Air was pulled continuously through each inlet to the base of the tower and then subsampled at 3 SLPM using a custom designed manifold. N_2O mixing ratios were measured using a tunable diode laser (TGA100A, Campbell Scientific Inc., Logan, Utah, USA) that was housed in a temperature-controlled building. Hourly calibrations were performed using a zero and span gas. The span gas was traceable to the National Oceanic and Atmospheric Administration-Earth System Research Laboratory (NOAA-ESRL). The hourly TDL calibration precision, estimated using the Allan variance technique [*Werle et al.*, 1993], was estimated to be 0.5 ppb [*Griffis et al*., 2013]. Following calibration, hourly average N_2O concentrations were computed. Previous work has employed N_2O measurements from the KCMP tower to estimate the total regional N_2O

source [*Griffis et al*., 2013], and as part of a global analysis to assess our current ability to constrain N_2O sources worldwide [*Wells et al*., 2015]. Further details regarding the sampling system and calibration can be found in *Griffis et al.* [2010, 2013].

Source Footprint Simulations

The source footprint function delineates areas that influence the tall tower observations [*Kim et al*., 2013]. The STILT model computes the upstream influence on a measurement site by releasing a suite of particles from the receptor (tall tower air intake at 100 m) and following their trajectory backwards in time. The time- and volume-integrated footprint function is quantified by tallying the total amount of time each particle spends in a volume element over a time step, normalized by the total amount of particles. Molar mass and density of air and the mixing height are accounted for in the function in order to scale all the particles to represent the entire vertical profile [*Lin et al*., 2003*; Gerbig et al*., 2003; *Lin and Gerbig*, 2005]. Multiplying the source footprint by the *a priori* emissions and summing over all locations provides an estimate of the tall tower mixing ratios.

We simulated wind fields using the WRF3.5, and interpolated them to the explicit location of each particle. The WRF3.5 was set up using 3 nests at 27 km, 9 km, and 3 km grid spacing. The outermost domain covers North America and each domain was centered on the tall tower. The simulations used WRF Single-Moment 3-class (WSM3) simple ice microphysics scheme [*Hong et al.*, 2004], Kain-Fritsch convective scheme [*Kain*, 2004] and the YSU scheme coupled to the Noah land surface model for the

planetary boundary layer (PBL) processes [*Hong et al*., 2006]. Initial and boundary conditions were provided by the National Center for Environmental Prediction (NCEP) Final Analysis (FNL; 1° × 1°), with a 6-h interval.

In this study, we released 500 particles per hour from the KCMP tall tower at a height of 100 m for the years of 2010 and 2011, and transported them backwards for 7 days to ensure that the trajectories adequately represented source contributions from within the US Corn Belt. Furthermore, we used observations from the NOAA Carbon Cycle and Greenhouse Gases (CCGG) program [*Dlugokencky et al*., 1994] near the outer edge of the source footprint, to represent the background mixing ratios. These observations are from discrete air samples collected approximately weekly in flasks at 77 sites, and are zonally and monthly averaged at 4° latitudinal resolution [*Wells et al*., 2015].

A Priori Emissions

We used EDGAR42 (Emission Database for Global Atmospheric Research, version 4.2, 2011, *http://edgar.jrc.ec.europa.eu*), to provide *a priori* annual N_2O emissions. Different sources in the inventory were tracked as separate tagged tracers in our simulation, with the sum of these equal to the total ambient N_2O mixing ratio. EDGAR42 represents the source emission at a spatial resolution of 0.1° × 0.1°. Here, we reorganized the anthropogenic sources of N_2O into 5 categories and included two natural source categories to represent emissions including:

(1) Direct emissions from agricultural soils (*dirA*), including synthetic N fertilizer, manure management, and crop residues;

(2) Indirect emissions from leaching/runoff in agriculture (*indA*), provided by CLM45-BGC (Community Land Model coupled to Biogeochemistry) [*Oleson et al*., 2013];

(3) Solid waste and wastewater (*waste*);

(4) Industrial processes (non-combustion) (*industry*);

(5) Fuel combustion and fugitive emissions from fuel (*energy*)

(6) Natural emissions (*natsoil*) from non-agricultural soil provided by EDGAR2; and

(7) Biomass Burning (*BB*) from GFED3 (Global Fire Emissions Database, version 3, 2011, *http://www.globalfiredata.org*).

Initially we used the indirect emissions from agriculture from the leaching/runoff category from EDGAR42. However, since *dirA* and *indA* from EDGAR42 are both calculated using the IPCC EF approach they were found to be highly correlated (r^2=0.99) , and therefore, cannot be considered independent variables in our optimization. To address this concern we used *indA* obtained from the CLM4.5-BGC. In CLM45-BGC, the Century N model [*Parton et al*., 1996, 2001; *Del Grosso et al*., 2000] was used to simulate the soil NO_3^- pool.

Since the CLM4.5-BGC model does not provide *indA* directly, we estimated it by multiplying the soil NO_3 pool losses to leaching and runoff by the IPCC EF_5 default value of 0.0075 [*de Klein et al*., 2006]. The CLM4.5-BGC model was run for 2010 with and without crops. The crop-on model simulated N leaching and runoff from both

agricultural and natural soils, while the crop-off model only simulated N leaching and runoff from natural soils. By taking the difference of these two simulations we estimated the N leaching and runoff attributed to agricultural soils. Next, we applied the EF_5 to obtain the indirect emissions from agriculture. Here we assume that N leaching and runoff from the CLM45-BGC is correct.

An important point is that indirect N_2O from volatilization and re-deposition is not explicitly represented in our *a priori* emissions. The IPCC methodology [*De Klein et al.,* 2006] does, however, account for agricultural N_2O emissions arising from the volatilization and re-deposition of reactive N. In our methodology all indirect emissions are allocated to one source category, *indA*. We attribute our indirect emission estimates to the leaching and runoff source category, because we have observed high fluxes from surface water systems within the region [*Turner et al*., 2015]. However, it is possible that these aquatic emissions arise from amplification of the N cycle owing to both leaching/runoff and deposition. Therefore, our emission factor may be overestimated and should be taken as a conservative upper bound.

Bayesian Inversion Methods

We defined the mixing ratio observed at the tall tower, *Y*, and subtracted the background value. Next, a scale factor Bayesian inverse (SFBI) method was applied for each month (April to October, in 2010 and 2011, respectively). As described in *Gerbig et al*., [2003], *Zhao et al*., [2009], and *Jeong et al*., [2012], *y* can be modeled as,

$$\mathbf{y} = \mathbf{K\Gamma} + \boldsymbol{\varepsilon} \qquad (2.1)$$

where y is the observed minus background mixing ratios, Γ is the scaling factors for different source types, K is the Jacobian matrix, representing the sensitivity of the observation variables to the specific source types, and ε is the system error, which consists of instrumental and model errors. In our case, the columns of K correspond to the mixing ratios for each of the source types being optimized, and Γ consists of the *a posteriori* scale factors for the 7 source types.

Applying Bayes' theorem, along with a normal distribution assumption, the *maximum a posteriori* (MAP) solution of Γ *is* to minimize the cost function $J(\Gamma)$:

$$\mathbf{2J(\Gamma) = (y - K\Gamma)^T S_\varepsilon^{-1}(y - K\Gamma) + (\Gamma - \Gamma_a)^T S_a^{-1}(\Gamma - \Gamma_a)} \qquad (2.2)$$

Where S_ε and S_a are the observational and *a priori* error covariance matrices, and each element of $\Gamma_a = 1$. The solution to $\nabla\Gamma J(\Gamma) = 0$ is then given by:

$$\mathbf{\Gamma_{post} = (K^T S_\varepsilon^{-1} K + S_a^{-1})^{-1}(K^T S_\varepsilon^{-1} y + S_a^{-1}\Gamma_a)} \qquad (2.3)$$

Observational errors consist of measurement and modelling errors. The measurement error ascribed to the TDL was based on the estimate of calibration precision, which was 0.5 ppb.

The measurement uncertainty from observations of background mixing ratios is 0.4 ppb, based on recommendations from the data providers [*Wells et al*., 2015]. Therefore, we assign an uncertainty of 0.4 ppb for background mixing ratios.

Since STILT releases a finite number of particles (500 particles in the simulations presented here), we follow previous work [*Gerbig et al*., 2003; *Miller et al*., 2008] and assign an uncertainty of 13% for the simulated back trajectories. The uncertainty associated with the simulation of PBL height (i.e., the effect of PBL height on concentrations) was estimated from the monthly mean values of $|H_{obs}-H_{model}|/H_{obs}$. H_{obs} is the mixing height inferred from radiosonde observations [*Matross et al*., 2006; *Miller et al*., 2008; *Kretschmer*, 2012; *Kim et al*., 2013]. H_{model} is the mixing height from the WRF-STILT simulations, using Yonsei University [*Hong et al*., 2006] and Mellor-Yamada-Janjic [*Janjic et al*., 1994] PBL schemes, respectively. The relative uncertainty estimated here is 21%. The modeling error ranges from 4-8 ppb and is much larger than the measurement error.

The uncertainties assigned to the *a priori* emissions shown in Figures 2a and 2b were obtained from the literature [*De Klein et al*., 2006; *Shcherbak et al*., 2014]. A recent meta-analysis found a nonlinear response in the EF for direct emissions from agricultural soils, however, they reported good agreement with the IPCC default value when N-rates were within the range of typical values for the Corn Belt albeit with reduced uncertainty, which is ~66% when the N application rate is ~150 kg ha^{-1} [*Shcherbak et al*., 2014]. The uncertainty in the indirect emissions was set at 401% by propagating assumed

uncertainties of 327% in the default EF_5 and 233% for the amount of N in leaching/runoff [*de Klein et al*., 2006]. Based on the IPCC recommendations, the uncertainties of *industry*, *waste*, *energy*, *natsoil* and *BB* to the *a priori* emissions were 30%, 30%, 30%, 38%, and 30% in the default EFs [*Gómez et al*., 2006]. Therefore, initial relative error estimates of 66%, 401%, 30%, 30%, 30%, 38%, and 30% were applied for *dirA*, *indA*, *industry*, *waste*, *energy*, *natsoil*, and *BB*, respectively to construct the a *priori* error covariance (see supporting information for more details related to the Bayesian inversion).

We conducted two tests to evaluate the sensitivity of the Bayesian inversion to the *a priori* flux distribution and spatial aggregation errors. First, we examined the differences when using a relatively high spatial resolution (*a priori* inventory with a $0.1^\circ \times 0.1^\circ$ resolution) versus a lower spatial resolution (*a priori* inventory with a de-gridded $0.5^\circ \times 0.5^\circ$ resolution) flux inventory. Second, we perturbed the original spatial distribution of the *a priori* flux inventory with a random variable that varies in space, having a normal distribution with a mean of unity, and a standard deviation of 30%. Both tests indicated that the sensitivity of the inversion results to spatial resolution and source distribution were not statistically significant at the 5% significance level.

We acknowledge that our modeling framework has limitations associated with the assumptions of spatial distribution and source aggregation. We are depending on measurements from one point in space, and therefore, cannot solve independently for

emissions from every model grid cell and time and individual source sector. This is a limitation of all similar atmospheric inversion studies and has been acknowledged previously [*Mikaloff Fletcher et al.*, 2004; *Carouge et al.*, 2010; *Bousquet et al.*, 2011].

Sensitivity to True Values

In order to identify the source types that contribute to the tall tower concentrations, we calculated the averaging kernel (AK) to quantify the sensitivity of the retrieved emissions to their true value [*Rodgers*, 2000; *Kim et al*., 2013].

The Bayesian inverse framework constrains the source categories, and provides optimized emissions using a cost function analysis. The AK represents the sensitivity of the MAP solution, $\hat{x}$, to the true state, x, (i.e., the true emissions from a specific source type):

$$\mathbf{A} = \mathbf{1} - \hat{\mathbf{S}}\mathbf{S}_{\mathbf{a}}^{-1} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} \qquad (2.4)$$

where A is the averaging kernel, $\hat{S}$ and S_a are the *a posteriori* and *a priori* error covariance matrices, respectively. The AK is used to gain insights regarding (1) How sensitive are the tall tower measurements to some specific source types; and (2) whether the source types will be resolved from one another. This type of sensitivity analysis allows identification of the source types contributing to the tall tower N_2O concentrations [*Heald et al*., 2004; *Kim et al*., 2013].

N_2O Emission Budgets

Following the optimization, we obtained the total N_2O budget (B_j) for each source type *j* according to:

$$\mathbf{B_j = \sum_{i=1}^{n} (f_j \times S \times t)} \qquad (2.5)$$

where *n* is the number of grid cells within the US Corn Belt. For each grid cell, f_j is the *a posteriori* N_2O emissions (kg N m^{-2} s^{-1}) for source type *j* , where the *a posteriori* emissions are the *a priori* emissions multiplied by the corresponding scaling factor (Γ), *S* is grid cell area (m^2), and *t* is time (s).

ArcGIS (v.10.1, ESRI, Redlands, California, USA) was used to identify grid cells within the Corn Belt and to compute the areas of each grid cell intersected by the Corn Belt. The intersection function was applied for the 9 States (South Dakota, Nebraska, Kansas, Missouri, Minnesota, Iowa, Illinois, Ohio, and Indiana), to obtain all the grid cell areas over the Corn Belt.

Results and Discussion

Source Footprint of the Tall Tower

The tall tower concentration source footprints (Figures 2c, 3) indicate that the measurements were influenced by sources from within the US Corn Belt. Figure 2.3 shows the season-averaged footprint function at the KCMP tall tower for 2010 as derived

from the WRF_STILT model. The source footprint at the tall tower reached north to Canada and extended south to the Gulf of Mexico. *Kim et al.* [2013] and *Hu et al.* [2015] also concluded that the KCMP tall tower measurements (at the 185 m level) were representative of surface influence at the continental scale.

Areas where the footprint strength was greater than 1e-4 ppm $\mu mol^{-1}\ m^2\ s^{-1}$, representing the dominant surface influence to the tall tower observations, were defined as intense footprint zones. These intense footprint zones encompassed the Corn Belt and extended north to southern Canada. Further, these analyses show that natural sources outside of the Corn Belt can also have an important influence on the tall tower observations. Figure 2.2d shows the land use categories of our study domain [US Geological Survey, 2006, *http://www.usgs.gov*]. Statistical analysis of the land use within the intense footprint zone indicates that agricultural soils, natural soils, urban areas, and water bodies accounted for 81.6%, 12.0%, 0.5%, and 1.4%, respectively. Based on source footprint analyses, the tall tower observations appeared to provide adequate representation of the emissions and transport of N_2O related to the US Corn Belt.

A Priori and A Posteriori Emissions

In the first Bayesian inversion, 7 source types were included in the optimization. The AK obtained from this inversion (Table 2.1) revealed extremely weak sensitivities for the industry, energy, waste, and biomass burning sources, indicating a very limited

contribution from these source types to the tall tower observations. These source types, therefore, were not resolved by the tall tower observations. However, the sensitivity is strong for the direct, and indirect agricultural sources and the natural sources. This is supported by the statistical analyses of the land use within the source footprint described above as well as previous investigations [*Miller et al*., 2012; *Saikawa et al*., 2013; *Griffis et al*., 2013]. The source categories to which measurements were insensitive were generally a result of extremely low emissions (e.g., biomass burning) or very limited area (e.g., industry). Agricultural soils and natural soils are key sources from within and outside of the intense footprint zones, respectively. Based on the low sensitivities as indicated by the AK and the land use analyses, the industry, energy, waste and biomass burning source categories were eliminated from further consideration, and a second Bayesian inversion was performed where we included only the direct and indirect agricultural and natural soil source categories. The annual average *a priori* emissions for this second inversion were 0.12 (direct), 0.04 (indirect), 0.04 (natural soils) nmol m^{-2} s^{-1}, respectively, which are represented spatially in Figure 2.2a and 2.2b. This second inversion yielded the final optimized *a posteriori* emission estimates.

To probe the degree that the *a priori* error construction influenced the optimized emissions, we performed a wide range of sensitivity studies by varying the *a priori* errors of 46-66% (direct), 300-400% (indirect) in the in the Bayesian inversion (Table 2.2). The *a posteriori* errors were significantly reduced in both direct and indirect emissions in the cost function analysis, indicating a robust constraint of the Bayesian inverse approach. Our best estimate suggests that annual mean (± the *a posteriori* error) emissions in total,

direct, and indirect agricultural sources were 0.30 ± 0.16, 0.18 ± 0.05, 0.12 ± 0.11 nmol m^{-2} s^{-1} in 2010, and 0.44 ± 0.27, 0.19 ± 0.05, 0.25 ± 0.22 nmol m^{-2} s^{-1} in 2011, respectively.

The relative uncertainty from natural soils reduced significantly from 38% to ~10%, suggesting a robust constraint. To investigate the impact of natural soils to agricultural sources, in the *a priori* emissions, a sensitivity test was performed by i) increasing the natural soil source by 50%; and ii) decreasing the natural soil source by 50%. The results indicated that natural soils had little impact on both direct and indirect emissions from agriculture in the optimization. The optimized natural soils emissions were ~1.8 times larger than indicated from EDGAR2 within the Corn Belt and is consistent with the findings of *Wells et al*. [2015].

The *a posteriori* direct emissions were 1.5- to 1.6- fold larger than the IPCC bottom-up approach, suggesting an EF ranging from 1.5% to 1.6%, in close agreement with recent studies using a non-linear response model [*Grace et al*., 2013; *Shcherbak et al*., 2014]. The *a posteriori* indirect emissions were 2.4- to 5.1- fold larger than from IPCC EF approach, providing further support [*Outram and Hiscock*, 2012; *Griffis et al*., 2013; *Turner et al*., 2015] that the indirect emissions are substantially underestimated by the IPCC bottom-up approach for this region.

Seasonal and Inter-annual Variations

As shown in Figure 2.4, there were clear seasonal and inter-annual variations in the *a posteriori* emissions. In both years, direct emissions began to increase after the snow melt, reached the peak in July and then decreased through to the end of October. Interestingly, indirect emissions began to increase after the snow melt and reached their maximum in June. We attribute this to a combination of spring-time fertilization throughout much of the Corn Belt and relatively high runoff and tile outflow. Summing over all seasons, our best estimates of the annual agricultural N_2O budgets for the Corn Belt were 319 ± 184 (total), 188 ± 66 (direct), and 131 ± 118 Gg N yr^{-1} (indirect) in 2010, versus 471 ± 326, 198 ± 80, and 273 ± 246 Gg-N yr^{-1} in 2011. The direct and indirect budgets in 2011 increased by 5% and 108% over 2010, respectively. Hydrometeorological factors are explored to identify drivers that result in the large difference in N_2O emission partitioning.

Figure 2.5 shows a comparison of the N_2O budget, and its partitioning between direct and indirect emissions, obtained from a variety of independent methods. The Bayesian inverse approach estimates the N_2O budget in 2010 (316 ± 183 Gg-N yr^{-1}), which is in reasonable agreement with tall tower atmospheric boundary layer approaches (420 ± 50 Gg-N yr^{-1}) [*Griffis et al*., 2013], as well as other top-down methodologies [*Kort et al*., 2008; *Miller et al*., 2012], but much larger than that derived from IPCC and other bottom-up approaches including EDGAR and Global Emission Initiative (GEIA) [*Bouwman et al*., 1995]. The direct N_2O budget from the IPCC EF approach (120 Gg-N yr^{-1}) is within

the uncertainty range of our approach (188 ± 66 Gg-N yr^{-1}). Therefore, for direct emissions, there is no significant difference between IPCC estimate and the Bayesian inversion. For indirect emissions, the IPCC estimate (52.3 Gg-N yr^{-1}) is within the uncertainty range from our Bayesian inversion (131 ± 118 Gg-N yr^{-1}). However, the optimization significantly reduces the uncertainties (Table 2.2) and places a better constraint on the indirect emissions. Furthermore, our findings also support the growing evidence [*Outram* and *Hiscock*, 2012; *Hinshaw* and *Dahlgren*, 2013; *Turner et al*., 2015] that indirect emissions are higher than that estimated using the IPCC EF approach.

The direct emissions from the Bayesian inverse model suggest an EF ranging from 1.5% to 1.6%, 1.5 to 1.6 times larger than the default IPCC EFs. *Shcherbak et al* [2014] showed that the N_2O emission factor increases nonlinearly as fertilizer N exceeds crop demand, and their nonlinear response model suggests a regional EF of 1.7%, assuming an average N application rate of 143 kg-N ha^{-1} for the Corn Belt [*Griffis et al*., 2013]. This is in excellent agreement with our Bayesian inverse analysis, and slightly higher than that suggested by *Griffis et al*. [2013] and *Fassbinder et al*. [2013]. Using an independent approach, our analyses indicate that the both direct and indirect emissions are underestimated in the IPCC inventories, and provide further support that the large disparity between top-down and bottom-up methodologies is likely due to poor constraints on indirect emissions.

It is estimated that on the order of 25–50% of the US Corn Belt has been drained to support agricultural production [*USDA*, 1987]. Although there is growing evidence that

fine-scale drainage features are hot spots of indirect N_2O emissions [*Outram* and *Hiscock*, 2012; *Turner et al*., 2015], quantifying the spatiotemporal variation in these emission remains a challenge. The episodic nature and high spatial variability of these hot spots make them difficult to characterize and give rise to large uncertainties when integrating over space and time to obtain annual emission estimates. The Bayesian inverse model approach used here suggests a regional EF ranging from 0.018 to 0.038 for indirect emissions. In order to directly compare to IPCC EF related to leaching and runoff, we subtracted N_2O emissions from volatilization and redeposition. *Griffis et al* [2013] used wet and dry N deposition (WDD) and the redeposition of local N (RDP) to represent indirect volatilization pathways. Multiplying WDD and RDP (a total of 3.1 Tg N yr^{-1}) by the IPCC indirect EF associated with volatilization (0.01, ranging from 0.002 to 0.05) gives an estimate of the N_2O budget from volatilization and redeposition of approximately 31 Gg N yr^{-1}. The indirect volatilization budget was then subtracted from the indirect emissions estimated from the Bayesian inversion to obtain the indirect emissions from leaching and runoff. The results indicate a regional EF ranging from 0.014 to 0.035 for indirect emissions related to leaching and runoff that is in relatively good agreement with that proposed by *Turner et al.* [2015]. They employed a chamber-based approach to measure riverine N_2O fluxes across stream orders ranging from fine-scale tile drainage to the Mississippi River. Interestingly, their data and analyses suggest a regional EF ranging from 0.007 to 0.03, which is in relatively close agreement with our inverse modeling approach. These independent approaches provide strong evidence that indirect emissions from drainage networks and streams within the US Corn Belt are important sources and that their EF should be revised accordingly.

Environmental Controls on Direct and Indirect N_2O Emissions

Nitrogen availability and environmental factors including air temperature, soil moisture, precipitation, surface runoff are important drivers of N_2O emissions [*Singurindy et al*., 2009; *Zona et al*., 2011; *Luo et al*., 2013]. Hourly weighted means of environmental variables of interest from WRF3.5, were computed for the entire study domain, with the weighting based on the intensity of the source footprint function for each grid cell and smoothed using a 24-hour running mean. For the study period (May to September, 2010 and 2011), the mean air temperature was 15.6 and 16.2 °C respectively. The cumulative precipitation was 496 and 545 mm, the mean soil water content was 0.32 and 0.39 kg/kg, and the cumulative surface runoff was 25.3 and 49.9 mm in 2010 and 2011, respectively. The larger cumulative precipitation in 2011 contributed to the high soil water content and surface runoff, with a soil water content and surface runoff that was 22% and 97% higher in 2011, respectively. Interestingly, both years showed similar direct emissions, while in 2011 the indirect emissions were 108% greater than in 2010.

To investigate how the inter-annual variations were driven with respect to important environmental factors, we applied biweekly Bayesian inverse analyses, from May to September in both 2010 and 2011. The Bayesian inversions permit retrieval of the scaling factors, indicating the amplitude of direct and indirect N_2O emissions, respectively. Combined with the 2-week mean environmental parameters, a linear regression analysis was conducted to investigate the correlations with direct/indirect emissions and the

environmental variables (Table 2.3). We derived a multi-linear regression model (MLR) that incorporated air temperature (T), soil water content (θ), and surface runoff (R) as explanatory variables using MATLAB (Matlab, Version R2015a, Mathworks, Natick, Massachusetts, USA) The MLR ($f(T,\theta,R)$) has the form,

$$\mathbf{\Gamma \approx f(T,\theta,R) = b_0 + b_1 T + b_2 \theta + b_3 R} \quad (2.6)$$

where Γ is the direct/indirect scaling factors.

The statistical analyses indicate that, biweekly variability in direct emissions had a positive linear correlation with air temperature (p=0.024, r^2=0.25), but was not correlated with soil water content (p=0.76) or surface runoff (p=0.97). Mean air temperature within the study period in 2010 and 2011 were similar and supports why direct emissions were similar in both years.

The biweekly variability in indirect emissions show a moderate correlation with air temperature (p=0.08, r^2=0.20), and strong correlation with soil water content (p=0.02, r^2=0.32), and surface runoff (p=0.01, r^2=0.33). The combined effect explains a significant amount (p=0.005, r^2=0.44) of the variability in the indirect emissions.

Short-term heavy precipitation is also an important factor contributing to wetter and anaerobic soil conditions and heavier surface runoff, and therefore larger direct and indirect emissions [*Fassbinder et al*., 2013]. For DOY 121-135 and DOY 136-152, 2011, the two time periods shared similar air temperature (16.8 °C and 17.3 °C, respectively),

but the intense precipitation event during DOY 136-152 increased the cumulative precipitation from 30 mm to 127.7 mm. This event increased the soil water content by 13%, as well as cumulative surface runoff significantly from 3.0 mm to 11.7 mm. During this time period, the Bayesian inverse analysis indicated that direct and indirect emissions increased by 40% and 83%, respectively.

These findings, and related results from *Turner et al*. [2015], demonstrate that the offsite transport of N from farm fields is a dual environmental threat. The role of nitrate in the generation of hypoxia in water bodies is well-known; its role as a precursor to substantial indirect emissions of N_2O has so far been less appreciated. The seriousness of both problems underscores the importance of reducing the leakage of reactive nitrogen in agricultural systems.

Conclusions

Based on high precision N_2O tall tower observations and a novel Bayesian inversion method that partitions N_2O emissions into its direct and indirect components, we have shown that:

1. N_2O emissions from agricultural systems within the Corn Belt were 319 ±184 (total), 188±66 (direct), and 131±118 Gg-N yr^{-1} (indirect) in 2010, versus 471±326, 198±80, and 273±246 Gg-N yr^{-1} in 2011. The direct and indirect emissions in 2011 increased 5% and 108 %, respectively, compared to 2010 due to increased precipitation.

2. Direct and indirect N_2O emissions were out of phase. Indirect emissions reached a maximum in June, while direct emissions reached a maximum in July. This phase shift is attributed to a combination of spring-time fertilization throughout much of the Corn Belt and relatively high runoff that peaks in June.
3. Inverse modeling analyses support that the indirect emission factor associated with runoff and leaching ranges from 0.014 to 0.035 for the US Corn Belt. This represents an upward adjustment of 1.9- to 4.6- fold relative to the Intergovernmental Panel on Climate Change and is in excellent agreement with recent bottom-up field studies.

Acknowledgements

Financial support for this research has been provided by the United States Department of Agriculture, Grant Number: USDA-NIFA 2013-67019-21364, USDA-ARS, NOAA (grant no. NA13OAR4310086), and MnDRIVE. We express our sincere thanks to Tom Nelson and Minnesota Public Radio for the logistical support at the KCMP tall tower. All models and analyses were performed at the Minnesota Supercomputing Institute for Advanced Computational Research. Data are hosted at http://www.biometeorology.umn.edu/research/data-archives.

