Carnegie Mellon University

CARNEGIE INSTITUTE OF TECHNOLOGY

Dissertation

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF Doctor of Philosophy

TITLE Atmospheric Impacts of Biofuel and Natural Gas Life Cycle Greenhouse Gas Emissions and Policy Implications

PRESENTED BY Stefan Schwietzke

ACCEPTED BY THE DEPARTMENT OF

Engineering & Public Policy

_____ _____

ADVISOR, MAJOR PROFESSOR DATE

_____ _____

DEPARTMENT HEAD DATE

APPROVED BY THE COLLEGE COUNCIL

_____ _____

DEAN DATE

Atmospheric Impacts of Biofuel and Natural Gas Life Cycle Greenhouse Gas Emissions and Policy Implications

Submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

in

Engineering and Public Policy

Stefan Schwietzke

Dipl.-Ing., Technology Management, Universität Stuttgart

Carnegie Mellon University
Pittsburgh, PA

December 2013

Abstract

Many studies have recently reported estimates of greenhouse gas (GHG) emissions and associated potential climate impacts of biofuel and natural gas (NG) use. U.S. corn ethanol production keeps increasing under federal mandates, and NG production soars due to successful tapping of unconventional resources in North America, particularly shale gas. Numerous life cycle assessment (LCA) studies document technology specific corn ethanol and NG GHG estimates. The estimates often include all life cycle stages from fuel supply to combustion, and point out potential for emissions reductions.

Several studies suggest that using GHG emissions as an evaluation metric underestimates corn ethanol's radiative forcing (RF) impact – a precursor and indicator for global temperature change – by 10-90% over the next few decades. This emissions timing effect may overestimate (i) ethanol's climate benefits over gasoline and (ii) the effectiveness of U.S. policies mandating and subsidizing ethanol. This work revisits the above studies, and builds upon existing models to quantify RF impacts across the corn ethanol life cycle. The emissions timing factor (ETF) is significantly smaller than previous estimates (2-13% depending on the chosen impact time frame), and the effect is dwarfed by uncertainty in the GHG emissions estimates. Nevertheless, ETF reduces ethanol's probability of meeting the federal target of 20% GHG reduction relative to gasoline from 53% (according to EPA GHG estimates) to 7-29%. However, the small potential climate impacts from U.S. ethanol use may not actually be observable based on estimated initial increases in global average surface temperature of < 0.001 °C.

About 25% of global primary energy production comes from NG, whose life cycle GHG emissions and potential future climate impacts from substituting coal are highly uncertain due to fugitive methane (CH_4) emissions from the NG industry. Accurately quantifying the NG fugitive emissions (FE) rate – the percentage of produced NG, mainly CH_4 and ethane (C_2H_6) – released to the atmosphere is challenging due to the size and complexity of the NG industry. Recent LCA estimates suggest that the current NG FE rate could be as high as 8% and 6%, from shale and conventional NG, respectively, and other bottom-up studies indicate even higher rates several decades ago. This work analyzes possible ranges of the global average NG FE rate based on atmospheric CH_4, C_2H_6, and carbon isotope ($\delta^{13}C$-CH_4) measurements recorded since 1984, and top-down modeling of their sources and sinks.

Box-model, $\delta^{13}C$-CH_4 mass balance, and 3D-modeling results agree on best estimate NG FE rates of 3-5% (of dry NG production and dry NG composition) globally over the past decade, and 5-8% around 1990. Upper bound FE rates are 5% and 7% in 2010 and 2000, respectively. Best estimate and upper bound values may be overestimated because both assume lower bound emissions from oil and coal production as well as complete absence of natural hydrocarbon seepage. While LCA studies are useful for identifying processes with the greatest NG FE reduction potential, the recent high bottom-up estimates do not appear representative of the U.S. national average based on top-down modeling results. Given the steadily declining NG FE rates one may expect that further emissions abatement is possible if industry practices are further improved.

Acknowledgements

This research was made possible through support from the Center for Climate and Energy Decision Making (CEDM). This Center has been created through a cooperative agreement between the National Science Foundation (SES-0949710) and Carnegie Mellon University. The ERM Foundation-North America Sustainability Fellowship has provided additional funding.

My first thanks go to my advisors and committee members: Mike Griffin, who patiently guided my development of critical thinking and clear communication, and for being an incessant source of subject knowledge; Scott Matthews, who accompanied my first steps through the web of academia, and for instilling a sense of confidence in my own work; Lori Bruhwiler, who volunteered to teach me everything I know about atmospheric inversions, and for putting in many hours of modeling, without which this dissertation would be incomplete; Klaus Keller, who contributed invaluable outside perspective, and for helping me improve the manuscript with a comprehensive review; Chris Weber, who nurtured a continuous exchange of data and relevant publications, and for taking the time to discuss even small details of my work.

Several other individuals and organizations supported my research throughout the years. I am grateful to Ken Davis and Thomas Lauvaux (Penn State University) for guidance in the early stages of this work. I would like to thank the researchers and staff from the Global Monitoring Division of the Earth System Research Laboratory at the National Oceanic and Atmospheric Administration (NOAA) in Boulder, CO. My visits to your lab were key for finishing my dissertation on time, and your feedback helped me adjust the directions of my research. Thanks to Gaby Pétron, John Miller, and Kenneth Masarie for providing guidance and various data. The faculty and staff of Carnegie Mellon's Department of Engineering and Public Policy (EPP) have created a warm, supportive, and challenging atmosphere, which made this place far more than just a place to work and study. I feel fortunate to have been a student in the Green Design Institute, a community that cultivates exemplary mutual support in many aspects of early academic life. Thanks to Xuejiao (Fiona) Zhang for meticulously translating and processing Chinese documents and data.

My time in EPP would not have been the same without Dan Schnitzer, Apurba Sakti, Enes Hosgor, Anu Narayanan, Austin Mitchell, Yeganeh Mashayekh, Kelly Klima, Patrick Agyapong, Julian Lamy, and Nektarios Leontiadis. What started as camaraderie and bonding over coursework assignments (and celebrating submissions thereof) soon became important to me as a home away from home. I deeply treasure this chapter of my life, and you were the main characters. To be continued.

Finally, I thank my family who supported me in many ways throughout 24 years of education. I'm grateful that I had the privilege to focus my attention on my own personal and academic development, which you've always backed with loving encouragement. Doing this over long distances for a decade is hard, but I know you've tried your best, and you did very, very well.

Ripened in the finest garden,
Supported by a trusted few,
Through the leafs a fruit is shining.
A blossom taps towards the blue.

It shall unfold its precious grain,
A source of guidance yet unknown,
So that before the petals drop
A garden elsewhere has been sown.

Contents

List of Figures

List of Tables

List of Abbreviations

BAU	Business as usual
BB	Biomass burning
BFC	Biomass fuel combustion
BLM	U.S. Bureau of Land Management
CBM	Coal-bed methane
CH_4	Methane
C_2H_6	Ethane
CO_2	Carbon dioxide
CO_2e	Carbon dioxide equivalent
CRF	Cumulative radiative forcing
CT-CH_4	CarbonTracker CH_4 model
$\delta^{13}C$-CH_4	Methane carbon isotope ^{13}C
DOE	U.S. Department of Energy
EDGAR	Emissions Database for Global Atmospheric Research
EF	Emissions factor
EIA	U.S. Energy Information Administration
EISA	Energy Independence and Security Act
EPA	U.S. Environmental Protection Agency
ETF	Emissions timing factor
FF	Fossil fuel
FSU	Former Soviet Union
GFED	Global Fire Emissions Database
GHG	Greenhouse gas
GHGA	Greenhouse gas accounting
GOM	Gulf of Mexico
IPCC	Intergovernmental Panel on Climate Change
LUC	Land use change
N_2	Nitrogen
N_2O	Nitrous oxide
NG	Natural gas
NGL	Natural gas liquids
NIR	National Inventory Report
NOAA	National Oceanic and Atmospheric Administration
OECD	Organisation for Economic Co-operation and Development
OH	Hydroxil radical
PNH	Polar Northern Hemisphere
PSH	Polar Southern Hemisphere
RF	Radiative forcing
RFS2	Renewable Fuels Standard 2
ROW	Rest of the world
SCM	Simple Climate Model
TNH	Temperate Northern Hemisphere
TSH	Temperate Southern Hemisphere
WTW	Well-to-wheel

Chapter 1: Introduction

Energy from renewable sources, such as biomass, wind, and solar radiation, increasingly contribute to the world's energy production, which mainly consists of coal, oil, and natural gas. While the fossil fuel (FF) contribution of global power generation has declined from 70% to 66% during 1980-2010, the biomass, wind, and solar contribution has increased 25-fold from 0.2% to 5% over the same period (*1*). This energy transition is the result of several factors including policy incentives and declining costs of renewable energy technologies (*2*, *3*). Policy incentives, such as renewable energy mandates, are partially driven by a wide scientific consensus that rising GHG concentrations associated with FF use pose a threat to global climate and the welfare of future generations (*4*, *5*). Evaluating the effectiveness of alternative energy transition pathways to reduce GHG emissions is important for efficiently allocating economic resources and designing energy and climate policy (*6*).

Life cycle assessment (LCA) is a tool, which can be applied to quantify the GHG emissions of alternative forms of energy, and it is increasingly used to inform policy-makers (*7*–*10*). LCA has traditionally been used to support complex environmental performance evaluations of emerging technologies due to its ability to represent metrics like GHG emissions at high resolution (*11*–*13*). A major benefit lies in producing a systematic inventory of the product discharges to the environment in every life cycle stage, e.g., material extraction, processing, use, and waste management (*14*). For instance, numerous LCA studies document the technology specific emissions during these stages and point out potential for emissions reductions, e.g., in the case of biofuels (*9*). While these emissions inventories can be extended to estimate impact measures, such as atmospheric RF as an indicator for temperature change, most LCA studies limit their analysis to emissions.

Several studies in 2009 and 2010 focusing on the climate implications of substituting corn ethanol for gasoline suggest that using GHG emissions as an evaluation metric underestimates ethanol's RF impacts, thereby overestimating (i) ethanol's climate benefits over gasoline and (ii) the effectiveness of U.S. policies mandating and subsidizing ethanol (*15*–*17*). U.S. corn ethanol production was 13.3 Bgal in 2012 (6.4% energy equivalent of U.S. gasoline production), and the U.S. Environmental Protection Agency's (EPA) Renewable Fuel Standard 2 includes a mandate of corn ethanol increasing to 16 Bgal in 2022 (*7*, *18*, *19*). The disagreement in ethanol's perceived climate benefits between the life cycle GHG inventory and RF impacts arises from ethanol's land use change (LUC) induced GHG emissions. This emissions source, which appears to be the largest single contributor to corn ethanol's life cycle GHGs, may be caused, directly or indirectly, by providing agricultural land for the biofuel feedstock (*20*, *21*). Emissions from LUC occur over a short period at the beginning of the life cycle (*15*), which may span a decade or more. Thus, the early LUC emissions may cause a greater amount of cumulative radiative forcing (CRF) over the next decades than the remaining life cycle emissions. Some estimate that the traditional metric used for policy support, i.e., GHG emissions per unit of produced fuel, underestimates RF impacts by

10-90% due to emissions timing, thereby whitewashing the effectiveness of biofuel development policies (*15–17*). This is particularly policy relevant because (i) most LCAs find little – if any – GHG benefits of corn ethanol over gasoline (*7*, *8*, *20*, *22*), and (ii) any renewable fuel supported by the 2007 Energy Independence and Security Act (EISA) must reduce GHG emissions by at least 20% over gasoline (*23*).

Chapter 2 of this dissertation revisits the above studies regarding the relevance of emissions timing for U.S. biofuel policies. It builds upon existing models to quantify RF impacts across the corn ethanol life cycle, and addresses several additional issues. An earlier version of this chapter was published in 2011 (*24*). Three research questions were formulated.

1) What is the magnitude of the emissions timing effect, i.e., underestimated RF impacts in traditional LCA, after addressing the limitations of previous models (described in more detail in section 2.1.3)?
2) By what percentage does adding the LUC induced emissions timing factor to the traditional LCA GHG metric reduce the probability of corn ethanol meeting the EISA target?
3) What is the absolute RF impact of U.S. corn ethanol production and use including LUC, and what are the implications for climate policy?

Chapters 3-5 address the need to better understand uncertainties in CH_4 emissions from the NG life cycle. CH_4 emissions are a key parameter for estimating climate implications of switching from coal to a greater share of natural gas (NG) production, which some consider a low cost transition fuel towards decarbonizing the economy (*25*). Two major reasons prompted this change of research focus. First, absolute potential climate impacts of transportation biofuels are small compared to NG, which contributed 1.1% and 26% to U.S. primary energy consumption in 2012, respectively (*1*). Second, given the successful tapping of unconventional NG sources in North America, NG production and associated GHG emissions are expected to grow rapidly in the near future (*26*). While many studies indicate associated reductions in GHG emissions (*27–34*), the actual RF impacts associated with the coal-to-gas transition are less clear. Some climate modeling studies suggest the distinct possibility that replacing coal with NG could lead to temporarily (decadal time scales) higher global warming due to increased emissions of more potent CH_4 in addition to greenhouse effects from CO_2 and sulfate aerosols (*35*, *36*).

Accurately quantifying the NG FE rate, i.e., the percentage of produced NG – mainly CH_4, but also C_2H_6 and higher hydrocarbons – released to the atmosphere, intentionally and unintentionally, is challenging due to the size and complexity of the NG industry, which includes extraction, processing, transport, and distribution. Recent U.S. estimates suggest that current NG FE could be as high as 8% and 6%, from shale and conventional NG, respectively (*28*). Other studies indicate possible NG FE over 9% from conventional sources several decades ago (*37*, *38*). In contrast to bottom-up studies such as the LCA studies discussed above that rely largely on emissions data from the NG industry, the objective of this research was to analyze possible ranges of the global

average NG FE rate based on atmospheric measurements recorded since 1984. This work draws upon top-down Bayesian inversion modeling approaches developed over the past two decades, which are used primarily to understand all sources and sinks of atmospheric CH_4 and other species. The traditional inversion process to model CH_4 includes (i) employing external *prior* best emissions estimates of each CH_4 or C_2H_6 source, (ii) simulating their atmospheric mixing using an emissions transport model, and (iii) constraining the *priors* using atmospheric measurements of CH_4 or C_2H_6 concentrations and $\delta^{13}C$-CH_4 from a global observation network (*39*). While previous studies have estimated absolute CH_4 emissions from NG and oil production (*39–44*), NG FE rates were not quantified. The main objective of this CH_4 and C_2H_6 modeling study is to quantify the range of NG FE rates over time that are consistent with atmospheric measurements. Four research questions were formulated, which will be addressed in Chapters 3-5.

1) What are the best estimate NG FE rates over the past two decades globally according to CH_4, C_2H_6, and $\delta^{13}C$-CH_4 tracers, and is there a trend over time?
2) What is the upper bound of global NG FE given the lower bound of all other CH_4 and C_2H_6 sources, and what are the *prior* assumptions (e.g., FF emissions factors and hydrocarbon composition) underlying this bound?
3) Can literature estimates of natural CH_4 and C_2H_6 seepage be reconciled with global measurements, and how does this affect the NG FE estimates?
4) What are the policy implications of global atmospheric observational constraints on NG FE rates?

Chapter 3 provides a more detailed description of the research motivation, objectives, and a literature review. Chapter 4 develops a transparent, global *prior* CH_4 and C_2H_6 emissions inventory needed to interpret NG FE rates based on inversion results. This inventory constitutes a bridge between the LCA and atmospheric inversion communities, which has not existed thus far. Chapter 5 describes (i) the global inversion box-model developed here, (ii) the existing 3D transport model employed to support the box-modeling, and (iii) the results. Chapter 6 concludes this dissertation with a summary of chapters 2-5, including main findings and implications for energy and climate policy.

Chapter 2: Relevance of Biofuel Greenhouse Gas Emissions Timing

2.1 Introduction

Corn ethanol is currently the prevailing biofuel in the U.S. Domestic production increased from 1.8 to 10.6 Bgal/yr over the past decade (*45*). The federal government established a mandate to blend gasoline with corn ethanol, which will increase to 15 Bgal/yr (7% of U.S. gasoline consumption) by 2016 (*23*). Before 2012, corn ethanol was subsidized at $0.45/gal (*2*, *23*). These policies are aimed towards mitigating climate change impacts from the transportation sector, and to overcome the dependence on foreign oil among others (*2*).

Currently, the perceived effectiveness of corn ethanol or other biofuels to mitigate climate change is mainly based on GHG emissions. Policy makers rely increasingly on greenhouse gas accounting (GHGA) or LCA to support complex performance evaluations of emerging technologies (*14*). However, as the accounting procedures within LCA are improved to provide a better representation of reality, LCA results sometimes change significantly (*7*). This can dramatically alter prior results and beliefs, influencing policies, and amending decisions that determine billions of dollars of public and private investments.

2.1.1 **Estimating corn ethanol life cycle GHG emissions**

Greenhouse gas accounting or LCA of corn ethanol and other biofuels includes establishing a GHG inventory of all emissions that occur during the fuel's production and use, and is usually compared to a gasoline baseline (*7*). Typically, emissions are accounted for during all life cycle stages including agricultural feedstock production, conversion into liquid fuel, combustion, transportation of the feedstock and the fuel, and ancillary processes such as production of fertilizer and machinery (*22*). These are sometimes referred to as well-to-wheel (WTW) emissions (*22*). Other effects may be taken into account including emissions from LUC (*21*).

Land use change emissions can occur directly when the biofuel feedstock displaces other biomass (e.g., previous crops or forests). The above- and belowground biomass may be burned and rapidly decayed, respectively. Emissions from LUC may also be released indirectly through market-mediated effects (see Appendix A for details). Throughout this paper, only WTW emissions and LUC emissions (direct plus indirect LUC) are distinguished.

Quantifying biofuel emissions from LUC is challenging, which leads to high uncertainty in LUC estimates. Literature values for LUC from corn ethanol range from 500 to 5,100 g CO_2-equivalent (CO_2e) per MJ of additional biofuel production capacity (*7*, *20*, *21*, *46*, *47*). Assuming a 30-year life cycle means 17 to 170 g CO_2e/MJ emissions per year.

EPA's mean estimate of corn ethanol WTW emissions is 46 g CO_2e/MJ emissions per year (*7*). Thus, LUC may increase the life cycle GHG emissions by 37 to 370%. EPA has estimated mean LUC emissions at 28 g CO_2e/MJ per year (*7*).

In LCA, the sum of all WTW emissions is weighted by their GHG specific global warming potentials, and LUC emissions are added after dividing by the total amount of fuel produced over the life cycle (*7*). The time horizon of the life cycle can be determined by the corn ethanol project time (*22*). This is the number of years the feedstock, i.e., corn, is produced on a given piece of land before the land is used for other purposes, e.g., conversion to pasture. A time horizon of 30 years is often used in LCA, but this is not an empirical value (*21*). In 2010, EPA found the mean life cycle GHG emissions to be 74 and 93 g CO_2e/MJ for corn ethanol and gasoline, respectively (see Appendix A for details) (*7*). This translates into a GHG balance of -20%. The GHG balance, the prevailing performance metric regarding climate change mitigation, describes how much more or less GHG emissions corn ethanol releases relative to gasoline (*22*).

2.1.2 **Emissions timing and the GHG balance**

Emissions from LUC occur in the first years of the life cycle until all new acreage from increased feedstock demand is converted to corn and indirect LUC has ceased through market equilibrium (*7*). A hypothetical emissions time profile in Figure 1 describes the CO_2e emissions released at any year of the fuel's life cycle for the functional unit, i.e., 1 MJ of fuel. The profile is illustrative for EPA emissions data, and includes emissions from land clearing, soil carbon, and foregone sequestration (*7*).

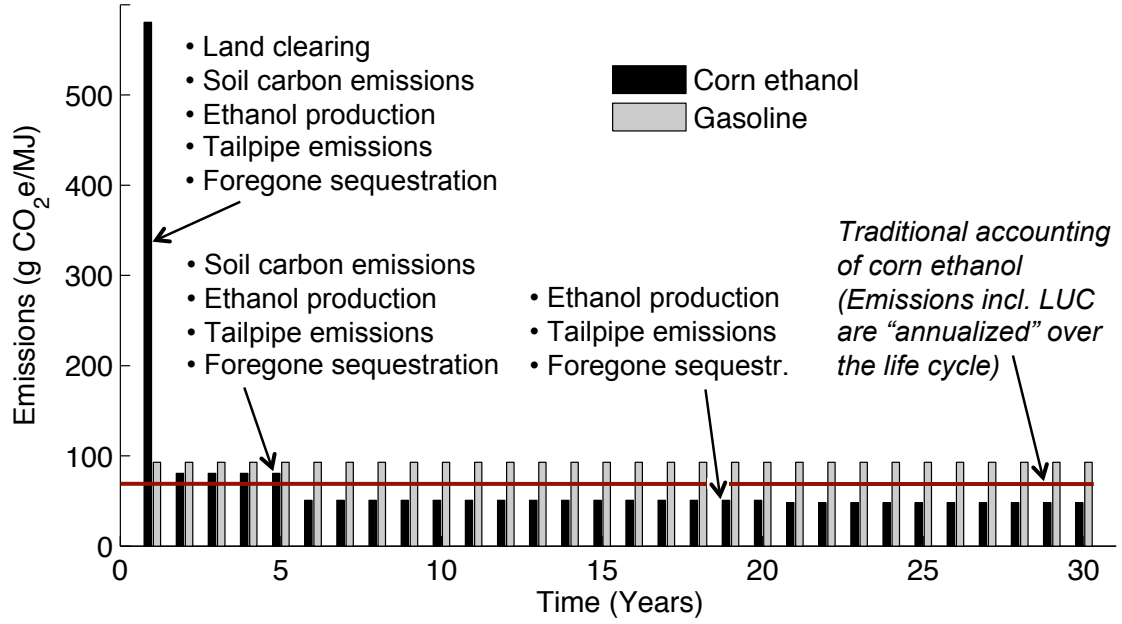

Figure 1: Illustrative GHG emissions of corn ethanol and gasoline (all EPA (*7*) data). Ethanol emissions are shown for both the emissions time profile and annualized emissions (traditional GHGA).

Section 2.2 provides details regarding the methods to estimate the emissions time profile. The GHG emissions of corn ethanol follow a different time pattern than those of gasoline because, in the model presented here, gasoline has no analogous large up-front emissions. The construction of the ethanol factory or the oil refinery, which may also cause early GHG emissions, is not considered. The varying emissions release over the corn ethanol life cycle is referred to as the emissions time profile. It is important to note that the GHG balance used by EPA and others does not address emissions timing because it allocates all LUC emissions evenly over time, i.e., emissions are annualized over the life cycle. EPA also considered a scenario in which emissions are discounted at 2% over 100 years to address emissions timing (see discussion in Appendix A).

2.1.3 **The significance of early emissions**

Releasing GHG emissions today may be more harmful than emissions in the future. The rationale for this statement was described previously (*15–17*), and is summarized here briefly. First, climate change poses the risk of potentially irreversible damages to humans and ecosystems, which calls for policies to mitigate climate change impacts from GHG emissions over the next decades (*4*). As a result, reducing GHG emissions today may be more valuable than a reduction in the future under certain assumptions about the social cost of carbon over time. Substituting gasoline with corn ethanol, however, is conceptually the opposite, i.e., reducing GHG emissions in the future, but increasing emissions today due to LUC.

Second, atmospheric CO_2 decays over a period of several centuries (*48*), causing emissions today to contribute to the greenhouse effect several hundred years in the future. Consider the difference between emitting CO_2 today and in 20 years. Over the next century, the cumulative climate impact of the former is greater than that of the latter. The GHG balance approach masks this difference in timing, possibly conveying fewer impacts than those actually occurring. As a result, ignoring emissions timing may whitewash the effectiveness of policies that support corn ethanol or other biofuels.

Several authors have estimated that the GHG balance undervalues the climate impacts of corn ethanol by 9-89% due to emissions timing using different ways of modeling RF (*15–17*). GHGs contribute to positive RF (and increasing global surface temperature) by absorbing some of the radiation and reflecting it back into the atmosphere (*48*).

These studies described new methods to address the concept of emissions timing in biofuel LCA. However, several model limitations oversimplify real world processes, which could affect the results significantly. First, marginal emission impulses, rather than actual emissions from annually varying ethanol quantities were used to estimate RF. The former implies that all LUC from the 15 Bgal/yr occurs instantaneously whereas in reality RF is a function of increasing ethanol production over time until full capacity is reached. Second, estimating RF from LUC emissions only, i.e., excluding WTW emissions (*16*), overlooks the fact that the CRF benefits from lower WTW ethanol emissions (relative to gasoline) increase with longer impact time frames. Third, estimating the influence of timing as the difference between the CRF balance (CRF of ethanol emissions relative to gasoline) and the GHG balance (based on life cycle

emissions) (*15*, *17*) may be misleading because of different underlying assumptions in the two metrics. For example, the CRF balance distinguishes two time horizons (the 30-year corn ethanol life cycle period and the year until which ethanol may be produced in the U.S.) while the GHG balance assumes both time horizons to be the same. Instead of isolating the influence of emissions timing, this approach quantifies the combined effect of emissions timing and assumed fuel quantities.

2.1.4 **Research objective**

This work builds on the methods and arguments established by O'Hare *et al.*, Kendall *et al.* and Levasseur *et al.* (*15*–*17*) through quantifying the influence of emissions timing on the RF impacts of corn ethanol. Realizing the limitations of the existing methods, a RF model was developed to address these limitations, and to update the results of the emissions timing effect. By modeling the total RF from actual fuel quantities, the order of magnitude effect of corn ethanol induced LUC emissions on global mean surface temperature (ΔT_s) was also estimated. The specific research questions are as follows:

1) What is the magnitude of the emissions timing effect, i.e., underestimated RF impacts in traditional LCA, after addressing the above-mentioned limitations of previous models?
2) By what percentage does adding the LUC induced emissions timing factor to the traditional LCA GHG metric reduce the probability of corn ethanol meeting the EISA target?
3) What is the absolute RF impact of U.S. corn ethanol production and use including LUC, and what are the implications for climate policy?

2.2 Methods

This analysis follows the general procedure in Figure 2. In the first step, literature GHG inventories are used to generate the emissions time profile of corn ethanol and gasoline (described in section 2.2.1). The time profile (g CO_2e/MJ/yr) is combined with the energy contents of annual fuel consumption (MJ/yr) to estimate the total GHG emissions for each year by using the same amounts of energy for corn ethanol and gasoline.

Next, the RF model (section 2.2.2) estimates RF for each fuel as a function of total GHG emissions over time using the emissions time profile (upper branch). CRF for each fuel is estimated over different impact time frames. The CRF balance of corn ethanol (CRF of ethanol relative to gasoline) is estimated analogously to the GHG balance (GHG emissions of ethanol relative to gasoline).

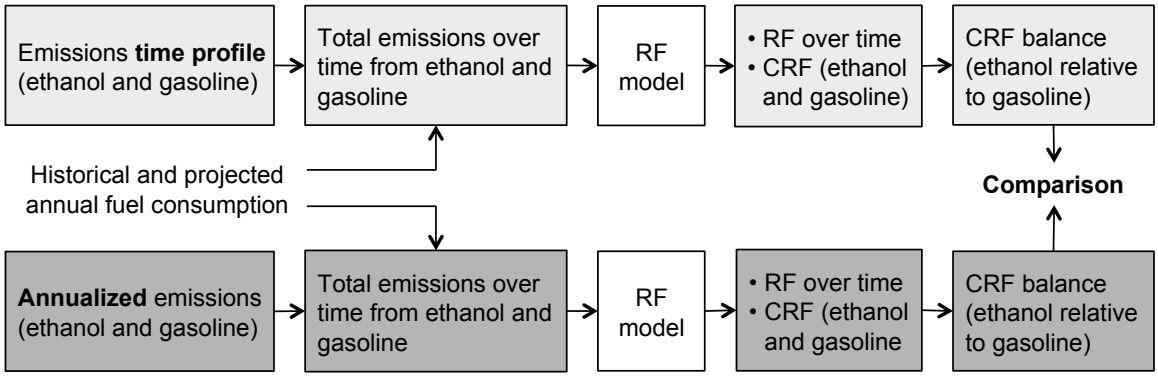

Figure 2: Overview of the analytical procedure of this research to quantify the influence of emissions timing on the RF impact of corn ethanol over a given time horizon. Fuel consumption starts at 1.8 Bgal/yr in year 2001, and reaches maximum capacity of 15 Bgal/yr in 2016 (see Appendix Figure A - 1).

Then, the simulation is repeated, but using the annualized emissions (lower branch), which is implied in traditional LCA (see Figure 1). Finally, the difference in CRF between the simulations for the emissions time profile and the annualized LUC emissions is used as a measure to quantify the influence of timing.

2.2.1 **GHG emissions inventory and time profile for corn ethanol and gasoline**

Radiative forcing is estimated using GHG emissions from four different LCA studies (*7*, *20*, *46*, *47*). The EPA study was used in support of the Renewable Fuels Standard 2 (RFS2). Emissions vary significantly among the studies for both LUC and WTW, i.e., the remaining life cycle emissions. A summary of data including the disaggregation of CO_2e emissions into CO_2, CH_4, and N_2O is presented in Appendix A. EPA data (*7*) was used as the base case (see Appendix Table A - 1). Emissions estimates for LUC from Hertel *et al*. and Tyner *et al*. (*46*, *47*) are used to test the sensitivity of the model. The results of a stochastic analysis of LUC uncertainty in Plevin *et al*. (*20*) were used to estimate upper and lower bounds of possible LUC values (see Appendix Table A - 2).

The allocation of emissions over time on a per MJ basis generally coincides with earlier work (*7*, *15*), and the results are summarized in Figure 1. However, absolute emissions are modeled (see Appendix Figure A - 1), which allows representing the actual ramp-up in annual corn ethanol production. The annual volume increases are used as a proxy for absolute LUC occurrence over time, which distributes LUC more evenly over time compared to previous work. It is still a conservative estimate since modeled upfront emissions may be shifted more towards future years in reality, thereby reducing the share of absolute upfront emissions (see Appendix A for details). Appendix A also provides details regarding the choice of the biofuel project time, the sources of LUC emissions, the gasoline emissions over time, and the calculation of total emissions from a given fuel volume in a particular year. In the base case, fuel use is simulated according to historical and projected corn ethanol production (*45*) using data from 2001 to 2030 (see Appendix Figure A - 1). Annual fuel use increases to reach the federal

mandate of 15 Bgal/yr in 2016 (Renewable Fuel portion of the RFS2), and is assumed to remain constant thereafter (*15*).

An accelerated LUC scenario was also calculated, which simulates ethanol production to reach 15 Bgal/yr in the first year, rather than over the course of 16 years, which causes all LUC to occur in one year. This reflects the hypothetically highest increase in GHG concentration from LUC in a single year. In this case, no more LUC occurs in the following years assuming that corn is grown on the lands already converted to agriculture in year 1. Accelerated LUC is implied in all prior emissions timing analyses (*15–17*).

2.2.2 **Radiative forcing over time**

A simplified RF model was developed based on the BERN model (*49*) and recommendations given in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) (*48*). BERN is a simple climate model (SCM) used by IPCC. It lacks the high resolution of the complex (atmosphere-ocean general circulation) models, which are used to study the fundamental processes governing climate change, and to allow projections of regional scale changes (*50*). However, BERN and other SCMs "replicate the global scale average behavior of complex models" (*50*), which was deemed sufficient for the modeling presented here. The model estimates RF and the magnitude of the change in global mean surface temperature ΔT_s given the emissions estimates above.

The RF model tracks the net radiative activity of the emissions over time. The only emissions considered in the RF model are from WTW (corn ethanol and gasoline) and LUC (corn ethanol only). All other anthropogenic GHG emissions are omitted because only the RF effects from fuel use are of interest. RF is estimated for each GHG (CO_2, CH_4, and N_2O) and fuel (corn ethanol and gasoline). In each impact year t (1 to 100), RF is the result of the emissions released in t and the non-decayed fraction of emissions that occurred in all previous release years r (1 to 100). For each GHG and fuel, an r by t matrix RF_{rt} is calculated using Eq. 1.

$$RF_{rt,Fuel,GHG} = E_{rt,Fuel,GHG} * \Delta C * RE_{rt,GHG} * D_{rt,GHG}, \quad \textbf{Eq. 1},$$

where $E_{rt,Fuel,GHG}$ is the matrix of the annual emissions release of a GHG (g), ΔC is the change in atmospheric concentration of the GHG due to the release (ppmv/g), $RE_{rt,\,GHG}$ is the radiative efficiency matrix, i.e., the instantaneous RF of a unit increase in atmospheric GHG concentration (W/m^2/ppmv), $D_{rt,\,GHG}$ is the decay matrix of the GHG (unit-less), and $*$ denotes element-by-element multiplication (as opposed to standard matrix multiplication). The details of the above matrices are described in Appendix A.

In order to calculate the total RF for each fuel and GHG, the rows r in the RF matrix are added to yield an RF vector with t elements for each impact year t (Eq. 2). Next, the three RF vectors representing each GHG are added to yield the total RF vector for each fuel (Eq. 3).

$$RF_{Fuel,GHG}(t) = \left[\sum_{r=1}^{100} RF_{rt,Fuel,GHG}(t=1), \sum_{r=1}^{100} RF_{rt,Fuel,GHG}(t=2), \cdots \sum_{r=1}^{100} RF_{rt,Fuel,GHG}(t=100)\right] \quad \textbf{Eq. 2}$$

$$RF_{Fuel}(t) = \sum_{GHG} RF_{Fuel,GHG}(t) \quad \textbf{Eq. 3}$$

Ethanol's CRF balance is estimated analogously to the GHG balance, i.e., relative to gasoline. In contrast to the GHG balance, however, the CRF balance is a function of impact time t (Eq. 4).

$$CRF\ balance(t) = \frac{\sum_t RF_{Ethanol}(t) - \sum_t RF_{Gasoline}(t)}{\sum_t RF_{Gasoline}(t)} \quad \textbf{Eq. 4},$$

This analysis defines the emissions timing factor ETF as the amount (in percent) by which the GHG balance underestimates actual RF impacts over a given t (Eq. 5).

$$ETF(t) = CRF\ balance_{Time\ profile}(t) - CRF\ balance_{Annualized}(t) \quad \textbf{Eq. 5},$$

According to IPCC (*48*), ΔT_s is estimated from RF in Eq. 6:

$$\Delta T_s = \frac{\lambda_{2xCO_2}}{\Delta F_{2xCO_2}} \times (RF_{Ethanol} - RF_{Gasoline}) \quad \textbf{Eq. 6},$$

where λ_{2xCO2} is the climate sensitivity, i.e., the increase in global mean surface temperature from a doubling of atmospheric CO_2 concentration from pre-industrial levels (from 280 ppmv to 560 ppmv). IPCC estimates the likely range of λ_{2xCO2} between 1.5 and 4.5 °C (*48*). ΔF_{2xCO2} is the RF response from this doubling of CO_2, commonly assumed to be 3.71 W/m^2 (*48*).

2.3 Results

Radiative forcing from 30 years of corn ethanol and gasoline use was estimated over a 100-year impact time frame based on literature life cycle emissions (including LUC) and fuel use scenarios in the U.S. The base case represents LUC distributed over 16 years (2001-2016) as annual ethanol production increases over this period. This represents a phased-in agricultural land approach to increasing U.S. ethanol production, thereby illustrating historical and projected annual ethanol volumes in the U.S. from 2001 to 2030.

2.3.1 **Radiative forcing over time**

Figure 3 shows RF from gasoline and corn ethanol for both annualized emissions and the time profile using base case estimates (EPA). Ethanol production increases to 15 Bgal/yr in 2016 according to historical and projected annual ethanol volumes in the U.S. (*45*). All three curves display increasing RF until year 30 because the net effect of the

GHG emissions release over time is an increase in atmospheric GHG concentration. After 30 years, RF declines in all three curves because GHG emissions from fuel use cease in the base case. However, a significant RF effect remains after 100 years due to the long atmospheric lifetime of the GHGs.

Initial (year 1) RF is highest for the corn ethanol time profile curve due to LUC. After about 10 years, the slope of the time profile curve decreases gradually because 80% of the overall LUC has occurred. After about 20 years, the curves of the time profile and gasoline intersect, which means a RF reduction from substituting gasoline with ethanol thereafter. Over the first 15 years, RF from annualized corn ethanol is higher than gasoline despite lower life cycle emissions because modeled fuel volumes are initially low. As a result, annualized LUC emissions have an over proportionate leverage on overall emissions.

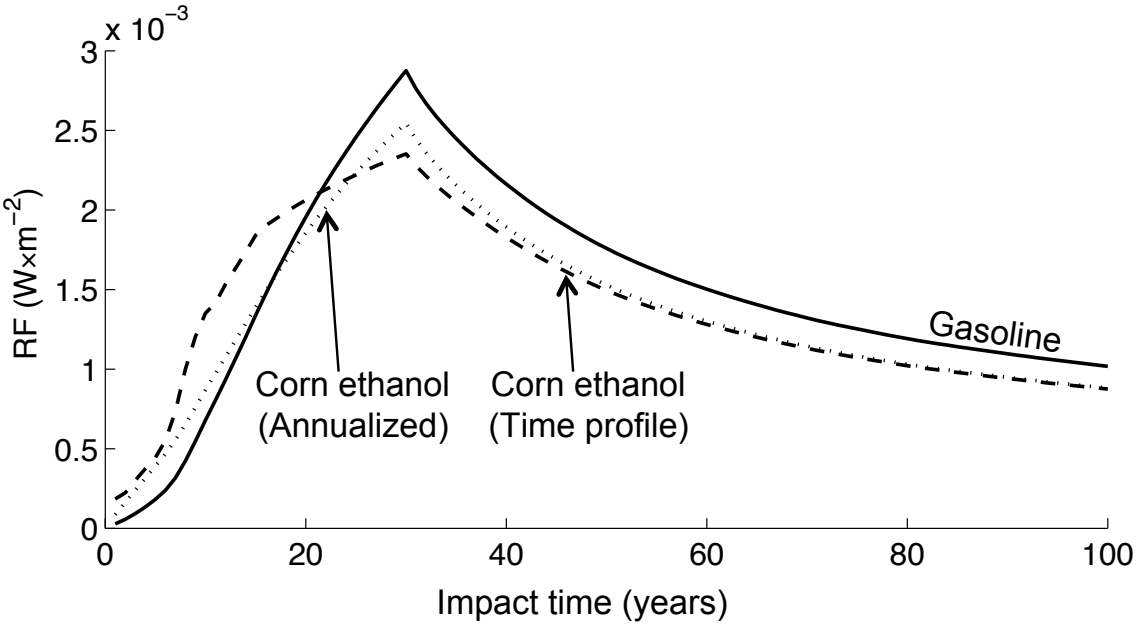

Figure 3: RF(t) of gasoline and corn ethanol (both annualized emissions and emissions time profile) using base case GHG emissions estimates (EPA (*7*), 30 years of production). See Appendix
Figure A - 2 and section 2.3.2 for a discussion of alternative production periods.

The fuel production increase at varying rates in the model contributes to the non-uniform slope over the first 30 years. For annualized corn ethanol, overall RF from LUC is distributed evenly over 30 years, creating a more uniform slope. The same is true for gasoline, which has no LUC effects in the model. RF and ΔT_s from electricity generation (*51*, *52*) and the transportation sector (*53*) was also estimated as reference points of what increases may be considered "high". The details provided in Appendix A show that the maximum effect of 15 Bgal/yr corn ethanol on RF and ΔT_s is less than 1% that of U.S. electricity consumption. Reducing U.S. gasoline consumption by 10% decreases

RF and ΔT_s by eight times the amount that RF and ΔT_s increases from 15 Bgal/yr corn ethanol.

2.3.2 **CRF balance**

Figure 4 shows the CRF balance of corn ethanol, i.e., the ethanol-gasoline ratio of RF integrated under the curves in Figure 3 for both the annualized emissions and the time profile over different impact time frames. The CRF provides a measure of ethanol's life cycle performance based on CRF up to a given impact year, as opposed to the traditional GHG balance.

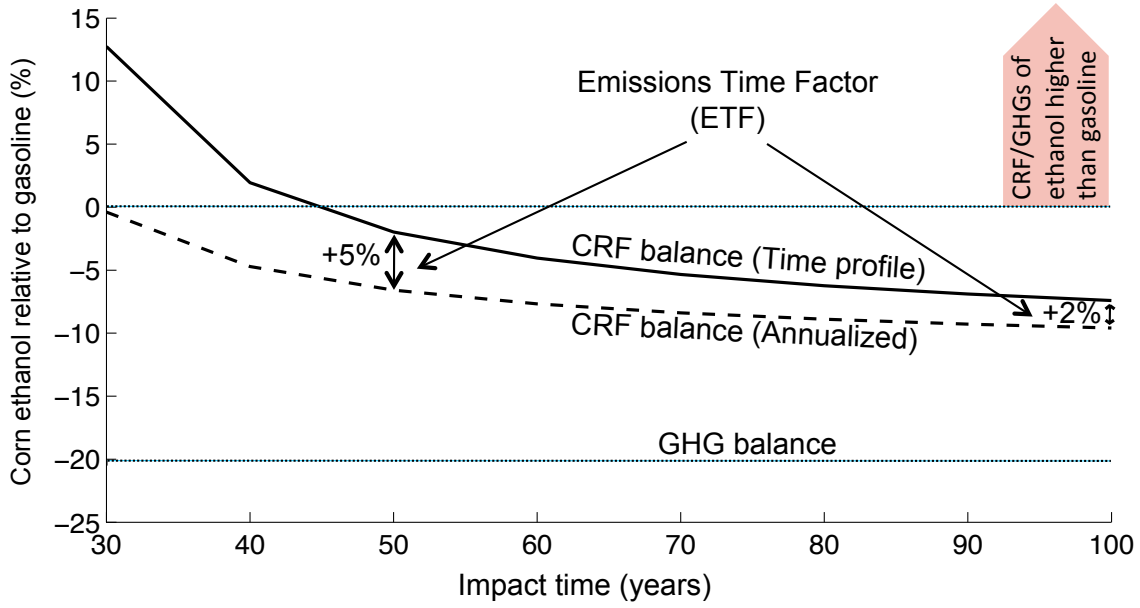

Figure 4: CRF balance of corn ethanol (both annualized emissions and emissions time profile) and ETF using base case GHG emissions estimates (EPA (*7*)). The GHG balance of corn ethanol is also depicted for reference. A lower CRF balance means lower RF impacts of corn ethanol relative to gasoline.

Both CRF curves decrease with higher impact time frames as the RF benefits from relatively lower WTW emissions from ethanol materialize over time. The curve for the time profile is above the annualized emissions because the LUC emissions of the former cause RF over a longer period within 100 years. The difference between both curves is ETF, which indicates by how much the RF impacts are underestimated when emissions timing is ignored. In the base case, ETF is 2 and 5 percentage points over 100 and 50 years, respectively, and increasing for shorter time frames. Modeling relatively low WTW emissions from the RFS2 (*7*) offsets some of the ethanol CRF increases due to modeling a ramp-up in ethanol production (see Appendix Figure A - 5).

The horizontal line at -20% in Figure 4 represents the GHG balance – the commonly used GHGA metric used to compare biofuels with gasoline – as a reference to show the gap between CRF and GHG balance. Previous analyses (*15*, *17*) quantified emissions timing as the gap between the GHG balance and the CRF balance (time profile). In

contrast, this method isolates the influence of (a) emissions timing due to significant upfront emissions and (b) anticipated fuel quantities on corn ethanol's climate performance. The latter is an independent policy issue, which should be addressed regardless of how harmful early emissions are to climate.

The CRF balance of the time profile is 12% to -7% (base case), which indicates that corn ethanol is 32 to 13 percentage points less beneficial than what the GHG balance conveys. Note that both the GHG balance and CRF balance shown in Figure 4 assume the same fuel production period of 30 years used in earlier studies (*7*, *15*–*17*, *21*, *46*, *47*) after which no more ethanol is produced to further offset early LUC emissions. Figure 5 illustrates the influence of alternative production periods on CRF. Further differences in implicit assumptions between GHG and CRF balance are illustrated in Appendix Figure A - 5.

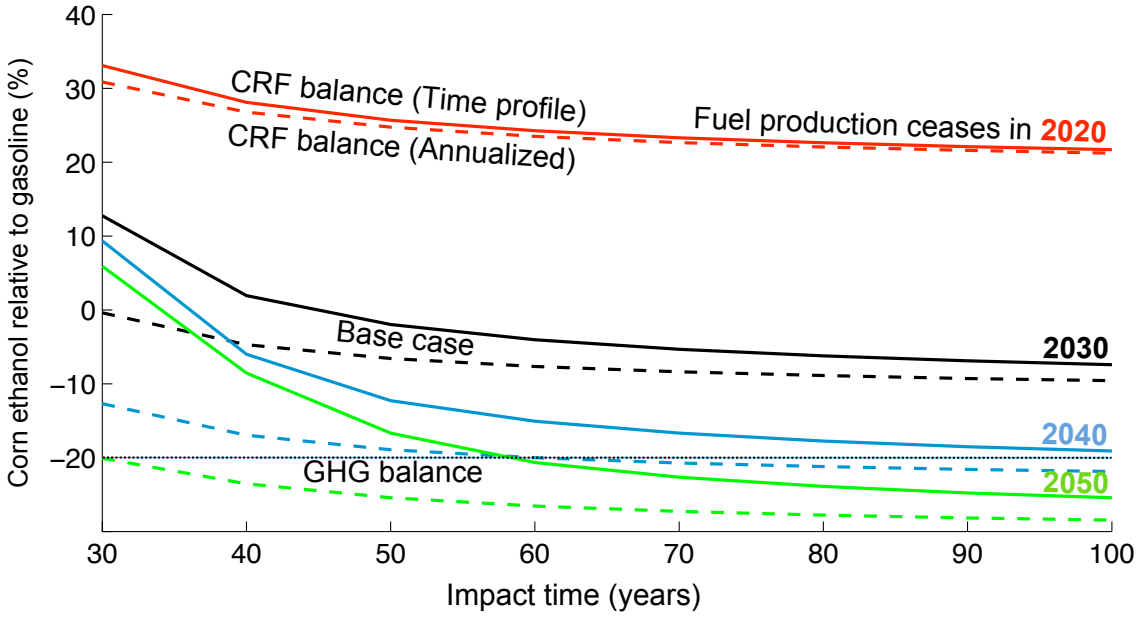

Figure 5: CRF balance of corn ethanol (both annualized emissions and emissions time profile) for varying fuel production periods using EPA (*7*) GHG emissions estimates.

Fuel production was simulated over 20, 30 (base case), 40 or 50 years, i.e., fuel production from 2001 to 2020, 2030, 2040, and 2050, respectively. As illustrated in Figure 5, longer fuel production allocates LUC over a greater amount of fuel volume, which improves ethanol's CRF balance. In fact, the CRF balance of the 2050 scenario is below the -20% mark (GHG balance) for 60 or more years of impacts. The model assumes no additional LUC after 2016 when the maximum capacity of 15 Bgal/yr is reached. For production periods less than 30 years, the CRF is shifted to higher values. As a result, the difference between the CRF balance and the GHG balance captures not only emissions timing, but also the influence of produced fuel volumes. Note that the fuel production period has a significantly higher influence on ethanol's climate performance than ETF. The model shows ETF is mainly a function of the scale of

assumed LUC, the impact time frame, and the production period, and ETF is fairly insensitive to other parameters. For instance, the sensitivity analysis in Appendix Figure A - 6 shows that ETF is only 1-6% for the relatively low LUC estimates in Tyner *et al.* (*47*), but it increases to 3-15% for higher LUC in Hertel *et al.* (*46*).

2.3.3 Influence of emissions timing on the GHG balance

The emissions timing factor may be added to the GHG balance of corn ethanol to reflect the influence of emissions timing. The value of ETF depends on the impact time frame chosen by the policy-maker. Figure 6 shows the GHG balance for several impact time frames.

The boxes with error bars pertain to LUC uncertainty values (95% confidence interval and minimum and maximum values, respectively) from Plevin *et al.* (*20*). The 95% confidence interval of -16% to 26% in Figure 5 compares to -32% to -7% in EPA (*7*). The triangles correspond to EPA base case estimates, which do not fall within the 95% confidence interval. According to this data, there is a less than 2.5% chance that corn ethanol will meet the EISA target of a 20% reduction in GHG emissions relative to gasoline, which confirms the conclusions of Mullins *et al.* (*9*).

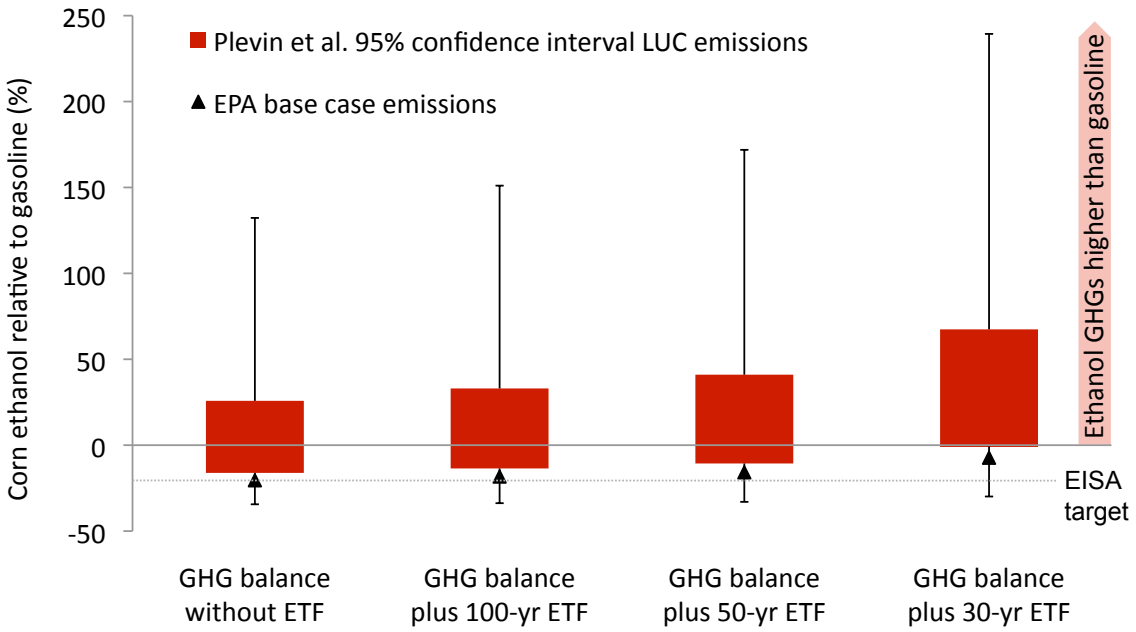

Figure 6: GHG balance of corn ethanol with and without ETF for several impact time frames. Illustrated estimates are EPA (*7*) base case and Plevin *et al.* (*20*). LUC uncertainty with 95% confidence intervals and min/max values (error bars). The EISA target is a 20% reduction in GHG emissions relative to gasoline.

Considering either Plevin *et al.* or EPA estimates (*7*, *20*), corn ethanol emissions uncertainty (LUC and fuel production period) has a much greater influence on the GHG balance than the ETF, which is triggered by LUC. However, emissions timing shifts the

95% confidence interval towards higher corn ethanol emissions relative to gasoline. That is, the chance of meeting the EISA target is reduced even further.

As mentioned above, the period over which corn ethanol will be produced in the U.S. also influences the performance of corn ethanol, which is not captured in the GHG balance. This effect is displayed in Appendix Figure A - 7. It shows that the CRF balance drops slightly below the GHG balance (in favor of ethanol) if ethanol is produced until 2050, but increases significantly if production ceases in 2020.

2.4 Conclusions

This study was designed to understand if and how the LCA of biofuels can accurately account for the fact that LUC emissions occur early in its life cycle. Adaptations to current LCA may be appropriate if traditional GHGA underestimates the RF impacts from biofuel emissions including LUC. The emissions timing effect of biofuels has been investigated previously. This research further develops the methods to quantify the timing effect in three ways. First, this study accounts for the annual ethanol production that increases over time, thereby distributing LUC over more than a decade. Second, the analysis accounts for total RF of both WTW and LUC emissions, rather than LUC alone. This takes into account the increasing CRF benefits of lower ethanol WTW emissions (than gasoline) at higher impact time frames. Finally, the difference between the CRF balance of the time profile and the annualized emissions (traditional LCA) was calculated, rather than the difference between the CRF balance of the time profile and the GHG balance. The former is necessary to isolate the effect of emissions timing from the effect of assumptions about annually varying fuel volumes.

In this model, the influence of emissions timing on RF is significantly smaller than 9-89% suggested by previous studies (*15*–*17*). In the base case, ETF is 2, 5, and 13% for impact time frames of 100, 50, and 30 years, respectively. Given a corn ethanol GHG balance of -20% (*7*), i.e., corn ethanol emitting 20% less GHGs than gasoline, accounting for time may increase the GHG balance to -18%, -15%, and -7% with respect to the above impact time frames. The results are sensitive to the uncertainty in the LUC emissions inputs as well as the time period over which the fuel is produced.

From a biofuel LCA perspective, emissions timing adds little to our understanding of the climate impacts of corn ethanol because the uncertainty inherent in the LUC estimates (that triggers ETF) contributes one to two orders of magnitude more to life cycle emissions uncertainty than ETF. To complicate matters, it may be inappropriate to select a single ETF for a given impact time frame because ETF depends significantly on the time period over which corn ethanol would be produced. The fuel production period is different from the life cycle period. The former is the period over which ethanol is produced in the U.S., irrespective of the location, whereas the latter is the period over which the feedstock is produced on a given piece of land (usually 30 years). The longer corn ethanol is produced at the mandated 15 Bgal/yr beyond 2016, the greater the CRF benefits from lower WTW GHG emissions. This assumes that LUC from corn ethanol is reduced dramatically for fuel produced at constant volumes, thereby avoiding an

increasing demand for agricultural land. It does not account for the possibility of accumulating carbon on land dedicated to corn after reducing ethanol and corn production.

From a policy maker's perspective, emissions timing is more important because, in the model, ETF increases the chance of corn ethanol missing the EISA target of 20% GHG reduction relative to gasoline. According to the EPA LCA, the mean (and median) GHG balance of corn ethanol is just below -20%, i.e., the probability that the target is met is just over 50%. Since EPA estimates that corn ethanol barely meets the EISA target (based on its mean value), even a small a skew in the distribution can significantly reduce the probability of meeting this target. A stochastic model was developed to quantify the extent to which adding the LUC induced ETF to the GHG balance reduces the probability to meet EISA (see methods and results in Appendix A). The analysis suggests that ETF reduces this probability from 53% to 7-29% depending on the chosen impact time frame. Thus, switching from the *mean GHG balance* to *probability of meeting the target* amplifies the influence of ETF from 2-13 percent points to 24-46% in reduced probability of meeting the EISA target. Furthermore, ETF reduces the probability of corn ethanol emitting equal or less GHG emissions than gasoline from 99% to 71-98%.

From a climate change mitigation perspective, the impacts from corn ethanol GHG emissions would be a very small contributor to reducing climate change as expressed by RF units, even when considering the relatively low EPA LUC emissions estimate. The model assumes that most of the emissions "cost" of corn ethanol has occurred over the last decade due to LUC effects from expanding ethanol production. This cost was estimated to be 0.0006 W/m^2 in the base case (15 Bgal/yr), which translates into a first order estimate of 0.0007 °C increase in ΔT_s. The model predicts a *reduction* in RF and ΔT_s at the end of the life cycle, which is in the same order of magnitude. It appears very unlikely that either temperature change (first positive, then negative) would entail increased or reduced damages of an order of magnitude that can actually be observed or quantified.

Chapter 3: Quantifying Natural Gas Fugitive Emissions – Motivation and Research Overview

This chapter provides the background for the second part of the dissertation, which aims at characterizing uncertainty of emissions of industry-wide NG FE. Throughout this work, NG FE refer to mainly CH_4 and C_2H_6 emissions from seals, equipment leaks, venting, and accidental leaks (*54*), which encompasses all industry segments and processes from NG production to end use. This chapter describes (i) the importance of the NG FE rate (in % of dry NG production and dry NG composition) from a climate and economic perspective, (ii) estimation methods and resulting NG FE ranges from the literature, and (iii) a conceptual overview of the approach taken here.

3.1 Importance of uncertainty in NG FE

Natural gas use is expected to increase significantly over the next decades both in the U.S. and worldwide due to an unprecedented expansion of unconventional NG production, specifically shale gas, coal-bed methane (CBM), and tight sands gas. In North America, proved NG reserves are estimated at 260 trillion cubic feet (Tcf) (*26*), and "remaining recoverable" unconventional NG resources are estimated at 3,200 Tcf (*55*). Global annual NG consumption is expected to increase from the current 110 Tcf to about 170 Tcf in 2035 (*25*). With increased NG production and overhauling or replacement of U.S coal-fired power plants, increased NG use for electricity generation will likely occur over the next 5-15 years (*56*). The U.S. The Energy Information Administration (EIA) projects that within 30 years an additional 100 GW NG-fired net electricity capacity will be installed compared to only 5 GW for coal (*26*).

The U.S. electricity sector is responsible for about 40% of energy related GHG emissions (*57*). Electricity generation from renewables offers large GHG reductions, but the high costs of a rapid transition to renewables may outweigh the potential benefits of avoiding climate change impacts (*58–61*). As a result, transitioning from coal fired electricity generation to a greater share of NG generation capacity is perceived as a low cost alternative to decarbonizing the energy system (*26*). However, the extent and timing to which a coal-to-gas transition will reduce GHG emissions and mitigate climate change is highly uncertain.

One of the most uncertain factors is the amount of NG emitted to the atmosphere during NG production, processing, transport, and local distribution, the latter for non-electricity related NG use, such as in the commercial and residential sectors (*36*). CH_4 is the main component of NG, which is a GHG – 25 times more potent than CO_2. Hence, relatively small amounts of NG FE can have considerable impacts on global climate. Some climate modeling studies suggest that – depending on the NG FE rate – using NG as a bridge fuel could lead to temporarily (decadal time scales) higher warming, partly due to increased emissions of CH_4 offsetting the potential reductions in CO_2 (*35*, *36*). Figure 7, which compares life cycle GHG emissions from coal and NG from the literature, illustrates the implications of these conclusions. It includes power generation with

assumed 39% and 50% power plant efficiency for coal and NG, respectively. Total GHG emissions from coal are significantly higher than NG, but (i) the CH_4 fraction of NG is much more uncertain compared to coal, and (ii) the CH_4 contribution in NG is up to 33% of total GHG emissions whereas this number is only about 5% for coal (not shown in Figure 7).

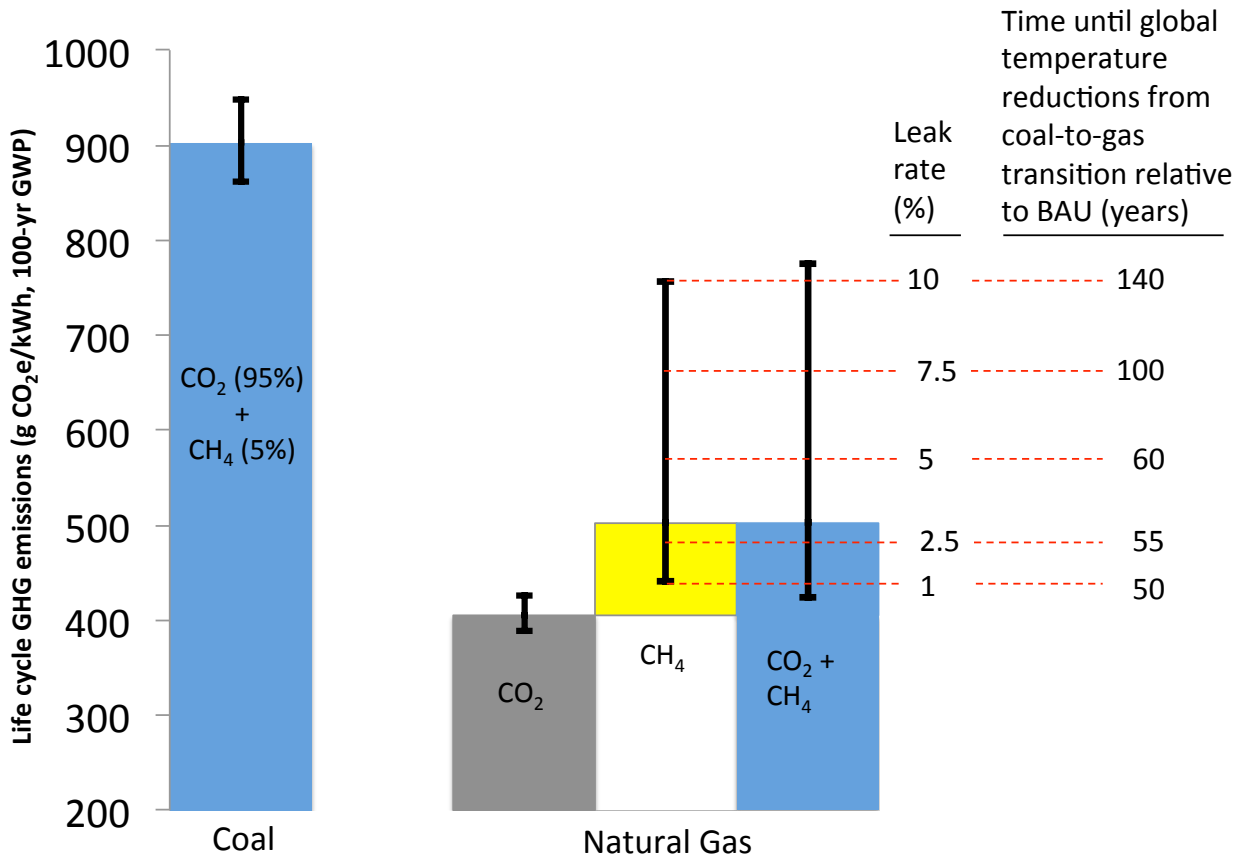

Figure 7: Comparison of GHG emissions between coal and NG life cycles including power generation at 39% and 50% efficiency, respectively. Literature estimates for coal (*28*, *31*, *36*), NG CO_2 (*28*, *30*, *34*, *36*), and the NG CH_4 base case (*10*) were used. The NG CH_4 uncertainties were calculated based on the range of NG FE rates used in Wigley (*36*) and the literature review in section 4.2.2. The temperature time lag estimates are adapted from Wigley (*36*) based on a global coal-to-gas transition as described above.

Greater NG CH_4 emissions (equivalent to up to 10% NG FE relative to production; see section 4.2.2 for a more detailed description) significantly increase the time until climate benefits can be achieved from a coal-to-gas transition. The numbers on the far right column summarize results by Wigley (*36*), which simulates global temperatures resulting from a business-as-usual (BAU) or "no-climate-policy" baseline emissions scenario as well as one where coal is gradually replaced with NG worldwide at a rate of 1.25% per year starting in 2000. At a 1% NG FE rate, global temperatures in the coal-to-gas transition scenario increase for 50 years relative to BAU before temperatures drop below BAU due to a combination of reduced coal related cooling aerosol emissions and

increased CH_4 from NG. In Wigley (*36*), this initial increase in temperatures (in the order of <0.1°C for a global transition) is mainly due to foregone emissions of cooling aerosols from reduced coal combustion. After 50 years, temperatures continue to decline below BAU due to reduced amounts of long-lived CO_2 emissions from NG. Yet, this time lag increases to about 140 years if the NG FE rate is as high as 10%. Alvarez *et al*. suggest that climate benefits from a coal-to-gas transition for power generation are achieved in a shorter period due to lower estimated (cooling) SO_2 emissions from U.S. coal plants (*62*). Reducing NG FE uncertainty is key for understanding the climate effects of a coal-to-gas transition over time.

Natural gas FE inflict economic losses on NG producers. Assuming a 2012 U.S. representative wellhead price of $3/Mcf NG (although prices have fluctuated from about $2-11/Mcf NG during 2000-2012), lost global NG has a market value of $15 billion/year (*63*). EPA's 2013 GHG inventory reports a 1.8% FE rate from NG production in the U.S. (*64*), which is equivalent to 0.36 Tcf/year or $1 billion/year. The NG industry can implement existing technologies (e.g., improved safety valves) and perform services (e.g., pipeline repairs) to reduce CH_4 emissions and maximize market output. However, these measures require upfront capital and O&M expenditures, which can represent a barrier for implementation and mitigation (*65*, *66*). The economic feasibility of these measures hinges critically on the estimated CH_4 savings potential, and thus on the estimated NG FE rate. While not necessarily linearly proportinate, higher FE (and therefore savings potential) reduces the net mitigation cost, which increases the range of economically feasible mitigation options. This research will provide valuable information for businesses in the NG industry to re-evaluate their CH_4 emissions reduction potential, which may increase the economic feasibility of available CH_4 mitigation technologies while simultaneously leading to more environmentally sustainable business practices.

3.2 Review of NG FE estimates, methods and uncertainties

The life cycles of coal and oil result in mainly CO_2 emissions from combustion and other processes, which are relatively easy to estimate (*34*). In contrast, NG FE emissions are difficult to quantify due to (i) challenges in accurately measuring fugitive CH_4 emissions from leaks and venting across a diverse industry and (ii) the fact that CH_4 arises from varied industries and natural sources. CH_4 emissions from the NG industry have been estimated through various approaches including bottom-up studies based on equipment-level measurement samples, global inversion modeling, and local air sampling. The following sub-sections will provide a brief description of each approach and summarize the literature results. It will also highlight the benefits and shortcomings of each approach in order to set the stage for the methods used in this work.

3.2.1 **Bottom-up analysis**

Bottom-up studies include direct leak measurement studies, which focus on specific parts of the NG industry such as pipeline transport (*67*), and state or national emissions inventories, which estimate absolute emissions based on emissions factors (EFs) and

total NG production (*54*). These bottom-up analyses for estimating NG FE essentially rely on measurements (leak detection & direct measurement of pipelines, processing facilities and other system components), engineering estimation (process specific emissions factors from manufacturers and industry models), and mass balance calculations (at the industry level). These data are provided mostly by the NG industry (*68*). The accuracy of these data is subject to limitations. First, EFs for equipment may vary by the chosen measurement method (*69*). Second, some measurement data sets represent a small sample that unlikely reflects the size and complexity of the NG industry. For instance, the number of well samples used to generate EFs by the EPA, which are referenced and used in LCAs, corresponds to only about 2% of the wells in the EPA GHG emissions inventory (*64*, *70*). The NG industry claims that these sample data are not representative of the industry, and that the EPA EFs are overestimated (*70*). While EFs based on small sample sizes could also be underestimated, a recent study using direct CH_4 measurements at 190 NG production sites by Allen *et al*. (*71*) indicate slightly smaller overall NG production related CH4 emissions compared to the EPA GHG inventory (*64*). Third, there may be significant spatial and temporal variability of FE rates. In 2013, EPA's Office of Inspector General concluded that EPA needs to improve air emissions data, since about half of its emission factors for oil and gas production are rated "below average or poor, or are unrated" (*72*). As a result, any analysis of these data is subject to considerable uncertainty.

Another type of bottom-up study is LCA, which estimates all emissions from NG extraction to end-use at a detailed level. LCA studies combine bottom-up data by producing a systematic inventory of the emissions in every life cycle stage. It is used to support complex environmental performance evaluations of competing technologies and fuels, such as NG and coal, given its ability to represent metrics like GHG emissions at high resolution (*11–13*). LCAs and direct measurement studies will be the focus throughout this chapter because these document their assumptions and system boundaries most transparently.

Numerous LCA studies document the life-cycle stage specific CH_4 emissions of NG production and point out potential for emissions reductions (*10*, *27*, *30–34*). In addition, other studies document CH_4 measurement data at various life cycle stages and world regions (*37*, *54*, *67*, *73*, *74*). These studies report NG FE rates from less than 1% to over 10% of NG lost during extraction, processing, transport, and local distribution. An in depth summary and discussion of the results of 16 studies targeting 15 countries from 1969-2011 is provided in section 4.2.2. While quantifying NG FE is technically challenging (see above), this wide range of FE estimates also results from differences in system boundary selection (see also Chapter 2) as well as assumptions and emissions data choices, i.e., industry and/or governmental data (*64*, *68*, *75*).

Bottom-up studies in general, and LCAs in particular, provide an in-depth analysis of the CH_4 emissions throughout different segments of the NG industry, and they point out emissions hotspots where abatement is most needed. From a policy perspective, however, the large uncertainties among studies create challenges for decision-makers

in estimating the potential climate benefits from a coal-to-gas transition. Hence, further analysis is needed to reduce NG FE uncertainties.