Table 2.1. Averaging Kernel of the source types

Source Categories	*dirA*	*indA*	*natsoil*	*industry*	*energy*	*waste*	*BB*
AK	0.81	0.90	0.93	0.002	0.002	0.005	0.002

Table 2.2. *A posteriori* Scale Factors and *a posteriori* errors for the Bayesian Inversion with various sensitivity tests

A priori errors		Optimized direct emissions		Optimized indirect emissions	
Direct Emissions	Indirect Emissions	Scaling Factor	*A posteriori* error	Scaling Factor	*A posteriori* error
46%	300%	2.0	28%	4.7	90%
46%	350%	1.9	29%	5.0	102%
46%	400%	1.8	29%	5.3	100%
56%	300%	2.2	33%	4.3	91%
56%	350%	2.1	35%	4.7	102%
56%	400%	2.0	35%	5.0	101%
66%	300%	2.5	37%	4.0	94%
66%	350%	2.2	39%	4.3	100%
66%	400%	2.1	38%	4.6	95%

Table 2.3. Linear regression analyses of environmental factors and the direct/indirect Scaling Factors

	Scaling Factors			
	Direct emissions		Indirect emissions	
	r^2	*p-value*	r^2	*p-value*
Air temperature(*T*)	0.25	0.024	0.2	0.08
Soil Moisture (θ)	0.007	0.76	0.32	0.02
Surface runoff (*R*)	0.0001	0.97	0.33	0.01
Combined effect (*T*, θ,*R*)	0.26	0.02	0.44	0.0047

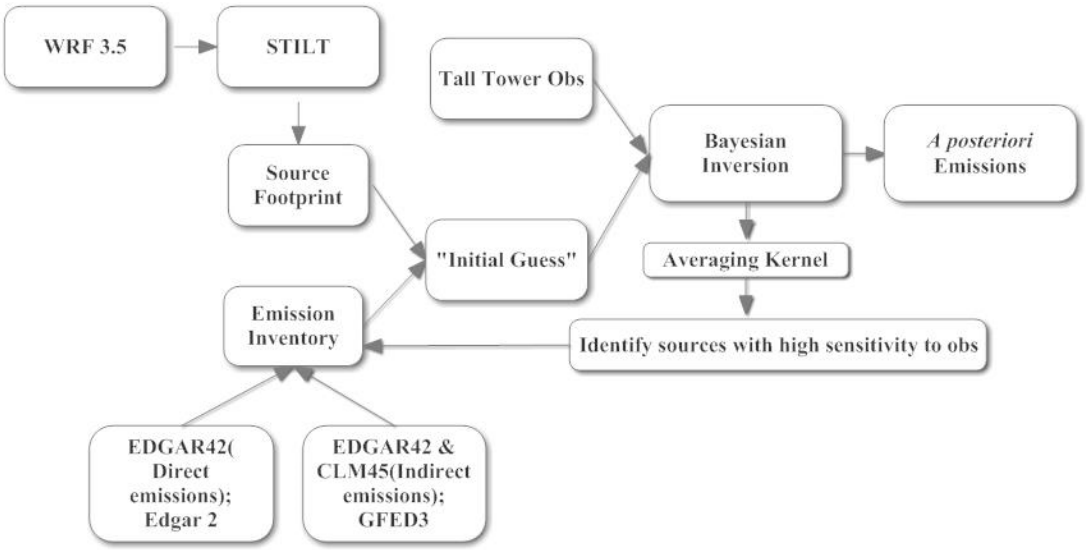


Figure 2.1. Overview of the inverse modeling approach used to estimate direct and indirect agricultural N_2O emissions.

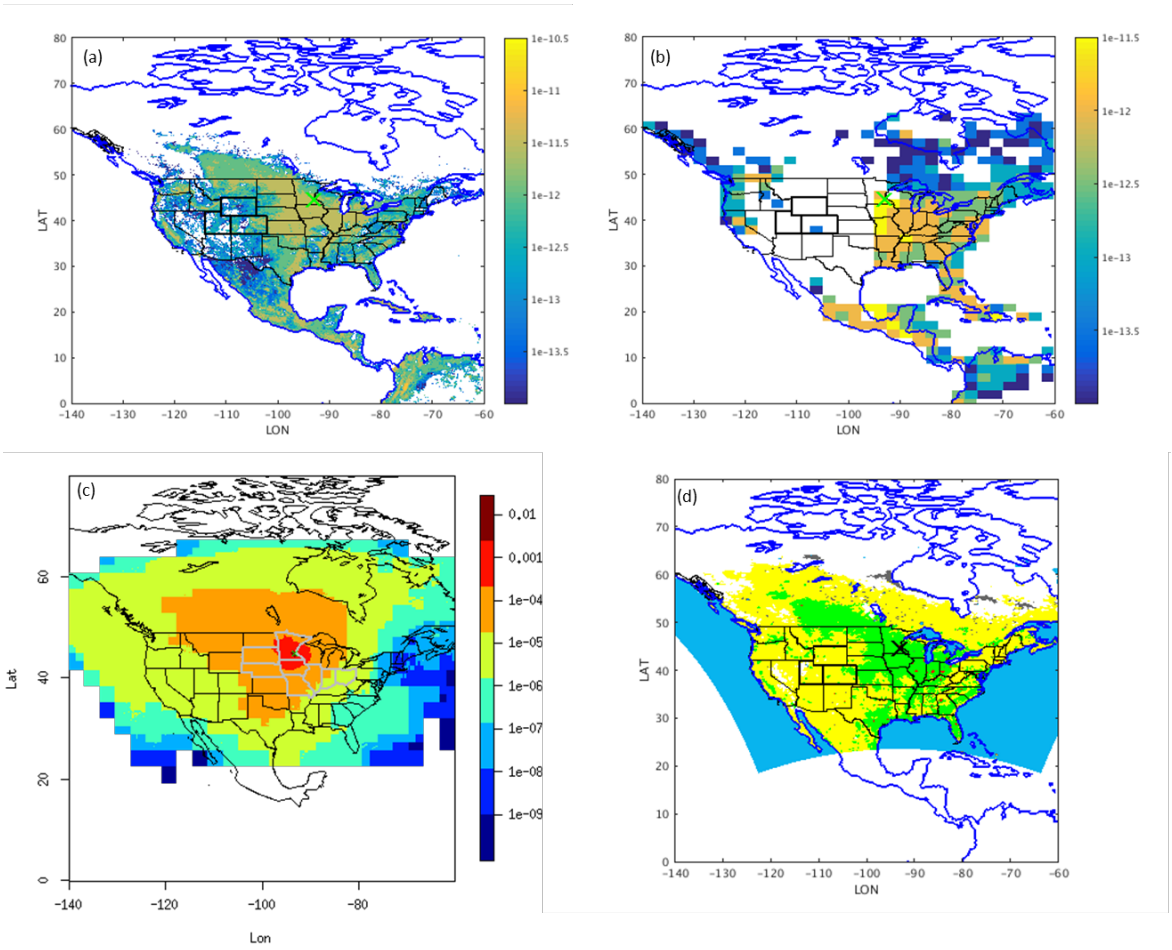


Figure 2.2. a) The *a priori* annual average direct N_2O emissions from agriculture [unit is $\log_{10}$ (kg m^{-2} s^{-1})]; b) *a priori* annual average indirect N_2O emissions from agriculture (the coastlines are denoted by blue lines, and the US States are denoted by black lines); c) annual average source footprint of the tall tower [units are $\log_{10}$ (ppm μmol^{-1} m^2 s)], the US Corn Belt is denoted by gray lines; d) land use distribution from USGS (red, green, yellow, gray, light blue colors denote the urban, agricultural soils, natural soils, wetlands, and water bodies, respectively).

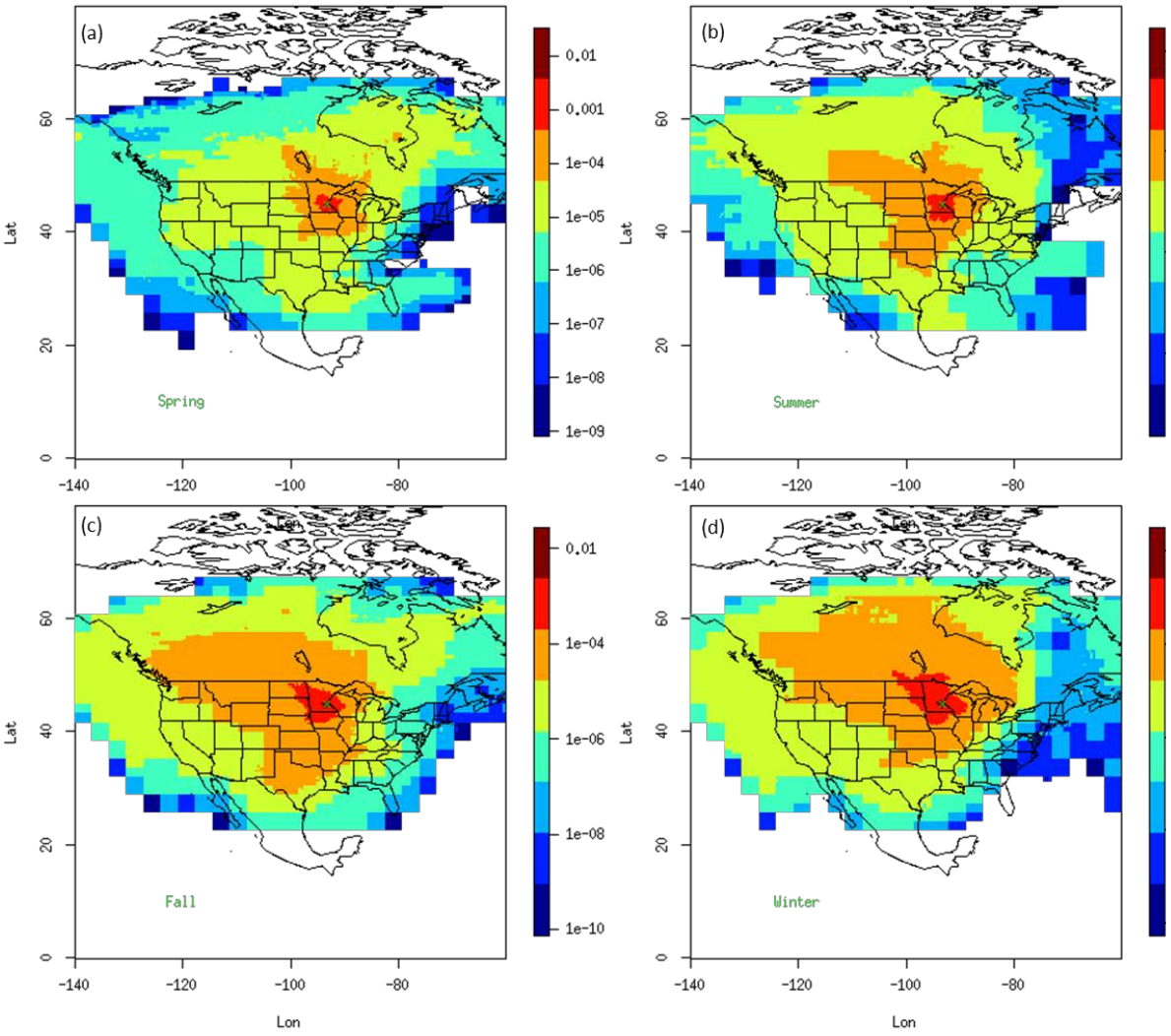


Figure 2.3. Average seasonal source footprints for measurements at the KCMP tall tower (indicated by symbol '×'), 2010 [units: $\log_{10}$(ppm μmol^{-1} m^2 s)] for a). Spring: Mar., April, May; b). Summer: June, July, Aug.; c). Fall: Sep., Oct., Nov.; d). Winter: Dec., Jan., Feb.

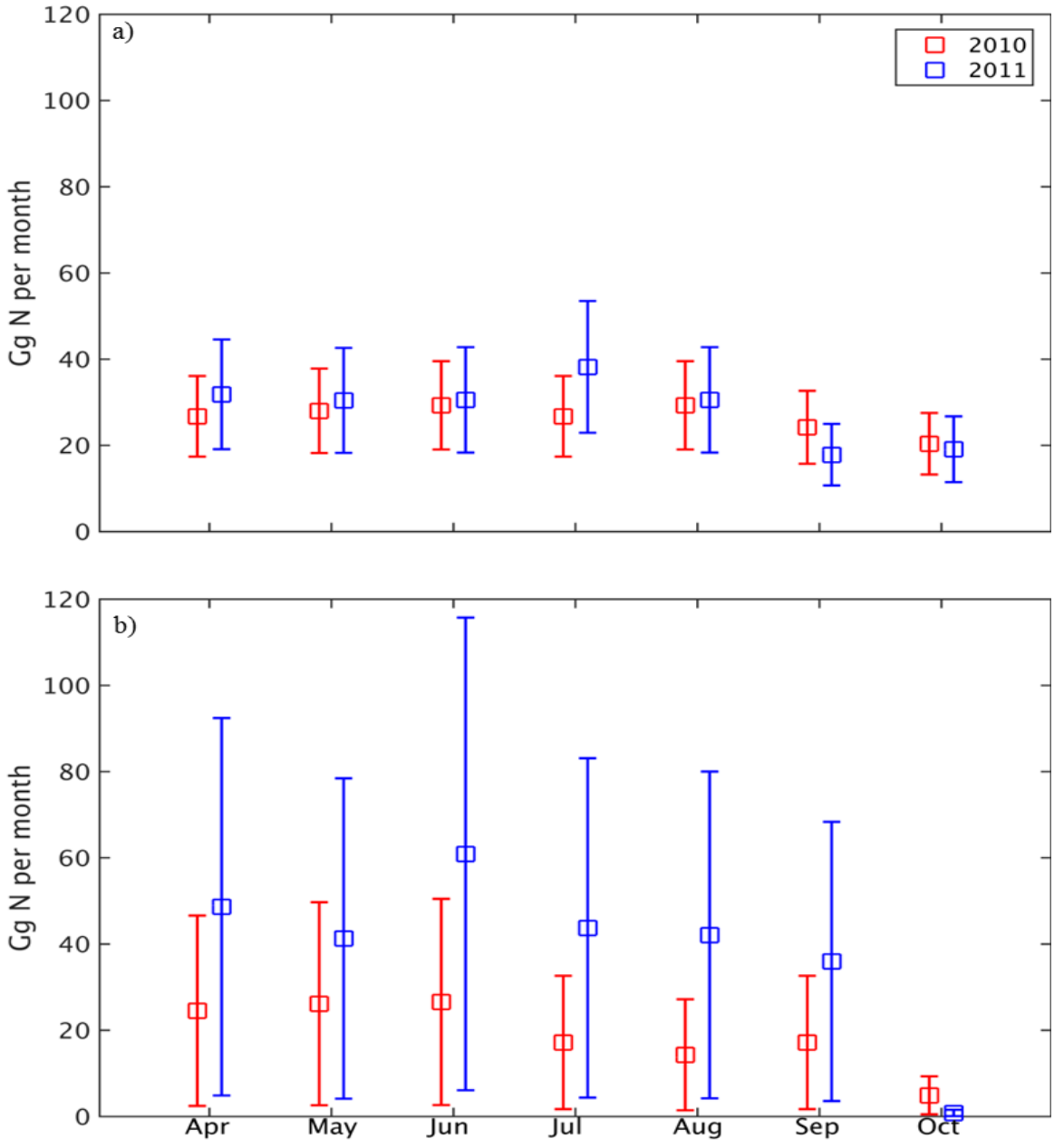


Figure 2.4. a) Direct and b) indirect N_2O emissions from agriculture within the US Corn Belt for 2010 and 2011, respectively. Error bars represent the *a posteriori* uncertainties after optimization, for direct and indirect emissions.

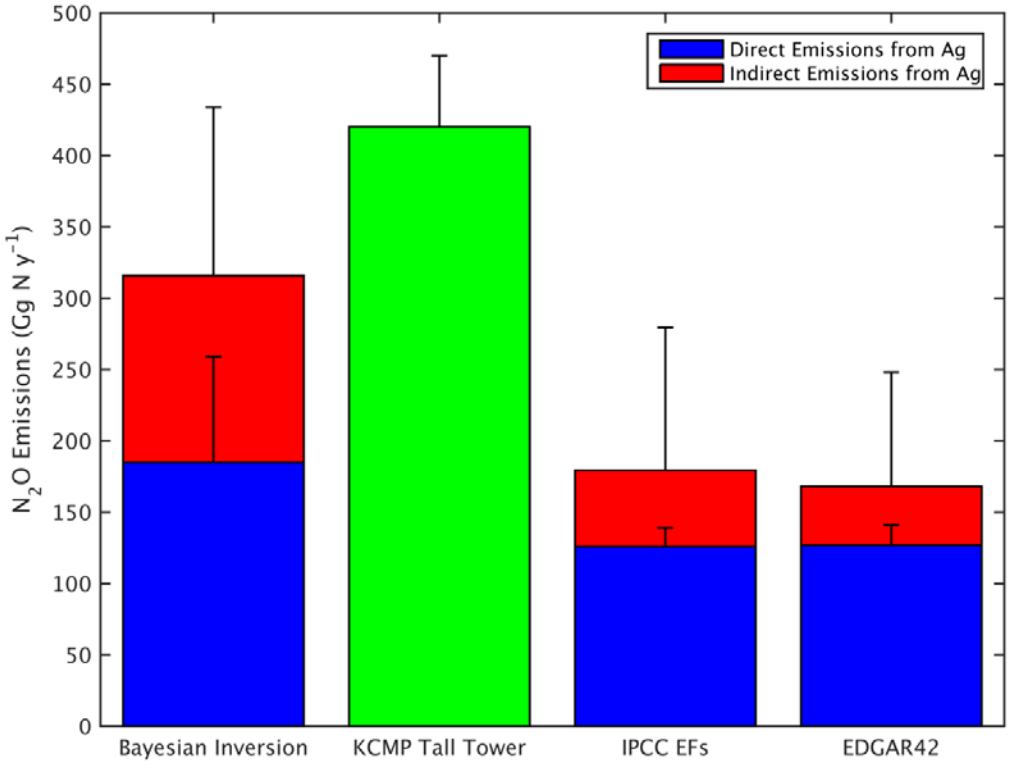


Figure 2.5. Comparison of N_2O budgets for the US Corn Belt in year 2010 estimated using different methods [tall Tower: boundary layer method at the KCMP tall tower [*Griffis et al.*, 2013], IPCC EFs: estimate from IPCC EF method]. Error bars indicate the uncertainties of regional budget estimate from direct and indirect emissions, respectively (the Bayesian inversion shows the N_2O budget from April to October in 2010, a total of 7 months; the KCMP tall tower, IPCC EFs, and EDGAR42 all show an annual budget).

Supporting Information for Chapter 2

Solutions of Bayesian Inversion

$$y = K\Gamma + \varepsilon \quad \text{(S2.1)}$$

Where y is the observed minus background mixing ratios, the columns of *K* correspond to the simulated mixing ratios for each of the source types being optimized, and Γ consists of the *a posteriori* scale factors for the 7 source types.

Γ_1 is the scaling factor for *dirA*;
Γ_2 is the scaling factor for *indA*;
Γ_3 is the scaling factor for *waste*;
Γ_4 is the scaling factor for *industry*;
Γ_5 is the scaling factor for *energy*;
Γ_6 is the scaling factor for *natsoil*;
Γ_7 is the scaling factor for *BB*.

Since the Bayesian inversion is conducted each month for each inversion, there are 24×30=720 hourly observed and simulated mixing ratios. The dimension of the matrix K is 720×7 .The cost function *J(Γ)* used to determine Γ is Equation 2.2. The *maximum a posteriori* (MAP) solution of Γ *is* to minimize the cost function *J(Γ)*, S_ε is the

observational error covariance matrix, the matrix equals $\sigma_0 \cdot I$, where I is the identity matrix. We are assuming the observational error (σ_0) is constructed and calculated via quadrative, $\sigma_o = \sqrt{\sigma_m{}^2 + \sigma_p{}^2 + \sigma_{mh}{}^2 + \sigma_b{}^2}$, where σ_m is the measurement error, σ_p is the error from particles, σ_{mh} is the error from simulated mixing height, σ_b is the error from background mixing ratios , in a unit of ppb. For simplicity, these individual errors are independent.

S_a is the *a priori* error covariance matrix, it's a diagonal matrix. Along the diagonal, the percentage uncertainty is given corresponding to *the a priori* uncertainty of each source type.

Chapter 3
Assessing future nitrous oxide emissions and mitigation priorities for the U.S. Corn Belt

Synopsis

There is an important need for advancing a measurement-modeling framework for assessing nitrous oxide (N_2O) emissions, identifying potential regional hotspots, and evaluating if implemented mitigation strategies are working. We combined multisite tall tower N_2O mixing ratio observations (2011 to 2017) and an Eulerian modeling approach to constrain N_2O emissions at relatively high temporal and spatial resolution and to project future emissions (2018 to 2050) in the US Midwest. The strongest N_2O emission trends were 0.13 nmol m^{-2} s^{-1} per year, highlighting that southern Minnesota and the lower Missouri Drainage Basin will likely represent regional hotspots and mitigation priorities. The emission factor for agricultural N_2O emissions is projected to increase 22-27% (i.e., 0.04 in 2011 to 0.05 by 2050) in response to changes in climate. A reduction of new nitrogen inputs of 0.04 to 0.05 Tg per year will be required to stabilize N_2O emissions at 2011-2017 values.

Introduction

Nitrous oxide (N_2O) mixing ratios are increasing 0.80 ppb y^{-1} in response to increased global agricultural production and the intensive use of synthetic nitrogen (N) fertilizers [Smith, 2017]. Projected increases in synthetic N use and changes in climate are expected to enhance N_2O emissions [Davidson et al., 2009; Griffis et al., 2017]. There is a timely need, therefore, for an improved measurement-modeling framework for assessing N_2O emissions, identifying potential regional hotspots, and evaluating the effectiveness of mitigation efforts. Developing such a framework for N_2O is particularly challenging

given the sparsity of field observations, the episodic nature of N_2O emissions, and the strong spatial variability associated with direct and indirect N_2O emissions. Assessing N_2O emissions from agricultural systems at relevant temporal and spatial scales is difficult because N_2O is produced directly from fertilized soils and indirectly from leaching and runoff, volatilization, and re-deposition processes. Indeed, the uncertainties associated with estimating emissions from these sources ranges from 10% to 400% [Chen et al., 2016]. Previous studies have attempted to constrain N_2O emissions from the U.S. Corn Belt using bottom-up and top-down approaches [Miller et al., 2012; Turner et al., 2015; Chen et al., 2016; Griffis et al., 2017]. These studies suggest that the indirect emissions from leaching and runoff play a more important role than previously thought. Indirect N_2O emissions were estimated to be 2.7- to 5.5-fold larger than emissions calculated using bottom-up inventories (e.g., EDGAR4.2) and were shown to dominate the inter-annual variability of total emissions [Griffis et al., 2017]. However, these recent inverse modeling efforts [Chen et al., 2016; Griffis et al., 2017] were applied on relatively coarse resolutions in both time (i.e., monthly) and space (i.e., regionally) and are, therefore, limited with respect to identifying 'when and where' large N_2O emissions occur within the region. There is a need, therefore, to capture the fine-scale N_2O variations in response to changes in climate and land management to enhance our capacity to identify critical emission locations and time periods to help develop specific and targeted GHG mitigation plans.

Over the past decade there has been an increase in the number of N_2O mixing ratios observations over the United States from tall towers [Dlugokencky et al.,2011; Griffis et

al., 2013], aircraft [Frankenberg et al., 2016], and satellites [Jacob et al., 2016]. These observations have been combined with advancements in atmospheric chemistry and transport models (CTMs) [Henze et al., 2007, 2009; Kopacz et al., 2009] to facilitate broader application of forward and inverse models [Turner et al., 2015; Wells et al., 2018] used to constrain N_2O emissions. Further, there have been advances in land surface models that can simulate N_2O emissions based on biophysical relationships. For example, the Community Land Model (CLM45-BGC-CROP) can provide dynamic N_2O emission estimates at relevant temporal and spatial scales independent of bottom-up inventories.

Building on our previous work [Griffis et al., 2013, 2017; Chen et al., 2016; Fu et al., 2017, 2018],we have developed a novel framework by coupling an Earth System Model (WRF-Chem version 3.8.1 and the land surface scheme CLM-BGC-CROP version 4.5) and tall tower network observations to simulate retrospective and future N_2O emissions for the US Corn Belt. Here, we use 7 years (2011-2017) of N_2O mixing ratio data from 6 tall tower sites across the US and performed coupled model simulations at a spatial resolution of $0.125^o \times 0.125^o$ to address the following objectives:

(1) Identify and quantify direct and indirect emissions at fine spatiotemporal scales ($0.125^o \times 0.125^o$, and daily to weekly, respectively);

(2) Probe the seasonal and inter-annual variability, and the emission budgets under future climatic and nitrogen fertilizer scenarios;

(3) Identify potential mitigation priority locations in guiding efficient and economical N_2O reduction efforts.

Methods

Approach Overview

The measurement-modeling framework employed in this study is outlined in Figure S3.1. Briefly, our approach involved three phases: First, hourly N_2O simulations from the WRF-Chem 3.8.1 model were driven by the chemical fields derived from the CLM45-BGC-CROP land surface scheme. In phase 1, model simulations were evaluated against observed N_2O mixing ratios from the NOAA sites and the KCMP tall towers for the year 2011 (Table S.S1). After examining the model-data biases, modeling tuning was performed *via* the adjustment of key parameters and algorithms (Supporting Information (SI), Table S.S3).The adjusted CLM45-BGC-CROP scheme was used again to drive the Eulerian model (WRF-Chem) for N_2O mixing ratio simulations, followed by another round of model-data bias evaluation. This model tuning was performed iteratively until reaching a minimum of the model-observation bias. Following the optimization of the CLM45-BGC-CROP scheme, we further tested and evaluated the model performance for the years 2012-2017 (Phase 2). Finally, based on the optimized model we simulated retrospective and future N_2O emissions for the US Corn Belt using various N fertilization management and climate (i.e., RCP 2.6 and RCP 8.5) scenarios (Phase 3).

Study domain and observation sites

The study domain is focused on the US Corn Belt and includes the major corn/soybean production systems of Minnesota, Illinois, Indiana, Iowa, Missouri, Ohio, South Dakota, Nebraska, and Wisconsin. Continuous N_2O mixing ratio measurements across the US

were made at hourly (KCMP tall tower) to near-weekly (NOAA CCGG system) time scales [Table S.S1]. Four of the tall towers, including KCMP, LEF, WBI, and BAO, are located within the Corn Belt or the inner domain of the WRF-Chem model. Three tall towers AMT, SCT and WKT, were located northeast, southeast and southwest of the domain of interest. The KCMP tall tower N_2O mixing ratios were measured using a tunable diode laser (TDL) technique (TGA100, Campbell Scientific Inc.) at 4 heights including 32, 56,100, and 185 m above the ground. The TDL calibrations were performed hourly with standard gases traceable to the National Oceanic and Atmospheric Administration-Earth System Research Laboratory (ESRL) 2006A N_2O mole fraction scale. Further details regarding the sampling procedures and calibrations can be found in Griffis et al [2013, 2017]. The surface observations are from discrete air samples collected approximately weekly in flasks in the NOAA CCGG program [Dlugokencky et al., 1994], which are analyzed using a gas chromatograph with an electron capture detector and reported on the NOAA 2006A calibration scale. An uncertainty of 0.4 ppb is assumed following previous studies [Wells et al., 2015, 2018; Chen et al., 2016]. Note that observations from the above NOAA sites are collected during the daytime.

CLM45-BGC-CROP land surface scheme

The CLM45-BGC-CROP scheme (CLM hereafter) is a process-based land surface model that provides an extensive description of the crop phenology and biophysical processes, and simulates energy, water, carbon and nitrogen budgets at field to regional scales [Chen

et al., 2015; 2018; Leng et al., 2016]. The CLM scheme provides a platform to advance a process-based approach to constraining regional emissions and to assess future N_2O emissions. A detailed description on the model spin-up, meteorological forcing and the algorithms and key parameterizations are provided in the supporting information (SI).

WRF-Chem model

An N_2O transport module was embedded in the WRF-Chem model (version 3.8.1) in order to simulate the N_2O mixing ratio for a 7-year period (2011-2017) at hourly resolution across the US Corn Belt. Briefly, the model set up consisted of the following key steps: (1) N_2O was treated as a passive tracer ; (2) the WRF-Chem simulations were performed using two nested domains at 70 km (outer) and 10 km (inner) grid spacing; (3) direct and indirect agricultural N_2O emissions were simulated using the CLM scheme to drive the WRF-Chem model; (4) the original initial and boundary conditions for the N_2O mixing ratios for each modeling period were obtained from the global chemical model, Model for Ozone and Related Chemical Tracers (MOZART) version 4 (http://www.acom.ucar.edu/wrf-chem/mozart.shtml). These values were found to be biased low compared to the N_2O background values from previous studies [Wells et al., 2015; Chen et al., 2016; Griffis et al., 2017] and were therefore corrected based on the continuous measurements from Niwot Ridge, Colorado (40.04^o N, 105.54^o W, elevation: 3018 m). Details regarding the model set up are provided in the Supporting Information (S1).

Future trend projections

Following the optimization of the CLM scheme we simulated future N_2O emissions to the year 2050 following different greenhouse gas forcings (i.e., RCP 2.6 and RCP 8.5) and N fertilization scenarios in order to probe the future spatiotemporal trends in regional agricultural N_2O emissions. Scenarios RCP 2.6 and RCP 8.5 indicate the representative concentration pathway of +2.6 and +8.5 watts per square meter radiative forcing by the year 2100. Furthermore, the future trend of nitrogen fertilizer applied in this region was estimated based on the FAO report (Food and Agriculture Organization of the United Nations, 2018), following the projected demand for food, fiber and biofuel energy. This corresponds to an increase of 0.08 Tg N per year within this region.

Results and Discussion

Spatiotemporal variability of N_2O emissions

Overall, the modeled posteriori N_2O mixing ratios at each tall tower location for 2011 were moderately to strongly correlated with the tall tower observations. Stronger correlations were observed at the KCMP ($r=0.79$, $p<0.01$), WBI ($r=0.66$, $p<0.01$), and BAO ($r=0.69$, $p<0.01$) sites compared to the LEF site($r=0.36$, $p<0.01$, Figure 3.1). This supports that the Eulerian modeling framework performs reasonably well in capturing both the transport processes and the biogeochemical processes and that the optimized CLM scheme has significantly improved the constraint on the N_2O emissions (Figure 3.1, Figure S3.2).

Both the model and observed values show strong temporal variability and relatively high mixing ratios at the KCMP ($r=0.79$, $p<0.01$) and the WBI ($r=0.66$, $p<0.01$) sites. Land use within the vicinity of these sites is dominated by agriculture including corn and soybeans [Griffis et al., 2013]. In contrast, the LEF site is surrounded by a mixed forest landscape and the model and observations show relatively low variability. Therefore, the optimized model revealed good overall performance in capturing the temporal variability, especially for the agriculture-dominated regions.

We extended our analyses to the period 2012-2017 to assess how well the model captured the inter-annual variability. Consistent with the year 2011, the optimized model showed very good agreement with the observations (Figure 3.1b, d, f, and h). The KCMP and WBI sites showed high seasonal and inter-annual variability, especially in the growing season, and indicated the strongest correlations (Table S.S2) within the Corn Belt. There was improved model accuracy for the LEF tower site when evaluated over the 7-year timespan, which is attributed to the slightly higher variability in both the modeled and observed N_2O mixing ratios relative to 2011 (Figure S3.4). Given the reasonably good performance of simulating N_2O mixing ratios across 4 sites and over 7 years, we use the optimized model to help understand the spatiotemporal variability in direct and indirect N_2O emissions.

Direct and indirect agricultural N_2O emissions

Figure 3.2 shows the monthly variation in the growing season direct and indirect emissions in year 2011. Direct emissions began to increase after the spring thaw, typically initiated in mid-March [Griffis et al., 2017], and increased following spring

fertilization in the early growing season. Direct emissions typically peaked over the period June to July, gradually spreading across the entire Corn Belt, i.e., from south to north. For each year of the 7-year time series the spatial distribution of direct emissions showed high coherence to 2011 and reflected the underlying pattern of synthetic N fertilization and agricultural land use (Figure S3.7). In contrast, the indirect emissions revealed more episodic features and high spatial variability (Figure 3.2) at seasonal to inter-annual scales. There were pronounced indirect emissions in April (years 2011-2017) and regional hotspots emerging in locations of strong precipitation, large stream flow, and spring N fertilization. Interestingly, there is also high spatial variability in indirect emissions year-to-year, corresponding to variability in precipitation and the resulting stream flow and drainage patterns [Turner et al., 2015; Griffis et al., 2017].

To further investigate how direct and indirect sources drive the N_2O spatiotemporal variability, we conducted sensitivity tests where the WRF-Chem model was driven by: (1) direct emissions only, (2) indirect emissions only; and (3) direct +indirect emissions (Figure S3.8). These tests were performed to evaluate the relative importance of these emissions and their influence on the tall tower mixing ratio observations in 2011.These tests revealed that indirect sources had a significant influence on the WBI and KCMP tall tower observations. For instance, the contribution from the indirect sources was about 3.9 ppb larger in May, 2011compared to direct sources, Further, the contribution from indirect sources were strongly correlated with observations from the KCMP and WBI (R^2=0.69 and 0.61, respectively) towers, indicating that they play a dominant role in

driving the spring-time variations of atmospheric N_2O mixing ratios within this agricultural landscape.

Future regional hotspots, mitigation priorities, and implications

Using the optimized CLM scheme, we simulated N_2O emissions from present to 2050 following the N fertilization management, as well as the RCP2.6 and RCP8.5 emissions scenarios. Over the 40-year period (RCP 8.5 scenario), direct emissions steadily increased over the cropland distribution in response to increasing air and soil temperature. The episodic characteristic and spatiotemporal variability of indirect emissions became more pronounced over the 40-year period when compared to the 2011-2017 average. This was driven by the increased precipitation and increased frequency of more extreme events leading to increased runoff and stream flow. The overall area and intensity of "hot regions" increased in response to the pronounced increase in air temperature and precipitation.

To identify locations with the strongest increasing trends associated with direct and indirect sources, i.e., hot grid cells, we calculated the slope of emissions over the 40-year period at grid scales (~10 km×10 km, Figure 3.3g-i for RCP8.5 scenarios). The largest slopes for N_2O emissions were 0.13 nmol m^{-2} s^{-1} per year, ~23 times greater than the average trend over the entire region (i.e., 0.0057 nmol m^{-2} s^{-1} per year). These analyses suggest that, indirect emissions showed a significantly stronger increasing trend for multiple regions (e.g. southern Minnesota; the boundary areas between Nebraska and

Missouri) (Figure 3.3h). The combined influence of both increased direct and indirect emissions contributed to hot regions located in the Lower Midwest, especially near the intersection of Nebraska, Kansas, Missouri and Iowa. This region is characterized by an agriculture-dominated landscape and is drained by the Missouri River, which forms one of the largest drainage basins in the US (Figure S3.7).

As expected, the future trends under the RCP2.6 scenarios (Figure S3.9a-c) indicated a more moderate response. Trends in direct emissions indicated larger areas near Missouri and Illinois that show a strong increasing slope (Figure S3.9d versus Figure S3.9a) under RCP8.5 compared to RCP2.6. Furthermore, the stronger increasing slope under RCP8.5 showed much larger spatial coverage than in RCP2.6 (FigureS9e versus Figure S3.9b), concentrated near the southern Minnesota and the Lower Missouri Drainage Basin. A comparison between the future trends under the RCP2.6 versus RCP8.5 scenarios further supports that there will be a strong positive feedback between climate change and N_2O emissions. Pronounced regional hotspots are expected for both the RCP8.5 and RCP2.6 scenarios with indirect emissions playing a dominant role.

To explore the relative importance of climate (i.e. air temperature and precipitation) on simulated N_2O emissions, we conducted another sensitivity test by keeping the synthetic N input constant over the 40-year period (Figure S3.10). Here, indirect emissions showed much stronger sensitivity to the change in climate (*p-value* <0.01), especially to precipitation (r^2 =0.59, *p-value*= 5.9E-7). Further, indirect emissions revealed an increasing trend (0.001 nmol m^{-2} s^{-1} mm^{-1}) that is ~10 times greater than direct emissions (8.1E-5 nmol m^{-2} s^{-1} mm^{-1}) in response to increasing precipitation. The dominant driver

of the increased emission forecast is the link between increasing precipitation (2.7 mm yr^{-1} under RCP 8.5 scenario) and indirect emissions.

We examined how the EF for the region might change in response to climate. Annual emissions increased by 0.0047 nmol m^{-2} s^{-1} yr^{-1} (i.e., 0.0007 (direct) and 0.004 nmol m^{-2} s^{-1} yr^{-1} (indirect)). By the year 2050, the EF for agricultural N_2O emissions is estimated to increase 22% to 27% (i.e., from ~0.04 in year 2011 to ~0.05 in year 2050). The newly fixed N (i.e. synthetic N + BNF) entering agricultural systems [Crutzen et al., 2008], was estimated to be 9.7 Tg N (2.6 Tg N from BNF (biological N fixation) and 7.1 Tg N from synthetic N) in 2011. Therefore, a reduction of new nitrogen inputs of 0.04 to 0.05 Tg per year will be required to stabilize N_2O emissions at 2011-2017 values *via* mitigation efforts associated with improvements in nitrogen use efficiency (NUE) or different cropping systems[Zhang et al., 2015].

Conclusions

Our measurements and optimized Eulerian model analyses support the following conclusions:

1. The a posteriori N_2O simulations were strongly correlated ($r=0.69$-0.82, $p<0.01$) with observations across the US Corn Belt, especially agriculture-dominated areas, supporting that the optimized CLM scheme captures fine-scale spatiotemporal variations from direct and indirect sources;
2. During the growing season, indirect emissions dominated the spatiotemporal variations within the Corn Belt, and contribute significantly larger emissions to the regional budget;

3. The episodic characteristic and spatiotemporal variability of indirect emissions became more pronounced of the 40-year period when compared to 2011-2017 average, driven by the increased precipitation and increased frequency of more extreme events leading to increased runoff and stream flow;
4. Under RCP8.5 scenario, the largest trends for N_2O emissions were 0.13 nmol m^{-2} s^{-1} per year, highlighting that southern Minnesota and the lower Midwest regions drained by the Missouri River will likely represent regional hotspots and mitigation priorities in the future;
5. Under the RCP8.5 scenario the EF for agricultural N_2O emissions is projected to increase 22-27% in response to climate change. A reduction of new nitrogen inputs of 0.04 to 0.05 Tg per year will be required to stabilize N_2O emissions at 2011-2017 values.

Acknowledgement

Financial support for this research has been provided by the National Science Foundation (grant number 1640337); and the USDA National Institute of Food and Agriculture (grant number 2018-67019-27808); and the USDA-ARS. We express our sincere thanks to Minnesota Public Radio and Tom Nelson for the logistical support at the KCMP tall tower. The Minnesota Supercomputing Institute provides key computing resources and assistance. Data are hosted at http://www.biometeorology.umn.edu/research/dataarchives and ESS-DIVE. We acknowledge the support of an MnDrive PhD fellowship to Z.C.

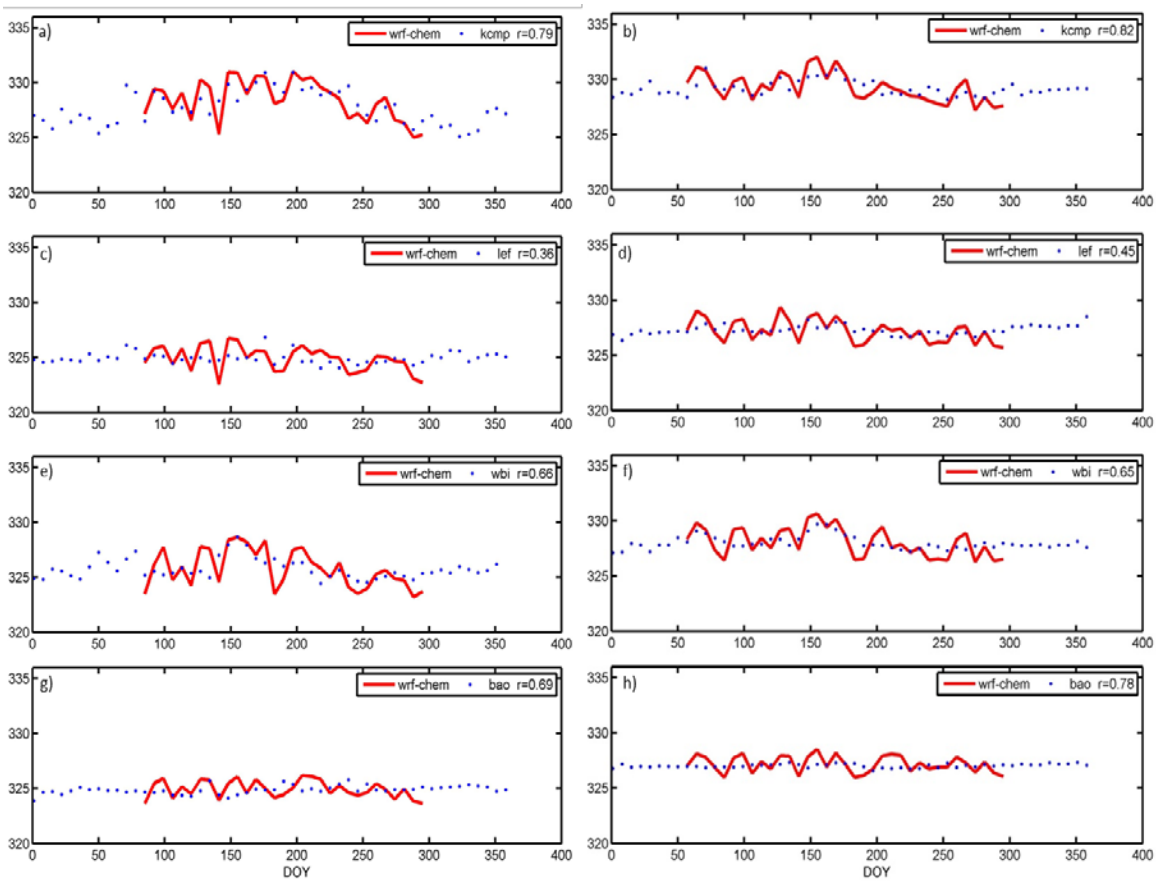


Figure 3.1. Weekly variations of N_2O mixing ratios from the tower-based observations and models in a) year 2011, and b) 7-year average (2011-2017).

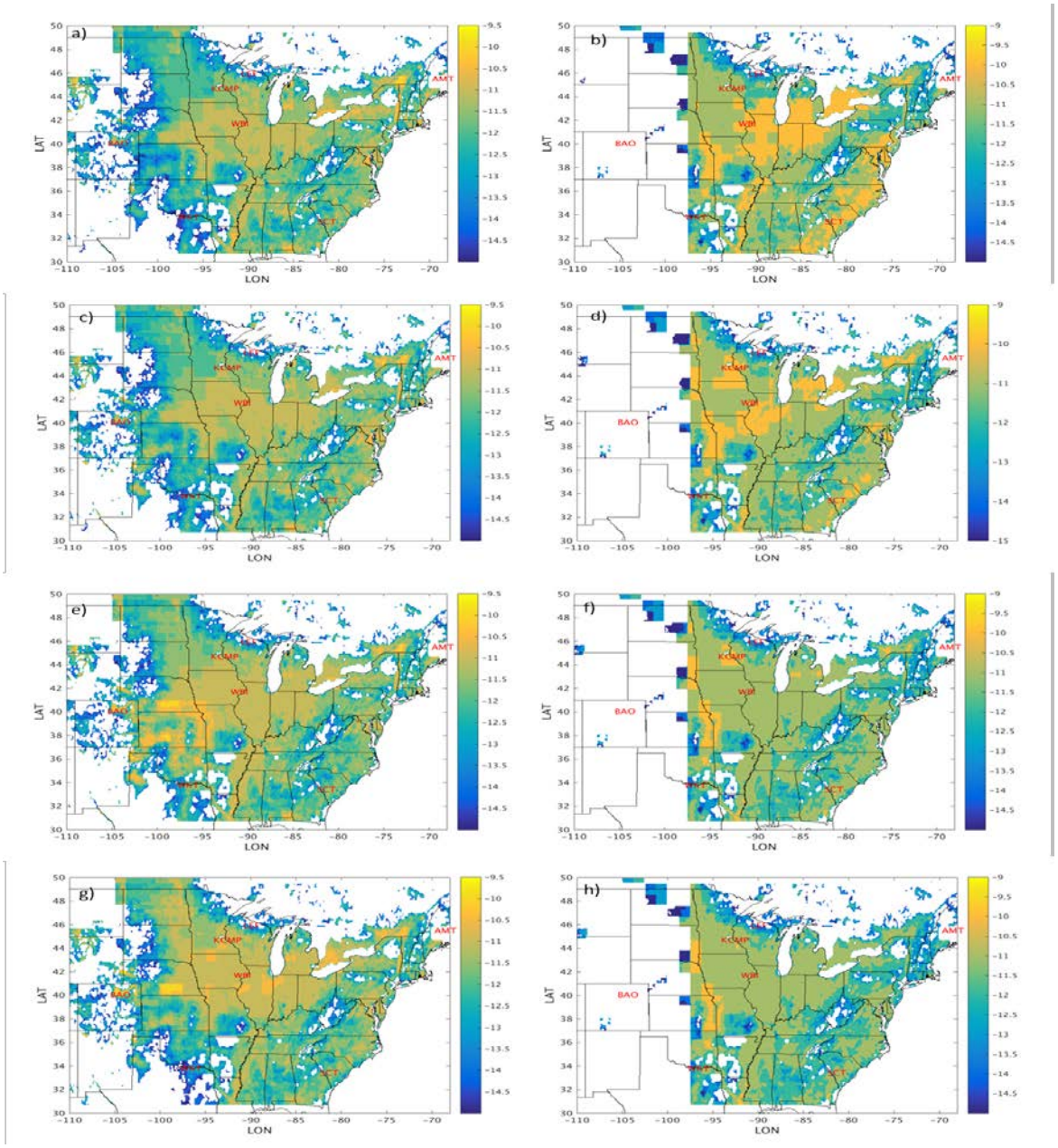


Figure 3.2. Monthly direct and indirect emissions for April, May, June and July in year 2011 (unit is $\log_{10}(\text{kg m}^{-2}\ \text{s}^{-1})$).

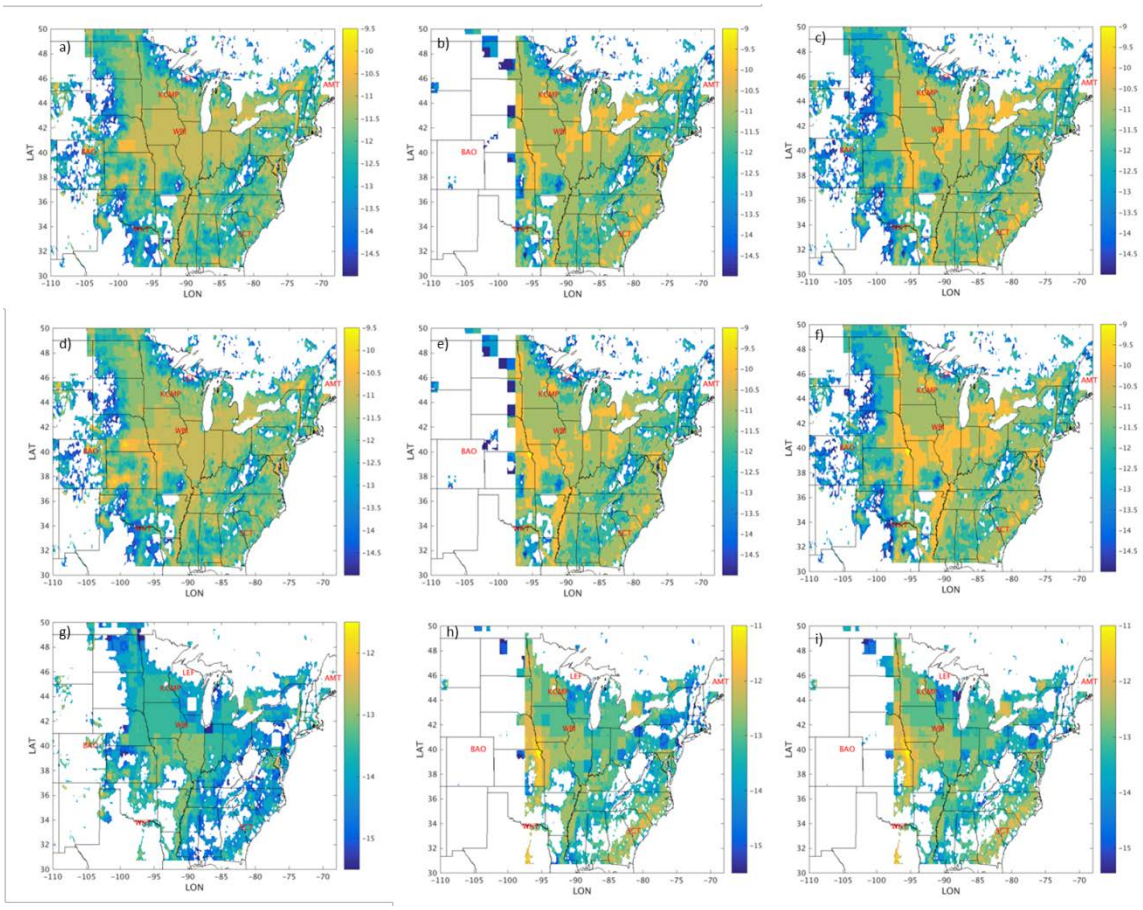


Figure 3.3. Average a) Direct, b) indirect and c)total agricultural N_2O emissions from year 2011; d) Direct, e)indirect and f)total agricultural N_2O emissions from year 2050 (unit of kg m^{-2} s^{-1}); Increasing slope (unit of kg m^{-2} s^{-1} per year) for g) direct, h) indirect, i) total emissions over a 40-yr period under the RCP 8.5 scenarios. Relative uncertainty of 7% and 19% for direct and indirect emissions, respectively, were applied for each grid cell across the US Corn Belt, derived from the sensitivity test based on the key parameter variations in the CLM model.

Supporting Information for Chapter 3

CLM45-BGC-CROP model

The CLM45-BGC-CROP scheme (CLM hereafter) is a process-based land surface model that provides an extensive description of the crop phenology and biophysical processes, and simulates energy, water, carbon and nitrogen budgets at field to reginal scales [Chen et al., 2015; Leng et al., 2016]. The CLM scheme incorporates the Century N-gas model [Parton et al., 1996; Del Grosso et al., 2001], which includes separate NH_4^+ and NO_3^- pools, as well as environmental controls on nitrification and denitrification rates. Further, *via* a 'hole in the pipe' approach [Davison et al., 2000], this model also predicts the N_2O emissions from nitrification and denitrification.

Spin-up process

The model was run offline and forced with the NCEP reanalysis meteorological dataset (*https://rda.ucar.edu/datasets/ds083.3/*). The spatial resolution and temporal resolution was $0.25^o \times 0.25^o$ and 6 h, respectively. The meteorological fields include solar radiation, air temperature, air humidity, air pressure, wind speed and precipitation. To ensure that the models reached a steady state, we performed two steps in our CLM45-BGC-CROP spin up processes, i.e., a 1000-year accelerated decomposition, followed by a 200-year additional normal run spinup mode [Thornton and Rosenbloom., 2005; Vertenstein et al., 2013].

Default CLM scheme and its limitations

The nitrogen availability, climate sensitivity and their combined effects, are the major factors that lead to the high bias in both the temporal variation and magnitude of agricultural N_2O emissions. In the default model the denitrification rate is unrealistically

high with approximately 50% of the unused available nitrogen lost (denitrified). This leads to an overestimated N_2O emission from denitrification as well as leaching/runoff. This effect would be even further enhanced during the spring time following synthetic nitrogen fertilizer application. Furthermore, a large nitrogen loss in the early growing season will cause relatively low N availability that will adversely affect crop growth. The default CLM scheme provides one solution, which is to give a 20-day period for the fertilizer application during the emergence phase that continues for 20 days.

In addition to the nitrogen availability, climate sensitivity plays a critical role in determining nitrification (Equation S1) and denitrification rates. Nitrification rate of NH_4^+ to NO_3^- is a function of soil temperature, soil moisture, and *p*H:

$$f_{nitr} = [NH_4] k_{nitr} f(T) f(H_2O) f(pH) \quad \text{(S3.1)}$$

where f_{nitr} is the nitrification rate, k_{nitr} is the maximum nitrification rate, and *f(T)* and *f(H_2O)* are rate modifiers accounting for soil temperature and soil water content. *f(pH)* is a rate modifier for *p*H. Note that the default model uses the same rate modifiers as are used in the decomposition routine for temperature and soil moisture.

Driven by the default CLM scheme, the *a priori* simulations of tower mixing ratios showed a significant underestimate in the early spring and a high overestimate in the later summer across all the observations sites within the Corn Belt (Figure S3.2), indicating an unreasonable seasonal variability and therefore very low statistical power (r=0.23). We attributed it to the inappropriate or over-simplified algorithms and parameterizations (Table S3.3) in the N_2O production mechanism, nitrification/denitrification, and crop phenology processes [Griffis et al., 2017].

Optimized CLM scheme

Previous studies [Parton et al., 1996, 2000; Del Grosso et al., 2001] under laboratory and field conditions that explore the climate sensitivity (i..e, air temperature, water-filled pore space (WFPS), and *p*H) were incorporated and adjusted into the algorithm..

$$f(H_2O) = (\frac{wfps - b}{a - b})^{d*(\frac{b-a}{a-c})} * (\frac{wfps - c}{a - c})^{d} \quad \text{(S3.2)}$$

Where $f(H_2O)$ is the effect of WFPS on nitrification fraction, *wfps* is the water-filled pore space (0-1), parameter *a*, *b*, *c*, *d* is 0.55, 1.70, -0.007, 3.22 for sandy soils, and 0.60, 1.27, 0.00012, 2.84 for medium-textured soils, respectively.

$$f(T) = 0.06 + 0.13e^{0.07*soilt} \quad \text{(S3.3)}$$

Where *f(T)* is the effect of soil temperature on nitrification fraction, *soilt* is the soil temperature at 10 cm depth (in ^{o}C).

Building on the updated algorithms, a number of key parameters were further evaluated and adjusted to reach optimal values (Table S3.3), in order to represent accurate processes related to N_2O production mechanism, nitrification/denitrification, and crop phenology. Detailed modifications to improve the simulations of N_2O were outlined in the Model optimization section below.

Modeling Optimization

The CLM scheme was optimized *via* a wide range of adjustments in the algorithms and parameterizations, based on th empirical evidence, and previous studies [Del Grosso et al., 2000; Parton et al., 2001; Griffis et al., 2017]. After adopting the empirical functions from Del Grosso et al [2000] and Parton et al [2000] (i.e., Equation S2 and S3), we

further optimized the model with efforts on key parameter adjustments (Table S3.3), e.g., Growing Degree Days, maximum nitrification rate, fertilization timing.

Table S3.3 indicated the original and optimal process-related parameters. Key parameters are mechanistically linked to N_2O productions based on our best expert knowledge and previous studies, followed by being tested and precisely optimized from a wide range of sensitivity test throughout. Here we want to show several representative examples to illustrate the optimization process.

The *a priori* mixing ratios were significantly lower (by 3-5 ppb) than the observations from April to early May (Figure S3.2), which indicated a large underestimate in N_2O emissions for this time period. We posit that this underestimate arises from the late fertilization timing and/or a low fertilizer amount. Therefore, *via* varying the minimum GDD (growing degree days) for planting, the fertilization timing was adjusted earlier by a range of 3-11 days across the Corn Belt in the sensitivity test.. For instance, to our knowledge, the spring fertilization timing is DOY 121 in the year of 2011 around Minnesota, which was DOY 120 from the adjusted model. Meanwhile, the optimized model, during the springtime, indicated increased correlation ($r=0.61$-0.71, $p<0.01$) with the observations at multiple tower sites.

In the CLM scheme, denitrification rate is high and would result in a 50% of the unused available nitrogen, therefore, N fertilizer is applied during the emergence phase and continues for 20 days. Under this setting, we still found too aggressive N loss during the early growing season.. Therefore, we further extended the fertilizer application duration to 25, 30 and 35 days, and conducted a sensitivity test under the measurement-modeling framework, and reached an optimal duration of 30 days.

Likewise, we continue iterating through the optimization procedure described above under a measurement-modeling framework, via adjusting a range of key parameters (e.g., maximum nitrified fraction of $NH4_{+}$, GDD for harvest) until we had maximized the model-data correlation and minimized the model-data bias across multiple observation sites (Figure 3.1). Taken together, the optimal setting for the CLM45-BGC-CROP model is listed as Equations S3.2-S3.3 and Table S3.3.

WRF-CHEM model setup

A N_2O module was embedded in the WRF-CHEM model (version 3.8.1) in order to simulate the N_2O concentrations across the US Corn Belt. In this study, N_2O is treated as a passive tracer. The WRF-CHEM model was set up using two nests at 70 km and 10 km grid spacing. The outermost domain covers North America, and each domain was centered on the KCMP tall tower site. The simulations used the WRF Single-Moment three-class (WSM3) simple ice microphysics scheme [Hong et al., 2004], Kain-Fritsch convective scheme [Kain, 2004], and the Yonsei University (YSU) scheme coupled to the Noah land surface model for the planetary boundary layer (PBL) processes [Hong et al., 2006]. Initial and boundary meteorological conditions were provided by the National Center for Environmental Prediction Final Analysis (1° × 1°), with a 6 h interval.