3.2.2 **Local and regional air sampling and analysis**

Natural gas FE can also be estimated based on collecting air samples near active oil and gas operation sites. These are sometimes referred to as field campaigns, in which air samples are collected via mobile labs installed in vehicles or airplanes over a period of several days or weeks (*76*). The field campaigns can be combined with data from regional tower-based observation networks (*77*). The collected air samples are analyzed for hydrocarbon composition and carbon isotope data in order to quantify air CH_4 and C_2H_6 concentrations, among others, and potential types of emissions sources, such as NG, landfills, and agriculture given their source specific carbon isotopic signature. A small fraction of atmospheric CH_4 contains the carbon isotope $\delta^{13}C$-CH_4, and CH_4 from the various emissions sources has different degrees of ^{13}C depletion, i.e., carbon isotope signature, due to differences in CH_4 formation processes (*78*). Measurements and analysis of $\delta^{13}C$-CH_4 thus provide an additional observational constraint for allocating CH_4 emissions among sources. After estimating the fraction of observed CH_4 from NG, this information is used in combination with local NG production data in order to estimate the site-specific NG FE rate. The methods are conceptually similar to global inversion modeling, which is described in more detail in the following section (3.2.3). Two recent studies with air sampling in Colorado and Utah suggested local NG FE rates near oil and gas basins ranging from 2.3-7.7% (mean 4.0%) and 6.2-11.7% (mean 8.9%), respectively (*77*, *79*).

The local and regional air sampling approach is complementary to bottom-up studies by providing atmospheric measurements to verify bottom-up estimates. While LCA and other bottom-up results are based on assumptions about specific business practices among others (see section 3.2.1 above), observational data reflects the emissions resulting from business practices at the site. Observations also reflect potential emissions sources, which may not be included in most bottom-up studies, such as leaking abandoned NG wells. However, the local and regional air sampling approach is subject to limitations. First, field campaigns lasting only several days or weeks may be viewed as a temporal snapshot. The measurement results depend on the individual NG production activities during the campaign, and NG FE estimates crucially depend on incorporating this information, which may not be fully available. For instance, if air sampling occurs during days when the number of well completions (pre-production) is high, NG FE will be overestimated because NG production volumes are used in the calculation. In a literature review of seven NG LCAs, Weber *et al*. (*10*) show that conventional well completions account on average 30% (range is 3-61%) of total on-site CH_4 emissions. This figure is higher for shale gas wells, but could be less when accounting for captured gas, e.g., green completions (*80*). Note that CH_4 from a well completion occurs over a short time of only several days or weeks whereas the remaining 70% of total on-site CH_4 emissions are released over several years. Of course, local air sampling measurements can also be skewed towards smaller NG FE rates for the same reason. Second, local and regional air sampling results may not be

representative nationally. However, nationally representative estimates are needed for informing energy and climate policy.

3.2.3 **Global inversion modeling**

Methane emissions can be estimated from the NG industry worldwide using global inversion modeling, which relies on atmospheric air samples distributed throughout the world and collected at varying frequencies since the 1980s (*81*). It is typically a Bayesian approach (*40*, *82*) in which *posterior* emissions locations, quantities, and sources are estimated using (i) *prior* emissions estimates, (ii) measurements of CH_4 or C_2H_6 concentrations, and/or $\delta^{13}C$-CH_4 isotope ratios from a global observation network, and (iii) atmospheric transport models to simulate how spatially and temporally distributed emissions change global concentrations (*81*). Global inversions typically rely on a global network of measurement stations. For instance, NOAA's network includes about 84 surface sites (*83*). The objectives of global inversion modeling are (i) to quantify the global CH_4 or C_2H_6 budget, i.e., the annual CH_4 or C_2H_6 emissions released to the atmosphere, and (ii) to distinguish the individual emissions sources comprising the global budget (*40*, *84*), which is illustrated in Figure 8 showing literature results of a CH_4 inversion.

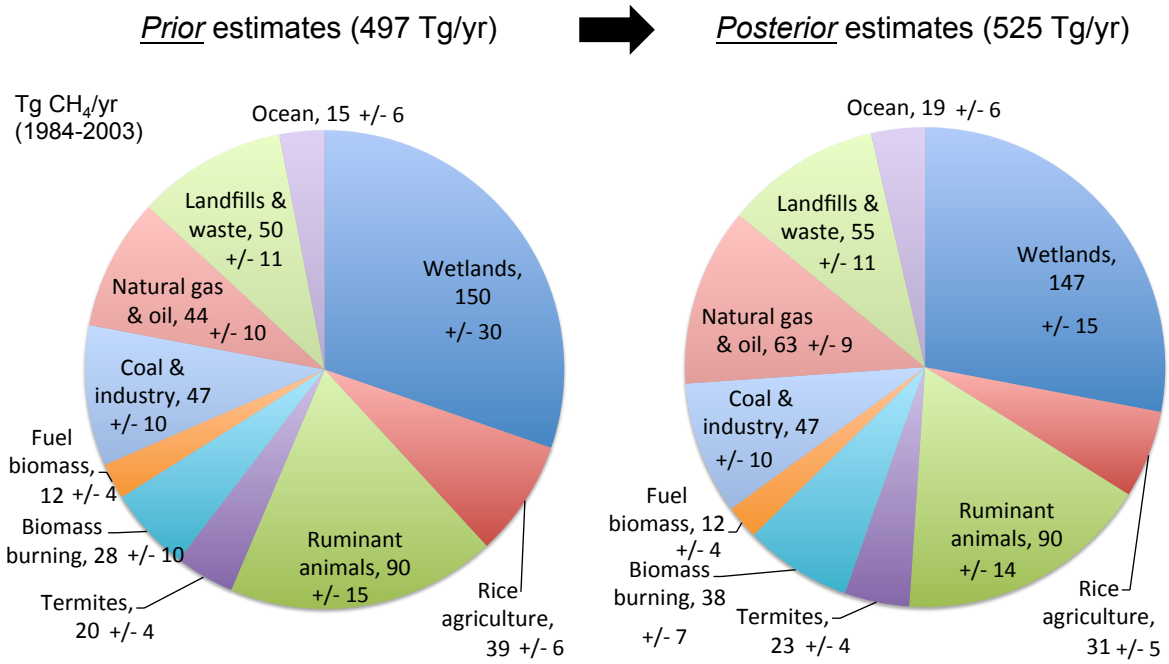

Figure 8: Literature results of global CH_4 inversion modeling. Global annual *prior* and *posterior* CH_4 emissions estimates averaged over 1984-2003; adapted from Bousquet *et al.* (*40*). *Prior* estimates are based on bottom-up studies. *Posteriors* are estimated by constraining the *priors* (budget and source contributions) using global atmospheric CH_4 and isotopic measurements. *Posterior* uncertainties indicate one standard deviation.

A conceptual overview of the inversion modeling process including tasks as well as involved data and models is given in Figure 9 for the example of CH_4. The *prior* source

emissions estimates (including uncertainties) shown in Figure 8 are developed at a country (anthropogenic sources) or regional (natural sources) level using a bottom-up approach including various data sources described in more detail in Chapter 4. Global maps including FF production sites and transportation routes, urban population, natural ecosystems and other proxies are used to allocate the total country and regional level emissions spatially across the globe (*85*). Emissions source specific carbon isotopic signatures (*78*), are linked to each source in the resulting emissions maps, which are generated at a given frequency, e.g., monthly. An emissions transport model is then used to simulate how the *prior* emissions change CH_4 concentrations and isotope ratios throughout the global atmosphere over time. Observational data is available since the 1980s (*83*), hence simulations can be performed over the past three decades.

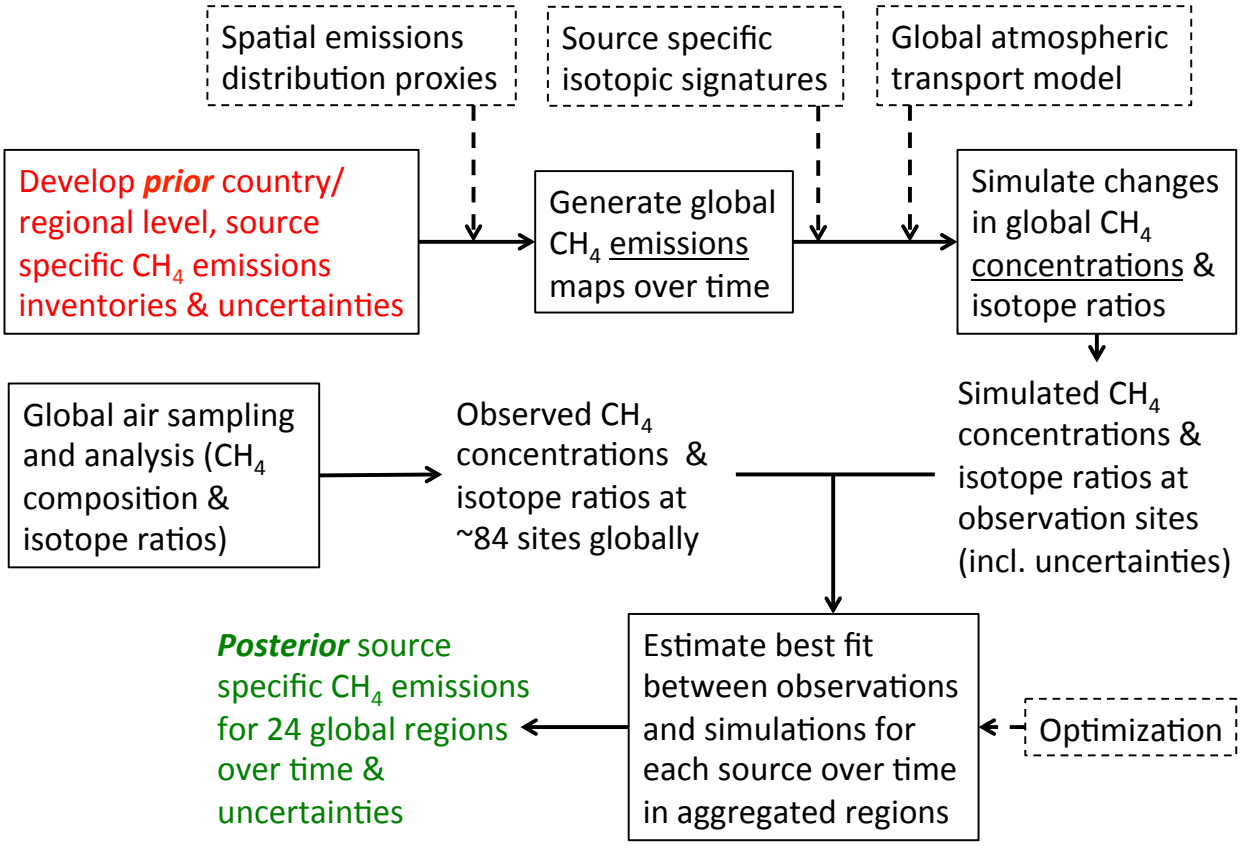

Figure 9: Conceptual overview of global CH_4 inversion modeling. CH_4 sources typically include those depicted in Figure 8. Solid and dashed frames symbolize tasks and data/models, respectively.

Next, the simulated concentrations, isotope ratios and associated uncertainties – including *prior* emissions and transport modeling uncertainties – are constrained by global measurements. The measurement data includes CH_4 concentrations and isotope ratios resulting from atmospheric air samples, i.e., mainly flask samples from an international network of surface stations and towers (*83*). In addition to flask samples, satellite data provide observational constraints using imaging spectrometers (*86*, *87*), which measures trace gases such as CH_4 by recording transmitted, backscattered, and

reflected radiation from the atmosphere. In contrast to the limited number of surface measurements, satellite data cover a geographically larger area, but it relies on reflected sunlight with less data retrieval at high latitudes in winter, at night, or cloudy skies (*81*). The quality of the obtained greenhouse gas data is subject to ongoing development (*42*, *87–89*).

An optimization approach is used to estimate the best fit between observations, i.e., measurements, and simulations for each source over space and time. The resulting spatially distinct concentrations from the atmospheric transport model are aggregated to so-called source regions. For instance, NOAA's CarbonTracker-CH_4 (CT-CH_4) model divides the atmosphere into 13 land regions and 11 ocean regions. CT-CH_4 then maximizes an objective function, which estimates *posterior* emissions by minimizing the sum of the squared residuals between measurement sites and simulation results (concentrations) in the corresponding source region over time (*82*). As will be discussed in more detail in Chapter 5, the resulting *posterior*s are influenced to a larger extent by the *prior* estimates in regions where observational data is sparse. The resulting *posterior* emissions for each source and region can be aggregated to yield global source strengths (see also right pie chart in Figure 8). The smaller the magnitude of total residuals, the smaller the difference between global atmospheric CH_4 burden and total *posterior* source strength.

The global atmospheric CH_4 burden used as an inversion constraint, i.e., total annual emissions, is determined based on CH_4 concentration measurements across the globe in combination with global CH_4 sink estimates (see section 5.2). The three major CH_4 removal processes from the atmosphere are reaction with the hydroxyl radical (OH), reaction with atomic chlorine (Cl) and excited-state oxygen (O^1D) in the stratosphere, and oxidation by methanotrophic bacteria, i.e., soil uptake (*90*). The OH sink is by far the largest sink, and its CH_4 removal processes as well as magnitude are discussed in Appendix B along with the stratosphere sink. In inversion modeling, the strength of the photochemical sink (OH, Cl, and O^1D) is often expressed as CH_4 lifetime τ (*39*, *81*, *91*). It represents the residence time of CH_4 in the atmosphere until it is removed through photochemical processes. Section 5.2 describes in detail how the CH_4 burden is estimated using CH_4 measurement time series and τ. In contrast to the photochemical sink, CH_4 removal through oxidation in soils is commonly accounted for by explicitly including a negative CH_4 source in the emissions *priors* (*39*, *40*). Recent estimates of soil uptake range between 25-40 Tg/yr, i.e., less than 10% of the global CH_4 burden (see section 4.6).

Several global inversion studies have reported FF related CH_4 emissions over the past decades (*39*, *40*, *42*, *92*, *93*). The global FF CH_4 emissions estimates in two recent inversion studies (*40*, *83*) are summarized in Figure 10. The best estimates of total FF CH_4 range from 78-110 Tg CH_4/yr (excluding *posterior* uncertainty ranges) and 62-123 Tg CH_4/yr (including *posterior* uncertainty ranges), which are constrained by *prior* information and measurements in the inversion model. Bousquet et al. (*40*) suggest that FFs account for 97-123 Tg CH_4/yr, and NG and oil production are responsible for 54-72 Tg CH_4/yr on average (1984-2003). Estimates for NG only were not provided, but can

be estimated. Accounting for the NG fraction of the NG and oil sum in the *prior* estimates, i.e., 71% and 29%, respectively (*94*), NG alone ranges from 39-51 Tg CH_4/yr. The NOAA estimates are lower compared to Bousquet *et al*. (*40*). According to the NOAA study FFs account for 62-94 Tg CH_4/yr. Accounting for the NG fraction in the FF *priors*, NG emissions range from 23-35 Tg CH_4/yr. Thus, considering both inversion studies, NG emissions range from 23-51 Tg CH_4/yr.

In order to compare these results to NG FE rates estimated in LCAs and other bottom-up studies, percent values of NG FE were converted to absolute (Tg CH_4) values in Figure 10. Using dry NG production and hydrocarbon composition data, which is described in more detail in Chapter 4, absolute NG CH_4 emissions were estimated globally for NG FE rates ranging from 1-10%, i.e., the uncertainty ranges in LCAs described above. Emissions were estimated over the period 1984-2011, which is covered by the inversion studies. FE rates of 1% and 10% translate into 15 and 150 Tg CH_4/yr on average over this period. The lower and upper bounds of both inversion studies (23-51 Tg CH_4/yr) then correspond to global average NG FE rates between 1.5-3.4%. Hence, these inversion studies suggest that the upper 2/3 range of LCA NG FE estimates are too high compared global atmospheric measurements.

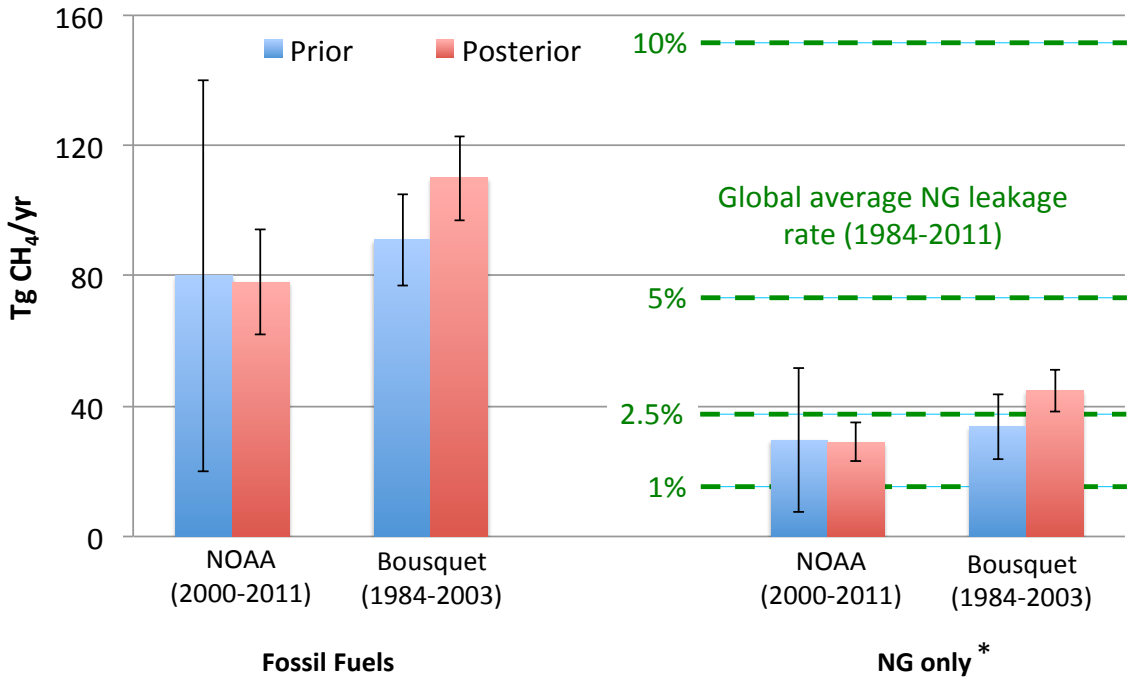

Figure 10: Summary of FF estimates in previous inversion studies (*39*, *40*). *Prior* uncertainties in NOAA are ±75% of mean value (no explanation provided in Bousquet *et al*.). *Posterior* FF uncertainties represent one standard deviation. The asterisk denotes that *posterior* uncertainties for NG only (excluding coal and/or oil) were not estimated in the studies. The error bars for NG only were estimated based on FF uncertainties in the inversion and the NG fraction of *prior* FF emissions, which is described in more detail in the text.

Global inversion modeling has been used to verify bottom-up emissions estimates, and current observation networks are suitable for assessing emissions over large scales (*81*). However, specifically quantifying NG FE using global inversion modeling is subject to the limitations discussed below.

Sparse observational network:

First, while distinguishing broad CH_4 sources, e.g., FFs, landfills, agriculture, and wetlands, is one of the main inversion objectives, it is also a great challenge. Source distinction is difficult because (i) isotope measurements (see sections 3.2.2 and 5.2 for details) are only available for a fraction of observation sites (only 13 stations globally in Bousquet *et al*. (*40*)), and (ii) the isotopic signature ranges of different sources in general, and FF sources in particular, overlap significantly (*95*). While monitoring CH_4 concentrations long-term (several decades) is an advantage over local and regional air sampling campaigns, the number of observation sites (~100) is relatively small for inferring globally distributed emissions. The spatial distribution of CH_4 observations is generally helpful in distinguishing some sources, e.g., FFs in the Northern hemisphere and equatorial and high latitude wetlands. However, it is often difficult to discriminate between other sources. For instance, CH_4 from landfills and enteric fermentation tends to be emitted in highly populated regions also geographically close to FF production and use. Enteric fermentation in particular is estimated to be one of the largest CH_4 sources globally in bottom-up and inversion studies, and some consider it the single most uncertain quantity (*96*). Moreover, there is a lack of measurements at sites with potentially high NG production activity, such as Russia. Given these conditions, observational CH_4 data can only provide a limited constraint in current inversion models. As a result, the *prior* source allocation can significantly influence the *posterior* source allocation, which is described in more detail in Chapter 5. This limits the amount of information that can be obtained about the total amount and spatial distribution of NG emissions from the inversion results above.

Influence of prior uncertainties:

Second, published inversion studies typically report *posterior* FF uncertainties, but rarely *posterior* uncertainties for NG only (*39*, *40*). While some studies quantify the extent to which the choice of *prior* estimates influence *posteriors* and their uncertainties for all sources in general (*82*), this information is typically not available for the NG source in particular. As discussed above, *prior* estimates may largely influence *posterior*s in regions with few measurements. Hence, documented *posterior* uncertainties may only represent a fraction of the total uncertainties. As a result, the earlier estimates of 23-51 Tg CH_4/yr corresponding to 1.5-3.4% global average NG FE rates could be too narrow, and further analysis may help quantify the larger range of uncertainty. Assigning *prior* uncertainties is a challenge for inversion modelers who rely on data sources, such as the Emissions Database for Global Atmospheric Research (EDGAR). EDGAR (*85*) is the commonly used *prior* data source for global CH_4 inversion modeling, which provides both country-level data and global emissions grid maps for atmospheric modeling. Due to insufficient transparency in parts of the database, only some of its data – including the spatial distribution of total emissions – is useful for the purposes of this dissertation

research. Since EDGAR (*85*) reports only point estimates for each source and location (see Figure 11), inversion modelers are likely to set *prior* uncertainties for a given emissions source without reference to bottom-up ranges. Yet, some bottom-up studies, such as LCA, analyze and report uncertainties based on data from different sources, e.g., process specific EFs collected from several industry and government reports. Results or approaches from LCA studies have not been used to generate inversion *priors* in the past, but LCA data can be useful for informing *prior* uncertainty ranges in future inversions.

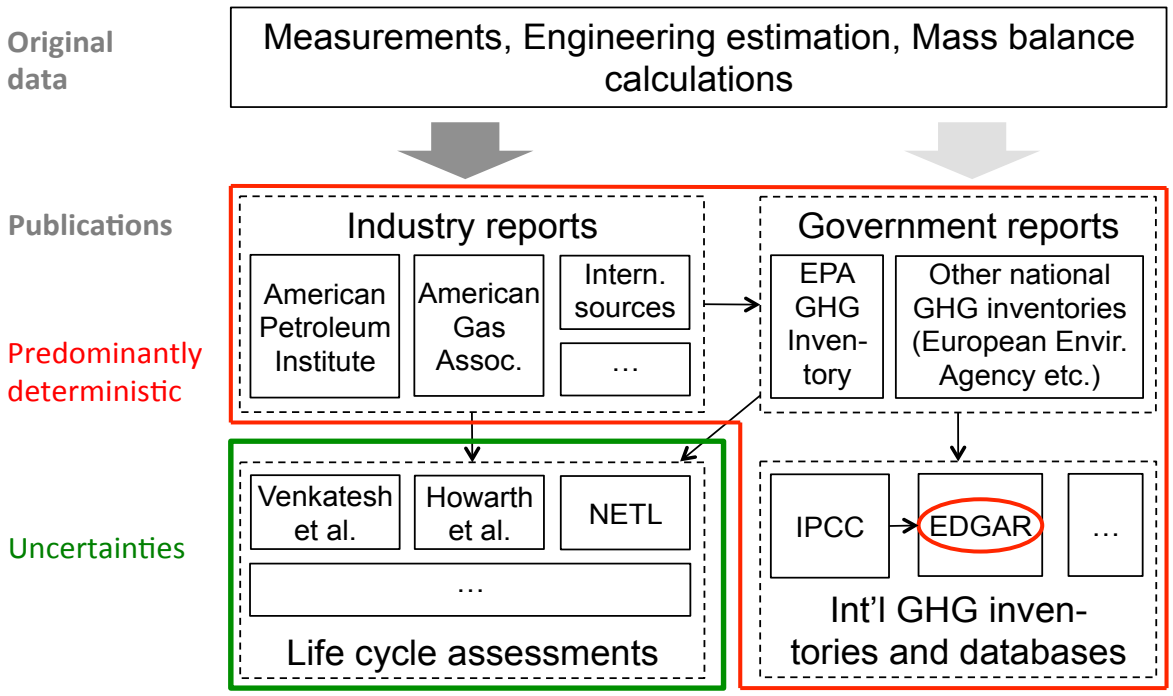

Figure 11: An approximate hierarchy of bottom-up data sources and publications used in LCA studies and as *priors* (EDGAR) in atmospheric inversion studies.

Transparency of prior estimates:

Third, estimating globally or regionally averaged NG FE rates based on inversion *posterior*s is limited by the transparency of the *prior* estimates. Since inversion *posterior*s (in Tg CH_4/yr) must be converted back to % NG FE rates, it is crucial that the information needed for this conversion is well documented. This information, which includes regionally specific NG production data and assumptions about NG hydrocarbon composition among others (see more in Chapter 4), is used to generate the inversion *priors*. Hence, developing transparently documented inversion *priors* is key to estimating NG FE rates through inversion modeling. Moreover, transparency is important for highlighting the assumptions underlying the inversion *priors*. A critical review of NG FE estimates based on global inversion modeling is only possible if these assumptions are well documented. Yet, *prior* emissions used in previous inversions lack this transparency. EDGAR uses National Inventory Reports (NIRs) of the United Nations Framework Convention on Climate Change (UNFCCC) member countries as the main

source among others (*85*). While this allows for collecting country specific emissions, it also creates challenges, such as the need to replace missing or incomplete data. Furthermore, the NIRs already aggregate data from industry and government reports (*64*, *75*), which renders data transparency a difficult task. Since EDGAR (*85*) reports include little information about their data compilation, assumptions regarding NG FE can only be inferred at a very coarse level. For example, assumptions about emissions allocation between oil and NG are not provided in EDGAR. A summary of the bottom-up data sources available for inversion *priors* is shown in Figure 11 to illustrate the reporting hierarchy of bottom-up estimates and the benefits of using LCA emissions estimates. Recent NG LCAs capture the emissions differences between bottom-up studies, and provide uncertainty ranges based on the literature source data (*10*, *28*, *31*, *34*).

3.2.4 **Summarizing the existing problem of quantifying global NG FE**

Research over the past few decades has contributed a large amount of data for quantifying NG FE in different parts of the world. This data has led to numerous in-depth analyses of NG industry processes and resulting CH_4 emissions. Despite this development, methodological and data limitations preclude robust conclusions regarding an average industry-wide NG FE rate. Natural gas FE estimates from bottom-up studies including LCA vary by an order of magnitude (between about 1% and 10% of NG production) because of measurement and engineering data choices, differences in system boundaries, and location variability. Moreover, measurement data may not be representative of the entire industry due to small sample sizes. Local and regional air sampling studies may only provide a temporal and spatial snapshot. Estimating NG FE based on measurements and limited data about on-site production activities creates additional challenges. Previous global CH_4 inversion results indicate that global average NG FE are at most one third of the upper estimates in the literature. The wide range of FE rates from LCA and local and regional air sampling studies and the considerable differences in estimates and uncertainty of both approaches warrants further investigation in order to better understand the climate impacts of NG production and use. While uncertainties need to be analyzed more carefully, global inversion modeling is a promising approach to re-assess LCA estimates. Future inversion studies may characterize possible NG FE uncertainties in more depth by (i) estimating NG FE for a wide range of *prior* estimates, and (ii) clearly documenting the data and assumptions underlying the *prior* FF estimates.

3.3 Research objectives and tasks

3.3.1 **Research objectives**

This part of the research can be considered a ground truth analysis of NG FE rates used in LCA studies by comparing bottom-up life cycle CH_4 emissions from global NG production (LCA results) and top-down CH_4 emissions from atmospheric concentration measurements. The research throughout this project will determine the likely ranges of global fugitive CH_4 emissions from NG production based on bottom-up estimates, and

constrain these ranges with atmospheric measurements. This will be achieved through combining LCA and global CH_4 inversion modeling, which is illustrated in Figure 12. While the objective of traditional inversion modeling is multifaceted, e.g., understanding the total budget of CH_4 and C_2H_6 as well as the interannual variability of the sources and sinks (*40*, *81*, *82*), this research focuses on characterizing CH_4 and C_2H_6 emissions uncertainty from the NG industry. This will be achieved by generating transparent *prior* emissions inventories using estimates from the LCA literature and other bottom-up studies. In order to link *prior* assumptions to *posterior* results, the emphasis lies on clarity and transparency of bottom-up estimates.

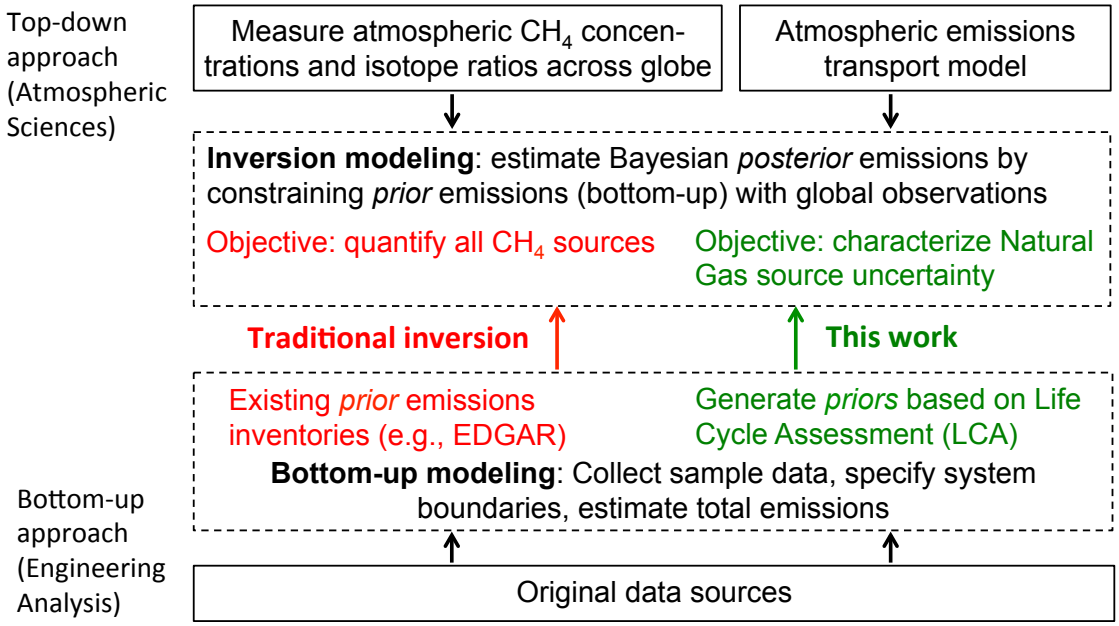

Figure 12: Research overview and differences to previous work in the literature.

The overarching research question throughout this work is: which range of NG FE (%) is reasonable globally, and which range is inconsistent given atmospheric measurements as a constraint? How would these refined ranges affect prior or future policy decisions? The specific research questions to be answered are:

1) What are the best estimate NG FE rates over the past two decades globally according to CH_4, C_2H_6, and $\delta^{13}C$-CH_4 tracers, and is there a trend over time?
2) What is the upper bound of global NG FE given the lower bound of all other CH_4 and C_2H_6 sources, and what are the *prior* assumptions (e.g., FF emissions factors and hydrocarbon composition) underlying this bound?
3) Can literature estimates of natural CH_4 and C_2H_6 seepage be reconciled with global measurements, and how does this affect the NG FE estimates?
4) What are the policy implications of global atmospheric observational constraints on NG FE rates?

This research does not explicitly quantify CH_4 emissions from shale gas operations, nor does it make projections about future trends in NG FE. As mentioned above, the modeling focuses entirely on emissions from the 1980s to present. However, this research is an important groundwork for future studies dealing with new technologies, such as hydraulic fracturing. Several LCA studies have compared NG FE from conventional and shale gas (*10*, *31*), the latter having reached the 10% of total U.S. production mark only in 2008 (*97*). While these studies provide a detailed analysis of the emissions differences between conventional and shale gas, robust conclusions about shale NG FE rates cannot be drawn due to the large uncertainties in NG FE. As a result, better knowledge about past and present NG FE from conventional sources is needed to infer FE of new technologies using LCA methods, and to make projections about future emissions.

3.3.2 **Research tasks**

The tasks required to meet the research objectives and specific research questions are outlined below. The list gives an overview of the tasks, which will be described in more detail in the following Chapters.

- Chapter 4: Develop Bayesian *prior* CH_4 and C_2H_6 emissions inventories
    a) Generate FF CH_4 emissions inventories from the 1980s to 2011.
        - Collect FF production data and review CH_4 and C_2H_6 EFs for NG, oil, and coal as well as gas hydrocarbon composition.
        - Allocate emissions between NG and oil production.
    b) Create global grid maps for spatial allocation of FF CH_4 and C_2H_6 emissions.
    c) Review *prior* CH_4 and C_2H_6 emissions and uncertainties from non-FF sources in the literature.
    d) Collect global CH_4 emissions grid maps for all other sources from existing databases.

- Chapter 5: Global CH_4 and C_2H_6 atmospheric modeling
    e) Develop a global box-model to carry out inversion modeling using global average CH_4, C_2H_6, and $\delta^{13}C$-CH_4 measurements to constrain NG FE.
    f) Simulate global emissions transport using the 3-dimensional CT-CH_4 model, compare results with the box-model, and analyze the spatial constraints for NG FE.

While similar to the conceptual overview of global CH_4 inversion modeling (see Figure 9), the *prior* emissions inventory tasks focus on generating FF inventories, and global totals for other CH_4 and C_2H_6 sources will be used directly from the literature. Furthermore, the analysis of inversion *posterior*s will focus on NG FE uncertainty in order to address the research questions above.

Chapter 4: Quantifying Natural Gas Fugitive Emissions – Bayesian *Prior* Estimates

4.1 Introduction

Over the past few decades, numerous bottom-up and field studies have estimated the amount of CH_4 emitted to the atmosphere from the NG industry. Yet, limitations in research methods and data lead to large uncertainties, i.e., between about 1% and 10% of NG production is lost due to leaks and venting (see Table 3). Recent global CH_4 inversion results indicate that uncertainty in global average NG FE may be reduced significantly using a global observation network. As described in Chapter 3, this approach requires a transparent set of Bayesian FF *priors* in addition to EDGAR (*85*), which provides only point estimates for both country-level data and global emissions grid maps rather than uncertainty ranges. Due to insufficient transparency in parts of the database, only some of its data – including the spatial distribution of total emissions – is useful for this research.

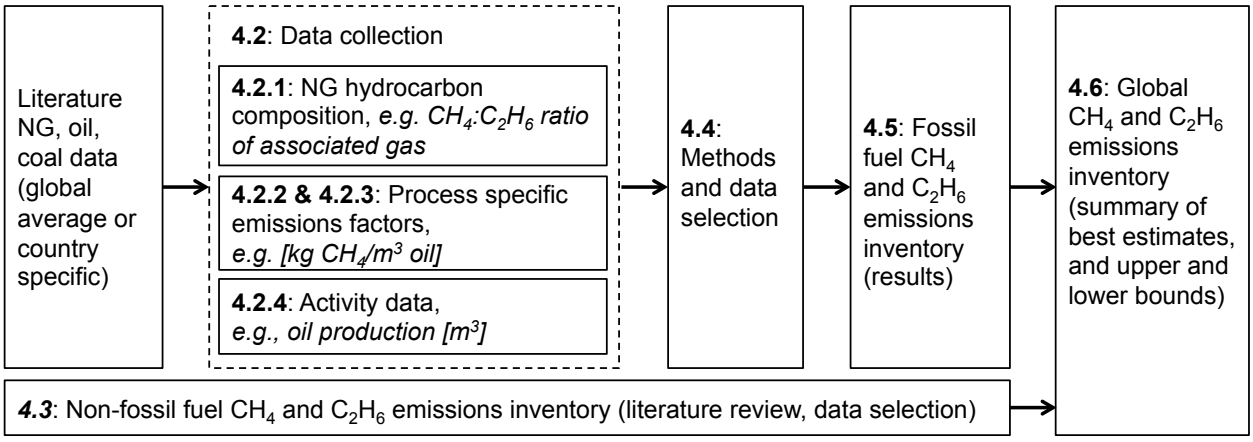

Figure 13: Overview and structure of Chapter 4 (section numbers are indicated in bold).

This chapter describes the emissions quantities and the data sources used to generate the FF *priors* for the atmospheric modeling part of this work. NG, oil, and coal related CH_4 and C_2H_6 emissions *priors* are estimated in detail using FF production data and EFs from the literature. Distinguishing the three FF CH_4 sources in observed atmospheric CH_4 concentrations is a challenging task in inversion modeling. Since the isotopic signatures of the three sources overlap significantly, and the emissions releases of these sources are also in close geographical proximity, the observational signal to detect each source individually is relatively weak (*40*). As a result, detailed analysis and transparent documentation of oil and coal industry *priors* helps isolating NG related *posterior* and associated uncertainties. Given a *posterior* FF CH_4 budget, the NG fraction is easily estimated (NG = FF – oil – coal) based on assumptions underlying the coal and oil *priors*. Other anthropogenic (e.g., agriculture, industry) and natural CH_4 emissions quantities were adopted as *priors* from the literature.

As shown in Figure 13, sections 4.2 and 4.3 review the literature and databases regarding emissions factors for NG, oil, and coal, gas hydrocarbon composition, activity data (FF production) for individual countries, which are needed to generate country-level emissions estimates, and non-FF sources. Section 4.4 describes the methods for generating fossil *priors* including their global spatial distribution based on aforementioned data collected. This section also summarizes the data sources used for other anthropogenic and natural CH_4 and C_2H_6 emissions. Section 4.5 summarizes the results, i.e., the *prior* emissions and uncertainties for each emissions source, and section 4.6 provides concluding remarks.

4.2 Fossil fuel CH_4 and C_2H_6 emissions data collection

4.2.1 Natural gas and coal-bed hydrocarbon gas composition

Life cycle FF hydrocarbon emissions depend on the amount of NG, oil, or coal produced, FE rate (or EFs), and CH_4 and C_2H_6 content in the released gas. For instance, the emissions at a given NG FE rate increase with higher CH_4 and C_2H_6 content in the leaking gas. Hydrocarbon composition across the NG industry, from oil production, and coal mining, particularly CH_4 and C_2H_6 content, is reviewed in this section. The terms in Table 1 will be used in this chapter to distinguish NG from different well types and locations, which determine the composition, use, and hydrocarbon emissions of NG as described below.

Table 1: Typical distinctions of NG based on well type and location (*98*).

Category	Description
Unassociated (or non-associated) gas	NG from gas and condensate wells, in which there is little or no crude oil. Gas wells typically produce raw natural gas by itself, while condensate wells produce free natural gas along with a semi-liquid hydrocarbon condensate.
Associated gas	NG from oil wells. This gas can exist separate from oil in the formation (free gas), or dissolved in the crude oil (dissolved gas).
Stranded gas	NG from deposits for which the construction of pipelines is uneconomical.

Natural gas from the well-head (gross NG) is processed to provide a consistent quality for the NG consumer. Processing removes higher hydrocarbons along with other impurities (*99*). As a result, NG composition varies between upstream (prior to central processing facilities) and downstream (see Table 2). NG CH_4 content upstream (gross and wet NG production) is generally lower compared to NG downstream (dry NG production) because the higher hydrocarbons and other NG components have not yet been removed. The opposite is true for C_2H_6. Apart from a few samples collected in the 1960s (see Table 2), there are no recent publicly available data providing specific NG hydrocarbon composition in different parts of the industry. This work determines NG FE rates – based on dry NG production and dry NG hydrocarbon composition – in order to compare results with some recent bottom-up studies, which are also based on dry NG production statistics (see Table 3 and discussion below for details). Since this requires composition data of dry NG, the CH_4 and C_2H_6 content of dry NG is estimated using upstream hydrocarbon composition data as well as wet, dry, and NGL production data

in a mass balance approach in sections 4.2.1 and 4.4.1. Table 2 summarizes available global upstream hydrocarbon composition data from the literature as well as downstream composition data from two sources. Literature data is given in different units (volume, molar, and weight percent) depending on sources.

Table 2: CH_4 and C_2H_6 content in NG across different parts of the industry, ordered by type of NG (unassociated/associated/mixed), process stage, sample size and location, and reference year.

Process		Sample size, Locations	Year	Units	CH_4	C_2H_6	Sources
Unassociated NG							
	Extraction (wet)	n/a	n/a	Vol-%	96.0	2.7	(*100*)
	Upstream of centralized processing plant	n/a (Canada)	1987	Mol-%	92.5	4.5	(*101*)
				Wt-%	85.0	7.8	
	Downstream of centralized process. plant	n/a (Canada)	< 2004	Mol-%	97.2	0.7	(*101*)
				Wt-%	95.0	1.3	
	Major transmission lines	28 (USA)	< 1966	Vol-%	84.6	5.7	(*102*)
	Distribution system	48 (USA)	1962	Vol-%	89.4	5.1	(*102*)
Associated NG							
	Extraction (wet)	47 (Williston, USA)	2006	Mol-%	54.0	17.0	(*103*)
				Wt-%	31.9	18.8	
		n/a (Canada)	1987	Mol-%	73.3	12.0	(*101*)
				Wt-%	52.3	16.0	
		n/a	n/a	Vol-%	30-40	n/a	(*100*)
Mixed unassociated & associated NG sources							
	Extraction	n/a (USA)	1996-2011	Vol-%	84.0	n/a	(*64*)
		6,109 (USA)	1948-1997	Vol-%	79.9	5.7	(*104*)
				Wt-%	64.7	8.6	
		625 (USA)	1976	Vol-%	88.3	4.7	(*105*)
		40 (USA)	1951	Vol-%	84.1	8.1	(*102*)
		8,405 (Uinta Basin, USA)	2012	Wt-%	76.4	8.9	(*106*)
		77 (Barnett Shale, USA)	2012	Wt-%	66.2	n/a	(*107*)
		n/a (Denver-Jules Basin, USA)	2006	Wt-%	57.6	15.5	(*108*)
		1,053 (globally)	1984-1997	Mol-%	81.8	3.7	(*109*)
				Wt-%	61.9	5.2	
		174 (globally)	1970	Vol-%	89.0	n/a	(*105*)
		32 (globally)	< 1966	Vol-%	81.5	9.8	(*102*)

As shown in Table 2 CH_4 and C_2H_6 values given in volume and molar percent are not significantly different from each other, but weight percent values are significantly different due to increasing specific weight for higher hydrocarbons. Hydrocarbon emissions from the oil industry are due to associated NG production at oil wells. Some

of these oil related fugitive emissions come from leaks, and the remainder is from vented or incompletely flared gas (see section 4.2.2 for details) due to process upsets and/or lack of infrastructure to deliver the gas to market, i.e., stranded gas. Unassociated wet NG generally has a higher CH_4 content than associated wet NG (>90 vol-% versus <80 vol-%), whereas the opposite is true for C_2H_6 (unassociated <5 vol-%, associated >10 vol-%). The CH_4 content of mixed associated and unassociated NG varies between approximately 80-90 vol-% (5-10 vol-% for C_2H_6). While the majority of data refers to NG production in North America, only three literature sources provide global composition data, including only one source after 1970. Two data sources comprise over 7,000 well samples (*104*, *109*) – rather than averages only – and were used to compare North American and global data (Figure 14).

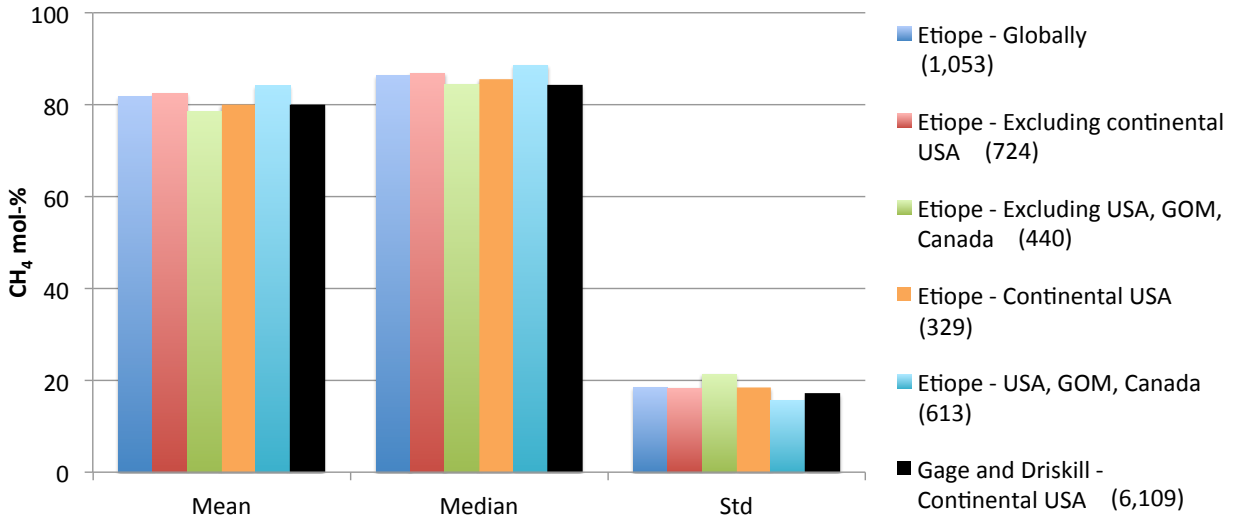

Figure 14: Comparison of CH_4 content in NG globally, in North America, and the U.S. between two data bases (*104*, *109*). Legend values in parentheses state the number of samples. GOM = Gulf of Mexico.

While the contents of the database in Etiope (*109*) are proprietary, permission to publish the above statistics was obtained. As shown in Figure 14, CH_4 content in the continental U.S. is not significantly different between both sources. Furthermore, CH_4 content globally (*109*) is not significantly different from the continental U.S. (*104*). Hence, it is concluded that on average the CH_4 content from the continental U.S. is representative globally.

The CH_4-C_2H_6 emissions ratio from coal mining is important for estimating coal related C_2H_6 emissions because most EFs for coal mining provide only CH_4 estimates (see section 4.2.3). Kotarba and Lewan (*110*) provide coal-bed gas composition from two hard coal (underground mining) basins in Poland. Based on 13 samples from different sites, the CH_4-C_2H_6 (wt-%) ratio ranges from 2 to 3,500, with a weighted average of 64 (weighted average C_2H_6 content is 1 wt-%). Gas composition of underground mines throughout the U.S. is available from 78 samples (*111*). The CH_4-C_2H_6 ratio ranges from 60 to 9,900 with a weighted average of 930, but the unit (vol-%, mol-%, or wt-%) is

omitted in the paper. If vol-% or mol-% are used, the weighted average CH_4-C_2H_6 ratio would be in the order of 500 (see calculations in section 4.4.1 for details). In summary, the weighted averages of the sample C_2H_6 contents in Poland and the U.S. is 1 wt-% or less.

4.2.2 **Emissions factors from the NG and oil industries**

Emissions factors describe the amount of hydrocarbon gas emitted to the atmosphere per unit of fuel produced. Metrics can be absolute, e.g., in units of kg CH_4/m^3 oil, or relative amounts, such as % produced dry NG. As discussed in Chapter 3, EFs in LCA are the basis for comparing GHG emissions among different fuels or technologies. EF values sometimes vary significantly depending on the data source. Numerous studies have estimated CH_4 emissions from the NG industry. This works reviews 16 studies, which have either (i) directly incorporated EFs from industry or government sources or LCA databases, or (ii) reported own measurements. Detailed EFs for some sub-processes (as opposed to aggregate EFs), such as gas flaring, are reviewed here in order to explain the differences between studies, and to justify the emissions inventory methods in section 4.4.

Table 3 summarizes the review results, which are shown in FE as a percent of produced NG. FE rates in some studies (*31*, *34*, *38*, *67*, *73*) are based on dry gas production, whereas this was not documented in some studies. In order to facilitate comparison of LCA and inversion results, FE rates for NG *priors* below are generated based on dry gas production. Most studies documented NG CH_4 emissions per MJ NG produced or per MWh generated power. In these cases, the data were converted to % NG FE using NG properties data from the literature, such as energy content of NG (see Appendix Table B - 8). One reason for the wide range of FE estimates (ranging from near zero to over 10%) is differences in the choice of system boundaries between studies. For example, potential emissions from local NG distribution networks are often omitted, and some studies include only pipeline FE. Furthermore, disagreements between studies arise due to data choices from different industry and government reports (*10*). Few studies have accounted for the full life cycle including production, processing, transmission, and local distribution.

Table 3: Literature values of NG FE ordered by location, reference year, and life cycle stage. Units are FE in percent of dry production (*31*, *34*, *38*, *67*, *73*); whether dry or gross production not reported in other studies. Studies reporting CH_4 emissions on a per MJ (or similar) basis (*113-117*) were converted to percent of dry NG production assuming 90 vol-% CH_4 content. Total values in parentheses indicate that study did not include all life cycle stages.

Country	Year [a]	Production	Processing	Transmission	Distribution	Total	Notes	Ref.
USA	2011	1.4	0.4	0.2	0.2	1.5-3.2	Convent. NG LCA	(*34*)
USA	2011	2.0		0.3	-	(2.3)	Convent. NG LCA	(*10*)
USA	2011	0.3-2.2	0-0.2	1.4-3.6		1.7-6.0	Convent. NG LCA	(*28*)
USA	2011	1.6	0.4	0.2	0.2	1.5-4.0	Shale gas LCA	(*31*)
USA	2011	2.0		0.3	-	(2.3)	Shale gas LCA	(*10*)
USA	2011	2.2-4.1	0-0.2	1.4-3.6		3.6-7.9	Shale gas LCA	(*28*)
USA	1992	0.4	0.2	0.5	0.3	0.9-1.9	-	(*112*)
Russia	2005	0.8-2.9			-	(0.8-2.9)	-	(*113*)
Russia	2003	0.1	-	0.4-1.6	0.5-0.8	(1.0-2.5)	Own measurements, transmission only	(*67*)
Russia	1996	-	0.1	0.9	-	(1.0)	Own measurements	(*73*)
Russia	Pre-1993	2.4-2.8		3.3-4.0	0.8-2.2	6.5-9.3	-	(*37*)
Thailand	2009	0.02	0.17	-	-	(0.19)	LCA database EFs	(*114*)
Romania	2007	2.9	0.8	1.9	-	(5.6)	-	(*115*)
Algeria	2000	0.1	0.1	-	-	(0.2)	LCA database EFs	(*116*)
Germany	2005	0.8		0.3		1.1	-	(*117*)
Germany	2000	0.04	-	-	-	(0.04)	LCA database EFs	(*116*)
Netherlands	2005	0.14		0.04		0.18		(*117*)
Netherlands	1987	-	-	-	0.6	(0.6)		(*118*)
Austria	2005	1.6		0.7		2.3		(*117*)
Belgium	2005	0.14		0.03		0.17		(*117*)
Spain	2005	0.22		0.19		0.41		(*117*)
France, Italy, Luxembourg	2005	0.7		0.3		1.0		(*117*)
Norway	1991	0.02			-	(0.02)		(*119*)
UK	1969	-	-	-	1.9-10.8	(1.9-10.8)	-	(*38*)

[a] Base year (if provided), otherwise year of publication.

As described in section 4.2.1, hydrocarbon emissions from the oil industry are mainly due to associated NG production at oil wells. The largest share of oil related emissions comes from the following sources: (i) unintentional fugitive emissions throughout the life cycle, (ii) CH_4 emitted from venting, mostly at the production site, and (iii) incomplete flaring. EFs related to the oil life cycle were reviewed from four studies (*64*, *120–122*) as shown in Table 4. The EFs span two orders of magnitude, and the IPCC EF for developing countries is four times the amount of developed countries. As a reference, EPA's EF assumes that 2% of produced NG is vented or released unintentionally (see Appendix B for details). In comparison, the IPCC study best reflects results from a 1990 U.S. Department of Energy (DOE) study by Barns and Edmonds (*123*, *124*), which suggests that 20% of all non-marketed associated NG escapes to the atmosphere through venting and fugitives, and 80% is flared. However, the authors acknowledge significant uncertainties in these figures. Furthermore, it is possible that the share of venting and fugitives has decreased significantly since 1990. More details about the EFs including their assumptions are discussed in Appendix B.

Table 4: Summary of oil CH_4 emissions factors from four different studies.

	EF [kg CH_4/m^3 oil]	**Notes (see Appendix B for details)**	**Sources**
EPA 2013			
Average (1990-2010)	2.9	Based on reported total emissions and EIA oil production statistics	(*64*)
95% C.I., low	2.2		
95% C.I., high	7.2		
Wilson et al. 2004, 2008			
low	0.8	Same as EPA; required allocation between oil/NG	(*121*), (*122*)
high	6.9		
IPCC 2006			
Developed world (weighted average)	11	As reported in source	(*120*)
Developing world low	11		
Developing world high	41		

A particular challenge in estimating CH_4 emissions from the oil life cycle is quantifying incompletely flared associated gas. Emissions inventories typically account for fugitive emissions, and use relatively high combustion efficiencies for flares, usually in the order of 99%, i.e., not all flared CH_4 is emitted as CO_2. For instance, the EF for flaring used in the Wilson *et al.* (2004) study (*121*) corresponds to a flaring efficiency of 99.93% (see Table B-2).

In Wilson *et al.* (2010), 99.65% of flared NG is emitted as CO_2 (*122*). In the EPA 2013 GHG inventory, flaring efficiency is 98% (*64*, *125*). However, a literature review indicates that such high efficiencies can be observed only in laboratory experiments (*126*), whereas industrial scale efficiencies may be significantly lower. Recent studies suggest flaring efficiencies significantly below 90%. See Appendix B for a more detailed discussion of flaring efficiency estimates in the literature. Given the low flaring EFs in

the studies in Table 4, CH_4 from incomplete flaring will be accounted for separately when using the EFs above for estimating total CH_4 emissions from the oil industry.

NOAA provides the latest estimates of worldwide NG flaring (*127*, *128*). The data over the period 1994-2010 is based on satellite sensor measurements with dedicated thermal bands, which can distinguish gas flares from anthropogenic lighting, such as cities. Figure 15 summarizes the NOAA data for the Top 10 flaring countries and rest of the world (ROW). Most of this NG and the contained CH_4 is combusted and released to the atmosphere as CO_2. As discussed above, however, industrial flares are not 100% efficient. Thus, a fraction of the total NG flared (on average 105 Tg NG/yr globally) is released as CH_4, adding to the overall CH_4 emissions from the oil industry, which is captured in the EFs in Table 4.

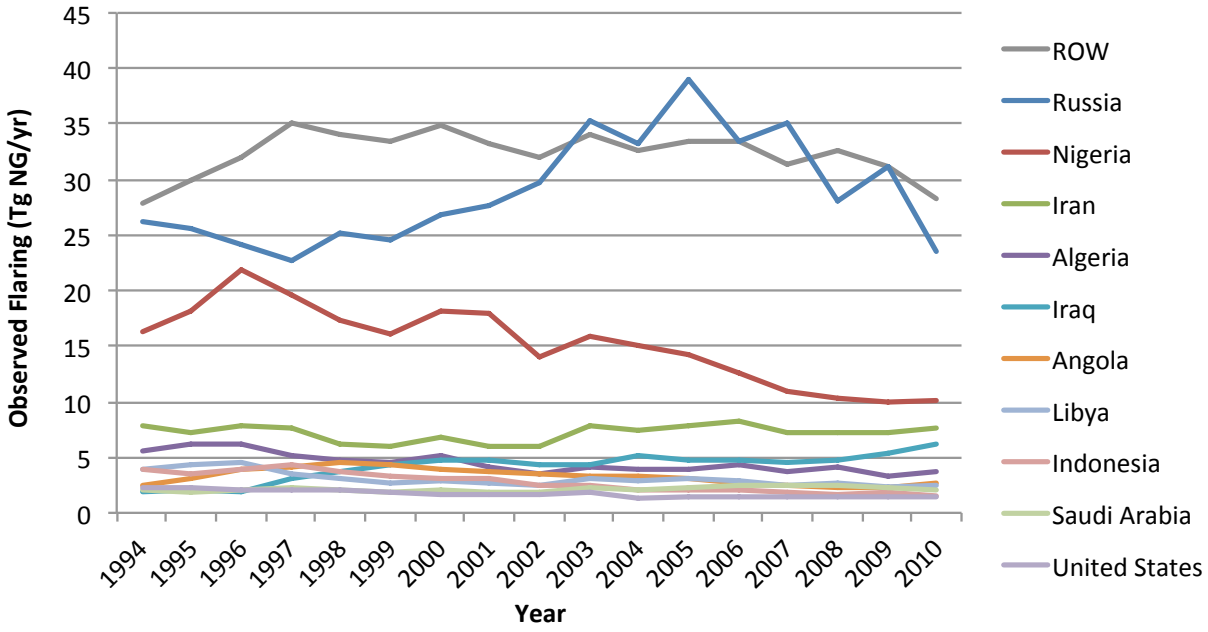

Figure 15: Observed NG flaring (from oil and gas fields) by country using NOAA data (*127*).

Flaring is particularly high in developing countries as shown in Figure 15. As described in detail in the literature, this phenomenon is mostly due to undeveloped infrastructure to transport and process the associated NG, and sell it to the market (*129*, *130*). Appendix B illustrates this phenomenon more clearly by showing the ratio of NG flaring to oil (and thus associated gas) production (Appendix Figure B - 1). Note that NOAA's flaring data only account for actually observed NG flaring and associated incomplete flaring CH_4 emissions. It does not account for CH_4 vented to the atmosphere, e.g., through improperly operated flares or intentional NG venting. Hence, the flaring estimates and the associated incomplete flaring CH_4 emissions represent a lower bound for oil CH_4 emissions in addition to CH_4 accounted for in the EFs above.

4.2.3 **Emissions factors from the coal industry**

Hydrocarbon emissions related to coal production depend on three main parameters (*131*): (i) the ratio of underground and surface mining because underground EFs are about one order of magnitude higher than for surface mining. This work establishes a database of coal production from underground and surface mining for major coal producing countries (see section 4.2.4 and Appendix Table B - 6) using literature sources. (ii) The amount of utilized coal CH_4 as an energy source, and the amount of flared CH_4 from the mine, which may change significantly over time (see Appendix B). (iii) EFs for each mining method in a specific coal basin, which depends mainly on the depth and rank of coal. Table 5 summarizes ranges of coal CH_4 EFs from underground and surface mining as well as post-mining activities and abandoned mine emissions (see Appendix B for definitions and further description). Given the magnitude and spatial variability of underground EFs, establishing a country-level production database of underground mining is important for estimating global CH_4 emissions from coal production. The country specific coal EFs in Table 5 cover 85% of world coal production, and the Tier 1 EFs may be used for the remaining countries. More details regarding the country-specific EFs and a discussion about the data sources are provided in Appendix B.

Table 5: Summary of literature coal CH_4 emissions factors from underground and surface mining as well as different life cycle stages. Units are m^3 CH_4/t (metric ton) coal.

	Range	Mean	Notes	Sources
Underground EFs				
IPCC Tier 1	10-25	n/a	Globally representative values	[1,2]
Major producers	7-24[a]	n/a	FSU[b], Germany, Poland, UK, Czech Republic, Australia	
USA	11-15	12	Range, IPCC; Mean, EPA	
China	10-12	11	Province-level averages; see discussion in Appendix B	[3,4]
Surface EFs				
IPCC Tier 1	0.3-2.0	1.2	Globally representative values	[1]
Post-mining EFs				
Underground	1.4-1.5	1.5	U.S. data, base year 2009; Range is due to temporal variability in reported data (1990-2010)	[1,2]
Surface	0.2-0.3	0.2		
Abandoned mine EFs				
Underground	1.1-1.5	1.3		

[a] See Appendix Table B - 3; [b] Former Soviet Union; [1] (*120*); [2] (*64*); [3] (*132*); [4] (*133*)

4.2.4 **Activity data**

The amount of fuel produced in a given year and/or country is referred to as activity data. This data is needed in an emissions inventory for generating emissions time series of a given fuel in combination with the associated emissions factors. EIA provides FF (NG, oil, coal) production data for years starting in 1980 until today including 230 countries

(*134*). EIA data are used in this work, and differences compared to other sources are discussed in Appendix B. Global production data since 1980 is summarized in Figure 16.

Country specific FF *production* data are only a proxy for the spatial distribution of emissions. Emissions may not always be released in the country where fuel *extraction* occurred. The vast majority of coal (Table 5) and oil related CH_4 emissions occur at the extraction site, thus using *production* statistics is reasonable. A fraction of CH_4 emissions from NG produced in Russia, for instance, may be released in Central European local distribution systems where it is transported via pipelines. A share of the pipeline emissions and all of the distribution related emissions of this NG is therefore associated with the NG consumption data of the importing country as well as any country where the pipeline transits (see Appendix B for details). It is assumed here that all emissions occur in the country where the FFs are produced due to lack of data, and given the uncertainties associated with the spatial allocation of CH_4 emissions in the grid maps as described in more detail in section 4.5.

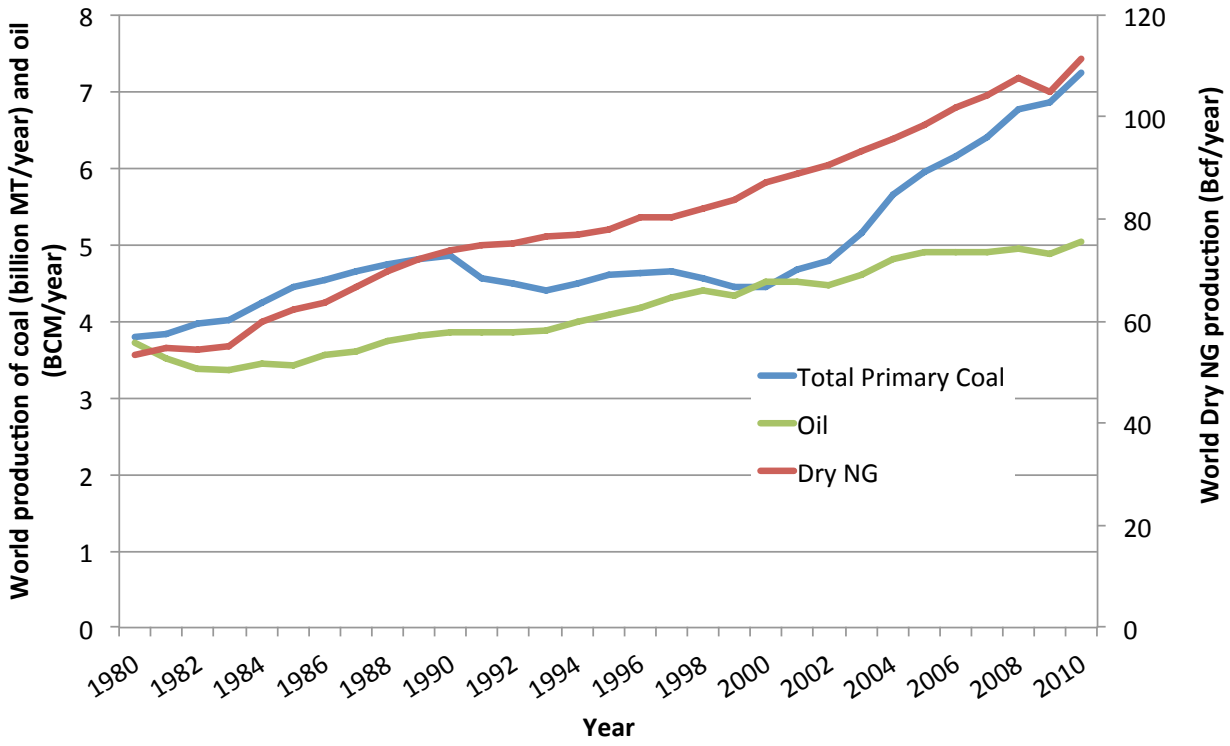

Figure 16: Summary of FF production data from the Energy Information Administration (*134*).

As shown in section 4.2.3, differentiating coal production by type of mining is important because EFs for underground mining are significantly higher compared to surface mining. Several sources document underground/surface breakdown for 23 countries representing 95% of world coal production throughout 1980-2010. As summarized in Appendix Table B - 6, underground coal production varies between 5-100% of total coal production among countries. Some data sources provide the underground/surface

breakdown for only a single year. In these cases, this value is used for the full time series from 1980-2011.

China is the world's largest coal producer accounting for 53% of global underground coal production over 1980-2010 (*134*). According to Raven Ridge Resources, illegal coal production is in the order of 10% of reported total production (*135*), i.e., total reported Chinese coal production (*134*) is likely underestimated. Furthermore, two sources (*136*, *137*) indicate a significant coal production data mismatch among different Chinese government agencies, which suggests that the EIA coal production data during 1997-2008 are underestimated. The Chinese sum of regional coal consumption data is ~20% higher than production data after accounting for import/export but consumption data may be over-reported to match politically desired gross domestic product output. This underestimate ranges from 1-23% for different years, and about 8% over the entire period. This is equivalent to 4% of world coal production during this period, and a higher value when only CH_4 emissions intensive underground coal production is considered. According to (*137*), Chinese coal production could be even higher, but no quantitative estimate is provided.

4.3 Non-FF CH_4 and C_2H_6 emissions data collection

This section summarizes the literature review of non-FF CH_4 and C_2H_6 emissions as well as the range of values chosen for atmospheric modeling in Chapter 5. *Prior* CH_4 emissions ranges were selected based on five of the most recent inversion studies (*39*–*41*, *43*, *44*) and two literature reviews (*138*, *139*) as shown in Figure 17. Both literature review studies encompass a total of 31 bottom-up and inversion studies, which cover most of the *posterior* ranges of the five recent inversion studies. This is particularly true for the lower bounds, which are important for constraining NG FE (see section 3.2.3). While the inversion studies suggest a wide range particularly for natural and agriculture/waste emissions (see all sub-categories in Table 10), relatively low estimates in one source category are balanced by higher estimates in another category.

The yellow vertical lines (and associated values in yellow boxes) represent the ranges chosen as *priors* for atmospheric modeling described in Chapter 5. The lower bounds were selected by including the lower bound *posterior*s of all source categories in the five inversion studies and the lower bounds of the literature reviews. Two exceptions were made. First, the lower bound estimates of the relatively small termites, oceans, and marine sediment sources as reported in Wuebbles and Hayoe (*139*) were excluded in the natural CH_4 category since these are one magnitude lower than best estimates. The adjusted lower bound natural emissions are the same as in IPCC (*138*). The selected lower bound natural emissions are consistent with a recent inter-model comparison of process-based wetland models (*140*) indicating a lower bound of 141 Tg CH_4/yr. Second, NOAA's biomass burning category was excluded, which is likely underestimated due to insufficient representation of small fires in its *prior* database and choice of *prior* uncertainties (*39*). Given the balance among source categories, e.g., relatively low natural emissions compensated by higher agriculture/waste emissions in the same study, it appears very unlikely that all *priors* chosen for modeling in Chapter 5

could be at the lower bound *simultaneously*. Yet, this scenario will be used in Chapter 5 for estimating upper bound NG FE rates. The *prior* means represent the average of the five inversion study *posteriors* minus 10% as a conservative estimate (the upper bound NG FE estimate increases as non-FF estimates decrease). The upper bound was selected such that lower and upper bound represent a normal distribution around the mean.

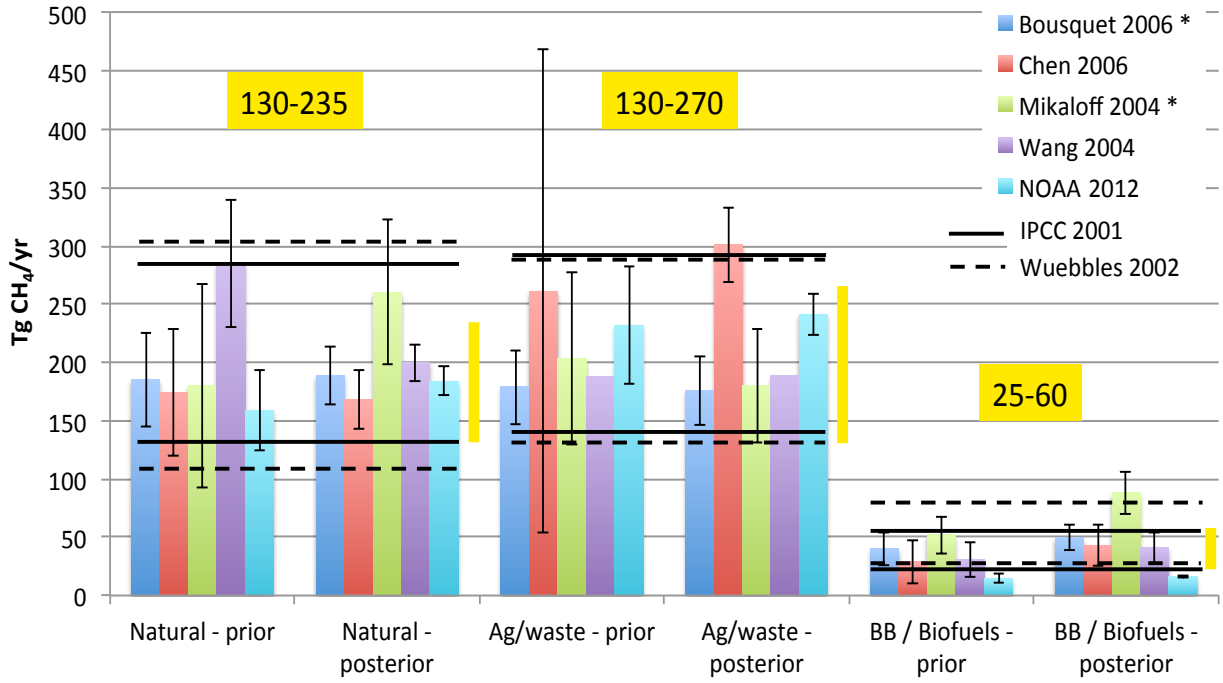

Figure 17: Summary of literature review for non-FF CH_4 sources from bottom-up and atmospheric modeling studies. Bars and horizontal lines indicate relatively recent CH_4 inversion studies (*39–41*, *43*, *44*) and literature reviews (*138*, *139*), respectively. The bounds of the literature reviews indicate the mean values of the studies with the lowest and highest values. *Posterior* uncertainties represent one std. dev. in Bousquet (std. dev. were added for each reported subcategory) and NOAA. Two std. dev. in Wang, but agriculture/waste uncertainties were only reported as the sum of FFs and soil sink uncertainties (details not reported in Chen and Mikaloff). Asterisk denotes that study includes isotopic analysis. Vertical yellow lines (with values in yellow boxes) show ranges chosen as priors for Chapter 5 (see text for selection criteria).