Agricultural N_2O emission drivers

In this study, we used direct and indirect agricultural N_2O emissions simulated using CLM45-BGC-CROP scheme, as well as adjusted natural soil emissions from EDGAR2, to drive the WRF-CHEM model. Natural soil emissions were adjusted based on the Bayesian inversion from Chen et al [2016]. Direct emissions were derived from the

CLM scheme using modified algorithms [Parton et al., 1996; Del Grosso et al., 2001], while indirect emissions were derived using model estimates of nitrate runoff and leaching and applying the most recent estimates of the runoff/leaching emission factor for the Corn Belt region [Chen et al., 2016]. The derived direct, indirect, and natural soil emissions were gridded into two domains of 70km ×70km and 10km ×10km horizontal resolution, respectively, to chemically force the regional N_2O simulations.

Initial and boundary N_2O concentrations

Originally, the initial and boundary conditions for the N_2O mixing ratio for each modeling period were obtained from the Global Chemical Model, Model for Ozone and Related Chemical Tracers (MOZART) version 4 (http://www.acom.ucar.edu/wrf-chem/mozart.shtml). However, we found that the background N_2O concentrations were significantly underestimated by 5-10 ppb in the MOZART model during year 2010-2017, *via* comparison to various previous studies across this region [Wells et al., 2015; Chen et al., 2016; Griffis et al., 2017]. A low initial and boundary N_2O concentration would lead to a high-bias in the surface emissions and large mismatch with the tall tower mixing ratio observations. Therefore, we made corrections to the initial and boundary concentrations using an observational approach. Observations from the NWR site (Niwot Ridge, Longitude -105.6^o W, Latitude 40.1^o N, 3523 m masl.), as well as from our ongoing aircraft measurement campaign (the Greenhouse Emissions in the Midwest (GEM) project) across the Corn Belt that measures N_2O concentration above the atmospheric mixing heights, are used to correct the background concentrations.

Table S3.1. Sites of N_2O Observations in this Study

Site	Location	Country	Latitude	Longitude
KCMP	St Paul, Minnesota	United States	44.689	-93.073
AMT	Argyle, Maine	United States	45.035	-68.682
LEF	Park Falls, Wisconsin	United States	45.945	-90.273
WBI	West Branch, Iowa	United States	41.725	-91.353
BAO	Boulder Atmospheric Observatory, Colorado	United States	40.050	-105.004
SCT	Beech Island, South Carolina	United States	33.406	-81.833
WKT	Moody, Texas	United States	31.315	-97.327

Table S3.2. Correlations (*R*) between tower-based observations and the a posteriori simulations

	Year 2011	Year 2012	Year 2013	Year 2014	Year 2015	Year 2016
KCMP	0.79	0.7	0.69	0.58	0.7	0.79
LEF	0.36	0.34	0.38	0.32	0.57	0.49
WBI	0.66	0.67	0.67	0.61	0.51	0.62
BAO	0.69	0.7	0.68	0.82	0.65	0.8

Table S3.3. Original and optimal values for key parameters

Parameters	Fertilizer Amount (kg N ha^{-1})		Fertilizer timing	Maximum nitrified rate	Estimated emission factor for leaching/ runoff	Soil pH	GDD_{min} (^{o}days) for emergence		%GDD from leaf emergence to beginning of grain fill		Fertilizer application duration (days)
	Corn	Soybean					Corn	Soybean	Corn	Soybean	
Original	150	25	DOY 128	0.1	0.020	6.5	50	50	55-65	55-65	20
Optimal values	142.6	3.2	DOY 121	0.075	0.025	6.5	35	35	50	45	30

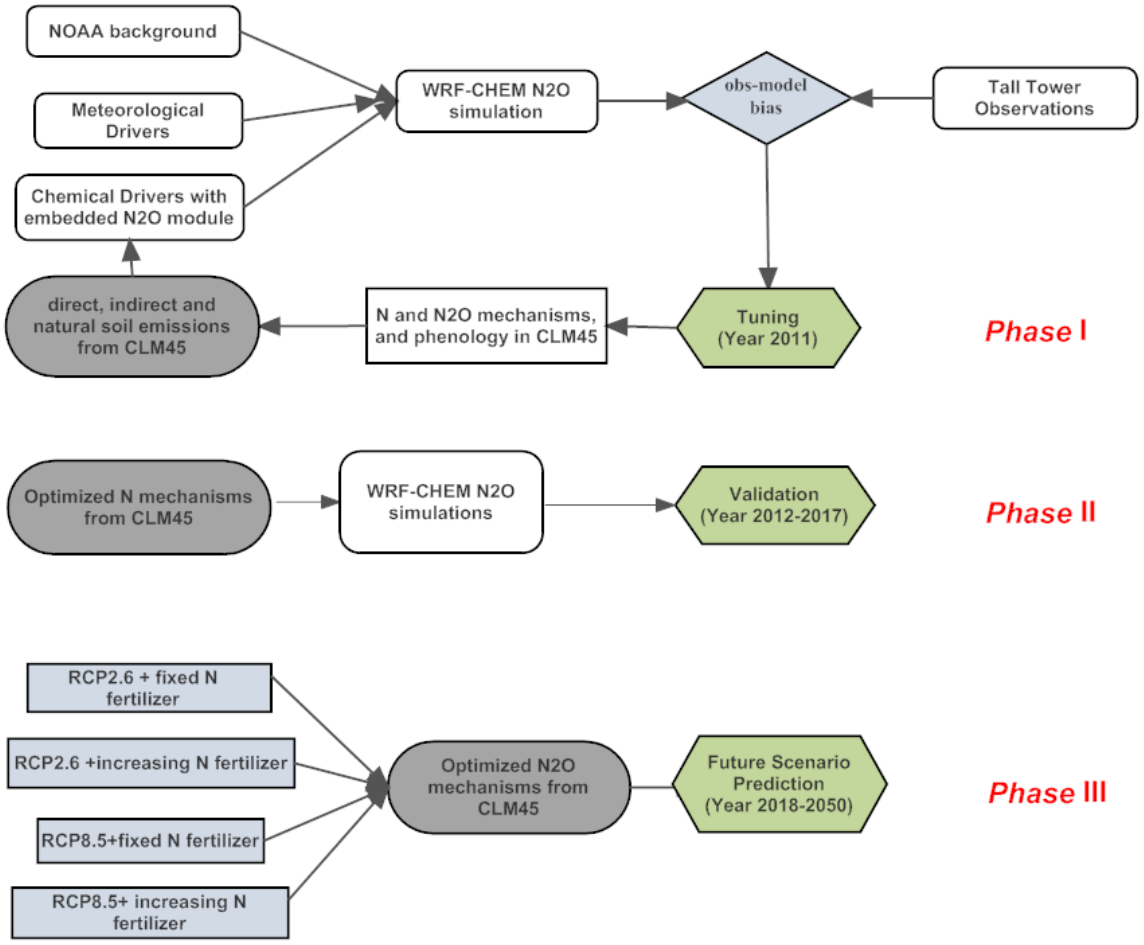


Figure S3.1. The flowchart of the Eulerian modeling configurations.

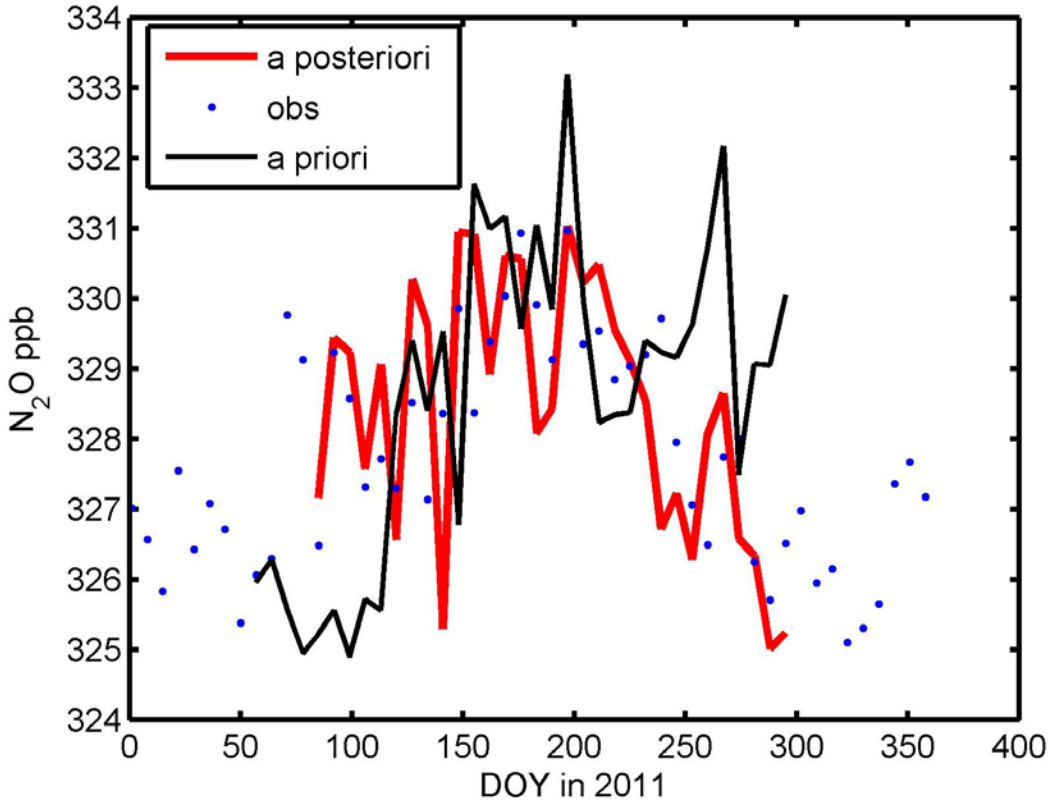


Figure S3.2. A priori, a posteriori and observed N_2O mixing ratios at the KCMP tall tower. (a priori, r=0.23)

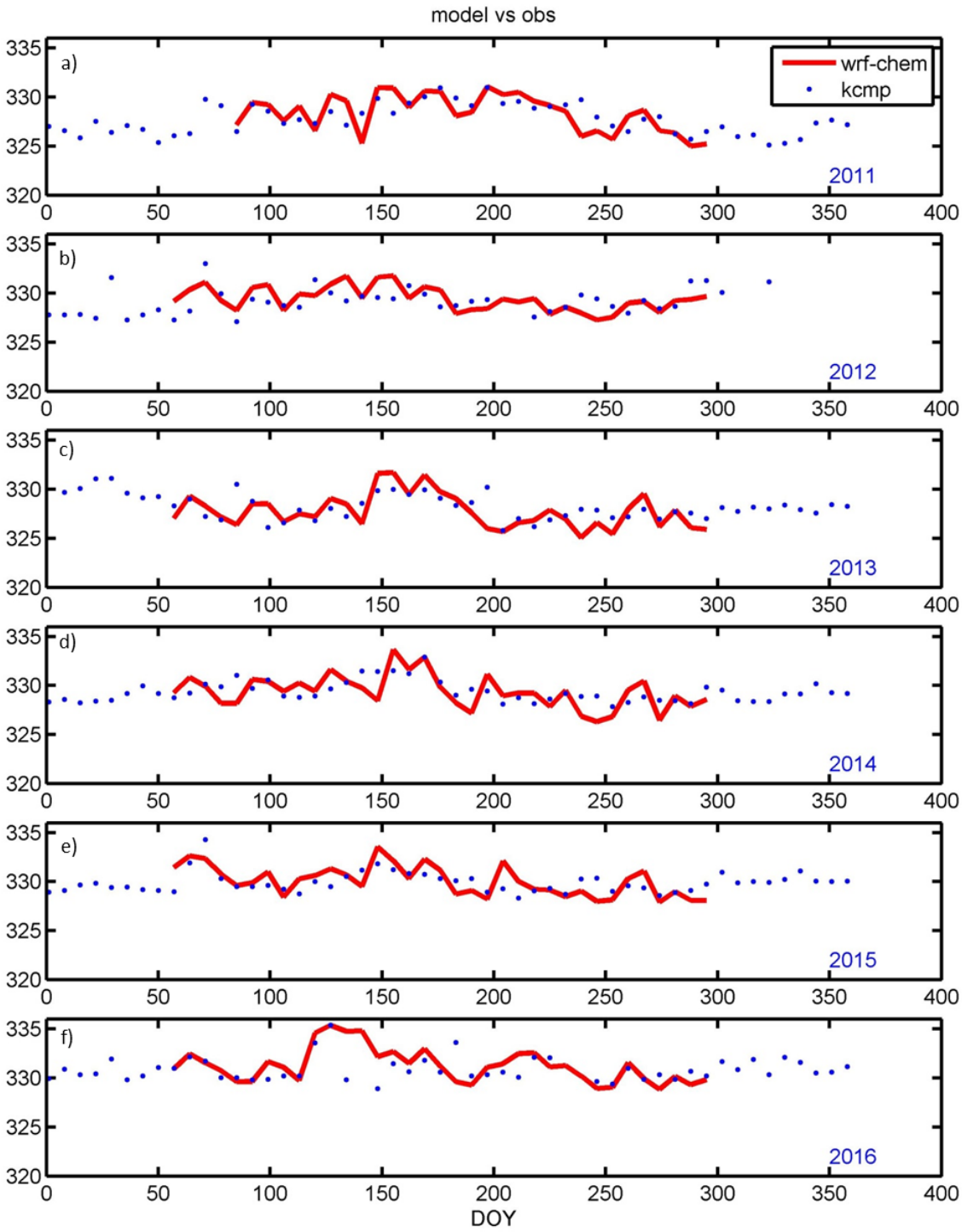


Figure S3.3. Weekly variations of N_2O mixing ratios from the tower-based models and observations in a) year 2011-2016 (a-f) at the KCMP site.

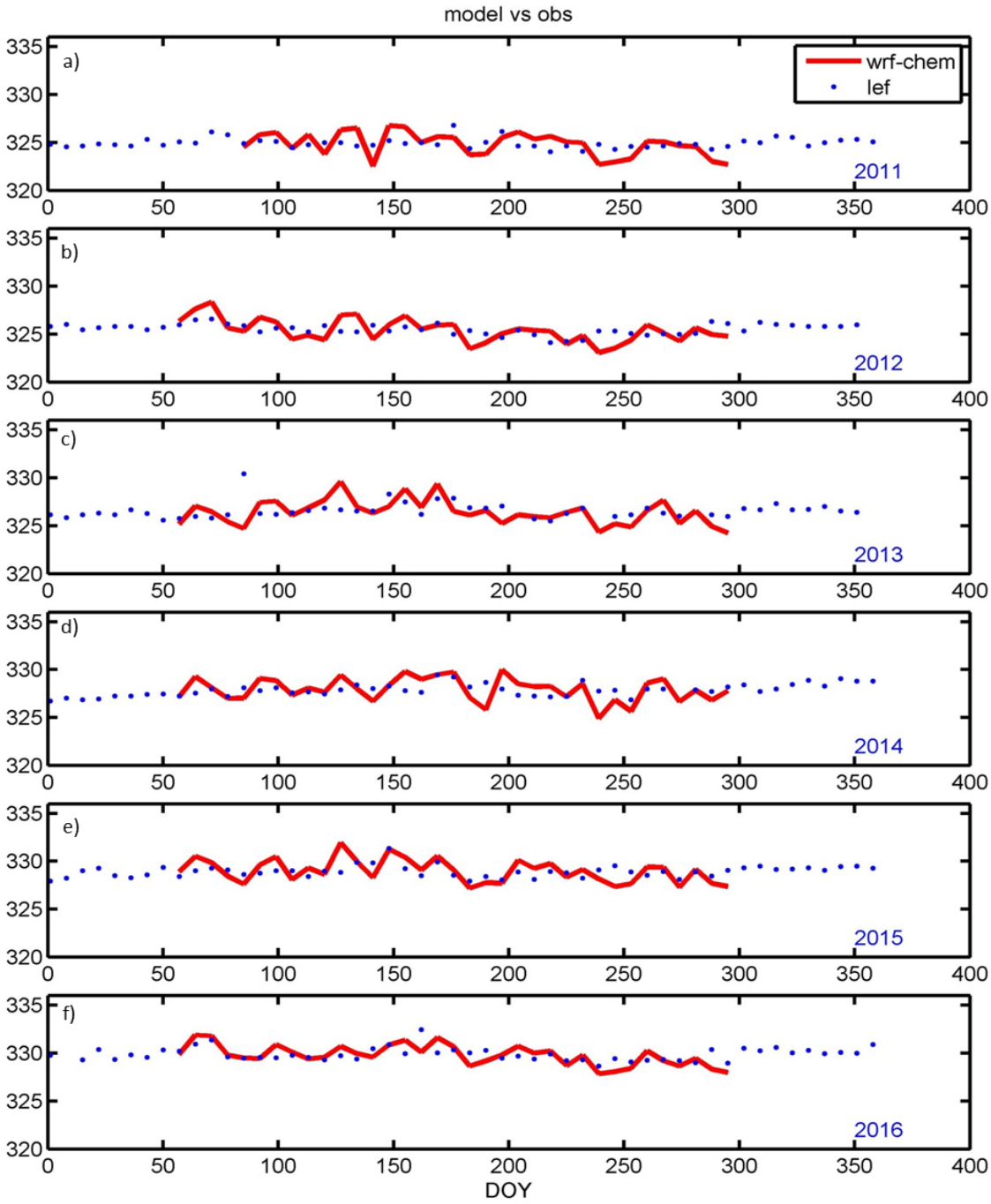


Figure S3.4. Weekly variations of N_2O mixing ratios from the tower-based models and observations in a) year 2011-2016 (a-f) at the LEF site.

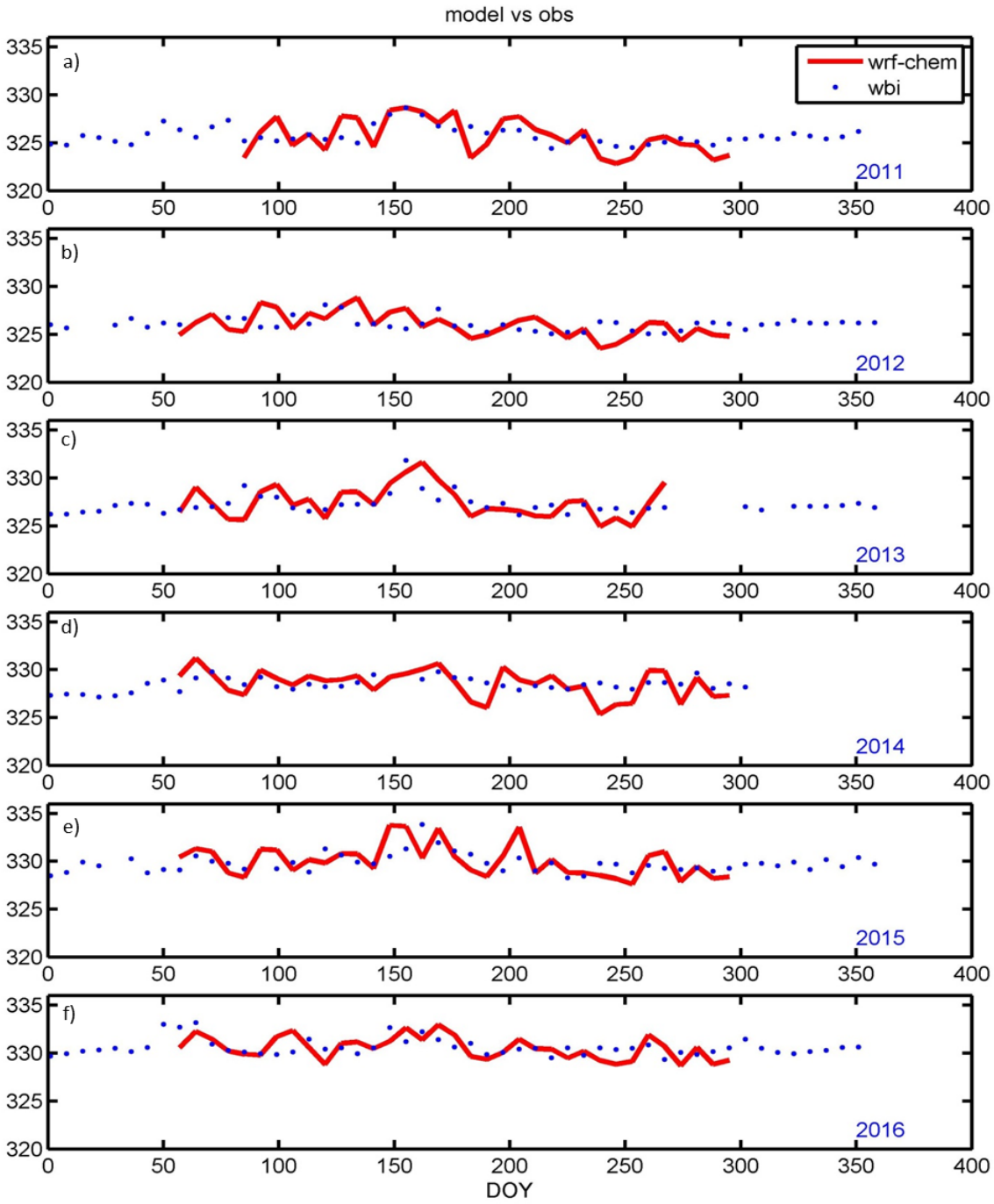


Figure S3.5. Weekly variations of N_2O mixing ratios from the tower-based models and observations in a) year 2011-2016 (a-f) at the WBI site.

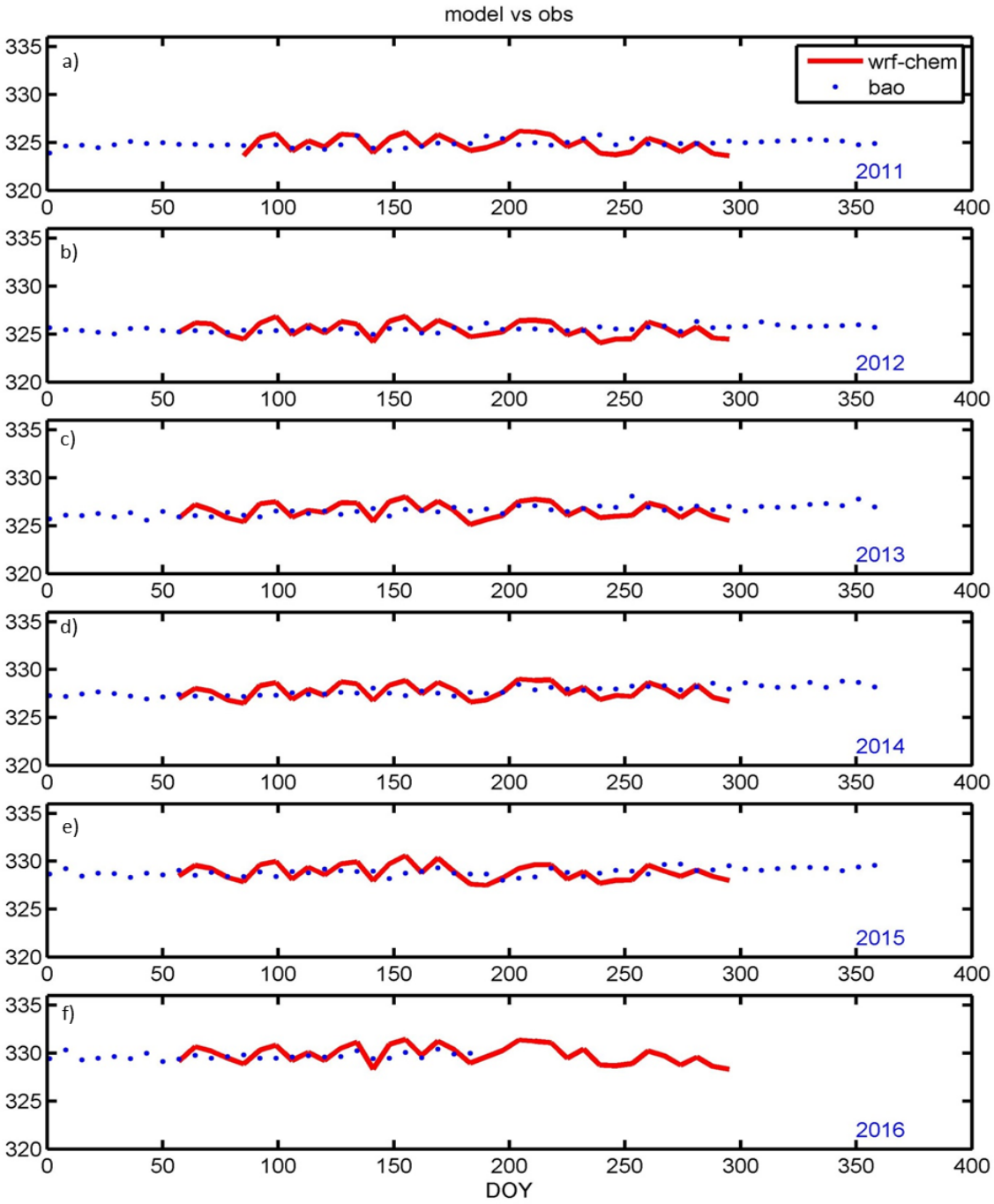


Figure S3.6. Weekly variations of N_2O mixing ratios from the tower-based models and observations in a) year 2011-2016 (a-f) at the BAO site.

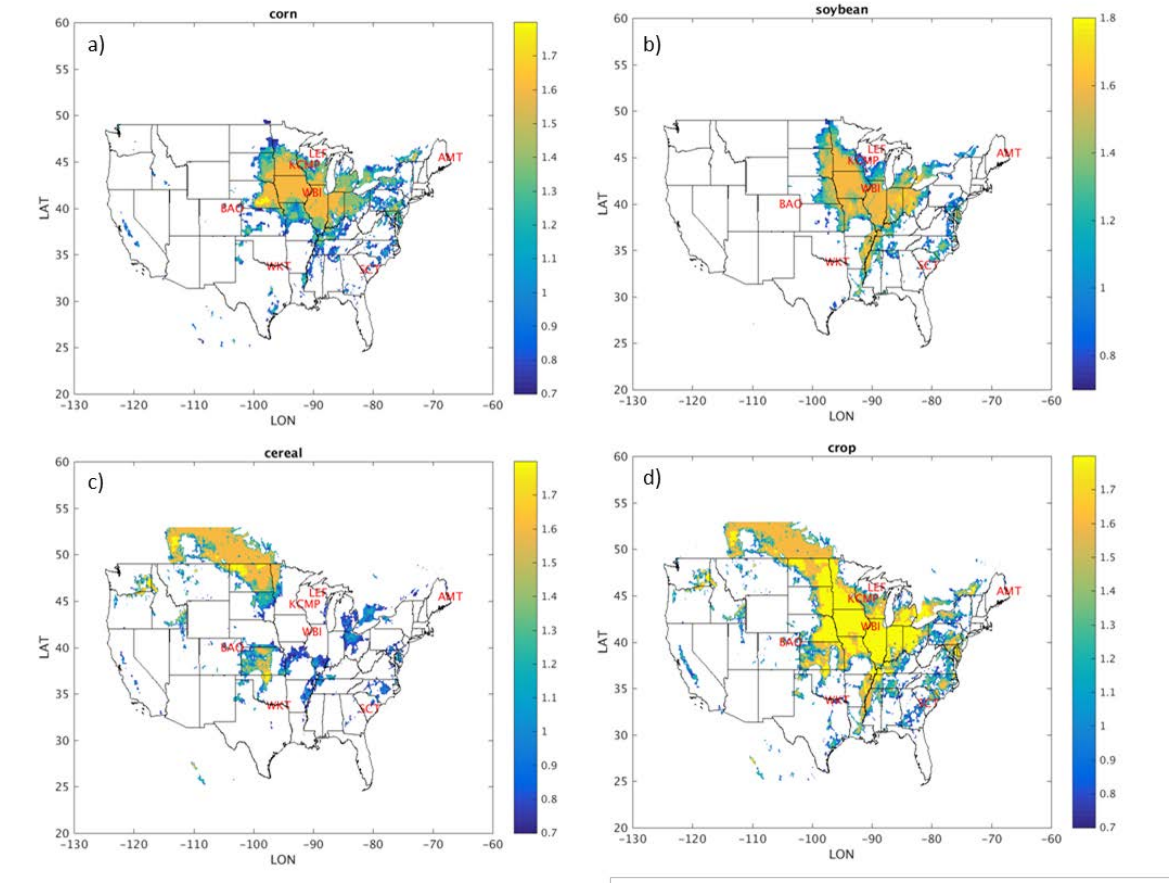


Figure S3.7. Crop Plant Function Types (PFTs) within the US.

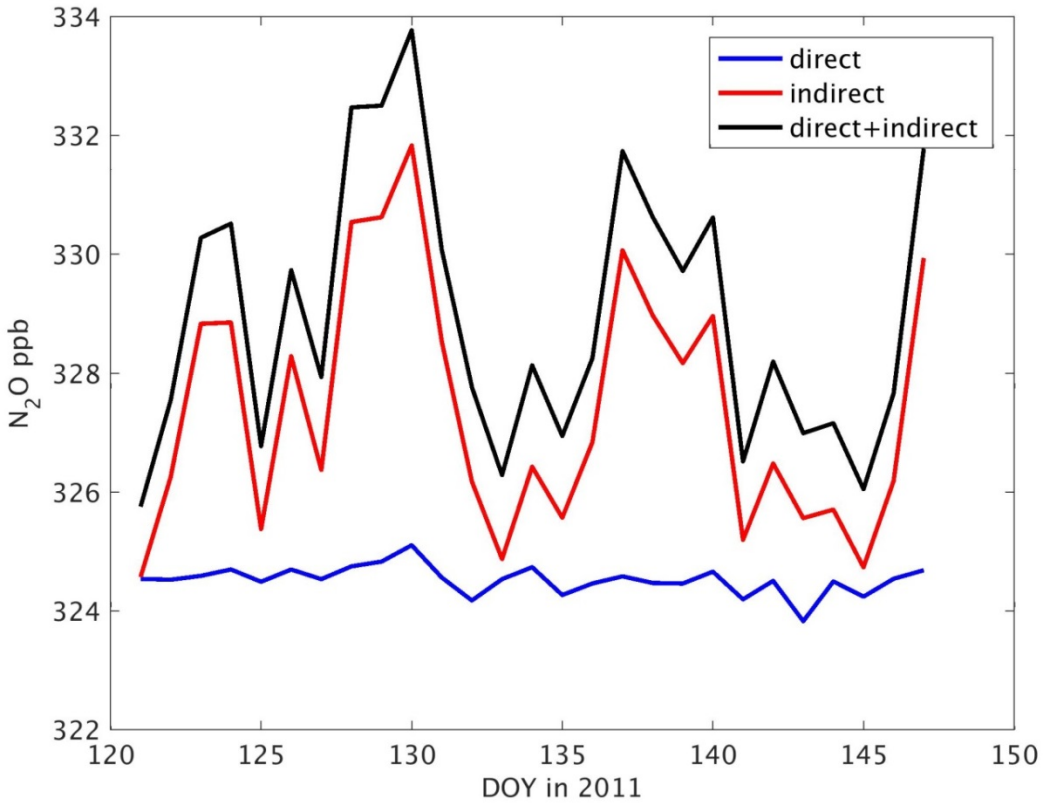


Figure S3.8. Simulated N_2O mixing ratios at the KCMP tall tower in May, 2011, driven by (a) direct source only, (b) indirect source only, and (c) direct + indirect sources in an otherwise identical fashion. The direct and indirect contributed to 1.2-2.9 and 2.2-9.7 ppb to the KCMP tall tower, respectively.

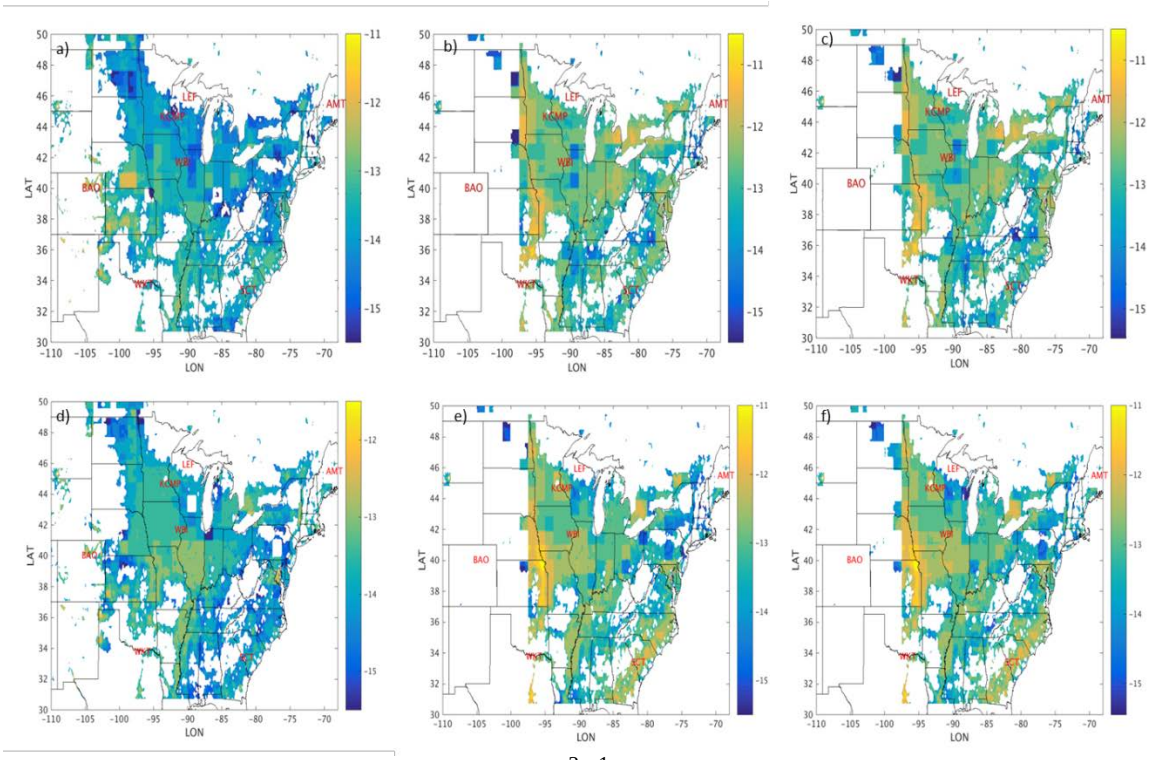


Figure S3.9. Increasing slope (unit of kg $m^{-2}s^{-1}$ per year) for a) direct, b) indirect, c) total emissions over a 40-yr period under the RCP2.6 scenarios, and increasing slope (unit of kg $m^{-2}s^{-1}$ per year) for d) direct, e) indirect, f) total emissions over a 40-yr period under the RCP8.5 scenarios.

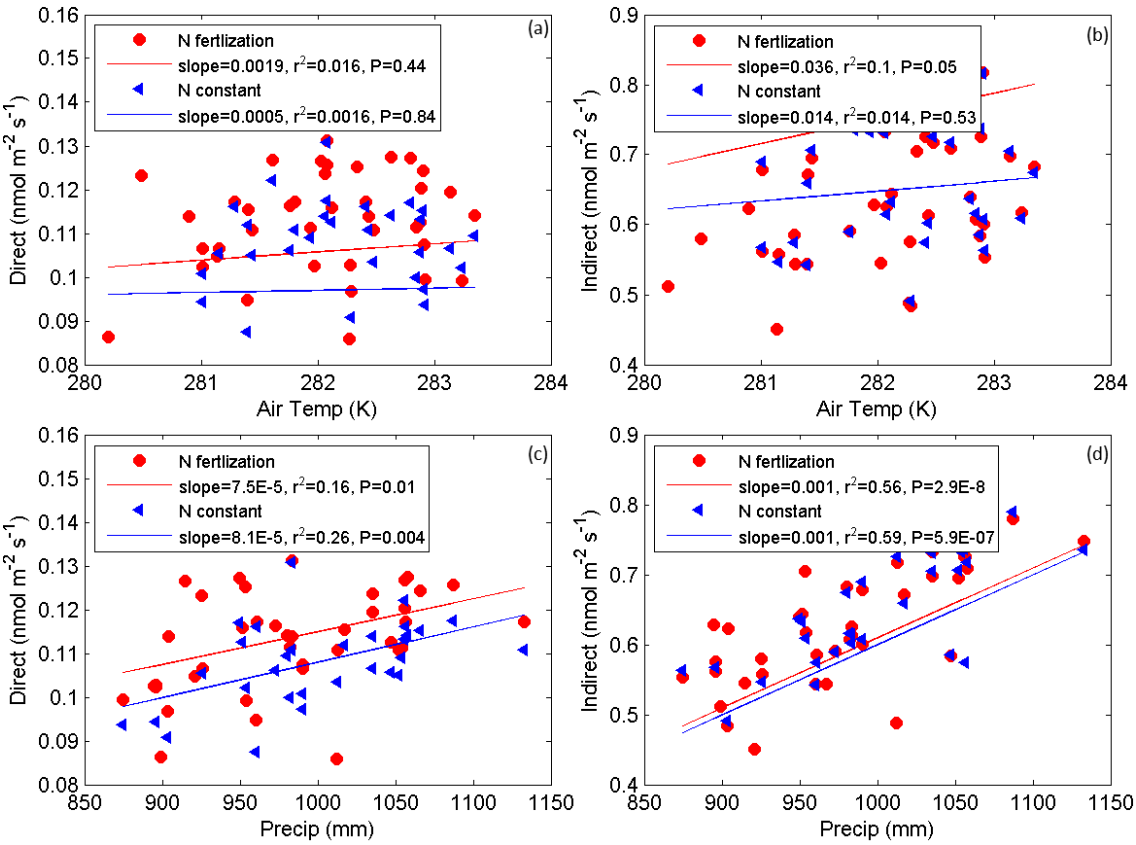


Figure S3.10. Direct emissions as a function of a) air temperature and c) precipitation under RCP85 scenario; indirect emissions as a function of c) air temperature and d) precipitation under RCP 8.5 scenarios. Direct and indirect emissions, air temperature and precipitation are annual averaged values across the Corn Belt.

Chapter 4
Source partitioning of methane emissions and its seasonality in the U.S. Midwest

Chen, Z., Griffis, T. J., Baker, J. M., Millet, D. B., Wood, J. D., Dlugokencky, E. J., et al. (2018). Source partitioning of methane emissions and its seasonality in the U.S. Midwest. *Journal of Geophysical Research: Biogeosciences*, 123, 646–659.

https://doi.org/10.1002/ 2017JG004356

Synopsis

The methane (CH_4) budget and its source partitioning are poorly constrained in the Midwestern United States. We used tall tower (185 m) aerodynamic flux measurements and atmospheric scale factor Bayesian inversions (SFBI) to constrain the monthly budget and to partition the total budget into natural (e.g. wetlands) and anthropogenic (e.g. livestock, waste, and natural gas) sources for the period June 2016 to September 2017. Aerodynamic flux observations indicated that the landscape was a CH_4 source with a mean annual CH_4 flux of +13.7 ± 0.34 nmol m^{-2} s^{-1} and was rarely a net sink. The SFBI analyses revealed a mean annual source of +12.3 ± 2.1 nmol m^{-2} s^{-1}. Flux partitioning revealed that the anthropogenic source (7.8 ±1.6 Tg CH_4 yr^{-1}) was 1.5 times greater than the bottom-up gridded EPA inventory, in which livestock and oil/gas sources were underestimated by 1.8- and 1.3-fold, respectively. Wetland emissions (4.0 ± 1.2 Tg CH_4 yr^{-1}) were the second largest source, accounting for 34% of the total budget. The temporal variability of total CH_4 emissions was dominated by wetlands with peak emissions occurring in August. In contrast, emissions from oil/gas and other anthropogenic sources showed relatively weak seasonality.

Introduction

The Midwestern United States is one of the most intensively managed agricultural regions in the world. This landscape is dominated by corn and soybean ecosystems that help support a livestock population of approximately 728 million animals within the US Corn Belt [USDA NASS, 2014]. This region is also one of the most wetland-rich landscapes across the United States [US Forest Service, 2016]. In addition, it includes a variety of urban and industrial complexes and major oil refineries. Short-term methane

(CH_4) measurements from a very tall tower within the region indicated that it is an important source of CH_4 [*Zhang et al*., 2014a]. However, the CH_4 budget, its source partitioning, and seasonality remain poorly constrained.

Recent space-based measurements imply large uncertainties in the United States anthropogenic CH_4 budget [*Kort et al*., 2014; *Wecht et al*., 2014; *Turner et al*., 2015; *Jacob et al*., 2016]. Bottom-up inventory emission databases, such as EDGAR42 [Emission Database for Global Atmospheric Research, version 4.2, 2011, http://edgar.jrc.ec.europa.eu], show that enteric fermentation, natural gas production, and manure management represent the top three anthropogenic CH_4 sources for the region. *Miller et al.* [2013] combined an atmospheric transport model and a geostatistical inverse modeling approach to estimate anthropogenic CH_4 emissions across the US for 2007 and 2008. They found that the CH_4 budget in the United States was underestimated by 1.5 and 1.7 times in the EPA [U.S. Environmental Protection Agency, https://www.epa.gov/] and EDGAR42 inventories, respectively. This was largely attributed to underestimates of the livestock and natural gas emissions. Furthermore, *Bruhwiler et al*. [2017] also highlighted the large uncertainties of regional to continental scale CH_4 emissions from using space-based observations, resulting from atmospheric transport variability, satellite sampling bias, and the choice of upwind background CH_4 concentration.

With the EDGAR42 bottom-up inventory used as *a priori*, *Wecht et al.* [2014] estimated North American CH_4 emissions at high spatial resolution by inversions of SCIAMACHY satellite observations using the GEOS-Chem chemical transport model and its adjoint.

They suggested that US livestock emissions were 40% greater than EDGAR42 and EPA inventories and that the low bias was associated with large underestimates in Iowa and Southern Minnesota. They also found that emissions associated with oil and gas were reasonably well-constrained by EPA. *Turner et al* .[2015] used 3 years of GOSAT satellite retrievals of column averaged CH_4 mole fraction (2009-2011) to constrain North American CH_4 emissions with high spatial resolution, with an inversion based on the GEOS-Chem adjoint. Their estimate of CH_4 emissions across the US was ~1.7 and ~1.5 times larger than the EDGAR4.2 and US EPA national inventories. They attributed the bias to oil/gas and livestock emissions, but they were unable to quantitatively separate the two, owing to their spatial overlap and limited observational coverage.

Wetland emissions remain one of the key sources of uncertainty in the regional to global atmospheric CH_4 budget, largely due to poor understanding of the biophysical processes controlling production and consumption in saturated soils [*Nisbet et al*., 2014; *Bloom et al*., 2017]. Within the US Corn Belt, *Zhang et al*. [2014a] used tall tower measurements to infer a CH_4 budget that was 5.8 times greater than EDGAR42, and hypothesized that the difference could be attributed to wetland emissions. A geostatistical inverse modeling study of *Miller et al*. [2014] indicated a large emission underestimate by the existing inventory (e.g., Kaplan model [*Kaplan*, 2002]) for Minnesota and Wisconsin wetlands. Recent inverse modeling studies using satellite observations also support major wetland CH_4 emissions for Florida as well as the Midwest [*Wecht et al*., 2014; *Turner et al*., 2015].

At finer spatial scales within the region, recent studies have quantified CH_4 fluxes from wetlands, agricultural crops, and livestock facilities. *Olson et al*.[2013] found that CH_4 emissions from a temperate peatland in north-central Minnesota showed instantaneous fluxes reaching a maximum of 290 nmol m^{-2} s^{-1} in August, with an annual budget ranging from 15.7 to 33.2 g CH_4 m^{-2} y^{-1} over the period 2009 to 2011. The instantaneous fluxes and annual budget were very sensitive to peat temperature and water table position. In contrast, CH_4 fluxes from individual corn and soybean plants were extremely small, ranging from about +0.4 nmol m^{-2} s^{-1} during the day to about -0.8 nmol m^{-2} s^{-1} during the night [*Zhang et al.*, 2014b]. *Bavin et al*. [2009] also showed that soil CH_4 fluxes from conventional and reduced tillage corn-soybean rotations were most often below the flux detection limit of static chambers.

There is a paucity of studies that have examined enteric CH_4 emissions and emissions related to manure management from within the region. Methane emissions have been reported for dairy manure storage facilities in Wisconsin (a simple storage basin) and Indiana (a storage lagoon, in which solids had previously been removed) [*Grant et al.*, 2015]. On a per animal basis, CH_4 emissions were larger from the storage basin than the lagoon with mean daily emissions of 295 g CH_4 $head^{-1}$ d^{-1} (374 g AU^{-1} d^{-1}) and 47 g CH_4 $head^{-1}$ d^{-1} (59 g AU^{-1} d^{-1}), respectively (AU represents Animal Unit, where 1 AU = 500 kg live weight). These daily emissions were shown to follow a positive linear relation with temperature. The relative lower emissions from the lagoon storage facility were attributed to the lower availability of carbon due to removal of solids prior to storage. The

area-based flux estimates were 94 kg CH_4 m^{-2} y^{-1} and 80 kg CH_4 m^{-2} y^{-1}, respectively and were estimated to be less than the enteric emissions from these farms [*Grant et al.*, 2015].

Enteric emissions from dairy cows can show substantial variability depending on size, sex, growth stage, activity (i.e. lactating vs dry), and diet [*Lassey*, 2007]. Because of limited studies conducted for the US Midwest, we draw on some recent studies from outside of the region. On-farm flux estimates, based on a backward Lagrangian stochastic technique, showed that enteric emissions ranged from 270 to 380 g (lactating cow d)$^{-1}$ for two dairy farms in Ontario, Canada, and accounted for about 40% of the total CH_4 farm emissions [*Vanderzaag et al.*, 2014]. Recent studies have shown that enteric emissions from dairy cows can be significantly reduced by using inhibitors designed to improve carbon use efficiency. *Hristov et al.* [2015] showed that enteric emissions varied from 400 to 500 g CH_4 d^{-1} and 290 to 390 g CH_4 d^{-1} over a 12-week period without/with the use of an inhibitor (3NOP), respectively.

Here, we build on our previous work [*Olson et al.*, 2013; *Zhang et al.*, 2014a, 2014b], and use tall tower aerodynamic flux measurements and SFBI analyses to obtain improved constraints on the CH_4 emissions from natural and anthropogenic sources within the region. The objectives were to: 1) Estimate the regional CH_4 budget; 2) Partition the emissions into natural and anthropogenic sources; and 3) Identify sources and time periods associated with high CH_4 emissions to improve our understanding of potential mitigation options.

Methods

Study site

The University of Minnesota tall tower trace gas observatory (KCMP tall tower; 244 m height) is located 25 km south of the Saint Paul – Minneapolis Metropolitan area in the Midwestern United States. The landscape is highly heterogeneous with agricultural lands predominant to the east, south, and west. To the north and northwest, there is a strong gradient from agricultural land to a relatively dense urban landscape. Detailed land use statistics as a function of distance and direction from the tall tower study site have been reported previously [*Griffis et al.,2013; Chen et al., 2016*]. The study domain contains 10 states in the US Midwest including Minnesota, Illinois, Indiana, Iowa, Missouri, Ohio, North Dakota, South Dakota, Nebraska, and Wisconsin.

Methane mixing ratio observations

The KCMP tall tower has been instrumented with meteorological and trace gas sensors since April 2007. Carbon dioxide, water vapor, and nitrous oxide have been measured at sample heights of 32, 56, 100, and 185 m [*Griffis et al.*, 2013, 2017; *Zhang et al.*, 2014a]. Turbulence is measured at 100 and 185 m using sonic-anemometers (model CSAT3; Campbell Scientific Inc. Logan, Utah). Near-continuous measurement of CH_4 mixing ratios was initiated in May 2016 using a Trace Gas Analyzer (TGA200A, Campbell Scientific Inc., Logan, Utah, USA). This system uses a state-of-the-art inter-band cascade laser that is thermoelectrically cooled to a temperature of 17°C. The manufacturer's reported measurement noise, based on the Allan variance, is 7.0 nmol mol^{-1} for an integration period of 100 ms. Our field tests indicated that the noise is approximately 9.5 nmol mol^{-1} for an integration period of 60-min (Figure 4.1). As shown later, this measurement noise is extremely small compared to the temporal fluctuations in the tall

tower observations. The TGA is calibrated hourly against an ultra-pure zero air standard and a working CH_4 span gas that was propagated from our Earth System Research Laboratory (NOAA ESRL) gold standard (Cylinder ID: CB11952, mole fraction = 1849.9 nmol mol^{-1}, reproducibility = 1 nmol mol^{-1}, WMO-CH_4-X2004A Calibration scale). Sample air is pulled continuously from inlets at 185 and 3 m at a total flow rate of approximately 15 SLPM at approximately 50 kPa. The 185 and 3 m sample inlets are subsampled at 30-s intervals. The air samples are dehumidified using a Nafion dryer system prior to analysis.

Flux-gradient method

The CH_4 mixing ratio gradients and fluxes provide information related to the landscape (i.e. mesoscale; ~100's of km^2) footprint, while the SFBI analyses constrain sinks and sources at the regional scale (~1000's of km^2) [*Zhang et al*., 2014a].

The aerodynamic flux was estimated from,

$$F_C = -K_C m_a \frac{dC}{dz} + F_s \qquad (4.1)$$

where the eddy diffusivity (K_C) was estimated from the momentum flux measured using eddy covariance at a height of 185 m [*Wood et al.,* 2016], m_a is the molar density of dry air, dC/dz is the CH_4 mixing ratio gradient, and F_s is an estimate of the change in hourly CH_4 storage integrated over a height of 185 m. An enhancement factor of 1.35 was applied to K_c to account for the difference between momentum and a trace gas scalar diffusivity [*Simpson et al.,* 1998].

Inverse modeling framework

Here we apply a Bayesian inversion to interpret the tall tower observations in terms of a constraint on the regional CH_4 sources in the US Midwest. The Stochastic Time-Inverted Lagrangian Transport (STILT) model [*Gerbig et al.*, 2003*; Lin et al*., 2003, 2004] was used to estimate the tall tower concentration source footprint using the sample inlet (receptor) at a height of 185 m and atmospheric drivers (e.g., planetary boundary layer (PBL) height, wind, atmospheric stability) obtained from the Weather Research and Forecasting (WRF) model version 3.8.1 [*Nehrkorn et al*., 2010]. The source footprint was then multiplied by a bottom-up emission inventory (*a priori* sources) and combined with the background mixing ratios to provide an initial estimate of the hourly CH_4 mixing ratios at the tall tower receptor.

Prior information

The Bayesian inverse analysis relies heavily on the prior spatial distribution of associated source categories, and it is therefore critical to use emission inventory data that have high spatial accuracy for proper attribution. EDGAR42 is one of the most comprehensive bottom-up inventories, but has been reported to show poor spatial accuracy in regional CH_4 emissions, particularly from the livestock and oil/gas sectors [*Miller et al.*, 2013; *Wecht et al*., 2014; *Turner et al*., 2015]. In EDGAR42, the oil/gas emissions are too heavily weighted by the spatial distribution and usage rather than the production [*Miller et al*., 2013; *Maasakkers et al*., 2016]. Further, the oil/gas sector has a strong correlation ($R^2 > 0.8$) with waste emissions because both are largely distributed according to human population. Therefore, a Bayesian inversion using EDGAR42 as the *a priori* estimate could wrongly attribute sources (e.g., assign CH_4 emissions from oil/gas production sites

to livestock). For these reasons, we decided to use the gridded EPA inventory [*Maasakkers et al.,* 2016] to represent the *a priori* anthropogenic sources.

The EPA inventory is available only as national totals for different source types [EPA, 2016]. *Maasakkers et al.* [2016] used a range of databases at the state to local source levels to disaggregate the inventory and allocated the spatial distribution of emissions for individual source types, and presented a gridded inventory of US anthropogenic CH_4 emissions with $0.1° \times 0.1°$ spatial resolution and detailed scale-dependent error characterization. Their estimate showed a significant spatial difference compared with EDGAR42, particularly for oil/gas systems and manure management. The gridded EPA inventory for the year 2012 placed higher emissions over oil/gas production areas and lower emissions over distribution areas, consistent with recent top-down constraints across the US [*Miller et al*., 2013; *Turner et al*., 2015].

Initially, (1) enteric fermentation, (2) manure management, (3) natural gas; (4) oil; (5) coal mining; (6) waste; (7) natural wetlands; and (8) others (i.e., forest fire hotspots and stationary combustion emissions), were adopted here as *a priori* source categories. However, spatially overlapping characteristics of enteric fermentation and manure management, as well as natural gas and oil, respectively, placed strong limitations on identifying and separating them individually within the inversion framework [*Turner and Jacob*, 2015]. Therefore, we combined enteric fermentation and manure management as **livestock** and combined natural gas and oil as **gas/oil.** Furthermore, *Bloom et al.* [2017] provided a full (2009-2010) and extended (2001-2015) estimate of wetland emissions, based on knowledge of regional to global wetland CH_4 sources and its biophysical

controls. We adopted the full estimate of wetland emissions with a gridded resolution of 0.1° × 0.1°, which was developed based on satellite-derived surface water content and precipitation reanalyses, as well as environmental parameterizations.

Other sources such as forest fire hotspots and stationary combustion were insignificant for the Midwest and not feasible to independently constrain from atmospheric data. Therefore, these sources were not included in the *a priori* estimate. The 5 source categories included in the *a priori* emission inventory were [*Maasakkers et al*., 2016; *Bloom et al*., 2017]:

(1) **livestock**, including enteric fermentation, and manure management;

(2) **gas/oil**, including natural gas production, processing, transmission, and distribution, as well as petrochemical production;

(3) **waste**, including landfills and wastewater treatment;

(4) **coal mining**, both surface and underground;

(5) monthly **natural wetlands**, and peatlands from *Bloom et al.* [2017].

We define the *a priori* errors for livestock, gas/oil, waste, coal mining sources at $0.1^{o}\times 0.1^{o}$ resolution based on the corresponding scale-dependent errors developed by *Maasakkers et al*. [2016] for the gridded EPA inventory. *Bloom et al*. [2017] derived 324 ensemble models of wetland CH_4 emissions in 2009 and 2010, based on 3 CH_4: C temperature dependencies (i.e., the temperature dependence of the ratio of C respired as CH_4), 9 heterotrophic respiration configurations, 4 wetland extent scenarios and 3 global scaling factor (i.e., a global budget estimate from wetland emissions) configurations.

Here, we used the ensemble mean of 324 model realizations for the year 2010 as the *a priori* estimate, and its coefficient of variation (CV) as the associated *a priori* error. Based on the derived *a priori* errors (i.e., 79 – 95%) we rounded them to 100% for each source category.

Concentration footprint functions

Concentration source footprint functions were determined based on the STILT model. For each hour between June 2016 and September 2017, we released 500 particles from the KCMP tall tower at a height of 185 m and transported them backwards for 7 days to ensure that the trajectories adequately represented source contributions from within the U.S. Furthermore, we used observations (Table 4.1) from the NOAA Carbon Cycle and Greenhouse Gases (CCGG) program near the outer edge of the source footprint to represent the background CH_4 mixing ratios. These observations are from discrete air samples collected approximately weekly in flasks [*Dlugokencky et al*., 2009, 2011], and are zonally and monthly averaged at 4° latitudinal resolution.

Examples of the concentration footprint functions during years 2016 and 2017 at the tall tower are shown in Figure 4.3a-c. For each month, based on the *a priori* inventory, we calculated the cumulative contribution of surface emissions to the tall tower mixing ratio at a range of concentration footprint scales and the corresponding distance from the tall tower (Figure 4.3d).These analyses indicated that areas where the footprint strength was greater than 1e-4 ppm μmol^{-1} m^2 s, contributed significantly to the tall tower mixing ratios (i.e. they accounted for about 75% of the total contribution). These intense areas contain Minnesota and a significant portion of the US Midwest. The tall tower

observations, therefore, are representative of the larger Midwest region [*Griffis et al*., 2013; *Zhang et al*., 2014a], and should provide a reasonably robust estimate of CH_4 emissions and their partitioning *via* the SFBI approach.