Biomass burning (BB) and biomass fuel combustion (BFC), the latter sometimes referred to as biofuels, are the two major non-fossil sources of atmospheric C_2H_6 (*84*, *141*). Table 6 summarizes literature values for these sources as well as the range of values chosen for atmospheric modeling. BB emissions were estimated using emissions factors (g CH_4 and C_2H_6 per kg dry matter burned) and total CH_4 emissions excluding BFC (Tg CH_4/yr). BFC emissions were estimated using the CH_4/C_2H_6 emissions ratio of BFC (*142*) and total BFC emissions (Tg CH_4/yr) in the developing world (*143*). Total estimates of combined BB and BFC emissions vary between 2-11 and around 5 Tg

C_2H_6/yr in bottom-up and atmospheric modeling studies, respectively. The *prior* values chosen in this work range from 2-9 Tg C_2H_6/yr, thereby giving slightly more weight to (*84*), which already constrains bottom-up values using atmospheric measurement data.

Table 6: Summary of literature review for non-FF ethane (C_2H_6) sources from bottom-up and atmospheric modeling studies as well as values chosen for this work. Units are Tg C_2H_6/yr.

	Biomass burning	Biomass fuel combustion	Total
Bottom-up studies	(*142*)	(*143*)	-
low	1.6	0.6	2.2
medium	3.6	1.7	5.3
high	7.5	4.0	11.5
Atmospheric modeling studies	(*84*)	(*84*)	-
low	2.4	2.6	5.0
high	2.8		5.4
This work			
low	1.6	0.6	2.2
medium	3.6	2.3	5.9
high	5.2	4.0	9.2

Natural CH_4 seepage may be a significant source of atmospheric CH_4, which is not currently accounted for in most of the bottom-up emissions inventories and global CH_4 inversion studies described above. The only exception is the Wuebbles *et al*. literature review (*139*) where the low estimate in the study accounts for 12 Tg CH_4/yr from natural seepage. Recent research suggests that geologic sources of CH_4 from earth's natural degassing represents an important factor in explaining atmospheric measurements (*144*). Seepage can occur as visible macro-seeps, marine seepage, micro-seepage, and in geothermal/volcanic areas. According to sample data – extrapolated globally – between 40-60 Tg CH_4 and 2-4 Tg C_2H_6 is released to the atmosphere from natural seeps (*144*), and about 80% of this amount is thermogenic (*145*). This is about 10% and 20% of the global CH_4 and C_2H_6 budget, respectively, and the source is isotopically and geographically very similar to FF CH_4. Hence, adding natural seepage as a source in atmospheric modeling may significantly reduce the ceiling of the NG CH_4 budget. Global grid maps of natural CH_4 seepage are not readily available to include in the 3D atmospheric modeling, and generating these grid maps is beyond the scope of this work. However, the above global estimates will be incorporated in the CH_4 and C_2H_6 atmospheric model simulations.

4.4 *Prior* FF emissions inventory methods

4.4.1 Emissions allocation between NG and oil industries and gas composition

In LCA or other bottom-up inventories, allocation of emissions is needed to account for the fact that some emitting processes generate more than one product (*8, 14*). Co-

product allocation is not standardized for specific products and processes (*12*). As reviewed by Weber *et al*. (*10*), allocation procedures in NG LCAs vary considerably among studies contributing to differences in results. In this study, emissions must be allocated between NG and oil since both fuels are sometimes extracted from the same well (see section 4.2.2). Unassociated and associated NG refers to gas produced from NG and oil wells, respectively (see Table 1). Since *posterior* NG related CH_4 and C_2H_6 emissions will be calculated as NG = FF – Oil – Coal (see section 4.1), the following assumptions for allocating CH_4 emissions between NG and oil industries are made here (see also Table 7):

i. Flaring observational data (and related CH_4 and C_2H_6 emissions from incomplete flaring) refers to oil wells, i.e., unassociated NG flaring is treated as negligible.
ii. In addition to flaring, fugitive CH_4 and C_2H_6 emissions occur during oil production as represented in oil EFs described above.
iii. Marketed associated and unassociated gas (oil and NG wells) is the denominator for the NG FE rate because both gas sources are included in dry gas production statistics used to estimate NG FE (CH_4 and C_2H_6 emissions as a percentage of dry gas production).

Table 7: Assumptions for allocating CH_4 and C_2H_6 emissions among NG and oil industries. The term *industry* is used here to distinguish emissions from the different *products* NG and oil.

Life cycle stages	Oil industry	NG industry	Notes
Pre-production and extraction of NG and oil	Literature fugitive EFs from oil production; incomplete flaring [a]	Remainder of leaks and venting emissions from oil and gas wells not covered in oil industry	Marketed associated and unassociated gas (oil and NG wells) is the denominator for NG FE rate
Processing	Literature fugitive EFs from oil refining	Fugitives from NG processing	-
Transport and distribution	Shipping emissions from oil tanks (*64*) [b]	Fugitives from pipelines and distribution system	NG related LNG emissions are assumed negligible (*64*)

[a] See section 4.2.2 and Appendix B for flaring emissions details

[b] Shipping related emissions are also included in EDGAR (*85*) inventories and grid maps described in section 4.4.3.

The U.S. Bureau of Land Management (BLM) database (*104*) contains routine analyses of U.S. oil and gas well hydrocarbon composition (among others) collected as part of a continuous survey for occurrences of helium. This survey was conducted from 1917 until 1996, and it consists of 17,167 samples. The database was first analyzed for missing entries in crucial fields such as hydrocarbon contents and sample date and/or entry errors, e.g., publication date before sample date (see details in the results section in Table 11). CH_4, C_2H_6, propane (C_3H_8), butane (C_4H_{10}), pentane (C_5H_{12}), hexane (C_6H_{14}), nitrogen (N_2), and carbon dioxide (CO_2) mol-% were averaged for each year and over the period 1948-1997 (this is the period for which data entries for all of the

above components were available; composition for the remaining years was not estimated due to missing data entries for some hydrocarbons). While other components are also present in NG, the above contribute on average 99.5 mol-%. The composition results were then used in combination with the chemical properties data of CH_4, C_2H_6, C_3H_8, and C_4H_{10} as summarized in Appendix Table B - 8 to calculate wt-% NG composition, which is needed to estimate CH_4 and C_2H_6 emissions based on NG FE rate scenarios. The data was also used to estimate total energy content of the samples and dry gas.

Downstream NG composition was estimated by performing a mass balance of NG upstream and downstream of processing while accounting for pipeline standards regarding energy content of processed NG. EIA data (*97*, *146*) is available for the U.S. on an annual basis for production of marketed (wet) NG (upstream of NG processing), dry NG (downstream), and NGLs (individually for each component; downstream). In the mass balance, the following condition is met:

- $Marketed\ NG = Dry\ NG + NGLs$ (in total and for each hydrocarbon)

The following assumptions were made:

- Of the total annual upstream NG volume, each hydrocarbon weight fraction is equal to EIA marketed (wet) NG multiplied by the respective weight fraction from the BLM samples of the same year. In other words, the BLM composition samples (converted to wt-%) are representative of wet NG.
- The difference between the reported marketed NG flows of higher HCs (C_2H_6, C_3H_8, and C_4H_{10}) and reported NGL outputs (also C_2H_6, C_3H_8, and C_4H_{10}) are returned to pipeline quality gas at the tail end of NG processing.

Non-hydrocarbons (especially N_2 and CO_2) in the BLM samples were excluded for this mass balance to represent the fact that non-hydrocarbons are already removed in marketed (wet) NG (*147*). Since dry NG must meet pipeline quality standards (see section 4.2.1), the energy content of the resulting dry NG composition from the mass balance was checked against these standards.

Mass balance Eq. 7 through Eq. 14 are given below, where $V_{wet,x,i}$ and $V_{wet,CH4,i}$ are absolute wet NG volume streams of component x (C_2H_6, C_3H_8, C_4H_{10}) and CH_4, respectively, in year i (to be processed); $VF_{wet,x,i}$ and $VF_{wet,CH4,i}$ are wet NG volume fractions from the BLM database; $P_{wet,i}$ and $P_{dry,i}$ are wet and dry NG production (in m^3), respectively; $V_{dry,x,i}$ and $V_{dry,CH4,i}$ are the absolute dry NG volume streams; $NGL_{x,i}$ is the absolute NG liquids volume stream of component x (*146*); $VF_{dry,x,i}$ and $VF_{dry,CH4,i}$ are dry NG volume fractions; $WF_{dry,x,i}$ and $WF_{dry,CH4,i}$ are dry NG weight fractions using the gas densities ρ in Table B - 8. Note that C_5H_{12} and C_6H_{14} are in liquid phase at standard temperature and pressure. Since these two components contribute less than 1 mol-% to NG in the BLM database, they were omitted from the mass balance.

$$V_{wet,x,i} = VF_{wet,x,i} \times P_{wet,i} \quad \textbf{Eq. 7}$$

$$V_{wet,CH4,i} = VF_{wet,CH4,i} \times P_{wet,i} \quad \textbf{Eq. 8}$$

$$V_{dry,x,i} = \left(V_{wet,x,i} - NGL_{x,i}\right) \quad \textbf{Eq. 9}$$

$$V_{dry,CH4,i} = V_{dry,i} - \sum_{x} V_{dry,x,i} \quad \textbf{Eq. 10}$$

$$VF_{dry,x,i} = V_{dry,x,i} \div V_{dry,i} \quad \textbf{Eq. 11}$$

$$VF_{dry,CH4,i} = V_{dry,CH4,i} \div V_{dry,i} \quad \textbf{Eq. 12}$$

$$WF_{dry,x,i} = \frac{VF_{dry,x,i} \times \rho_x}{VF_{dry,CH4,i} \times \rho_{CH4} + \sum_x VF_{dry,x,i} \times \rho_x} \quad \textbf{Eq. 13}$$

$$WF_{dry,CH4,i} = \frac{VF_{dry,CH4,i} \times \rho_{CH4}}{VF_{dry,CH4,i} \times \rho_{CH4} + \sum_x VF_{dry,x,i} \times \rho_x} \quad \textbf{Eq. 14}$$

4.4.2 Country-specific CH_4 and C_2H_6 emissions from NG, oil, and coal

Prior NG and oil CH_4 and C_2H_6 emissions were estimated for those countries representing the top 96% and 91% of 1980-2011 world production, respectively (see Appendix Table B - 7). NG and oil estimates represent 37 and 26 countries, respectively, out of 233 countries and world regions represented in EDGAR (*85*). EIA production statistics were used as described in section 4.2.4.

Prior emissions estimates for the remaining countries were adopted from EDGAR (*85*). Choosing this relatively small set of countries allows focusing on finding the best available data for the key emissions areas of the world, and comparing those with EDGAR. To provide a context, even if focusing on this small set of countries resulted in overlooking a hypothetical 50% error in EDGAR emissions for the sum of the remaining countries (less than 10% of world production), this would change global emissions only in the order of 5%. *Prior* emissions were generated from 1980 until 2011 to match the time series for which NOAA's observational CH_4 data is available.

In order to bound *prior* NG FE estimates from the literature, three NG FE scenarios were established as shown in Table 8. FE rates were distinguished between OECD countries, Russia (or FSU), and ROW, and included a FE reduction factor over time. Both distinctions (spatial and temporal) are meant to represent potential improvements in infrastructure that reduce emissions from NG production over time, but at different intervals depending on world region. The emphasis is on keeping assumptions transparent in order to facilitate interpretation of inversion results. Based on the literature review in Table 3, EFs are mainly available for the U.S., some European countries, and Russia. Note that values in Table 8 include lower and upper bounds from individual studies (see footnotes of Table 8), whereas Table 3 summarizes only mean values. The IPCC Tier 1 recommendations indicate higher NG FE and associated emissions in developing countries compared to OECD (*120*). While data for other developing countries is limited, the scenarios assume highest FE in FSU while FE in the ROW are averaged between OECD and FSU. FE reductions in FSU in this model

started in 1990 with the collapse of the Soviet Union, which is consistent with conclusions from previous CH_4 inversion studies and reviews, e.g., Dlugokencky *et al.* (*81*). In addition to the FE rates in Table 8, atmospheric modeling was carried out for scenarios using global average, constant FE rates over time, such as 1%, 3%, 5%, 7%, and 9%.

Table 8: NG FE scenarios used to bound the literature values from section 4.2.2.

	OECD*		FSU		ROW[i]	
	1980	**2010**	**1980-1990**	**2010**	**1980**	**2010**
Low	3.0%	1.5%[a]	6.5%[d]	2.9%[f]	4.8%	2.2%
Medium	4.0%	2.3%[b]	7.25%	4.5%[g]	5.6%	3.4%
High	6.0%	6.0%[c]	8.0%[e]	6.0%[h]	7.0%	6.0%

[a] (*34*), low; [b] (*10*), mean; [c] (*28*), high (conventional NG); [d] (*37*), low; [e] (*37*), high; [f] Gas production/processing 2% from (*34*), mean, plus transport/distribution 0.9%, (*67*), low; [g] Gas production/processing 2% (*34*), mean, plus transport/distribution 2.5% (*67*), high; [h] (*28*), high (conventional NG); [i] Average between OECD and FSU; * OECD countries as listed in Appendix Table B - 7.

Country-specific absolute NG industry CH_4 and C_2H_6 emissions $E_{NG,CH4,i}$ and $E_{NG,C2H6,i}$ in year i were estimated as in Eq. 15 and Eq. 16 based on FE rate $FE_{NG,i}$ and dry NG production $P_{NG,dry,i}$ (in units of Tg/year). The dry NG CH_4 weight fraction $WF_{dry,CH4}$ is calculated in Eq. 14 (average of $WF_{dry,CH4,i}$ over time). Using dry NG weight fraction for CH_4 is appropriate in combination with dry NG production. It is compatible from a mass balance perspective, and it allows comparing NG FE results with most other studies, which also used dry NG as a basis. A weighted average dry and wet NG C_2H_6 weight fraction would be needed to reflect that (i) a significant portion of C_2H_6 is removed in the dry gas stream, and (ii) the ratio between upstream and downstream FE is not necessarily 1:1, and highly uncertain (see Table 3). If dry NG C_2H_6 weight fraction was used instead, the *prior* C_2H_6 emissions may be underestimated significantly due to upstream NG FE and associated higher C_2H_6 content. As a compromise, the NG C_2H_6 weight fraction was chosen based on wet composition data, and used in combination with dry NG production data. As shown in Table 2, the wet NG C_2H_6 weight fraction ranges from 5.2-8.7% in the U.S. and internationally. In order to reflect this large range and associated uncertainty in the literature, three scenarios for the NG C_2H_6 weight fraction WF_{C2H6} were chosen: low (5.2%), medium (7.0%) and high (8.7%). Applying the lower dry NG production data compensates for the fact that using wet NG C_2H_6 weight fraction may overestimate C_2H_6 emissions from FE downstream.

$$E_{NG,CH4,i} = FE_{NG,i} \times P_{NG,dry,i} \times WF_{dry,CH4} \quad \textbf{Eq. 15}$$

$$E_{NG,C2H6,i} = FE_{NG,i} \times P_{NG,dry,i} \times WF_{C2H6} \quad \textbf{Eq. 16}$$

Out of the different literature values for oil EFs, which span two orders of magnitude, the relatively low EPA CH_4 EFs in Table 4 (*64*) were chosen as a representative global average (2.9 kg CH_4/m^3 oil). As described in section 4.2.2, the 2012 EPA EFs are

equivalent to 2% of non-marketed associated NG released to the atmosphere from venting and fugitive sources, as opposed to the 20% by DOE in 1990. Thus, EFs may in fact be higher than the EPA estimate, but this choice provides a lower bound for oil industry emissions. Ignoring potentially unaccounted CH_4 from oil well venting and other fugitives in the inversion *priors* also provides a lower bound for the total *prior* FF CH_4 budget. Hence, considering a given *posterior* FF CH_4 budget from the inversion in Chapter 5 as well as the relatively low 2012 EPA EFs generates an upper bound for NG industry emissions.

In addition to EPA EFs, NOAA flaring observations were used to estimate CH_4 emissions from incomplete flaring. NOAA observations are available only from 1994 to 2010, but *prior* emissions start in 1984. Flaring data were extrapolated based on the ratio of flaring to oil production. The ratio is relatively constant over time for most countries, and some have higher ratios in earlier years (see Appendix Figure B - 1). The 1994 ratio was used for pre-1994, which is a conservative estimate, since the trend indicates that the ratio could be higher. Based on the literature review in Table 2 and section 4.2.2, an optimistic 95% flaring efficiency and 40 wt-% associated (flared) wet NG CH_4 content were chosen as conservative values, i.e., leading to a lower bound of flaring emissions. The influence of a more pessimistic 80% flaring efficiency estimate on NG FE will be discussed in Chapter 5.

Country-specific absolute oil industry CH_4 emissions $E_{Oil,CH4,i}$ in year i (Eq. 17) were estimated based on emissions factor $EF_{Oil,CH4,i}$, NG production $P_{Oil,i}$ (in m^3 oil/year), flaring efficiency $EF_{Flare,i}$, associated (flared) wet NG CH_4 content $WF_{assoc,flare,CH4}$, and observed flaring amount $P_{Flare,i}$ (in Tg NG/year, see Figure 15). Oil industry C_2H_6 emissions $E_{Oil,C2H6,i}$ (Eq. 18) were scaled from $E_{Oil,CH4,i}$ using the associated wet NG composition (wt-%) ratio of CH_4 and C_2H_6, $R_{oil,CH4/C2H6}$, from the literature. Given the uncertainty among sources (1.7-3.3, see Table 2), three ratio scenarios were chosen: low (3.3), medium (2.5) and high (1.7) C_2H_6 content.

$$E_{Oil,CH4,i} = EF_{Oil,CH4,i} \times P_{Oil,i} + EF_{Flare,i} \times WF_{assoc,flare,CH4} \times P_{Flare,i} \quad \textbf{Eq. 17}$$

$$E_{Oil,C2H6,i} = E_{Oil,CH4,i} \div R_{Oil,CH4/C2H6} \quad \textbf{Eq. 18}$$

Prior coal CH_4 emissions were estimated for 22 countries for which data regarding the share of underground and surface mining was available (see Appendix Table B - 6). These countries represent 95% of world primary coal production during 1980-2010. EIA production statistics were used as described in 4.2.4, except for Chinese coal throughout 1997-2008 where (*136*) was used. Despite this adjustment, it should be noted that Chinese coal production may in fact be higher (*137*).

Country-specific absolute coal industry CH_4 emissions $E_{Coal,CH4,i}$ in year i (Eq. 19 through Eq. 21) were estimated based on underground mining emissions factor $EF_{Coal,i,u}$, underground coal production $P_{Coal,i,u}$, surface mining emissions factor $EF_{Coal,i,s}$, and surface coal production $P_{Coal,i,s}$. $EF_{Coal,i,u}$ is the sum of mining emissions factor $EF_{Coal,i,u,m}$ and post-mining and abandoned mines emissions factor $EF_{Coal,i,u,pm\&am}$. $EF_{Coal,i,s}$ is the sum of mining emissions factor $EF_{Coal,i,s,m}$ and post-mining emissions factor $EF_{Coal,i,u,pm}$.

Coal industry C_2H_6 emissions $E_{Coal,C2H6,i}$ (Eq. 18) were scaled from $E_{Coal,CH4,i}$ using the coal-bed gas composition (wt-%) ratio of CH_4 and C_2H_6, $R_{Coal,CH4/C2H6}$, from the literature. Given the uncertainty among sources (see section 4.2.3), three ratio scenarios were chosen: low C_2H_6 (1000), medium C_2H_6 (100) and high (50) C_2H_6 content.

$$E_{Coal,i} = EF_{Coal,i,u} \times P_{Coal,i,u} + EF_{Coal,i,s} \times P_{Coal,i,s} \quad \textbf{Eq. 19}$$

$$EF_{Coal,i,u} = EF_{Coal,i,u,m} + EF_{Coal,pm\&am} \quad \textbf{Eq. 20}$$

$$EF_{Coal,i,s} = EF_{Coal,i,s,m} + EF_{Coal,pm} \quad \textbf{Eq. 21}$$

$$E_{Coal,C2H6,i} = E_{Coal,CH4,i} \div R_{Coal,CH4/C2H6} \quad \textbf{Eq. 22}$$

Emissions factors $EF_{Coal,i,u}$ and $EF_{Coal,i,s}$ are based on Table 5 (either averages between low/high estimates or best estimates) and summarized in Table 9. Note that these country specific EFs cover 86% of world underground coal production in 2010, and that the more uncertain IPCC Tier 1 EFs are incorporated for the remaining 14% of coal producing countries.

Table 9: Summary of EFs chosen for coal CH_4 emissions *priors* (in units of m^3 CH_4/t coal).

	Share of 2010 under-ground production	Emissions factors, mining only	Incl. post-mining & abandoned mines[a]	EF Sources
China	53%	11.4	14.1	(*133*), (*132*)
USA	13%	11.7	14.5	(*148*)
FSU	5%	20	22.8	(*120*)
Poland	5%	9.4	12.2	
Australia	4%	15.6	18.4	
Germany	2%	22.4	25.2	
Czech Republic	2%	23.9	26.7	
United Kingdom	2%	15.3	18.1	
Total	**86%**			
ROW	14%	17.5	20.3	(*120*)
Surface	n/a	1.2	1.4	

[a] (*64*)

4.4.3 Spatial distribution of country-level emissions using EDGAR CH_4 emissions grid maps

The spatial distribution of country-level emissions in EDGAR's global emissions grid maps (*85*) was adopted, which is based on population and other proxies (see section 3.2.3). Note that the grid maps only provide approximate spatial information of emissions sources on a country or regional level. However, the spatial resolution of the

inversion models is usually significantly lower than the grid maps due to computational efficiency constraints. Furthermore, the significance of lacking spatial detail on a country or regional level may be limited when analyzing simulation results for zonal averages as performed in Chapter 5. The absolute emissions in the grid maps were scaled based on the FF estimates generated in section 4.4.2. In order to scale the countries in the NG, oil, and coal maps individually in each year, the grid cells belonging to each country were identified. A national identifier map was generated using raw data provided by EDGAR (*85*) (land-based national boundaries) and the Flanders Marine Institute maritime boundaries geo-database to perform this step (*149*). Boundaries for maritime sovereign regions were required to allocate offshore emissions to individual countries, e.g., U.S. NG and oil drilling in the Gulf of Mexico. The EDGAR emissions raster was scaled for each country individually, such that the total emissions of each country in the grid map matches the country's emissions estimated in section 4.3. The procedure for updating EDGAR's grid maps was followed for all countries estimated individually, which account for 96%, 91%, and 95% of global NG, oil, and coal production, respectively (see Appendix Table B - 7 for details). Emissions for the remaining countries were left unchanged from the original EDGAR dataset. Individual grid maps for NG, oil, and coal sources were generated for each year of the inversion time frame.

A summary of the collected spatial emissions data sources for all emissions sources is provided in Table 10. All CH_4 sources and estimates will be used in the atmospheric modeling part described in Chapter 5. The "Other energy and industry" category was used directly as provided by EDGAR (*85*). Annual grid maps for NG, oil, coal, and "Other energy and industry" were linearly interpolated between each year to yield monthly grid maps. In contrast to EDGAR, estimates from the Global Fire Emissions Database (GFED) contain daily grid maps, which allows modeling the annual cycle of agriculture/waste CH_4. GFED annual cycle "Agriculture/Waste" emissions were adopted in order to scale EDGAR annual totals and grid maps of the same sources. "Other biomass burning" grid maps were used from the GFED database. "Other energy and industry" and "Agriculture/Waste" annual totals were linearly extrapolated from 2008 (last year in EDGAR) to 2011 using the last 10 years available in EDGAR. A more detailed description of the grid map scaling process using Matlab is provided in Appendix B.

Table 10: Data sources for emissions and emissions distributions used in 3D modeling.

	Emissions source data	Spatial distribution	Seasonal cycle	IPCC source codes (*120*)
FF				
NG	This work (section 4.4)	(*94*)	no	1B2b
Oil				1B2a
Coal				1B1
Other energy and industry				
Public power and heat, other energy industries, transportation, residential and other sectors, industrial processes, FF fires	(*94*)	(*94*)	no	1A1/2/3/4, 2, 7A
Natural sources				
Wetlands, enteric fermentation of termites and wild animals, oxidation in dry soils, oceans	(*88*)	(*150*), (*151*)	(*88*)	-
Agriculture/Waste				
Enteric fermentation, manure management, rice cultivation, other agriculture, agricultural waste burning, solid waste disposal, wastewater handling, waste incineration	(*94*)	(*94*)	(*152*)	4A/B/C/D/F, 6A/B/C
Other biomass burning				
Grassland and open savanna fires, woodland fires, deforestation and degradation fires, forest fires, peat fires	(*153*), (*154*)	(*155*)	(*155*)	4 E, 5 A/C/D/F

(*94*): EDGAR v4.2

4.5 *Prior* FF emissions inventory results

This section summarizes the results of the NG composition data analysis, the absolute country-specific CH_4 emissions from the NG, oil, and coal industries, and the spatial allocation of emissions. The emissions results (to be used as *priors* in Chapter 5) will be compared with EDGAR (*85*), which is widely used in atmospheric modeling. As described in Chapter 3, prior FF emissions were estimated in this chapter to provide more detail and transparency compared to EDGAR, which will be useful for interpreting inversion results in Chapter 5. The original BLM database (*104*) consisting of 17,167 samples was analyzed and samples with the following criteria were eliminated: (i) samples without sample date, (ii) where sample date was after publication date, (iii) where CH_4 content was lower than C_2H_6 content, (iv) samples not containing CH_4 and/or C_2H_6 data, (v) samples without C_3+ data, and (vi) samples not containing N_2 and/or CO_2 data. The remaining sample size is 6,109 (see Table 11).

Table 11: Summary of gas samples used for analyzing hydrocarbon composition from BLM database (*104*).

	Removed samples	Sample size
Original number of samples	-	17,167
No samples date	298	16,869
Sample date after publication date	1,525	15,344
CH_4 content < C_2H_6 content	322	15,022
No CH_4 and/or C_2H_6 data available	20	15,002
No C_3+ data available	8,480	6,522
No N_2 and/or CO_2 data available	413	**6,109 (Final sample size)**

Table 12 shows hydrocarbon composition of the 6,109 samples from 1948-1997, the period over which all hydrocarbon and N_2 and CO_2 data was available. The energy content of the BLM data (1,091 Btu/cft) is consistent with a recent survey among the U.S. NG industry (*156*) on inlet gas to processing plant (1,101 Btu/cft), i.e., between well gas and wet gas. Appendix Figure B - 2 shows annual averages over the period 1948-1997 including possible explanations for the inter-annual variations.

Table 12: Results of well gas sample analysis using BLM data (*104*).

	Units	CH_4	C_2H_6	C_3H_8	C_4+	CO_2	N_2	Total
Weighted avg. 1948-1997	mol-%	79.9	5.7	3.0	2.5	1.8	6.6	99.4
Std (all samples)		17.2	4.0	3.0	3.1	8.3	12.2	n/a
Weighted avg. 1948-1997	wt-%	64.7	8.6	6.5	8.1	2.9	9.1	100

The results of the mass balance are summarized in Table 13. It covers the period 1981-1997, during which both well samples and NGL production data was available. The upstream CH_4 mass flow was on average 0.3% lower than downstream. Since this is physically not possible, two phenomena may explain this result. First, this discrepancy may represent uncertainties in the average wet NG composition from Table 2 and Figure 14 as well as dry gas and NGL production data (*97*, *146*). Second, the difference may represent remaining N_2 and CO_2 in the dry gas, which was not accounted for in the mass balance.

Table 13: Summary of dry gas composition using an upstream-downstram mass balance approach.

	Units	CH_4	C_2H_6	C_3H_8	C_4+	ΔCH_4 (upstream-downstream)[a]	Energy content (Btu/cft)	
							Mass balance	Literature[b]
Weighted avg. 1981-1997	vol-%	93.0	4.6	1.4	0.9	-0.3	1,089	950-1,150
	wt-%	85.3	8.0	3.7	3.0	-0.3		

[a] Difference between upstream and downstream CH_4 mass flow, in % of dry NG production (see explanation above).
[b] See Appendix B for details.

Dry gas composition is consistent with pipeline quality standards, which includes a range of pipeline gas energy content between 950 and 1,150 Btu/cft (*157*). Dry NG composition results coincide reasonably well with NREL data (*158*), which suggests pipeline CH_4 and C_2H_6 contents of 94.4 mol-% and 3.1 mol-%, respectively. Given that CH_4 and C_2H_6 contents in Table 13 are lower than NREL, total NG emissions (and FE) estimates in this work may be conservative, i.e., underestimated, which serves the study objective of finding an upper bound of NG FE.

Figure 18 summarizes global CH_4 emissions over time based on the NG FE scenarios described in section 4.4.2. All three scenarios assume that FE decrease significantly over time due to replacement of old leaking infrastructure and/or management improvements. It will be analyzed in Chapter 5 whether this temporal decline in FE rates is consistent with observational constraints. Nevertheless, absolute emissions increase due to growing NG production.

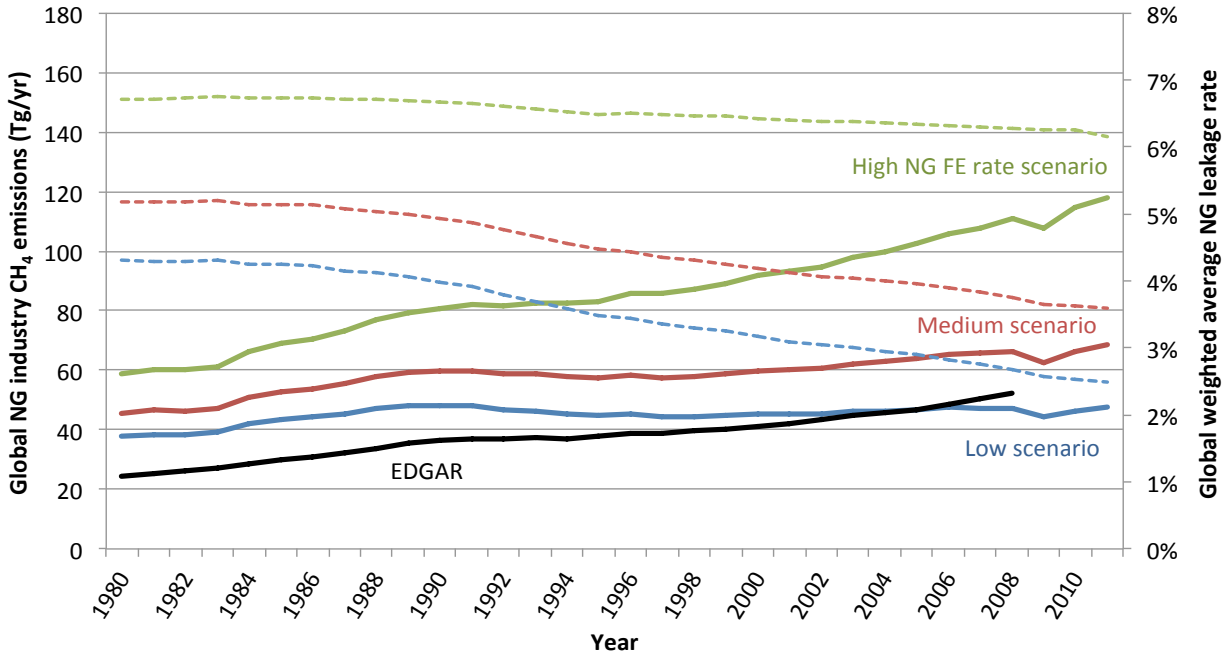

Figure 18: Global NG industry CH_4 emissions based on regionally and temporally specific FE rate scenarios in Table 8. Solid lines are emissions for each scenario (left axis), dashed lines are FE rates associated with each scenario (right axis), and EDGAR (*85*) (black) is shown for comparison.

It also shows that CH_4 emissions in all three scenarios are significantly higher than EDGAR (*85*) until about 2006, after which EDGAR overtakes the low FE scenario. Figure 19 shows the results for NG FE rates applied globally and constant over time. The comparison shows that EDGAR emissions assume a global average NG FE rate in the order of 3% provided that NG production data and NG CH_4 content are not significantly different from this analysis. However, note that while assuming the same NG FE rate in Figure 19, country-level emissions vary significantly between this work and EDGAR as shown in Appendix Table B - 9. For instance, prior CH_4 emissions in this

work are about 3, 5, and 7 times higher than EDGAR in Canada, the Netherlands, and Norway, respectively. Prior CH_4 emissions in other countries, such as the Ukraine, are up to 0.8 times lower than in EDGAR. As discussed in section 3.2.3, posteriors in previous inversion studies appeared to be strongly influenced by prior emissions. By using a larger range of potential NG FE rates, this work will clarify whether high FE rates are possible given a limited global CH_4 budget.

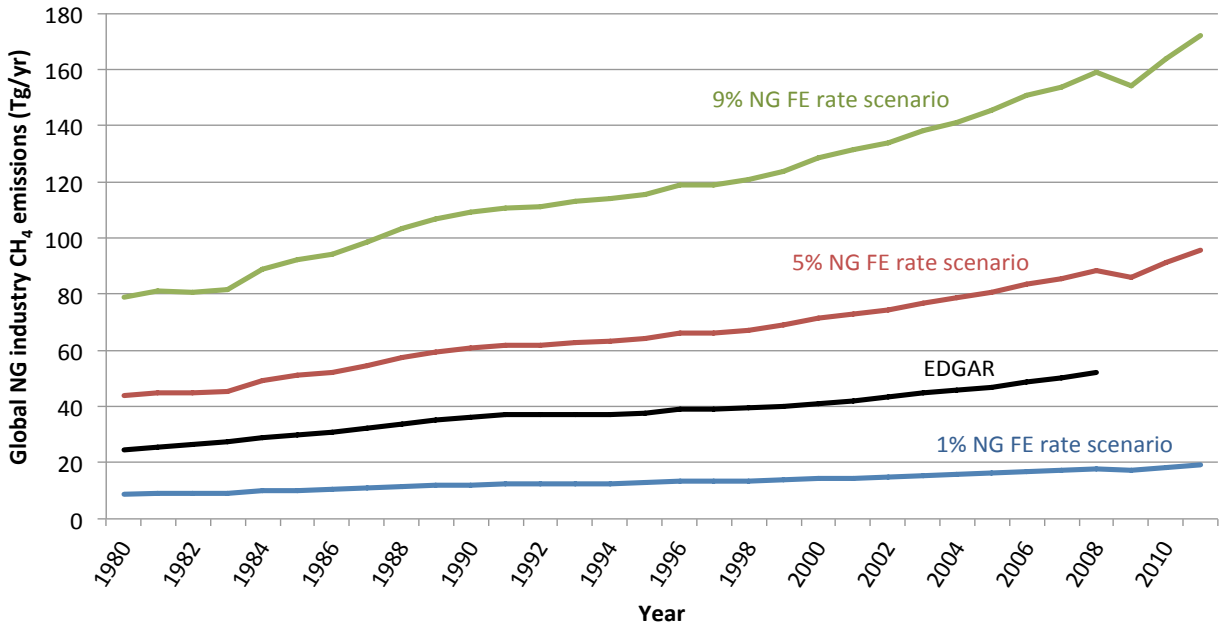

Figure 19: Global NG industry CH_4 emissions based on FE rates applied globally and constant over time (see section 4.4).

Figure 20 and Figure 21 show results for the global C_2H_6 emissions equivalent to Figure 18 and Figure 19 for CH_4, i.e., based on the NG FE scenarios described in section 4.4.2. Both Figures represent C_2H_6 emissions assuming the "low" NG C_2H_6 content scenario in Eq. 16, and the results for all NG C_2H_6 content scenarios are illustrated in Appendix Figure B - 3. C_2H_6 emissions over time increase at the same rate as CH_4 emissions given the same underlying NG production data. EDGAR (*85*) does not provide C_2H_6 emissions estimates, thus a comparison cannot be made.

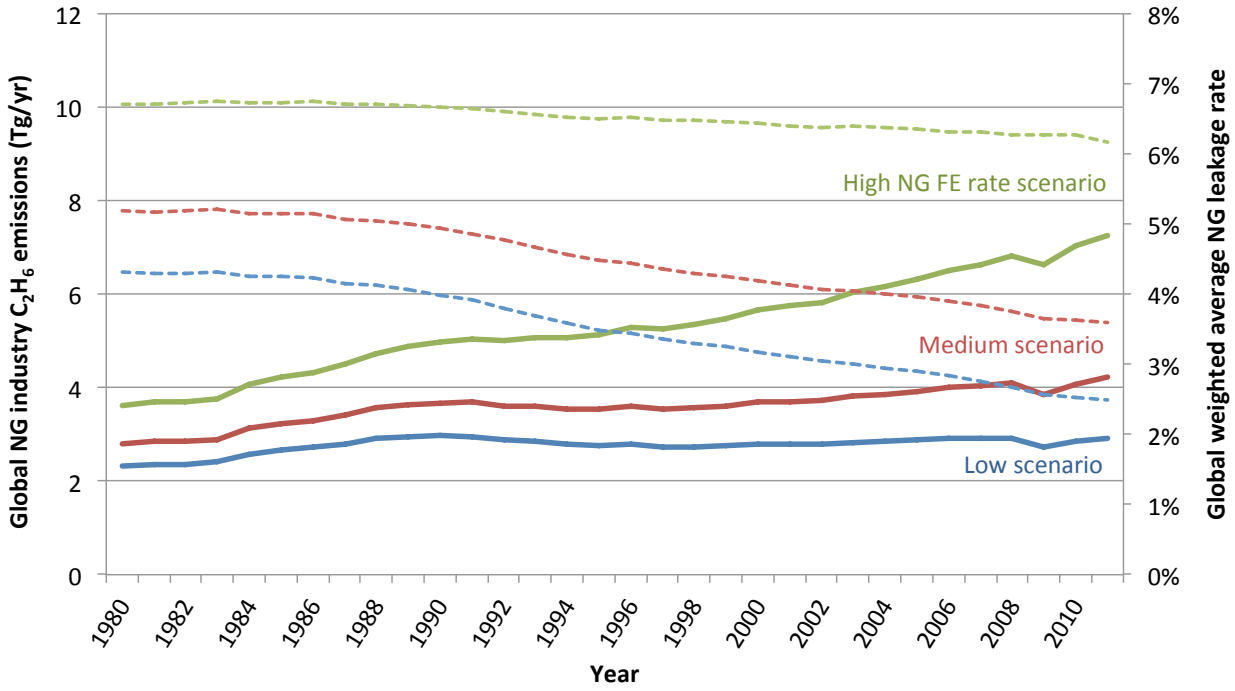

Figure 20: Global NG industry C_2H_6 emissions based on regionally and temporally specific FE rate scenarios in Table 8. Solid lines are emissions for each scenario (left axis), dashed lines are FE rates associated with each scenario (right axis). Note these results represent the "low" NG C_2H_6 content scenario in Eq. 16. The results for all NG C_2H_6 content scenarios are illustrated in Appendix Figure B - 3.

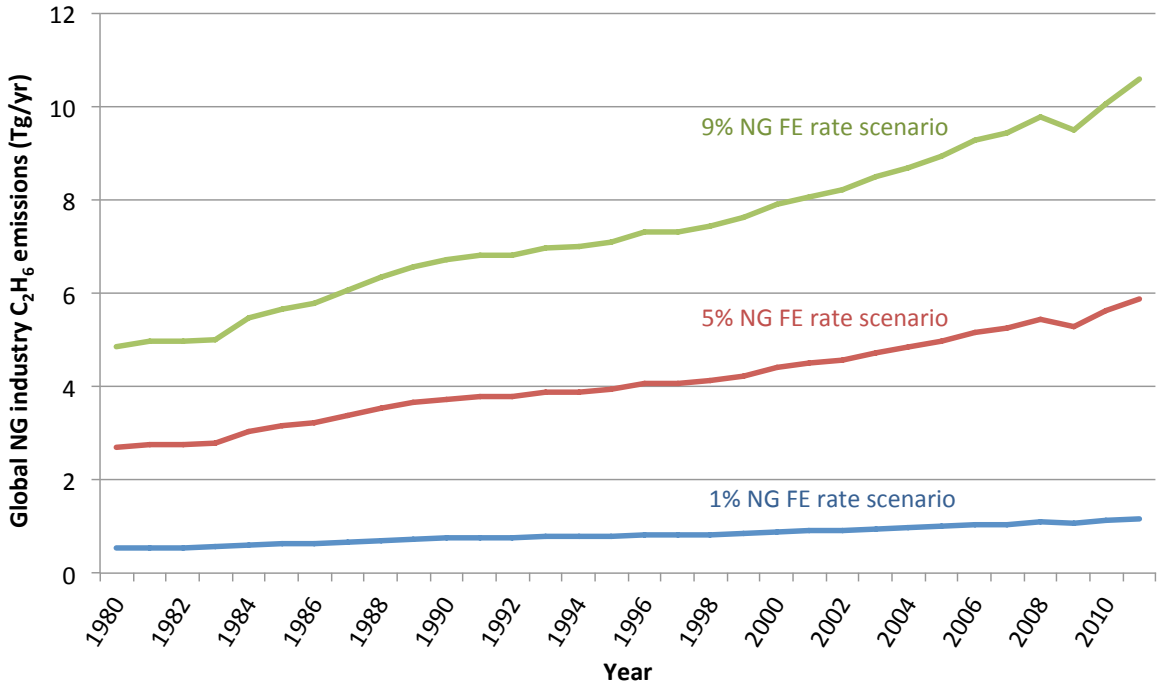

Figure 21: Global NG industry C_2H_6 emissions based on FE rates applied gobally and constant over time (see section 4.3) under the "low" NG C_2H_6 content scenario in Eq. 16.

Global oil CH_4 and C_2H_6 emissions based on Eq. 17 and Eq. 18 are summarized in Figure 22 in comparison with emissions using other EFs described in Table 4. As mentioned above, the EPA emissions factors (*64*) were used in this study to provide a lower bound for oil industry emissions in order to quantify an upper bound for NG FE. The results of this work coincide very closely with EDGAR (*85*) estimates, which are flanked by values from the Wilson *et al*. studies (*121*, *122*) in the Gulf of Mexico (GOM), i.e., GOM EFs applied globally. The low IPCC estimate (*120*) is significantly higher than all other sources while the high IPCC estimate is a factor of 4 greater than low IPCC for developing countries (see 4.2.2). While oil CH_4 emissions from EDGAR and this work are very similar, country-level emissions vary significantly between both studies as shown in Appendix Table B - 9. For instance, *prior* CH_4 emissions in this work are over 0.7 and 2.5 times higher than EDGAR in the U.S. and the U.K., respectively. *Prior* CH_4 emissions in other countries, such as Canada and Nigeria, are up to 0.6 times lower than in EDGAR.

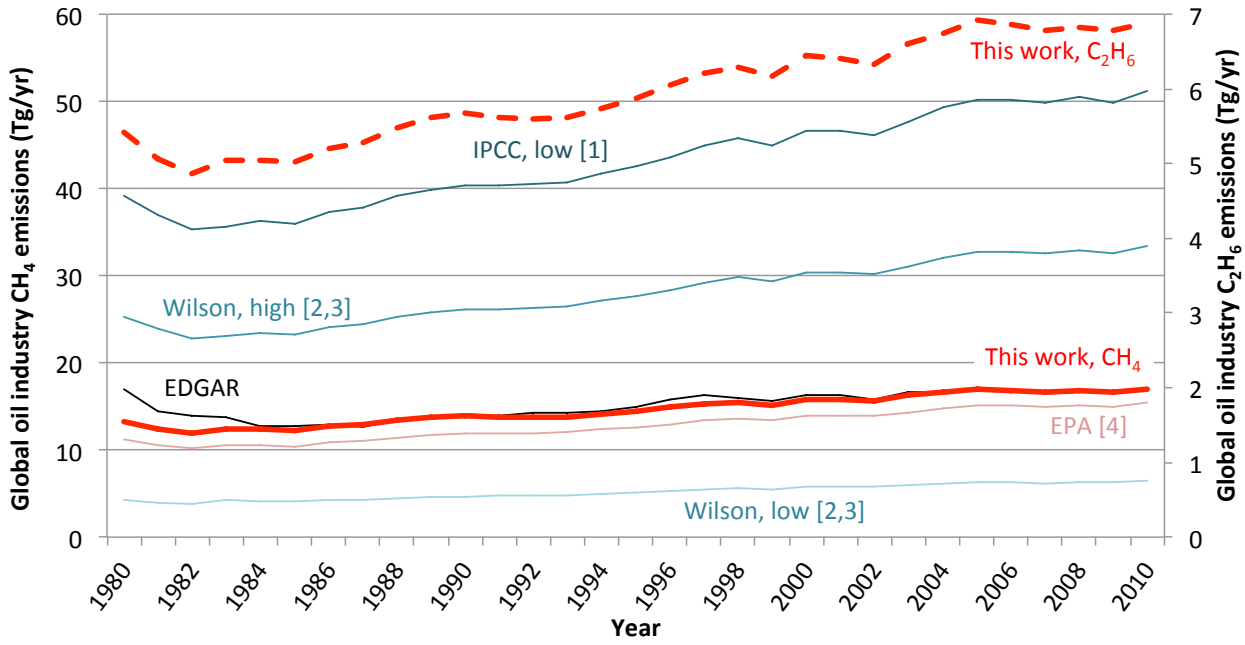

Figure 22: Global oil industry CH_4 and C_2H_6 (medium well C_2H_6 content scenario, dashed red line, right axis) emissions from this work based on (*64*) EFs plus emissions from incomplete flaring, and comparison with emissions using other literature EFs and data from EDGAR (*94*). [1] (*120*); [2] (*121*); [3] (*122*); [4] (*64*).

Global coal industry CH_4 and C_2H_6 emissions are summarized in Figure 23. The global results of this work are nearly identical to EDGAR between 1980-2002. Past 2002, EDGAR emissions increase at a greater rate, largely due to higher emissions estimates in China, the U.S., India, and Russia. Similar to NG and oil, differences in coal emissions for individual countries vary significantly between this work and EDGAR. In fact, emissions for 17 (12) out 22 countries estimated here differ by at least 10% (25%) from EDGAR (see Appendix Table B - 9).

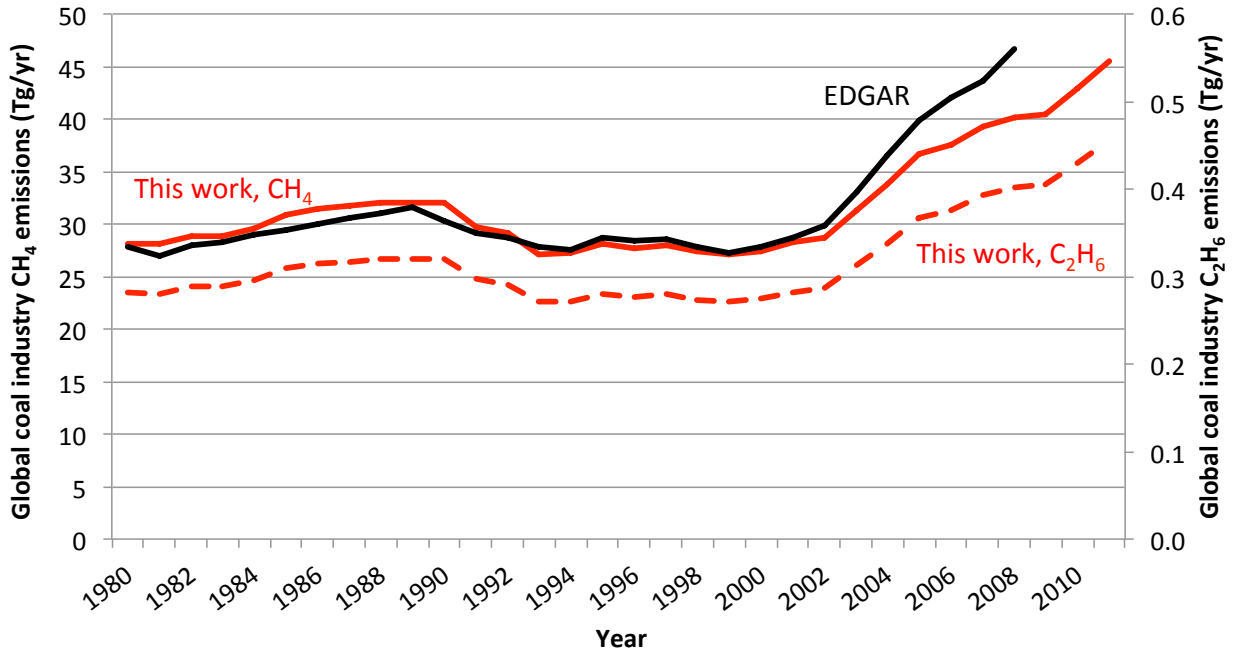

Figure 23: Global coal industry CH_4 and C_2H_6 (dashed red line, right axis) emission results in comparison with data from EDGAR (*94*).

Grid maps used for atmospheric modeling were developed for CH_4 and C_2H_6 for each emissions source and year based on the spatial distribution provided by EDGAR (see Chapter 3). The spatial distribution of emissions in the grid maps were adopted from EDGAR, and scaled to match the country level emissions totals estimated above. Example grid maps are shown in Figure 24 illustrating NG CH_4 emissions in 2008. Panels a (this work, high NG FE scenario) and b (EDGAR) show results at the higher 0.1° x 0.1° grid resolution.

Finally, the grid maps were converted to a lower 1° x 1° grid resolution (panels c and d), which is required for use with CT-CH_4 in Chapter 5. While emissions are plotted on a log scale, the higher emissions on the left panels can already be distinguished for some countries and regions, such as the U.S., the Middle East, and Norway. Grid maps for oil and coal CH_4 emissions are shown in Appendix Figure B - 4 and Figure B - 5. Oil grid maps include emissions from tanker transport across the oceans. Coal grid maps appear incomplete due to missing information regarding the spatial emissions distribution and because coal industry emissions are mainly point sources from individual mines, which are difficult to visualize on a grid map (*159*). Grid maps for C_2H_6 are not shown as these differ from CH_4 only in the scale of the legend.

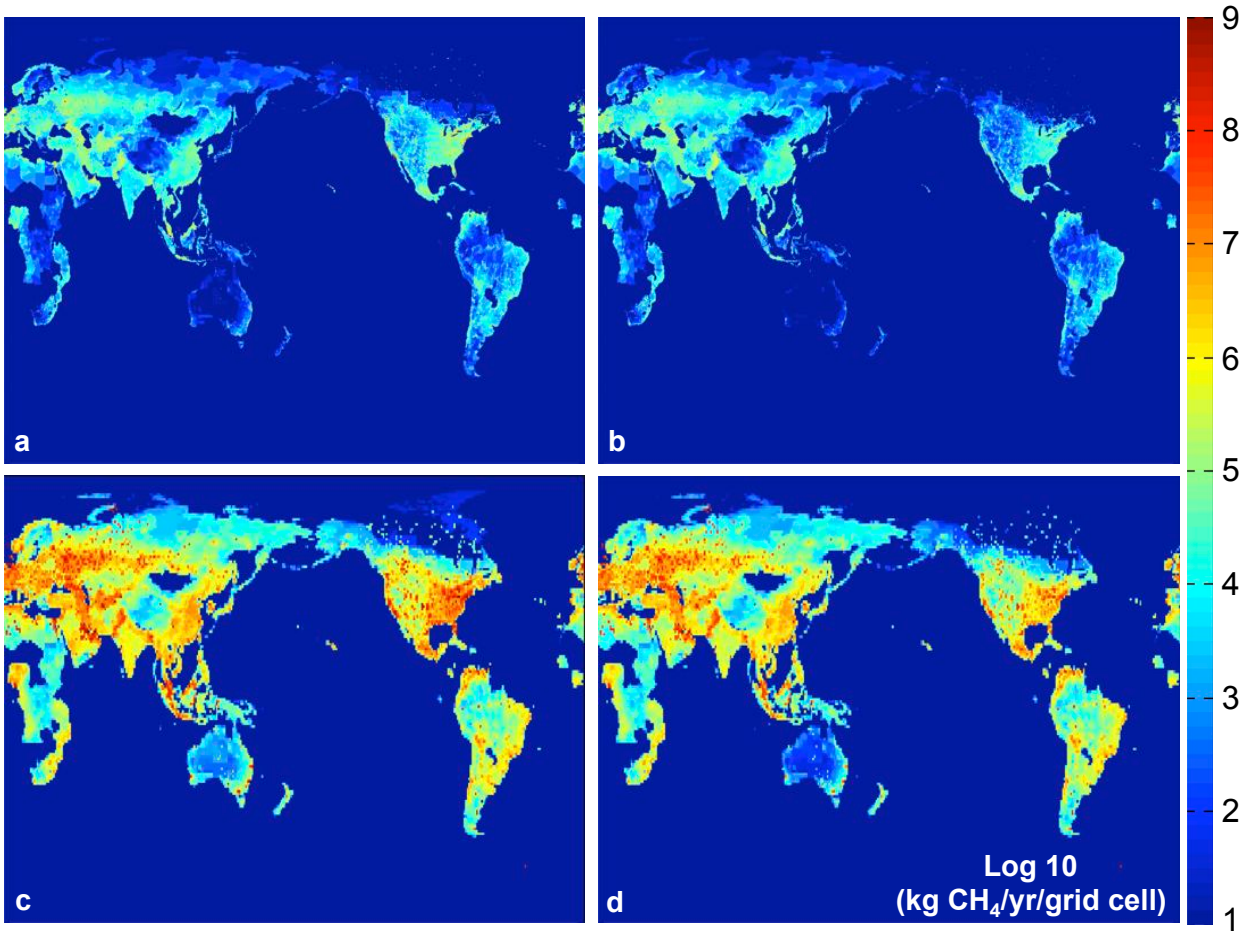

Figure 24: CH4 emissions grid maps in 0.1° x 0.1° resolution (upper panels) for the high NG FE scenario from this work (panels a, c) in comparison with data from EDGAR (*94*) (panels b, d) for the year 2008, and in 1° x 1° resolution (lower panels) for use in CarbonTracker-CH4. Note that the legend units are in kg CH4/yr/grid cell on a logarithmic scale (numbers indicating exponents to base 10). The values in panels c and d are higher due to fewer grid cells.

4.6 Summary

The objective of this chapter was to explain spatially coded global CH_4 and C_2H_6 emissions estimates to be used for atmospheric modeling. The emphasis was on striking a balance between using detailed country-level data for the major emitting countries and allowing transparency in methods and results in order to facilitate interpretation of atmospheric modeling results, particularly NG FE, based on assumptions underlying the emissions *priors*.

Emissions were estimated based on NG FE rate scenarios (NG industry), country-level and other emissions factors (oil and coal industries), hydrocarbon gas composition, and country-level FF production data from the literature. Wet gas composition was quantified based on over 6,000 well samples in order to estimate CH_4 and C_2H_6 emissions from

NG FE as well as oil and coal industries. Dry gas composition was quantified through a mass balance approach using wet gas composition and wet gas, dry gas, and NGL production in the U.S. Table 14 summarizes the *prior* CH_4 emissions results (global totals for each FF industry; low, medium, and high NG FE scenario) in comparison with EDGAR (*85*) estimates. The data indicates significant upward NG corrections compared to EDGAR, but decadal averages of oil and coal in this work are very similar to EDGAR. The relative differences to EDGAR are amplified for individual countries, which influences the spatial distribution of emissions in the grid maps, and potentially also the inversion results.

Table 14: Global *prior* FF emissions estimates compared to EDGAR (*94*). Note that differences for individual countries are significantly larger (in some cases over +/-100%) than the global averages.

	This work (Tg CH_4/yr)	Difference relative to EDGAR (%)		
	Avg. 1980-2008	Avg. 1980-2008	1980	2008
NG (by FE scenario)				
low	45	+19	+55	-9
medium	57	+53	+86	+27
high	83	+122	+141	+113
Oil, best estimate	14	-4	-22	+1
Coal, best estimate	31	-2	+1	-14

Table 15 summarizes the *prior* CH_4 and C_2H_6 emissions results (global totals for each FF industry) for annually constant NG FE rates ranging from 1-9% applied globally as well as oil and coal emissions using a range of C_2H_6 contents in the emitted hydrocarbon gases. While NG is the main source for CH_4 emissions for most FE rate scenarios, the oil industry dominates C_2H_6 emissions except for 9% NG FE, and the coal industry plays only a minor role in total FF C_2H_6 emissions. “Other energy and industry” CH_4 emissions are based on EDGAR (see Table 10).

Table 15: Summary of global *prior* FF emissions estimates for different FE rate and gas composition scenarios as well as “other energy and industry” emissions referenced in Table 10.

	CH_4 (Tg/yr)		**C_2H_6 (Tg/yr)**	
	Avg. 1980-2008	**2011**	**Avg. 1980-2008**	**2011**
***NG** (annually const. NG FE rate applied globally)*[a]				
1%	13	19	1.0	1.6
3%	38	57	3.1	4.7
5%	64	96	5.2	7.8
7%	90	134	7.3	11
9%	115	172	9.4	14

	CH_4 (Tg/yr)		C_2H_6 (Tg/yr)	
	Avg. 1980-2008	**2011**	**Avg. 1980-2008**	**2011**
Oil				
low C_2H_6 content			4.6	5.4
medium C_2H_6 content	14	17	5.9	6.9
high C_2H_6 content			8.3	9.8
Coal				
low C_2H_6 content			0.03	0.05
medium C_2H_6 content	31	46	0.3	0.5
high C_2H_6 content			0.6	0.9
Other energy and industry	18	20	-	-

[a] NG values assume medium C_2H_6 content

Table 16 summarizes the *prior* non-FF emissions based on the literature review in section 4.3. Note that the 3D-modeling values for agriculture / waste based on EDGAR and GFED (see Table 10), which vary annually and seasonally, coincide well with the medium values chosen from the literature. Moreover, the medium box-model *prior* total non-FF emissions are identical with *priors* used in the 3D simulation (385 Tg/yr), which allows direct comparison of box-model results with CT-CH_4 (FF *priors* used in both models are also identical). Note that the gobal soil CH_4 sink *prior* used in both models is slightly higher than in the literature in Figure 17: 25 Tg/yr (*40*), 30 Tg/yr (*44*), and 38 Tg/yr (*41*). (*43*) did not include a soil sink in their inversion study.

Table 16: Summary of global *prior* non-FF emissions estimates used in the 3D-forward modeling and ranges for box-modeling.

	CH_4 (Tg/yr)					C_2H_6 (Tg/yr)		
	Natural	**Agriculture / waste**	**BB**	**Soil sink**	**Total**	**BB**	**BFC**	**Total**
3D-model *(avg. 1980-2011)*	215	194	16	-40	385	n/a	n/a	n/a
Box-model *(const. over time)*								
low	130	130	25	-40	250	1.6	0.6	2.2
medium	182	200	43	-40	385	3.6	2.3	5.9
high	235	270	60	-40	525	5.2	4.0	9.2

BB - Biomass burning; BFC - Biomass fuel combustion

Chapter 5: Quantifying Natural Gas Fugitive Emissions – Global Atmospheric Modeling

5.1 Introduction

Fossil fuel, and other anthropogenic and natural emissions *priors* in Chapter 4 were developed based on detailed EFs and literature estimates. The atmospheric modeling approach in this work contains three main steps. First, NG FE uncertainty will be reduced using the total global atmospheric CH_4 budget as a constraint. This is achieved by estimating *posterior* CH_4 emissions of NG, oil, coal, and other non-FF emissions source categories using a global CH_4 inversion box-model. A box-model treats the global atmosphere as a single parameter rather than a 3-dimensional space as occurs in the inversion studies reviewed in Chapters 3 and 4. The box-model uses an annually weighted average of all available atmospheric measurements worldwide and thereby neglects the spatial differences in measurements and emissions. As a result, *posterior* emissions from different emissions categories are constrained by the global CH_4 budget and the *prior* uncertainties. Simulations with this box-model are computationally efficient and allow more model simulations than an inversion model with 3-dimensional emissions transport. Based on rapidly calculated box-model estimates, *priors* are selected for more refined CT-CH_4 simulations.

Second, NG FE uncertainty will be reduced using C_2H_6 as an alternative trace gas. As discussed and quantified in Chapter 4, C_2H_6 is the second most abundant hydrocarbon component of NG, oil, and coal. In contrast to CH_4, the major global C_2H_6 emissions sources consist of only FFs, biomass burning, and biomass fuel combustion. It is less difficult to constrain individual processes given this small number of sources. Global C_2H_6 measurements and the *prior* C_2H_6 emissions estimates from Chapter 4 will be implemented in the inversion box-model. The resulting *posterior* C_2H_6 emissions will thus provide an additional constraint for NG FE alternative to the global CH_4 budget. Moreover, $\delta^{13}C$-CH_4 measurements will be used in a CH_4 mass balance as an additional constraint, which is not currently available in CT-CH_4. Using $\delta^{13}C$-CH_4 measurements constrains *posterior* emissions (and thus NG FE) due to differences in isotopic signatures between source categories as discussed in 3.2.

Finally, the NG FE scenarios from Table 8, which are in agreement with the most likely emissions scenarios based on global CH_4, C_2H_6, and $\delta^{13}C$-CH_4 constraints, will be used in global 3-dimensional CH_4 forward modeling. This approach is used to simulate atmospheric concentrations across space and time resulting from the above emissions scenarios and the global emissions grid maps described in Chapter 4. The simulated concentrations can then be compared to measurements at each site. While the scenarios may not be equally likely, simulating each scenario individually is useful for analyzing whether and how different NG FE rates can explain potential spatial inconsistencies between simulations and observations. The rationale of this approach is twofold. The 3-dimensional transport model will demonstrate the validity of the box-model results by comparing the global average concentrations of the 3D simulation with

those in the box-model. Furthermore, the 3D-simulations provide the data to calculate the inter-hemispheric gradients resulting from each emissions scenario, which can be compared with observations. Since about 95% of FF emissions are released in the northern hemisphere, the N-S gradient is an important constraint for distinguishing sources (*81*), particularly for analyzing NG FE uncertainty. While forward modeling generates modeled CH_4 concentrations to be compared with observations, it does not result in optimized *posteriors*. However, forward modeling comes at a reduced computational cost, and results can be used as a basis for inversion modeling in the future.

5.2 Global atmospheric box- and 3D-model simulation methods

The 1-box-model treats the atmosphere as a single box. The model conserves the global mass of the emissions, and it accounts for global annually observed trace gas concentrations, which eliminates the need for complex global emissions transport. The box-model was developed based on Bruhwiler *et al*. (*82*) by considering the probability $p(z|s)$ that the simulation (with associated transport errors) of the emissions sources s yields the total emissions burden z derived from the observations (see below). Invoking Bayes' Theorem yields the probability of the sources given the observations, $p(s|z)$. Maximizing $p(s|z)$ is achieved by minimizing the (spatial/temporal) differences between simulation and observations. The solution to this problem yields *posterior* emissions and uncertainties in Equations 23 and 24.

$$s' = s_p + QH^T(R + HQH^T)^{-1}\left(z - Hs_p\right) \quad \textbf{Eq. 23}$$

$$Q' = Q - QH^T(R + HQH^T)^{-1}HQ \quad \textbf{Eq. 24}$$

The *posterior* and *prior* emissions vectors s' and s_p, respectively, have the form

$$s' = \begin{bmatrix} s'_1 \\ s'_2 \\ \vdots \\ s'_j \end{bmatrix} \quad \textbf{Eq. 25,} \qquad s_p = \begin{bmatrix} s_{p_1} \\ s_{p_2} \\ \vdots \\ s_{p_j} \end{bmatrix} \quad \textbf{Eq. 26,}$$

and the *prior* and *posterior* covariance matrices Q and Q', respectively, have the form

$$Q = \begin{bmatrix} Q_{11} & 0 & \dots & 0 \\ 0 & Q_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q_{jj} \end{bmatrix} \quad \textbf{Eq. 27,} \quad Q' = \begin{bmatrix} Q'_{11} & 0 & \dots & 0 \\ 0 & Q'_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q'_{jj} \end{bmatrix} \quad \textbf{Eq. 28,}$$

where j is the number of emissions years, and each element is itself a vector. Each element s'_j and s_{pj} is a vector equal to the length of the number m of emissions sources (e.g., NG). Each element in Q and Q' is a square diagonal matrix of size $m \times m$ with diagonal elements for each emissions year. The off-diagonal elements in Eq. 28 are

zero because each element in Q' only represents total uncertainties of all sources in a given year *j*, and is itself a matrix. The sub-matrices for each year represent *prior* (Q) and *posterior* (Q') emissions uncertainty, i.e., variance for each source m, and covariance between sources. Specifically,

$$\begin{bmatrix} s'_1 \\ s'_2 \\ \vdots \\ s'_m \end{bmatrix}, \begin{bmatrix} s_{p_1} \\ s_{p_2} \\ \vdots \\ s_{p_m} \end{bmatrix}, \begin{bmatrix} Q_{11} & 0 & \dots & 0 \\ 0 & Q_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q_{mm} \end{bmatrix}, \begin{bmatrix} Q'_{11} & Q'_{12} & \dots & Q'_{1m} \\ Q'_{21} & Q'_{22} & \dots & Q'_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ Q'_{m1} & Q'_{m2} & \dots & Q'_{mm} \end{bmatrix}.$$

H is the emissions transport matrix (and H^T its transpose) in the form

$$H = \begin{bmatrix} H_{11} & 0 & \dots & 0 \\ 0 & H_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & H_{jj} \end{bmatrix}, \qquad \textbf{Eq. 29,}$$

where each element in H is itself a vector with length m

$$[H_1 \quad H_2 \quad \dots \quad H_m],$$

and each element $H_m = 1$, thus representing no transport, since the box-model treats the atmosphere as a single box. A more sophisticated model may represent the fact that the non-decayed CH_4 fraction in year j is carried over to year $j+1$. In this case, H, H^T and Q' would be non-diagonal. However, an approximate model solution showed that the annual emissions differences between the models are only in the order of 1-2%, which was deemed insignificant for the purpose of this work.

The model-data mismatch error covariance matrix R is an estimate of the uncertainty of the measured and globally and annually averaged concentrations z. This model assumes a 95% confidence interval for z, i.e., approximately ±1% (or ±20 ppb for CH_4). In a more complex model, i.e., at least two boxes, R would also include uncertainty of atmospheric transport. However, this can be ignored in this 1-box-model. The covariance matrix R represents the observational error in the form

$$R = \begin{bmatrix} R_{11} & 0 & \dots & 0 \\ 0 & R_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R_{jj} \end{bmatrix} \qquad \textbf{Eq. 30,}$$

where each element R_{jj} is the variance of the annual global emissions burden z_j, from the emissions burden vector z in the form

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_j \end{bmatrix} \qquad \textbf{Eq. 31.}$$

The annual emissions burden z_j for CH_4 is estimated using Eq. 33, which is the solution to differential Eq. 32, solved for z. The burden z_j for C_2H_6 is estimated using Eq. 34:

$$\frac{dC_{CH4}}{dt} = z_{CH4} - \frac{1}{tau} * C_{CH4} \quad \textbf{Eq. 32,}$$

$$z_{CH4}(t) = \frac{C_{CH4}(t) - C_{CH4}(t-1) * e^{-\frac{1}{tau}}}{tau * \left(1 - e^{-\frac{1}{tau}}\right)} \quad \textbf{Eq. 33,}$$

$$z_{C2H6}(t) = C_{C2H6}(t) \times SF_{C2H6} \quad \textbf{Eq. 34,}$$

where $C_{CH4}(t)$ is the annually observed global average CH_4 concentration at year t multiplied by 2.767 Tg/ppb (*160*), and tau is the global average lifetime of CH_4. For tau, a range of 9.1-9.7 years was chosen based on the literature review summarized in Appendix B. The scaling factor SF_{C2H6} converts the annually observed global average C_2H_6 concentration $C_{C2H6}(t)$ into the annual emissions burden $z_{C2H6}(t)$, which is based on 3D-modeling in (*161*) and has been applied recently in (*84*). Given uncertainties of up to 45% due to the reaction rate with and concentration of OH (*161*), its best estimate and upper value (corresponding to a higher global budget for estimating upper bound NG FE rates), 0.018 and 0.026 Tg C_2H_6/ppt, respectively, will be used.