Methane budget and partitioning

Here we used the SFBI method to constrain the regional budget and to partition it into its source contributions [*Kim et al*., 2013; *Chen et al*., 2016]. The SFBI method was applied monthly from June 2016 to September 2017,

$$\mathbf{y} = \mathbf{K\Gamma} + \boldsymbol{\varepsilon} \tag{4.2}$$

where **y** is the mixing ratio observed at the tall tower minus the background mixing ratios, **Γ** contains the scaling factors for different source types, **K** is the Jacobian matrix representing the sensitivity of the observation variables to the specific source types, and **ε** is the system error, which consists of measurement uncertainties and model uncertainties. The columns of **K** correspond to the mixing ratios for each of the source types being optimized, and **Γ** consists of the *a posteriori* scale factors for the different source types. Applying Bayes' theorem, along with a normal distribution assumption, the *maximum a posteriori* (MAP) solution of **Γ** *is* to minimize the cost function **J(Γ)**:

$$2\mathbf{J}(\mathbf{\Gamma}) = (\mathbf{y} - \mathbf{K\Gamma})^{\mathbf{T}}\mathbf{S}_{\varepsilon}^{-1}(\mathbf{y} - \mathbf{K\Gamma}) + (\mathbf{\Gamma} - \mathbf{\Gamma}_{\mathbf{a}})^{\mathbf{T}}\mathbf{S}_{\mathbf{a}}^{-1}(\mathbf{\Gamma} - \mathbf{\Gamma}_{\mathbf{a}}) \tag{4.3}$$

where S_ε and S_a are the observational and *a priori* error covariance matrices, and each element of $\mathbf{\Gamma_a} = 1$. The solution to $\nabla\mathbf{\Gamma J(\Gamma)} = 0$ is then given by:

$$\mathbf{\Gamma}_{\mathbf{post}} = (\mathbf{K}^{\mathbf{T}}\mathbf{S}_{\varepsilon}^{-1}\mathbf{K} + \mathbf{S}_{\mathbf{a}}^{-1})^{-1}(\mathbf{K}^{\mathbf{T}}\mathbf{S}_{\varepsilon}^{-1}\mathbf{y} + \mathbf{S}_{\mathbf{a}}^{-1}\mathbf{\Gamma}_{\mathbf{a}}) \tag{4.4}$$

Here, observational errors include measurement and modelling uncertainties. The mixing ratio measurement error for the TDL was based on the Allan variance test described above (9.5 nmol mol^{-1} for an integration period of 60-min). Further, *Gerbig et al* [2003]

conducted a wide range of sensitivity tests by varying the released number of particles from 50 to 1000 in the STILT model, and each test was run 100 times to analyze the particle number dependence and reproducibility. Following the work of *Gerbig et al*. [2003], a relative uncertainty of 13% was assigned for the number of particles used in the backward trajectories. In this study, the transport error associated with the number of particles used ranged from ~6-9 ppb. We used different PBL schemes, i.e., Yonsei University [*Hong et al*., 2006] and Mellor-Yamada-Janjic [*Janjic*, 2002] to simulate the mixing height from the WRF-STILT simulations. The modeled mixing heights were then compared to that inferred from the radiosonde observations [*Miller et al*., 2008; *Kim et al*., 2013] to derive the relative uncertainty associated with the simulated PBL height, which was estimated to be 27%. The transport error associated with the PBL height simulations ranged from ~10-16 ppb. Finally, it is critical to use mass conserving wind fields to drive the transport [*Gerbig et al*., 2003]. The WRF simulations used in this study are constructed specifically for mass conservation and have been found to perform better than other meteorological products [*Nehrkorn et al*., 2010; *Miller et al*., 2013]. A detailed uncertainty analysis associated with aggregation error is presented in the Results and Discussion section.

Results and discussion

Methane mixing ratios and fluxes

Mean CH_4 mixing ratios at 3 m and 185 m (June 01, 2016 to September 30, 2017) were 2024.4 ± 132.4 μmol mol^{-1} and 1982.0 ± 54.4 nmol mol^{-1}, respectively (Figure 4.4). Wind direction had a relatively weak influence on the CH_4 mixing ratios at 3 m (Figure 4.4). For example, advection from the urban airshed to the north resulted in CH_4 mixing ratios that were about 0.1 μmol mol^{-1} higher than the ensemble mean mixing ratio. There is also

a small but noticeable increase in CH_4 mixing ratio from the southwest wind sector, which is predominantly agricultural land use. Hourly CH_4 mixing ratios at the 185 m and 3 m levels were highly correlated ($r^2 = 0.976$, n = 5305, $p < 0.0001$). The large deviations from the 1:1 line (Figure 4.4d) were generally associated with weak turbulent mixing (i.e. when friction velocity was less than 0.1 m s^{-1}).

Vertical CH_4 gradients were very small in the winter (December through March) compared to the mid-summer (July through August, Figure 4.5a), indicating that within the aerodynamic flux footprint, microbial activity was the main driver of emissions (i.e. not fuel production or consumption). Further, there was a relatively low frequency of negative mixing ratio gradients, implying that the agricultural landscape was rarely a net CH_4 sink.

The mean aerodynamic CH_4 flux was +13.7 ± 0.34 (mean ± standard error) nmol m^{-2} s^{-1} (Figure 4.5), which equates to a net CH_4 source of about 6.9 g CH_4 m^{-2} y^{-1}, and is in good agreement with previous estimates at the landscape scale [*Zhang et al*., 2014b].

Bayesian analyses

After the first inversion, the averaging kernel (AK) was calculated to quantify the sensitivity of retrieved emissions to their true values [*Rodgers*, 2000; *Kim et al*., 2013; *Chen et al*., 2016]. The AK values indicated a very low sensitivity of tall tower observations to the coal mining source, probably due to the limited near-field distribution relative to the tall tower receptor (Figure 4.2c). Therefore, we eliminated the coal mining source term from further consideration and applied a second inversion where we only

included the livestock, oil/gas, waste and natural wetlands and peatlands source categories. This second inversion produced the final optimized emission estimates.

Figure 4.6a illustrates the linear regressions between the measured and simulated CH_4 mixing ratios based on the *a priori* and *a posteriori* estimates. The optimized CH_4 mixing ratios showed a much stronger correlation (R^2=0.72, and p=6.2e-5) than the *a priori* estimate (R^2=0.17, and p=0.21), and the slope increased from 0.29 ± 0.17 to 0.72 ± 0.12, indicating that the optimization helped to reduce the model bias and significantly improved the constraint on the emissions (Figure 4.6b). Based on the cost function analysis (equation (3)), the optimization is constrained to a certain degree by the error construction of observations and the a priori estimate, e.g., decreasing the observational errors or increasing the *a priori* errors leads to an increase in the optimized scaling factors, and acted to increase the slope of the linear regression between the modeled and observed CH_4 mixing ratios.

From the second inversion, there was a mean scaling factor of 1.5 relative to the gridded EPA inventory for anthropogenic sources (Table 4.2). Among them, optimized livestock and oil/gas sources were 1.8 and 1.4 times greater (Figure 4.7a), respectively, in close agreement with other recent studies [*Miller et al*., 2013; *Turner et al*., 2015]. The optimized mean annual anthropogenic CH_4 emission was 12.3 ± 2.1 nmol m^{-2} s^{-1} for the Midwest region. Compared to the source attribution for the national-scale budget, there was a higher contribution from livestock and lower contribution from oil/gas sources for the US Midwest, corresponding to its agriculture-dominated landuse characteristics.

Wetlands within our study domain are largely distributed in central and northern Minnesota as well as Wisconsin [*Cohen et al.,* 2016], among the most sensitive areas to the tall tower receptor based on the concentration footprint functions (Figure 4.3). The optimized wetland emissions are in excellent agreement with the estimate from *Bloom et al.* [2017] (Figure 4.7b), and represent the second largest source (4.0 Tg CH_4 yr^{-1}), accounting for 34% of total CH_4 emissions within the Midwest region (Table 4.2).

Compared to the CH_4 emission estimate at the national scale from the most recent space-based study [*Turner et al*., 2015], our work suggests that the natural, anthropogenic, and total CH_4 emissions in the Midwest account for 41.7 ± 12.5%, 18.2 ± 3.7%, and 22.5 ± 3.8% of the US budget, respectively (Figure 4.8). This further supports the importance of the US Midwest to the national and global CH_4 budget and highlights the need to understand how these emissions are likely to responds to changes in land management and climate.

Seasonality analysis

The temporal variability of total optimized CH_4 emissions was dominated by wetlands (Figure 4.9). Since emissions from oil/gas, livestock and waste showed relatively weak seasonality, the contribution from wetland emissions well explained the seasonal variation of the tall tower CH_4 mixing ratios. The wetland emissions were very weak from November to February, and began to rise after snowmelt in March, and peaked in August. This seasonality agrees closely with eddy covariance observations and retrospective modeling studies at Bog Lake Fen, Marcell Experiment Forest Station in northern Minnesota [*Olson et al.,* 2013].

To further explore the climate sensitivity of the regional methane emissions, the weighted means of selected environmental variables from the National American Regional Reanalysis [NARR, *Mesinger et al*., 2006] data were computed for the entire study domain. Here, the weighting is based on the intensity of the source footprint function for each grid cell. The seasonal and inter-annual variation of optimized wetland emissions appear to be consistent with variations in air temperature. We found that the peak CH_4 emissions in August was coincident with maximum source footprint-weighted soil temperature, which also agreed with earlier observations at the Marcell Experiment Forest Station , where CH_4 emissions peaked with peat temperature [*Olson et al*., 2013]. *Olson et al.* [2013] reported increased CH_4 emissions at Bog Lake Fen in warm years including 1994, 2001 and 2005, and suggested that temperature dominated the temporal variability, while water table depth (WTD) played a lesser role.

The notably higher temperature (1.5 ^{o}C within the growing season) in 2016 compared to 2017 likely enhanced wetland emissions. For instance, the weighted September air temperature was 2.3 ^{o}C higher in 2016, and the corresponding regional wetland emissions were 62% higher compared to 2017. Conversely, there appears to be very limited impact of precipitation differences on the wetland emissions. For example, June 2016 and September 2017 experienced dramatically different precipitation amounts (106.0 versus 24.7 mm, respectively), yet the regional wetland emissions (4.16 versus 4.22 nmo m^{-2} s^{-1}) were not statistically different.

Finally, it is possible that the warmer conditions stimulated higher CH_4 emissions from manure. For instance, the June and September air temperatures were 1.1 ^oC and 2.3 ^oC higher in 2016, and the corresponding regional livestock emissions were 56% and 60% higher than in 2017. Since enteric emissions are less dependent on air temperature (i.e. animal body temperature is regulated), stronger air temperature sensitivity for manure emissions is expected [*Sommer et al.,* 2004; *Dong et al., 2006*; *Wood et al.,* 2013].

Uncertainty analysis

To probe how aggregation error influenced the Bayesian inversion, we decreased the spatial resolution of the *a priori* emissions and meteorological fields from 0.1° × 0.1° to 0.5° × 0.5° for each source category. We then re-ran STILT to obtain the source footprint functions at 0.5° × 0.5° resolution and applied the SFBI in an otherwise identical fashion to our previously described approach. The low resolution inversion showed that when decreasing the resolution: (1) the correlation among source categories became stronger, e.g., the correlation coefficient (*R*) between livestock and oil/gas increased from 0.35 to 0.56; (2) the scale factor for the livestock source increased while that for the oil/gas source decreased. Further, we found that the AK for oil/gas source decreased from 0.89 to 0.55, indicating a weak sensitivity to the tall tower observations. As expected, averaging the inventory over a coarser grid caused spatial overlapping among source categories, leading to higher correlation among them. Using *a priori* information with a coarser resolution reduces the amount of independent information contained in the a priori inventory. The source(s) of interest cannot be separated and results in a lower sensitivity to the tall tower observations, providing a weak constraint within the inverse modeling framework.

Further, we examined the potential impact of having an upward/downward bias in the background CH_4 concentrations on the Bayesian inversion. Here we conducted a range of sensitivity tests by varying the monthly background mixing ratios (from mean - 2 SD to mean + 2 SD) based on the discrete flask-air observations from the NOAA CCGG program. These sensitivity analyses revealed that the annual budget ranged from 10.3–12.4 Tg CH_4 yr^{-1}, which showed only a minor variation (5% – 12%) compared to the best estimate (Table 4.3). These analyses suggest that the optimized annual budget and the source attribution show low sensitivity to the uncertainties in the background mixing ratios, and support that background CH_4 values are reasonably constructed and provide a reliable estimate for the Bayesian inversion framework.

Finally, by assembling key parameterizations in the models from *Bloom et al.* [2017], a range of *a priori* wetland emission scenarios were applied for the sensitivity tests (Table 4.4), to assess how well the inversion constrained the wetland source. The sensitivity studies showed similar seasonality, and an annual wetland budget that ranged from 3.41 to 4.38 Tg CH_4 yr^{-1}, or 86.1% to 110.7% of the best estimate (4.0 Tg CH_4 yr^{-1}). This supports a relatively robust constraint of the wetland source in the Bayesian inversion.

An important question is, to what extent can we improve the CH_4 budget estimate for the US Midwest? Currently there are two critical limitations, including the spatially sparse tall tower concentration observations and the lack of direct flux measurements of key land use categories. Building on previous inverse studies [*Turner et al*., 2015; *Chen et al.*,

2016; *Griffis et al*., 2017; *Michalak et al*., 2017], the following recommendations are made in order to help reduce regional scale uncertainties. There is a need for (1) improved high resolution information on activity data to better specify the *a priori* source distribution; (2) improved spatial representation of atmospheric CH_4 concentrations; and (3) direct measurement of CH_4 fluxes from underrepresented and important landuse categories. To address some of these concerns our ongoing research is making use of aircraft measurement campaigns to provide new insights regarding the spatial patterns of CH_4 concentrations during key times of the year. In light of the importance of enteric emissions for the region, we have planned intensive flux measurement campaigns to improve emission estimates from large representative livestock facilities within the region. Furthermore, the measurement of the stable isotopes of CH_4 at the tall tower has significant potential to help with source attribution partitioning, error reduction in the inversion, and could provide an efficient way to assess systematic biases in the atmospheric inversion methodology.

Conclusions

Hourly CH_4 observations from a tall tower in the Upper Midwest United States were used to constrain the CH_4 emissions from natural and anthropogenic sources for the region based on aerodynamic flux measurements and SFBI analyses. The data and analyses support the following conclusions:

1. The mean annual landscape and regional scale CH_4 emissions (13.7 ± 0.34 versus 12.3 ± 2.1 nmol m^{-2} s^{-1}) agreed reasonably well within the margin of uncertainty;

2. Regional natural, anthropogenic and total CH_4 emissions were 4.0 ± 1.2, 7.8 ± 1.6, and 11.8 ± 2.0 Tg CH_4 yr^{-1}, accounting for 41.7 ± 12.5%, 18.2 ± 3.7%, and 22.5 ± 3.8% of the overall US natural, anthropogenic and total CH_4 budgets, respectively;

3. Wetlands were the second largest regional source, contributing 34% to the regional budget;

4. The seasonality of total CH_4 emissions was dominated by wetlands. Wetland emissions increased significantly following snowmelt and reached a maximum in August;

5. The anthropogenic source categories were 1.5 times greater than the bottom-up inventory, with livestock and oil/gas sources underestimated by 1.8- and 1.3-fold, respectively.

Acknowledgements

This research was supported by the United States Department of Agriculture, USDA grant number 2013-67019-21364; United States National Science Foundation (grant number 1640337); USDA-ARS and NASA (grant number NNX17AK18G). We express our sincere thanks to Minnesota Public Radio and Tom Nelson for the logistical support at the KCMP tall tower. The Minnesota Supercomputing Institute provides key computing resources and assistance. Data are hosted at http://www.biometeorology.umn.edu/research/data-archives and ESS-DIVE . We acknowledge the support of an MnDrive PhD fellowship to Z.C.. JDW acknowledges support from the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research Program, through Oak Ridge National Laboratory's Terrestrial Ecosystem Science (TES) Science Focus Area (SFA). ORNL is managed by UT-Battelle, LLC, for the U.S. DOE under contract DE-AC05-00OR22725.

Table 4.1. Sites of flask CH_4 observations used in this study

NOAA code	Location	Country	Latitude	Longitude
AMT	Argyle, Maine	United States	45.035	-68.682
DND	Dahlen, North Dakota	United States	47.5	-99.24
ESP	Estevan Point, British Columbia	Canada	49.383	-126.544
ETL	East Trout Lake, Saskatchewan	Canada	54.35	-104.983
HIL	Homer, Illinois	United States	40.07	-87.91
LEF	Park Falls, Wisconsin	United States	45.945	-90.273
MWO	Mt. Wilson Observatory	United States	34.225	-118.059
NWR	Niwot Ridge, Colorado	United States	40.053	-105.586
SGP	Southern Great Plains, Oklahoma	United States	36.607	-97.489
THD	Trinidad Head, California	United States	41.054	-124.151
UTA	Wendover, Utah	United States	39.902	-113.718
WBI	West Branch, Iowa	United States	41.725	-91.353
WGC	Walnut Grove, California	United States	38.265	-121.491

Table 4.2. Annual mean a priori and a posteriori emissions and scaling factors for each source category (from October 2016 to September 2017)

Emissions ($Tg\ CH_4\ yr^{-1}$)	**Wetlands**	**Livestock**	**Oil/gas**	**Waste**	**Total**	**Anthropogenic**
a priori	3.8	2.6	1.2	1.3	8.9	5.1
a posteriori	4.0 ± 1.2	4.8 ± 1.5	1.6 ± 0.6	1.4 ± 0.6	11.8 ± 2.0	7.8 ± 1.6
Scale factor	1.1	1.8	1.3	1.1	1.3	1.5

Table 4.3. The a posteriori emissions with various sensitivity tests on background mixing ratios

Background mixing ratios	*a posteriori* emissions ($Tg\ CH_4\ yr^{-1}$)					
	Wetlands	**Livestock**	**Oil/gas**	**Waste**	**Total**	**Anthropogenic**
mean-2 SD	4.5	5.4	1.7	1.5	13.1	8.6
mean -1 SD	4.3	5.3	1.6	1.2	12.4	8.1
mean	4.0	4.8	1.6	1.4	11.8	7.8
mean + 1 SD	3.6	4.5	1.6	1.4	11.1	7.5
mean + 2 SD	3.3	4.2	1.4	1.4	10.3	7.0

Table 4.4. Wetland CH_4 model ensemble configurations

Tests	Global scale factor (Tg CH4 yr^{-1})	CH_4 : C temperature dependence (Q_{10})	Heterotrophic respiration	Wetland extent scenario	Wetland emissions (Tg CH_4 yr^{-1})
S1	124.5	1.0	Ensemble mean from the Multi-scale synthesis and Terrestrial Model Intercomparison Project (MsTMIP)[a,b]	GLOBCOVER spatial extent[c] and SWAMPS[d] inundation temporal variability parameterization	3.76
S2	124.5	2.0			3.74
S3	124.5	3.0			3.54
S4	166.0	1.0			3.63
S5	166.0	2.0			4.26
S6	166.0	3.0			3.41
S7	207.5	1.0			4.0
S8	207.5	2.0			4.38
S9	207.5	3.0			4.16

Note. [a]*Huntzinger et al*., 2013. [b]*Wei et al*., 2014. [c]*Bontemps et al*., 2011. [d]*Schroeder et al*., 2015

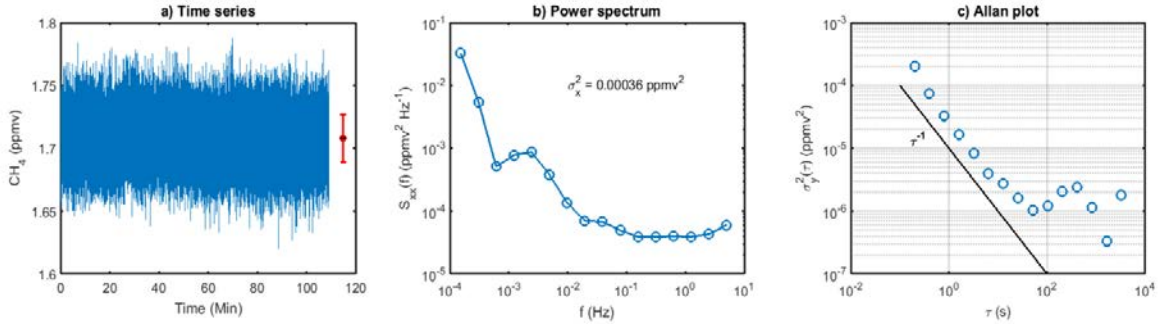


Figure 4.1. Results of sampling a standard gas cylinder including a) the 10 Hz time-series with the mean ± 1 SD (red symbols/bars) and mean ± 1 SE (black); b) the power spectrum of the concentration series (as spectral densities); and c) an Allan variance plot.

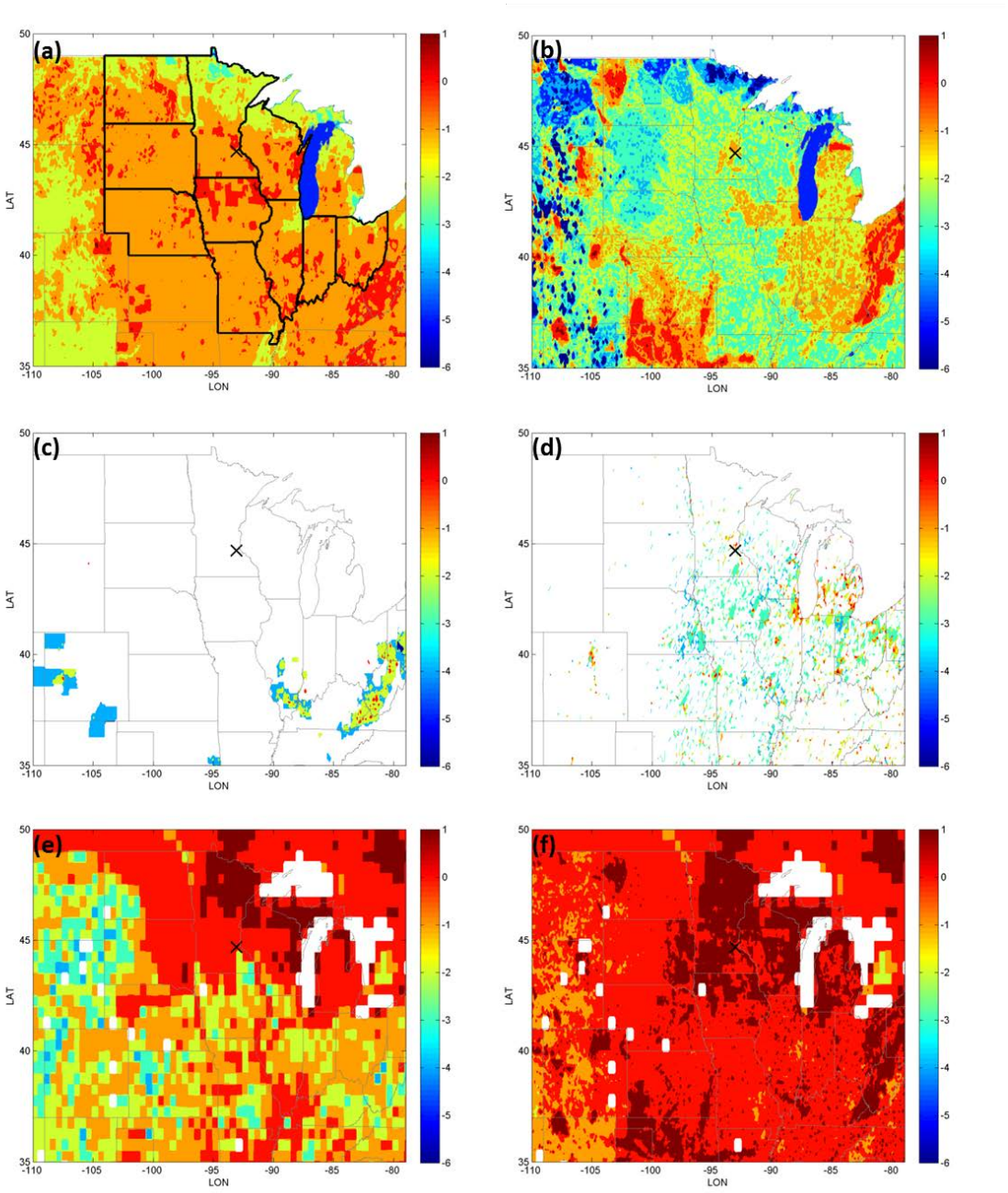


Figure 4.2. The *a priori* annual average emissions from (a) livestock; (b) natural gas +oil; (c) coal mining; (d) waste; (e) natural wetlands and (f) total (unit is $\log_{10}$(nmol m^{-2} s^{-1})). Anthropogenic source categories including livestock, oil/gas, coal mining, and waste were from *Maasakkers et al*. [2016], and the natural wetland emissions were derived from *Bloom et al*. [2017]. The KCMP tall tower is indicated by the black crosses in each panel.

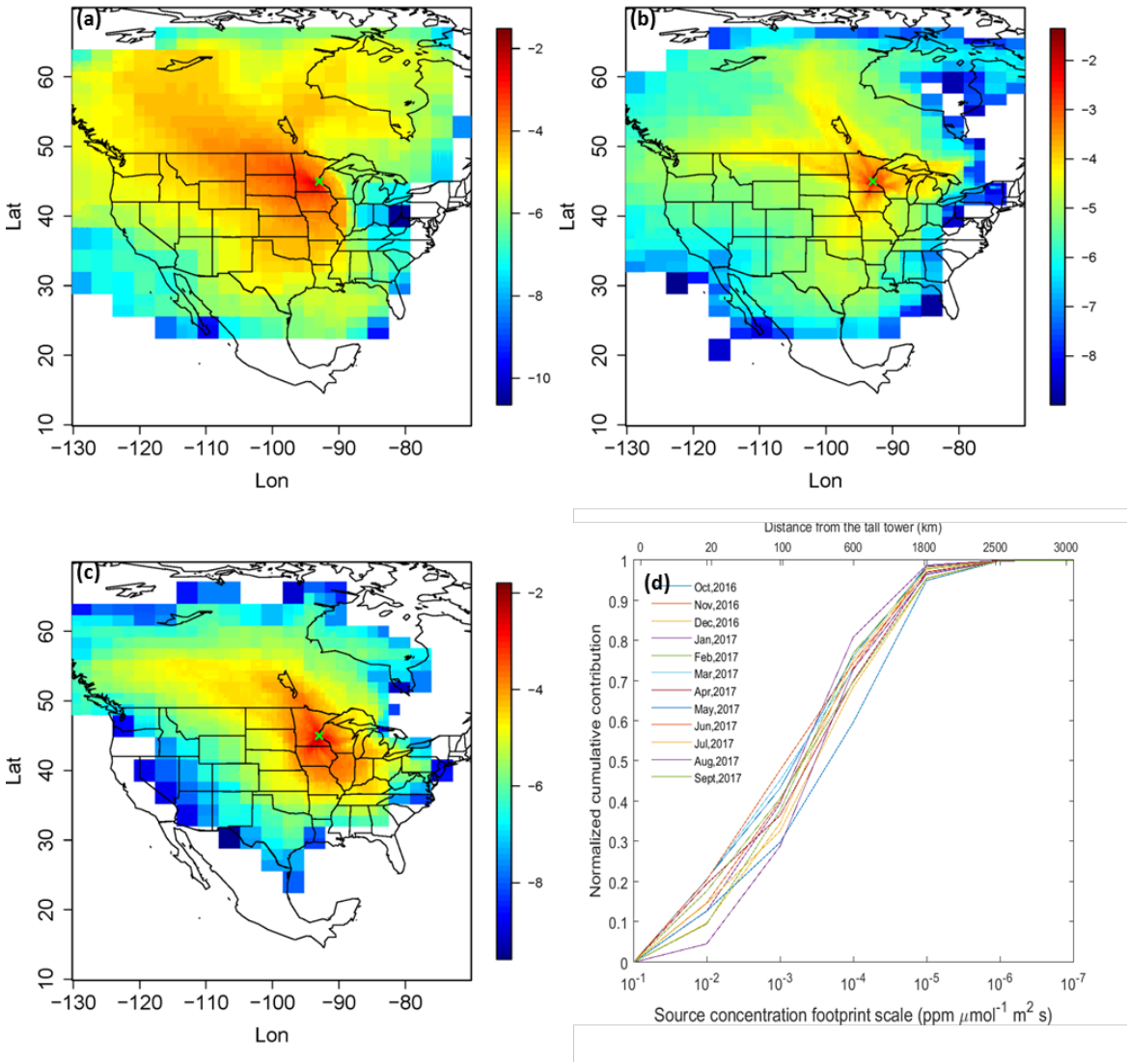


Figure 4.3. Source footprint functions (units: $\log_{10}$(ppm μmol^{-1} m^2 s)) for measurements at the KCMP tall tower (indicated by crosses). (a) December, 2016; (b) March. 2017; (c) August, 2017. (d) Normalized cumulative contribution as a function of source footprint scale (ppm μmol^{-1} m^2 s) and the corresponding distance from the tall tower (km).