Methane measurements from air samples collected at 84 surface sites in the NOAA Cooperative Global Air Sampling Network are available for the years 1983-2012 (*162*). In addition to these weekly samples, in situ quasi-continuous CH_4 time series from towers of external networks are included. The global average annual CH_4 concentrations used in Eq. 33 were calculated by NOAA (*163*), and are based on measurements from a subset of network sites. These sites are typically at remote marine sea level locations, which are representative of well-mixed marine boundary layer air (*163*). A smooth curve is fitted to the weekly measurements, and a relative weighting scheme is used to fit global averages and latitudinal distribution, which favors sites with relatively low noise and consistent sampling (*164*). The uncertainty of NOAA's calculated global average CH_4 concentrations is 1.1 ppb (one std. dev.) averaged across years (*165*).

A map of the global measurement sites is presented in Figure 25, and a more detailed description of the data, sampling methods, and observational network including a list of all sites is available at NOAA's online documentation (*162*). Global annually averaged C_2H_6 measurements were provided by (*165*) based on air-sample measurements from (*166*), and described in more detail elsewhere (*84*). As discussed in Appendix B, the box-model reproduces the globally averaged CT-CH_4 3D-simulation results reasonably well including the inter-annual variations as well as the allocation among sources.

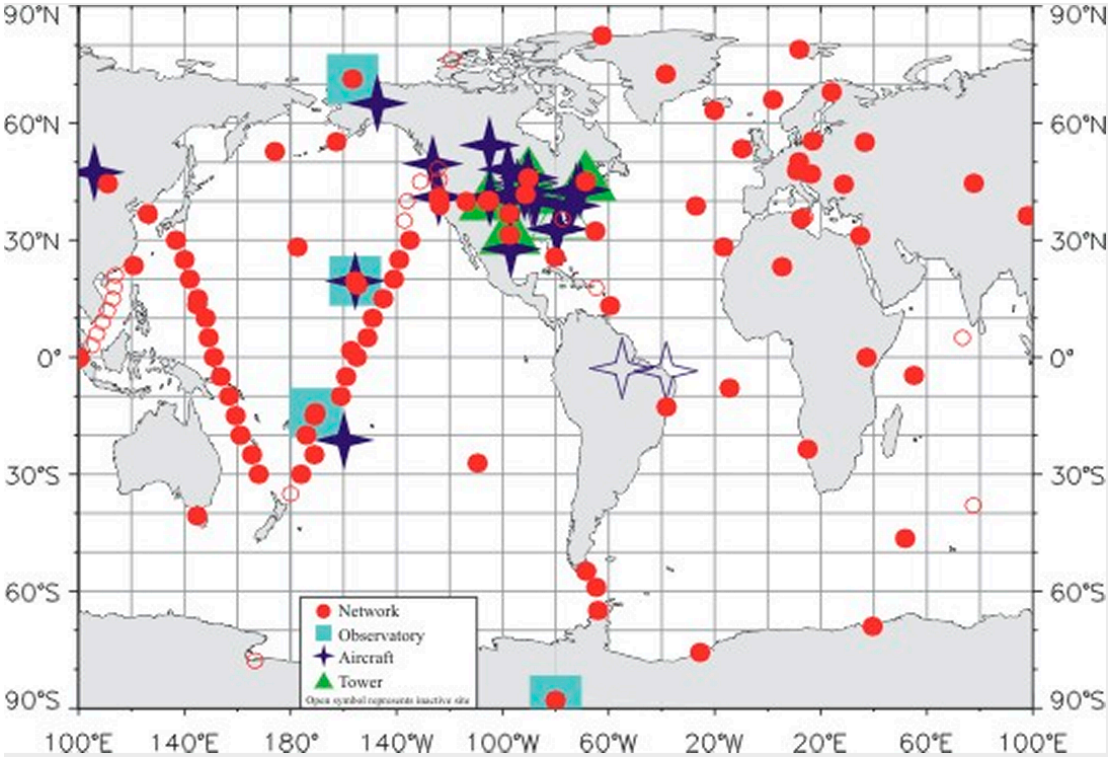

Figure 25: Sampling network of NOAA's Global Monitoring Division providing CH_4 measurements at each site (graph from (*83*)). Aircraft data are used to evaluate 3D inversion results, and are disregarded in this work.

Ethane observations starting in 1984 come from the UC-Irvine Global Monitoring Program as described in (*84*) (via personal communication with a co-author (*165*)), which consists of 45 remote surface locations in the Pacific basin (Alaska to New Zealand). The globally averaged C_2H_6 mixing ratio provided by the authors and used in this work is based on separating the Earth's surface into 16 latitudinal bands, averaging the C_2H_6 mixing ratios within each band, and then averaging the 16 band means. Each global annual C_2H_6 mixing ratio is the average of four consecutive seasonal means (*84*). The global representativeness of the measurements were previously evaluated by the authors through comparison of monthly global and zonal C_2H_6 averages with those from NOAA and the University of Colorado's Institute of Arctic and Alpine Research (INSTAAR) cooperative global air sampling network. The NOAA/INSTAAR network comprises a wider spatial sampling range during the period 2006–2010. The authors determined that the Pacific-based measurements typically represent the globally averaged background concentration at the Earth's surface to within 5-10%, and that the global background air data from the NOAA/INSTAAR network are similar to the background air composition in the Pacific Basin (*84*). Globally and annually averaged CH_4 and C_2H_6 measurements are summarized in Appendix Table B - 9.

In addition to CH_4 and C_2H_6 box-model inversions, a global mass balance of $\delta^{13}C$-CH_4 in atmospheric CH_4 was carried out in order to test NG FE rate scenarios using a third

observational constraint. The ^{13}C:^{12}C ratio of CH_4, δ, in per mil (‰) or part per thousand can be expressed as (*167*):

$$\delta = \left(\frac{R_{Sample}}{R_{Standard}} - 1\right) \times 1000‰ \quad \textbf{Eq. 35,}$$

where $R = (Rare\ isotope \div Abundant\ isotope)$. The global mass balance for three CH_4 sources can be formulated for each year as (*168*):

$$z = z_{Mic} + z_{FF} + z_{BB} \quad \textbf{Eq. 36,}$$

$$\delta_q z = \delta_{Mic} \times z_{Mic} + \delta_{FF} \times z_{FF} + \delta_{BB} \times z_{BB} \quad \textbf{Eq. 37,}$$

where z_{Mic}, z_{FF}, and z_{BB} refer to the microbial, FF, and BB fraction of total annual CH_4 emissions, respectively, as described in Table 10, and z_{Mic} includes all natural and agriculture/waste sources. The different CH_4 emissions sources are aggregated to only three emissions categories in order to avoid an under-constrained system of two linear equations (Eq. 36 and Eq. 37). The equation system is solved for z_{Mic} and z_{FF} as an optimization problem as described in more detail in Appendix B, while z_{BB} is considered known to be at least 25 Tg/yr based on the discussion in section 4.3. The isotopic signatures δ_{Mic}, δ_{FF}, and δ_{BB} are considered known to lie within the range of -59 to -63 ‰, -38 to -42 ‰, and -22 to -26 ‰, respectively, based on weighted averages of each emissions category from 13 literature sources (*168*). The total annual CH_4 emissions burden z is the same as in Eq. 33, and the flux weighted mean isotopic ratio of all CH_4 sources (*167*) is:

$$\delta_q = \alpha\delta_a + \varepsilon - \frac{\varepsilon\left(1 + {}^{\delta_a}/_{1000}\right)}{z} \times \frac{dC_{CH4}}{dt} + \frac{d\delta_a}{dt} \times \frac{C_{CH4}}{z} \quad \textbf{Eq. 38,}$$

where globally and annually averaged measurements δ_a are -47.1 ± 0.05 ‰ over the past decade (*169*), ε = -6.3 ‰, the observed growth rate of δ_a is 0.01‰/yr, $\alpha = \varepsilon \div 1000 + 1$ (*168*), and C_{CH4} is the same as in Eq. 33.

Three-dimensional forward simulations of CH_4 emissions using CT-CH_4 complement the global box-model simulations and the $\delta^{13}C$-CH_4 mass balance. Emissions were simulated for 11 individual source categories, which are described in more detail in Table 10. FF (NG, oil, coal) and natural (wetlands, soils, oceans, termites, and wild animals) CH_4 emissions were modeled as separate sources given the existing CT-CH_4 model code, which allows tracking the individual contributions of total concentrations. The resulting source-specific concentrations are added to background concentrations to yield total concentrations. The input emissions grid maps are as described in section 4.4.

Atmospheric transport in CT-CH_4 is based on Transport Model 5 (TM5) (*170*), which is developed and maintained jointly by the Institute for Marine and Atmospheric Research Utrecht (IMAU, The Netherlands), the Joint Research Centre (JRC, Italy), the Royal Netherlands Meteorological Institute (KNMI, The Netherlands), and NOAA ESRL (USA).

The meteorological fields used for transport in TM5 come from the European Center for Medium range Weather Forecast (ECMWF) operational forecast model. For CT-CH_4, TM5 is run at a global 6°x 4° resolution, 60-90 layers in the vertical depending on the year, and at time steps of less than 3 hours. Forward simulations in this work cover the period 1989-2011 for which meteorological fields were already integrated in CT-CH_4, and measurements are the same as used for the box-model. Zonal integrals of CH_4 concentrations were estimated analogously to global averages described above. The following five different zones were distinguished in order to analyze the spatial differences ignored in the box-model: polar Northern hemisphere (PNH, 53.1°N-90°N), temperate Northern hemisphere (TNH, 17.5°N-53.1°N), the tropics (17.5°S-17.5°N), temperate Southern hemisphere (TSH, 17.5°S-53.1°S), and polar Southern hemisphere (PSH, 53.1°S-90°S).

5.3 Atmospheric CH_4 and C_2H_6 modeling results

5.3.1 Global box-model simulation results

The results for CH_4 box modeling are summarized in Figure 26, which shows *prior* and *posterior* emissions (left and right bars, respectively) for a range of NG FE rates (upper, middle, lower panels) and different CH_4 lifetimes (left and right panels). Inversion models typically assign *prior* uncertainties to all sources, and report *posterior* means and uncertainties for all sources (see Figure 17). A different approach is used in this work, which is equivalent to a mass balance with the objective to identify best estimate and upper bound NG FE rates. The mass balance is carried out in order to estimate NG FE using only the global total *posterior* emissions burden as a constraint in combination with *prior* emissions scenarios. Note that this mass balance could have been calculated without developing an *inversion* box-model. However, the inversion 1-box-model described here will provide the basis for future model development, such as a 2-box-model and/or linked CH_4 and C_2H_6 inversions.

A visual illustration of the box-model mass balance is provided in Appendix Figure B - 11. For each year, the inversion model forces FF CH_4 *posterior*s to equal *priors* by assigning nearly zero *prior* uncertainties. By assigning fixed *posterior* FF emissions, i.e., NG FE rate scenarios and lower bounds for oil and coal, only the non-FF *posterior*s will be adjusted to match the annual CH_4 budget based on measurements. As mentioned in more detail below and in Appendix Figure B - 11, an upper bound of NG FE rates can thus be quantified by identifying NG FE scenarios where total *priors* exceed *posterior* emissions. Figure 26 shows that total *prior* emissions exceed *posterior*s increasingly at higher NG FE, which is more pronounced in the low CH_4 budget scenario. Moreover, the rate at which *prior* emissions increase due to surging FF production over time exceeds *posterior*s, i.e., measurements, particularly throughout the past decade, which will be discussed in more detail below.

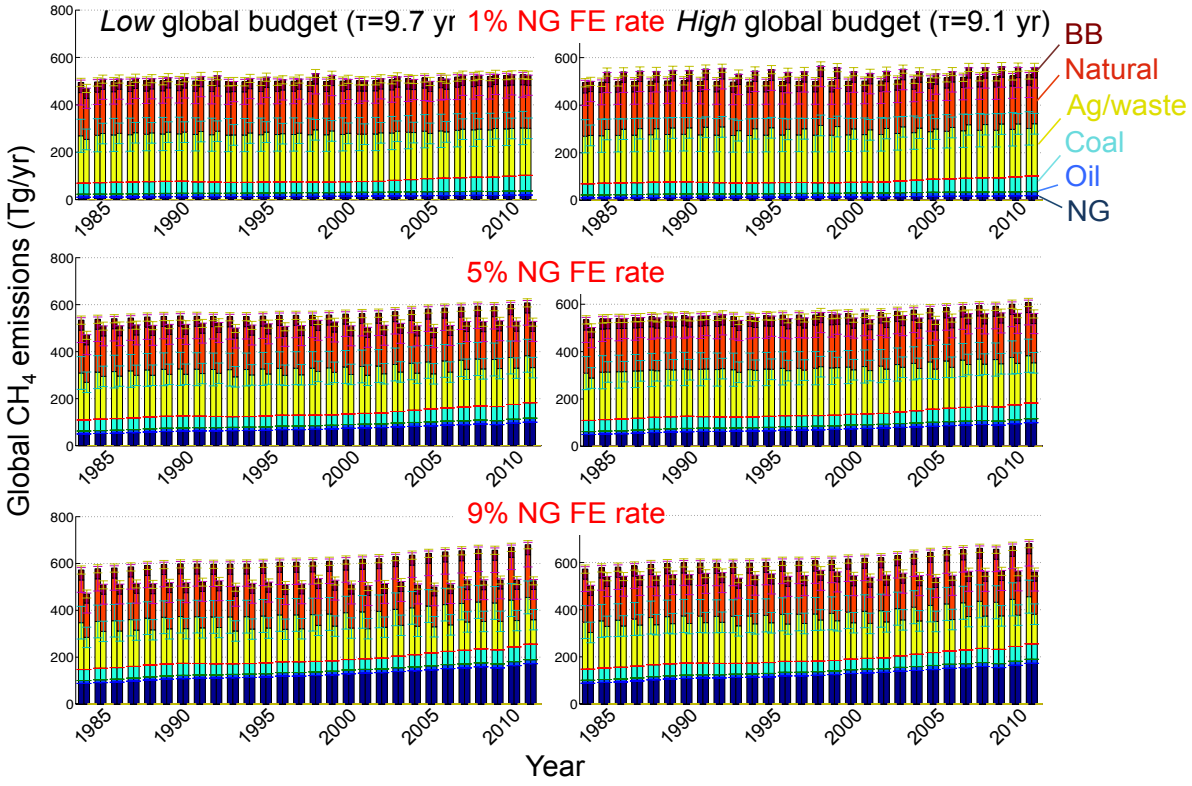

Figure 26: Box-model CH_4 inversion results using global, annually constant NG FE rates for a range of CH_4 lifetimes. The left and right bars in each year are *prior* and *posterior* emissions, respectively.

The implications of the box-model CH_4 inversion are summarized in Figure 27, which shows the differences between *prior* and *posterior* emissions relative to FF CH_4 for years ranging from 1985 to 2010. Analogously to Figure 26, results are plotted for different NG FE rates and two global CH_4 budget scenarios. Positive values indicate that NG FE rates are incompatible with measurements and *priors*. In this case, NG related emissions are too high to match the total CH_4 budget while maintaining the *prior* estimates of all other emissions sources. The upper panel shows *prior mean* minus *posterior* mean. In this metric, positive values indicate that NG FE are too high assuming that all other emissions sources are in reality at least as high as the *prior* best estimate (*prior mean*). If this assumption holds and CH_4 lifetime is at least 9.1 years, then the NG FE rate cannot be higher than about 5% since 1985, and must be 4% or less since 2000. The vertical axis shows that, for example, in the 9% NG FE and low global budget scenario (upper left panel), the total FF emissions would have to be over 50% lower in order to match the global CH_4 budget. While *prior* oil and coal emissions are likely a low estimate (see section 4.4), *prior* non-FF emissions may be significantly lower than the best estimate as discussed in section 4.3. For example, assuming 80% associated gas-flaring efficiency (instead of 95% as described in section 4.4.2) increases *prior* oil emissions by 5.5 Tg CH_4/yr (mean value 1984-2011). Given the *posterior* FF budget estimated in the inversion box-model, the lower flaring efficiency would reduce NG FE by 0.3 and 0.6 percentage points in 1984 and 2011, respectively.

The lower panel in Figure 27 shows which NG FE rates are compatible with measurements if all non-FF sources (biomass burning, natural, and agriculture/waste) were in reality at their lower bound of *prior* estimates. In this case, NG FE could be 9% or higher, and only slightly less in 2010 if CH_4 lifetime was relatively high (lower left panel). Moreover, global average CH_4 observations would not be sufficient to reduce NG FE uncertainty in the LCA literature. It should be emphasized, however, that this assumes all non-FF sources to be at their lowest *prior* values *simultaneously*.

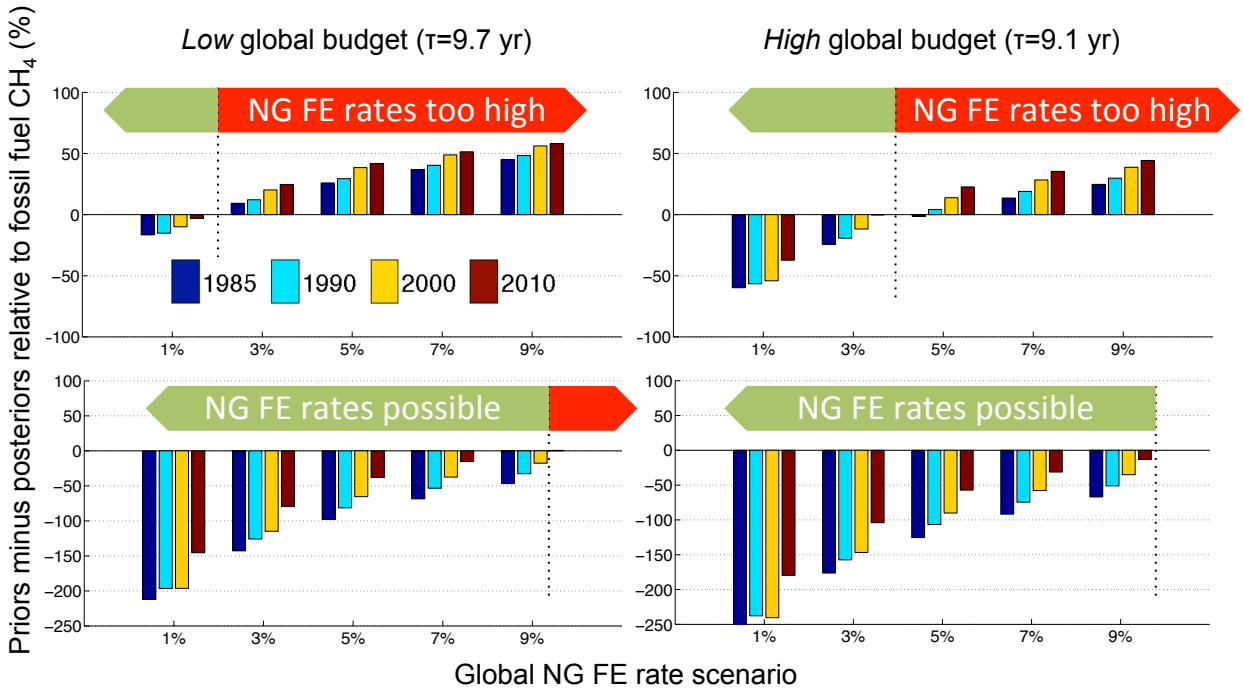

Figure 27: *Prior* minus *posterior* CH_4 emissions for different years and CH_4 lifetimes. Positive values indicate incompatibility of NG FE rates with measurements and *priors*. Upper panel: *prior* mean – *posterior* mean (best estimate NG FE). Lower panel: *prior* lower bound – *posterior* mean (upper bound NG FE).

Global average CH_4 observations are only sufficient to reduce NG FE uncertainty under the scenarios described above because a large fraction of the total CH_4 budget is from non-FF sources, which are also highly uncertain. Box-model inversions using C_2H_6 measurements were performed to provide additional observational evidence. While the C_2H_6 content of FF use is more uncertain than CH_4 (see section 4.4), non-FF sources constitute a smaller fraction of the global C_2H_6 budget compared to CH_4. This renders C_2H_6 a favorable tracer for FF CH_4 and C_2H_6 emissions including the NG FE rate. Figure 28 and Figure 29 show C_2H_6 inversion results analogously to Figure 26 and Figure 27 for the medium FF C_2H_6 content. The low and high global C_2H_6 budget scenarios are derived using the methods described in section 5.2. Missing years indicate that measurements were not available. NG FE rates refer to % of NG production (same as in the CH_4 inversion), i.e., interpretation of results is exactly the same as in the CH_4 inversion. Note that *posterior* BB and BFC emissions in Figure 28 are negative in some years. This indicates that FF C_2H_6 emissions alone are higher than the global C_2H_6

budget, and the model forces the non-FF sources to be negative in order to maintain the global budget as a whole. Hence, negative values are not to be understood as an emissions sink. Negative values signal that a NG FE rate is inconsistent with measurements for a given scenario in this model. Similarly to the CH_4 inversion, Figure 28 emphasizes that the rate at which *prior* emissions increase due to surging FF production over time exceeds *posterior*s, which may signify a declining NG FE rate over time, and/or significant emissions from natural seepage and/or leaking abandoned gas wells.

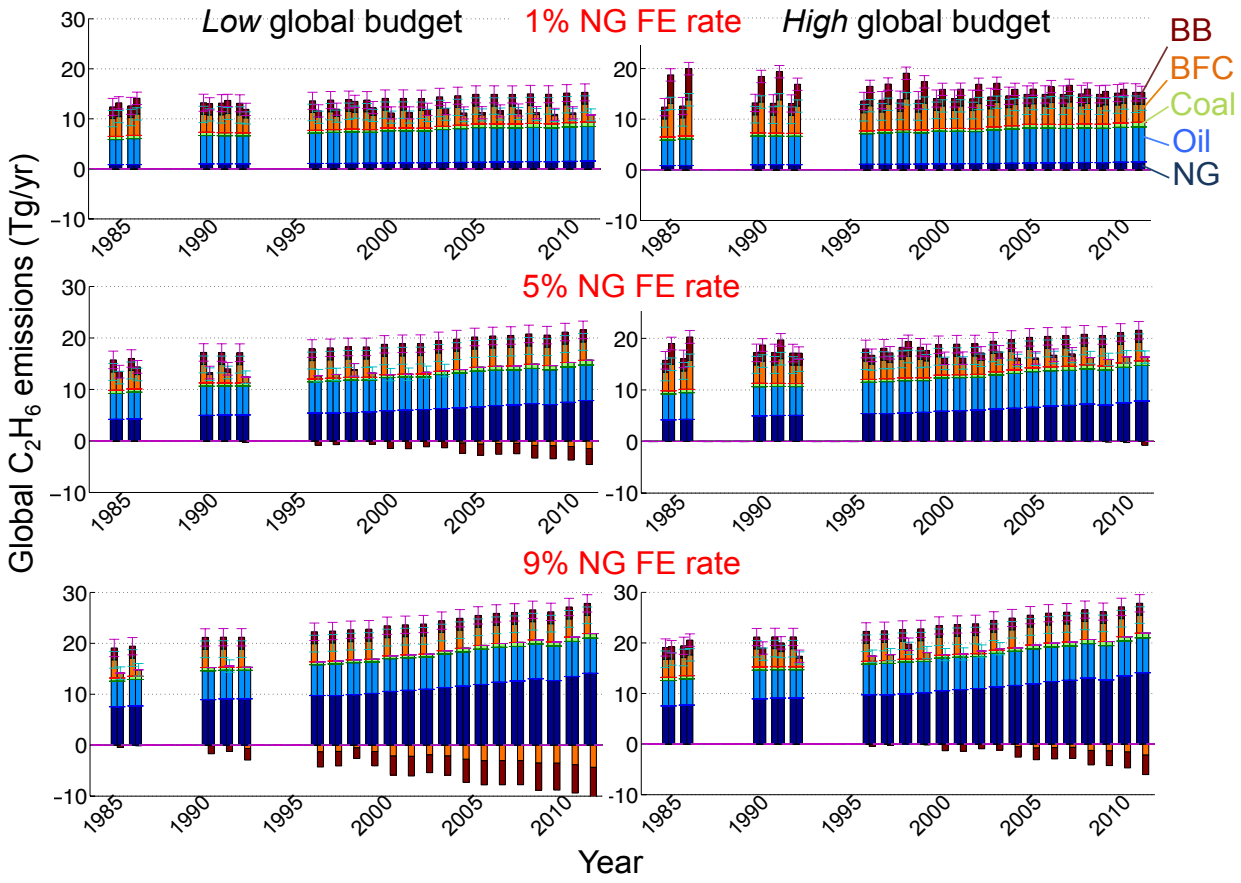

Figure 28: Box-model C_2H_6 inversion results using global, annually constant NG FE rates and medium FF C_2H_6 content. The low and high budget scenarios refer to C_2H_6 measurements scaled (Tg/ppt) using (*161*) best estimate and uncertainties, respectively. Negative *posterior* emissions indicate FF *priors* are too high to accommodate non-fossil emissions in the budget.

Figure 29 shows that current NG FE rates do not exceed 4% in this calculation, even if the global budget is high and BB and BFC emissions are at their lowest *prior* levels (lower right panel). Only under the assumption of the lowest threshold C_2H_6 content of all FFs (see section 4.4) could current NG FE be as high as 5% (see Appendix Figure B - 12 and Figure B - 13) and as high as 7% in 2000. FE rate uncertainty before 2000 (dark and light blue bars) is more difficult to constrain with confidence. If the global C_2H_6

budget is high and/or BB and BFC emissions are low, NG FE could be 9% or higher (lower right panel in Figure 29 and right panel in Appendix Figure B - 13). However, note that the global C_2H_6 burden modeled for estimating upper bound NG FE are higher than any literature estimates. As shown in Figure 28, the global C_2H_6 burden between 1985 and 1992 ranges from 17-21 Tg C_2H_6/yr compared to 10-18 Tg C_2H_6/yr over the same period from 11 studies reviewed by Xiao *et al*. (*171*). This suggests that – all else equal – it is very unlikely that NG FE rates are as high as 5% and 7% in 2010 and 2000, respectively.

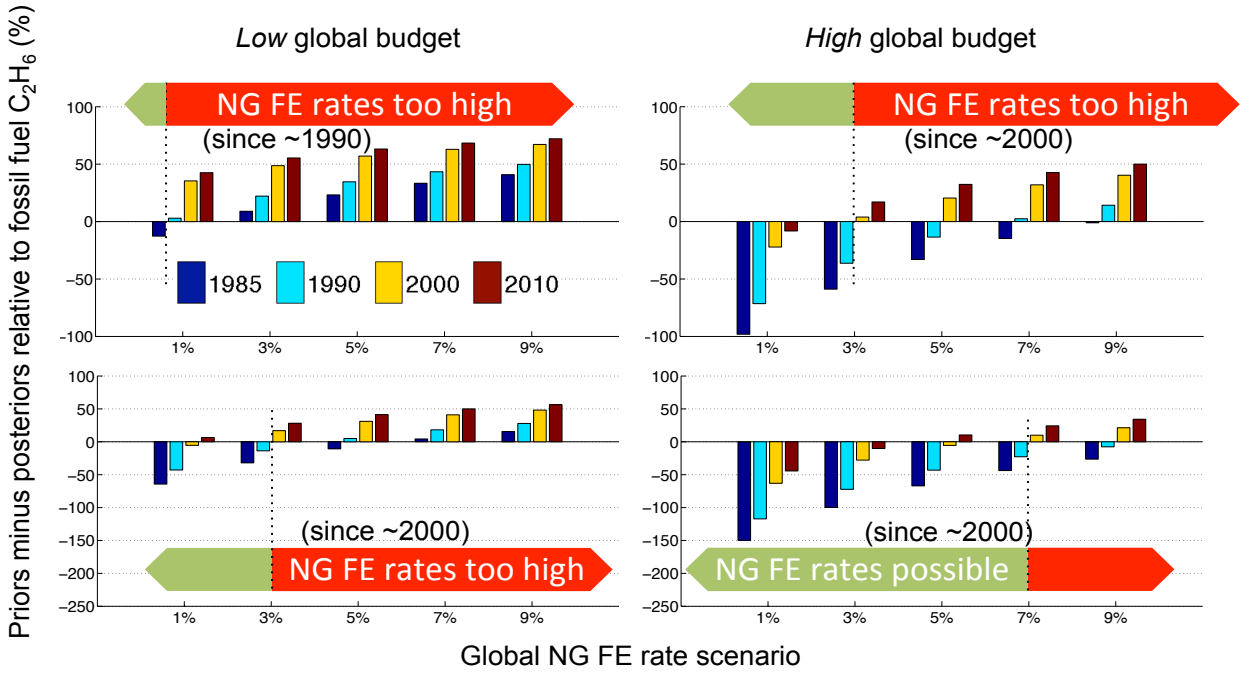

Figure 29: *Prior* minus *posterior* C_2H_6 emissions for different years and global C_2H_6 budgets, and medium FF C_2H_6 content. Positive values indicate incompatibility of NG FE rates with measurements and *priors*. Upper panel: *prior* mean – *posterior* mean (best estimate NG FE). Lower panel: *prior* lower bound – *posterior* mean (upper bound NG FE).

As summarized in Table 17, current NG FE rates under 9% cannot be ruled out using CH_4 measurements alone, but best estimates here are 3-4% over the past decade, which is consistent using both CH_4 and C_2H_6 measurements. NG FE are much more uncertain before 2000. Best estimate NG FE estimates in 1990 are 2-5% based on CH_4 data, and up to 8% using C_2H_6 data. It should be emphasized that the assumptions going into the box-model and FF *prior* emissions were intentionally skewed towards upper bound scenarios in order to analyze which NG FE rates can be ruled out with confidence. The upper bound scenarios assume that (i) all non-FF sources are at their minimum simultaneously, (ii) relatively small amounts of associated gas venting is considered in oil EFs, and (iii) no natural seepage is included in the *priors*. NG FE rates over 9% are possible in the upper bound scenario using CH_4 measurements, but C_2H_6 measurements indicate that current NG FE rates over 5% are inconsistent with the observations given any scenario and the described assumptions therein.

Table 17: Summary of possible NG FE rates (in % of dry NG production) under different scenarios using the box-model inversion.

Tracer	CH_4		C_2H_6			
FF C_2H_6 content	-		Low		Medium	
Global burden	Low	High	Low	High	Low	High
Best NG FE estimate (mean non-FF priors)						
2010	2	3	1	3	1	2
2000	2	4	1	4	1	3
1990	2	5	2	8	1	7
Upper bound NG FE (low non-FF priors)						
2010	9	>9	2	5	1	4
2000	>9	>9	3	7	2	6
1990	>9	>9	6	>9	5	>9

The results of the $\delta^{13}C$-CH_4 mass balance are summarized in Table 18, which shows emissions estimates for each category and corresponding isotopic signatures. The best NG FE estimates are based on literature mean value isotopic signatures described in section 5.2, a mean biomass burning source of 43 Tg CH_4/yr from the literature (see Figure 17), and a medium CH_4 lifetime of 9.4 years, which yields NG FE rates between 3-5% given the FF CH_4 bottom-up inventory in section 4.5. This is in agreement with the best estimate from CH_4 and C_2H_6 measurements in Table 17. The high NG FE estimate assumes a short CH_4 lifetime of 9.1 years, a lower biomass burning threshold of 25 Tg CH_4/yr, and an isotopic signature uncertainty range of ± 2‰. For comparison, the equivalent literature BB estimate in Table 6 is 32 Tg CH_4/yr including a conservative (low) BFC value.

Table 18: CH_4 emissions source attribution and associated NG FE rate based on global isotope mass balance, averaged over the period 1990-2010.

	Microbial	BB	FF	NG FE rate
Units (unless otherwise noted)	**Tg CH_4/yr**			**%**
Best NG FE estimate				
Emissions, 2010	367	43	133	3
Emissions, 2000	355	43	126	5
Emissions, 1990	357	43	128	5
$\delta^{13}C$-CH_4 source signature (‰)	-61	-24	-40	-
Upper bound NG FE				
Emissions, 2010	310	25	225	8
Emissions, 2000	300	25	216	>9
Emissions, 1990	303	25	218	>9
$\delta^{13}C$-CH_4 source signature (‰)	-63	-26	-42	-

Also note that the isotopic signatures in the high NG FE scenario are skewed to significantly lighter (more negative) values compared to those used in the recent literature (*172*, *173*), which yields higher NG FE rates. Maximizing the FF source in the mass balance optimization problem yields a 7% NG FE rate in 2010, but 9% or higher is possible in earlier decades. In this case, all isotopic signatures must be at the low end of the uncertainty range. While the best estimate from isotopes agrees with the best estimate from CH_4 and C_2H_6, isotopes constrain current NG FE to 8% or lower, which cannot be achieved using CH_4 measurements alone.

As reviewed in section 4.3, emissions from natural seepage may range between 40-60 Tg CH_4/yr and 2-4 Tg C_2H_6/yr, the $\delta^{13}C$-CH_4 fraction being indistinguishable from FFs. Additional box-model simulation results using the lower bound natural seepage values reduces NG FE estimates (both best estimates and upper bound) by about two percentage points (see Appendix Table B - 10). In this case, best estimates range from less than 1% to less than 3% over the past decade. The best estimate for 2010 is 1% or less, i.e., smaller than the current EPA bottom-up estimate, which appears unlikely. Hence, in order to defend the magnitude of the modeled natural seepage source, best estimate natural and/or agriculture/waste, and biomass burning *priors* are too high. Natural seepage reduces the 2010 upper bound NG FE estimate to 6% and 3% using $\delta^{13}C$-CH_4 and C_2H_6 constraints, respectively.

As shown above, CH_4 *priors* increase at a greater rate than CH_4 measurements, and C_2H_6 *priors* increase while C_2H_6 measurements decrease over time. Possible explanations are: (i) NG FE rates have decreased significantly since the 1980s. (ii) NG FE rates are at the lower end of box-model results, and there are significant emissions from natural seepage and/or leaking abandoned gas wells. Since seepage and leaking abandoned well emissions are relatively independent of NG and oil production, total *prior* emissions are expected to increase at a lower rate. Indeed, a 1% global average NG FE rate in combination with a natural seepage source would eliminate the inconsistency between *prior* CH_4 and C_2H_6 growth rates and measurements (see Appendix Figure B - 14 and Figure B - 16). However, such low FE rate is unlikely as discussed above. Instead, these results support the evidence in Tables Table 17 and Table 18 that NG FE rates have decreased over time. This conclusion is consistent with several other studies. Simpson *et al*. and Aydin *et al*. (*84*, *141*) concluded that the reduction in global C_2H_6 concentrations is due to decreased flaring and venting of NG. Monteil *et al*. (*172*) suggest that reductions in FF emissions are a likely explanation for a decline in both CH_4 concentrations and $\delta^{13}C$-CH_4 since 1980. Recent direct CH_4 measurements at 190 NG production sites in the U.S. by Allen *et al*. (*71*) indicate lower overall CH_4 emissions from production (well pad) activities than previous measurement data used in EPA's 2013 GHG inventory (*64*).

5.3.2 **Global CT-CH_4 model forward simulation results**

NOAA's CT-CH_4 system - driven by the TM5 global transport model - was used to simulate atmospheric concentrations of each *prior* scenario for the years 1989-2011. Observational CH_4 data is available from 84 NOAA sites distributed globally for this

period. Global averages of the forward simulation output from CT-CH_4 are consistent with box-model results as shown in Figure 30, which displays differences between the high, medium and low NG FE scenarios (see Table 8 and Figure 18) and measurements. The medium FE scenario is a reasonable fit throughout the 1990s (~ 4% FE) compared to 2% and 4-5% in the box-model for CH_4 lifetimes τ of 9.7 and 9.1, respectively. In the 2000s, CT-CH_4 suggests a best estimate NG FE rate of just over 4% dropping to less than 3% in 2010 compared 2-4% (2000) and 2-3% (2010) in the box-model depending on τ. Given that τ used in CT-CH_4 is approximately 9.45, best estimates of both models agree within one percentage point NG FE. Forward results and measurements are presented separately in Appendix Figure B - 18, which shows similar seasonal patterns in simulations and measurements.

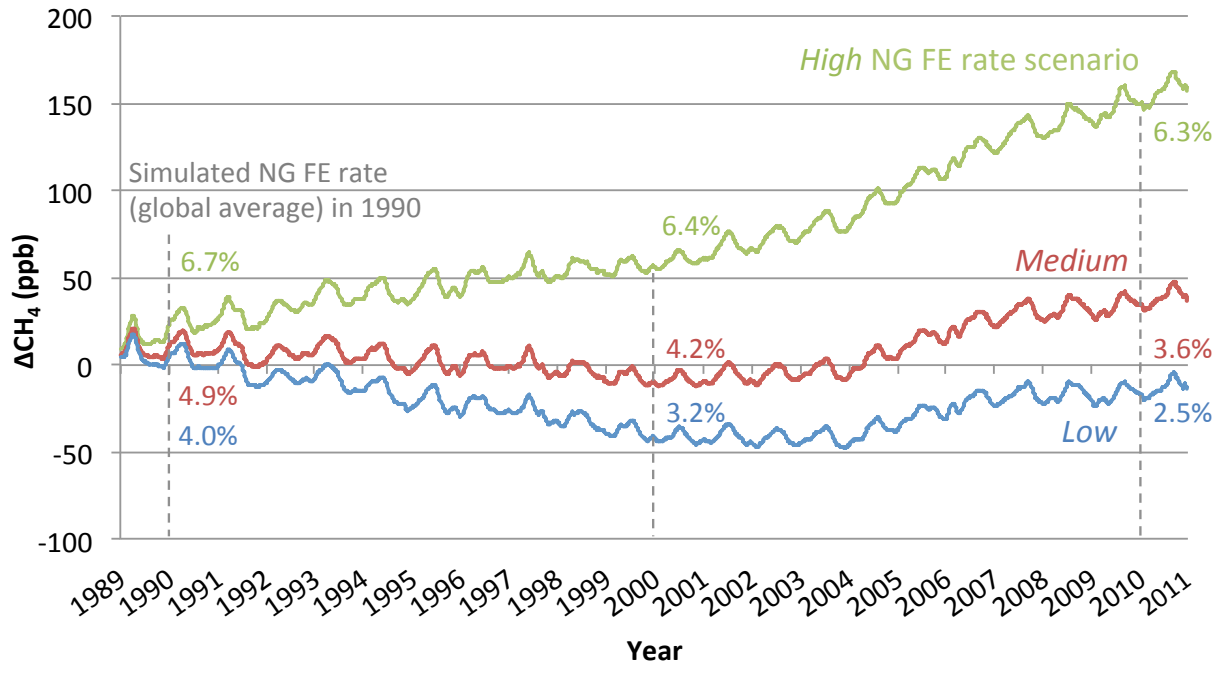

Figure 30: CarbonTracker-CH_4 global average forward modeling results (simulation minus measurement) for three NG FE scenarios with decreasing FE rates over time.

Simulations and measurements across 41 latitudinal bands (intervals of 0.05 sine of latitude) are plotted in Figure 31 as an indicator of the inter-hemispheric gradient for the year 2000. Other years are shown in Appendix Figure B - 19 revealing a qualitatively similar pattern. The gradient provides information about the spatial distribution of CH_4 emissions, which cannot be obtained from the box-model. The spatial fit of simulations and measurements can be used as a proxy for the attribution of sources. About 96% of NG CH_4 emissions in the *prior* grid maps simulated with CT-CH_4 are released in the Northern hemisphere. The equivalent values for oil, coal, agriculture/waste, and natural sources are 91%, 88%, 82%, and 54%, respectively. Figure 31 shows that CH_4 concentrations at the North Pole are about 130 ppb higher compared to most Southern latitudes. Simulations exhibit a similar, but more pronounced North-South gradient, i.e., 150, 165, and 180 ppb for low, medium, and high NG FE scenarios, respectively. While

Figure 30 shows that the medium FE scenario best matches global total CH_4 abundance, the inter-hemispheric gradient indicates that total medium *prior* emissions are too high in the North and too low in the South. Since (i) reducing NG FE alone is not sufficient to match the observed inter-hemispheric gradient, and (ii) coal and oil CH_4 *priors* are considered a low estimate, misallocation of non-FF CH_4 *priors* across latitudes must at least partially explain the mismatch between simulations and measurements. This is consistent with previous atmospheric inversions, which tend to reduce high latitudinal sources compensated by increases at lower latitudes, and tropical wetlands may be underestimated in particular (*44*, *92*).

The magnitude of the mismatch is in the order of one percentage point NG FE per hemisphere, i.e., the difference between the low and medium FE scenario, which is equivalent to 10-20 Tg CH_4/yr. As a hypothetical first order estimate, reducing 10-20 Tg CH_4/yr *prior* emissions from a given Northern latitudinal band, and increasing emissions in the equivalent Southern band by the same amount may eliminate the mismatch. As described above, particularly natural emissions *priors* have a strong Southern hemispheric contribution, i.e., ~95 and ~80 Tg CH_4/yr in the North and South, respectively. Re-allocating the above mismatch in the natural source alone would therefore reverse the North-South contributions. Instead of re-allocating 10-20 Tg CH_4/yr within the same latitudinal bands in the North and South, the mismatch can also be eliminated by re-allocating emissions larger or smaller than 10-20 Tg CH_4/yr depending on the chosen locations. This could be established by conducting a full inversion.

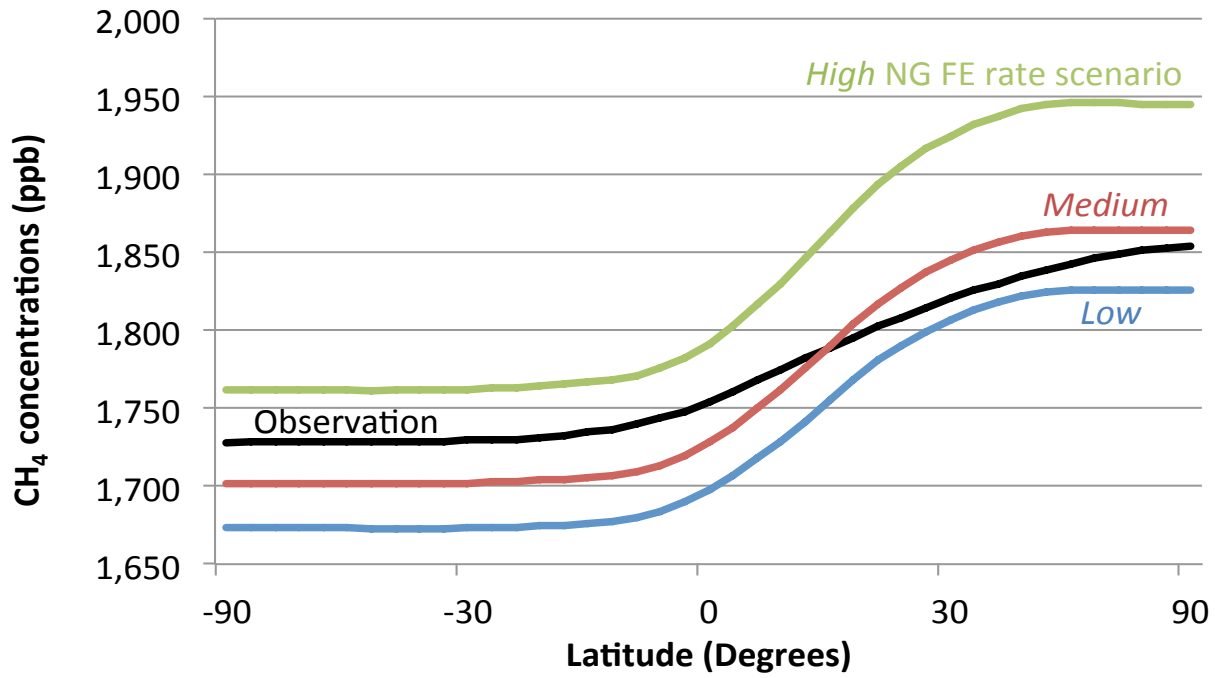

Figure 31: CT-CH_4 simulated CH_4 concentrations across 41 latitudinal bands for the year 2000 compared with NOAA's measurements (*174*). Observational uncertainties for the individual latitudinal bands were not available at the time of dissertation submission. The uncertainties are likely between those for global average (±1.1 ppb) and individual sites (up to ±20 ppb) with highest values where data are sparse, i.e., the tropics.

Seasonally averaged forward simulation output from CT-CH_4 is shown in Figure 32, which displays differences between the high, medium and low NG FE scenarios and measurements for five different latitudinal bands. As shown in Figure 31, higher NG FE rates provide a better fit in Southern latitudes compared to Northern latitudes. Additionally, the seasonal data indicate that NG FE scenarios in the polar and temperate Northern hemispheric zones do not provide the same fit (simulation minus measurement) in summer and winter, as opposed to nearly identical fit in the Southern hemisphere.

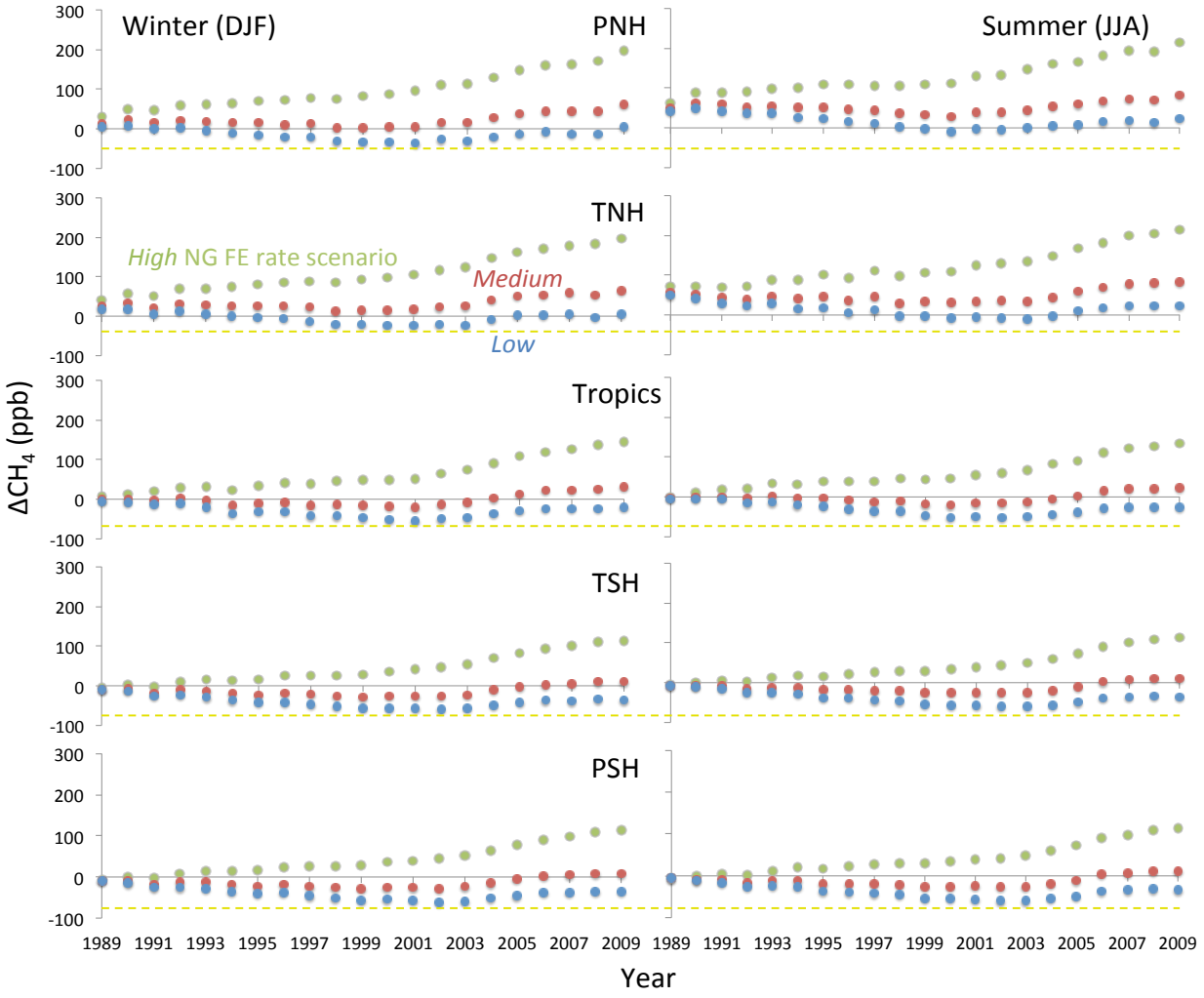

Figure 32: CT-CH_4 seasonal forward modeling residuals (simulation minus measurement) for five latitudinal zones, i.e., polar northern hemisphere (PNH), temperate northern hemisphere (TNH), tropics, temperate southern hemisphere (TSH), and polar southern hemisphere (PSH), and high (green dotted line), medium (red), and low (blue) NG FE scenarios. Winter and summer plots show averages over December, January, February, and June, July, August, respectively. The dashed yellow line is a reference for comparing simulation-measurement fit in winter and summer.

The best-fit ($\Delta CH_4 = 0$) summer-winter difference is ~25-30 ppb in PNH and decreases to nearly zero in PSH. This indicates that given any FE rate scenario, *prior* Northern

emissions are too high in the summer and/or too low in the winter. If Northern seasonal emissions sources with summer peaks were smaller, Northern summer concentrations would fall by a given amount, and Northern winter concentrations would fall to a much lower extent. As a result, the best-fit NG FE scenario would provide a better agreement across seasons.

Agriculture/waste and wetlands are the two major source categories of similar magnitude (~190 and ~170 Tg CH_4/yr, respectively) with a distinct seasonal pattern, which may explain this phenomenon. Particularly, modeled *prior* Northern wetland CH_4 emissions peak around July/August and are nearly zero from November to March (see also (*140*)). The seasonal differences (peak minus trough) of wetlands (~50 ppb) are roughly twice as strong as for agriculture/waste (~25 ppb) at Barrow, and smaller at lower latitudes. This is shown in Figure 33, where agriculture/waste and wetland contributions to CH_4 concentrations from CT-CH_4 forward simulations are plotted for three sites across a wide latitudinal spectrum. At Barrow, agriculture/waste concentrations are countercyclical to wetlands, with the latter peaking around July/August. This behavior may be explained by a combination of atmospheric transport of agriculture/waste emissions from lower latitudes, a reduced winter OH sink at higher latitudes (as agriculture/waste emissions are transported North) as well as local high latitude wetland sources during the summer. The countercyclical behavior was also observed at other remote Northern sites across the longitudinal spectrum, e.g., Summit, Greenland, and Pallas-Sammaltunturi, Finland, with wetland peaks around July (see Appendix Figure B - 20). Moving towards lower latitudes, the cyclic differences between agriculture/waste and wetland sources become less pronounced, e.g., three months at Mauna Loa and one month at the South Pole.

The behavior in Figure 33 is consistent with at least two possible explanations for the seasonal mismatch in Figure 32. First, *prior* Northern wetland emissions may be too high. Reducing this source would lower Northern summer concentrations to a greater extent than in winter, and thus provide a better seasonal fit. The magnitude of seasonal differences of wetlands (~ 50 ppb at Barrow) appears sufficient to explain Northern summer-winter discrepancy (~25-30 ppb). In this case, *prior* Southern emissions, e.g., wetlands, need to be increased simultaneously in order to maintain the global budget. This also means that increased NG emissions (predominantly Northern) cannot balance reduced Northern wetland emissions in the global budget. It should be noted that some studies suggest increased CH_4 emissions from Siberian wetlands due to thawing permafrost as a feedback to climate change. For instance, based on lake measurement extrapolated to North Siberian lakes, Walter *et al*. (*175*) conclude that previous Northern wetland emissions were underestimated by up to 64%. Shakova *et al*. (*176*) find about 10 Tg CH_4/yr from sub-sea permafrost in the East Siberian Arctic Shelf, which have not been accounted for in global CH_4 inventories. Results from this work as well as NOAA's previous inversion (*177*) do not find a missing CH_4 source in Northern Polar Hemisphere based on historical measurements. In any case, if there is a significant unaccounted permafrost CH_4 source, the above NG FE estimates would be overestimated given the global CH_4 burden as a constraint. Nevertheless, increases in global temperatures have the potential to contribute large amounts of CH_4 emissions from permafrost and other

Northern soils in the future. According to some estimates, these sources hold twice the amount of carbon that is present in the atmosphere today (*178*).

Second, agriculture/waste *priors* may be too low given their countercyclical pattern relative to wetlands. Increasing agriculture/waste priors would increase Northern winter concentrations more than Northern summer concentrations, thereby reducing the seasonal summer-winter mismatch in the North. In order to compensate for the resulting increase in global concentrations as well as the inter-hemispheric gradient mismatch (Figure 31), reductions in Northern wetland and/or NG emissions, and increased Southern emissions are needed. In either case, NG *priors* would have to be reduced from the best fit (3-4% over the last decade), if any.

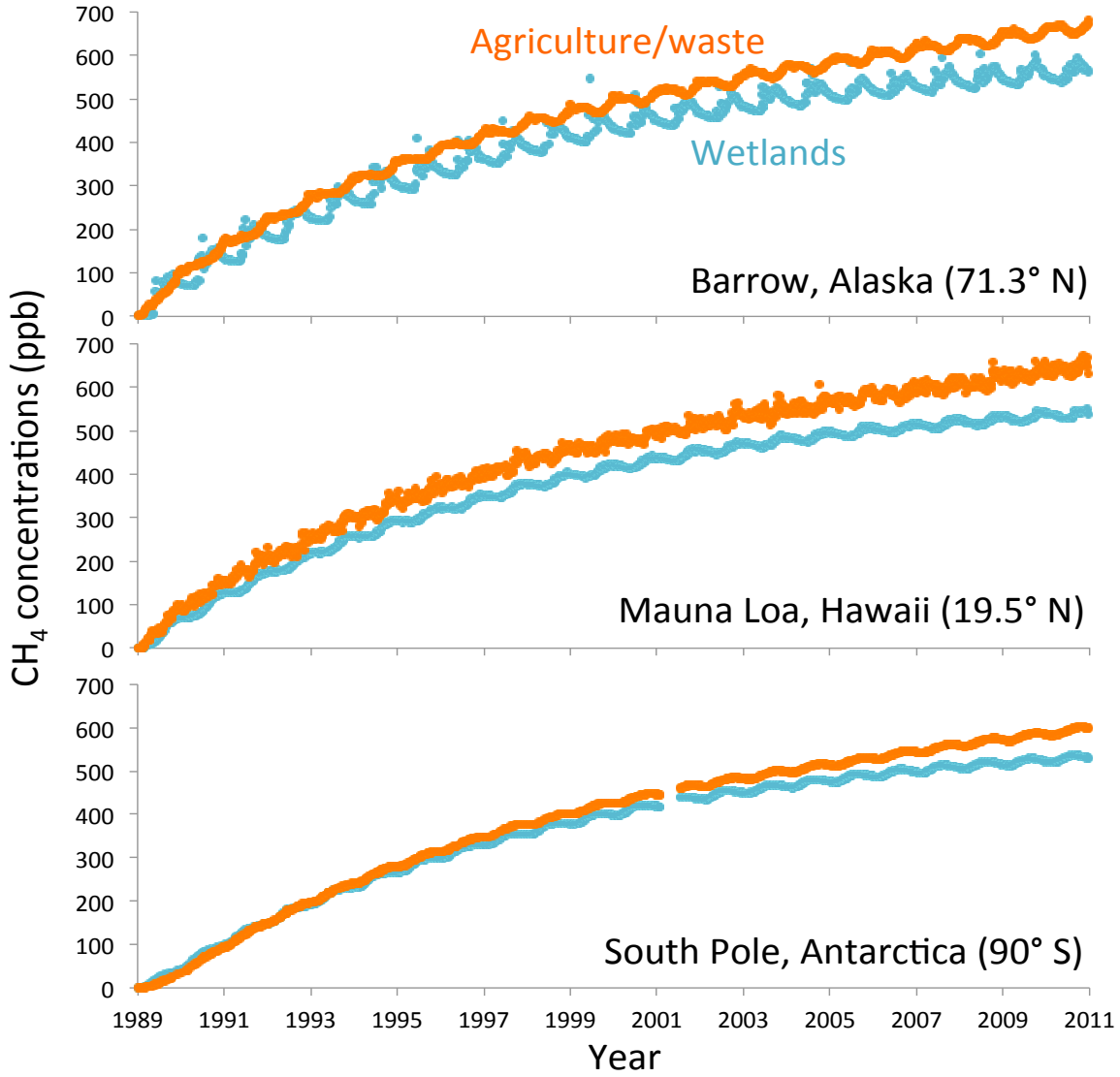

Figure 33: CarbonTracker-CH_4 simulated CH_4 concentration contributions (excluding CH_4 sink) from agriculture/waste and wetland sources at three select sites. The gap in Antarctica around 2001 is due to an error in data retrieval from the simulation results with no effects on results in the following years.

The low magnitude of the inter-hemispheric simulation-measurement mismatch, less than 5% of the global budget, indicates that the spatial misallocation of the *priors* is small. Since the spatial fit is largely determined by the allocation of emissions sources, the relatively good fit of simulation and measurement supports the best estimate FE rates in the box-model (2-5% from 1990-2010) as opposed to the upper bound estimates (7% in 2000 and higher in earlier years). The inter-hemispheric gradient and seasonal comparisons show that an improved spatial emissions allocation includes (i) a *prior* emissions transfer from Northern to Southern wetland emissions and/or (ii) increased *prior* Northern agriculture/waste emissions in combination with reduced NG FE relative to the best estimate (3-4%). Thus neither the inter-hemispheric gradient nor the seasonal comparisons suggest that a global average NG FE rate of 3-4% over the period 1990-2010 is too low. This conclusion is subject to potential imprecision of the TM5 emissions transport model, which may lead to uncertainties in the simulated spatial allocation of CH_4 concentration. This could affect the perceived accuracy of the best estimate NG FE rates, but evaluating TM5 is beyond the scope of this work.

5.4 Summary and Conclusions

This chapter described the observational constraints on the *prior* emissions estimates and NG FE scenarios developed in Chapter 4. A global box-model was built as a computationally efficient way to analyze a range of emissions scenarios and global CH_4 and C_2H_6 emissions burdens, the latter representing uncertainty in atmospheric chemistry. A global $\delta^{13}C$-CH_4 isotope mass balance was carried out in order to use an additional observational constraint on the NG FE rate. Finally, the CH_4 emissions grid maps developed in Chapter 4 were applied in NOAA's CT-CH_4 model to add spatial emissions constraints.

Box-model, isotope mass balance, and CT-CH_4 results agree on best estimate NG FE rates of 3-5% globally over the past decade. C_2H_6 measurements indicate upper bound FE rates of 5% and 7% in 2010 and 2000, respectively, assuming an upper bound global emissions burden based on measurements and literature as well as lower bound literature C_2H_6 FF content. In contrast to C_2H_6, NG FE rates below 9% over the past decade cannot be ruled out with global average CH_4 and $\delta^{13}C$-CH_4 measurements alone due to uncertainty in non-FF sources. However, spatial constraints derived from CT-CH_4 support the best estimate range of 3-5% over the past decade. The inter-hemispheric gradient shows that the best estimates non-FF *prior* CH_4 emissions in Chapter 4 are too high in the Northern hemisphere. Furthermore, simulations and measurements were inconsistent across seasons for any NG FE scenario. Adjusting the best estimate 3-5% NG FE rate upwards would only aggravate the inter-hemispheric disagreement between simulations and measurements. Taking into account the global CH_4 budget, it was concluded that overestimated Northern wetland emissions and underestimated Southern wetland emissions best explain both the inter-hemispheric and the seasonal discrepancy.

Model results suggest that FE rates have decreased over time. The best estimate NG FE rate around 1990 ranges from 5-8%, which is consistent using global average CH_4,

C_2H_6, and $\delta^{13}C$-CH_4 measurements. This is supported by CT-CH_4 unless non-FF CH_4 emissions have changed significantly prior to 2000. Box-modeling suggests that even conservative estimates of natural seepage (40 Tg CH_4/yr and 2 Tg C_2H_6/yr) reduce NG FE rates by about two percentage points. In this case, the best estimate NG FE rate would be less than 1% in 2010. Such low global rate is lower than any recent U.S. bottom-up estimate, and is therefore very unlikely. Hence, in order to defend the magnitude of the modeled natural seepage source, best estimate natural and/or agriculture/waste, and biomass burning *priors* might be too high.

Best estimate global average NG FE rates of 3% in 2010 in this work suggest that recent high LCA estimates of 6-8% FE rates in the U.S. (*28*) may be possible at individual sites, but do not appear representative as a national average. This conclusion assumes that U.S. industry practices do not lead to significantly higher emissions compared to the global average. Given the declining NG FE rates over time estimated here, it is also possible that the above high LCA estimates are due to outdated literature EFs. The observed steady decline in NG FE rates suggests incremental improvements of industry practices and/or gradual adaptation throughout the industry rather than immediate innovations at a particular point in time. Based on the results presented here, it is difficult to identify particularly effective NG FE mitigation practices or technologies to be adopted elsewhere in the industry. Nevertheless, the steadily declining FE rates indicate that further reductions are possible, and the referenced LCA studies point out the processes with greatest potential for improvement. In this light, requests from policy-makers to improve industry practices, such as EPA's New Source Performance Standards (*179*), appear promising and enforcement should be ensured.

As pointed out in previous inversion studies, the global observational network is currently too sparse to distinguish CH_4 sources at a country level with confidence, including NG FE. The large *posterior* uncertainties in this work support this finding, particularly for CH_4, which calls for an expansion of the current network. The upper bound global average NG FE was constrained most effectively using C_2H_6 measurements. However, uncertainties could be reduced even further if more associated and unassociated NG composition sample data were publicly available from basins across the world given the large variability of C_2H_6 content. Policies aimed at providing this composition data, e.g., requiring drilling companies to report the data or supporting research to generate the data, would improve the accuracy of NG FE estimates.

Local and regional inversion studies using field measurements can complement global inversion modeling, which may provide basin specific NG FE rates unattainable with the current global observation network. Some of these studies are currently underway or planned, and further research support is needed to enable their realization across basins in the U.S. The industry average NG FE estimates from this work can be used as a reference, and basin specific studies may point out areas with local or regional hot spots.

The potential *initial* adverse climate impacts of a global coal-to-gas transition are likely at the less severe end of the spectrum as formulated by Wigley (*36*). According to the

author's model, and assuming the best estimate global average NG FE rates of 3-5% over the past decade suggest that global temperatures will *increase* by less than 0.05°C *before* 2060, and *decrease* by a larger amount *thereafter*. Initial temperature increases will be smaller and time scales will be shorter after incorporating further reductions in NG FE rates in the future. As a comparison, a recent analysis by Smith and Mizrahi suggests that reducing all anthropogenic CH_4 emissions by 54% (from 257 to 119 Tg CH_4/yr) over the 2005 to 2030 period would reduce global temperatures by about 0.1°C in 2050 (*180*). While Smith and Mizrahi conclude that policies to reduce short-lived climate forcers, such as CH_4, only have moderate impacts on near-term global climate warming, the comparison with Wigley (*36*) emphasizes the significance of NG FE for global anthropogenic CH_4 reductions.

Chapter 6: Conclusions

The climate implications of biofuel and NG use have been two of the most widely discussed research topics in the LCA community over the past few years. Research interest has increased as biofuel and NG production in the U.S. grew rapidly over the last decade. The perceived desirability of corn ethanol as a renewable transportation fuel ebbed as studies indicated significant GHG emissions from direct and indirect LUC. Detrimental climate impacts from emissions timing related to LUC were also suggested, but conclusions were based on only a few studies in an emerging field of research. The NG FE rate and the corresponding CH_4 emissions have been highly uncertain for several decades. Bottom-up studies – including LCA – are highly dependent on relatively small samples of industry emissions data, which is a major contributor to uncertainty in NG FE rates. This work has addressed both issues by taking an atmospheric science perspective. The contribution of this dissertation is twofold. First, existing atmospheric modeling approaches were developed further. The emissions timing model presented incorporates several real-world processes, which were not accounted for in previous studies. The inversion box-model was designed and the CT-CH_4 model was adopted to explicitly quantify uncertainty in NG FE as opposed to deriving best estimates for all CH_4 sources. Second, this work bridges the gap between the atmospheric modeling and LCA communities. The careful selection of metrics and biofuel GHG emissions estimates over time was key for estimating the effect of emissions timing in the RF model. Establishing more transparent Bayesian *prior* CH_4 and C_2H_6 emissions inventories were key for quantifying NG FE rates based on observational constraints.

6.1 Research questions and brief answers

Chapter 2: Biofuel emissions timing

1) *What is the magnitude of the emissions timing effect, i.e., underestimated RF impacts in traditional LCA, after addressing the above-mentioned limitations of previous models?*

The emissions timing factor (ETF), i.e., the influence of emissions timing on RF is significantly smaller than 9-89% suggested in previous studies. In the base case, ETF is 2, 5, and 13% for impact time frames of 100, 50, and 30 years, respectively. Accounting for time may increase corn ethanol's GHG balance from -20% (no ETF) to -18%, -15%, and -7% with respect to the above impact time frames. The results are sensitive to the uncertainty in the LUC estimates and the time period over which the fuel is produced. Emissions timing adds little to our understanding of the climate impacts of corn ethanol because the uncertainty inherent in the LUC estimates (that triggers ETF) contributes one to two orders of magnitude more to life cycle emissions uncertainty than ETF.

2) *By what percentage does adding the LUC induced emissions timing factor to the traditional LCA GHG metric reduce the probability of corn ethanol meeting the EISA target?*

From a policy maker's perspective, emissions timing is more important because ETF increases the chance of corn ethanol missing the EISA target of 20% GHG reduction relative to gasoline. According to EPA, the probability of meeting the target is 53%, but even a small skew in the LCA GHG distribution can significantly reduce the probability of meeting this target. This work suggests that ETF reduces this probability to 7-29% depending on the chosen impact time frame. Thus, switching from the *mean GHG balance* to *probability of meeting the target* amplifies the influence of ETF from 2-13 percent points to 24-46% in reduced probability of meeting the EISA target.

3) *What is the absolute RF impact of U.S. corn ethanol production and use including LUC, and what are the implications for climate policy?*

The impacts from corn ethanol emissions represent a very small contributor to reducing climate change as expressed in RF units, even when considering the relatively low EPA LUC emissions estimate. Most of the emissions "cost" of corn ethanol has occurred over the last decade due to LUC effects from expanding ethanol production. This cost was estimated to be 0.0006 W/m^2 in the base case (15 Bgal/yr), which translates into a first order estimate of 0.0007 °C increase in global average surface temperature. The model predicts a *reduction* in RF and T_s at the end of the life cycle, which is in the same order of magnitude. It appears very unlikely that either temperature change (first positive, then negative) would entail increased or reduced damages of an order of magnitude that can actually be observed or quantified. Both the low probability of meeting the EISA target as well as the small magnitude of climate benefits (if any) should be taken into account when allocating climate change mitigation funds in support of corn ethanol.

Chapters 3-5: Uncertainty in NG FE rates

1) *What are the best estimate NG FE rates over the past two decades globally according to CH_4, C_2H_6, and $\delta^{13}C$-CH_4 tracers, and is there a trend over time?*

Box-model, isotope mass balance, and 3D-modeling results agree on best estimate NG FE rates of 3-5% globally over the past decade. Discrepancies in the inter-hemispheric gradient and among seasons between simulations and measurements suggest overestimated *prior* Northern wetland emissions. Adjusting the best estimate NG FE rate upwards would only aggravate the inter-hemispheric disagreement. FE rates have likely decreased over time. The best estimate NG FE rate around 1990 ranges from 5-8%, which is consistent using global average CH_4, C_2H_6, and $\delta^{13}C$-CH_4 measurements. This is supported by CT-CH_4 unless non-FF CH_4 emissions have changed significantly prior to 2000.

2) *What is the upper bound of global NG FE given the lower bound of all other CH_4 and C_2H_6 sources, and what are the prior assumptions (e.g., FF emissions factors and hydrocarbon composition) underlying this bound?*

Ethane measurements indicate upper bound FE rates of 5% and 7% in 2010 and 2000, respectively, assuming an upper bound global emissions burden based on measurements and literature as well as lower bound literature C_2H_6 FF content. In contrast to C_2H_6, NG FE rates below 9% over the past decade cannot be ruled out with global average CH_4 and $\delta^{13}C$-CH_4 measurements alone due to uncertainty in non-FF

sources. Best estimate and upper bound values may be overestimated because both assume lower bound emissions from oil and coal production as well as complete absence of natural seepage.

3) *Can literature estimates of natural CH_4 and C_2H_6 seepage be reconciled with global measurements, and how does this affect the NG FE estimates?*

Box-modeling suggests that even conservative estimates of natural seepage (40 Tg CH_4/yr and 2 Tg C_2H_6/yr) reduce NG FE rates by about two percentage points. In this case, the best estimate NG FE rate would be less than 1% in 2010. Such low global rate is even lower than any recent U.S. bottom-up estimate, and seems hence very unlikely. Hence, in order to defend the magnitude of the modeled natural seepage source, best estimate natural and/or agriculture/waste, and biomass burning *priors* must be considered too high.

4) *What are the policy implications of global atmospheric observational constraints on NG FE rates?*

Best estimate global average NG FE rates of 3% in 2010 in this work suggest that recent high LCA estimates of 6-8% FE rates in the U.S. (*28*) may be possible at individual sites, but do not appear representative as a national average. This assumes that U.S. industry practices do not lead to significantly higher emissions compared to the global average. The steadily declining NG FE rates suggest that further reductions are possible, and the referenced LCA studies point out the processes with greatest potential for improvement. In this light, requests from policy-makers to improve industry practices, such as EPA's New Source Performance Standards (*179*), appear promising and enforcement should be ensured. The global observational network is currently too sparse to distinguish CH_4 sources at a country level with confidence, including NG FE, which calls for an expansion of the current network. Policies aimed at providing a greater sample size of associated and unassociated NG composition data (specifically C_2H_6 content), e.g., requiring drilling companies to report the data or supporting research to generate the data, would improve the accuracy of NG FE estimates. Research support for basin specific inversion studies using field measurements is useful for complementing global inversion modeling. This may yield results indicating areas with local or regional hot spots where reductions in NG FE are most needed. Combining the best estimates NG FE rates of 3-5% and previous analysis by Wigley (*36*), a global coal-to-gas transition scenario may increase temperatures by less than 0.05°C before 2060, and decrease by a larger amount thereafter.

6.2 Future work

Biofuel emissions timing

1) Employing a set of more sophisticated climate models would allow quantifying ETF based on temperature changes. This may be important to account for (i) the time gap between RF and temperature changes (in the order of a decade), and (ii) the added uncertainty inherent in the climate response, both of which are masked in RF based ETF. Also, the climate model did not account for non-GHG emissions, such as aerosols

and black carbon, which are also released during LUC from biomass burning. Aerosol and black carbon emissions have a negative and positive RF impact, respectively, and modeling these emissions may significantly change RF from biofuel GHG emissions. While this may be important for biofuel LCA in general, it could also increase or reduce the influence of emissions timing. However, it does not address other issues of GHG emissions, such as the acidification of the oceans.

2) This case study looked at relatively small emissions or emissions reductions. Changing traditional LCA methodology by accounting for emissions timing may be reasonable only if timing actually impacts climate on an absolute scale. While GHG emissions clearly add up among different energy technologies, this accumulation effect is not obvious for emissions timing. Future research should explore scenarios in which different energy technologies exhibiting large upfront emissions generate significant climate impacts due to emissions timing. As an example, such scenarios could include carbon capture and sequestration or NG production from shale formations, whose temporal emissions patterns may vary significantly.

Uncertainty in NG FE rates

1) A formal statistical analysis is needed to estimate probability distribution functions (PDF) of the global average NG FE rate (over time) in addition to best estimates and upper bound scenarios. This can be achieved by running the inversion box-model in a Monte Carlo simulation, thereby explicitly incorporating *prior* probability distributions of the most uncertain model parameters. Specific parameters should include hydrocarbon composition of associated and unassociated NG, CH_4 lifetime, the global C_2H_6 burden, and non-FF CH_4 and C_2H_6 sources. This approach allows sampling a wide range of *prior* emissions estimates based on distributions derived from a more comprehensive literature review compared to this work. The PDFs will provide researchers and policy-makers with more detailed estimates of NG FE rates.

2) The box-model results can be further refined using 3D-modeling not included in this work. First, C_2H_6 forward modeling, e.g., using CT-CH_4, may provide important insights about NG FE rates because the vast majority of C_2H_6 emissions is due to FFs, which are released in the Northern hemisphere. Second, generating and simulating global grid maps of CH_4 and C_2H_6 from natural seepage is needed to better understand its magnitude, and thus its influence on NG FE rates. Coordinates of global natural seepage sites are readily available (*109*, *144*), but have not yet been included in global inversion modeling. Finally, the box- and 3D-model results can be used to inform *prior* distributions for a full inversion. This also includes comparing simulation and measurement at specific sites of interest, e.g., Southern Great Plains (SGP) with relatively small non-FF emissions influence, as a first indicator of regional differences in NG FE rates.

References

1. EIA, *International Energy Statistics* (2013), http://www.eia.gov/cfapps/ipdbproject/iedindex3.cfm?tid=2&pid=38&aid=12&cid=regions&syid=1980&eyid=2011&unit=BKWH.

2. U.S. Congress, Using biofuel tax credits to achieve energy and environmental policy goals (2010), available at www.cbo.gov/doc.cfm?index=11477&type=1.

3. IEA, *World Energy Outlook 2010* (2010).

4. IPCC, *Climate Change 2007: Synthesis Report* T. Barker, Ed. (Cambridge, United Kingdom and New York, NY, USA, 2007), pp. 12–17.

5. K. Keller, D. McInerney, D. F. Bradford, Carbon dioxide sequestration: how much and when?, *Climatic Change* **88**, 267–291 (2008).

6. S. Dietz, N. Stern, Note--On the Timing of Greenhouse Gas Emissions Reductions: A Final Rejoinder to the Symposium on "The Economics of Climate Change: The Stern Review and its Critics," *Review of Environmental Economics and Policy* **3**, 94–113 (2008).

7. EPA, *Renewable Fuel Standard Program (RFS2) Regulatory Impact Analysis* (Environmental Protection Agency, 2010).

8. A. E. Farrell *et al.*, Ethanol can contribute to energy and environmental goals., *Science* **311**, 506–508 (2006).

9. K. A. Mullins, W. M. Griffin, H. S. Matthews, Policy Implications of Uncertainty in Modeled Life-Cycle Greenhouse Gas Emissions of Biofuels, *Environmental Science & Technology* **45**, 132–138 (2010).

10. C. L. Weber, C. Clavin, Life Cycle Carbon Footprint of Shale Gas: Review of Evidence and Implications, *Environmental Science Technology* **46** (11), pp 5688–5695 (2012).

11. M. A. Curran, Environmental life-cycle assessment, *The International Journal of Life Cycle Assessment* **1**, 231–237 (1996).

12. ISO, 14040: Environmental management–life cycle assessment—Principles and framework. Vol. 3, *International Organization for Standardization* **3** (2006).

13. B. W. Vigon *et al.*, Life Cycle Assessment: Inventory Guidelines and Principles (1993).

14. C. T. Hendrickson, L. B. Lave, H. S. Matthews, *Environmental Life Cycle Assessment of Goods and Services: An Input-Output Approach* (Resources for the Future, 2006.

15. M. O'Hare *et al.*, Proper accounting for time increases crop-based biofuels' greenhouse gas deficit versus petroleum, *Environmental Research Letters* **4**, 024001 (2009).

16. A. Kendall, B. Chang, B. Sharpe, Accounting for time-dependent effects in biofuel life cycle greenhouse gas emissions calculations., *Environmental Science & Technology* **43**, 7142–7 (2009).

17. A. Levasseur, P. Lesage, M. Margni, L. Deschênes, R. Samson, Considering time in LCA: dynamic LCA and its application to global warming impact assessments., *Environmental Science & Technology* **44**, 3169–74 (2010).

18. EIA, *Petroleum statistics* (2013; http://www.eia.gov/oog/info/twip/twip_gasoline.html#production).

19. RFA, *2013 Ethanol Industry Outlook* (2013).

20. R. J. Plevin, The greenhouse gas emissions from biofuels' indirect land use change are uncertain, but may be much greater than previously estimated, *Environmental Science & Technology* **44**, 8015–8021 (2010).

21. T. Searchinger *et al.*, Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change., *Science* **319**, 1238–40 (2008).

22. M. Wang, *GREET Model: The Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation Model* (Argonne, IL, 2010; http://greet.es.anl.gov/).

23. U.S. Congress, Energy Independence and Security Act of 2007*Washington, DC: Public Law* **20544**, 110–140 (2007).

24. S. Schwietzke, W. M. Griffin, H. S. Matthews, Relevance of emissions timing in biofuel greenhouse gases and climate impacts, *Environmental Science and Technology* **45**, 8197–8203 (2011).

25. EIA, *International Energy Outlook 2011* (US Energy Information Administration (EIA), 2011; http://www.eia.gov/forecasts/ieo/index.cfm).

26. EIA, *Annual Energy Outlook 2011* (Office of Integrated Analysis and Forecasting, Energy Information Agency, US Department of Energy, 2011.

27. A. Burnham *et al.*, Life-Cycle Greenhouse Gas Emissions of Shale Gas, Natural Gas, Coal, and Petroleum, *Environmental Science & Technology* **46**, 619–627 (2012).

28. R. W. Howarth, R. Santoro, A. Ingraffea, Methane and the greenhouse-gas footprint of natural gas from shale formations, *Climatic Change* **106**, 679–690 (2011).

29. N. Hultman, D. Rebois, M. Scholten, C. Ramig, The greenhouse impact of unconventional gas for electricity generation, *Environmental Research Letters* **6**, 044008 (2011).

30. P. Jaramillo, W. M. Griffin, H. S. Matthews, Comparative life-cycle air emissions of coal, domestic natural gas, LNG, and SNG for electricity generation., *Environmental Science Technology* **41**, 6290–6296 (2007).

31. M. Jiang *et al.*, Life cycle greenhouse gas emissions of Marcellus shale gas, *Environmental Research Letters* **6**, 034014 (2011).

32. NETL, *Life Cycle Greenhouse Gas Inventory of Natural Gas Extraction, Delivery and Electricity Production* (National Energy Technology Laboratory, 2011).

33. T. Stephenson, J. E. Valle, X. Riera-Palou, Modeling the Relative GHG Emissions of Conventional and Shale Gas Production, *Environmental Science & Technology* **45** (24), 10757–10764 (2011).

34. A. Venkatesh, P. Jaramillo, W. M. Griffin, H. S. Matthews, Uncertainty in Life Cycle Greenhouse Gas Emissions from United States Natural Gas End-Uses and its Effects on Policy., *Environmental Science and Technology* **45**, 8182–8189 (2011).

35. K. Hayhoe, H. S. Kheshgi, A. K. Jain, D. J. Wuebbles, Substitution of Natural Gas for Coal: Climatic Effects of Utility Sector Emissions, *Climatic Change* **54**, 107–139 (2002).

36. T. M. L. Wigley, Coal to gas: the influence of methane leakage, *Climatic Change* **108**, 601–608 (2011).

37. A. I. Reshetnikov, N. N. Paramonova, A. A. Shashkov, An evaluation of historical methane emissions from the Soviet gas industry, *Journal of Geophysical Research* **105**, 3517–3529 (2000).

38. C. Mitchell, J. Sweet, T. Jackson, A study of leakage from the UK natural gas distribution system, *Energy Policy* **18**, 809–818 (1990).

39. NOAA, CarbonTracker-CH4. *http://www.esrl.noaa.gov/gmd/ccgg/carbontracker-ch4/index.html* (2012).

40. P. Bousquet *et al.*, P. Ciais *et al.*, Eds. Contribution of anthropogenic and natural sources to atmospheric methane variability, *Nature* **443**, 439–43 (2006).

41. J. S. Wang *et al.*, A 3-D model analysis of the slowdown and interannual variability in the methane growth rate from 1988 to 1997, *Global Biogeochemical Cycles* **18** (2004).

42. P. Bergamaschi *et al.*, Inverse modeling of global and regional CH4 emissions using SCIAMACHY satellite retrievals, *J. Geophys. Res.* **114**, D22301 (2009).

43. Y.-H. Chen, R. G. Prinn, Estimation of atmospheric methane emissions between 1996 and 2001 using a three-dimensional global chemical transport model, *Journal of Geophysical Research, Atmospheres* **111** (2006).

44. S. E. Mikaloff Fletcher, P. P. Tans, L. M. Bruhwiler, J. B. Miller, M. Heimann, CH4 sources estimated from atmospheric observations of CH4 and its 13C/12C isotopic ratios: 1. Inverse modeling of source processes, *Global Biogeochemical Cycles* **18**, GB4004 (2004).

45. RFA, *Historic U.S. fuel Ethanol Production* (Renewable Fuels Association, 2010; http://ethanolrfa.org/pages/statistics/#A).

46. T. W. Hertel *et al.*, Effects of US Maize Ethanol on Global Land Use and Greenhouse Gas Emissions: Estimating Market-mediated Responses, *BioScience* **60**, 223–231 (2010).

47. W. E. Tyner, F. Taheripour, Q. Zhuang, D. Birur, U. Baldos, *Land Use Changes and Consequent CO2 Emissions due to US Corn Ethanol Production: A Comprehensive Analysis* (2010).

48. IPCC, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Chapter 2* S. Solomon *et al.*, Eds. (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007).

49. F. Joos *et al.*, Global warming feedbacks on terrestrial carbon uptake under the Intergovernmental Panel on Climate Change (IPCC) emission scenarios, *Global Biogeochemical Cycles* **15**, 891–907 (2001).

50. J. T. Houghton, L. G. Meira Filho, D. J. Griggs, K. Maskell, *An introduction to simple climate models used in the IPCC second assessment report* (Intergovernmental Panel on Climate Change, 1997; http://www.phys.port.ac.uk/units/2007/global/ipcctr2.pdf).

51. EIA, *Emissions of Greenhouse Gases Report* (Energy Information Administration, 2009; http://www.eia.doe.gov/oiaf/1605/ggrpt/carbon.html).

52. EIA, *U.S. Electricity demand projections.* (Energy Information Administration, 2010; http://www.eia.doe.gov/oiaf/aeo/electricity.html).

53. EIA, *U.S. Petroleum consumption* (Energy Information Administration, 2010; http://tonto.eia.doe.gov/dnav/pet/pet_cons_psup_dc_nus_mbbl_a.htm).

54. EPA, *Greenhouse Gas Emissions Reporting from the Petroleum and Natural Gas Industry* (2010); http://www.epa.gov/ghgreporting/documents/pdf/2010/Subpart-W_TSD.pdf.

55. IEA, *World Energy Outlook 2011 - Special Report: Are we entering a golden age of gas?* (OECD Publishing, 2011; http://www.oecd-ilibrary.org/energy/world-energy-outlook-2006_weo-2006-en), p. 329.

56. G. Morgan, J. Apt, L. Lave, The U.S. Electric Power Sector and Climate Change Mitigation (2005; http://wpweb2.tepper.cmu.edu/ceic/papers/USElectricPower.pdf).

57. EIA, *January 2012 Monthly Energy Review* (US Energy Information Administration (EIA), 2012).

58. D. Anthoff, R. S. J. Tol, The Impact of Climate Change on the Balanced-Growth-Equivalent: An Application of FUND, *Environmental and Resource Economics* **43**, 351–367 (2009).

59. W. D. Nordhaus, *A Question of Balance: Weighing the Options on Global Warming Policies* (Yale University Press, 2008).

60. A. Grübler, S. Messner, Technological change and the timing of mitigation measures, *Energy Economics* **20**, 495–512 (1998).

61. S. Kallbekken, N. Rive, Why delaying emission reductions is a gamble, *Climatic Change* **82**, 27–45 (2007).

62. R. A. Alvarez, S. W. Pacala, J. J. Winebrake, W. L. Chameides, S. P. Hamburg, Greater focus needed on methane leakage from natural gas infrastructure., *Proceedings of the National Academy of Sciences of the United States of America* **109**, 6435–40 (2012).

63. EIA, *NG well-head prices in the U.S.* (2013; http://www.eia.gov/dnav/ng/ng_pri_sum_dcu_nus_m.htm).

64. EPA, *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990 – 2011* (2013; http://www.epa.gov/climatechange/ghgemissions/usinventoryreport.html).

65. EPA, *Natural Gas STAR Program: Recommended Technologies and Practices* (2013; http://www.epa.gov/gasstar/tools/recommended.html).

66. EPA, *Analysis of Methane Mitigation Options using the MARKAL Model for the US* (U.S. Environmental Protection Agency. Coalbed Methane Outreach Program, 2004).

67. J. Lelieveld *et al.*, Low methane leakage from gas pipelines: A switch from coal or oil to natural gas could mitigate climate effects in the short term, *Nature* **434**, 841–842 (2005).

68. ANGA, *Greenhouse Gas Emission Estimation Methodologies, Procedures, and Guidelines for the Natural Gas Distribution Sector* (Washington, D.C., 2008; American Gas Association).

69. EPA, *Natural Gas Industry Methane Emission Factor Improvement Study* (Washington D.C., 2011).

70. API/ANGA, *Characterizing Pivotal Sources of Methane Emissions from Natural Gas Production: Summary and Analysis of API and ANGA Survey Responses* (2012).

71. D. T. Allen *et al.*, Measurements of methane emissions at natural gas production sites in the United States, *Proceedings of the National Academy of Sciences* (2013).

72. EPA, *EPA Needs to Improve Air Emissions Data for the Oil and Natural Gas Production Sector* (2013).

73. J. V Dedikov *et al.*, Estimating methane releases from natural gas production and transmission in Russia, *Atmospheric Environment* **33**, 3291–3299 (1999).

74. EPA, *Global Mitigation of Non-CO2 Greenhouse Gases* (United States Environmental Protection Agency, http://www.epa.gov/climatechange/economics/international.html, 2006; international.html), pp. 1–438.

75. API, *Compendium of Greenhouse Gas Eemissions Methodologies for the Oil and Natural Gas Industry* (2009).

76. G. Pétron, J. Kofler, G. Frost, Emissions from oil and natural gas operations in northeastern Utah; *AGU Fall meeting*, (San Francisco, 2013).

77. G. Petron *et al.*, Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study, *Journal of Geophysical Research* **117**, D04304 (2012).

78. M. Schoell, Genetic Characterization of Natural Gases, *American Association of Petroleum Geologists Bulletin* **67**, 2225–2238 (1983).

79. A. Karion *et al.*, Methane emissions estimate from airborne measurements over a western United States natural gas field, *Geophysical Research Letters* (2013).

80. F. O'Sullivan, S. Paltsev, Shale gas production: potential versus actual greenhouse gas emissions, *Environmental Research Letters* **7**, 44030 (2012).

81. E. J. Dlugokencky, E. G. Nisbet, R. Fisher, D. Lowry, Global atmospheric methane: budget, changes and dangers., *Philosophical Transactions of the Royal Society - Series A: Mathematical, Physical and Engineering Sciences* **369**, 2058–2072 (2011).

82. L. M. P. Bruhwiler, A. M. Michalak, W. Peters, D. F. Baker, P. Tans, An improved Kalman Smoother for atmospheric inversions, *Atmospheric Chemistry and Physics* **5**, 2691–2702 (2005).

83. NOAA, ESRL Global Monitoring Division, *http://www.esrl.noaa.gov/gmd/ccgg/iadv/* (2012).

84. I. J. Simpson *et al.*, Long-term decline of global atmospheric ethane concentrations and implications for methane, *Nature* **488**, 490–494 (2012).

85. G. Janssens-Maenhout *et al.*, *EDGAR-HTAP: a harmonized gridded air pollution emission dataset based on national inventories* (Luxembourg, 2012).

86. J. F. Meirink, H. J. Eskes, A. P. H. Goede, Sensitivity analysis of methane emissions derived from SCIAMACHY observations through inverse modelling, *Atmospheric Chemistry and Physics Discussions* **5**, 9405–9445 (2005).

87. P. Bergamaschi *et al.*, Inverse modelling of national and European CH4 emissions using the atmospheric zoom model TM5, *Atmospheric Chemistry and Physics Discussions* **5**, 2431–2460 (2005).

88. P. Bergamaschi *et al.*, Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations, *Journal of Geophysical Research* **112**, 1–26 (2007).

89. O. Schneising *et al.*, Atmospheric greenhouse gases retrieved from SCIAMACHY: comparison to ground-based FTS measurements and model results, *Atmospheric Chemistry and Physics* **12**, 1527–1540 (2012).

90. C. D. Holmes, M. J. Prather, O. A. Sovde, G. Myhre, Future methane, hydroxyl, and their uncertainties: key climate and emission parameters for future predictions, *Atmospheric Chemistry and Physics* **13**, 285–302 (2013).

91. S. Houweling, Global Modeling of atmospheric methane sources and sinks, Ph. D. thesis, Utrecht University (2000).

92. S. Houweling, T. Kaminski, F. Dentener, J. Lelieveld, M. Heimann, Inverse modeling of methane sources and sinks using the adjoint of a global transport model, *Journal of Geophysical Research* **104**, 26137–26160 (1999).

93. M. Kandlikar, Bayesian inversion for reconciling uncertainties in global mass balances, *Tellus Series B Chemical And Physical Meteorology* **49B**, 123–135 (1997).

94. EDGAR, *Global CH4 emissions grid maps* (2012; http://edgar.jrc.ec.europa.eu/datasets_grid_list.php?v=42&edgar_compound=CH4).

95. L. J. Molofsky, J. A. Connor, S. K. Farhat, A. S. Wylie Jr., T. Wagner, Methane in Pennsylvania water wells unrelated to Marcellus shale fracturing, *Oil & Gas Journal* **109**, 54+ (2011).

96. EDGAR, Uncertainties in EDGAR3, *http://themasites.pbl.nl/tridion/en/themasites/edgar/documentation/uncertainties/index-2.html* (2010).

97. EIA, *U.S. NG production statistics* (2013; http://www.eia.gov/dnav/ng/ng_prod_sum_dcu_NUS_m.htm).

98. NaturalGas.org, *Overview of Natural Gas* (2013; http://www.naturalgas.org/lng/lng.asp).

99. EIA, *Natural Gas Processing: The Crucial Link Between Natural Gas Production and Its Transportation to Market* (2006).

100. M. V. Okrepilov, Importance and main areas of metrological provision for associated gas recovery and utilization, *Measurement Techniques* **55** (2012).

101. CAPP, *A National Inventory of Greenhouse Gas (GHG), Criteria Air Contaminant (CAC) and Hydrogen Sulphide (H2S) Emissions by the Upstream Oil and Gas Industry* (Calgary, Alberta, Canada, 2004).

102. G. C. Segeler, *Gas Engineers Handbook* American Gas Association, Ed. (The Industrial Press, New York, NY, ed. 1, 1966).

103. U. Rick, *Personal communication, with data from WRAP Phase III report* (Denver, CO, 2013).

104. B. D. Gage, B. of L. M. D. L. Driskill, *Analyses of natural gases 1917-2007* (2008).

105. G. Marland, R. M. Rotty, Carbon dioxide emissions from fossil fuels: a procedure for estimation and results for 1950-1982, *Tellus* **36B**, 32–261 (1984).

106. G. Pétron, Personal communication, with data from WRAP Phase III report, (2013).

107. J. Logan *et al.*, *Natural Gas and the Transformation of the U.S. Energy Sector: Electricity* (Golden, CO, 2012).

108. A. Bar-Ilan, *Development of baseline 2006 emissions from oil and gas activity in the Denver-Julesburg Basin, WRAP Phase III report* (Fort Collins, CO, 2008; http://www.wrapair.org/forums/ogwg/documents/2008-04_06_Baseline), p. 34.

109. G. Etiope, Personal communication (2013).

110. M. J. Kotarba, M. D. Lewan, Characterizing thermogenic coalbed gas from Polish coals of different ranks by hydrous pyrolysis, *Organic Geochemistry* **35**, 615–646 (2004).

111. A. G. Kim, *The Composition of Coalbed Gas* (Washington, D.C., 1973).

112. M. R. Harrison, T. M. Shires, J. K. Wessels, R. M. Cowgill, *Methane Emission from the Natural Gas Industry* (Chicago, IL, 1996).

113. C. Dienst *et al.*, *Treibhausgasemissionen des russischen Exportpipeline System – Ergebnisse und Hochrechnungen empirischer Untersuchungen in Russland* (Mainz, Germany, 2004).

114. K. Phumpradab, S. H. Gheewala, M. Sagisaka, Life cycle assessment of natural gas power plants in Thailand, *The International Journal of Life Cycle Assessment* **14**, 354–363 (2009).

115. C. Dinca, P. Rousseaux, A. Badea, A life cycle impact of the natural gas used in the energy sector in Romania, *Journal of Cleaner Production* **15**, 1451–1462 (2007).

116. M. Faist-Emmenegger, T. Heck, N. Jungbluth, M. Tuchschmid, *Project "ecoinvent data v2.0", Teil V Erdgas* (Villingen, 2007).

117. R. Dones, T. Heck, M. Faist Emmenegger, N. Jungbluth, Life Cycle Inventories for the Nuclear and Natural Gas Energy Systems, and Examples of Uncertainty Analysis (14 pp), *The International Journal of Life Cycle Assessment* **10**, 10–23 (2005).

118. G. J. van den Born, A. F. Bouwman, J. G. J. Olivier, R. J. Swart, *The emission of greenhouse gases in the netherlands* (Bilthoven, The Netherlands, 1991).

119. K. K. Bakkane, *Life Cycle Data for Norwegian Oil and Gas* (Akademika Publishing, 1995), p. 313.

120. IPCC, *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme* H. S. Eggleston, L. Buendia, K. Miwa, T. Ngara, K. Tanabe, Eds. (IGES, 2006; http://www.ipcc-nggip.iges.or.jp/public/2006gl/index.htm).

121. D. Wilson, J. Fanjoy, R. Billings, *Gulfwide Emission Inventory Study for the Regional Haze and Ozone Modeling Effort* (New Orleans, LA, 2004).

122. D. Wilson *et al.*, *Year 2008 Gulfwide Emission Inventory Study* (New Orleans, LA, 2010).

123. D. W. Barns, J. A. Edmonds, *Evaluation of the Relationship Between the Production and Use of Energy and Atmospheric Methane Emissions* (1990; http://www.ntis.gov/search/product.aspx?ABBR=DE90007982), p. 264.

124. EIA, *Emssions of Greenhouse Gases in the United States 1987-1992* (1994; http://worldcat.org/arcviewer/2/LEGAL/2009/12/22/H1261515294409/viewer/file2.html).

125. EPA, *Parameters for Properly Designed and Operated Flares* (2012).

126. P. Gogolek, A. Caverly, J. Seebold, R. Schwartz, J. Pohl, *Emissions from elevated flares – A survey of the literature* (Ottawa, Canada, 2010).

127. C. D. Elvidge, K. E. Baugh, D. Ziskin, S. Anderson, T. Ghosh, *Estimation of Gas Flaring Volumes Using NASA MODIS Fire Detection Products* (Boulder, CO, 2011).

128. C. D. Elvidge *et al.*, A Fifteen Year Record of Global Natural Gas Flaring Derived from Satellite Data, *Energies* **2**, 595–622 (2009).

129. J. S. P. Loe, O. Ladehaug, Reducing gas flaring in Russia: Gloomy outlook in times of economic insecurity, *ENERGY POLICY* **50**, 507–517 (2012).

130. GAO, *Natural Gas Flaring and Venting - Opportunities to Improve Data and Reduce Emissions* (Washington, D.C., 2004).

131. C. M. Boyer II, J. R. Kelafant, V. A. Kuuskra, K. C. Manger, *Methane emissions from coal mining: issues and opportunities for reduction* (1990).

132. EPA, *Reducing Methane Emissions From Coal Mines in China: The Potential for Coalbed Methane Development* (1996).

133. CIRI, *China Coal Industry Yearbook 2009* (2011).

134. EIA, *International Energy Statistics* (2012; http://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm?tid=1&pid=7&aid=1).

135. R. Pilcher, *Personal communication* (2012).

136. N. Aden, *Initial Assessment of NBS Energy Data Revisions* (2010; http://china.lbl.gov/publications/initial-assessment-nbs-energy-data-revisions).

137. D. Guan, Z. Liu, Y. Geng, S. Lindner, K. Hubacek, The gigatonne gap in China's carbon dioxide inventories, *Nature Climate Change* **2**, 1–4 (2012).

138. IPCC, *Climate Change 2001: The Physical Science Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Chapter 6* J. T. Houghton *et al.*, Eds. (Cambridge, United Kingdom and New York, NY, USA, 2001; http://www.grida.no/publications/other/ipcc_tar/?src=/climate/ipcc_tar/).

139. D. J. Wuebbles, K. Hayhoe, Atmospheric methane and global change, *Earth-Science Reviews* **57**, 177–210 (2002).

140. J. R. Melton *et al.*, Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP), *Biogeosciences* **10**, 753–788 (2013).

141. M. Aydin *et al.*, Recent decreases in fossil-fuel emissions of ethane and methane derived from firn air., *Nature* **476**, 198–201 (2011).

142. M. O. Andreae, P. Merlet, Emission of trace gases and aerosols from biomass burning, *Global Biogeochemical Cycles* **15**, 955–966 (2001).

143. R. Yevich, J. A. Logan, An assessment of biofuel use and burning of agricultural waste in the developing world, *Global Biogeochemical Cycles* **17** (2003), doi:10.1029/2002GB001952.

144. G. Etiope, P. Ciccioli, Earth's degassing: a missing ethane and propane source., *Science* **323**, 478 (2009).

145. G. Etiope, K. R. Lassey, R. W. Klusman, E. Boschi, Reappraisal of the fossil methane budget and related emission from geologic sources, *Geophysical Research Letters* **35**, 1–5 (2008).

146. EIA, *Natural Gas Plant Field Production* (2013; http://www.eia.gov/dnav/pet/PET_PNP_GP_DC_NUS_MBBLPD_A.htm).

147. EIA, *Definitions, Sources and Explanatory Notes* (2013; http://www.eia.gov/dnav/ng/TblDefs/ng_prod_sum_tbldef2.asp).

148. EPA, State Inventory and Projection Tool (2012) (available at http://www.epa.gov/statelocalclimate/resources/tool.html).

149. S. Claus, N. De Hauwere, B. Vanhoorne, F. Hernandez, J. Mees, *Maritime Boundaries Geodatabase* (2012; http://www.marineregions.org/downloads.php).

150. I. Fung, E. Matthews, J. Lerner, Atmospheric Methane Response to Biogenic Sources - Results from a 3-D Atmospheric Tracer Model, *Abstracts of Papers of the American Chemical Society* **193**, 6–GEOC (1987).

151. J. O. Kaplan, Wetlands at the Last Glacial Maximum: Distribution and methane emissions, *Geophysical Research Letters* **29**, 3–6 (2002).

152. E. Matthews, I. Fung, J. Lerner, Methane emission from rice cultivation: Geographic and seasonal distribution of cultivated areas and emissions, *Global Biogeochemical Cycles* **5**, 3 (1991).

153. L. Giglio, G. R. Van Der Werf, J. T. Randerson, G. J. Collatz, P. Kasibhatla, Global estimation of burned area using MODIS active fire observations, *Atmospheric Chemistry and Physics* **6**, 957–974 (2006).

154. G. R. Van Der Werf *et al.*, Interannual variability of global biomass burning emissions from 1997 to 2004, *Atmospheric Chemistry and Physics Discussions* **6**, 3175–3226 (2006).

155. GFED, *Global Fire Emissions Database* (2013; http://www.globalfiredata.org/).

156. EIA, *Natural Gas Annual Respondent Query System* (2013; http://www.eia.gov/cfapps/ngqs/ngqs.cfm?f_report=RP9&f_sortby=&f_items=&f_year_start=&f_year_end=&f_show_compid=&f_fullscreen).

157. M. Michot Foss, *Interstate Natural Gas Quality Specifications and Interchangeability* (Sugar Land, TX, 2004; www.beg.utexas.edu/energyecon/lng).

158. P. L. Spath, M. K. Mann, N. R. E. Laboratory, Ed. Life Cycle Assessment of Hydrogen Production via Natural Gas Steam Reforming, *Energy* , NREL/TP–570–27637 (2001).

159. G. Janssens-Maenhout, *Personal communication* (2012).

160. I. Fung, M. Prather, J. John, J. Lerner, E. Matthews, Three-dimensional model synthesis of the global methane cycle, *Journal of Geophysical Research* **96**, 13033–13065 (1991).

161. J. Rudolph, The tropospheric distribution and budget of ethane, *Journal of Geophysical Research: Atmospheres* **100**, 11369–11381 (1995).

162. ESRL, *CarbonTracker-CH4 documentation* (Boulder, CO, 2013; http://www.esrl.noaa.gov/gmd/ccgg/carbontracker-ch4/documentation_obs.html#ct_doc).

163. ESRL, *NOAA/ESRL calculation of global means* (Boulder, CO, 2013; http://www.esrl.noaa.gov/gmd/ccgg/about/global_means.html).

164. P. P. Tans, T. J. Conway, T. Nakazawa, Latitudinal distribution of the sources and sinks of atmospheric carbon dioxide derived from surface observations and an atmospheric transport model, *Journal of Geophysical Research* **94**, 5151–5172 (1989).

165. L. Bruhwiler, Personal communication (2013).

166. D. Blake, *Methane, Nonmethane Hydrocarbons, Alkyl Nitrates, and Chlorinated Carbon Compounds including 3 Chlorofluorocarbons (CFC-11, CFC-12, and CFC-113) in Whole-air Samples* (2010).

167. K. R. Lassey, D. C. Lowe, M. R. Manning, The Trend in Atmospheric Methane Delta C-13 Implications for Isotopic Constraints on the Global Methane Budget, *Global Biogeochemical Cycles* **14**, 41–49 (2000).

168. J. B. Miller, in *Stable isotopes and biosphere-atmosphere interactions*, L. B. Flanagan, J. R. Ehleringer, D. E. Pataki, Eds. (Elsevier, 2005), pp. 288–310.

169. I. Levin *et al.*, No inter-hemispheric δ13CH4 trend observed, *Nature* **486**, E3–E4 (2012).

170. M. Krol *et al.*, The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, *Atmospheric Chemistry and Physics* **5**, 417–432 (2005).

171. Y. Xiao *et al.*, Global budget of ethane and regional constraints on U.S. sources, *Journal of Geophysical Research* **113**, 1–13 (2008).

172. G. Monteil *et al.*, Interpreting methane variations in the past two decades using measurements of CH4 mixing ratio and isotopic composition, *Atmospheric Chemistry and Physics* **11**, 9141–9153 (2011).

173. S. Houweling *et al.*, Atmospheric constraints on global emissions of methane from plants, *Geophysical Research Letters* **33**, L15821 ST – Atmospheric constraints on global emi (2006).

174. K. Masarie, Personal communication (2013).

175. K. M. Walter, S. A. Zimov, J. P. Chanton, D. Verbyla, F. S. Chapin III, Methane bubbling from Siberian thaw lakes as a positive feedback to climate warming, *Nature* **443**, 71–75 (2006).

176. N. Shakhova, I. Semiletov, A. Salyuk, V. Yusupov, D. Kosmach, Extensive methane venting to the atmosphere from sediments of the East Siberian Arctic Shelf, *Science* **327**, 1246–1250 (2010).

177. L. Bruhwiler, CarbonTracker-CH4: A new assimilation system for understanding the budget of atmospheric methane, *Unpublished work* (2013).

178. E. A. G. Schuur, B. Abbott, P. C. Network, High risk of permafrost thaw, *Nature* **480**, 32–33 (2011).

179. EPA, *EPA's Air Rules for the Oil & Natural Gas Industry* (2012; http://www.epa.gov/airquality/oilandgas/pdfs/20120417summarywellsites.pdf).

180. S. J. Smith, A. Mizrahi, Near-term climate mitigation by short-lived forcers, *Proceedings of the National Academy of Sciences* (2013), doi:10.1073/pnas.1308470110.

181. R. J. Plevin, *Review of final RFS2 analysis. Available at http://plevin.berkeley.edu/docs/Plevin-Comments-on-final-RFS2-v7.pdf* (2010).

182. K. E. Trenberth, L. Smith, The Mass of the Atmosphere: A Constraint on Global Analyses, *Journal of Climate* **18**, 864–875 (2004).

183. IPCC, Data Distribution Centre (2010) (available at http://www.ipcc-data.org/ancillary/tar-bern.txt.).

184. G. Myhre, E. Highwood, K. Shine, F, New estimates of radiative forcing due to well mixed greenhouse gases, *Geophysical Research* **24**, 2715 (1998).

185. J. Lelieveld, F. J. Dentener, W. Peters, M. C. Krol, On the role of hydroxyl radicals in the self-cleansing capacity of the troposphere, *Atmospheric Chemistry and Physics* **4**, 2337–2344 (2004).

186. R. G. Prinn, Evidence for variability of atmospheric hydroxyl radicals over the past quarter century, *Geophysical Research Letters* **32**, 2–5 (2005).

187. M. J. Prather, C. D. Holmes, J. Hsu, Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry, *Geophysical Research Letters* **39** (2012), doi:10.1029/2012GL051440.

188. V. Naik *et al.*, Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), *Atmospheric Chemistry and Physics* **13**, 5277–5298 (2013).

189. A. Voulgarakis *et al.*, Analysis of present day and future OH and methane lifetime in the ACCMIP simulations, *Atmospheric Chemistry and Physics* **13**, 2563–2587 (2013).

190. NOAA, *Map of oil platforms in Norther Gulf of Mexico* (2006; http://oceanexplorer.noaa.gov/explorations/06mexico/background/oil/media/platform_600.html).

191. M. T. Strosher, Characterization of emissions from diffusion flare systems, *Journal of the Air & Waste Management Association* **50**, 1723–1733 (2000).

192. P. Gogolek, *Experimental studies on methane emissions from associated gas flares* (Ottawa, Canada, 2012).

193. P. Gogolek, *Personal communication* (2012).

194. EPA, Global Methane International Coal Mine Methane Projects Database (2012) (available at http://www2.ergweb.com/cmm/projects/ProjectFindResultsAll.aspx?mode=new).

195. EPA, *Global Anthropogenic Non-CO2 Greenhouse Gas Emissions: 1990 - 2030* (Washington, D.C., 2012).

196. IEA, *Coal Mine Methane in China: A Budding Asset with the Potential to Bloom* (Paris, France, 2009).

197. G. B. Stracher, T. P. Taylor, Coal fires burning out of control around the world: thermodynamic recipe for environmental catastrophe, *International Journal of Coal Geology* **59**, 7–17 (2004).

198. WEC, *Sustainable Global Energy Development: The Case of Coal* (2004).

199. Euracoal, *Coal industry across Europe 2011* (2011).

200. IEA, *Medium Term Coal Market Report* (2011).

201. VDKi, *Annual Report 2011* (2011).

202. Euracoal, *Personal communication* (2012; http://www.euracoal.org/pages/home.php?idpage=1).

203. WEC, *2007 Survey of Energy Resources* (2007).

204. DECC, *Historical coal data: coal production, availability and consumption 1853 to 2011* (2012; https://www.gov.uk/government/statistical-data-sets/historical-coal-data-coal-production-availability-and-consumption-1853-to-2011).

205. BP, *Statistical Review of World Energy 2013* (2013; http://www.bp.com/en/global/corporate/about-bp/statistical-review-of-world-energy-2013/downloads.html).

206. OECD, *OECD Factbook Statistics* (2013; http://www.oecd-ilibrary.org/economics/data/oecd-factbook-statistics_factbook-data-en).

207. OECD, *Energy Statistics of OECD Countries* (2013; http://www.lexisnexis.com/statuniv; http://www.oecd-ilibrary.org/energy/energy-statistics-of-oecd-countries_19962827-en).

208. FERC, Interstate Pipelines under the Natural Gas Act (2013) (available at http://www.ferc.gov/industries/gas/gen-info/reg-ent.asp).

209. Air Liquide, *Gas Encyclopedia* (2013; http://encyclopedia.airliquide.com/Encyclopedia.asp?GasID=41#GeneralData).

210. Engineering ToolBox, *Fuel Gases - Heating Values* (2013; http://www.engineeringtoolbox.com/heating-values-fuel-gases-d_823.html).

Appendices

A. Additional data for Chapter 2

Contents

Life cycle assessment

Life cycle assessment is an evolving tool, originally developed in the soda-manufacturing industry for comparing the environmental burdens that glass and plastic containers represent over their lifetime (*14*). This is sometimes referred to as “cradle-to-grave” or WTW analysis. The tool is now often used to track energy use and environmental impacts of different products and processes.

Modeling assumptions

Corn ethanol WTW and LUC emissions

Land use change causes GHG emissions by changing the carbon stock of the land where the biofuel feedstock is cultivated (*21*). For instance, clearing a forest for agriculture may cause GHG emissions from burning biomass during clearing and disrupting the soil during cultivation, which can lead to decay of below-ground biomass. Furthermore, LUC can prevent CO_2 uptake through foregone accumulation of carbon over years or decades (the feedstock production period) (*21*). Since 2008, LUC emissions induced by biofuel production have been studied widely (*7*, *20*, *21*, *46*, *47*). Biofuels may induce LUC when feedstock production displaces other crops (e.g., soybean) or biomass (e.g., forests). Crops or biomass may be displaced either directly by the feedstock (direct LUC) or through biofuel induced price changes of agricultural commodities (indirect LUC) (*21*). Biofuel production may cause indirect LUC domestically or internationally, reducing the desirability of biofuels (*7*, *20*, *21*, *46*, *47*). While some sources estimate direct and indirect LUC emissions separately, others do not. Throughout this paper, I only distinguish WTW emissions and LUC emissions (direct plus indirect LUC).

EPA has estimated mean LUC emissions at 28 g CO_2e/MJ per year (*7*, *20*, *21*, *46*, *47*). Despite the high uncertainty, EPA included this value in their Regulatory Impact Analysis for the Renewable Fuel Standard 2 (*7*). In 2010, EPA found the mean life cycle GHG emissions to be 74 and 93 g CO_2e/MJ for corn ethanol and gasoline, respectively (*7*). These figures, which refer to EPA’s highlighted results, are for the year 2022 including projections about technology and yield improvements. They also estimated values for 2012 and 2017, which are considerably higher (*181*).

Alternative emissions estimates, annual fuel volumes, and time series

Table A - 1 summarizes literature values for corn ethanol and gasoline GHG emissions used to establish the annualized emissions and the time profile and to calculate climate change impacts. The emissions are divided into LUC (ethanol only) and WTW, i.e., the remaining life cycle emissions, for corn ethanol and gasoline.

Table A - 1: Summary of GHG emissions and emission factors for gasoline and corn ethanol from the literature: EPA (*7*), Tyner (*47*), and Hertel (*46*).

	Units	Sources		
		EPA, 2010	Tyner, 2010	Hertel, 2010
Gasoline				
WTW [a]	g CO_2e/MJ	93	92	92
Corn ethanol				
Total LUC [b,c]	g CO_2e/MJ	840	502	950
Annualized LUC [d]		28	17	32
WTW [a]		46	64	64
Total		74	81	96
GHG balance corn ethanol relative to gasoline	%	-20	-12	+4
Economic model used for LUC	-	FAPRI-CARD, FAMOS	GTAP	GTAP
Increase in annual fuel production causing the LUC	Bgal/yr	12.3-15.0	1.8-15.0	1.8-15.0
Time period of LUC [e]	years	12	14	14

[a] WTW CO_2e emissions are disaggregated into CO_2, CH_4, and N_2O using EPA RFS2 (*7*) data as described below.

[b] Values exclude carbon sequestration (above- or belowground) after termination of biofuel feedstock production. However, Hertel *et al.* (*46*) calculate a scenario in which 10% of all lost carbon is sequestered within 30 years after feedstock production through reforestation reducing LUC by 80 g CO_2e/MJ.

[c] EPA (*7*) calculates LUC emissions as CO_2, CH_4, and N_2O, and the total LUC emissions are given in g CO_2e/MJ. All other sources assume all lost carbon is converted to CO_2. Disaggregation in the three GHGs was performed as described below. The disaggregated values were converted to g CO_2e/MJ in the table for comparison.

[d] In all sources, LCA allocates LUC emissions evenly over the ethanol volume that is produced in 30 years, the biofuel project period. This time period is used to determine the marginal LUC emissions per unit of ethanol.

[e] Time period in which the physical LUC is estimated to take place. This is necessary for modeling the LUC over time instead of assuming all LUC occurring in the same year.

A stochastic analysis of LUC emissions by Plevin *et al*. (*20*) provides the lower and upper bounds of LUC emissions values for the RF model. Table A - 2 summarizes the 95% confidence interval as well as the minimum and maximum LUC values.

Table A - 2: Summary of upper and lower bounds of LUC emissions (g CO_2e/MJ) using stochastic data from Plevin *et al.* (*20*). Values pertain to a 30-yr life cycle period and uniformly distributed model parameters (e.g., emissions per ha converted land).

		Lower bound	Upper bound
95% confidence interval	Total LUC	945	2130
	Annualized LUC	32	71
Minimum and maximum	Total LUC	450	5,100
	Annualized LUC	15	170

For use in the RF model, WTW CO_2e emissions were disaggregated into CO_2, CH_4, and N_2O using EPA data (*7*). Disaggregation is critical to capture the actual impacts because the GWPs of CH_4 and N_2O are one to two orders of magnitude higher than CO_2. Well-to-wheel mass fractions for CO_2, CH_4, and N_2O are 94.1%, 0.201%, and 0.00143% of CO_2e emissions, respectively, for ethanol, and 96.0%, 0.115%, and 0.00377% for gasoline (*7*). Multiplying all fractions by the respective 100-year GWPs adds up to 100%.

EPA analysis (*7*) provides LUC emissions individually for all three GHGs. In contrast, the other sources assume that all C is converted in CO_2. Carbon dioxide equivalent emissions were disaggregated following the methodology in EPA (*7*), which states that CH_4 and N_2O are small fractions of the LUC CO_2 emissions, and the fractions vary by the converted land type. The averages of these mass fractions over all land types, weighted by the LUC occurring in each country and land type, are 0.21% and 0.011% for CH_4 and N_2O, respectively.

The time step in this analysis is one year. The ethanol emissions of a single year include all WTW emissions. Ethanol is produced over the 30-year project time frame during which agricultural land is dedicated for feedstock production. This value is often used in LCA including EPA analysis (*7*). Additionally, there are four sources of LUC emissions whose time pattern assumptions were adopted from the BTIME model (*15*). In the first year of LUC, all standing (above-ground) biomass is burned. Soil carbon (below-ground) is released over the first 20 years since land clearing. Eighty percent of these emissions occur in the first 5 years, the rest by year 20. Foregone sequestration is accounted for as a carbon debt because the displaced biomass can no longer accumulate carbon as long as corn for ethanol is produced (see Table A - 3). The end of the 30-year period concludes the analysis of the base case.

Table A - 3: Parameters of the BTIME model for allocating LUC emissions over time. Source: (*15*).

LUC-Parameter	Value	Unit
Forest fraction	80	% of LUC emissions
Pasture fraction	20	% of LUC emissions
Above-ground emissions (immediate loss)		
Forest	77.5	% of Forest LUC
Pasture	14.3	% of Pasture LUC
Below-ground emissions		
Forest	22.5	% of Forest LUC

LUC-Parameter	Value	Unit
Pasture	85.7	% of Pasture LUC
Early loss	80	% of Below-ground emissions
End of early loss period	5	Years after LUC occurrence
End of later loss period	15	Years after early loss period
Foregone sequestration		
Annual emissions during production period	4.7	g CO_2/MJ/yr

The time profile of ethanol emissions with annualized LUC is also presented in Figure 1. The annual values for gasoline emissions represent the entire life cycle emissions per unit of gasoline. These values are constant over all years because LUC due to gasoline production is assumed negligible. In order to fairly compare the temporal effects of gasoline and ethanol production (including LUC), gasoline production is modeled over 30 years as well. Ethanol emissions with annualized LUC are also constant over time, but lower than gasoline due to the lower overall life cycle emissions.

Total WTW and LUC emissions in a given year were calculated separately before adding to get the total GHG emissions over time. Total WTW emissions per year are the product of the WTW portion of the time profile (g/MJ), annual fuel use (gal), and energy content of the fuel (88.9 and 132 MJ/gal for ethanol and gasoline, respectively). Total LUC emissions at a given year were calculated as the product of the LUC portion of the emissions time profile (g/MJ) and the incremental ethanol production of the same year (MJ), which triggers the LUC. Table A - 3 summarizes the assumptions of the BTIME model going into establishing the emissions time profile.

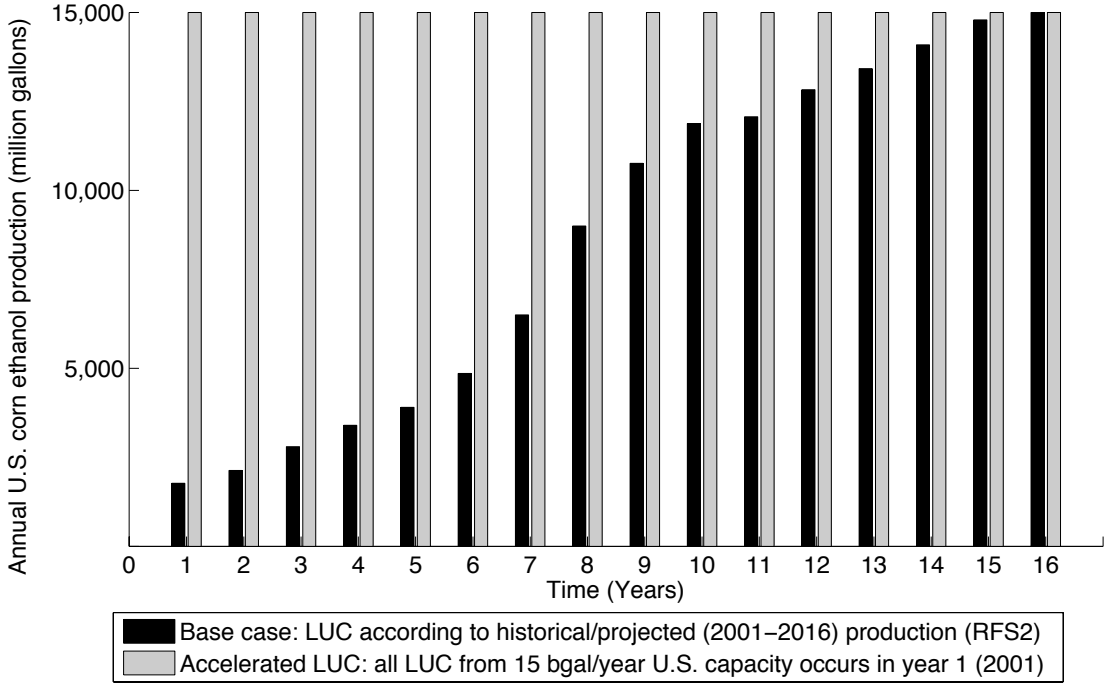

Figure A - 1: LUC scenarios according to annual corn ethanol production considered in this analysis.

Figure A - 1 depicts the annual ethanol volumes considered in the base case and the accelerated LUC scenario. Using annual increases in corn ethanol production as an indicator for LUC occurrence over time (particularly regarding ILUC) is a simplification. However, it is very unlikely that LUC (even ILUC) will occur *before* the associated increase in corn ethanol production in each year. Thus, the method should be regarded as allocating LUC at the earliest possible points in time. LUC could in reality be shifted more towards the future, thereby reducing the share of upfront emissions. As a result, ETF should be regarded as a conservative estimate (lower ETF possible).

Accounting for emissions timing in the LCA literature

Several authors have estimated that the GHG balance undervalues the climate impacts of corn ethanol by 9-89% due to emissions timing (*15*–*17*). In order to correct the underestimation, O'Hare *et al.* (*15*) calculate the ratio of CRF of corn ethanol and gasoline GHG emissions. This ratio is then multiplied by the corn ethanol GHG balance. Moreover, the authors examined the use of a discount rate, which weighs emissions today greater than those in the future (*15*). Depending on the chosen time preference, i.e., discount rate, life cycle emissions of corn ethanol may be higher or lower than gasoline. For instance, counting emissions over 100 years, corn ethanol emissions may be 5% lower than gasoline, but when discounted at 7% corn ethanol emissions would be 16% greater (*15*). Discounting of emissions in the LCA of RFS2 was performed using 0% and 2% discount rates. However, the final results ignored the 2% scenario, partly due to the "lack of consensus in the scientific community about the best way to translate GHG emissions into a proxy for economic damages" (*7*). In fact, choosing the 2% discount rate would have resulted in corn ethanol not meeting the "Renewable Fuel" requirement of EISA for facilities built after December 2007 (*7*). In turn, this may have interfered with meeting the EISA biofuel production targets.

Kendall *et al.* estimate a penalty factor for LUC emissions (*16*). The factor is calculated as the increase in CRF of a LUC emission impulse today relative to annualizing LUC over 30 years. Radiative forcing (RF) is a measure of the imbalance between the earth's incoming solar radiation, and its reflection back into space (*48*). GHGs contribute to positive RF (and increasing surface temperature) by absorbing some of the radiation and reflecting it back into the atmosphere (*48*). The penalty factor is multiplied by the LUC emissions, which increases the life cycle emissions of corn ethanol by 46%, and its GHG balance by 89%.

Levasseur *et al.* describe a method to compare the performance of a biofuel and the FF incumbent based on climate change impacts of a marginal increase in GHG concentration over time, in RF per MJ of fuel (*17*). The RF impacts are simulated using GHG emissions as they occur over time, including emissions from LUC in the beginning of the life cycle. The authors find that the GHG balance is 9-19% lower than the CRF using the actual emissions time profile. They conclude that performance should be assessed based on impacts over time, e.g., in CRF, as opposed to GHG emissions. However, they also note that the simplified RF model may not accurately represent actual impacts, which is an important potential limitation of the study.

Radiative forcing model

The following equations are used in Eq. 1. The elements in the emissions matrix $E_{rt,\,Fuel,\,GHG}$ are as in Eq. A - 1. Emissions are simulated for fuel use in the release years r from 1 to 30 (the end of the life cycle) in the base case. There will be some ethanol emissions after 30 years due to the final increment of LUC (from reaching full capacity of 15 Bgal/yr) that causes foregone sequestration beyond the model's 30-year time horizon of fuel production. Gasoline emissions will cease for r >30.

$$E_{rt,Fuel,GHG} = \begin{bmatrix} e(r=1) & e(r=1) & \dots & e(r=1) \\ 0 & e(r=2) & \dots & e(r=2) \\ : & : & & : \\ 0 & 0 & \dots & e(r=100) \end{bmatrix} \qquad \textbf{Eq. A - 1,}$$

where $e(r)$ is the amount of the total emissions in the release year r described in section 2.2.1. The GHG emissions are converted into changes in atmospheric concentrations using Eq. A - 2.

$$\Delta C = \frac{10^6\ ppmv \,/\, M_{GHG}}{m_{air} \,/\, M_{air}} \qquad \textbf{Eq. A - 2,}$$

where M_{GHG} is the molar mass of the GHG, which is 44, 16, and 44 g/mol for CO_2, CH_4, and N_2O, respectively. The constants m_{air} ($5.14*10^{21}$ g) and M_{air} (28.9 g/mol) are the total mass and the molar mass of air in the atmosphere, respectively (*182*). The elements in the radiative efficiency matrix $RE_{rt,\,GHG}$ for each GHG are as in Eq. A - 3.

$$RE_{rt,GHG} = \begin{bmatrix} re(t=1) & re(t=2) & \dots & re(t=100) \\ re(t=1) & re(t=2) & \dots & re(t=100) \\ : & : & & : \\ re(t=1) & re(t=2) & \dots & re(t=100) \end{bmatrix} \qquad \textbf{Eq. A - 3,}$$

where $re(t)$ is the radiative efficiency (in $W/m^2/ppmv$) of a GHG in the impact year t. The radiative efficiencies of each GHG are given in Eq. A - 4, Eq. A - 5, and Eq. A - 6. The following three scenarios for future atmospheric background CO_2 concentration BC_{CO2} were obtained from the IPCC Data Distribution Centre (*183*). The A1B scenario, representing moderate increases in CO_2 concentration, is used as the base case, and B1 and A1FI are considered as low and high concentration scenarios, respectively. The radiative efficiencies of CH_4 and N_2O were treated as constant over time as a conservative estimate (Eq. A - 5 and Eq. A - 6), which deviates from IPCC methods (*48*). This simplification has little impact on the results because both GHGs represent only small portions of the overall RF in the model.

For CO_2, *re(t)* decreases with increasing future atmospheric CO_2 concentration according to Eq. A - 6 (*184*).

$$re_{CO_2}(t) = \frac{5.35}{BC_{CO_2}(t)} \quad \textbf{Eq. A - 4}$$

$$re_{CH_4} = 0.37 \quad \textbf{Eq. A - 5}$$

$$re_{N_2O} = 3.03 \quad \textbf{Eq. A - 6}$$

The elements in the decay matrix $D_{rt,\,GHG}$ for each GHG are as in Eq. A - 7.

$$D_{rt,GHG} = \begin{bmatrix} d(t=1) & d(t=2) & \dots & d(t=100) \\ 0 & d(t=1) & \dots & d(t=99) \\ : & : & & : \\ 0 & 0 & \dots & d(t=1) \end{bmatrix} \quad \textbf{Eq. A - 7,}$$

The decay rates in Eq. A - 8, Eq. A - 9, and Eq. A - 10 assume a background concentration of 378 ppmv (*184*).

$$d_{CO_2}(t) = 0.217 + 0.259 * e^{-(t-1)/172.9} + 0.338 * e^{-(t-1)/18.51} + 0.186 * e^{-(t-1)/1.186} \quad \textbf{Eq. A - 8,}$$

$$d_{CH_4}(t) = e^{-(t-1)/11.3} \quad \textbf{Eq. A - 9,}$$

$$d_{N_2O}(t) = e^{-(t-1)/114} \quad \textbf{Eq. A - 10}$$

Modeling results using alternative data and scenarios

Figure A - 2 shows RF time series individually for gasoline and corn ethanol. In contrast to the base case, fuel production is simulated over 50 years. As a result, the positive spread of absolute RF between gasoline and corn ethanol is significantly larger (see Chapter 2).

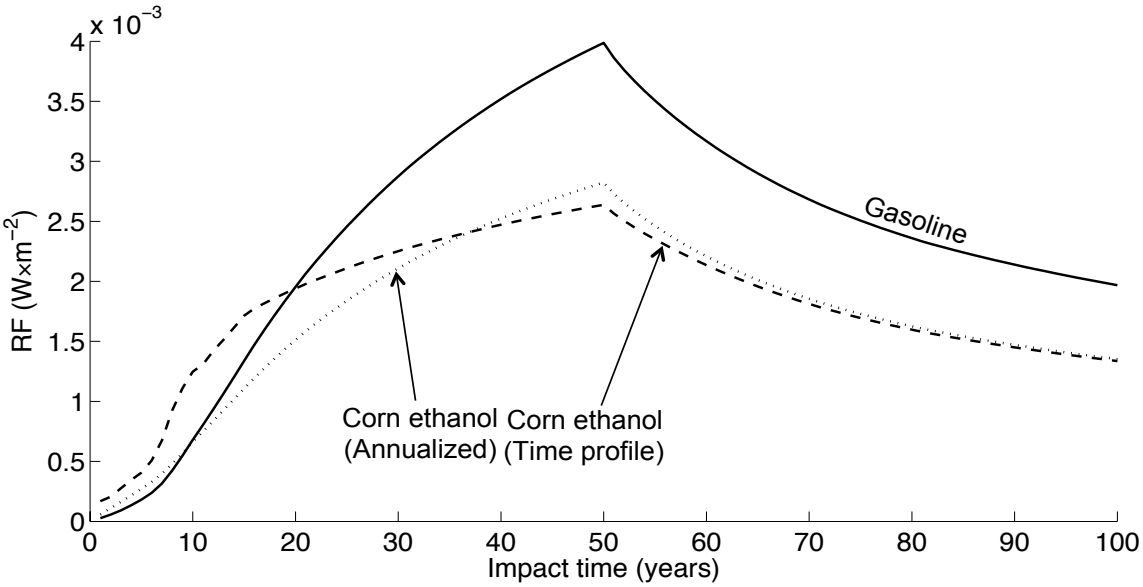

Figure A - 2: RF(t) of gasoline and corn ethanol (both annualized emissions and emissions time profile) for 50 years of production (EPA (*7*) emissions estimates).

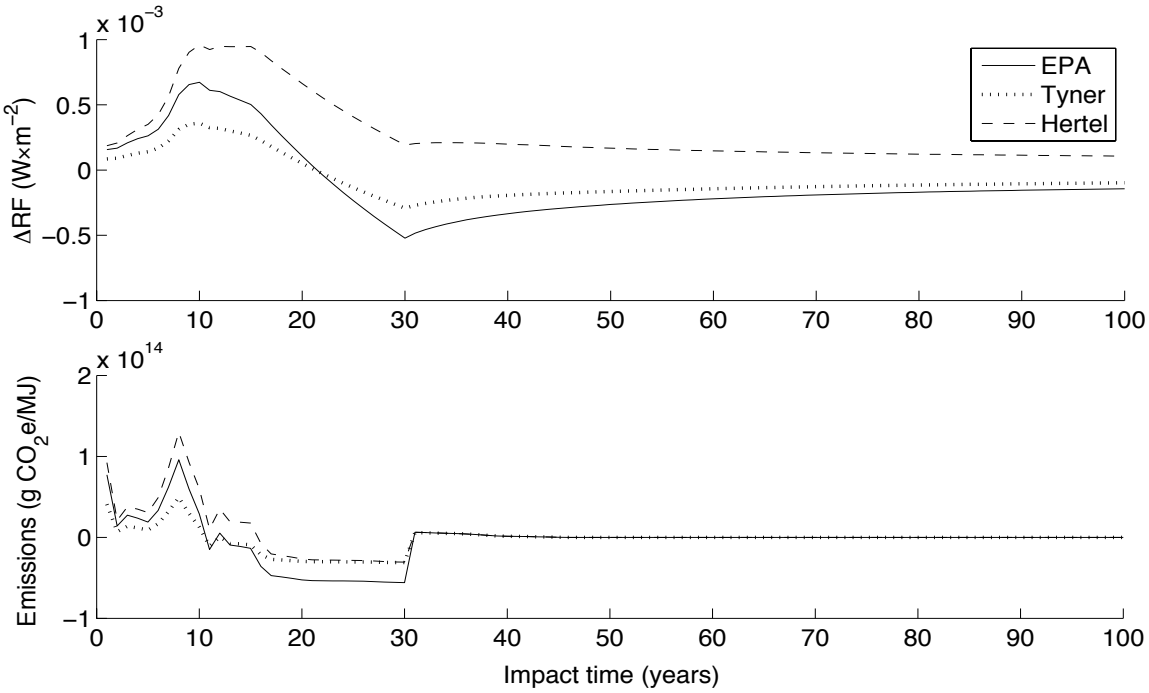

Figure A - 3. ΔRF(t) and GHG emissions (corn ethanol minus gasoline, 15 Bgal/yr capacity, EPA (*7*) emissions estimates).

Figure A - 3 depicts ΔRF (corn ethanol minus gasoline) and GHG emissions over the impact time frame, which shows the time lag between changes in emissions and RF over time due to the relatively slow decomposition of GHGs in the atmosphere. In addition to the main paper, it also shows RF using emissions estimates in Tyner *et al.* (*47*) and Hertel *et al*. (*46*).

Radiative forcing in the accelerated LUC scenario

Figure A - 4 shows RF(t) of corn ethanol and gasoline individually in the accelerated LUC scenario, which assumes all LUC from 15 Bgal/yr corn ethanol to occur in year 1. This assumption is implicit in the previous analyses (*15*–*17*). The ethanol curves are initially higher than gasoline due to LUC emissions all occurring in the first time step, but then increases at a lower rate due to lower life cycle emissions during the production and use phase modeled during the remainder of the 100 years. After 30 years, both curves decline because this analysis time frame for production ceases, and GHG decay reduces atmospheric concentration of these gases and associated RF.

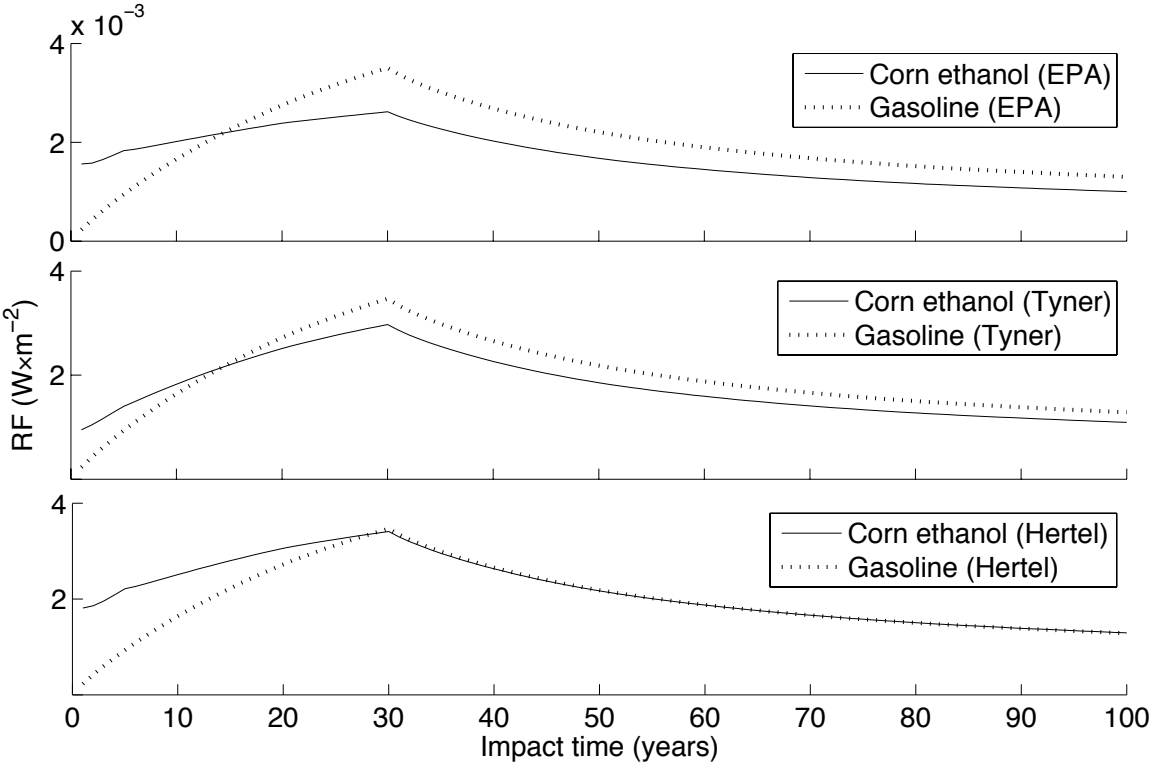

Figure A - 4: RF(t) for corn ethanol (15 Bgal/yr) and gasoline individually (using EPA, Tyner *et al.*, and Hertel *et al.* (*7*, *46*, *47*) emissions estimates, accelerated LUC scenario).

Figure A - 5 shows the CRF balance for both the base case and the accelerated LUC case. Note that the CRF balance is significantly lower in the accelerated LUC scenario compared to the base case. The accelerated LUC scenario is equivalent to previous analyses (*15*–*17*), which assume most LUC (all above-ground LUC emissions) to occur in year 1. Furthermore, note that annual fuel use in the base case increases over time (Figure A - 1), whereas the GHG balance implicitly assumes full capacity from the first year. As a result, LUC impacts are leveraged over a smaller amount of WTW emissions benefits compared to the GHG balance. Indeed, the CRF curves in the accelerated LUC scenario, which assumes full capacity in the first year, are shifted down towards the GHG balance (Figure A - 5).

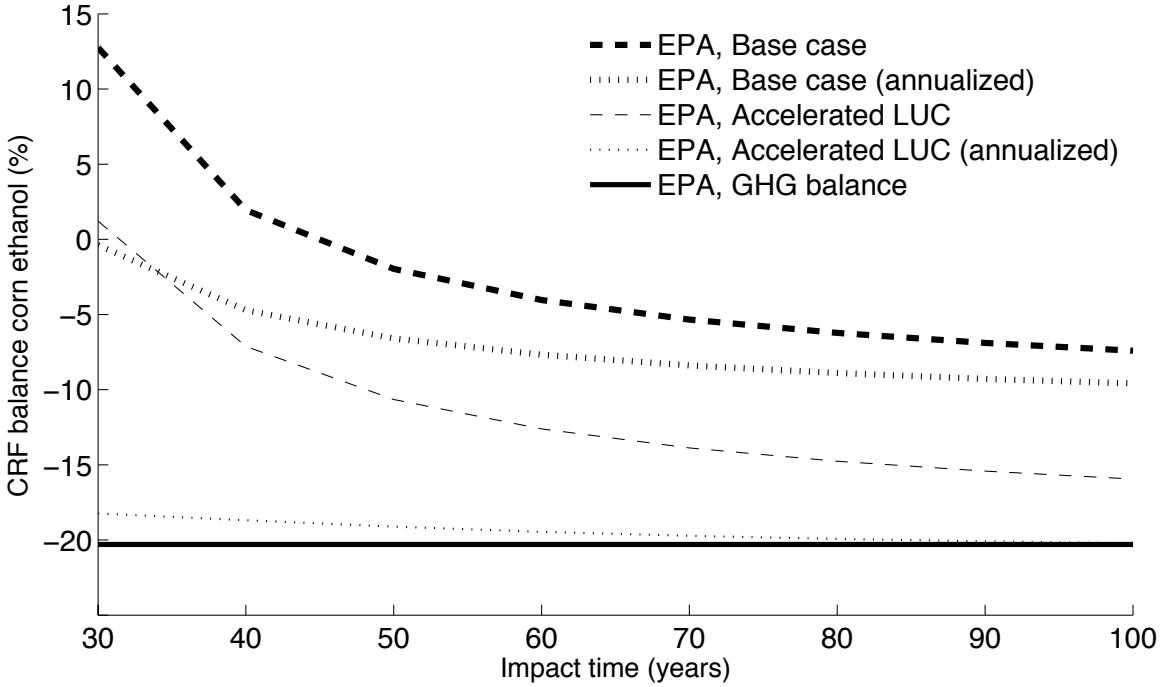

Figure A - 5: CRF balance of corn ethanol for both the base case and the accelerated LUC scenario.