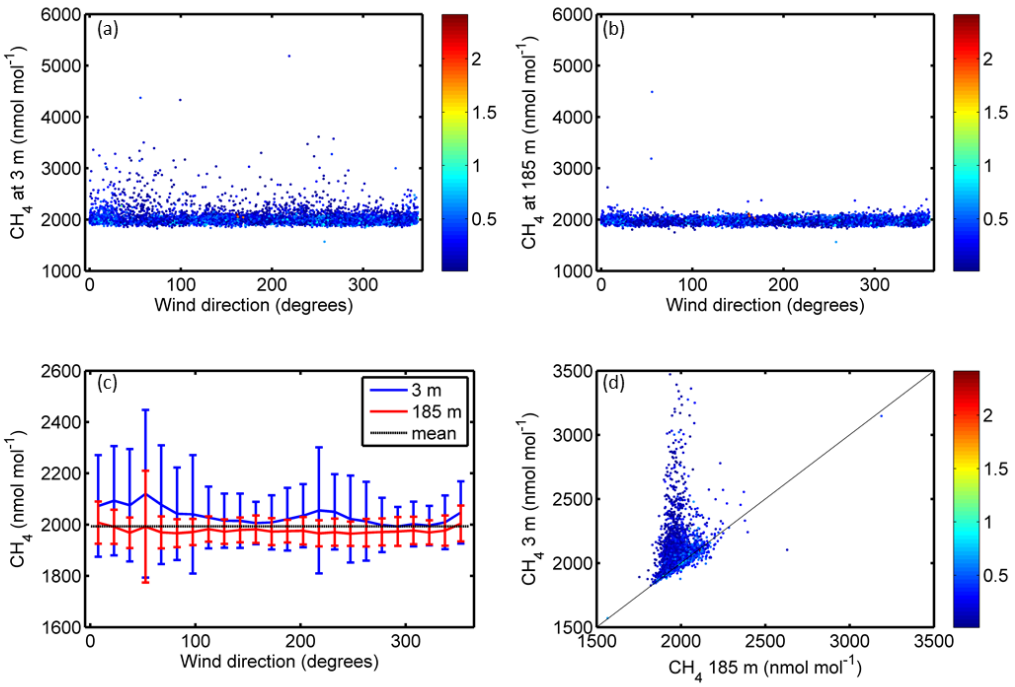


Figure 4.4. Hourly mean CH_4 mixing ratios measured at 3 m (panel a) and 185 m (panel b) as a function of wind direction. The color bar represents the friction velocity value (m s^{-1}) Panel c shows the mean methane mixing ratios (+/- 1 standard deviation) as a function of wind direction binned using 15 degree intervals. Panel d shows the relation between CH_4 mixing ratios measured at 3 m versus 185 m. The color bar indicates the friction velocity value. Each data point represents an hourly average value. Note that the 1:1 line is obscured by the best-fit linear regression line.

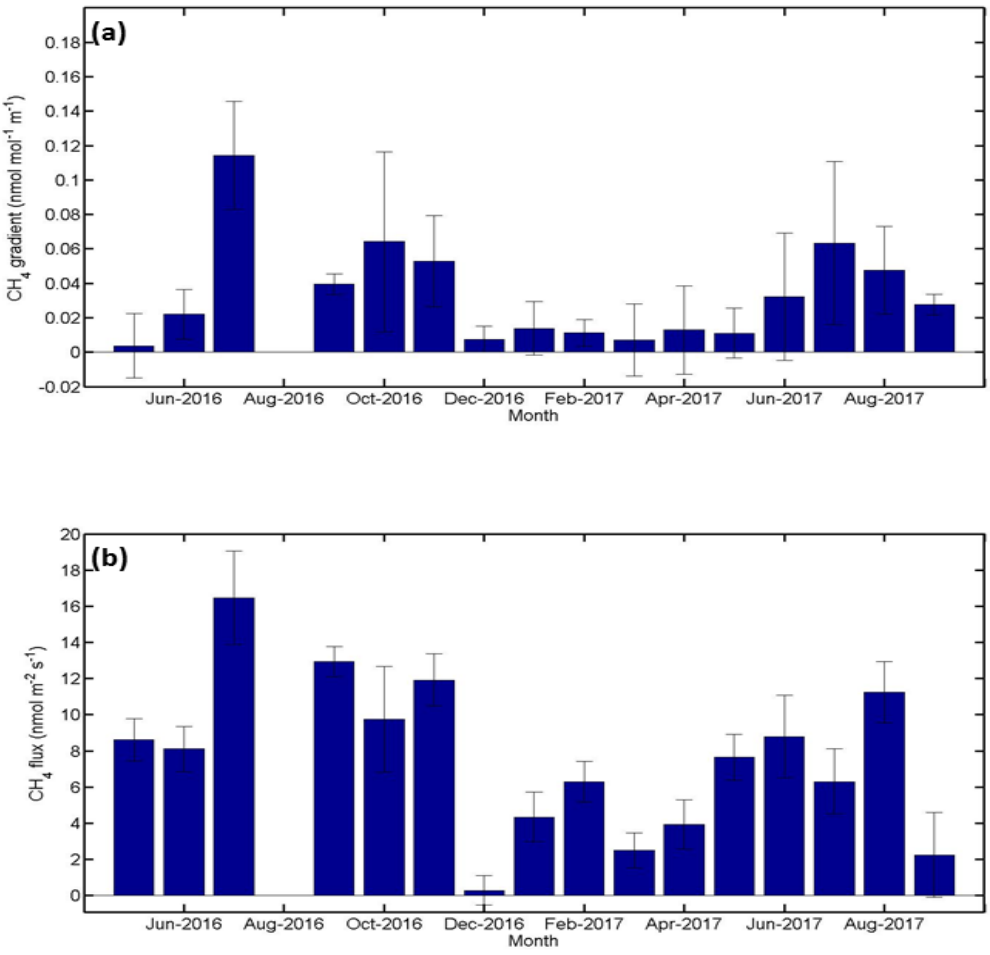
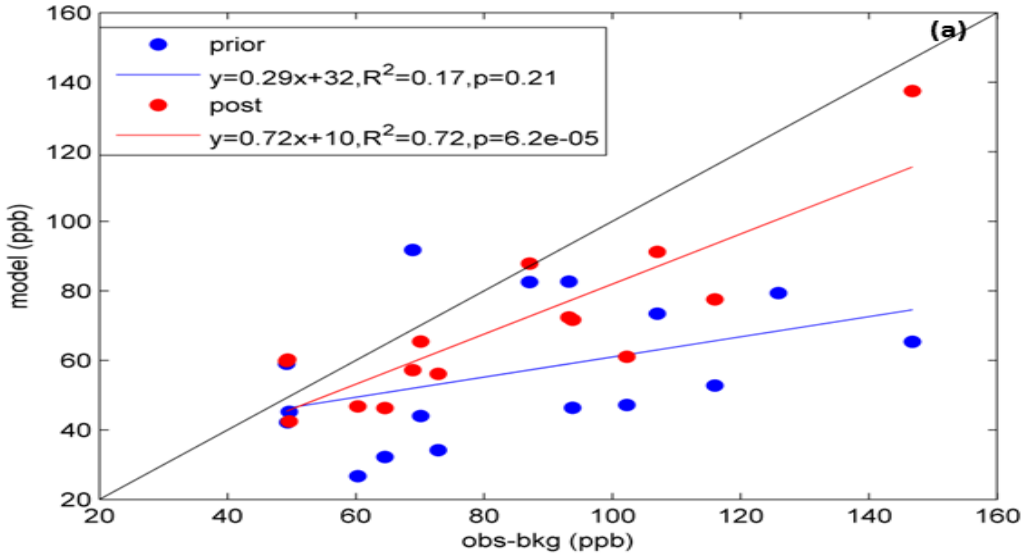


Figure 4.5. Median monthly vertical CH_4 mixing ratio gradient (Panel a), and median monthly flux-gradient estimate (Panel b). Error bars represent the standard error. Note that August, 2016 is not reported due to a large fraction of missing hourly observations.



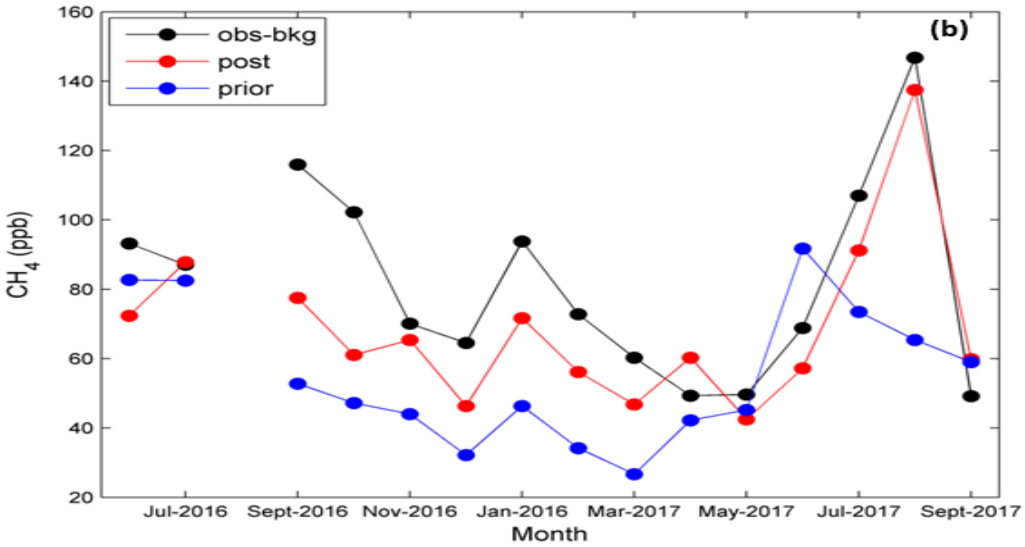


Figure 4.6 a). Monthly linear regression between the observed and simulated CH_4 mixing ratios for *a priori* and *a posteriori* simulations between June 2016 and September 2017 (Uncertainty values indicate a 95% confidence interval); b) *a priori*, *a posteriori* estimate and the tall tower observations of CH_4 mixing ratios.

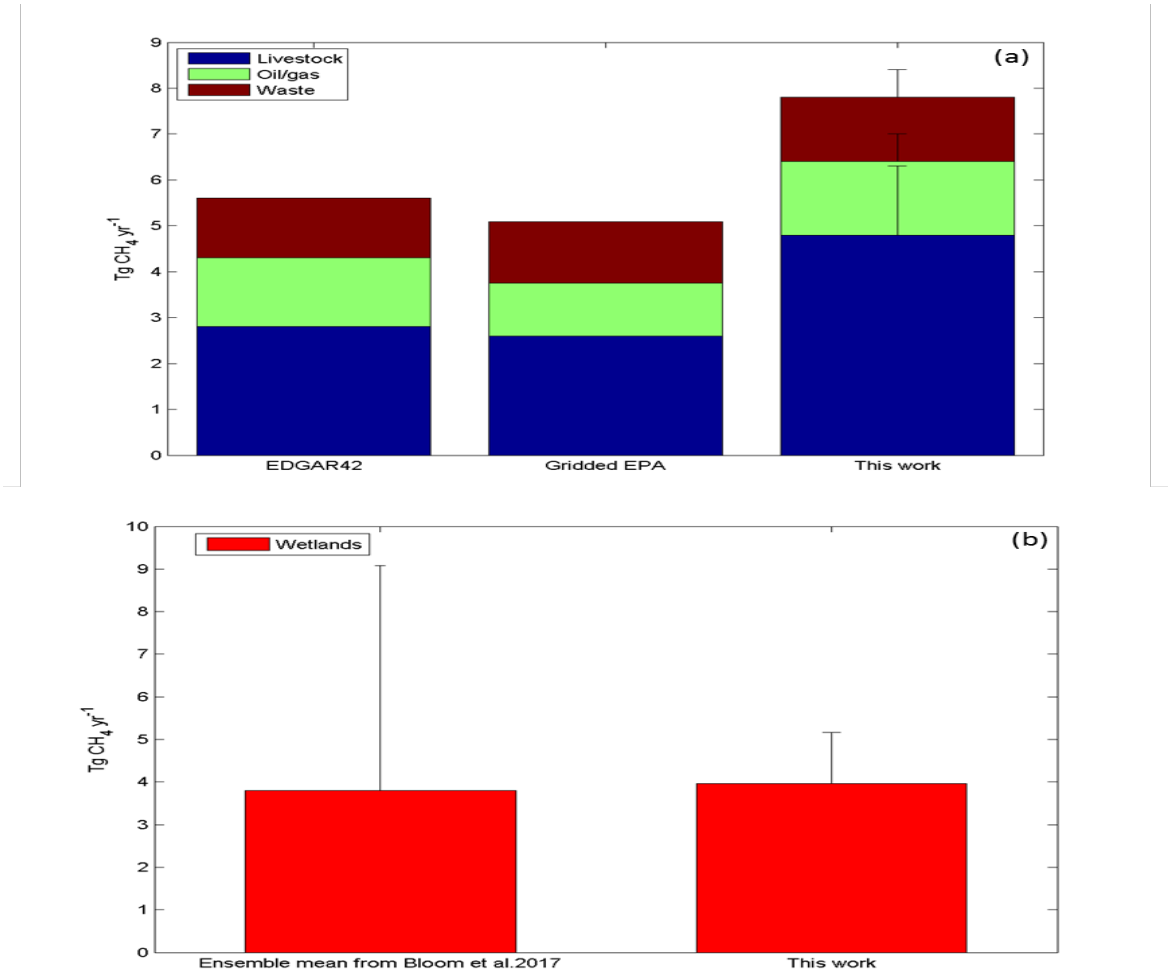


Figure 4.7. Annual regional CH_4 emissions within the US Midwest from different studies including a) anthropogenic and b) natural sources. Error bars indicate the uncertainties of the regional emission estimate from each source, respectively.

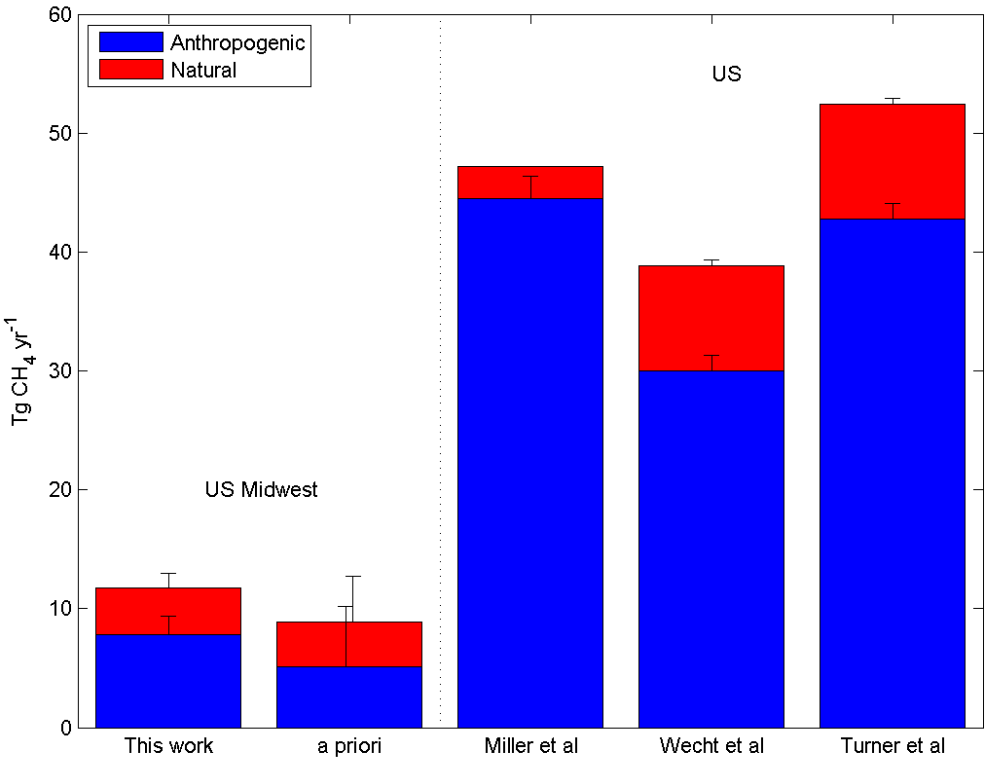


Figure 4.8. Regional and national methane budgets for natural and anthropogenic sources. Error bars indicate the uncertainties of the regional emission estimate from anthropogenic and natural sources, respectively.

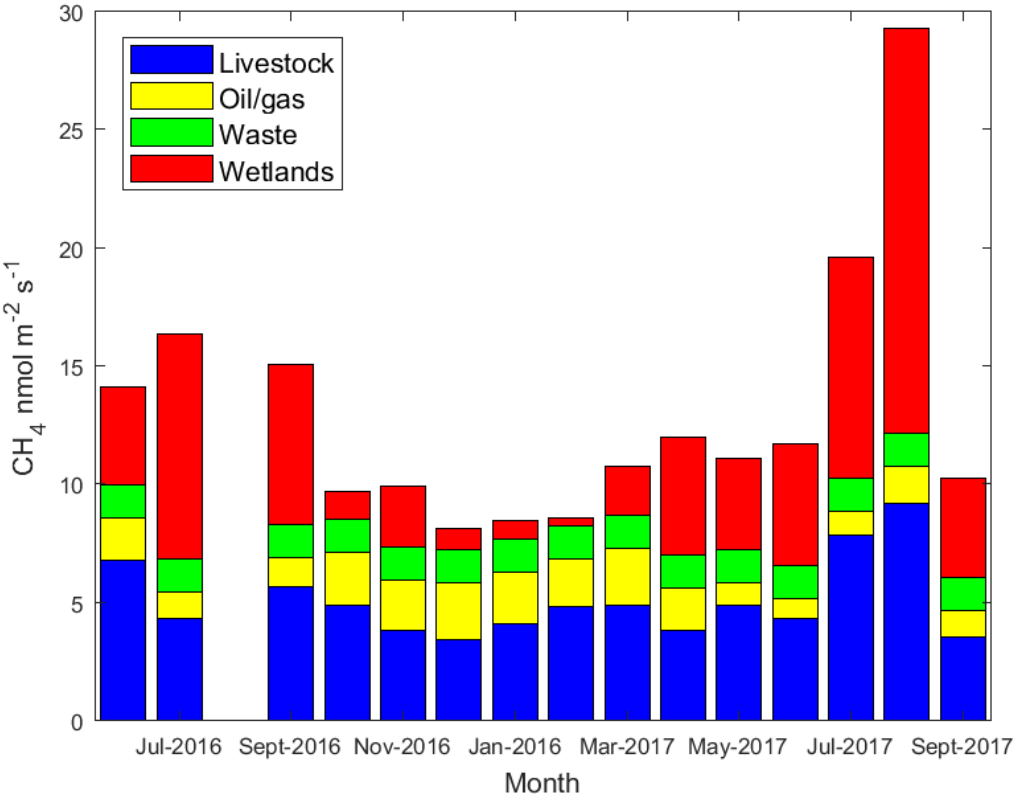


Figure 4.9. Optimized monthly variation of CH_4 emissions for each source category, i.e., the livestock, oil/gas, waste and the wetlands.

Chapter 5
Conclusions

This study sought to: 1) quantify the relative importance of direct versus indirect N_2O emissions within the US Corn Belt and to explore its seasonality and inter-annual variability; 2) assess retrospective and future N_2O emissions at fine spatiotemporal scales to help identify hotspots, hot moments, and mitigation priorities for the US Corn Belt; and 3) partition CH_4 emissions into natural (e.g. wetlands) and anthropogenic (e.g. livestock, waste, and natural gas) sources and explore its temporal variability.

In chapter 2, a novel inverse modeling technique paired with high-precision tall tower observation data was employed, to simultaneously constrain the direct and indirect N_2O sources within the Corn Belt. The data and analyses support that indirect emission factor ranges from 1.4-3.5%, representing an upward adjustment of 1.9- to 4.6-fold compared to the Intergovernmental Panel on Climate Change. The seasonality of the direct and indirect emissions was out of phase, with indirect emissions peaking in May and June, reflecting the temporal dynamics and interaction of synthetic N application with runoff and stream flow.

Findings from Chapter 2 have provided further motivation determining the accurate quantification of N_2O emission at higher spatial and temporal resolution. This is essential in facilitating effective mitigation efforts with knowledge of high emissions in terms of locations (hot spots) and time (hot moments). In Chapter 3, Eulerian modeling approach and multiple tower-based measurements within the Corn Belt were employed to assess N_2O emissions at fine spatiotemporal scales (i.e., $0.125^o \times 0.125^o$, and daily to weekly, respectively). The a posteriori N_2O simulations were strongly correlated ($r=0.69 - 0.82$, $p<0.01$) with mixing ratio observations of multiple tall towers across the US Corn Belt

supporting that the optimized land surface scheme captured the fine-scale spatiotemporal variations from direct and indirect emissions.

The largest slopes for N_2O emissions were 0.13 nmol m^{-2} s^{-1} per year, highlighting that southern Minnesota and the lower Midwest regions drained by the Missouri River will likely represent regional hotspots and mitigation priorities in the future. The dominant driver of the increased emissions is the link between the increasing precipitation and the indirect emissions. The emission factor for agricultural N_2O emissions is projected to increase 22-27% (i.e., 0.04 in 2011 to 0.05 by 2050) in response to changes in climate. A reduction of new nitrogen inputs of 0.04 to 0.05 Tg per year will be required to stabilize N_2O emissions at 2011-2017 values.

Furthermore, it's of critical importance to delve into the relative importance of natural and anthropogenic CH_4 sources within this region. In Chapter 4, tall tower (185 m) aerodynamic flux measurements and a scale factor Bayesian inversion were used to constrain the monthly budget and to partition the CH_4 emissions into natural (e.g. wetlands) and anthropogenic (e.g. livestock, waste, and natural gas) sources within this region. The data and analyses suggested that the livestock and oil/gas sources were 80% and 30% higher than the EPA estimates, respectively. Wetlands dominated the temporal variability with peaking emissions in August, and were the second largest source in the regional CH_4 budget. Flux partitioning revealed that the regional natural, anthropogenic, and total CH_4 emissions were 4.0 ± 1.2, 7.8 ± 1.6, and 11.8 ± 2.0 Tg CH_4 yr^{-1}. The total regional CH_4 emissions for the Corn Belt accounted for ~23% of the overall US budget.

With an increasing observation coverage (i.e., towers, aircrafts, and satellites) in both surface fluxes and atmospheric mixing ratios, as well as the advancement in chemistry and transport modeling techniques, I firmly believe, that our abilities to perform improved constraints on the GHGs in magnitudes, seasonality, and mechanisms are only going to grow over time. Ultimately, these efforts will further advance our scientific understanding of the GHGs, and lead to successful mitigation planning in guiding the GHG reductions.

Future Research Directions

Simulations of regional N_2O with WRF-Chem and its adjoint

As described in Chapter 3, In the recent decades, the increasing observation coverage from towers, aircrafts and satellites, paired with the advancements in chemistry and transport models (CTMs), have facilitated the wide application of forward and inverse models [*Henze et al*., 2007, 2009; *Turner et al*., 2015; *Wells et al*., 2018] in evaluating the surface fluxes. An adjoint model, has shown the advantage in incorporating multiple observations and coupling with a chemistry transport model (CTM) to efficiently reach optimal estimation of the underlying state factors (e.g., surface fluxes, gridded concentration field).

In the recent decade, GEOS-Chem and its adjoint have been extensively used to constrain aerosols [*Henze et al*., 2009], GHGs [*Turner et al*., 2015; *Wells et al*., 2015, 2018] and reactive gas-phase species [*Kaiser et al*., 2017] at regional to global scales. GEOS-Chem CTM (*www.geos-chem.org*) is applied with GEOS-5 meteorological data from the NASA Goddard Earth Observing System. To our knowledge, the finest horizontal resolution in

the most recent version of GEOS-Chem (version 11-01) is $0.25^{o} \times 0.3125^{o}$, approximately 25 km ×25 km over North America. However, previous studies [*Pillai et al*., 2012; *Fu et al*., 2017] suggested high horizontal resolution of 2 to 20 km in the CTMs, in order to accurately capture top-down tower-based observation variability made in the mixed layer. Therefore, the GOES-Chem and its adjoint, is insufficient to exactly probe tower-based observations in response to surface flux variations at finer scales.

Compared to the GEOS-chem model, the WRF-Chem has the advantage in simulating N_2O at higher resolutions. The WRF-Chem model has been widely used in investigating atmospheric compositions at a high horizontal resolution [e.g., 2-10 km, *Tie et al*., 2007; *Chapman et al*., 2009; *Misenis et al*., 2010] .For example, in Chapter 3, the WRF-Chem was applied with an inner domain resolution of 9 km×9 km. Therefore, the development of WRF-Chem and its adjoint, shows a potential in efficiently reaching the optimal estimation of surface N2O fluxes at finer resolutions. Furthermore, Chapter 3 used WRF-Chem model to perform forward simulations, which has been shown to be computationally expensive. The development of WRF-Chem adjoint model can largely improve the simulation efficiency *via* a numerical solution in the construction of Jacobian Matrix, and the model iteration based on the quasi-Newton minimization routine [*Byrd et al*., 1995; *Henze et al*., 2007].

Reference

Alexander, R. B., R. A. Smith, and G. E. Schwarz (2000), Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico, *Nature*, *403*(February), 758–762, doi:10.1038/35001562.

Baker, J. M., T. J. Griffis, and T. E. Ochsner (2012), Coupling landscape water storage and supplemental irrigation to increase productivity and improve environmental stewardship in the U.S. Midwest, *Water Resour. Res.*, *48*(5), 1–12, doi:10.1029/2011WR011780.

Beck, V., C. Gerbig, T. Koch, M. M. Bela, K. M. Longo, S. R. Freitas, J. O. Kaplan, C. Prigent,P. Bergamaschi, and M. Heimann (2013), WRF-Chem simulations in the Amazon region during wet and dry season transitions: Evaluation of methane models and wetland inundation maps, *Atmos. Chem. Phys.*, *13*(16), 7961–7982, doi:10.5194/acp-13- 7961-2013.

Beaulieu, J. J. et al. (2011), Nitrous oxide emission from denitrification in stream and river networks, *Proc. Natl. Acad. Sci. U. S. A.*, *108*(1), 214–219, doi:10.1073/pnas.1011464108.

Bloom, A., W. K. Bowman, M. Lee, J. A. Turner, R. Schroeder, R. J. Worden, R. Weidner, C. K. McDonald, and J. D. Jacob (2017), A global wetland methane emissions and uncertainty dataset for atmospheric chemical transport models (WetCHARTs version 1.0), *Geosci. Model Dev.*, *10*(6), 2141–2156, doi:10.5194/gmd-10-2141-2017.

Bontemps, S., Defourny, P., Bogaert, E. V., Arino, O., Kalogirou, V., and Perez, J. R.(2011),

Globcover Products Description and Validation Report, *Tech. rep.*, ESA, 2011.

Bousquet, P. et al. (2011), Source attribution of the changes in atmospheric methane for 2006-2008, *Atmos. Chem. Phys.*, *11*(8), 3689–3700, doi:10.5194/acp-11-3689-2011.

Bruhwiler, L. M., et al. (2017), U.S. CH_4 emissions from oil and gas production: Have recent large increases been detected?, *J. Geophys. Res. Atmos*., 122, 4070–4083, doi:10.1002/2016JD026157.

Bouwman, A. F., K. W. Vanderhoek, and J. G. J. Oliver (1995), Uncertainties in the global source distribution of nitrous-oxide, *J. Geophys. Res.*, *100*(D2), 2785–2800,

doi:10.1029/94JD02946.

Carouge, C., P. Peylin, P. J. Rayner, P. Bousquet, F. Chevallier, and P. Ciais (2010), What can we learn from European continuous atmospheric CO2 measurements to quantify regional fluxes—Part 2: Sensitivity of flux accuracy to inverse setup, *Atmos. Chem. Phys.*, 8(5), 3119–3129, doi:10.5194/acpd-8-18621-2008.

Chapman, E. G., W. I. Gustafson, R. C. Easter, J. C. Barnard, S. J. Ghan, M. S. Pekour, and J. D. Fast (2009), Coupling aerosol-cloud-radiative processes in the WRF-Chem model: Investigating the radiative impact of elevated point sources, *Atmos. Chem. Phys.*, *9*(3), 945–964, doi:10.5194/acp-9-945-2009.

Chen, M., T. J. Griffis, J. Baker, J. D. Wood, and K. Xiao (2015), Simulating crop phenology in the Community Land Model and its impact on energy and carbon fluxes, *J. Geophys. Res. Biogeosci*., 120, 310–325, doi:10.1002/ 2014JG002780.

Chen, Z. et al. (2016), Partitioning N_2O emissions within the US Corn Belt using an inverse modeling approach, *Global Biogeochem. Cycles*, *30*, 1192–1205, doi:10.1002/2015GB005313.