The base case CRF results are incidentally similar to those in O'Hare *dissertation*. (*15*) (0% discount rate) because current (RFS2) corn ethanol WTW emissions are employed, which are 46 g CO_2/MJ, i.e., 25% lower than in O'Hare *dissertation*. (*15*) (60 g CO_2/MJ).

Sensitivity analysis

A sensitivity analysis was performed to test the robustness of the model results to a number of input parameters, which is shown in Figure A - 6. ETF changes only marginally for different background concentrations, different CH_4 and N_2O emissions, higher ethanol WTW emissions, and for shifting all LUC emissions from the emissions time profile to the first year of the life cycle (accelerated LUC). Future average gasoline emissions may increase if currently untapped fossil sources may add more GHG intensive fuel to the mix. Considering an extreme case where average gasoline emissions increase by 50%, ETF would decrease from 2-13% to 1-8%. A 50% drop in corn ethanol WTW emissions would not change ETF significantly. Lower overall LUC estimates in Tyner *et al.* (*47*) reduce ETF significantly, particularly over shorter impact time frames. Accordingly, higher LUC emissions in Hertel *et al.* (*46*) increase ETF (see also doubling of EPA LUC estimates). Figure A - 6 shows that ETF is lower when accounting for a phased-in ethanol production over time (base case) compared to production at full capacity from the first year.

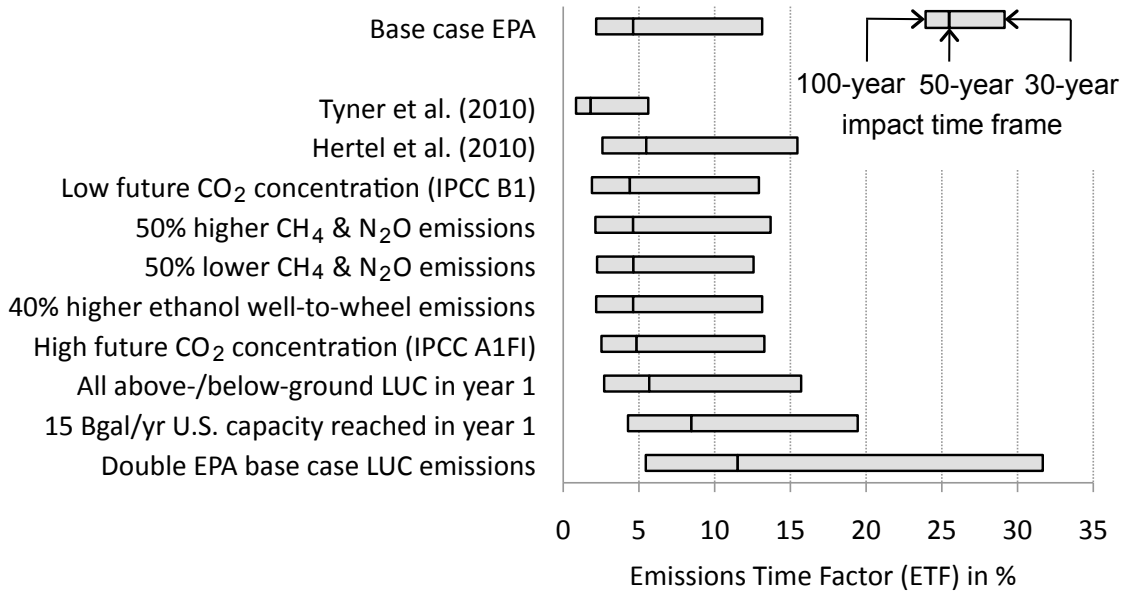

Figure A - 6: Sensitivity analysis results for ETF over 100, 50, and 30-year impact time frames. EPA (*7*) emissions estimates are used as the base case.

Figure A - 7 shows how the production period influences the CRF balance of corn ethanol with respect to uncertainty. Over a 100-year impact time frame and fuel production until 2030 (base case), the GHG balance is lower than the CRF balance, i.e., the GHG balance overestimates ethanol's benefits.

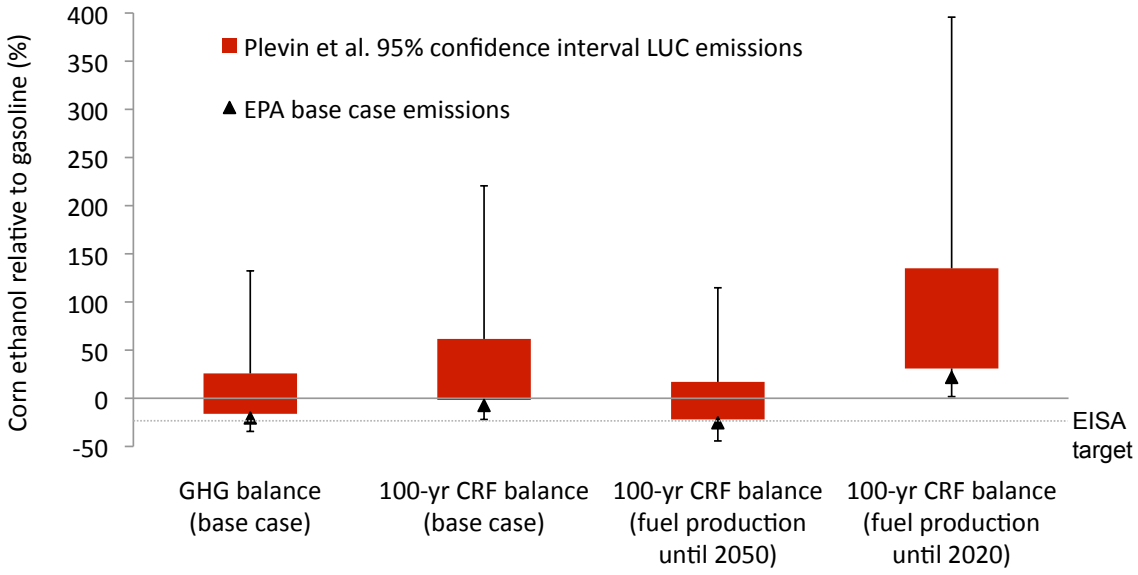

Figure A - 7: Direct comparison of the GHG balance and 100-yr CRF balance (time profile) of corn ethanol over different fuel production time frames. Only the CRF balance accounts for emissions timing. Illustrated estimates are EPA (*7*) base case and Plevin *et al.* (*20*) LUC uncertainty with 95% confidence intervals and min/max values (error bars).

However, the GHG balance underestimates ethanol's benefits if fuel production is counted until 2050. This is because the GHG balance ignores the policy cap of 15 Bgal/yr corn ethanol in 2016, and corn ethanol may not cause further LUC after maximum production is reached. As a result, corn ethanol may reduce emissions beyond 30 years of production (life cycle period), which is not captured by the GHG balance. Similarly, the GHG balance would overestimate ethanol's benefits by about 40 percentage points if corn ethanol production would cease in 2020. The model assumes that all LUC occurs until that year without further emissions benefits from lower WTW emissions. Analogously to Figure 6 in the main article, the uncertainty inherent in the LUC estimates clearly dominates the potential emissions benefits or costs of corn ethanol.

Reference points for climate impacts

Additional analysis for similar climate decisions in other industries beyond biofuels was done to understand the relative significance of ΔT_s results. Emissions estimates from U.S. electricity consumption (*15*) was used in combination with U.S. electricity demand forecasts (*16*) to estimate RF impacts of power generation over the same time frame as corn ethanol (2001-2030). Since electricity generation is the largest source of GHGs in the U.S., the results can serve as a reference of the magnitude of impacts that may be considered "high". Figure A - 8 compares total RF and ΔT_s for electricity generation, gasoline reduction, and substitution of gasoline with corn ethanol (*51–53*).

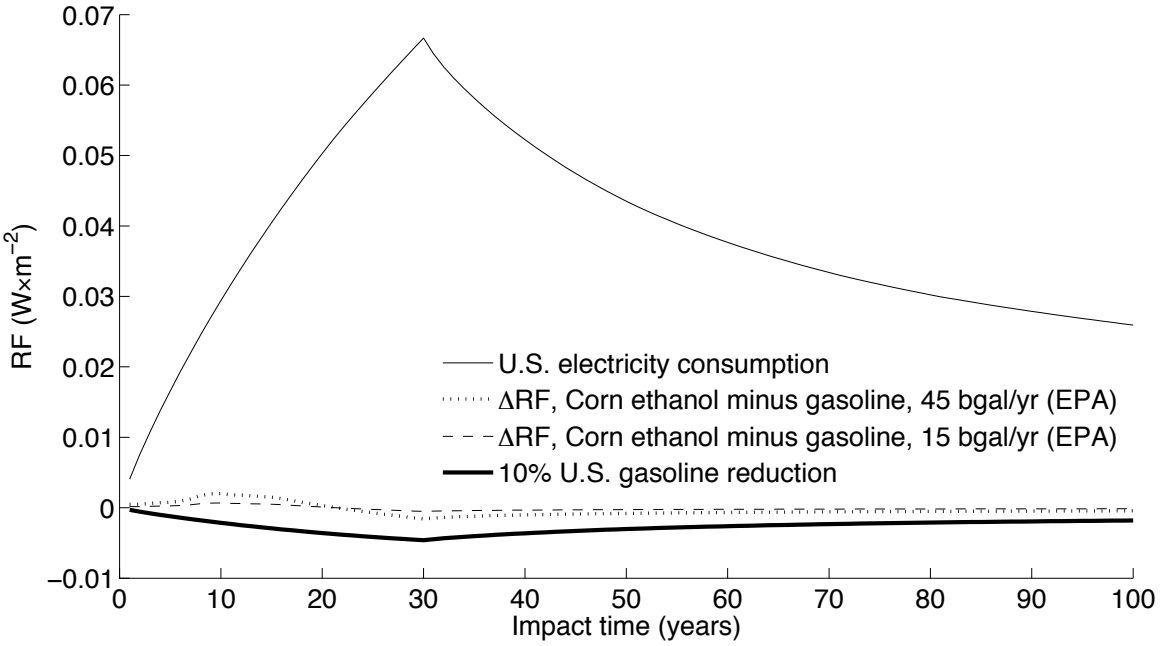

Figure A - 8: RF(t) from U.S. electricity consumption, U.S. gasoline reduction, and corn ethanol use (base case).

Maximum RF from corn ethanol is 0.0006 W/m^2. Between 1950 and 2000, global RF has increased by 1.5 W/m^2 (*48*). The magnitude of global mean surface temperature change associated with maximum RF from corn ethanol at maximum 15 Bgal/yr capacity, i.e., 7% of U.S. gasoline, is 0.0007 °C. As shown in Table A - 4, the maximum effect of this corn ethanol volume on RF and ΔT_s is less than 1% that of U.S. electricity consumption. Reducing U.S. gasoline consumption by 10% decreases RF and ΔT_s by eight times the amount that RF and ΔT_s increases from 15 Bgal/yr corn ethanol.

Table A - 4: Maximum RF and ΔT_s (at 4.5°C climate sensitivity) resulting from GHG emissions due to U.S. electricity generation and reduction in gasoline use as reference points for substitution of gasoline with corn ethanol (all 2001-2030).

		Reference points [b]	
	U.S. corn ethanol minus gasoline [a]	**U.S. Electricity consumption**	**10% U.S. gasoline reduction**
RF (W/m^2)	0.0006	0.067	-0.0046
ΔT_s (°C) [c]	0.0007	0.081	-0.0056

[a] EPA base case (15 Bgal/yr max. capacity). This maximum difference in RF between corn ethanol and gasoline occurs in year 10 (see also Figure 3 in Chapter 2).

[b] Historical (*51*) and projected (*52*) CO_2 emissions from electricity. Gasoline consumption based on 2009 (*53*). Shown maximum RF occurs in year 30 (end of the simulated impact period).

[c] ΔT_s values should be considered as an order of magnitude estimate.

Corn ethanol's probability of meeting the Renewable Fuel target in the Energy Independence and Security Act of 2007

Adding ETF to the GHG balance could reduce the desirability of corn ethanol in terms of GHG emissions. Moreover, the Energy Independence and Security Act of 2007 (EISA) requires mandated "Renewable Fuels" to reduce GHG emissions by at least 20% relative to gasoline. Given EPA's life cycle assessment (LCA) (and assuming that the mean and median value of the GHG balance are identical), there is a 50% chance that corn ethanol meets this target (*7*). Adding ETF may reduce this probability even further. The objective of the final part of this analysis is to quantify how ETF influences the probability of corn ethanol meeting the EISA target.

Data and methods

Data from EPA's Regulatory Impact Analysis for the RFS2 is used (*7*). Figure A - 9 shows the probability distribution function (PDF) for corn ethanol's GHG balance, which was generated by EPA using Monte Carlo analysis. It is approximately normally distributed with a mean of -20.4%.

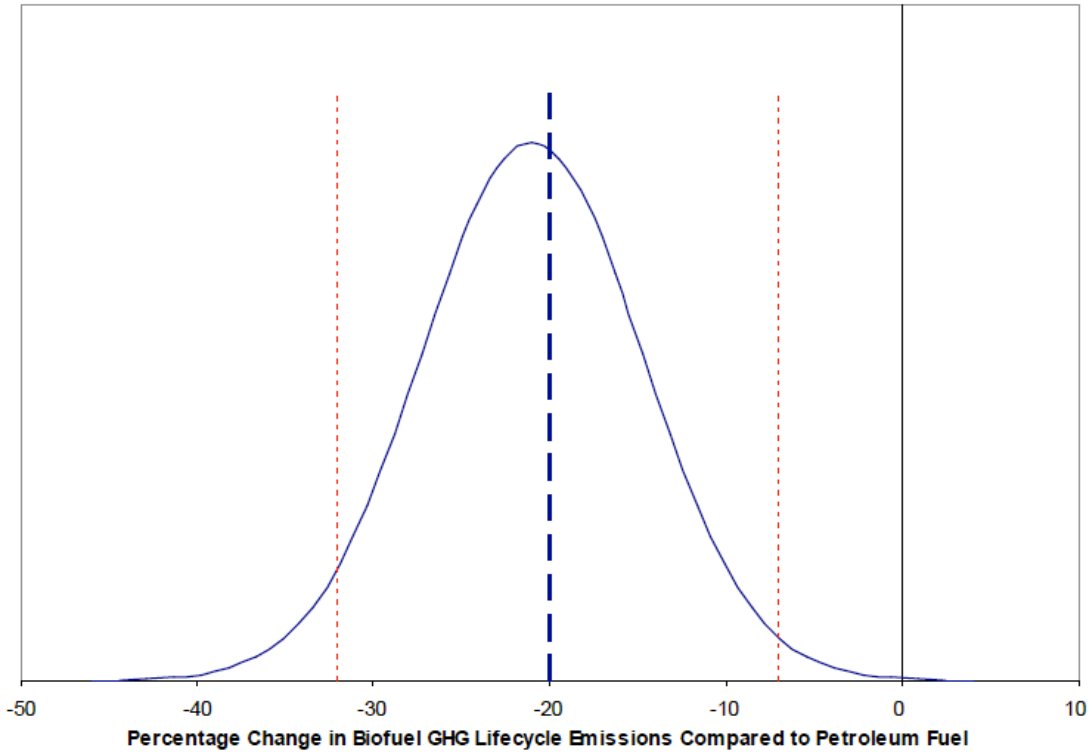

Figure A - 9: EPA result for the GHG balance of corn ethanol (graph from (*7*)). The -20% line represents the EISA target. Ethanol emissions are shown relative to gasoline (mean 92.9 g CO_2e/MJ).

The RFS2 also provides estimates for the breakdown between LUC and the remaining WTW emissions, which is summarized in Table A - 5.

Table A - 5: Distribution statistics (adopted from (7)) of corn ethanol GHG emissions (n/a indicates no data is available).

Emissions source	[g CO_2e/MJ]		
	2.5th percentile	**Mean**	**97.5th percentile**
LUC	19.9	30.3	43.6
WTW	n/a	44.5	n/a
Total	51.2	74.8	91.9

GHG balance	[%]		
(Gasoline 92.9 g CO_2e/MJ)	-32	-20	-7

Adding ETF to the distribution should skew the original distribution to the right, i.e., to higher GHG balance values. ETF is only applied to corn ethanol emissions (keeping gasoline constant) because gasoline emissions are assumed to be constant over time (no LUC). First, a Monte Carlo simulation is performed in Analytica to generate estimates of the original distributions (Figure A - 9 and Table A - 5). The LUC emissions distribution was generated using a Lognormal distribution Log(median=29.7, geometric stdev=1.22) and a sample size of 10,000. The median and geometric standard deviations were chosen to achieve the best possible fit with the estimates in Table A - 5.

A similar process was followed to generate an approximate distribution for the WTW emissions. Since no percentile values were available for WTW emissions, a normal distribution with Normal(44, 2.5) was chosen (again a sample size of 10,000). This distribution achieved the best fit for the total corn ethanol emissions distribution. The Analytica model was designed such that LUC and WTW distributions are added to yield the total corn ethanol emissions. Then, the GHG balance associated with each pair of LUC and WTW emissions is calculated using Eq. A - 11.

$$GHG\ balance = \frac{Emissions_{EtOH,LUC} + Emissions_{EtOH,WTW} - Emissions_{Gasoline}}{Emissions_{Gasoline}} \quad \textbf{Eq. A - 11,}$$

where $Emissions_{Gasoline}$ equals 92.9 g CO_2e/MJ. This yields the distribution of the GHG balance of corn ethanol. Next, the MATLAB climate model is run for random samples of LUC and WTW emissions to calculate ETF values. This is important because ETF increases with increasing LUC emissions. A regression analysis is performed to estimate ETF as a function of LUC emissions. The ETF independent of WTW emissions for the small ranges of WTW emissions considered here. Finally, the Analytica model is designed to add ETF to the GHG balance to account for the emissions timing effect that

is being investigated. This yields new distributions of the emissions timing corrected GHG balance - one for each impact time horizon. Analytica uses Eq. A - 12 to calculate the corrected GHG balance for each data point in the simulation.

$$GHG\ balance_{corrected} = \frac{Emissions_{EtOH,LUC} + Emissions_{EtOH,WTW} - Emissions_{Gasoline}}{Emissions_{Gasoline}} + ETF \quad \textbf{Eq. A - 12,}$$

where ETF is regressed as a function of LUC emissions (see above). The new, corrected distributions are then used to calculate the probabilities of corn ethanol (i) meeting the EISA target of -20% emissions reductions and (ii) emitting equal or less than gasoline. This is accomplished by calculating the percentile values of each observation based on the simulated estimates. This means first adding the number of observations (e.g., for -20% GHG balance and below) and then dividing by the sample size.

Results

Reproduction of EPA's PDF of corn ethanol's GHG balance

The PDF of corn ethanol's GHG balance in EPA's RFS2 was reproduced because only the graph in Figure A - 9 was available. The simulation result for LUC emissions is illustrated in Figure A - 10 (all Analytica simulation figures contain 600 data points). The mean, 2.5th and 97.5th percentiles are 30.4, 20.1 and 43.9 g CO_2e/MJ, respectively, which is a reasonable approximation to the original data (Table A - 5). The result of the WTW distribution is illustrated in Figure A - 11. The mean, 2.5th and 97.5th percentiles are 44.0, 39.1 and 48.9 g CO_2e/MJ, respectively.

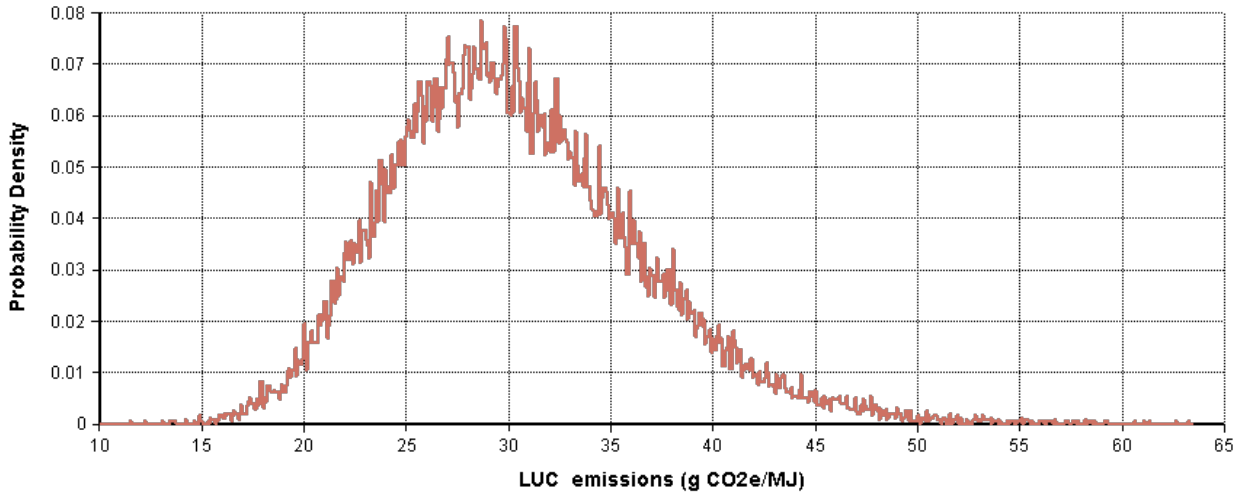

Figure A - 10: PDF of corn ethanol LUC emissions using Analytica (sample size 10,000) for a Lognormal distribution with Log(median=29.7, geometric stdev=1.22).

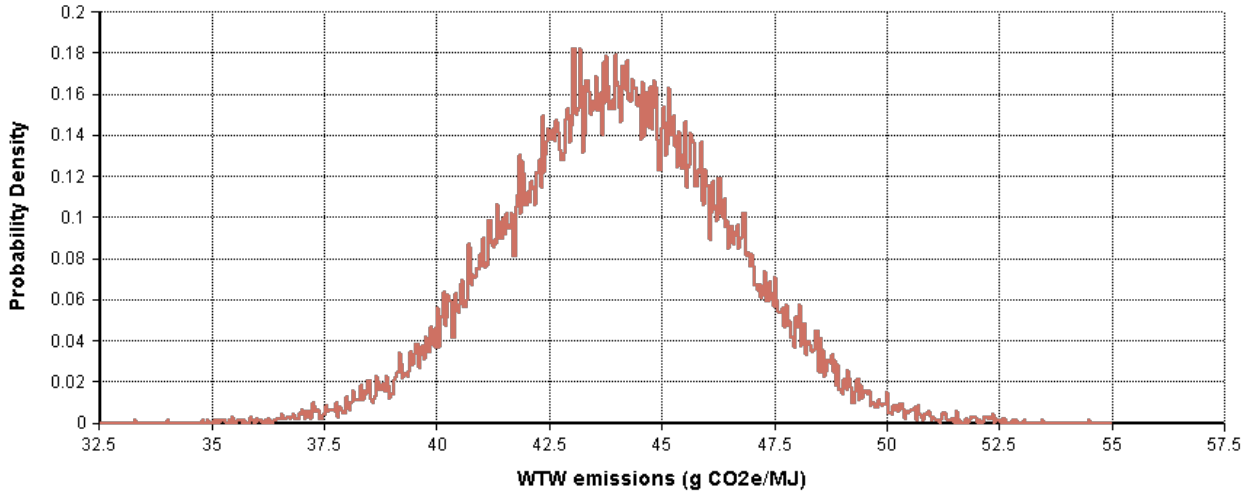

Figure A - 11: PDF of corn ethanol WTW emissions using Analytica (sample size 10,000) for a Normal distribution with N(44, 2.5).

Running the Monte Carlo simulation for both LUC and WTW distributions results in the total corn ethanol emissions distribution shown in Figure A - 12. The distribution shows the GHG balance on the abscissa, which was converted from the emissions value using Eq. A - 11. This PDF is the model for the curve in Figure A - 9. The mean, 2.5th, and 97.5th percentile values of the (here) simulated data are -20.0%, -32.3%, and -4.7%, respectively. The 97.5th percentile value is off by 2.3 percent points relative to the original data in Table A - 5.

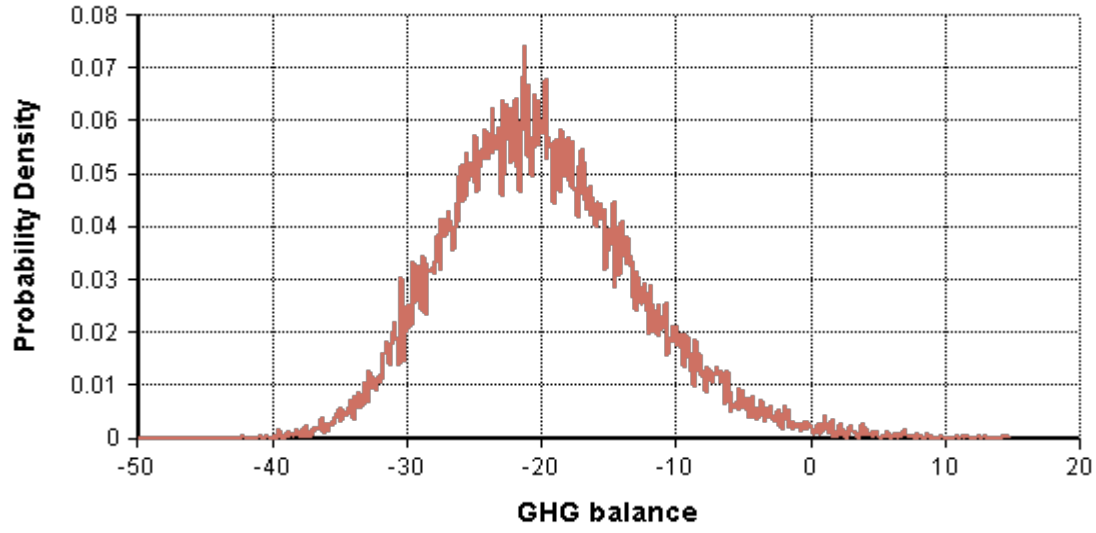

Figure A - 12: PDF of the corn ethanol GHG balance (in %) from adding LUC and WTW emissions.

MATLAB climate model results

Figure A - 13 shows ETF as a function of the LUC emissions. The ETF is mainly a function of the LUC value, and ETF increases with higher LUC. The shorter the impact

time horizon, the greater the relative influence of LUC emissions on overall emissions, and the higher rate of increasing ETF. The ETF was initially calculated for 6 random LUC values which showed ETF to be a linear function of the GHG balance. A regression analysis was performed in Excel to test the significance of the results, and to extrapolate ETF for higher GHG balances. For instance, the 6 values of the 100-year ETF have an R^2 of 1, and P-values of 10^{-9} and 10^{-11} for intercept and slope, respectively (i.e., very high confidence in regression result). After extrapolation, ETF was calculated again for higher GHG balances to check the linearity of the climate model (with respect to LUC emissions values), which confirmed the extrapolated values.

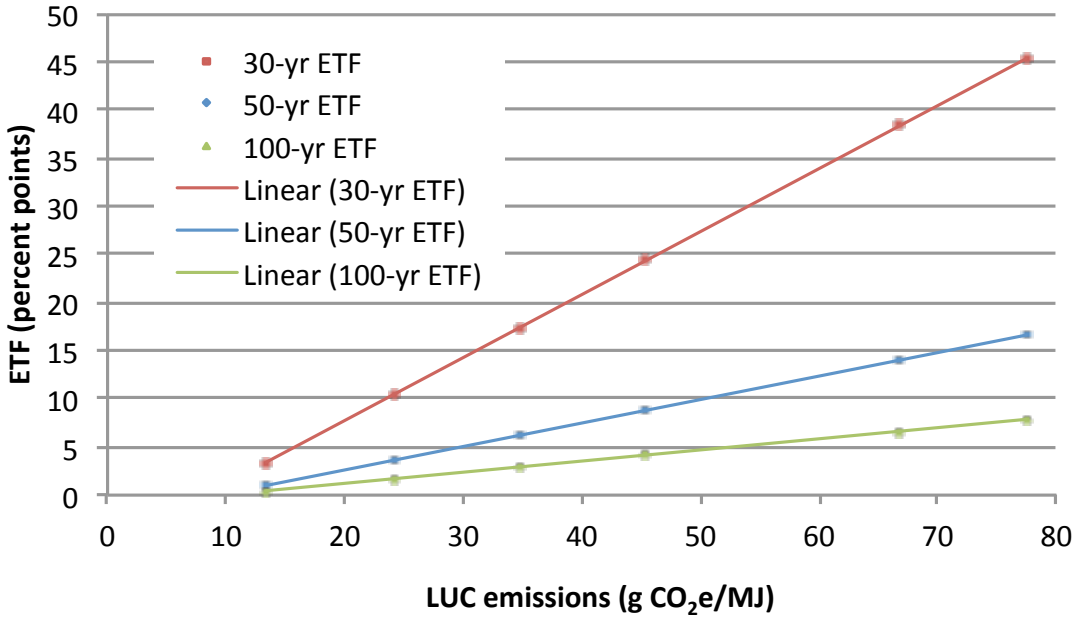

Figure A - 13: MATLAB climate model results of ETF for 30, 50, and 100-year impact time frames.

Figure A - 14 shows the influence of ETF on the GHG balance as a function of the impact time frame. The upper line represents ETF for bin values that generate a GHG balance of 0%. For very high impact time frames, the GHG balance including ETF will approach 0%, but for short impact time frames such as 30 years ETF increases the GHG balance to 25%. The lower line represents ETF for bin values that generate a GHG balance of -20%.

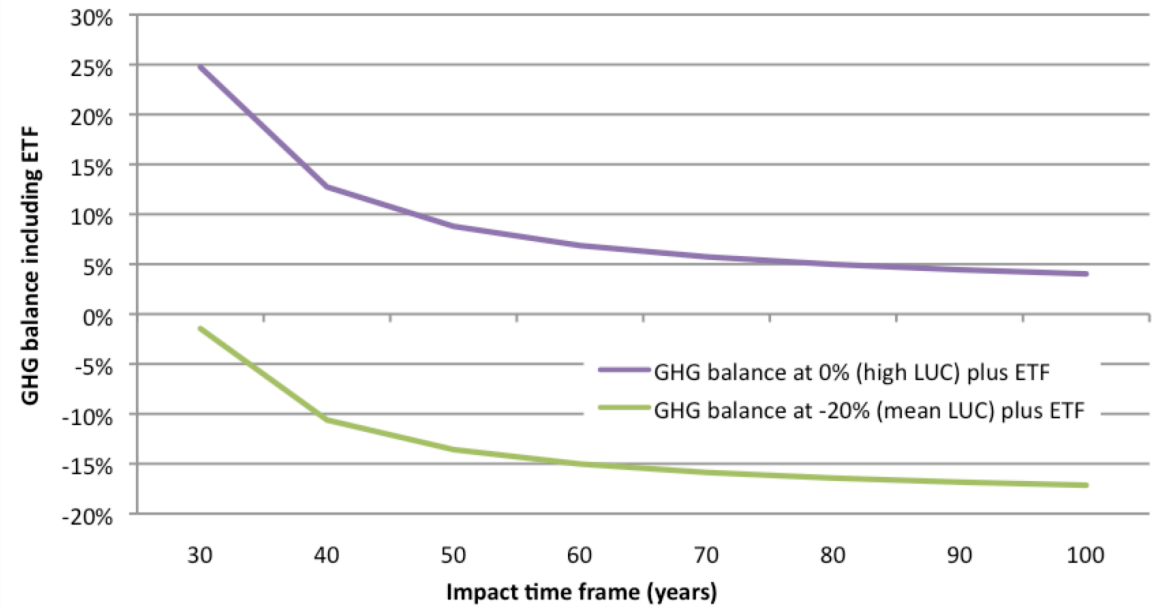

Figure A - 14: The influence of ETF on the GHG balance for different LUC values and impact time frames.

GHG balance distributions corrected with ETF

Figure A - 15 and Figure A - 16 show how the GHG balance distribution changes as ETF (for different impact time frames) is added to correct for emissions timing. The GHG balance distribution skews increasingly towards the right for shorter ETF impact time frames.

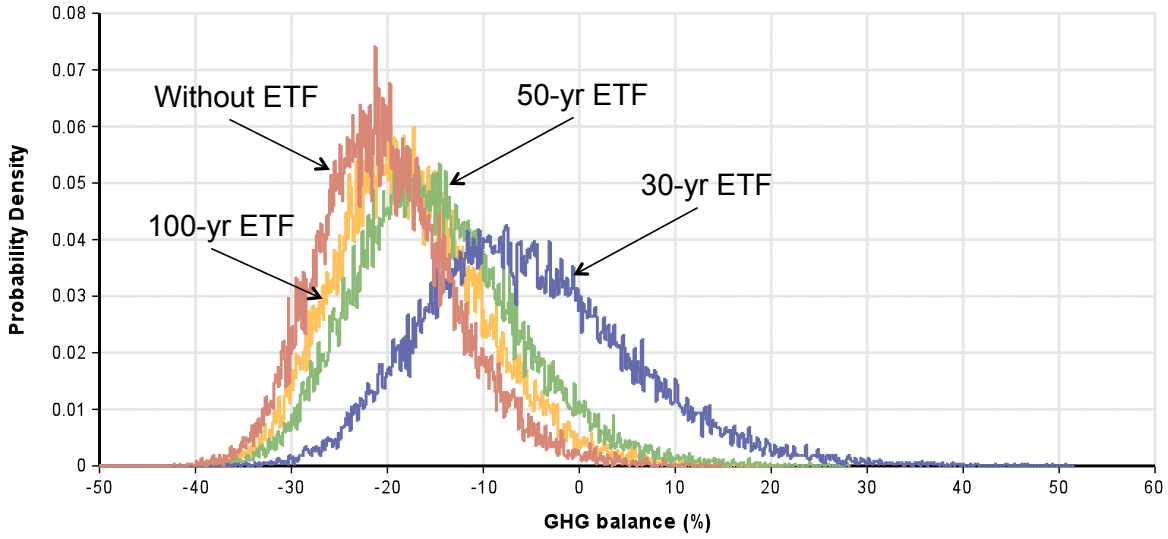

Figure A - 15: PDF of the GHG balance (no ETF) and for the adjusted GHG balance with 30-yr, 50-yr, and 100-yr ETF correction factor.

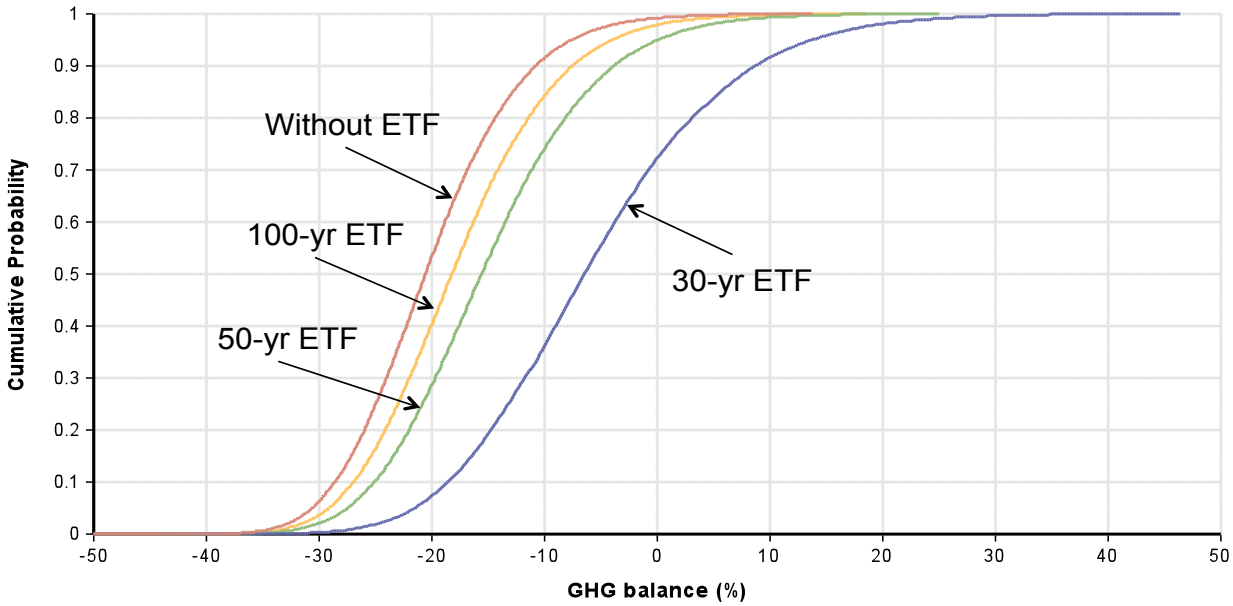

Figure A - 16: CDF of the GHG balance (no ETF) and for the adjusted GHG balance with 30-yr, 50-yr, and 100-yr ETF correction factor.

Figure A - 17 shows how ETF influences the probability of corn ethanol reducing GHG emissions relative to gasoline. The lower two lines (green) show the probabilities of corn ethanol meeting the EISA target of 20% GHG emissions reduction without ETF (dotted line) and including ETF (solid line). For impact time frames of 30-50 years, the probability may decline from 53% to about 7-29%. For longer time frames, the probability is still only 40% or less. The upper two lines (violet) show the probabilities of corn ethanol emitting equal or less than gasoline. This value is about 99% without ETF. However, including ETF reduces this probability to 71-98% depending on the impact time frame.

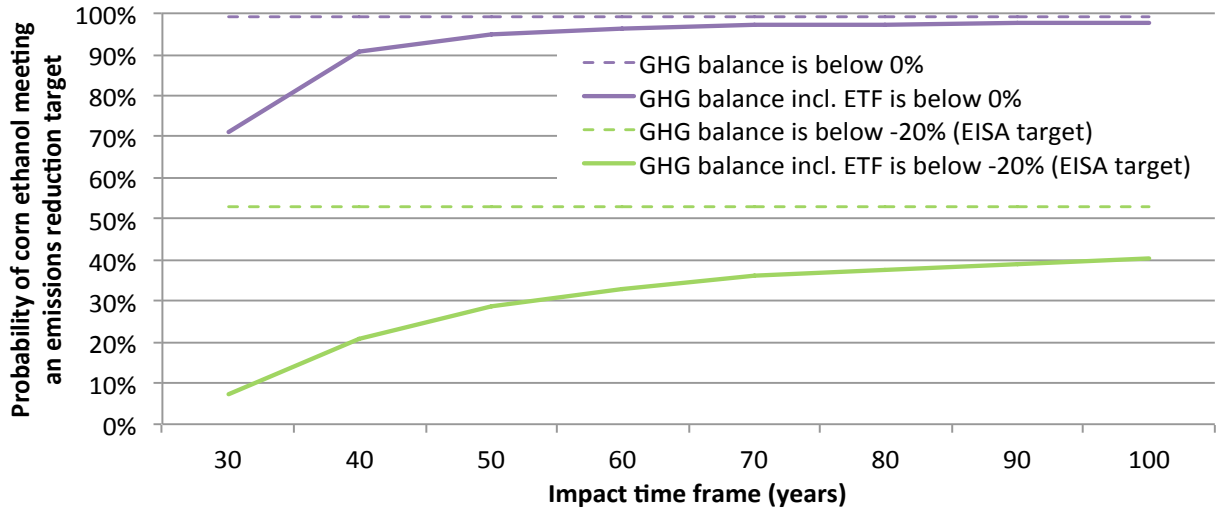

Figure A - 17: Influence of ETF on the probability that corn ethanol either meets the EISA target of 20% GHG emissions reduction (green) or has the same or less GHG emissions than gasoline (violet).

Summary of results

A statistical model was developed to investigate the influence of emissions timing on the probability that corn ethanol meets certain GHG reductions relative to gasoline. The EPA has estimated a distribution for the GHG balance of corn ethanol, and uses the mean value to test whether corn ethanol meets the 20% GHG reduction target of EISA. Perhaps a more meaningful metric to evaluate the GHG performance of corn ethanol is the probability that corn ethanol meets the target. Moreover, EPA does not account for the fact that LUC emissions from corn ethanol occur early in the life cycle, thereby constituting a greater CRF impact than the GHG balance conveys. The combination of the climate model and the statistical model addresses both the issue of evaluation metric and emissions timing.

The results show that from an LCA perspective, i.e., how much does ETF increase the GHG balance, ETF is fairly insignificant for LUC values of 30 g CO_2e/MJ or less and impact time frames of 50 years or higher. The small impact of ETF (2-13 percent points) is particularly apparent when considering the large uncertainties of LUC itself. It must be noted that other analyses (apart from EPA) estimate much larger LUC uncertainties than EPA (ranging within an order of magnitude).

However, ETF is much more significant from a policy perspective. Since EPA estimates that corn ethanol barely meets the EISA target (based on its mean value), even a small a skew in the distribution can significantly reduce the probability of meeting this target. Based on my model, ETF reduces this probability from 53% to 7-29% depending on the chosen impact time frame. Thus, switching from the “mean GHG balance” to “probability of meeting the target” amplifies the influence of ETF from 2-13 percent points to 24-46% in reduced probability of meeting the EISA target. Furthermore, ETF reduces the probability of corn ethanol emitting equal or less GHG emissions than gasoline from 99% to 71-98%.

The results presented here are subject to limitations. The distribution of total corn ethanol emissions that was generated using random number generation and Monte Carlo simulation is only a coarse reproduction of EPA estimates. The 97.5th percentile value of the final GHG balance distribution has a significant error relative to the original EPA estimates. Particularly, the WTW emissions distribution is uncertain because little data was available. This analysis could be improved if EPA sample distributions would be available. Furthermore, the bin size could be increased from currently 600 to the number of samples (10,000) to avoid the coarse PDFs, and to obtain more precise Monte Carlo simulation results.

B. Additional data for Chapters 3-5

Contents

Literature survey of OH and stratospheric CH_4 removal processes

Atmospheric CH_4 is oxidized through a series of reactions including, primarily, OH. The chain reaction initiated via CH_4 and OH is described in detail in the literature (*91*). Since OH is highly reactive and has a very short atmospheric residence time (a few seconds), it is difficult to measure its concentration throughout the globe. In order to characterize atmospheric OH, methyl chloroform (CH_3CCl_3 or MCF) is often used as a trace species, which has relatively well-known anthropogenic sources. Models can help quantify the abundance of OH in the atmosphere because MCF has a longer residence time than OH, and it is removed only by reaction with OH (*185*). MCF has been measured since 1978 revealing a steep increase until it reached a maximum in 1992, after which it declined rapidly because it has been phased out under the Montreal Protocol to protect the ozone layer (*185*). However, there is uncertainty regarding the precise amounts of MCF emitted to the atmosphere, e.g., due to potentially delayed MCF emissions relative to industry accounting (*186*). Uncertainty of MCF production is a major contributor to uncertainty in global OH abundance and thus CH_4 lifetime.

Prinn (*186*) supports the theory that industry emissions estimates were under-reported or at least delayed into the 1990s, which means that given observed MCF concentrations, higher or delayed MCF emissions suggest higher OH concentrations during this period. Given the observed CH_4 concentration during the same time, CH_4 emissions were relatively high in that period, which suggests CH_4 lifetime is relatively small (9.3 ± 0.7 years) in order to fit model simulations and measurements. Prather *dissertation*. (*187*) estimate CH_4 lifetime based on Monte Carlo simulations of an atmospheric/tropospheric chemistry model. The model includes abundances, growth rate, and multiple sinks for different species as parameters, and suggests a CH_4 lifetime of 9.1 ± 0.9 years. Holmes *dissertation*. (*90*) simulate CH_4 lifetime over the 1997-2009 period using three different chemical transport models, and then use the results to construct a parametric model to predict CH_4 lifetime since 1980. The authors focus on the tropospheric OH sink, which is the largest, and has the greatest impact on lifetime variability (other sinks include oxidation in the stratosphere, oxidation by tropospheric chlorine, and uptake into soil). Their CH_4 lifetime ranges from 8.5-10.1 years, which is within the range of "contemporary tropospheric chemistry models (ACCMIP)", i.e., 9.7 ± 1.5 years (*188*) and 9.3 ± 0.9 yr (*189*). A range of 9.1-9.7 years is chosen in this work based on the best estimates from the above studies.

Estimation of oil CH_4 EFs in Table 4

Total CH_4 emissions for a base year were given in the EPA (*64*) and Wilson *dissertation*. (*121*, *122*) studies, and the EFs were calculated using total oil produced in the same year as the total emissions. The EPA study estimated CH_4 emissions from the NG and oil systems separately. Assuming 65% vol-% CH_4 content in associated NG (see Table 2), the EPA EF suggests that 2% of gross NG production from oil wells is released to the atmosphere in 2011, which includes both venting and fugitives (*64*). Note that this number is one order of magnitude smaller than the DOE estimate (section 4.2.2). The Wilson *dissertation*. studies reported total CH_4 emissions occurring at GOM platforms, which produce both NG and oil. EFs for individual processes at the oilrigs were also provided on a component-day basis (not per barrel oil or BCF NG), but component-day activity data was not documented. This data was also not available upon request. Hence, allocation of emissions between NG and oil was needed to calculate EFs for oil. Two different allocation methods were used to yield a range of EFs: (i) allocating emissions according to the oil-to-NG emissions ratio from the EPA 2012 study (low estimate), and (ii) allocating all reported emissions to oil (high estimate). Detailed information about the allocation process and data is provided in Table B - 1. GOM total CH_4 emissions (2010) are based on activity data reported from platform operators and EFs from EPA authors. The emissions are likely underestimated because only about 75% of GOM platform operators (3,026 platforms (*122*) out of ~4,000 active platforms in 2006 (*190*)) reported activity. Furthermore, allocating all emissions to NG would yield a NG FE rate of only 0.9%, which is small compared to most recent U.S. estimates in Table 3. This suggests that the EFs used in the Wilson *dissertation*. study (summarized in Table 4) are either very small, or total production was underreported by the platform operators, thus leading to lower total emissions in the Wilson *dissertation*. study. Tier 1 EFs were readily available in the IPCC study for developing and developed countries. The largest fraction of oil EFs are due to extraction activities. For instance, according to EPA (*64*), oil refining and transport account for only 1% and 0.3% of total CH_4 emissions, respectively.

Table B - 1: Allocation of CH_4 emissions to oil and NG based on Wilson *dissertation*. (*121*, *122*)

	Unit	2004	2010
Reported emissions	MT CH_4	573,753	422,707
	Gg CH_4	574	423
GOM FE, gas only	%	0.94	1.47
GOM EF, oil only	Gg CH_4/barrel	1.10E-06	9.99E-07
Emissions from oil allocated	Gg CH_4	76.3	56.2
GOM EF, oil allocated	Gg CH_4/barrel	1.46E-07	1.33E-07
Platforms	-	2873	3026
Emissions ratio oil to NG (2005)	%	13.3	

Table B - 2: Calculation of flaring efficiency based on Wilson *dissertation*. (*121*, *122*)

0.126	lb CH_4/MMBtu flared gas
57.2	g CH_4/MMBtu flared gas
3,165,891,305	g CH_4/Tg flared CH_4
0.00032	Tg CH_4/Tg flared CH_4
0.032%	uncombusted CH_4 to atmosphere

Field measurements in Alberta indicated only 60-80% combustion efficiencies for industrial flares (*191*). The major variables affecting flare performance are flare gas composition, gas flow rate, wind speed, steam or air rate for assisted flares, flare tip size, design, configuration, exit velocity, and pilot stabilization. A recent experimental flare test study using industrial scale flares indicates 95-98% flaring efficiency with flare tips ranging from 3-6 inches in diameter (*192*). Flaring efficiency decreases with greater crosswind speed and pipe diameter and lower fuel mass flow (fuel exit velocity). Pipe diameter of >6 inches are possible in Canada (and up to 12 inches in other countries) because greater velocity means greater flaring capacity, which is needed in regions where less NG is transported to market (*193*). The larger pipe diameter causes an efficiency drop. Fuel mass flow may vary due to fixed capacity of installed NG infrastructure (NG equipment capacity is optimized for a given NG flow), i.e., mass flow at the flare tip is small when production is low and less NG is sent to flare. If NG flow drops while maintaining pipeline pressure, fuel mass flow in flare drops, which causes efficiency drop. Furthermore, flaring efficiency may decrease by corroding of pipes due to high NG sulfur content (*193*). As a result, the above flaring efficiency of 95-98% may be considered a high estimate compared to what is obtained in practice.

As discussed in section 4.2.2, flaring is particularly high in developing countries due to missing infrastructure to transport and process the associated NG, and sell it to the market (*129*, *130*). Figure B - 1 illustrates this phenomenon more clearly by showing the ratio of NG flaring to oil production. Countries with high NG flaring relative to oil (and therefore associated NG) production make relatively little use of their associated NG, likely due to the inability to capture this gas. No OECD country is among the Top 10 flaring countries in Figure B - 1, likely due to existing infrastructure to capture associated NG (*129*, *130*). As described in more detail in section 4.3, the ratio of NG flaring to produced NG over time is used to estimate flaring amounts for years prior to the start of flaring observations in 1994.

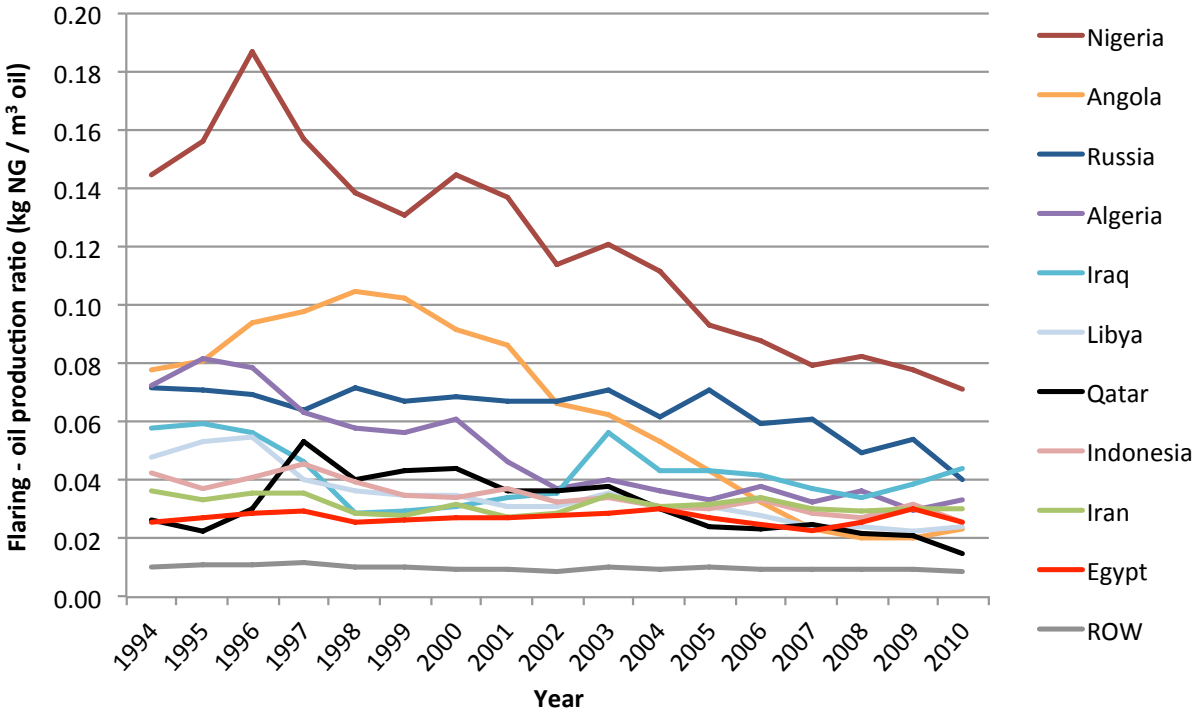

Figure B - 1: Ratio of NG flaring to oil production as an indicator of missing infrastructure to transport and process associated NG based on data from Elvidge *dissertation*. and EIA (*128*, *134*).

EFs of coal emissions literature review (section 4.2.3)

Emissions factors from coal mining depend mainly on the CH_4 content of coal. Global Tier 1 EFs are provided where country-specific EFs are not available. For comparison, the 2009 EPA estimate for the U.S. is 12 m^3 CH_4/t weighted average (range is 4-60 m^3/t) based on state-level data (*148*). This includes avoided emissions from CH_4 recovery and utilization, and excludes post-mining and abandoned mine emissions, which are shown in Table B - 4 and Table B - 5.

Table B - 3: CH_4 emissions factors for underground coal production in major coal producing countries, and Tier 1 emissions factors globally (*120*). Units are m^3 CH_4/t coal.

	Low	High
Former Soviet Union	17.8	22.2
U.S.	11.0	15.3
Germany	22.4	22.4
UK	15.3	15.3
Poland	6.8	12.0
Czech Republic	23.9	23.9
Australia	15.6	15.6
Tier 1 recommended	10	25

Average U.S. post-mining emissions are 1.5 m^3/t (underground mining) and 0.2 m^3/t (surface mining), and abandoned mine emissions are 1.3 m^3/t (Table B - 4). The recommended Tier 1 post-mining emissions in IPCC range from 0.9-4.0 m^3/t (*120*), which are due to coal processing, transportation, and final crushing and pulverization at the coal plant (*131*). Tier 1 abandoned mines EFs depend largely on the number, dates, and types (vented, sealed, flooded) of mines closed, of which little data is available (*120*).

Table B - 4: Estimates of U.S. coal post-mining CH_4 emissions (*64*).

	Units	1990	2005	2010	Average
Emissions					
UG	Million m^3 CH_4	550	457	403	
Surface		139	154	152	
U.S. underground coal production					
UG	Million MT	385	296	265	
Surface		548	730	719	
Emissions Factors					
UG	m^3 CH_4/MT coal	1.43	1.54	1.52	**1.50**
Surface		0.25	0.21	0.21	**0.23**

As a proxy for global abandoned mines emissions, Table B - 5 calculates an EF based on reported total U.S. abandoned mines emissions and coal production over the period 1990-2010. The EF is in the order of only 5-10% compared to coal mining activities.

This is likely a low estimate with respect to the period 1980-2011 given that CH_4 recovery from U.S. abandoned mines was substantial over the last decade (*64*, *194*). U.S. post-mining and abandoned mine emissions in the EPA 2012 GHG inventory are 10% and 7% of total CH_4 emissions from the coal industry (*64*). Uncertainty range in the EPA 2012 GHG inventory is -19% to +30% (95% C.I.).

Table B - 5: Estimates of U.S. abandoned coal mine CH_4 emissions (*64*).

	Units	1990	2005	2010	Average
Emissions					
Abandoned UG	Million m^3 CH_4	430	499	543	
Recovered & used		-	104	188	
Net emissions		430	394	355	
Emissions Factors					
Abandoned UG	m^3 CH_4/MT coal	1.1	1.7	2.0	
Recovered & used		-	0.3	0.6	
Net emissions		1.1	1.4	1.5	**1.3**

Coal-bed methane estimates are accounted separately from CH_4 recovery (*134*). 1989 CBM production is equal to 50% of coal mine CH_4 emissions in the same year (2.6 billion m^3 CBM). According to EPA, 1.1 billion m^3 produced from a single mine in Alabama using gob wells over 6 years at a 30-40% capture rate at this mine (*131*). 2010 CBM production is equal to 1,400% of coal mine CH_4 emissions in the same year (*134*). Recent CBM production may be an order of magnitude greater than coal mine CH_4 emissions because CBM can be produced at much larger rates than coal on a per mine basis, i.e., the annual coal seam volume tapped for CBM production is larger than that for coal production. Vertical wells are used to drain CH_4 from coal beds 10 years prior to mining in order to produce CBM (*131*).

China is the world's largest coal producer accounting for 53% of global underground coal production over 1980-2010 (*134*). Since a Chinese EF is not listed in (*120*), further research was necessary. EPA provides average EFs for state-owned coal mines in all Chinese provinces, which account for about one third of Chinese coal production in 1994 (*132*, *133*). Average EFs range from 2 to 38 m^3/t by province with a weighted average ranging from 10 to 12 m^3/t. The majority of produced coal in 1994 came from low CH_4 emissions mines. While coal production shifted among provinces from 1996 to 2009, the weighted average EFs did not change significantly (11.0 and 11.7 m^3/t in 1994 and 2009, respectively). The share of underground production is 90% (*133*) as opposed to 95% elsewhere (*132*). According to Raven Ridge Resources, the contractor for the above EPA report, individual province level EFs are trustworthy because this information is critical for safety reasons, i.e., avoiding hazardous mine CH_4 concentrations (*135*).

EFs and data for coal emissions inventory methods (section 4.4.2)

International hard and lignite production data were used for underground and surface production, respectively, unless if better data was not available. Abandoned mines emissions were included only if countries reported them. Estimated emissions mitigation is in the order of 7% of emissions globally (*195*, *196*), but U.S. CH_4 recovery increases from 8% to 26% from 1990-2010 (*64*, *148*). A challenge in establishing accurate emissions inventories is the lack of data or anecdotal nature of reported data. For example, China and India, among other countries, have extensive uncontrolled fires in their coal mining regions which may add to fugitive emissions, but are not included in the estimates (*197*). A database listing global coal CH_4 recovery projects with data on mitigated CH_4 exists online (*194*), but only 14 out 315 listed projects document the amount of CH_4 drained and/or vented. Table B - 6 summarizes underground and surface mining by country for major coal producing countries.

Table B - 6: Literature survey of underground and surface mining of major coal producing countries

	Percentage of world primary coal production during 1980-2010	Underground mining	Surface mining	Year, if available	Source	Additional source notes
China	28.9%	90%	10%	-	(*133*), (*198*)	Reserves mineable in opencast pits are relatively low, and most of the mining is pursued in underground mines
United States	18.7%	31%	69%	1990-2009	(*148*)	-
Germany	6.6%	Hard coal/lignite split		-	(*198*), (*199*)	-
India	6.5%	10%	90%	2010	(*198*), (*200*)	25%/75% split in 2004
Russian Federation	6.4%	31%	69%	2010	(*201*)	-
Australia	5.2%	23%	77%	-	(*198*), (*201*)	43%/57% hard coal split in 2002, and all lignite is surface mining (that's 34.5% underground of total in 2002)
South Africa	4.1%	47%	53%	2002	(*198*)	Nearly half of the production comes from surface mining operations (2010)
Poland	4.0%	Hard coal/lignite split		-	(*199*), (*202*)	Poland's lignite deposits are exclusively mined in opencast mines. All hard coal is deep mined at an average working depth of some 600 metres.
Kazakhstan	2.4%	see notes		-	(*198*)	By the end of 2002, many high-cost underground coal mines have been closed, and its more competitive surface mines have been purchased, and are now operated
Ukraine	2.0%	98%	2%	-	(*198*)	Only three out of 149 active coal mines are surface mines
Czech Republic	1.8%	Hard coal/lignite split		-	(*202*)	-
Indonesia	1.7%	5%	95%	-	(*198*)	Coal is almost entirely extracted in opencast operations
Canada	1.3%	5%	95%	-	(*198*), (*203*)	Coal mining in the western provinces is confined to opencast pits. Over 90% of all deposits are concentrated in the western provinces of Saskatchewan, Alberta and British Colombia. The Canadian coal industry is privately owned. Output is mainly from surface mines.
United Kingdom	1.3%	65%	35%	1980-2011	(*204*)	-
Greece	1.1%	see notes		-	(*199*)	Lignite is mostly mined […] exclusively in opencast mines.
Turkey	1.0%	see notes		-	(*199*)	Almost 90% of Turkey's total lignite production is from opencast mines.
Romania	0.8%	see notes		-	(*199*)	95% of lignite deposits are situated in the Oltenia mining basin and more than 80 % of these can be mined in opencast mines. The remaining lignite deposits have low economic potential, explaining why the extraction in most other areas has stopped.
Spain	0.6%	see notes		-	(*199*)	Over 60% of the hard coal is mined in opencast mines
Brazil	0.1%	35%	65%	2005	(*203*)	-
Venezuela	0.1%	5%	95%	-	(*198*)	Most of the coal is mined in opencast pits
Slovenia	0.1%	100%	0%	-	(*202*)	-
Slovakia	0.0%	100%	0%	-	(*202*)	-
Sum	94.7%					

Country-specific activity data used in the emissions inventory

British Petroleum (BP) data is available for 49 countries, the remaining countries being grouped in world regions. Oil, NG, and coal production statistics start in 1965, 1970, and 1981, respectively (*205*). The complete International Energy Agency/OECD database is available through paying a fee (*206*). E-books are free, but the temporal resolution is sparse (*207*). BP and IEA/OECD data mostly matches EIA data, and the largest observed discrepancies are in the order of 10-20%.

Country-specific FF production and consumption data was analyzed from 1980-2010. For coal, the average difference between production and consumption in a given country is up to 75% of production (Australia, Indonesia) due to import and export activities. The difference worldwide weighted by production is 10%, which does not account for larger differences in individual years and countries. For NG, the average difference is up to 93% of production (Norway, Turkmenistan, Algeria), and worldwide weighted by production is 22%. For oil, the average difference is well above 50% of production for most countries, and worldwide weighted by production is 56%.

Table B - 7: List of countries included in the emissions inventory using country-specific EFs. Percentage refers to fraction of global production (coal: *primary* coal production).

	NG		Oil		Coal	
	Country	**Percentage**	**Country**	**Percentage**	**Country**	**Percentage**
1	United States	22.9%	United States	13.4%	China	30.0%
2	Russia	24.7%	Saudi Arabia	11.9%	United States	18.7%
3	Canada	6.0%	Russia	12.5%	Germany	6.6%
4	Netherlands	3.5%	Iran	4.6%	India	6.1%
5	United Kingdom	2.9%	Mexico	4.4%	Russia	6.4%
6	Algeria	2.7%	China	4.3%	Australia	5.2%
7	Indonesia	2.3%	Venezuela	3.6%	South Africa	4.1%
8	Iran	2.4%	Canada	3.5%	Poland	4.0%
9	Norway	2.2%	United Arab Emirates	3.2%	Kazakhstan	2.4%
10	Saudi Arabia	1.9%	United Kingdom	3.1%	Ukraine	2.0%
11	Uzbekistan	2.1%	Norway	3.0%	Czech Republic	1.8%
12	Mexico	1.5%	Nigeria	2.7%	Indonesia	1.7%
13	Malaysia	1.4%	Kuwait	2.6%	Canada	1.3%
14	United Arab Emirates	1.4%	Iraq	2.4%	United Kingdom	1.3%
15	China	1.4%	Algeria	2.0%	Greece	1.1%
16	Turkmenistan	2.1%	Libya	2.0%	Turkey	1.1%
17	Argentina	1.2%	Indonesia	1.9%	Romania	0.8%
18	Australia	1.2%	Brazil	1.8%	Spain	0.6%
19	Qatar	1.2%	Egypt	1.1%	Brazil	0.1%
20	Romania	1.0%	Angola	1.1%	Venezuela	0.1%

	NG		Oil		Coal	
	Country	**Percentage**	**Country**	**Percentage**	**Country**	**Percentage**
21	Venezuela	1.0%	Oman	1.0%	Slovenia	0.1%
22	Pakistan	0.9%	Qatar	1.0%	Slovakia	0.0%
23	Egypt	0.9%	India	1.0%		
24	India	0.8%	Argentina	1.0%		
25	Italy	0.7%	Malaysia	0.9%		
26	Trinidad and Tobago	0.7%	Australia	0.8%		
27	Thailand	0.6%				
28	Germany	0.9%				
29	Ukraine	0.8%				
30	Nigeria	0.5%				
31	Brunei	0.4%				
32	Oman	0.4%				
33	Bangladesh	0.4%				
34	Libya	0.3%				
35	Bahrain	0.3%				
36	Kuwait	0.3%				
37	Poland	0.2%				
Sum		**96.2%**		**90.8%**		**95.4%**

Procedure for scaling EDGAR grid maps

Individual grid maps for NG, oil, and coal in NetCDF file format (*94*) were imported to Matlab. A 1,800 x 1 array and a 1 x 3,600 array was generated in Matlab, with each row and column including and signifying the number of the latitude and longitude position of a 0.1° x 0.1° world grid. The next step was generating a matrix representing this 0.1° x 0.1° world grid, which identifies the country belonging to each grid cell. This is important for scaling the grid maps where each country has a different scaling factor. A country identifier text file was provided by EDGAR (*85*), which included latitude and longitude (i.e., a rectangle on a world grid), the area percentage of the particular country in the rectangle (i.e., bordering countries share the same rectangle), and the country identifier code (latitude, longitude, and percentage were separated into one file "numdata", and the country identifier code into one file "country" in order to avoid Matlab data type conflicts when using strings and numerics). EDGAR's country identifier data was used, and the latitude and longitude entries with the associated grid rows and columns from the aforementioned 1,800 x 1 and 1 x 3,600 arrays were replaced (convert_numdata_lat.m and convert_numdata_lon.m; this converts numdata to numdata_matrix). Next, a 1,800 x 3,600 matrix was created for each country to be scaled (see Table B - 7). In each country matrix, every cell belonging to this country contains a value between 0 and 1, where 1 is a cell belonging to this country, decimal values representing the percentage of a borderline country described above, and 0 for all other cells (make_country.m). A similar procedure was conducted for adding country identifiers to grid cells referring to maritime areas, which were not included in EDGAR's country identifier file using instead the VLIZ Maritime Boundaries Geodatabase Version v6.1_20110512 (add_maritime.m). The resulting grid maps are contained in gridded_NG_countries.mat, gridded_oil_countries.mat, and gridded_coal_countries.mat.

For each FF source (NG, oil, coal) and year, each country grid map was scaled individually in Matlab (convert_emissions_coal.m, convert_emissions_oil.m etc.) using the scaling factors generated in Excel spreadsheets (Global CH_4 emissions from FFs.xlsx). The scaling factors were generated by calculating the relative difference between the total annual CH_4 emissions developed in section 4.3 (for each FF source, country and year) and EDGAR (note this is EDGAR v4.2). For each FF source (NG, oil, coal) and year, all scaled country grids were added along with the remaining un-scaled countries from EDGAR to yield a single grid map containing the emissions from all countries. There were two major challenges in obtaining the scaling factors due to missing or erroneous data in the EDGAR database. First, the global sum of the country totals in the EDGAR coal emissions spreadsheet does not match the totals of the grid maps indicating that the individual country emissions in the spreadsheet are erroneous. Second, country totals for oil and natural gas are aggregated in the EDGAR spreadsheet, and need to be disaggregated to calculate scaling factors for oil and NG individually. The EDGAR country totals for all three FF sources were determined in a "re-engineering" process by using the emissions from the grid maps in combination with national identifier maps, and then checking the world totals with the numbers provided by EDGAR.

GFED daily grid maps were first aggregated into monthly grid maps to yield an annual cycle defined by monthly total values (aggregate_daily_to_monthly.m). Next, the aggregated monthly grid maps were scaled to match annual totals from EDGAR (ag_waste_Lori_scale.m). In order to maintain compatibility between Matlab grid map files and NOAA's CT-CH_4, the 0.1° x 0.1° grid resolution was compressed to 1° x 1° grid (aggregate_to_1x1.m), and converted to text files.

Annual NG well sample analysis results, pipeline composition standards, and chemical properties of NG

Figure B - 2 shows hydrocarbon composition of the 6,109 samples from 1948-1997, the period over which all hydrocarbon and N_2 and CO_2 data was available. The large inter-annual variation is likely due to differences in sampling sites each year and/or small sample size in particular years, rather than composition changes over time. For instance, the CH_4 drop in 1960 is partially due to only 15 samples in that year. Other components not shown, e.g., oxygen and argon, are generally below 1 mol-% of NG.

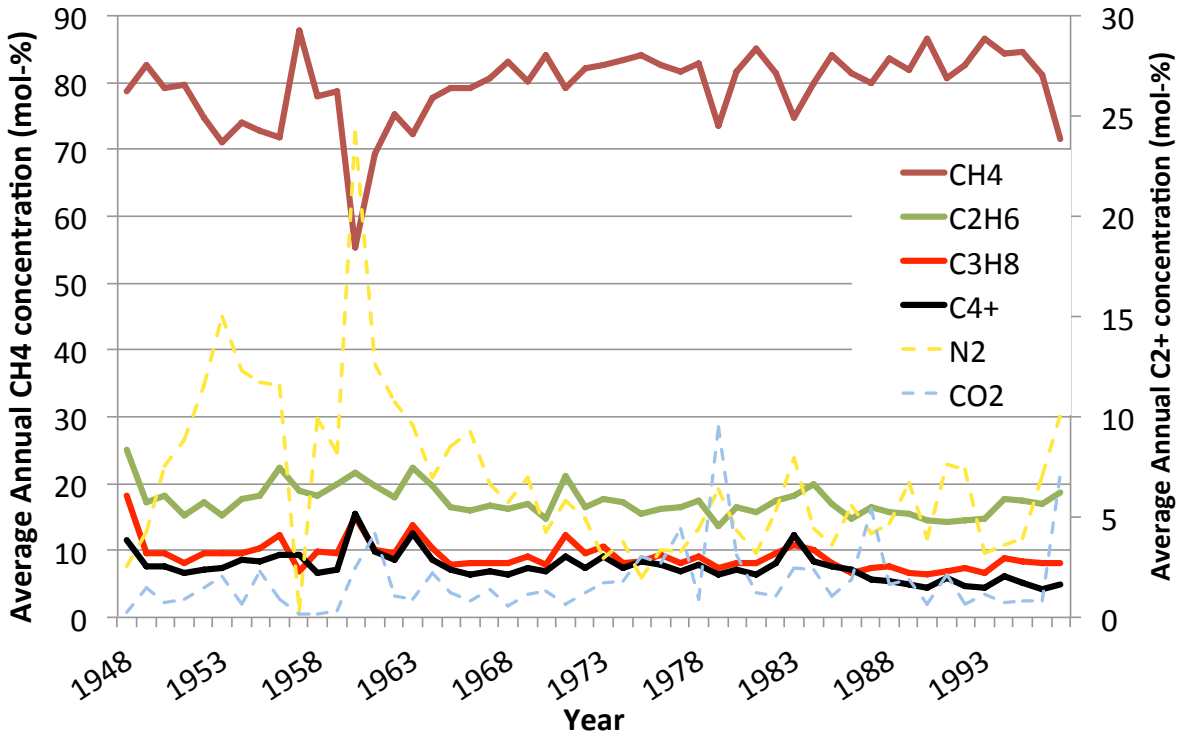

Figure B - 2: U.S. well sample gas composition analysis results by year using BLM data (*104*).