Cohen, M.J., I.F. Creed, L. Alexander, N.B. Basu, A.J.K. Calhoun, C. Craft, E. D'Amico, E. DeKeyser, L. Fowler, H.E. Golden, J.W. Jawitz, P. Kalla, L.K. Kirkman, C.R. Lane, M. Lang, S.G. Leibowitz, D.B. Lewis, J. Marton, D.L. McLaughlin, D.M. Mushet, H. Raanan-Kiperwas, M.C. Rains, L. Smith, and S.C. Walls (2016), Do geographically isolated wetlands influence landscape functions?, *P. Natl. Acad. Sci. USA*, *113*, 1978-1986, doi:10.1073/pnas.1512650113.

Crutzen, P. J., A. R. Mosier, K. A. Smith, and W. Winiwarter (2007), N2O release from agro-biofuel production negates global warming reduction by replacing fossil fuels, *Atmos. Chem. Phys. Discuss.*, *7*(4), 11191–11205, doi:10.5194/acpd-7-11191-2007.

Davidson, E. A. (2009), The contribution of manure and fertilizer nitrogen to atmospheric nitrous oxide since 1860, *Nat. Geosci.*, *2*(9), 659–662, doi:10.1038/ngeo608.

Davidson, E. A., and D. Kanter (2014), Inventories and scenarios of nitrous oxide

emissions, *Environ. Res. Lett.*, *9*(10), doi:10.1088/1748-9326/9/10/105012.

DelGrosso, S. J., W. J. Parton, a R. Mosier, D. S. Ojima, a E. Kulmala, and S. Phongpan (2000), General model for N_2O and N_2 gas emissions from soils when comparing observed and gas emission rates from irrigated field soils used for model testing NO_2, *Global Biogeochem. Cycles*, *14*(4), 1045–1060, doi:10.1029/1999GB001225.

Del Grosso, S. J., T. Wirth, S. M. Ogle, and W. J. Parton (2008), Estimating agricultural nitrous oxide emissions, *Eos (Washington. DC).*, *89*(51), 529, doi:10.1029/2008EO510001.

De Klein, C., R. S. A. Novoa, S. Ogle, K. A. Smith, P. Rochette, T. C.Wirth, B. G. McConkey, A. Mosier, and K. Rypdal (2006), N_2O emissions from managed soils, and CO_2 emissions from lime and urea application, in *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Vol 4: Agriculture, Forestry and Other Land Use*, edited by H. S. Eggleston et al., pp. 11.11–11.54, Institute for Global Environmental Strategies (IGES), Intergovernmental Panel on Climate Change (IPPC), Kanagawa, Japan.

Dlugokencky, E.J., L. Bruhwiler, J.W.C. White, L.K. Emmons, P.C. Novelli, S.A. Montzka, K.A. Masarie, P.M. Lang, A.M. Crotwell, J.B. Miller, and L.V. Gatti (2009), Observational constraints on recent increases in the atmospheric CH_4 burden, *Geophys. Res. Lett.*, *36*, doi:10.1029/2009gl039780.

Dlugokencky, E.J., E.G. Nisbet, R. Fisher, and D. Lowry (2011), Global atmospheric methane: budget, changes and dangers, *Philos. T. Roy. Soc. A*, *369*, 2058-2072, doi:10.1098/rsta.2010.0341.

Dong, H., Mangino, J., McAllister, T.A., Hatfield, J.L., Johnson, D.E., Lassey, K.R., de Lima, M.A., Romanovskaya, A., (2006) Emissions from livestock and manure management. In: Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T., Tanabe, K. (Eds.), 2006 IPCC Guidelines for National Greenhouse Gas Inventories, vol. 4. IGES, Japan, pp. 1-87.

EPA. Inventory of US Greenhouse Gas Emissions and Sinks: 1990−2014, 2016. https://www.epa.gov/ghgemissions/usgreenhouse- gas-inventory-report-1990-2014.

Fassbinder, J. J., N. M. Schultz, J. M. Baker, and T. J. Griffis (2013), Automated, low-power chamber system for measuring nitrous oxide emissions., *J. Environ. Qual.*, *42*, 606–14, doi:10.2134/jeq2012.0283.

Fu, C., X. Lee, T. J. Griffis, E. J. Dlugokencky, and A. E. Andrews (2017), Investigation of the N_2O emission strength in the U. S. Corn Belt, *Atmos. Res.*, *194*(April), 66–77, doi:10.1016/j.atmosres.2017.04.027.

GALLOWAY, J. N., J. D. ABER, J. W. ERISMAN, S. P. SEITZINGER, R. W. HOWARTH, E. B. COWLING, and B. J. COSBY (2003), The Nitrogen Cascade, *Bioscience*, *53*(4), 341, doi:10.1641/0006-3568(2003)053[0341:TNC]2.0.CO;2.

Gerbig, C., J. C. Lin, S. C. Wofsy, B. C. Daube, a. E. Andrews, B. B. Stephens, P. S. Bakwin, and C. a. Grainger (2003), Toward constraining regional-scale fluxes of CO 2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, *J. Geophys. Res. Atmos.*, *108*(D24), n/a–n/a, doi:10.1029/2002JD003018.

Gómez, D. R., J. D. Watterson, B. B. Americano, C. Ha, G. Marland, E. Matsika, L. N. Namayanga, B. Osman-Elasha, J.D. Kalenga Saka, and K. Treanton (2006), IPCC guidelines for national greenhouse gas inventories. *Institute for Global Environmental Strategies, Kamiyamaguchi Hayama, Japan*.

Grace, P. R., G. Philip Robertson, N. Millar, M. Colunga-Garcia, B. Basso, S. H. Gage, and J. Hoben (2011), The contribution of maize cropping in the Midwest USA to global warming: A regional estimate, *Agric. Syst.*, *104*(3), 292–296, doi:10.1016/j.agsy.2010.09.001.

Griffis, T. J., J. M. Baker, S. D. Sargent, M. Erickson, J. Corcoran, M. Chen, and K. Billmark (2010), Influence of C4 vegetation on 13CO2 discrimination and isoforcing in the upper Midwest, United States, *Global Biogeochem. Cycles*, *24*(4), 1–16, doi:10.1029/2009GB003768.

Griffis, T. J. et al. (2016), Investigating the source, transport, and isotope composition of water vapor in the planetary boundary layer, *Atmos. Chem. Phys.*, *16*(8), 5139–5157, doi:10.5194/acp-16-5139-2016.

Griffis, T. J., X. Lee, J. M. Baker, M. P. Russelle, X. Zhang, R. Venterea, and D. B. Millet (2013), Reconciling the differences between top-down and bottom-up

estimates of nitrous oxide emissions for the U.S. Corn Belt, *Global Biogeochem. Cycles*, *27*(3), 746–754, doi:10.1002/gbc.20066.

Griffis, T. J., Z. Chen, J. M. Baker, J. D. Wood, D. B. Millet, X. Lee, R. T. Venterea, and P. A. Turner (2017), Nitrous oxide emissions are enhanced in a warmer and wetter world, *Proc. Natl. Acad. Sci.*, *114*(45), 12081–12085, doi:10.1073/pnas.1704552114.

Groffman, P. M., K. Butterbach-Bahl, R. W. Fulweiler, A. J. Gold, J. L. Morse, E. K. Stander, C. Tague, C. Tonitto, and P. Vidon (2009), Challenges to incorporating spatially and temporally explicit phenomena (hotspots and hot moments) in denitrification models, *Biogeochemistry*, *93*(1–2), 49–77, doi:10.1007/s10533-008-9277-5.

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, I. V., Brönnimann, S., Charabi, Y., Dentener, F. J., Dlugokencky, E. J., Easterling, D. R., Kaplan, A., Soden, B, J., Thorne, P. W., Wild, M., and Zhai, P. M.: Observations: atmosphere and surface, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to *he Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, 2013.

Heald Colette L., C. L., D. J. Jacob, D. B. A. Jones, P. I. Palmer, J. A. Logan, D. G. Streets, G. W. Sachse, J. C. Gille, R. N. Hoffman, and T. Nehrkorn (2004), Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide, *J. Geophys. Res. D Atmos.*, *109*(23), 1–17, doi:10.1029/2004JD005185.

Henze, D. K., A. Hakami, and J. H. Seinfeld (2007), Development of the adjoint of GEOS-Chem, *Atmos. Chem. Phys.*, *7*(9), 2413–2433, doi:10.5194/acp-7-2413-2007.

Hinshaw, S. E., and R. A. Dahlgren (2013), Dissolved nitrous oxide concentrations and fluxes from the eutrophic San Joaquin River, California, *Environ. Sci. Technol.*, *47*(3), 1313–1322, doi:10.1021/es301373h.

Hong, S.-Y., J. Dudhia, and S.-H. Chen (2004), A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation,

Mon. Weather Rev., *132*(1), 103–120, doi:10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.

Hong, S.-Y., Y. Noh, and J. Dudhia (2006), A new vertical diffusion package with an explicit treatment of entrainment processes., *Mon. Weather Rev.*, *134*(9), 2318–2341, doi:10.1175/MWR3199.1.

Hu, L. et al. (2015), Emissions of C6-C8 aromatic compounds in the United States: Constraints from tall tower and aircraft measurements, *J. Geophys. Res. D Atmos.*, *120*(2), 826–842, doi:10.1002/2014JD022627.

Hristov, A. N. et al. (2015) An inhibitor persistently decreased enteric methane emission from dairy cows with no negative effect on milk production, *Proc. Natl. Acad. Sci.*, *112*(37), E5218–E5218, doi:10.1073/pnas.1515515112.

Huntzinger, D. N., Schwalm, C., Michalak, A. M., Schaefer, K., King, A. W., Wei, Y., Jacobson, A., Liu, S., Cook, R. B., Post, W. M., Berthier, G., Hayes, D., Huang, M., Ito, A., Lei, H., Lu, C., Mao, J., Peng, C. H., Peng, S., Poulter, B., Riccuito, D., Shi, X., Tian, H., Wang, W., Zeng, N., Zhao, F., and Zhu, Q.: The North American Carbon Program Multi-Scale Synthesis and Terrestrial Model Intercomparison Project – Part 1: Overview and experimental design, Geosci. Model Dev., 6, 2121–2133, doi:10.5194/gmd-6-2121-2013, 2013.

Jacob, D. J., A. J. Turner, J. D. Maasakkers, J. Sheng, K. Sun, X. Liu, K. Chance, I. Aben, J. McKeever, and C. Frankenberg (2016), Satellite observations of atmospheric methane and their value for quantifying methane emissions, *Atmos. Chem. Phys.*, *16*(22), 14,371–14,396, doi:10.5194/acp-16-14371-2016.

Janjic, Z. (2002), Nonsingular Implementation of the Mellor-Yamada Level 2.5 Scheme in the NCEP Meso model, *Natl. Centers Environ. Predict.*, 61.

Jeong, S., C. Zhao, A. E. Andrews, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012), Seasonal variation of CH_4 emissions from central California, *J. Geophys. Res.*, *117*(D11), D11306, doi:10.1029/2011JD016896.

Kaiser, J. et al. (2017), High-resolution inversion of OMI formaldehyde columns to quantify isoprene emission on ecosystem-relevant scales: application to the Southeast US, *Atmos. Chem. Phys. Discuss.*, (x), 1–27, doi:10.5194/acp-2017-1137.

Kain, J. S. (2004), The Kain–Fritsch Convective Parameterization: An Update, *J. Appl. Meteorol.*, *43*(1), 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Kaplan, J. (2002), Wetlands at the Last Glacial Maximum: Distribution and methane emissions,*Geophys. Res. Lett.*, *29*(6), 1079,doi:10.1029/2001GL013366.

Kirschke, S. et al. (2013), Three decades of global methane sources and sinks, *Nat. Geosci.*, *6*(10), 813–823, doi:10.1038/ngeo1955.

Kim, S. Y., D. B. Millet, L. Hu, M. J. Mohr, T. J. Griffis, D. Wen, J. C. Lin, S. M. Miller, and M. Longo (2013), Constraints on carbon monoxide emissions based on tall tower measurements in the U.S. upper midwest, *Environ. Sci. Technol.*, *47*(15), 8316–8324, doi:10.1021/es4009486.

Kopacz, M. et al. (2010), Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES), *Atmos. Chem. Phys.*, *10*(3), 855–876, doi:10.5194/acp-10-855-2010.

Kort, E. A., J. Eluszkiewicz, B. B. Stephens, J. B. Miller, C. Gerbig, T. Nehrkorn, B. C. Daube, J. O. Kaplan, S. Houweling, and S. C. Wofsy (2008), Emissions of CH4 and N2O over the United States and Canada based on a receptor-oriented modeling framework and COBRA-NA atmospheric observations, *Geophys. Res. Lett.*, *35*(18), 1–5, doi:10.1029/2008GL034031.

Kort, E. A., C. Frankenberg, K. R. Costigan, R. Lindenmaier, M. K. Dubey, and D. Wunch (2014), Four corners: The largest US methane anomaly viewed from space, *Geophys. Res. Lett.*, *41*(19), 6898–6903, doi:10.1002/2014GL061503.

Kretschmer, R., F. T. Koch, D. G. Feist, G. Biavati, U. Karstens, and C. Gerbig (2012), Toward assimilation of observation-derived mixing heights to improve atmospheric tracer transport models, *Geophys. Monogr. Ser.*, *200*, 185–205, doi:10.1029/2012GM001255.

Lassey, K. R. (2007), Livestock methane emission: From the individual grazing animal through national inventories to the global methane cycle, *Agric. For. Meteorol.*, *142*(2–4), 120–132, doi:10.1016/j.agrformet.2006.03.028.

Leng, G., et al. (2016), Simulating county-level crop yields in the Conterminous United States using the Community LandModel: The effects of optimizing irrigation and

fertilization, *J. Adv. Model. Earth Syst.*, (13 SEP 2016), 1411–1431, doi:10.1002/2016MS000660.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., & Grainger, C. A. (2003). A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *Journal of Geophysical Research: Atmospheres (1984–2012)*, *108*(D16).

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Grainger, C. A., & Hollinger, D. Y. (2004). Measuring fluxes of trace gases at regional scales by Lagrangian observations: Application to the CO_2 Budget and Rectification Airborne (COBRA)study. *Journal of Geophysical Research: Atmospheres (1984–2012)*, *109*(D15).

Lin, J. C., C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, C. A. Grainger, B. B. Stephens, P. S. Bakwin, and D. Y. Hollinger (2004), Measuring fluxes of trace gases at regional scales by Lagrangian observations: Application to the CO2 Budget and Rectification Airborne (COBRA) study, *J. Geophys. Res. D Atmos.*, *109*(15), 1–23, doi:10.1029/2004JD004754.

Luo, G. J., R. Kiese, B. Wolf, and K. Butterbach-Bahl (2013), Effects of soil temperature and moisture on methane uptake and nitrous oxide emissions across three different ecosystem types, *Biogeosciences*, *10*(5), 3205–3219, doi:10.5194/bg-10-3205-2013.

Matross, D. M. et al. (2006), Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data, *Tellus, Ser. B Chem. Phys. Meteorol.*, *58*(5), 344–358, doi:10.1111/j.1600-0889.2006.00206.x.

McClain, M. E. et al. (2003), Biogeochemical Hot Spots and Hot Moments at the Interface of Terrestrial and Aquatic Ecosystems, *Ecosystems*, *6*(4), 301–312, doi:10.1007/s10021-003-0161-9.

Maasakkers, J. D., D. J. Jacob, M. P. Sulprizio, A. J. Turner, M. Weitz, T. Wirth, C. Hight, M. DeFigueiredo, and A. A. Bloom (2012), A gridded version of the EPA national methane inventory, *Environ. Sci. Technol.*,, *50* (23),13123–13133, **doi:**10.1021/acs.est.6b02878.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H. and Ek, M.B. (2006). North American regionalreanalysis. *Bull. Am. Meteorol. Soc.*, *87*(3), 343-360.

Mikaloff Fletcher, S. E., P. P. Tans, L. M. Bruhwiler, J. B. Miller, and M. Heimann (2004), CH4 sources estimated from atmospheric observations of CH4 and its 13C/12C isotopic ratios: 2. Inverse modeling of CH4 fluxes from geographical regions, *Global Biogeochem. Cycles*, *18*(4), 1–15, doi:10.1029/2004GB002224.

Millar, N., G. Philip Robertson, P. R. Grace, R. J. Gehl, and J. P. Hoben (2010), Nitrogen fertilizer management for nitrous oxide (N 2O) mitigation in intensive corn (Maize) production: An emissions reduction protocol for US Midwest agriculture, *Insectes Soc.*, *57*(2), 185–204, doi:10.1007/s11027-010-9212-7.

Miller, S. M. et al. (2008), Sources of carbon monoxide and formaldehyde in North America determined from high-resolution atmospheric data, *Atmos. Chem. Phys. Discuss.*, *8*(3), 11395–11451, doi:10.5194/acpd-8-11395-2008.

Miller, S. M. et al. (2013), Anthropogenic emissions of methane in the United States., *Proc. Natl. Acad. Sci. U. S. A.*, *110*(50), 20018–22, doi:10.1073/pnas.1314392110.

Miller, S. M. et al. (2014), Observational constraints on the distribution, seasonality, and environmental predictors of North American boreal methane emissions, *Global Biogeochem. Cycles*, *28*(2), 146–160, doi:10.1002/2013GB004580.

Misenis, C., and Y. Zhang (2010), An examination of sensitivity of WRF/Chem predictions to physical parameterizations, horizontal grid spacing, and nesting options, *Atmos. Res.*, *97*(3), 315–334, doi:10.1016/j.atmosres.2010.04.005.

Myhre, G. et al. (2013), Anthropogenic and Natural Radiative Forcing: In Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, *Cambridge Univ. Press. Cambridge, United Kingdom New York, NY, USA*, 659–740, doi:10.1017/ CBO9781107415324.018.

Nehrkorn, T., J. Eluszkiewicz, S. C. Wofsy, J. C. Lin, C. Gerbig, M. Longo, and S. Freitas (2010), Coupled weather research and forecasting-stochastic time-inverted lagrangian transport (WRF-STILT) model, *Meteorol. Atmos. Phys.*, *107*(1), 51–64,

doi:10.1007/s00703-010-0068-x.

Nevison, C. (2000), Review of the IPCC methodology for estimating nitrous oxide emissions associated with agricultural leaching and runoff, *Chemosph. - Glob. Chang. Sci.*, *2*(3-4), 493–500, doi:10.1016/S1465-9972(00)00013-1.

Nisbet, E. G.,E. J. Dlugokencky, and P. Bousquet (2014), Methane on the Rise – Again, *Science*, 343, 493–495, https://doi.org/10.1126/science.1247828.

Oleson, K. W., D. M. Lawrence, B. Gordon, M. G. Flanner, E. Kluzek, J. Peter, S. Levis, S. C. Swenson, E. Thornton, and J. Feddema (2013), Technical description of version 4.5 of the Community Land Model (CLM), *NCAR/TN-503+STR NCAR Tech. Note*, (July), 266, doi:10.5065/D6RR1W7M.

Olson, D., T. J. Griffis, A. Noormets, R. Kolka, and J. Chen (2013), Interannual, seasonal, and retrospective analysis of the methane and carbon dioxide budgets of a temperate peatland, *J. Geophsyical Res. - Biogeosciences*, *118*, 226–238, doi:10.1002/jgrg.20031.

Outram, F. N., and K. M. Hiscock (2012), Indirect nitrous oxide emissions from surface water bodies in a lowland arable catchment: A significant contribution to agricultural greenhouse gas budgets?, *Environ. Sci. Technol.*, *46*(15), 8156–8163, doi:10.1021/es3012244.

Parton, W., E. Holland, S. Del Grosso, D. Hartman, M. Martin, A. Mosier, D. Ojima, and D. Schimel (2001), Generalized model for NOx and N2O emissions from soils, *J. Geophys. Res.*, *106*(D15), 17403–17419, doi:10.1029/2001JD900101.

Pillai, D., C. Gerbig, R. Kretschmer, V. Beck, U. Karstens, B. Neininger, and M. Heimann (2012), Comparing Lagrangian and Eulerian models for CO₂ transport – a step towards Bayesian inverse modeling using WRF/STILT-VPRM, *Atmos. Chem. Phys.*, *12*(19), 8979–8991, doi:10.5194/acp-12-8979-2012.

Ravishankara, A. R., J. S. Daniel, and R. W. Portmann (2009), Nitrous Oxide (N 2 O): The Dominant Ozone-Depleting Substance Emitted in the 21st Century, *Science (80-.).*, *123*(October), 2007–2010, doi:10.1126/science.1176985.

Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: theory and practice* (Vol. 2, p. 256). Singapore: World scientific.

Saikawa, E., C. A. Schlosser, and R. G. Prinn (2013), Global modeling of soil nitrous oxide emissions from natural processes, *Global Biogeochem. Cycles*, *27*(3), 972–989, doi:10.1002/gbc.20087.

Schroeder, R., McDonald, K. C., Chapman, B. D., Jensen, K., Podest, E., Tessler, Z. D., Bohn, T. J., and Zimmermann, R (2015), Development and Evaluation of a Multi-Year Fractional Surface Water Data Set Derived from Active/Passive Microwave Remote Sensing Data, *Remote Sensing*, 7, 16688–16732, https://doi.org/10.3390/rs71215843..

Shcherbak, I., N. Millar, and G. P. Robertson (2014), Global metaanalysis of the nonlinear response of soil nitrous oxide (N2O) emissions to fertilizer nitrogen., *Proc. Natl. Acad. Sci. U. S. A.*, *111*(25), 9199–204, doi:10.1073/pnas.1322434111.

Singurindy, O., M. Molodovskaya, B. K. Richards, and T. S. Steenhuis (2009), Nitrous oxide emission at low temperatures from manure-amended soils under corn (Zea mays L.), *Agric. Ecosyst. Environ.*, *132*(1-2), 74–81,doi:10.1016/j.agee.2009.03.001.

Smith, K. a, A. R. Mosier, P. J. Crutzen, and W. Winiwarter (2012), The role of N2O derived from crop-based biofuels, and from agriculture in general, in Earth's climate., *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, *367*(1593), 1169–74, doi:10.1098/rstb.2011.0313.

Sommer, S., S. Petersen, and H. Møller (2004), Algorithms for calculating methane and nitrous oxide emissions from manure management, *Nutr. Cycl. Agroecosystems*, *69*(2), 143–154, doi:10.1023/B:FRES.0000029678.25083.fa.

Stehfest, E., and L. Bouwman (2006), N2O and NO emission from agricultural fields and soils under natural vegetation: Summarizing available measurement data and modeling of global annual emissions, *Nutr. Cycl. Agroecosystems*, *74*(3), 207–228, doi:10.1007/s10705-006-9000-7.

Tian, H. et al. (2012), Century-Scale Responses of Ecosystem Carbon Storage and Flux to Multiple Environmental Changes in the Southern United States, *Ecosystems*, *15*(4), 674–694, doi:10.1007/s10021-012-9539-x.

Tie, X., S. Madronich, G. H. Li, Z. Ying, R. Zhang, A. R. Garcia, J. Lee-Taylor, and Y. Liu (2007), Characterizations of chemical oxidants in Mexico City: A regional chemical dynamical model (WRF-Chem) study, *Atmos. Environ.*, *41*(9), 1989–2008, doi:10.1016/j.atmosenv.2006.10.053.

Thompson, R. L. et al. (2014), Nitrous oxide emissions 1999 to 2009 from a global atmospheric inversion, *Atmos. Chem. Phys.*, *14*(4), 1801–1817, doi:10.5194/acp-14-1801-2014.

Turner, A. J. et al. (2015), Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, *Atmos. Chem. Phys.*, *15*(12), 7049–7069, doi:10.5194/acp-15-7049-2015.

Turner, A. J. and D. J. Jacob (2015), Balancing aggregation and smoothing errors in inverse models, *Atmos. Chen. Phys*., 15, 7039–7048, doi:10.5194/acp-15-7039-2015.

Turner, P. A., T. J. Griffis, X. Lee, J. M. Baker, R. T. Venterea, and J. D. Wood (2015), Indirect nitrous oxide emissions from streams within the US Corn Belt scale with stream order, *Proc. Natl. Acad. Sci.*, *112*(32), 201503598, doi:10.1073/pnas.1503598112.

Turner, P. A., T. J. Griffis, J. M. Baker, X. Lee, J. T. Crawford, L. C. Loken, and R. T. Venterea (2016), Regional-scale controls on dissolved nitrous oxide in the Upper Mississippi River, *Geophys. Res. Lett.*, *43*(9), 4400–4407, doi:10.1002/2016GL068710.

United States Department of Agriculture – National Agricultural Statistics Service, Census of Agriculture, 2012. *Technical Report* 2014, 2014.

US Forest Service (2016), National Wetlands Inventory, http://www.fws.gov/wetlands.

Vanderzaag, A. C., T. K. Flesch, R. L. Desjardins, H. Baldé, and T. Wright (2014), Agricultural and Forest Meteorology Measuring methane emissions from two dairy farms : Seasonal and manure-management effects, *Agric. For. Meteorol.*, *194*, 259–267, doi:10.1016/j.agrformet.2014.02.003.

Valentine, D. W., and K. Weier (1996), Generalized model for Nz and NzO production from nitrification, *Global Biogeochem. Cycles*, *10*(3), 401–412.

Wania,R et al. (2013), Present state of global wetland extent and wetland methane modelling: Conclusions from a model inter-comparison project (WETCHIMP),

Wecht, K.J, Daniel J. Jacob, Christian Frankenberg, Zhe Jiang, and D. R. Blake (2014), Journal of Geophysical Research : Atmospheres of SCIAMACHY satellite data, *J. Geophys. Res. Atmos. Res.*, 7741–7756, doi:10.1002/2014JD02155.

Wei, Y., et al. (2014), The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project—Part 2: Environmental driver data, *Geosci. Model Dev.*, **7**, 2875–2893, doi:10.5194/gmd-7-2875-2014.

Wells, K. C. et al. (2015), Simulation of atmospheric N₂O with GEOS-Chem and its adjoint: evaluation of observational constraints, *Geosci. Model Dev.*, *8*(10), 3179–3198, doi:10.5194/gmd-8-3179-2015.

Wells K. C. et al. (2018), Top-down constraints on global N_2Oemissions at optimal resolution: application of a new dimension reduction technique, *Atmos. Chem. Phys.*, *18*, 735–756, doi:10.1111/geb.12166.

Werle, P., R. Mcke, and F. Slemr (1993), The limits of signal averaging in atmospheric trace-gas monitoring by tunable diode-laser absorption spectroscopy (TDLAS), *Appl. Phys. B Photophysics Laser Chem.*, *57*(2), 131–139, doi:10.1007/BF00425997.

Wood, J.D., R.J. Gordon, and C. Wagner-Riddle (2013), Biases in discrete CH_4 and N_2O sampling protocols associated with temporal variation of gas fluxes from manure storage systems, *Agr. Forest Meteorol.*, *171*, 295-305, doi:10.1016/j.agrformet.2012.12.014.

Wood, J. D., T. J. Griffis, J. M. Baker, C. Frankenberg, M. Verma, and K. Yuen (2016), Multi-scale analyses reveal robust relationships between gross primary production and solar induced fluorescence, *Geophys. Res. Lett.*, *44*, 533-541, doi:10.1002/2016GL070775.

Zhang, X., X. Lee, T. J. Griffis, J. M. Baker, and W. Xiao (2014a), Estimating regional greenhouse gas fluxes: An uncertainty analysis of planetary boundary layer techniques and bottom-up inventories, *Atmos. Chem. Phys.*, *14*, 10705–10719, doi:10.5194/acp-14-10705-2014.

Zhang, X., X. Lee, T. J. Griffis, J. M. Baker, M. D. Erickson, N. Hu, and W. Xiao (2014b), The influence of plants on atmospheric methane in an agriculture-dominated landscape, *Int. J. Biometeorol.*, *58*(5), 819–833, doi:10.1007/s00484-013-0662-y.

Zhao, C., A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T. Nehrkorn, and M. L. Fischer (2009), Atmospheric inverse estimates of methane emissions from Central California, *J. Geophys. Res. Atmos.*, *114*(16), 1–13, doi:10.1029/2008JD011671.

Zona, D., I. a. Janssens, M. S. Verlinden, L. S. Broeckx, J. Cools, B. Gioli, A. Zaldei, and R. Ceulemans (2011), Impact of extreme precipitation and water table change on N2O fluxes in a bio-energy poplar plantation, *Biogeosciences Discuss.*, *8*(2), 2057–2092, doi:10.5194/bgd-8-2057-2011.