In order to check whether the mass balance calculations in sections 4.2.1 and 4.4.1 are reasonable, the energy content of the resulting dry gas composition can be compared with literature values of pipeline quality standards. As described above, NG from the wellhead or field facility meets pipeline quality standards in only few cases (*99*). Mostly, NGLs and other components are separated in a processing plant, which changes the composition of the NG. Condensates are often removed from the gas stream at the wellhead through the use of mechanical separators. NGL rich gas is often found in deep-water GOM and the Rocky Mountain area, which must be processed to meet pipeline company standards. Standards include a specific energy content range, e.g., 1,035 Btu per cubic feet, +/- 50 Btu (*99*). Other sources give a wider range including 950 - 1,150 Btu/scf (*157*, *208*). To achieve the standards, processing facilities sometimes blend some heavy hydrocarbons into the gas stream (extract less NGLs) in order to bring it within acceptable Btu levels and/or absorb some of the NGLs if their market

prices are too low. The quantity of NGLs and other byproducts actually produced depends on current market prices (*99*).

Table B - 8: Chemical properties of NG components used for conversion from mol-% to wt-% (*209*, *210*).

Conversion metric	Units	CH_4	C_2H_6	C_3H_8	C_4H_{10}	CO_2	N_2
Molar mass	g/mol	16.04	30.07	44.1	65.3	32	28
Gas density[a] (ρ)	g/m^3	680	1,282	1,910	2,510	1,870	1,165
Energy content	Btu/cft	1,011	1,783	2,572	3,603	n/a	n/a

[a] At 1.013 bar and 15 °C

Global NG industry C_2H_6 emissions for all FE rate and C_2H_6 scenarios

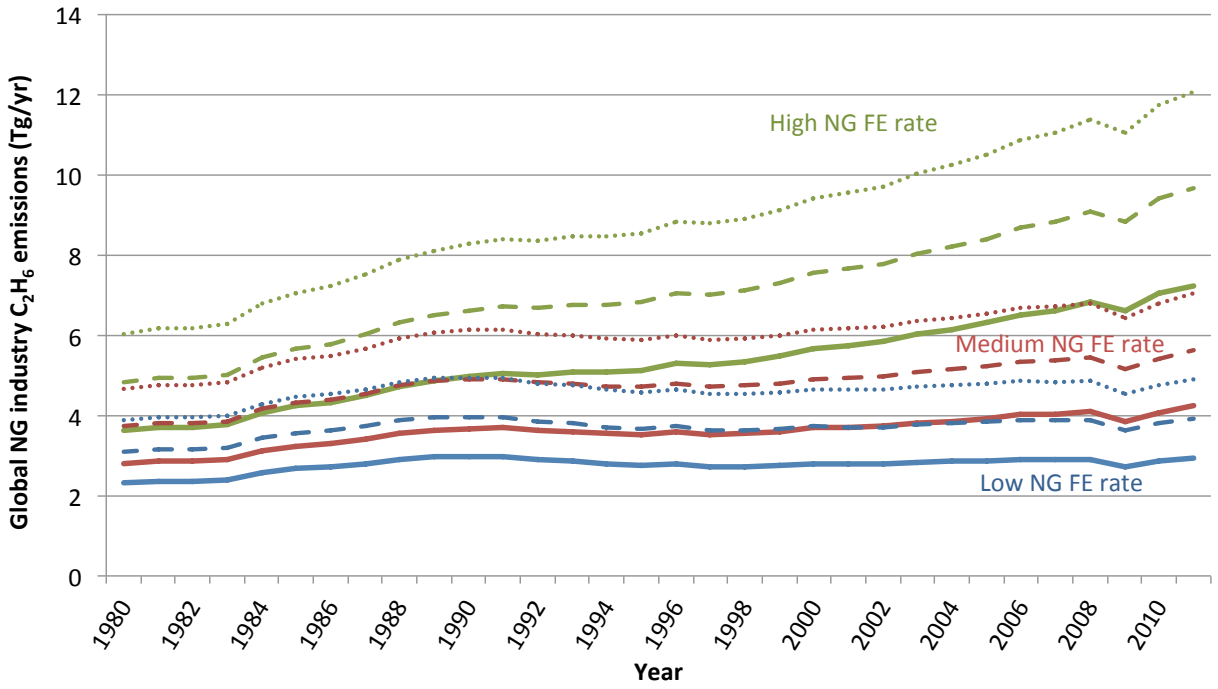

Figure B - 3: Global NG industry C_2H_6 emissions based on FE rates applied gobally and constant over time (see section 4.4). Dotted, dashed, and solid lines are high, medium, and low NG C_2H_6 content scenarios.

Country-level *prior* emissions differences between this work and EDGAR

Table B - 9: Country-level *prior* emissions differences between this work and EDGAR (*85*) for NG, oil, and coal. Results are differences relative to EDGAR, i.e., positive percentage values indicate higher emissions in this work compared to EDGAR. The NG column assumes a 3% FE rate applied globally.

NG		Oil		Coal	
United States	63%	United States	71%	China	-10%
Russia	-22%	Saudi Arabia	14%	United States	-4%
Canada	316%	Russia	32%	Germany	41%
Netherlands	552%			India	4%
United Kingdom	95%	Iran	-32%	Russia	-13%
Algeria	28%	Mexico	6%	Australia	30%
Indonesia	15%	China	12%	South Africa	44%
Iran	-29%	Venezuela	12%	Poland	-55%
Norway	697%	Canada	-63%	Kazakhstan	-16%
Saudi Arabia	53%	United Arab Emirates	24%	Ukraine	171%
Uzbekistan	3%	United Kingdom	266%	Czech Republic	289%
Mexico	-32%	Norway	46%	Indonesia	-57%
Malaysia	-8%	Nigeria	-52%	Canada	4%
United Arab Emirates	58%	Kuwait	17%	United Kingdom	22%
China	-44%	Iraq	-28%	Greece	6%
Turkmenistan	10%	Algeria	-40%	Turkey	-8%
Argentina	-9%	Libya	-31%	Romania	31%
Australia	138%	Indonesia	-27%	Spain	-10%
Qatar	26%	Brazil	26%	Brazil	-37%
Romania	-28%	Egypt	-15%	Venezuela	-87%
Venezuela	-24%	Angola	-46%	Slovenia	41%
Pakistan	-55%	Oman	-37%	Slovakia	579%
Egypt	-6%	Qatar	-16%		
India	-11%	India	-19%		
Italy	1%	Argentina	3%		
Trinidad and Tobago	77%	Malaysia	-22%		
Thailand	35%	Australia	-9%		
Germany	27%				
Ukraine	-80%				
Nigeria	3%				
Brunei	169%				
Oman	-16%				
Bangladesh	-2%				
Libya	-27%				

NG		Oil		Coal	
Bahrain	92%				
Kuwait	26%				
Poland	-37%				
Total	**10%**	**Total**	**-4%**	**Total**	**-2%**

Oil and coal emissions grid maps

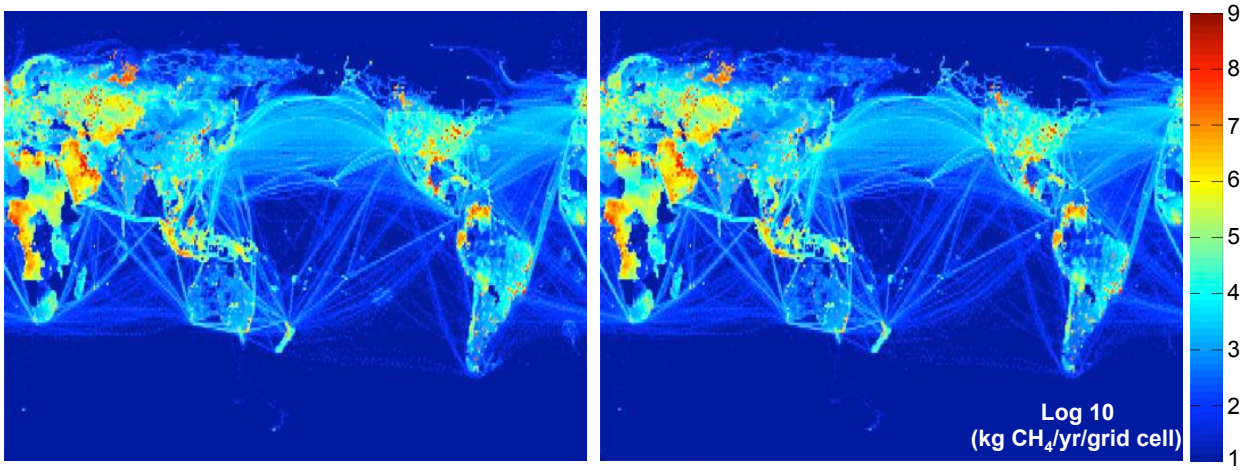

Figure B - 4: Grid maps (1° x 1°) for oil CH_4 emissions from this work (left) in comparison with data from EDGAR (*94*) (right) for the year 2008. Note that the legend units are in kg CH_4/yr/grid cell on a logarithmic scale (numbers indicating exponents to base 10).

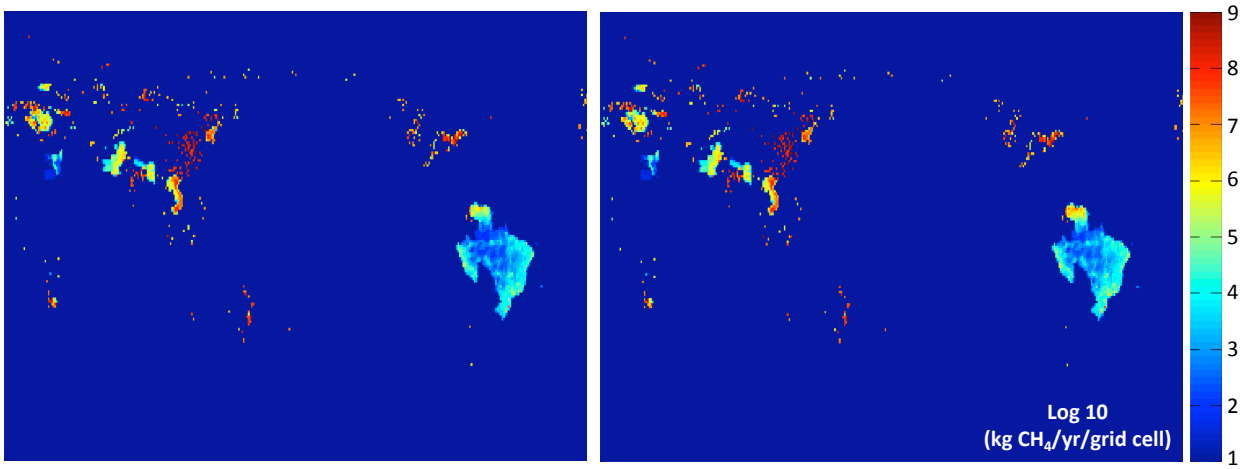

Figure B - 5: Grid maps (1° x 1°) for coal CH_4 emissions from this work (left) in comparison with data from EDGAR (*94*) (right) for the year 2008. Note that the legend units are in kg CH_4/yr/grid cell on a logarithmic scale (numbers indicating exponents to base 10).

Observational data and comparison of box-model and CT-CH_4 results

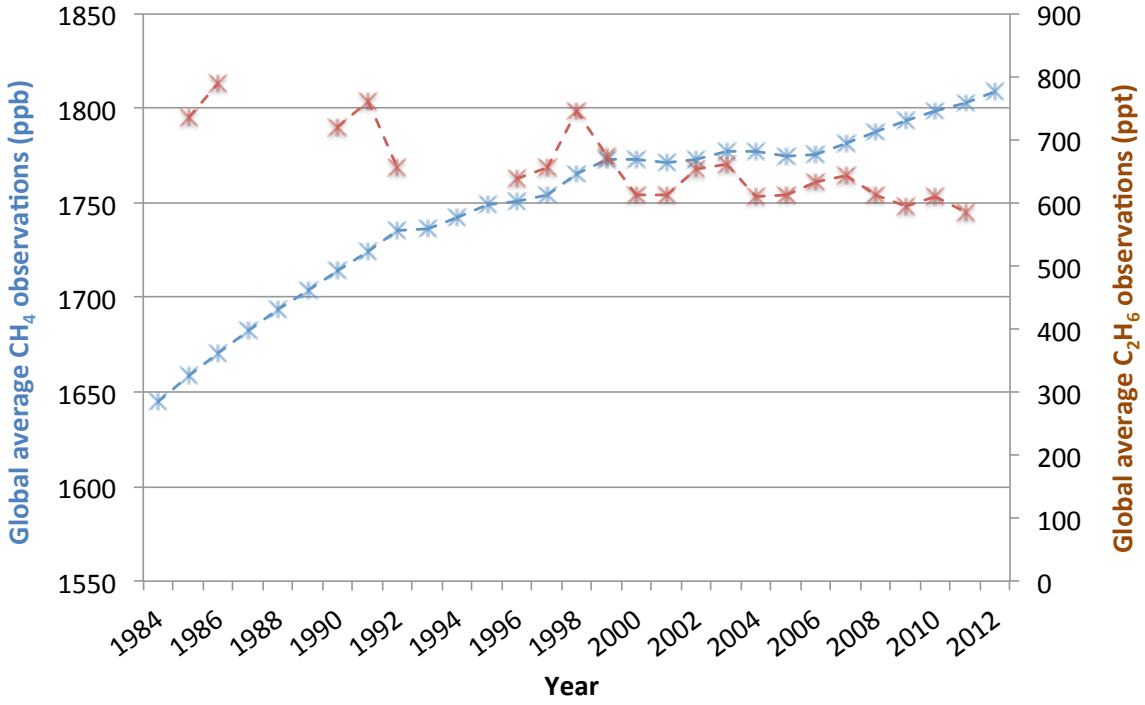

Figure B - 6: Observational CH_4 (blue, left axis) and C_2H_6 (red, right axis) data from NOAA (*83*) and UC-Irvine (via personal communication with Lori Bruhwiler at NOAA; see (*84*) for details), respectively, and trends (dashed lines) used for box-model simulations. Uncertainty of global average CH_4 concentrations is 1.1 ppb (one std. dev.) averaged across years (*165*). Corresponding C_2H_6 data was not available at time of dissertation submission.

Box-model simulations were performed using the same *prior* emissions dataset as in a previous global inversion study using CT-CH_4 (*39*) in order to analyze how well the box-model reproduces the globally averaged 3D-simulation results. The global and annual averages of the observational data used in CT-CH_4 were applied in the box-model. Figures Figure B - 7 and Figure B - 8 show that the total *posterior*s, the inter-annual variations as well as the allocation among sources is well reproduced in the box-model. The box-model *posterior*s are slightly higher than CT-CH_4 *posterior*s because Figure B - 7 excludes the ocean source, which would another 10-15 Tg CH_4/yr. The *posterior* uncertainties in the box-model are significantly larger than in CT-CH_4, which was expected given that there is more observational data in the 3D model to constrain each source. As shown in Figures Figure B - 9 and Figure B - 10, the box-model also provides a good reproduction of the long-term trends of concentrations, while seasonal variations cannot be represented, which was expected as well. Figure B - 10 shows that despite using reasonable observational uncertainties (+/- 20 ppb), the box-model *posterior*s will closely match the global budget, i.e., the measurements. This was also expected because *posterior*s in CT-CH_4 are the result of minimizing differences between *priors* and measurements at all sites, which causes measurements at some sites to be

weighted less. Since the box-model works with only a single globally averaged observation per year, such weighting scheme is absent in the box-model.

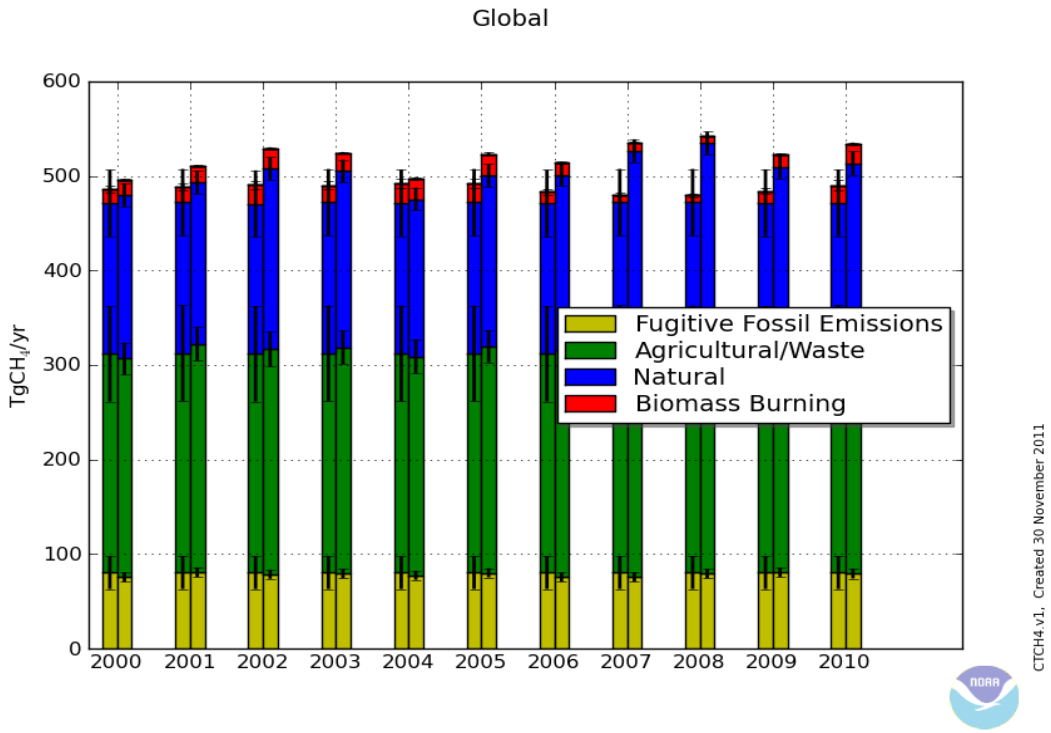

Figure B - 7: Globally averaged CT-CH_4 inversion results (graph from (*39*)), which uses an approximate CH_4 lifetime τ=9.45. Note that the ocean source (~10-15 Tg CH_4/yr) is not included in this graph.

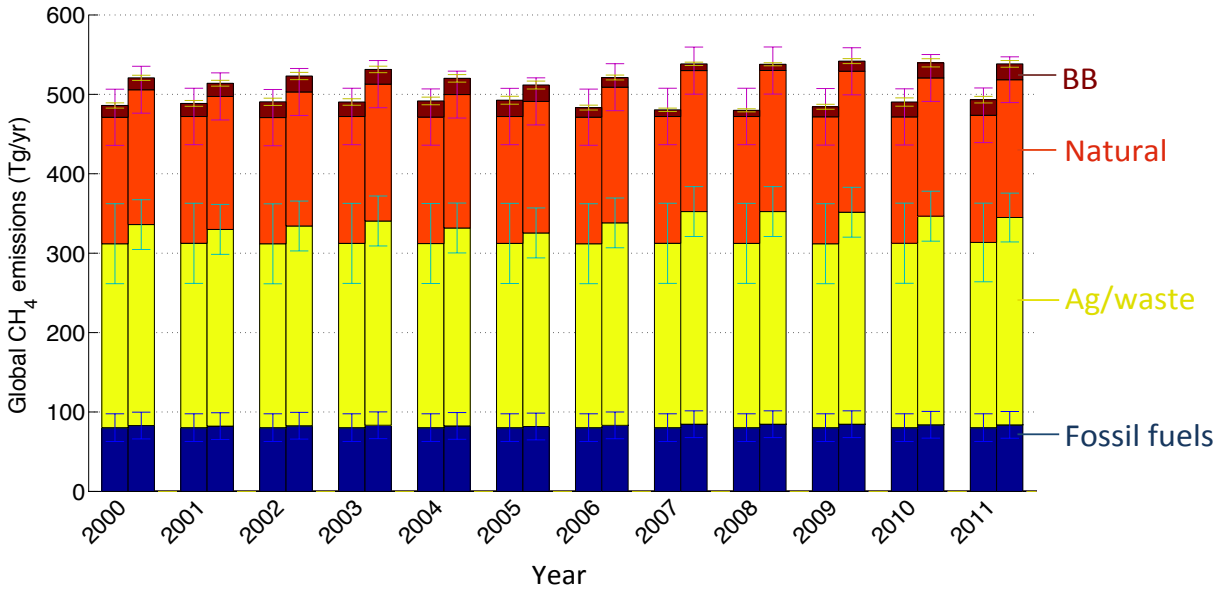

Figure B - 8: Box-model inversion results with CH_4 lifetime τ=9.45, observational uncertainties of +/- 20 ppb, and using the same *prior* emissions dataset as in the CT-CH_4 inversion (see Figure B - 7).

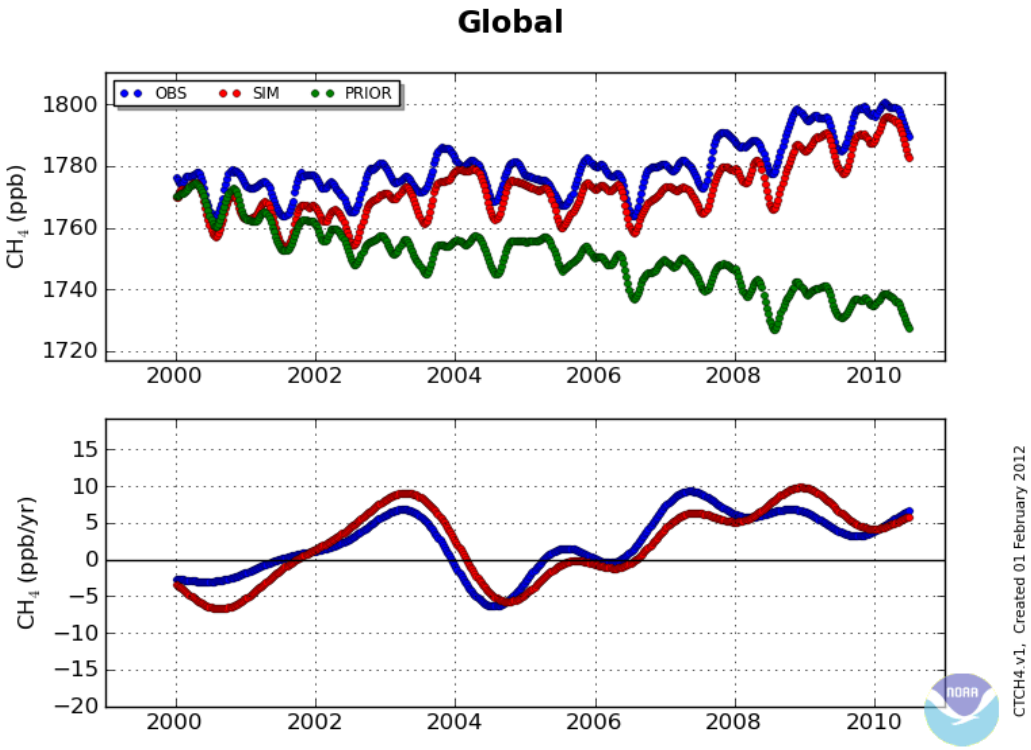

Figure B - 9: Globally averaged CT-CH_4 forward modeling results (graph from (*39*)), which uses an approximate CH_4 lifetime τ=9.45.

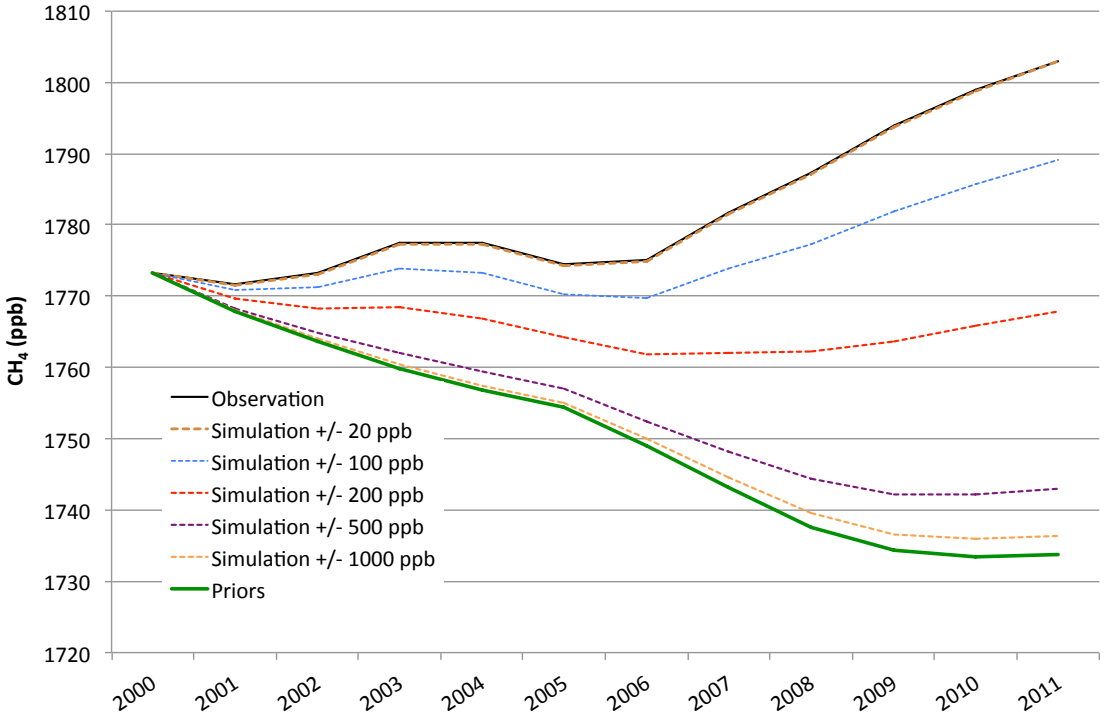

Figure B - 10: Box-model forward simulation results using the *posterior* emissions from the box-model inversion (see Figure B - 8) for a range observational uncertainties.

Linear optimization problem for isotope mass balance

Eq. 36 and Eq. 37 were re-arranged to give:

$$z_{FF} = \frac{\delta_q \times z - \delta_{Mic} \times (z - z_{BB}) - \delta_{BB} \times z_{BB}}{\delta_{FF} - \delta_{Mic}} \quad \textbf{Eq. B - 1,}$$

$$z_{Mic} = z - z_{FF} - z_{BB} \quad \textbf{Eq. B - 2},$$

where units for z and δ are in Tg/yr and ‰, respectively. The optimization problem is to minimize Eq. B - 1, such that:

$$z_{BB} \geq 25 \quad \textbf{Eq. B - 3,}$$

$$-38 \geq \delta_{FF} \geq -42 \quad \textbf{Eq. B - 4},$$

$$-59 \geq \delta_{Mic} \geq -63 \quad \textbf{Eq. B - 5},$$

$$-22 \geq \delta_{BB} \geq -26 \quad \textbf{Eq. B - 6}.$$

Eq. B - 3 ensures that there are only two unknowns in the problem of two linear equations with δ_q and Q known as described in section 5.2. The solutions in Table 18 were found using the Generalized Reduced Gradient method in Microsoft Excel Solver.

Illustration of the box-model mass balance approach

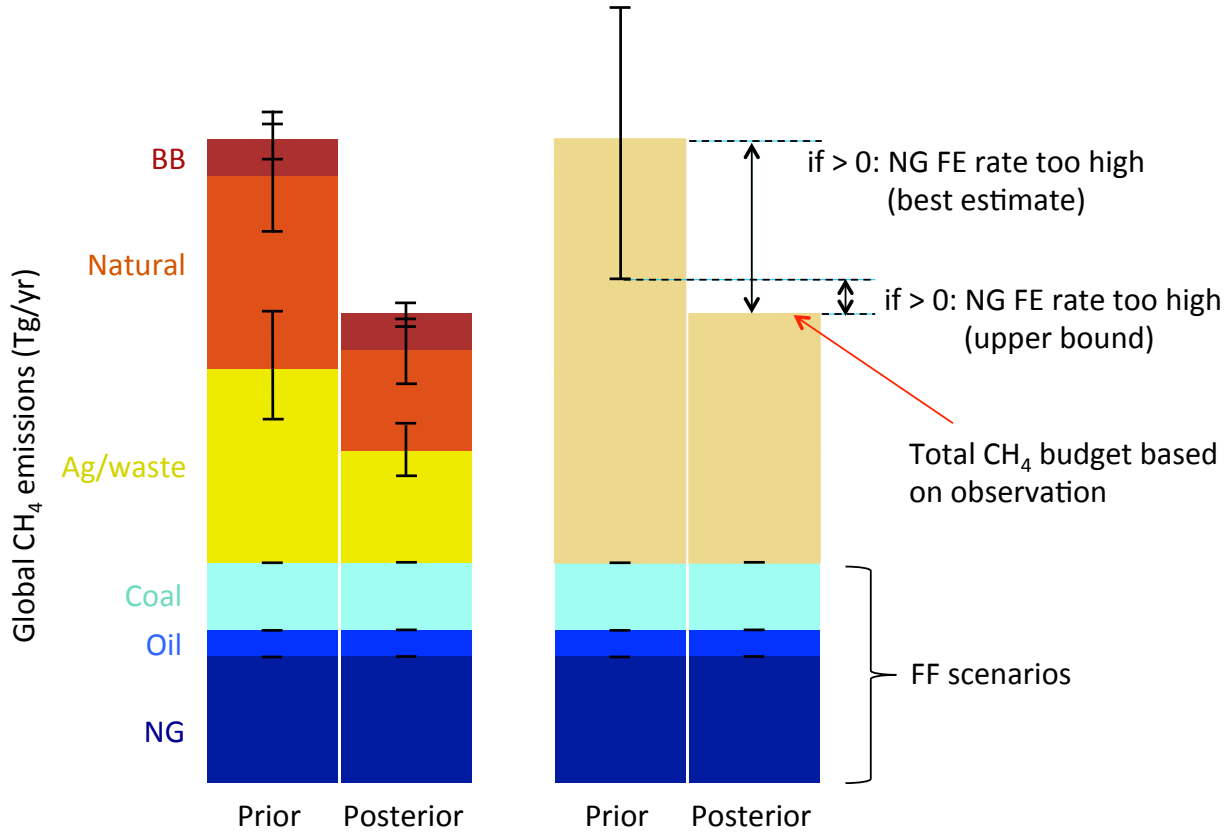

Figure B - 11: Qualitative illustration of mass balance approach using inversion box-model results. Priors and posteriors on the left represent a single year from the box-model. FF prior and posterior uncertainties are nearly zero, and FF estimates represent conservative (i.e., low) estimates for oil and coal (see Chapter 4) and one of several modeled scenarios for NG. Reducing the NG FE rate scenario will reduce the total prior emissions, thereby reducing the mismatch between total priors and total observation based posteriors. The combined prior non-FF uncertainties on the right are the sum of Ag/waste, natural, and BB uncertainties as described in Figure 17 and Table 16. For a given FF scenario, the associated NG FE rate is too high (as a best estimate) if the prior mean is greater than the posterior mean. For a given FF scenario, the associated NG FE rate is too high (as an upper bound) if the prior lower bound is greater than the posterior mean. Note that both the best estimate and upper bound NG FE rates estimated here are skewed towards higher values because conservative oil and coal EFs were chosen in Chapter 4.

Additional box-model and CT-CH_4 results

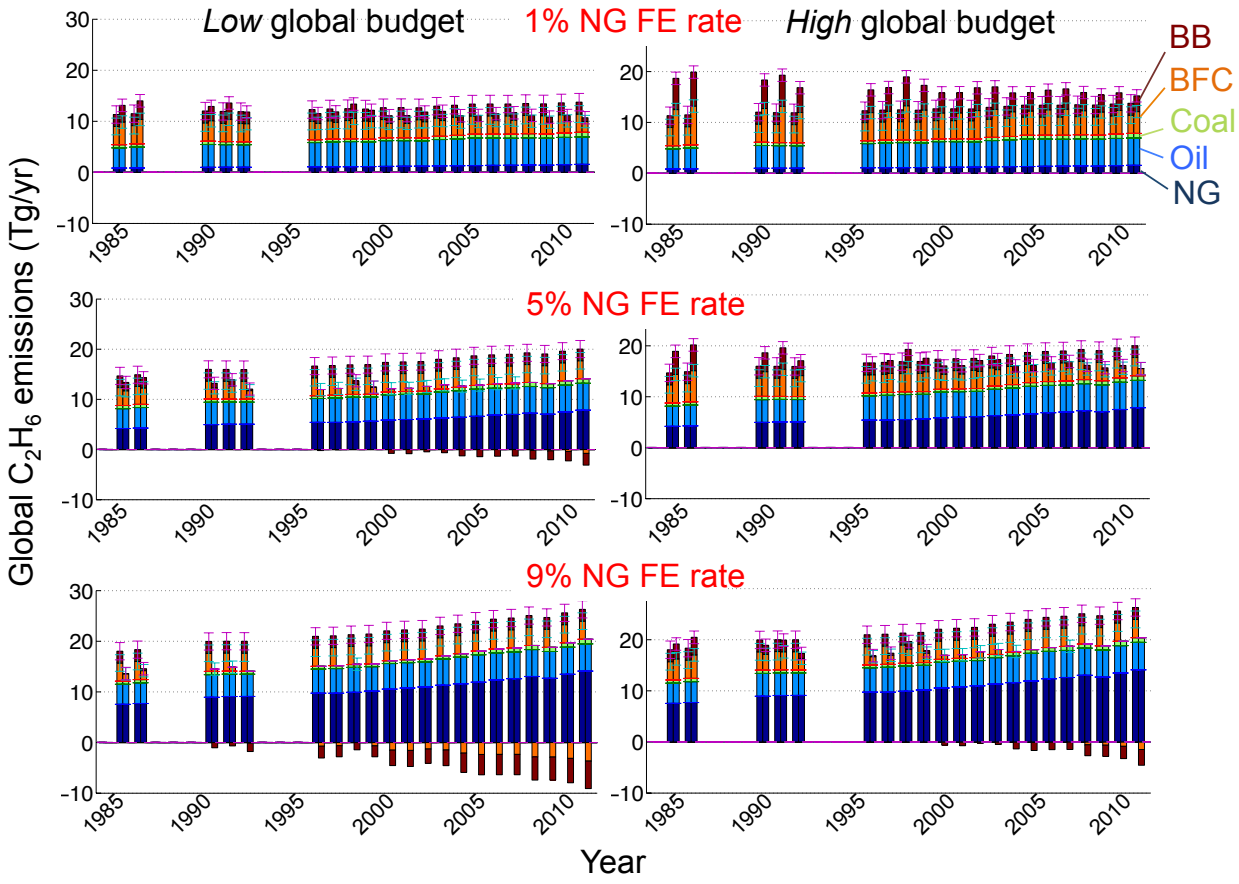

Figure B - 12: Box-model C_2H_6 inversion results using global, annually constant NG FE rates and low C_2H_6 hydrocarbon gas content. The low and high budget scenarios refer to C_2H_6 measurements scaled (Tg/ppt) using (*161*) best estimate and uncertainties, respectively. Negative *posterior* emissions indicate FF *priors* are too high to accommodate non-fossil emissions in the budget.

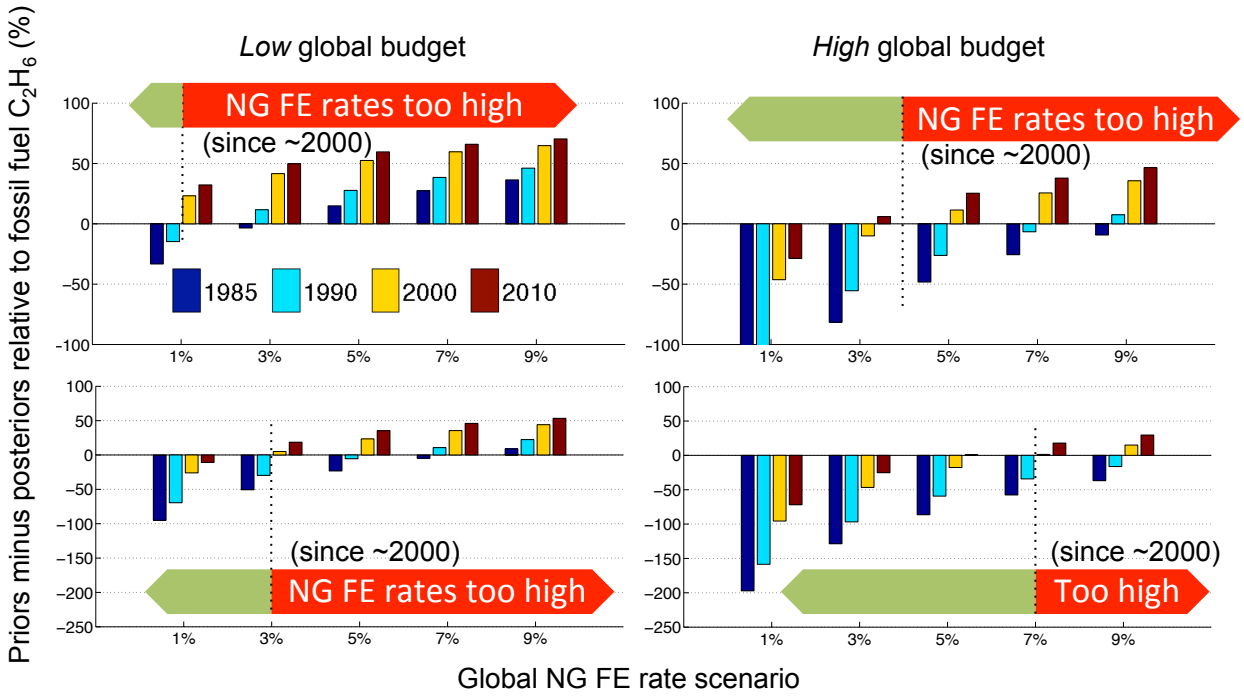

Figure B - 13: *Prior* minus *posterior* C_2H_6 emissions for different years and global C_2H_6 budgets and low C_2H_6 hydrocarbon gas content. Positive values indicate incompatibility of NG FE rates with measurements and *priors*. Upper panel: *prior* mean – *posterior* mean (best estimate NG FE). Lower panel: *prior* lower bound – *posterior* mean (upper bound NG FE).

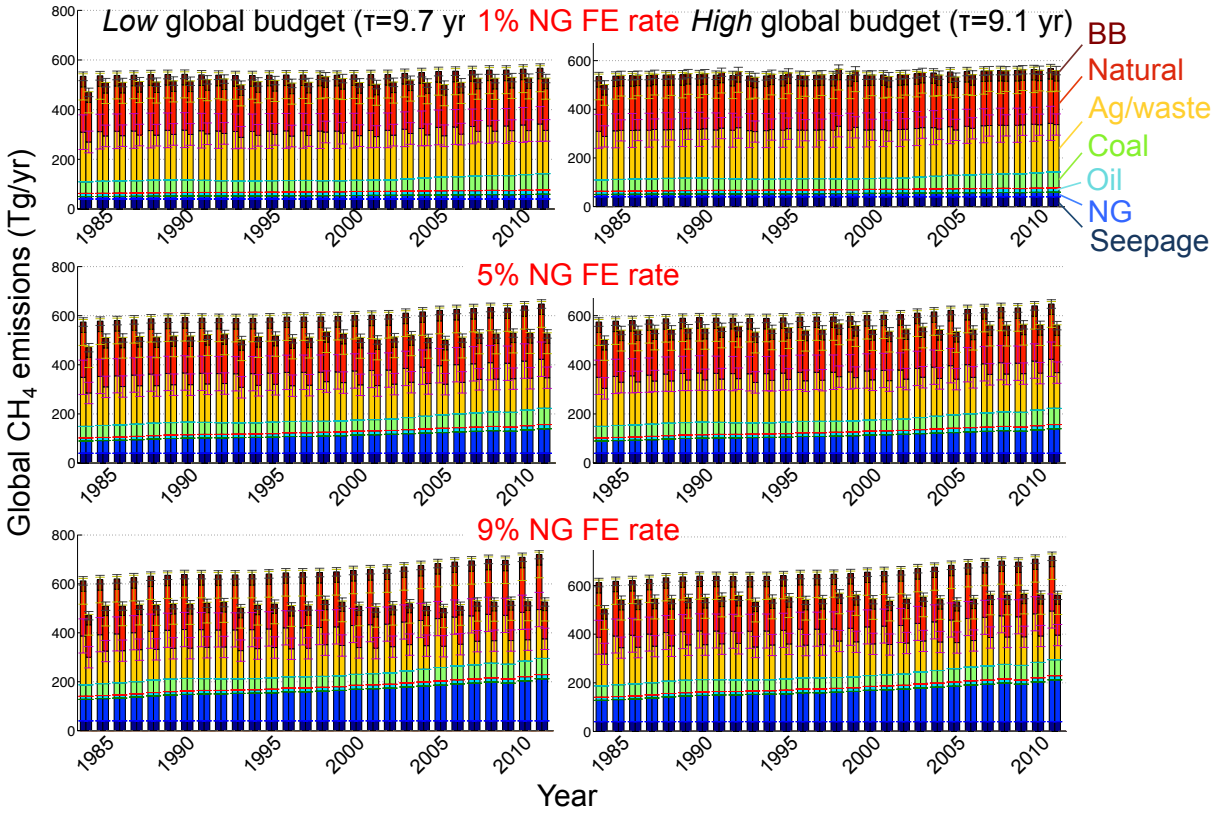

Figure B - 14: Box-model CH_4 inversion results including a global natural seepage emissions source of 40 Tg CH_4/yr.

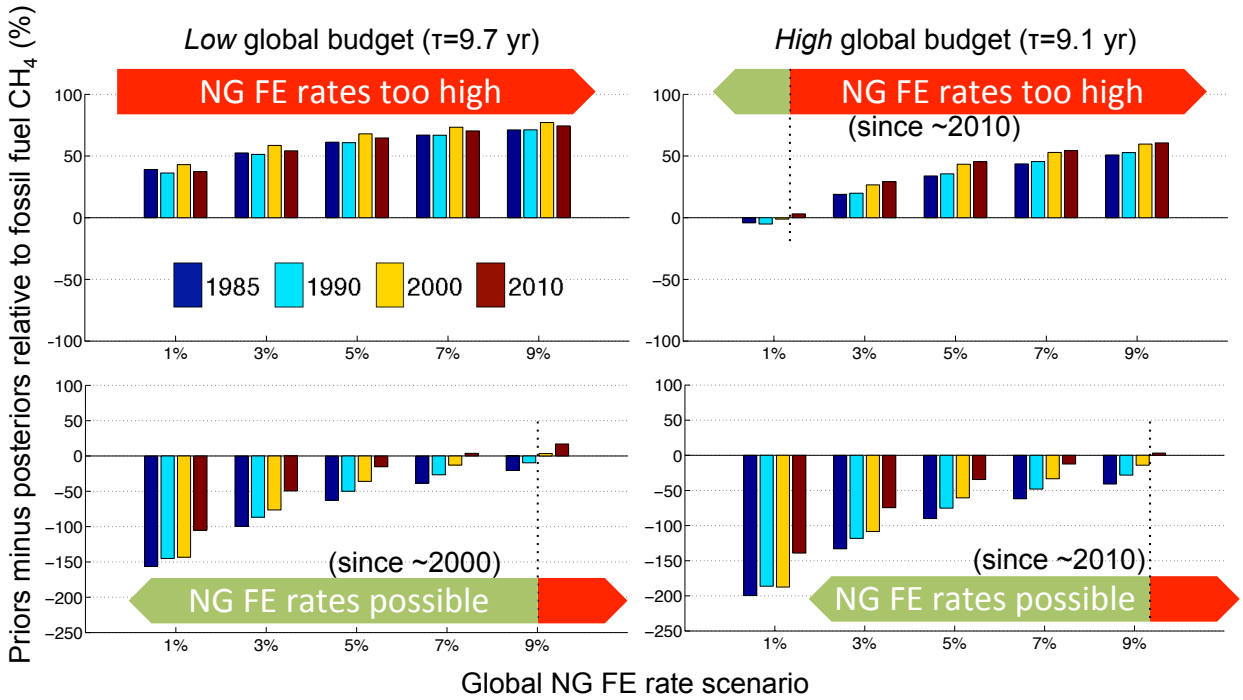

Figure B - 15: *Prior* minus *posterior* CH_4 emissions for different years including a global natural seepage emissions source of 40 Tg CH_4/yr.

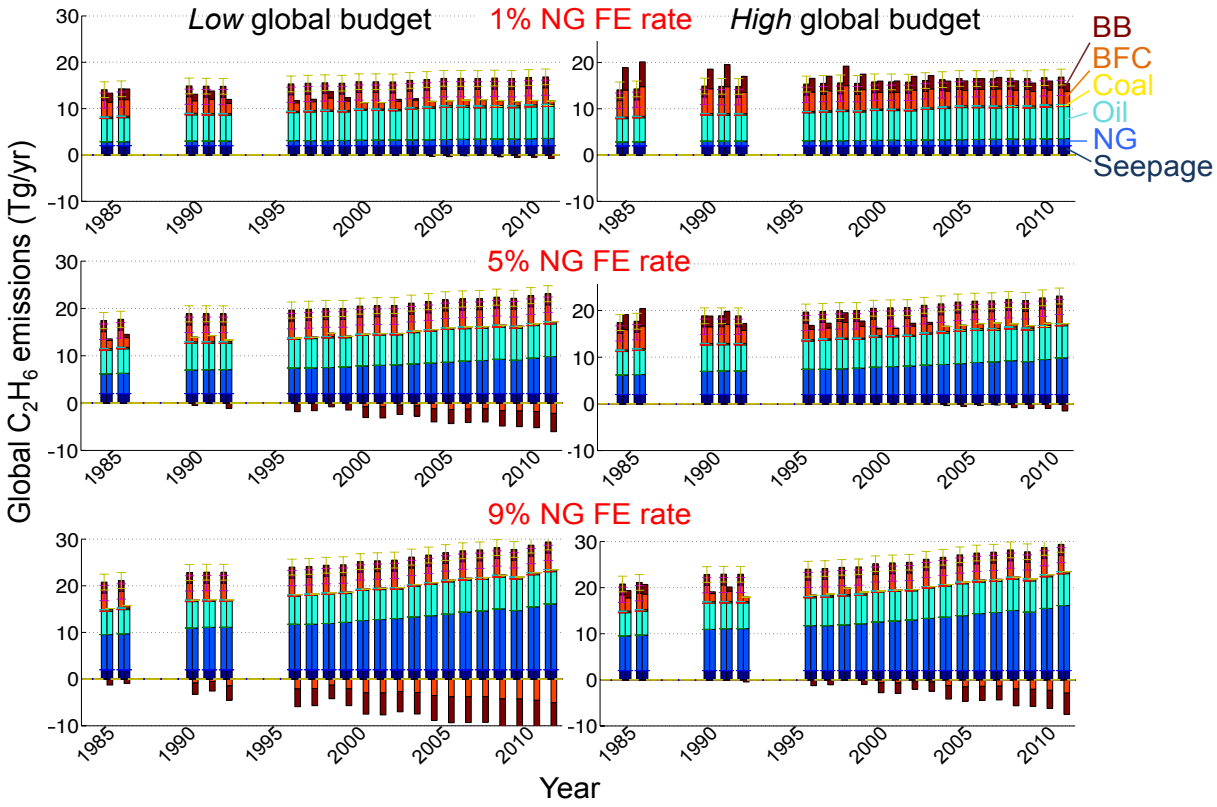

Figure B - 16: Box-model C_2H_6 inversion results for medium C_2H_6 hydrocarbon gas content including a global natural seepage emissions source of 2 Tg C_2H_6/yr.

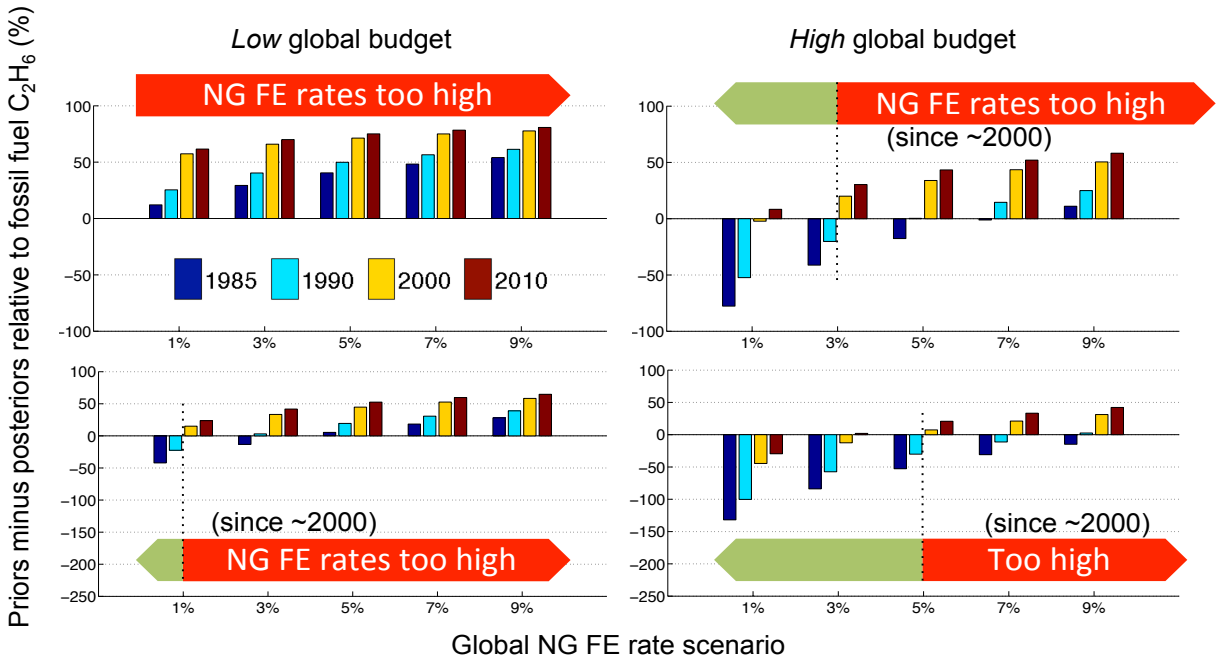

Figure B - 17: *Prior* minus *posterior* C_2H_6 emissions for different years and medium C_2H_6 hydrocarbon gas content including a global natural seepage emissions source of 2 Tg C_2H_6/yr.

Table B - 10: Summary of possible NG FE rates (in % of production) based on box-model inversion $\delta^{13}C$-CH_4 isotope mass balance under different scenarios including a natural seepage source of 40 Tg CH_4/yr and 2 Tg C_2H_6/yr. The isotope results assume that 80% of seeped CH_4 is isotopically indistinguishable from FFs (*145*).

	CH_4		C_2H_6 (medium C2 content)		$\delta^{13}C$-CH_4
Global budget	**Low**	**High**	**Low**	**High**	**-**
Best estimate (mean non-FF priors)					
2010	<1	1	<1	<1	1
2000	<1	1	<1	<2	<3
1990	<1	<2	<1	5	<3
High confidence (low non-FF priors)					
2010	7	9	<1	3	6
2000	9	>9	<1	4	<9
1990	>9	>9	<3	<9	>9

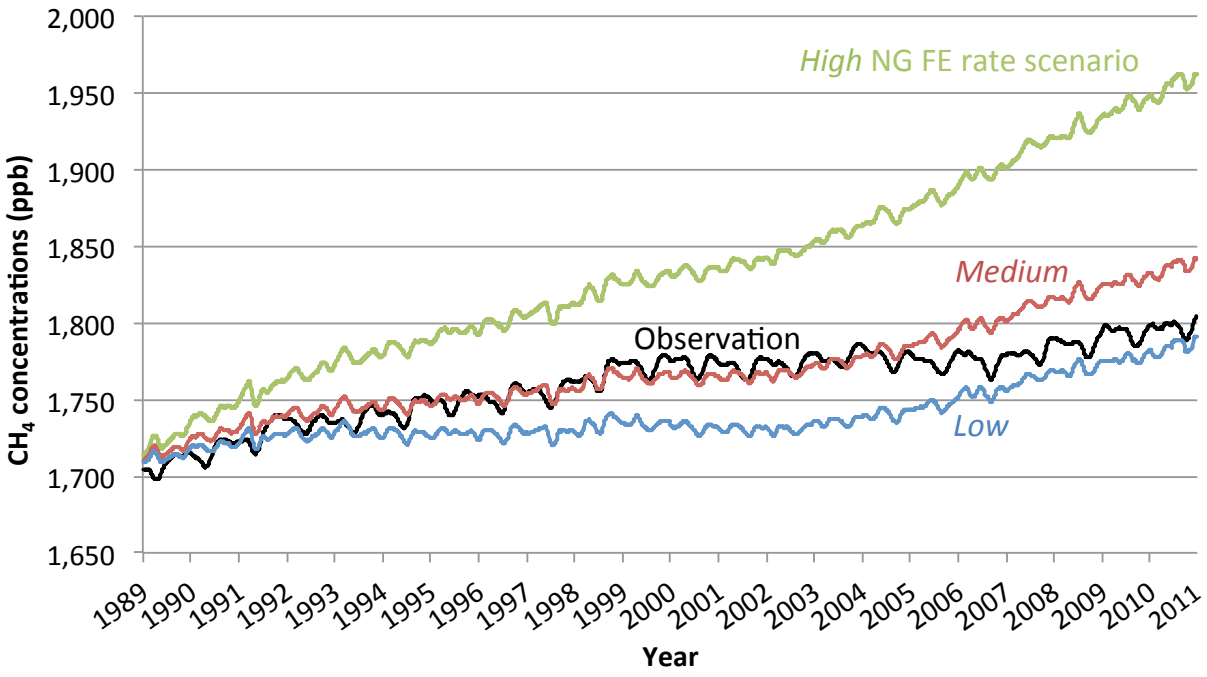

Figure B - 18: CT-CH_4 global average forward modeling results for three NG FE scenarios with decreasing FE rates over time as well as NOAA's measurements (*165*).

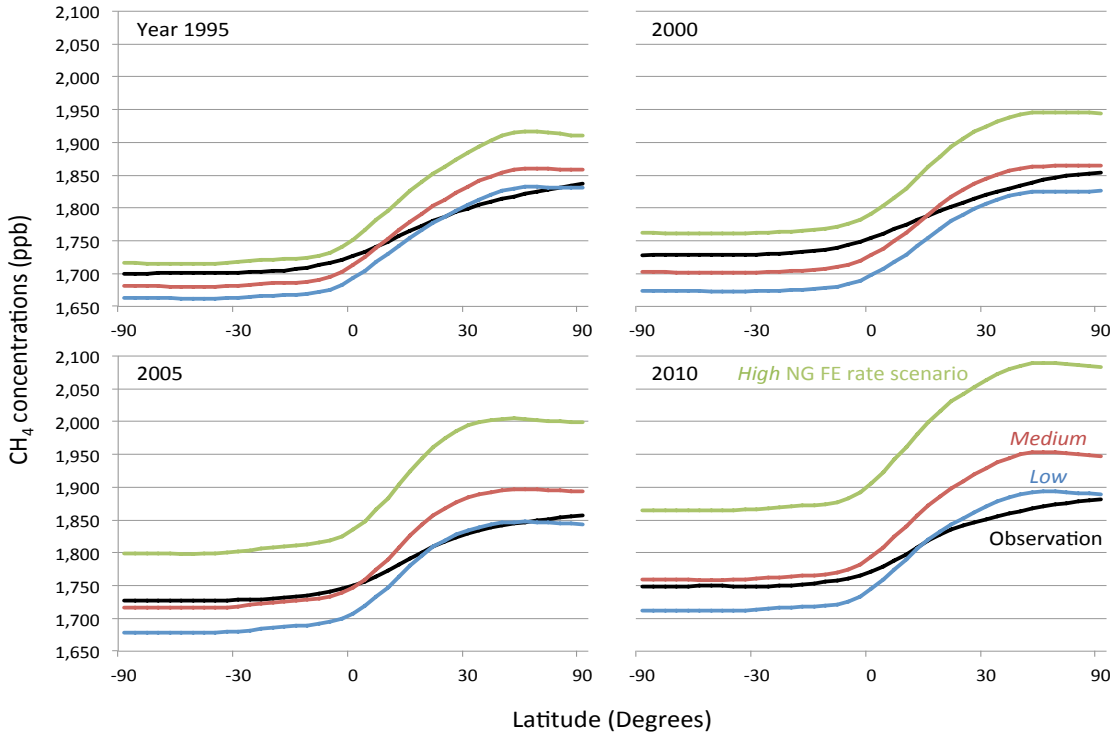

Figure B - 19: Simulated CH_4 concentrations across 41 latitudinal bands and time compared with NOAA's measurements (*174*).

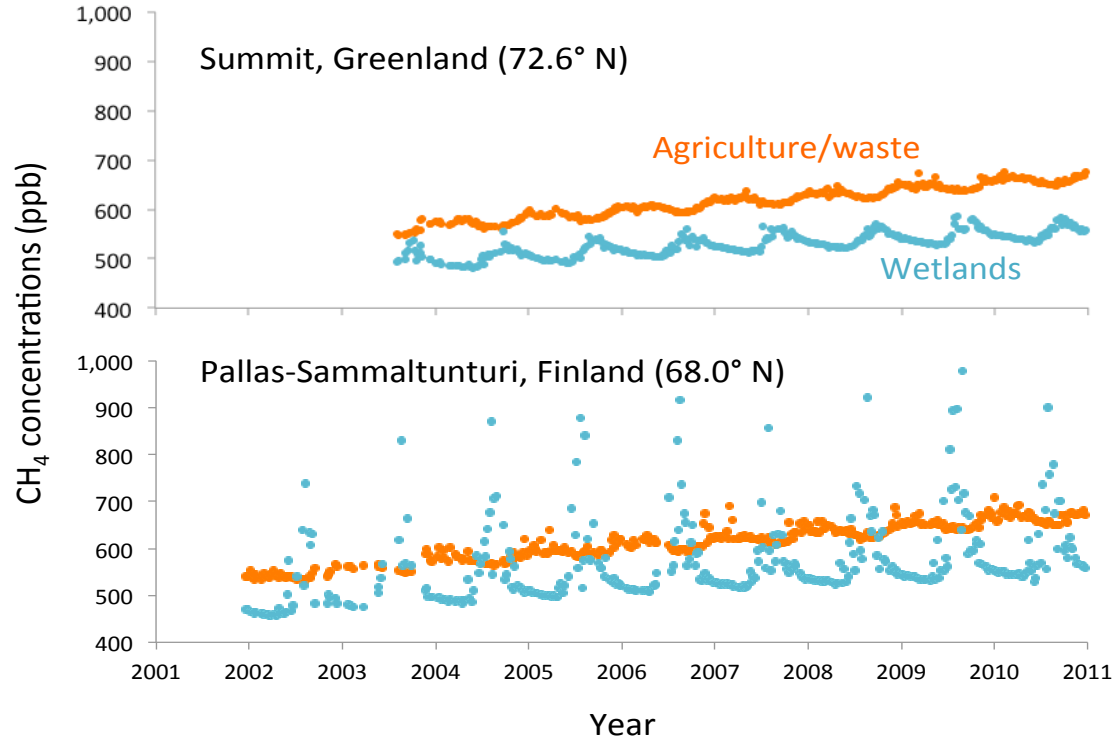

Figure B - 20: CarbonTracker-CH_4 simulated CH_4 concentration contributions from agriculture/waste and wetland sources at two sites of similar latitute as Barrow Alaska in Figure 33.

C. Matlab code

Contents

Matlab code for emissions timing model (Chapter 2)

Note: Matlab file imports emissions data generated in Excel spreadsheets using methods described in Chapter 2, but not shown here.

```
% ********************************************************
% *                                                      *
% *          Emissions timing model for Biofuels         *
% *                                                      *
% *                  Stefan Schwietzke                   *
% *                                                      *
% *              Carnegie Mellon University              *
% *                                                      *
% *  Equations for estimating RF from literature sources *
% *             as described in Chapter 2                *
% *                                                      *
% ********************************************************

% +++ Basic formulas used in the model+++
% +
% + RF_biofuel(t) = Radiative forcing due to biofuel emissions in year t [W/m2]
% +
% + add_CO2_conc(t)  = Additional atmosph. CO2 concentration due to biofuel emissions in
% + year t [ppmv]
% +
% + RF_CO2(t) = Radiative forcing from a unit increase in CO2 concentration in year t
% + [W/m2/ppmv]
% +
% + RF_CH4(t) = Radiative forcing from a unit increase in CH4 concentration in year t
% + [W/m2/ppmv]
% +
% + RF_N2O(t) = Radiative forcing from a unit increase in N2O concentration in year t
% + [W/m2/ppmv]
% +
% + Bern_model(t) = Decay function of atmospheric CO2
% +
% + Decay_CH4_function(t) = Decay function of atmospheric CH4
% +
% + Decay_N2O_function(t) = Decay function of atmospheric N2O
% +
% + Profile_CO2(t) = Emission flux in year t [g CO2]
% +
% + Profile_CH4(t) = Emission flux in year t [g CH4]
% +
% + Profile_N2O(t) = Emission flux in year t [g N2O]
% +
% + Rad_eff(CO2(t)) = Radiative efficiency of CO2 [W/m2/ppmv] as a function of
% + atmospheric CO2 concentration
% +
% + Rad_eff(CH4) = Radiative efficiency of CH4 [W/m2/ppmv]; constant (simplified)
% +
% + Rad_eff(N2O) = Radiative efficiency of N2O [W/m2/ppmv]; constant (simplified)
% +
% + CO2(t) = Future atmospheric CO2 concentration as projected by IPCC
% +
% + RF_biofuel(t) = (add_CO2_conc(t) * RF_CO2(t)) + (add_CH4_conc(t) * RF_CH4(t) +
% + (add_N2O_conc(t) * RF_N2O(t))
% +
% + add_CO2_conc(t) = Profile_CO2(t) * (10e+6 ppmv/44 g CO2/mol) / (5.14e+21 g air/28.9 g
% + air/mol); Mass of atmosph. air: Lide, David R. Handbook of Chemistry and Physics.
% + Boca Raton, FL: CRC, 1996: 14-7.
% +
% + add_CH4_conc(t) = Profile_CH4(t) * (10e+6 ppmv/16 g CH4/mol) / (5.14e+21 g air/28.9 g
% + air/mol)
% +
% + add_N2O_conc(t) = Profile_N2O(t) * (10e+6 ppmv/44 g CH4/mol) / (5.14e+21 g air/28.9 g
% + air/mol)
% +
% + RF_CO2(t) = Rad_eff(CO2(t)) * Bern_model(t)
```

```
% +
% + RF_CH4(t) = Rad_eff(CH4) * Decay_CH4_function(t)
% +
% + RF_N2O(t) = Rad_eff(N2O) * Decay_N2O_function(t)
% + Rad_eff(CO2(t) = 5.35/C [W/m2/ppmv]; C = atmosph. CO2 concentration as projected in
% + t; d(delta_F)/d(C) with delta_F = 5.35 ln(C/C0)
% +
% + Rad_eff(CH4) = 0.37 [W/m2/ppmv]
% +
% + Rad_eff(N2O) = 3.03 [W/m2/ppmv]
% +
% + Bern_model(t) = a0+a1*exp(-t/theta1)+a2*exp(-t/theta2)+a3*exp(-t/theta3); Bern model
% + with a0=0.217, a1=0.259, a2=0.338, a3=0.186, theta1=172.9, theta2=18.51, theta3=1.186
% +
% + Decay_CH4_function(t) = 1*exp(-t/theta); theta=11.3
% +
% + Decay_N2O_function(t) = 1*exp(-t/theta); theta=114
% +
% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
% +++ The impact horizon and concentration scenario will be determined by the user +++
% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

disp('Which emissions scenario should be used?')
disp('For "All LUC in first year", enter 1')
disp('For "Distributed LUC", enter 2')
emiss_scenario = input('Emissions scenario:   ')

disp('Which future IPCC atmospheric CO2 concentration scenario should be used?')
disp('For low scenario B1,     enter 1')
disp('For medium scenario A1B, enter 2')
disp('For high scenario A1FI,  enter 3')
conc_scenario = input('Concentration scenario:   ')

if emiss_scenario == 1
    Emissions = xlsread('Emissions_file.xlsx', 5, 'A1:O500') % file including emissions
profiles for ethanol (6 scenarios) and gasoline (6 scenarios), and instantaneous RF of
CO2
else
    Emissions = xlsread('Emissions_file.xlsx', 6, 'A1:O500')
end

if conc_scenario == 1
    conc_column=10
elseif conc_scenario == 2
    conc_column=11
else
    conc_column=12
end

% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
% + Create Radiative efficiency matrices, Decay matrices (using Bern model function and +
% + first order exponential decay), and conversion from emissions to concentrations     +
% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

TH = 100 % determines number of years in the emissions profile (Time Horizon),
Emissions_file has rows stacked on top of each other 5 times
t = 100 % impact time frame

% Radiative efficiency of CO2 [W/m2/ppmv]
Rad_eff_matrix_CO2=zeros(TH,t) % create matrix containing radiative efficiency of CO2 of
each year (correlating to atmospheric CO2 concentration from IPCC); each row has the same
values
for column=1:t % for all columns of the radiative efficiency matrix
    Rad_eff_matrix_CO2(:,column)=Emissions(column,conc_column) % read in the data from
file
end

% same for CH4, but CH4 is assumed constant (due to relatively fast decay)
Rad_eff_matrix_CH4=ones(TH,t)
```

```
Rad_eff_matrix_CH4=Rad_eff_matrix_CH4*Emissions(1,14)
% same for N2O
Rad_eff_matrix_N2O=ones(TH,t)
Rad_eff_matrix_N2O=Rad_eff_matrix_N2O*Emissions(1,15)
Decay_matrix_CO2_noshift=zeros(TH,t) % create matrix containing the decay rate of CO2;
this is then shifted so that decay does not apply to the first impact year of emissions
released in the 2nd year and so on
Decay_CO2=Bern_model(t) % Bern model contains the formula to calculate decay of each year
    for column=1:t % for all rows of the decay matrix
        Decay_matrix_CO2_noshift(:,column)=Decay_CO2(column-1) % calculates RF of a year
within the TH from a_C_year formula; read in the data from file
    end
Decay_matrix_CO2_shift_long=zeros(TH,2*t)% Shift Decay matrix to count emissions starting
from the year they occur
shift=0 % column in which RF starts counting for each emission year
for THyear=1:TH
    shift=shift+1
    Decay_matrix_CO2_shift_long(THyear,shift:shift+t-
    1)=Decay_matrix_CO2_noshift(THyear,:)
end
Decay_matrix_CO2=Decay_matrix_CO2_shift_long(:,1:t) % truncate the columns after t

% same for CH4 and N2O
Decay_matrix_CH4_noshift=zeros(TH,t);
Decay_matrix_N2O_noshift=zeros(TH,t)
Decay_CH4=Decay_CH4_function(t)
Decay_N2O=Decay_N2O_function(t)
    for column=1:t
        Decay_matrix_CH4_noshift(:,column)=Decay_CH4(column-1)
        Decay_matrix_N2O_noshift(:,column)=Decay_N2O(column-1)
    end
Decay_matrix_CH4_shift_long=zeros(TH,2*t)
Decay_matrix_N2O_shift_long=zeros(TH,2*t)
shift=0
for THyear=1:TH
    shift=shift+1
    Decay_matrix_CH4_shift_long(THyear,shift:shift+t-
1)=Decay_matrix_CH4_noshift(THyear,:)
    Decay_matrix_N2O_shift_long(THyear,shift:shift+t-
1)=Decay_matrix_N2O_noshift(THyear,:)
end
Decay_matrix_CH4=Decay_matrix_CH4_shift_long(:,1:t)
Decay_matrix_N2O=Decay_matrix_N2O_shift_long(:,1:t)

% Multiply both matrices for further multiplication with CO2 conc. matrices
% This is stacked 5 times to match CO2 conc. matrices
RF_per_ppmv_CO2_TH_by_t=Rad_eff_matrix_CO2.*Decay_matrix_CO2
Rad_forc_per_ppmv_CO2=vertcat(RF_per_ppmv_CO2_TH_by_t, RF_per_ppmv_CO2_TH_by_t,
RF_per_ppmv_CO2_TH_by_t, RF_per_ppmv_CO2_TH_by_t, RF_per_ppmv_CO2_TH_by_t)

% same for CH4 and N2O
RF_per_ppmv_CH4_TH_by_t=Rad_eff_matrix_CH4.*Decay_matrix_CH4
Rad_forc_per_ppmv_CH4=vertcat(RF_per_ppmv_CH4_TH_by_t, RF_per_ppmv_CH4_TH_by_t,
RF_per_ppmv_CH4_TH_by_t, RF_per_ppmv_CH4_TH_by_t, RF_per_ppmv_CH4_TH_by_t)
RF_per_ppmv_N2O_TH_by_t=Rad_eff_matrix_N2O.*Decay_matrix_N2O
Rad_forc_per_ppmv_N2O=vertcat(RF_per_ppmv_N2O_TH_by_t, RF_per_ppmv_N2O_TH_by_t,
RF_per_ppmv_N2O_TH_by_t, RF_per_ppmv_N2O_TH_by_t, RF_per_ppmv_N2O_TH_by_t)

% additional CO2 concentration [ppmv]; see above
add_CO2_conc_per_gCO2=(1e+6/44)/(5.14e+21/28.9)

% same for CH4 and N2O
add_CH4_conc_per_gCH4=(1e+6/16)/(5.14e+21/28.9)
add_N2O_conc_per_gN2O=(1e+6/44)/(5.14e+21/28.9)
```

```
% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
% ++++++++++++++ Create RF matrices for biofuel emissions  ++++++++++++++++++
% +                                                                          +
% +            M1 (LCA Model 1 = Tyner 2010 (GTAP)                           +
% +            M2 (LCA Model 2 = USEPA 2010 (FAPRI)                          +
% +            M3 (LCA Model 3 = Searchinger 2008 (FAPRI)                    +
% +                                                                          +
% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

profile_M1_CO2=zeros(5*TH,t) % will contain CO2 emissions released each year for all 3 M1
scenarios (low volume, low annualized, high volume) and 2 according gasoline scenarios
(low, high volume)
profile_M2_CO2=zeros(5*TH,t)
profile_M3_CO2=zeros(5*TH,t)
profile_M1_CH4=zeros(5*TH,t) % same for CH4
profile_M2_CH4=zeros(5*TH,t)
profile_M3_CH4=zeros(5*TH,t)
profile_M1_N2O=zeros(5*TH,t) % same for N2O
profile_M2_N2O=zeros(5*TH,t)
profile_M3_N2O=zeros(5*TH,t)
profile_reference_CO2_elec=zeros(TH,t) % emissions profile of first reference case (US
electricity emissions)
profile_reference_CO2_fuel_red=zeros(TH,t) % emissions profile of second reference case
(10% gasoline reduction)

for column=1:t % for each column of the profile
    % profile_M1_CO2 has the CO2 emissions
    % (Ethanol low [row 1:100], ann. [row 101:200], high [row 201:300], and Gasoline low
[row 301:400], high [row 401:500])
    % of each year (i.e., each loop) written in every row
    % (as a check instead of just using Emissions directly), so it can be
    % multiplied (after converting into conc.) with rad. eff. and decay matrices
    % to get RF impacts for each impact year
    profile_M1_CO2(1:5*TH,column)=Emissions(1:5*TH,1)
    % do the same for M2, M3, and reference case
    profile_M2_CO2(1:5*TH,column)=Emissions(1:5*TH,2)
    profile_M3_CO2(1:5*TH,column)=Emissions(1:5*TH,3)
    % same for CH4
    profile_M1_CH4(1:5*TH,column)=Emissions(1:5*TH,4)
    profile_M2_CH4(1:5*TH,column)=Emissions(1:5*TH,5)
    profile_M3_CH4(1:5*TH,column)=Emissions(1:5*TH,6)
    % same for N2O
    profile_M1_N2O(1:5*TH,column)=Emissions(1:5*TH,7)
    profile_M2_N2O(1:5*TH,column)=Emissions(1:5*TH,8)
    profile_M3_N2O(1:5*TH,column)=Emissions(1:5*TH,9)
    % same for reference CO2 emissions
    profile_reference_CO2_elec(1:TH,column)=Emissions(101:200,13)
    profile_reference_CO2_fuel_red(1:TH,column)=Emissions(11:110,15)
end
add_CO2_conc_M1=profile_M1_CO2.*add_CO2_conc_per_gCO2 % convert emissions in additional
CO2 concentration
add_CO2_conc_M2=profile_M2_CO2.*add_CO2_conc_per_gCO2
add_CO2_conc_M3=profile_M3_CO2.*add_CO2_conc_per_gCO2
add_CH4_conc_M1=profile_M1_CH4.*add_CH4_conc_per_gCH4
add_CH4_conc_M2=profile_M2_CH4.*add_CH4_conc_per_gCH4
add_CH4_conc_M3=profile_M3_CH4.*add_CH4_conc_per_gCH4
add_N2O_conc_M1=profile_M1_N2O.*add_N2O_conc_per_gN2O
add_N2O_conc_M2=profile_M2_N2O.*add_N2O_conc_per_gN2O
add_N2O_conc_M3=profile_M3_N2O.*add_N2O_conc_per_gN2O
add_CO2_reference_elec=profile_reference_CO2_elec.*add_CO2_conc_per_gCO2
add_CO2_reference_fuel_red=profile_reference_CO2_fuel_red.*add_CO2_conc_per_gCO2

% Multiply (element by element) CO2 concentrations matrix with Rad. eff. matrix and Decay
matrix
RF_M1_CO2=add_CO2_conc_M1.*Rad_forc_per_ppmv_CO2
RF_M2_CO2=add_CO2_conc_M2.*Rad_forc_per_ppmv_CO2
RF_M3_CO2=add_CO2_conc_M3.*Rad_forc_per_ppmv_CO2
% same for CH4
RF_M1_CH4=add_CH4_conc_M1.*Rad_forc_per_ppmv_CH4
RF_M2_CH4=add_CH4_conc_M2.*Rad_forc_per_ppmv_CH4
RF_M3_CH4=add_CH4_conc_M3.*Rad_forc_per_ppmv_CH4
% same for N2O
```

```
RF_M1_N2O=add_N2O_conc_M1.*Rad_forc_per_ppmv_N2O
RF_M2_N2O=add_N2O_conc_M2.*Rad_forc_per_ppmv_N2O
RF_M3_N2O=add_N2O_conc_M3.*Rad_forc_per_ppmv_N2O
% Add RFs from CO2, CH4, and N2O
RF_M1=RF_M1_CO2+RF_M1_CH4+RF_M1_N2O
RF_M2=RF_M2_CO2+RF_M2_CH4+RF_M2_N2O
RF_M3=RF_M3_CO2+RF_M3_CH4+RF_M3_N2O
% RF for reference case
RF_reference_CO2_elec=add_CO2_reference_elec.*RF_per_ppmv_CO2_TH_by_t
RF_reference_CO2_fuel_red=add_CO2_reference_fuel_red.*RF_per_ppmv_CO2_TH_by_t

% Add RFs of each year of the impact time (t); every line is a different
% scenario (ethanol low, ann., high) and fuel (gas. low, high)
Total_RF_M1=zeros(5,t)
Total_RF_M1(1,:)=sum(RF_M1(1:100,:),1)
Total_RF_M1(2,:)=sum(RF_M1(101:200,:),1)
Total_RF_M1(3,:)=sum(RF_M1(201:300,:),1)
Total_RF_M1(4,:)=sum(RF_M1(301:400,:),1)
Total_RF_M1(5,:)=sum(RF_M1(401:500,:),1)
% same for M2, M3, reference case
Total_RF_M2=zeros(5,t)
Total_RF_M2(1,:)=sum(RF_M2(1:100,:),1)
Total_RF_M2(2,:)=sum(RF_M2(101:200,:),1)
Total_RF_M2(3,:)=sum(RF_M2(201:300,:),1)
Total_RF_M2(4,:)=sum(RF_M2(301:400,:),1)
Total_RF_M2(5,:)=sum(RF_M2(401:500,:),1)
Total_RF_M3=zeros(5,t)
Total_RF_M3(1,:)=sum(RF_M3(1:100,:),1)
Total_RF_M3(2,:)=sum(RF_M3(101:200,:),1)
Total_RF_M3(3,:)=sum(RF_M3(201:300,:),1)
Total_RF_M3(4,:)=sum(RF_M3(301:400,:),1)
Total_RF_M3(5,:)=sum(RF_M3(401:500,:),1)
Total_RF_reference(1,:)=sum(RF_reference_CO2_elec,1)
Total_RF_reference(2,:)=sum(RF_reference_CO2_fuel_red,1)

% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
% ++++++ Calculating the Delta of RFs (biofuel-gasoline) and CRF   +++++++
% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

% RF absolute difference between biofuel and gasoline
Total_RF_diff_to_gas(1,:)=Total_RF_M1(1,:)-Total_RF_M1(4,:) % M1 low
Total_RF_diff_to_gas(2,:)=Total_RF_M1(2,:)-Total_RF_M1(4,:) % M1 annualized
Total_RF_diff_to_gas(3,:)=Total_RF_M1(3,:)-Total_RF_M1(5,:) % M1 high
Total_RF_diff_to_gas(4,:)=Total_RF_M2(1,:)-Total_RF_M2(4,:) % M2 low
Total_RF_diff_to_gas(5,:)=Total_RF_M2(2,:)-Total_RF_M2(4,:) % M2 annualized
Total_RF_diff_to_gas(6,:)=Total_RF_M2(3,:)-Total_RF_M2(5,:) % M2 high
Total_RF_diff_to_gas(7,:)=Total_RF_M3(1,:)-Total_RF_M3(4,:) % M3 low
Total_RF_diff_to_gas(8,:)=Total_RF_M3(2,:)-Total_RF_M3(4,:) % M3 annualized
Total_RF_diff_to_gas(9,:)=Total_RF_M3(3,:)-Total_RF_M3(5,:) % M3 high

% Truncate total RF vector to 30 ... 90 years (to compare with GHG balance over 30 to 100
years)
Total_RF_diff_to_gas_30=Total_RF_diff_to_gas(:,1:30)
Total_RF_diff_to_gas_40=Total_RF_diff_to_gas(:,1:40)
Total_RF_diff_to_gas_50=Total_RF_diff_to_gas(:,1:50)
Total_RF_diff_to_gas_60=Total_RF_diff_to_gas(:,1:60)
Total_RF_diff_to_gas_70=Total_RF_diff_to_gas(:,1:70)
Total_RF_diff_to_gas_80=Total_RF_diff_to_gas(:,1:80)
Total_RF_diff_to_gas_90=Total_RF_diff_to_gas(:,1:90)

% Make reference vector for plotting electricity, ethanol, and ethanol difference (EPA)
RF_reference_vector(1,:)=Total_RF_reference(1,:) % electricity
RF_reference_vector(2,:)=Total_RF_M2(3,:) % M2 high
RF_reference_vector(3,:)=Total_RF_M2(1,:) % M2 low
RF_reference_vector(4,:)=Total_RF_diff_to_gas(6,:) % M2 ethanol-gasoline high
RF_reference_vector(5,:)=Total_RF_diff_to_gas(4,:) % M2 ethanol-gasoline low
RF_reference_vector(6,:)=Total_RF_M1(3,:) % M1 high
RF_reference_vector(7,:)=Total_RF_M1(1,:) % M1 low
RF_reference_vector(8,:)=Total_RF_diff_to_gas(3,:) % M1 ethanol-gasoline high
RF_reference_vector(9,:)=Total_RF_diff_to_gas(1,:) % M1 ethanol-gasoline low
RF_reference_vector(10,:)=Total_RF_M3(3,:) % M3 high
RF_reference_vector(11,:)=Total_RF_M3(1,:) % M3 low
```

```
RF_reference_vector(12,:)=Total_RF_diff_to_gas(9,:) % M3 ethanol-gasoline high
RF_reference_vector(13,:)=Total_RF_diff_to_gas(7,:) % M3 ethanol-gasoline low
RF_reference_vector(14,:)=Total_RF_reference(2,:) % gasoline reduction

% Calculate CRF balance of biofuels relative to gasoline over 30 to 100 years (in %)
Total_RF_gas_vector=[Total_RF_M1(4,:); Total_RF_M1(4,:); Total_RF_M1(5,:);
Total_RF_M1(4,:); Total_RF_M1(4,:); Total_RF_M1(5,:); Total_RF_M1(4,:); Total_RF_M1(4,:);
Total_RF_M1(5,:)]
Total_RF_gas_vector_30=Total_RF_gas_vector(:,1:30)
Total_RF_gas_vector_40=Total_RF_gas_vector(:,1:40)
Total_RF_gas_vector_50=Total_RF_gas_vector(:,1:50)
Total_RF_gas_vector_60=Total_RF_gas_vector(:,1:60)
Total_RF_gas_vector_70=Total_RF_gas_vector(:,1:70)
Total_RF_gas_vector_80=Total_RF_gas_vector(:,1:80)
Total_RF_gas_vector_90=Total_RF_gas_vector(:,1:90)

Total_RF_balance(:,1)=sum(Total_RF_diff_to_gas_30,2)./sum(Total_RF_gas_vector_30,2)*100
Total_RF_balance(:,2)=sum(Total_RF_diff_to_gas_40,2)./sum(Total_RF_gas_vector_40,2)*100
Total_RF_balance(:,3)=sum(Total_RF_diff_to_gas_50,2)./sum(Total_RF_gas_vector_50,2)*100
Total_RF_balance(:,4)=sum(Total_RF_diff_to_gas_60,2)./sum(Total_RF_gas_vector_60,2)*100
Total_RF_balance(:,5)=sum(Total_RF_diff_to_gas_70,2)./sum(Total_RF_gas_vector_70,2)*100
Total_RF_balance(:,6)=sum(Total_RF_diff_to_gas_80,2)./sum(Total_RF_gas_vector_80,2)*100
Total_RF_balance(:,7)=sum(Total_RF_diff_to_gas_90,2)./sum(Total_RF_gas_vector_90,2)*100
Total_RF_balance(:,8)=sum(Total_RF_diff_to_gas,2)./sum(Total_RF_gas_vector,2)*100

% Calculate CRF balance of biofuels relative to electricity use and relative to 10%
gasoline reduction over 100 years (in %)
RF_reference_balance_elec=sum(RF_reference_vector,2)./sum(RF_reference_vector(1,:),2)*100
RF_reference_balance_fuel_red=sum(RF_reference_vector,2)./sum(RF_reference_vector(14,:),2
)*100

% Export data to Excel
if emiss_scenario == 1
    xlswrite('RFWI.xlsx', Total_RF_balance, 'Total_RF_balance', 'A1:H9')
else
    xlswrite('RFWI.xlsx', Total_RF_balance, 'Total_RF_balance', 'A13:H21')
end
GHG_balance=zeros(3,8)
GHG_balance(1,:)=Emissions(3,15)
GHG_balance(2,:)=Emissions(4,15)
GHG_balance(3,:)=Emissions(5,15)
xlswrite('RFWI.xlsx', GHG_balance, 'Total_RF_balance', 'A10:H12')

% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
% ++++++++++++++++++++++++++   Plotting   ++++++++++++++++++++++++++++++++
% ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

% plot RF difference for low volume scenario
RF_difference = figure;
axes1 = axes('Parent',RF_difference,'GridLineStyle','--');
% ylim(axes1,[-0.001 0.002])
plot1 = plot(Total_RF_diff_to_gas([4 1 7],:)','Parent',axes1,'Color',[0 0 0]);
set(gca,'box','off')
set(plot1(1),'DisplayName','EPA');
set(plot1(2),'LineStyle',':','DisplayName','Tyner');
set(plot1(3),'LineStyle','--','DisplayName','Hertel');
xlabel('Time since first emission release (years)')
ylabel('\DeltaRF, corn ethanol minus gasoline (W\timesm^{-2})')
% title('Delta RF, (corn ethanol minus gasoline), 15 bgal')
legend(axes1,'show');

% plot RF difference for actual and annualized emissions (EPA, low volume scenario)
RF_difference = figure;
axes1 = axes('Parent',RF_difference,'GridLineStyle','--');
% ylim(axes1,[-0.001 0.002])
plot1 = plot(Total_RF_diff_to_gas([4 5],:)','Parent',axes1,'Color',[0 0 0]);
set(gca,'box','off')
set(plot1(1),'DisplayName','EPA, actual emissions profile');
set(plot1(2),'LineStyle','--','DisplayName','EPA, annualized LUC emissions');
xlabel('Time since first emission release (years)')
ylabel('\DeltaRF, corn ethanol minus gasoline (W\timesm^{-2})')
```

```
% title('Delta RF, (corn ethanol minus gasoline), 15 bgal')
legend(axes1,'show');

% plot RF reference case
if emiss_scenario == 1
    RF_reference = figure;
axes1 = axes('Parent',RF_reference,'GridLineStyle','--');
plot2 = plot(RF_reference_vector([1 4 5 14],:)','Parent',axes1,'Color',[0 0 0]);
set(gca,'box','off')
set(plot2(1),'DisplayName','U.S. electricity consumption');
set(plot2(2),'LineStyle',':','DisplayName','\DeltaRF, Corn ethanol minus gasoline, 102
bgal/yr (EPA)');
set(plot2(3),'LineStyle','--','DisplayName','\DeltaRF, Corn ethanol minus gasoline, 15
bgal/yr (EPA)');
set(plot2(4),'DisplayName','10% U.S. gasoline reduction','LineWidth',2.5);
xlabel('Time since first emission release (years)')
ylabel('RF (W\timesm^{-2})')
% title('RF of U.S. electricity consumption and corn ethanol use')
legend(axes1,'show');
else
    RF_reference = figure;
axes1 = axes('Parent',RF_reference,'GridLineStyle','--');
plot2 = plot(RF_reference_vector([1 4 5 14],:)','Parent',axes1,'Color',[0 0 0]);
set(gca,'box','off')
set(plot2(1),'DisplayName','U.S. electricity consumption');
set(plot2(2),'LineStyle',':','DisplayName','\DeltaRF, Corn ethanol minus gasoline, 45
bgal/yr (EPA)');
set(plot2(3),'LineStyle','--','DisplayName','\DeltaRF, Corn ethanol minus gasoline, 15
bgal/yr (EPA)');
set(plot2(4),'DisplayName','10% U.S. gasoline reduction','LineWidth',2.5);
xlabel('Time since first emission release (years)')
ylabel('RF (W\timesm^{-2})')
% title('RF of U.S. electricity consumption and corn ethanol use')
legend(axes1,'show');
end

% plot RF of ethanol and gasoline individually, low volume
if emiss_scenario == 2
    RF_individual_EPA = figure;
axes1 = axes('Parent',RF_individual_EPA,'GridLineStyle','--','FontSize',16);
plot3 = plot(Total_RF_M2([1 2 4],:)','Parent',axes1,'Color',[0 0 0]);
set(gca,'box','off')
set(plot3(1),'LineStyle','--','DisplayName','Corn ethanol, actual
(EPA)','LineWidth',1.5);
set(plot3(2),'LineStyle',':','DisplayName','Corn ethanol, annualized
(EPA)','LineWidth',1.5);
set(plot3(3),'DisplayName','Gasoline (EPA)','LineWidth',1.5);
xlabel('Time since first emission release, t (years)','FontSize',16)
ylabel('RF (W\timesm^{-2})','FontSize',16)
% title('RF of corn ethanol (EPA) and gasoline, 15 bgal')
legend(axes1,'show');
RF_individual_Tyner_Hertel = figure;
axes(1) = subplot(2,1,1);
plot1 = plot(Total_RF_M1([1 4],:)','Parent',axes(1),'Color',[0 0 0]);
set(gca,'box','off')
set(plot1(1),'DisplayName','Corn ethanol (Tyner)');
set(plot1(2),'LineStyle',':','DisplayName','Gasoline (Tyner)');
ylabel('RF (W\timesm^{-2})')
legend(axes(1),'show');

axes(2) = subplot(2,1,2);
plot2 = plot(Total_RF_M3([1 4],:)','Parent',axes(2),'Color',[0 0 0]);
set(gca,'box','off')
set(plot2(1),'DisplayName','Corn ethanol (Hertel)');
set(plot2(2),'LineStyle',':','DisplayName','Gasoline (Hertel)');
xlabel('Impact time (years)')
ylabel('Emissions (g CO_{2}e/MJ)')
legend(axes(2),'show');
else
    RF_individual_all = figure;
axes(1) = subplot(3,1,1);
plot1 = plot(Total_RF_M2([1 4],:)','Parent',axes(1),'Color',[0 0 0]);
```

```
set(gca,'box','off')
set(plot1(1),'DisplayName','Corn ethanol (EPA)');
set(plot1(2),'LineStyle',':','DisplayName','Gasoline (EPA)');
ylabel('RF (W\timesm^{-2})')
legend(axes(1),'show');

axes(2) = subplot(3,1,2);
plot2 = plot(Total_RF_M1([1 4],:)','Parent',axes(2),'Color',[0 0 0]);
set(gca,'box','off')
set(plot2(1),'DisplayName','Corn ethanol (Tyner)');
set(plot2(2),'LineStyle',':','DisplayName','Gasoline (Tyner)');
ylabel('RF (W\timesm^{-2})')
legend(axes(2),'show');

axes(3) = subplot(3,1,3);
plot3 = plot(Total_RF_M3([1 4],:)','Parent',axes(3),'Color',[0 0 0]);
set(gca,'box','off')
set(plot3(1),'DisplayName','Corn ethanol (Hertel)');
set(plot3(2),'LineStyle',':','DisplayName','Gasoline (Hertel)');
xlabel('Impact time (years)')
ylabel('RF (W\timesm^{-2})')
legend(axes(3),'show');
end
```

Matlab code for generating emissions grid maps (Chapter 4)

Matlab file names are as described in Appendix B.

Note: Matlab files import emissions data generated in Excel spreadsheets using methods described in Chapter 4, but not shown here.

convert_numdata_lat.m:

```
for row = 1:1633314
    search = numdata(row,1);
    rows_found = find(numdata(:,1)==search);
    for x = 1:length(rows_found)
        numdata_new(row,1) = find(abs(lat(:,1)-search) <= 0.5e-1);
    end
end
```

convert_numdata_lon.m:

```
for row = 1:1633314
    search = numdata(row,2);
    if search == 0
        numdata_matrix_lon(row,1)=1;
    else
        rows_found = find(numdata(:,2)==search);
        for x = 1:length(rows_found)
            numdata_matrix_lon(row,1) = find(abs(lon(1,:)-search) <= 0.5e-1);
        end
    end
end
```

make_country.m

```
country_code = 'DEU'; % reads in country code, e.g., Germany (DEU)
DEU = zeros(1800,3600);
country_cell = find(ismember(country,country_code)==1); % finds rows in "country"
containing that code
y = length(country_cell); % determines length of vector above
for x = 1:y % do for each index containing the country code
    row_value = country_cell(x,1); % determine value of the index (pertaining to a grid
    cell)
    cell_lat = numdata_matrix(row_value,1); % latitudinal value of cell
    cell_lon = numdata_matrix(row_value,2); % longitudinal value of cell
    DEU(cell_lat,cell_lon) = numdata_matrix(row_value,3)/100;
end
```

add_maritime.m

```
% data
data = input('Grid: ');
max_data = max(data(:));
max_round = roundsd(max_data,1);
max_order = order(max_data);
min_data = min(nonzeros(data(:)));
min_round = roundsd(min_data,1);
min_order = order(min_data);
x=linspace(0, 359.9, 3600);
y=linspace(-89.9, 90, 1800);
[X,Y] = meshgrid(x,y);
% myscale=[10^min_order 10^max_order];
% myscale=[10^1 10^9];
myscale=[10^1 10^9];
```

```
% log plot
figure
pcolor(X,Y,log10(data))
shading flat
caxis(myscale)
h=colorbar;
% ticks_wanted=10.^(min_order:max_order);
ticks_wanted=10.^(1:max_order);
caxis(log10(myscale))
set(h,'YTick',log10(ticks_wanted));
set(h,'YTickLabel',ticks_wanted);
```

convert_emissions_oil.m (for example of oil grid maps)

```
% Convert gridded emissions from EDGAR to LCA

% Year of emissions
Year = input('Year: ');
Y = Year - 1979;
% Input file
File = input('EDGAR file: ');
EDGAR_oil_prior = File;
% read in calculated LCA emissions for each of the top 26 countries from spreadsheet
Excel_emissions_oil = xlsread('Global CH4 emissions from fossil fuels.xlsx', 17,
'C145:AH171');
Excel_emissions_oil = Excel_emissions_oil * 1e9;
% read in the emissions differences of the top 37 countries relative to EDGAR
Oil_rel_diff = xlsread('Global CH4 emissions from fossil fuels.xlsx', 17, 'AJ145:BP171');

% Gridded relative LCA emissions factors of only the top 37 countries
EF = USA*Oil_rel_diff(1,Y) + SAU*Oil_rel_diff(2,Y) + RUS*Oil_rel_diff(3,Y) +
IRN*Oil_rel_diff(5,Y) + MEX*Oil_rel_diff(6,Y) + CHN*Oil_rel_diff(7,Y) +
VEN*Oil_rel_diff(8,Y) + CAN*Oil_rel_diff(9,Y) + ARE*Oil_rel_diff(10,Y) +
GBR*Oil_rel_diff(11,Y) + NOR*Oil_rel_diff(12,Y) + NGA*Oil_rel_diff(13,Y) +
KWT*Oil_rel_diff(14,Y) + IRQ*Oil_rel_diff(15,Y) + DZA*Oil_rel_diff(16,Y) +
LBY*Oil_rel_diff(17,Y) + IDN*Oil_rel_diff(18,Y) + BRA*Oil_rel_diff(19,Y) +
EGY*Oil_rel_diff(20,Y) + AGO*Oil_rel_diff(21,Y) + OMN*Oil_rel_diff(22,Y) +
QAT*Oil_rel_diff(23,Y) + IND*Oil_rel_diff(24,Y) + ARG*Oil_rel_diff(25,Y) +
MYS*Oil_rel_diff(26,Y) + AUS*Oil_rel_diff(27,Y);
% gridded LCA emissions factors
LCA_oil_prior = EDGAR_oil_prior + EDGAR_oil_prior .* EF;
% gridded absolute LCA emissions, world
LCA_oil_prior_sum = sum(LCA_oil_prior(:));
% gridded absolute LCA emissions, top 37 countries only
LCA_oil_prior_Top26 = LCA_oil_prior .* Top_26;
LCA_oil_prior_Top26_sum = sum(LCA_oil_prior_Top26(:));
```

aggregate_to_1x1.m

```
% data
data = input('Grid: ');
max_data = max(data(:));
max_round = roundsd(max_data,1);
max_order = order(max_data);
min_data = min(nonzeros(data(:)));
min_round = roundsd(min_data,1);
min_order = order(min_data);
x=linspace(0, 359.9, 3600);
y=linspace(-89.9, 90, 1800);
[X,Y] = meshgrid(x,y);
% myscale=[10^min_order 10^max_order];
% myscale=[10^1 10^9];
myscale=[10^1 10^9];

% log plot
figure
```

```
pcolor(X,Y,log10(data))
shading flat
caxis(myscale)
h=colorbar;
% ticks_wanted=10.^(min_order:max_order);
ticks_wanted=10.^(1:max_order);
caxis(log10(myscale))
set(h,'YTick',log10(ticks_wanted));
set(h,'YTickLabel',ticks_wanted);
```

heatplot.m (for plotting grid maps)

```
% data
data = input('Grid: ');
max_data = max(data(:));
max_round = roundsd(max_data,1);
max_order = order(max_data);
min_data = min(nonzeros(data(:)));
min_round = roundsd(min_data,1);
min_order = order(min_data);
x=linspace(0, 359.9, 3600);
y=linspace(-89.9, 90, 1800);
[X,Y] = meshgrid(x,y);
% myscale=[10^min_order 10^max_order];
% myscale=[10^1 10^9];
myscale=[10^1 10^9];

% log plot
figure
pcolor(X,Y,log10(data))
shading flat
caxis(myscale)
h=colorbar;
% ticks_wanted=10.^(min_order:max_order);
ticks_wanted=10.^(1:max_order);
caxis(log10(myscale))
set(h,'YTick',log10(ticks_wanted));
set(h,'YTickLabel',ticks_wanted);
```

Matlab code for inversion box-model (Chapter 5)

Note: Matlab file imports emissions data generated in Excel spreadsheets using methods described in Chapter 4, but not shown here.

Inversion box-model:

```
% ************************************************************
% *                                                          *
% *          Global CH4 and C2H6 inversion box model         *
% *                                                          *
% *                   Stefan Schwietzke                      *
% *                                                          *
% *              Carnegie Mellon University                  *
% *                                                          *
% *   Equations for estimating posteriors from literature    *
% *          sources as described in Chapter 5               *
% *                                                          *
% ************************************************************

disp('Which emissions scenario should be used?')
disp('For "NOAA" (CH4 only), enter 1')
disp('For "LCA low to high", enter 2')
disp('For "LCA 1% NG FE rate flat", enter 3')
disp('For "LCA 3% NG FE rate flat", enter 4')
disp('For "LCA 5% NG FE rate flat", enter 5')
disp('For "LCA 7% NG FE rate flat", enter 6')
disp('For "LCA 9% NG FE rate flat", enter 7')
emiss_scenario = input('Emissions scenario:   ')
disp('For CH4 lifetime 9.7 years and ethane budget "best guess",             enter 1')
disp('For CH4 lifetime 9.7 years and ethane budget "high uncertainty bound", enter 2')
disp('For CH4 lifetime 9.1 years and ethane budget "best guess",             enter 3')
disp('For CH4 lifetime 9.1 years and ethane budget "high uncertainty bound", enter 4')
ethane_obs = input('CH4 and C2H6 budget scenario:   ')
disp('For plotting since 2000, enter 1')
disp('For plotting since 1984, enter 2')
plotting = input('Plotting choice:   ')

Emissions = xlsread('Emissions and observations inputs box model v6.xlsx', 1, 'E2:AF82');

sp = zeros(364,1); % NG_CH4, Oil_CH4, Coal_CH4, Ag/Waste_CH4, Nat_CH4, BB_CH4,
Seepage_CH4; NG_C2H6, Oil_C2H6, Coal_C2H6, BB_C2H6, Biofuel_C2H6, Seepage_C2H6 (for 1984-
2011, i.e., 28 years); so 13*28=364
Emis_total_CH4 = zeros(28,2); % Total emissions for 28 years and two species (CH4 and
C2H6)
Q = zeros(364,364); % Uncert_NG_CH4, Uncert_oil_CH4, Uncert_coal_CH4,
Uncert_Ag/Waste_CH4, Uncert_Nat_CH4, Uncert_BB_CH4; Uncert_NG_C2H6, Uncert_Oil_C2H6,
Uncert_Coal_C2H6, Uncert_BB_C2H6, Uncert_Biofuel_C2H6 (11*28 diagonal)
R = zeros(56,56); % Observational uncertainty (28 years for CH4 and 28 years for C2H6,
diagonal)

row = 1;
if emiss_scenario == 1 % NOAA
    for i = 1:28 % years
        sp(row,1) = Emissions(26,i); % Seepage_CH4 (pseudo)
        sp(row+1,1) = Emissions(15,i); % FF_CH4
        sp(row+2,1) = Emissions(16,i); % Oil_CH4 (pseudo)
        sp(row+3,1) = Emissions(17,i); % Coal_CH4 (pseudo)
        sp(row+4,1) = Emissions(1,i); % Ag/Waste_CH4
        sp(row+5,1) = Emissions(2,i); % Nat_CH4
        sp(row+6,1) = Emissions(3,i); % BB_CH4
        Emis_total_CH4(i,1) = Emissions(24,i); % total CH4 emissions
        Q(row,row) = Emissions(46,i); % Uncert_Seepage_CH4 (pseudo)
        Q(row+1,row+1) = Emissions(42,i); % Uncert_FF_CH4
        Q(row+2,row+2) = Emissions(43,i); % Uncert_oil_CH4 (pseudo)
        Q(row+3,row+3) = Emissions(44,i); % Uncert_coal_CH4 (pseudo)
        Q(row+4,row+4) = Emissions(28,i); % Uncert_Ag/Waste_CH4
        Q(row+5,row+5) = Emissions(29,i); % Uncert_Nat_CH4
```

```
        Q(row+6,row+6) = Emissions(30,i); % Uncert_BB_CH4
        R(i,i) = Emissions(48,i); % CH4
        row = row + 7;
    end
    for i = 1:28 % years
        R(i+28,i+28) = Emissions(81,i); % C2H6
    end
    % for plotting priors
    sp_Seepage_CH4(:,1) = Emissions(26,:); % Seepage_CH4
    sp_NG_CH4(:,1) = Emissions(15,:); % FF_CH4
    sp_Oil_CH4(:,1) = Emissions(16,:); % Oil_CH4 (pseudo)
    sp_Coal_CH4(:,1) = Emissions(17,:); % Coal_CH4 (pseudo)
    sp_AW_CH4(:,1) = Emissions(1,:); % Ag/Waste_CH4
    sp_Nat_CH4(:,1) = Emissions(2,:); % Nat_CH4
    sp_BB_CH4(:,1) = Emissions(3,:); % BB_CH4
    Q_NG_CH4(:,1) = Emissions(39,:); % Uncert_FF
    Q_Oil_CH4(:,1) = Emissions(40,:); % Uncert_oil_CH4 (pseudo)
    Q_Coal_CH4(:,1) = Emissions(41,:); % Uncert_coal_CH4 (pseudo)
    Q_AW_CH4(:,1) = Emissions(25,:); % Uncert_Ag/Waste_CH4
    Q_Nat_CH4(:,1) = Emissions(26,:); % Uncert_Nat_CH4
    Q_BB_CH4(:,1) = Emissions(27,:); % Uncert_BB_CH4
else
    if emiss_scenario == 2 % LCA low to high
        for i = 1:28 % years
            sp(row,1) = Emissions(25,i); % Seepage_CH4
            sp(row+1,1) = Emissions(7,i); % NG_CH4
            sp(row+2,1) = Emissions(13,i); % Oil_CH4
            sp(row+3,1) = Emissions(14,i); % Coal_CH4
            sp(row+4,1) = Emissions(4,i); % Ag/Waste_CH4
            sp(row+5,1) = Emissions(5,i); % Nat_CH4
            sp(row+6,1) = Emissions(6,i); % BB_CH4
            Emis_total_CH4(i,1) = Emissions(18,i); % total CH4 emissions
            Q(row,row) = Emissions(45,i); % Uncert_Seepage_CH4
            Q(row+1,row+1) = Emissions(34,i); % Uncert_NG_CH4
            Q(row+2,row+2) = Emissions(40,i); % Uncert_oil_CH4
            Q(row+3,row+3) = Emissions(41,i); % Uncert_coal_CH4
            Q(row+4,row+4) = Emissions(31,i); % Uncert_Ag/Waste_CH4
            Q(row+5,row+5) = Emissions(32,i); % Uncert_Nat_CH4
            Q(row+6,row+6) = Emissions(33,i); % Uncert_BB_CH4
            R(i,i) = Emissions(48,i);
            row = row + 7;
        end
        row = 197;
        for i = 1:28 % years
            sp(row,1) = Emissions(65,i); % Seepage_C2H6
            sp(row+1,1) = Emissions(51,i); % NG_C2H6
            sp(row+2,1) = Emissions(57,i); % Oil_C2H6
            sp(row+3,1) = Emissions(58,i); % Coal_C2H6
            sp(row+4,1) = Emissions(49,i); % BB_C2H6
            sp(row+5,1) = Emissions(50,i); % Biofuel_C2H6
            Emis_total_C2H6(i,1) = Emissions(59,i); % total C2H6 emissions
            Q(row,row) = Emissions(78,i); % Uncert_Seepage_C2H6
            Q(row+1,row+1) = Emissions(70,i); % Uncert_NG_C2H6
            Q(row+2,row+2) = Emissions(76,i); % Uncert_oil_C2H6
            Q(row+3,row+3) = Emissions(77,i); % Uncert_coal_C2H6
            Q(row+4,row+4) = Emissions(68,i); % Uncert_BB_C2H6
            Q(row+5,row+5) = Emissions(69,i); % Uncert_Biofuel_C2H6
            R(i+28,i+28) = Emissions(81,i);
            row = row + 6;
        end
        % for plotting priors
        sp_Seepage_CH4(:,1) = Emissions(25,:); % Seepage_CH4
        sp_NG_CH4(:,1) = Emissions(7,:); % NG_CH4
        sp_Oil_CH4(:,1) = Emissions(13,:); % Oil_CH4
        sp_Coal_CH4(:,1) = Emissions(14,:); % Coal_CH4
        sp_AW_CH4(:,1) = Emissions(4,:); % Ag/Waste_CH4
        sp_Nat_CH4(:,1) = Emissions(5,:); % Nat_CH4
        sp_BB_CH4(:,1) = Emissions(6,:); % BB_CH4
        sp_Seepage_C2H6(:,1) = Emissions(65,:); % Seepage_C2H6
        sp_NG_C2H6(:,1) = Emissions(51,:); % NG_C2H6
        sp_Oil_C2H6(:,1) = Emissions(57,:); % Oil_C2H6
        sp_Coal_C2H6(:,1) = Emissions(58,:); % Coal_C2H6
        sp_BB_C2H6(:,1) = Emissions(49,:); % BB_C2H6
```

```
    sp_Biofuel_C2H6(:,1) = Emissions(50,:); % Biofuel_C2H6
    Q_Seepage_CH4(:,1) = Emissions(45,:); % Uncert_Seepage_CH4
    Q_NG_CH4(:,1) = Emissions(34,:); % Uncert_NG_CH4
    Q_Oil_CH4(:,1) = Emissions(40,:); % Uncert_oil_CH4
    Q_Coal_CH4(:,1) = Emissions(41,:); % Uncert_coal_CH4
    Q_AW_CH4(:,1) = Emissions(31,:); % Uncert_Ag/Waste_CH4
    Q_Nat_CH4(:,1) = Emissions(32,:); % Uncert_Nat_CH4
    Q_BB_CH4(:,1) = Emissions(33,:); % Uncert_BB_CH4
    Q_Seepage_C2H6(:,1) = Emissions(78,:); % Uncert_Seepage_C2H6
    Q_NG_C2H6(:,1) = Emissions(70,:); % Uncert_NG_C2H6
    Q_Oil_C2H6(:,1) = Emissions(76,:); % Uncert_oil_C2H6
    Q_Coal_C2H6(:,1) = Emissions(77,:); % Uncert_coal_C2H6
    Q_BB_C2H6(:,1) = Emissions(68,:); % Uncert_BB_C2H6
    Q_Biofuel_C2H6(:,1) = Emissions(69,:); % Uncert_Biofuel_C2H6
else
    if emiss_scenario == 3 % LCA 1% NG FE rate flat
        for i = 1:28 % years
            sp(row,1) = Emissions(25,i); % Seepage_CH4
            sp(row+1,1) = Emissions(8,i); % NG_CH4
            sp(row+2,1) = Emissions(13,i); % Oil_CH4
            sp(row+3,1) = Emissions(14,i); % Coal_CH4
            sp(row+4,1) = Emissions(4,i); % Ag/Waste_CH4
            sp(row+5,1) = Emissions(5,i); % Nat_CH4
            sp(row+6,1) = Emissions(6,i); % BB_CH4
            Emis_total_CH4(i,1) = Emissions(19,i); % total CH4 emissions
            Q(row,row) = Emissions(45,i); % Uncert_Seepage_CH4
            Q(row+1,row+1) = Emissions(35,i); % Uncert_NG_CH4
            Q(row+2,row+2) = Emissions(40,i); % Uncert_oil_CH4
            Q(row+3,row+3) = Emissions(41,i); % Uncert_coal_CH4
            Q(row+4,row+4) = Emissions(31,i); % Uncert_Ag/Waste_CH4
            Q(row+5,row+5) = Emissions(32,i); % Uncert_Nat_CH4
            Q(row+6,row+6) = Emissions(33,i); % Uncert_BB_CH4
            R(i,i) = Emissions(48,i);
            row = row + 7;
        end
        row = 197;
        for i = 1:28 % years
            sp(row,1) = Emissions(65,i); % Seepage_C2H6
            sp(row+1,1) = Emissions(52,i); % NG_C2H6
            sp(row+2,1) = Emissions(57,i); % Oil_C2H6
            sp(row+3,1) = Emissions(58,i); % Coal_C2H6
            sp(row+4,1) = Emissions(49,i); % BB_C2H6
            sp(row+5,1) = Emissions(50,i); % Biofuel_C2H6
            Emis_total_C2H6(i,1) = Emissions(60,i); % total C2H6 emissions
            Q(row,row) = Emissions(78,i); % Uncert_Seepage_C2H6
            Q(row+1,row+1) = Emissions(71,i); % Uncert_NG_C2H6
            Q(row+2,row+2) = Emissions(76,i); % Uncert_oil_C2H6
            Q(row+3,row+3) = Emissions(77,i); % Uncert_coal_C2H6
            Q(row+4,row+4) = Emissions(68,i); % Uncert_BB_C2H6
            Q(row+5,row+5) = Emissions(69,i); % Uncert_Biofuel_C2H6
            R(i+28,i+28) = Emissions(81,i);
            row = row + 6;
        end
        % for plotting priors
        sp_Seepage_CH4(:,1) = Emissions(25,:); % Seepage_CH4
        sp_NG_CH4(:,1) = Emissions(8,:); % NG_CH4
        sp_Oil_CH4(:,1) = Emissions(13,:); % Oil_CH4
        sp_Coal_CH4(:,1) = Emissions(14,:); % Coal_CH4
        sp_AW_CH4(:,1) = Emissions(4,:); % Ag/Waste_CH4
        sp_Nat_CH4(:,1) = Emissions(5,:); % Nat_CH4
        sp_BB_CH4(:,1) = Emissions(6,:); % BB_CH4
        sp_Seepage_C2H6(:,1) = Emissions(65,:); % Seepage_C2H6
        sp_NG_C2H6(:,1) = Emissions(52,:); % NG_C2H6
        sp_Oil_C2H6(:,1) = Emissions(57,:); % Oil_C2H6
        sp_Coal_C2H6(:,1) = Emissions(58,:); % Coal_C2H6
        sp_BB_C2H6(:,1) = Emissions(49,:); % BB_C2H6
        sp_Biofuel_C2H6(:,1) = Emissions(50,:); % Biofuel_C2H6
        Q_Seepage_CH4(:,1) = Emissions(45,:); % Uncert_Seepage_CH4
        Q_NG_CH4(:,1) = Emissions(35,:); % Uncert_NG_CH4
        Q_Oil_CH4(:,1) = Emissions(40,:); % Uncert_oil_CH4
        Q_Coal_CH4(:,1) = Emissions(41,:); % Uncert_coal_CH4
        Q_AW_CH4(:,1) = Emissions(31,:); % Uncert_Ag/Waste_CH4
        Q_Nat_CH4(:,1) = Emissions(32,:); % Uncert_Nat_CH4
```

```
            Q_BB_CH4(:,1) = Emissions(33,:); % Uncert_BB_CH4
            Q_Seepage_C2H6(:,1) = Emissions(78,:); % Uncert_Seepage_C2H6
            Q_NG_C2H6(:,1) = Emissions(71,:); % Uncert_NG_C2H6
            Q_Oil_C2H6(:,1) = Emissions(76,:); % Uncert_oil_C2H6
            Q_Coal_C2H6(:,1) = Emissions(77,:); % Uncert_coal_C2H6
            Q_BB_C2H6(:,1) = Emissions(68,:); % Uncert_BB_C2H6
            Q_Biofuel_C2H6(:,1) = Emissions(69,:); % Uncert_Biofuel_C2H6
        else
            if emiss_scenario == 4 % LCA 3% NG FE rate flat
[…] (code truncated here; repeats until emissions scenario 7 (9% NG FE)
end

if ethane_obs == 1
    z = xlsread('Emissions and observations inputs box model v6.xlsx', 2, 'C1:C56');
else
    if ethane_obs == 2
    z = xlsread('Emissions and observations inputs box model v6.xlsx', 2, 'D1:D56');
    else
        if ethane_obs == 3
        z = xlsread('Emissions and observations inputs box model v6.xlsx', 2, 'E1:E56');
        else
            z = xlsread('Emissions and observations inputs box model v6.xlsx', 2,
            'F1:F56');
        end
    end
end

H = zeros(56,364);
H_value_CH4 = ones(1,7); % 7 sources (Seepage, NG, oil, coal, ag/waste, nat, BB)
column = 1;
for i = 1:28
    H(i,column:column+6) = H_value_CH4;
    column = column + 7;
end
H_value_C2H6 = ones(1,6); % 6 sources (Seepage, NG, oil, coal, BB, biofuel)
column = 197;
for i = 29:56
    H(i,column:column+5) = H_value_C2H6;
    column = column + 6;
end

Ht = transpose(H);

s_post = sp + Q * Ht * (R + H * Q * Ht)^(-1) * (z - H * sp);
s_post_Seepage_CH4 = zeros(28,1);
s_post_NG_CH4 = zeros(28,1);
s_post_Oil_CH4 = zeros(28,1);
s_post_Coal_CH4 = zeros(28,1);
s_post_AW_CH4 = zeros(28,1);
s_post_Nat_CH4 = zeros(28,1);
s_post_BB_CH4 = zeros(28,1);
s_post_sum_CH4 = zeros(28,1);
s_post_Seepage_C2H6 = zeros(28,1);
s_post_NG_C2H6 = zeros(28,1);
s_post_Oil_C2H6 = zeros(28,1);
s_post_Coal_C2H6 = zeros(28,1);
s_post_BB_C2H6 = zeros(28,1);
s_post_Biofuel_C2H6 = zeros(28,1);
s_post_sum_C2H6 = zeros(28,1);
Q_post = Q - Q * Ht * (R + H * Q * Ht)^(-1) * H * Q;
Q_post_Seepage_CH4 = zeros(28,1);
Q_post_NG_CH4 = zeros(28,1);
Q_post_Oil_CH4 = zeros(28,1);
Q_post_Coal_CH4 = zeros(28,1);
Q_post_AW_CH4 = zeros(28,1);
Q_post_Nat_CH4 = zeros(28,1);
Q_post_BB_CH4 = zeros(28,1);
Q_post_Seepage_C2H6 = zeros(28,1);
Q_post_NG_C2H6 = zeros(28,1);
Q_post_Oil_C2H6 = zeros(28,1);
Q_post_Coal_C2H6 = zeros(28,1);
Q_post_BB_C2H6 = zeros(28,1);
Q_post_Biofuel_C2H6 = zeros(28,1);
```

```
row = 1; % CH4
for i = 1:28 % years
    s_post_Seepage_CH4(i,1) = s_post(row,1);
    s_post_NG_CH4(i,1) = s_post(row+1,1);
    s_post_Oil_CH4(i,1) = s_post(row+2,1);
    s_post_Coal_CH4(i,1) = s_post(row+3,1);
    s_post_AW_CH4(i,1) = s_post(row+4,1);
    s_post_Nat_CH4(i,1) = s_post(row+5,1);
    s_post_BB_CH4(i,1) = s_post(row+6,1);
    s_post_sum_CH4(i,1) = s_post_Seepage_CH4(i,1) + s_post_NG_CH4(i,1) +
    s_post_Oil_CH4(i,1) + s_post_Coal_CH4(i,1) + s_post_AW_CH4(i,1) + s_post_Nat_CH4(i,1)
    + s_post_BB_CH4(i,1);
    Q_post_Seepage_CH4(i,1) = Q_post(row,row);
    Q_post_NG_CH4(i,1) = Q_post(row+1,row+1);
    Q_post_Oil_CH4(i,1) = Q_post(row+2,row+2);
    Q_post_Coal_CH4(i,1) = Q_post(row+3,row+3);
    Q_post_AW_CH4(i,1) = Q_post(row+4,row+4);
    Q_post_Nat_CH4(i,1) = Q_post(row+5,row+5);
    Q_post_BB_CH4(i,1) = Q_post(row+6,row+6);
    row = row + 7; % 7 sources (Seepage, NG, oil, coal, ag/waste, nat, BB)
end
row = 197; % C2H6
for i = 1:28 % years
    s_post_Seepage_C2H6(i,1) = s_post(row,1);
    s_post_NG_C2H6(i,1) = s_post(row+1,1);
    s_post_Oil_C2H6(i,1) = s_post(row+2,1);
    s_post_Coal_C2H6(i,1) = s_post(row+3,1);
    s_post_BB_C2H6(i,1) = s_post(row+4,1);
    s_post_Biofuel_C2H6(i,1) = s_post(row+5,1);
    s_post_sum_C2H6(i,1) = s_post_Seepage_C2H6(i,1) + s_post_NG_C2H6(i,1) +
    s_post_Oil_C2H6(i,1) + s_post_Coal_C2H6(i,1) + s_post_BB_C2H6(i,1) +
    s_post_Biofuel_C2H6(i,1);
    Q_post_Seepage_C2H6(i,1) = Q_post(row,row);
    Q_post_NG_C2H6(i,1) = Q_post(row+1,row);
    Q_post_Oil_C2H6(i,1) = Q_post(row+2,row+1);
    Q_post_Coal_C2H6(i,1) = Q_post(row+3,row+2);
    Q_post_BB_C2H6(i,1) = Q_post(row+4,row+3);
    Q_post_Biofuel_C2H6(i,1) = Q_post(row+5,row+4);
    row = row + 6; % 6 sources (Seepage, NG, oil, coal, BB, biofuel)
end

% write CH4 priors and posteriors in (84,7) matrix (7 columns for 7 emissions sources
% and 28 years for prior and posterior each + 1 empty row each year as space between
years)

s_CH4 = zeros(84,7);
row = 1;
for i = 1:28 % years
    s_CH4(row,1) = sp_Seepage_CH4(i);
    s_CH4(row,2) = sp_NG_CH4(i);
    s_CH4(row,3) = sp_Oil_CH4(i);
    s_CH4(row,4) = sp_Coal_CH4(i);
    s_CH4(row,5) = sp_AW_CH4(i);
    s_CH4(row,6) = sp_Nat_CH4(i);
    s_CH4(row,7) = sp_BB_CH4(i);
    s_CH4(row+1,1) = s_post_Seepage_CH4(i);
    s_CH4(row+1,2) = s_post_NG_CH4(i);
    s_CH4(row+1,3) = s_post_Oil_CH4(i);
    s_CH4(row+1,4) = s_post_Coal_CH4(i);
    s_CH4(row+1,5) = s_post_AW_CH4(i);
    s_CH4(row+1,6) = s_post_Nat_CH4(i);
    s_CH4(row+1,7) = s_post_BB_CH4(i);
    row = row + 3;
end

if plotting == 1
    s_CH4 = s_CH4(49:84,:);
end

% write C2H6 priors and posteriors in (84,6) matrix (6 columns for 6 emissions sources
% and 28 years for prior and posterior each + 1 empty row each year as space between
years)
```

```
if emiss_scenario ~= 1 % is not 1 (don't need C2H6 for NOAA emissions)

s_C2H6 = zeros(84,6);
row = 1;
for i = 1:28 % years
    s_C2H6(row,1) = sp_Seepage_C2H6(i);
    s_C2H6(row,2) = sp_NG_C2H6(i);
    s_C2H6(row,3) = sp_Oil_C2H6(i);
    s_C2H6(row,4) = sp_Coal_C2H6(i);
    s_C2H6(row,5) = sp_BB_C2H6(i);
    s_C2H6(row,6) = sp_Biofuel_C2H6(i);
    s_C2H6(row+1,1) = s_post_Seepage_C2H6(i);
    s_C2H6(row+1,2) = s_post_NG_C2H6(i);
    s_C2H6(row+1,3) = s_post_Oil_C2H6(i);
    s_C2H6(row+1,4) = s_post_Coal_C2H6(i);
    s_C2H6(row+1,5) = s_post_BB_C2H6(i);
    s_C2H6(row+1,6) = s_post_Biofuel_C2H6(i);
    row = row + 3;
end

if plotting == 1
    s_C2H6 = s_C2H6(49:84,:);
end

end % end if

% calculate CH4 uncertainties for checking with spreadsheet (priors) and output graphs
(posteriors)

Uncert_prior_Seepage_CH4 = zeros(28,2);
Uncert_prior_NG_CH4 = zeros(28,2);
Uncert_prior_Oil_CH4 = zeros(28,2);
Uncert_prior_Coal_CH4 = zeros(28,2);
Uncert_prior_AW_CH4 = zeros(28,2);
Uncert_prior_Nat_CH4 = zeros(28,2);
Uncert_prior_BB_CH4 = zeros(28,2);
Uncert_post_Seepage_CH4 = zeros(28,2);
Uncert_post_NG_CH4 = zeros(28,2);
Uncert_post_Oil_CH4 = zeros(28,2);
Uncert_post_Coal_CH4 = zeros(28,2);
Uncert_post_AW_CH4 = zeros(28,2);
Uncert_post_Nat_CH4 = zeros(28,2);
Uncert_post_BB_CH4 = zeros(28,2);
Uncertainty_CH4 = zeros(84,7); % difference high-mean (or mean-low) for 7 columns (7
emissions sources) and 84 rows (28 years for prior and posterior each + 1 empty row each
year as space between years
row = 1;

for i = 1:28 % years
Uncert_prior_Seepage_CH4(i,:) = norminv([0.05 0.95], sp_Seepage_CH4(i),
sqrt(Q_Seepage_CH4(i)));
    Uncert_prior_NG_CH4(i,:) = norminv([0.05 0.95], sp_NG_CH4(i), sqrt(Q_NG_CH4(i)));
    Uncert_prior_Oil_CH4(i,:) = norminv([0.05 0.95], sp_Oil_CH4(i), sqrt(Q_Oil_CH4(i)));
    Uncert_prior_Coal_CH4(i,:) = norminv([0.05 0.95], sp_Coal_CH4(i),
    sqrt(Q_Coal_CH4(i)));
    Uncert_prior_AW_CH4(i,:) = norminv([0.05 0.95], sp_AW_CH4(i), sqrt(Q_AW_CH4(i)));
    Uncert_prior_Nat_CH4(i,:) = norminv([0.05 0.95], sp_Nat_CH4(i), sqrt(Q_Nat_CH4(i)));
    Uncert_prior_BB_CH4(i,:) = norminv([0.05 0.95], sp_BB_CH4(i), sqrt(Q_BB_CH4(i)));

    Uncert_post_Seepage_CH4(i,:) = norminv([0.05 0.95], s_post_Seepage_CH4(i),
    sqrt(Q_post_Seepage_CH4(i)));
    Uncert_post_NG_CH4(i,:) = norminv([0.05 0.95], s_post_NG_CH4(i),
    sqrt(Q_post_NG_CH4(i)));
    Uncert_post_Oil_CH4(i,:) = norminv([0.05 0.95], s_post_Oil_CH4(i),
    sqrt(Q_post_Oil_CH4(i)));
    Uncert_post_Coal_CH4(i,:) = norminv([0.05 0.95], s_post_Coal_CH4(i),
    sqrt(Q_post_Coal_CH4(i)));
    Uncert_post_AW_CH4(i,:) = norminv([0.05 0.95], s_post_AW_CH4(i),
    sqrt(Q_post_AW_CH4(i)));
    Uncert_post_Nat_CH4(i,:) = norminv([0.05 0.95], s_post_Nat_CH4(i),
    sqrt(Q_post_Nat_CH4(i)));
    Uncert_post_BB_CH4(i,:) = norminv([0.05 0.95], s_post_BB_CH4(i),
```

```
    sqrt(Q_post_BB_CH4(i)));

    Uncertainty_CH4(row,1) = Uncert_prior_Seepage_CH4(i,2) - sp_Seepage_CH4(i);
    Uncertainty_CH4(row,2) = Uncert_prior_NG_CH4(i,2) - sp_NG_CH4(i);
    Uncertainty_CH4(row,3) = Uncert_prior_Oil_CH4(i,2) - sp_Oil_CH4(i);
    Uncertainty_CH4(row,4) = Uncert_prior_Coal_CH4(i,2) - sp_Coal_CH4(i);
    Uncertainty_CH4(row,5) = Uncert_prior_AW_CH4(i,2) - sp_AW_CH4(i);
    Uncertainty_CH4(row,6) = Uncert_prior_Nat_CH4(i,2) - sp_Nat_CH4(i);
    Uncertainty_CH4(row,7) = Uncert_prior_BB_CH4(i,2) - sp_BB_CH4(i);
    Uncertainty_CH4(row+1,1) = Uncert_post_Seepage_CH4(i,2) - s_post_Seepage_CH4(i);
    Uncertainty_CH4(row+1,2) = Uncert_post_NG_CH4(i,2) - s_post_NG_CH4(i);
    Uncertainty_CH4(row+1,3) = Uncert_post_Oil_CH4(i,2) - s_post_Oil_CH4(i);
    Uncertainty_CH4(row+1,4) = Uncert_post_Coal_CH4(i,2) - s_post_Coal_CH4(i);
    Uncertainty_CH4(row+1,5) = Uncert_post_AW_CH4(i,2) - s_post_AW_CH4(i);
    Uncertainty_CH4(row+1,6) = Uncert_post_Nat_CH4(i,2) - s_post_Nat_CH4(i);
    Uncertainty_CH4(row+1,7) = Uncert_post_BB_CH4(i,2) - s_post_BB_CH4(i);
    row = row + 3;
end

if plotting == 1
    Uncertainty_CH4 = Uncertainty_CH4(49:84,:);
end

% calculate C2H6 uncertainties for checking with spreadsheet (priors) and output graphs
(posteriors)

[…] truncated here; analogous to CH4 above

% generate emissions files for plotting both positive and negative values
% (http://www.mathworks.com/matlabcentral/answers/17728)
% errorbarbar still gets called with s_C2H6, but s_C2H6_neg and s_C2H6_pos is used in
errorbarbar

s_CH4_neg = s_CH4;
s_CH4_neg(s_CH4_neg>0) = 0;
s_CH4_pos = s_CH4;
s_CH4_pos(s_CH4_pos<0) = 0;

Uncertainty_CH4(s_CH4<0) = 0;

if plotting == 1
    errorbarbar_since_2000_v5(s_CH4,Uncertainty_CH4);
else
    errorbarbar_v5(s_CH4,Uncertainty_CH4);
end

if emiss_scenario ~= 1 % is not 1 (don't need C2H6 for NOAA emissions)
    % generate emissions files for plotting both positive and negative values
    % (http://www.mathworks.com/matlabcentral/answers/17728)
    % errorbarbar still gets called with s_CH4, but s_CH4_neg and s_CH4_pos is used in
errorbarbar

    s_C2H6_neg = s_C2H6;
    s_C2H6_neg(s_C2H6_neg>0) = 0;
    s_C2H6_pos = s_C2H6;
    s_C2H6_pos(s_C2H6_pos<0) = 0;

    Uncertainty_C2H6(s_C2H6<0) = 0;

    plot_C2H6 = 1;
    if plotting == 1
    errorbarbar_since_2000_v5(s_C2H6,Uncertainty_C2H6);
else
    errorbarbar_v5(s_C2H6,Uncertainty_C2H6);
    end
end

% CH4 policy graphs

% plotting policy graphs in units of % of FF

if emiss_scenario == 3 % 1% NG FE
    Summary_mean_CH4 = zeros(5,4);
```

```
    Summary_uncertainty_CH4 = zeros(5,4);
    Summary_mean_CH4(1,1) = (sp_AW_CH4(2)+sp_Nat_CH4(2)+sp_BB_CH4(2)-s_post_AW_CH4(2)-
    s_post_Nat_CH4(2)-s_post_BB_CH4(2))/(sp_NG_CH4(2)+sp_Oil_CH4(2)+sp_Coal_CH4(2))*100;
    Summary_mean_CH4(1,2) = (sp_AW_CH4(7)+sp_Nat_CH4(7)+sp_BB_CH4(7)-s_post_AW_CH4(7)-
    s_post_Nat_CH4(7)-s_post_BB_CH4(7))/(sp_NG_CH4(7)+sp_Oil_CH4(7)+sp_Coal_CH4(7))*100;
    Summary_mean_CH4(1,3) = (sp_AW_CH4(17)+sp_Nat_CH4(17)+sp_BB_CH4(17)-
    s_post_AW_CH4(17)-s_post_Nat_CH4(17)-
    s_post_BB_CH4(17))/(sp_NG_CH4(17)+sp_Oil_CH4(17)+sp_Coal_CH4(17))*100;
    Summary_mean_CH4(1,4) = (sp_AW_CH4(27)+sp_Nat_CH4(27)+sp_BB_CH4(27)-
    s_post_AW_CH4(27)-s_post_Nat_CH4(27)-
    s_post_BB_CH4(27))/(sp_NG_CH4(27)+sp_Oil_CH4(27)+sp_Coal_CH4(27))*100;
    Summary_uncertainty_CH4(1,1) =
    (Uncert_prior_AW_CH4(2,1)+Uncert_prior_Nat_CH4(2,1)+Uncert_prior_BB_CH4(2,1)-
    s_post_AW_CH4(2)-s_post_Nat_CH4(2)-
    s_post_BB_CH4(2))/(sp_NG_CH4(2)+sp_Oil_CH4(2)+sp_Coal_CH4(2))*100;
    Summary_uncertainty_CH4(1,2) =
    (Uncert_prior_AW_CH4(7,1)+Uncert_prior_Nat_CH4(7,1)+Uncert_prior_BB_CH4(7,1)-
    s_post_AW_CH4(7)-s_post_Nat_CH4(7)-
    s_post_BB_CH4(7))/(sp_NG_CH4(7)+sp_Oil_CH4(7)+sp_Coal_CH4(7))*100;
    Summary_uncertainty_CH4(1,3) =
    (Uncert_prior_AW_CH4(17,1)+Uncert_prior_Nat_CH4(17,1)+Uncert_prior_BB_CH4(17,1)-
    s_post_AW_CH4(17)-s_post_Nat_CH4(17)-
    s_post_BB_CH4(17))/(sp_NG_CH4(17)+sp_Oil_CH4(17)+sp_Coal_CH4(17))*100;
    Summary_uncertainty_CH4(1,4) =
    (Uncert_prior_AW_CH4(27,1)+Uncert_prior_Nat_CH4(27,1)+Uncert_prior_BB_CH4(27,1)-
    s_post_AW_CH4(27)-s_post_Nat_CH4(27)-
    s_post_BB_CH4(27))/(sp_NG_CH4(27)+sp_Oil_CH4(27)+sp_Coal_CH4(27))*100;
else
    if emiss_scenario == 4 % 3% NG FE

[…] truncated here; analogous to 1% NG FE above

end

figure2 = figure;
set(figure2, 'Position', [1000 400 700 360])
    % Create axes
    axes2 = axes('Parent',figure2,'YGrid','on',...
    'XTickLabel',{'1%','3%','5%','7%','9%'},...
    'XTick',[1 2 3 4 5],...
    'Position',[0.0617577197149644 0.114232209737828 0.923990498812352
0.814606741573034],...
    'FontSize',20);
    % box(axes2,'on');
    hold(axes2,'all');
    b = bar(Summary_mean_CH4,'grouped');
    ylim(axes2,[-100 100]);
%     % Create ylabel
%     ylabel('CH_{4} emissions relative to FF prior (%) ','FontSize',20);
%     % Create xlabel
%     xlabel(NG FE rate','FontSize',20);
%     l = cell(1,4);
%     l{1}='1985'; l{2}='1990'; l{3}='2000'; l{4}='2010';
%     legend_mean = legend(b,l,'Orientation','horizontal');
%     set(legend_mean,'Location','SouthEast');
%     title('Prior mean - Posterior mean ');

figure3 = figure;
set(figure3, 'Position', [500 400 700 360])
    % Create axes
    axes3 = axes('Parent',figure3,'YGrid','on',...
    'XTickLabel',{'1%','3%','5%','7%','9%'},...
    'XTick',[1 2 3 4 5],...
    'Position',[0.0707162284678151 0.114232209737828 0.915031990059501
0.811776758872037],...
    'FontSize',20);
    % box(axes3,'on');
    hold(axes3,'all');
    b = bar(Summary_uncertainty_CH4,'grouped');
    ylim(axes3,[-250 100]);
%     % Create ylabel
%     ylabel('CH_{4} emissions relative to FF prior (%) ','FontSize',20);
%     % Create xlabel
```

```
%     xlabel('NG FE rate','FontSize',20);
%     l = cell(1,4);
%     l{1}='1985'; l{2}='1990'; l{3}='2000'; l{4}='2010';
%     legend_mean = legend(b,l);
%     set(legend_mean,'Location','SouthEast');
%     title('Lower bound prior uncertainty - Posterior mean ');

% C2H6 policy graphs

[…] truncated here; analogous to CH4 graphs above
```

errorbar_v5 (function used in box-model code above; adapted using external model code as referenced below):

```
function [b,e] = errorbarbar(x,y,E,barSettings,lineSettings)

% load workspace into function (need that for variable plot_C2H6)
evalin('base','save myvars.mat');
load myvars.mat

% function to plot bar plots with error bars
%   [b,e] = errorbarbar(y,E)
%   [b,e] = errorbarbar(x,y,E)
%   [b,e] = errorbarbar(x,y,E, barSetings, lineSettings)
% Inputs barSettings and lineSettings are cells containing settings for
% bars and lineSettings. If you want to set only one of them, set the other
% to empty.
% Outputs b and e are the handles to the bar and the errorbar plotted.
%
% NOTE: Currently does NOT support errorbar(x,y,L,U) directly. It is
% possible as a trick using lineSettings appropriately.
%
%  See Also:
%       bar, errorbar
%
%  Dependencies:
%       No additional files are required.
%

% Created by Venn on 2009-JUL-13 (vennjr@u.northwestern.edu)
% Modified:
%   2011-Jun-22. Removed the legends for the error bars.
%   2009-Jul-15. Now works with stacked bars too. Woot!
%   2009-Jul-14. Fixed minor bug in the input settings and fixed a bug when
%                either barSettings or lineSettings are empty.

%% use the appropriate setting
if nargin<5 || isempty(lineSettings)
    lineSettings = {'linestyle','none'};
end
if nargin<4
    barSettings = {};
end

%% plot the bars CH4

% in this case "errorbarbar(x,E)", x and E are both 84x3 matrices

if plot_C2H6 == 0 % plot CH4

if nargin<3 % if no x is provided
    % Create figure
    figure1 = figure;
    set(figure1, 'Position', [10 400 720 300])
    % Create axes
    axes1 = axes('Parent',figure1,'YGrid','on','XTick',zeros(1,0),...
    'Position',[0.053585500394011 0.048 0.924747832939323 0.923119133574007],...
    'FontSize',20);
%     axes1 = axes('Parent',figure1,'YGrid','on',...
      'XTickLabel',{'1984','1985','1986','1987','1988','1989','1990','1991','1992',
      '1993','1994','1995','1996','1997','1998','1999','2000','2001','2002','2003',
      '2004','2005','2006','2007','2008','2009','2010','2011'},...
```

```
%     'XTick',[1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 64 67 70 73 76
      79 82],...
%     'FontSize',16);
%     xlab = get(gca,'XTickLabel');
%     set(gca,'XTickLabel',xlab);
%     rotateXLabels(gca, 45)
%     box(axes1,'on');
    hold(axes1,'all');
%     b = bar(x,1,'stacked','Parent',axes1); % "1" reduces space between bars to zero
    hold on
    bar(s_CH4_neg,1,'stacked','Parent',axes1);
    b = bar(s_CH4_pos,1,'stacked','Parent',axes1);
    hold off
%     Create ylabel
%     ylabel('Global CH_{4} emissions (Tg/yr)   ','FontSize',16);
%     Create xlabel
%     xlabel('Year  ','FontSize',16);
%     % Create legend
%     lg = cell(1,6);
%     lg = {'NG' 'Oil' 'Coal' 'Ag/waste' 'Natural' 'BB'};
%     order=[6 5 4 3 2 1];
%     legend_mean = legend(b(order),lg{order});
%     set(legend_mean,'Location','NorthWest');
    xlim(axes1,[0 84.5]);
    ylim(axes1,[0 800]);
%     Resize the axes in order to prevent it from shrinking.
    set(axes1,'Position',[0.053585500394011 0.137658227848101 0.838734372923487
0.833460905725906]);
    E = y;
    y = x;
else
%   b = bar(x,y,barSettings{:});
    b = bar(x,y,1,'stacked');
end

end

%% plot the bars C2H6

[…] truncated here; analogous to CH4 above

%% get the xdata to plot the error plots
c = get(b,'Children');
if iscell(c)
    for i = 1:length(c)
        xdata(:,i) = mean(get(c{i},'xdata'));
        tempYData  = get(c{i},'ydata');
        ydata(:,i) = mean(tempYData(2:3,:))';
    end
else
    xdata = mean(get(c,'xdata'));
    tempYData  = get(c,'ydata');
    ydata = mean(tempYData(2:3,:))';
end

%% plot the errorbars
hold on;
e = errorbar(xdata,ydata,E,lineSettings{:});

for i = 1:length(e)
    hasbehavior(e(i), 'legend', false);
end
hold off;
```