Modeling $\text{N}_2\text{O}$ Emissions from Agricultural Soils Using a Multi-level Linear Regression

Kyle Carey
August, 2007

MP Advisor:
Dr. William H. Schlesinger

Secondary Advisor:
Dr. Song Qian

Masters project submitted in partial fulfillment of the requirements for the Master of Environmental Management degree in the Nicholas School of the Environment and Earth Sciences of Duke University

2007
Table of Contents

Abstract

Background/Introduction .................................................................4

Data and Methods
  Data Collection.............................................................................8
  Data Summary..............................................................................10
  The Multi-level Linear Model......................................................17

Results
  Overall Model..............................................................................22
  Crop Type Models........................................................................24
  Fertilizer Type Models.................................................................32

Analysis/Discussion
  Overall Model..............................................................................43
  Crop Type Models........................................................................44
  Fertilizer Type Models.................................................................45

Conclusion......................................................................................52

References.......................................................................................53

Database Reference........................................................................56

Figure Appendix
Nitrous oxide (N\textsubscript{2}O) is a potent greenhouse gas emitted from soils through the microbial processes of denitrification/nitrification. In agricultural soils these natural processes and thus N\textsubscript{2}O emission are significantly enhanced by the use of nitrogen fertilizers. Use of fertilizer has grown rapidly since the 2\textsuperscript{nd} half of the 20\textsuperscript{th} century and will likely continue to grow as more countries develop advanced agricultural management strategies and demand for food increases. Though fertilizer enhanced N\textsubscript{2}O emissions from agriculture contribute only a small amount to the current anthropogenic global climate warming, it is expected that these emissions will play a larger role in the future.

This study consists of a statistical meta-analysis of an agricultural N\textsubscript{2}O emission database made up of data taken from peer-reviewed literature. A multi-level linear regression is used to investigate the relationship between N fertilizer input and N\textsubscript{2}O emissions; (1) independent of variation between each study, (2) as a categorical function of crop type and (3) as a categorical function of fertilizer type. An understanding of these relationships could help to establish management strategies to more efficiently use N fertilizers, reducing N\textsubscript{2}O emissions and lowering expenses for agricultural producers.

The results of the multi-level linear analysis of the dataset indicate that the relationship between N input and N\textsubscript{2}O emissions is not independent of the conditional variation between studies. The categorical analysis of differences in crop type also did not have a significant influence on N\textsubscript{2}O emissions. The categorical analysis of N fertilizer forms did show a significant influence on emissions. Differences in the slopes of the fertilizer type models provide relative comparability of expected N\textsubscript{2}O emissions of different chemical forms of N fertilizer for a given N input. The analysis performed in this study yielded important insight into the factors influencing N\textsubscript{2}O emissions from agricultural soils. These findings can be used to guide both management strategies and further research dealing with an increasingly important topic.
Background/Introduction

To effectively mitigate further anthropogenic climate change, greenhouse gases other than carbon dioxide (CO$_2$) must be included in the regulatory approach. Nitrous oxide (N$_2$O), a potent greenhouse gas with a global warming potential (GWP) of 296 compared to CO$_2$, currently contributes 6% of the enhanced greenhouse effect (Houghton, 2004, p. 44, 247). Its greenhouse warming strength stems from its long atmospheric lifetime (120 years) due in part to its low reactivity and the lack of a destruction mechanism for N$_2$O in the troposphere (Ibid, 44). The only known sink for N$_2$O is found in the stratosphere, through photo-dissociation or reactions with excited oxygen atoms (Ibid, and Schlesinger, 1997, 71). The destruction of N$_2$O creates NO (Schlesinger, 1997, 71), which destroys ozone allowing more ultraviolet light to enter the troposphere. This increases the amount of solar radiation reaching the earth’s surface indirectly enhancing global warming (IPCC, 2001, 4.2.3.3). The atmospheric concentration of N$_2$O has been increasing at a rate of 0.25% per year (Houghton, 2004, 44) leading to the current globally averaged concentration of 314 ppb compared to the pre-industrial atmospheric concentration of 270 ppb from natural N$_2$O emissions (IPCC, 2001, 4.2.1.2).

Yearly emissions of N$_2$O are estimated by the IPCC (2001) and Mosier et al. (1998) to total 16.4 Tg N/yr. Of the total, 10.7 Tg N/yr are thought to be from natural sources (Ibid). This estimate implies that current anthropogenic emissions are 5.7 Tg N/yr assuming there has been no change in natural emissions over time (Ibid). While anthropogenic N$_2$O emissions are produced by a number of industrial sources, including nylon production, nitric acid production, fossil fuel fired power plants, and vehicular emissions (IPCC, 2001, 4.2.1.2), agriculture is believed to account for a little over 50% of the anthropogenic emissions (Ibid, table 4.4).

N$_2$O is produced in soils by microbes through the processes nitrification and denitrification (EPA, 2007, 6-16). Nitrification is the aerobic microbial oxidation of ammonium (NH$_4^+$) to nitrate (NO$_3^-$) (IPCC, 11.5, 2006). N$_2$O is a by-product of nitrification that leaks from microbial cells into the soil and ultimately into the atmosphere. N$_2$O flux associated with nitrification is generally a small fraction of total nitrification flux, but can make a major contribution to total soil N$_2$O emissions (Barnard et al., 2005).

Denitrification is the anaerobic microbial reduction of nitrate (NO$_3^-$) to nitrogen gas (N$_2$) (IPCC, 11.5, 2006). Nitrous oxide is produced as a gaseous intermediate in the reaction sequence of denitrification (Ibid), and leaks to the atmosphere.

Nitrification and denitrification and thus N$_2$O emissions are controlled by complex factors. Nitrification is favored by the availability of NH$_4^+$, at moderate pH and in well aerated soils (Barnard et al., 2005). Denitrification in soils is favored by a high availability of NO$_3^-$ and organic C as an energy source in poorly aerated soils with a pH close to neutrality (Ibid). Although denitrification is considered to be the major source of N$_2$O in most conditions, nitrification can contribute a substantial N$_2$O flux under aerobic conditions (Williams et al., 1989). Both processes occur in soils simultaneously, making it difficult to determine the specific contribution of nitrification and denitrification to observed N$_2$O emissions (Arah, 1997). The emission of N$_2$O from soils is also indirectly influenced by climatic and environmental conditions that affect soil O$_2$ concentrations.
such as temperature, soil characteristics, soil water content and biological activity (Barnard et al., 2005).

Agricultural activities increase N\textsubscript{2}O emissions from soils in a number of ways. Increased use of chemical fertilizer, application of organic materials, including manure from domesticated animals, production of N-fixing crops, retention of crop residues and the cultivation of organic soils directly increase the availability of mineral N and organic C (EPA, 2007, 6-16). Greater mineral N and organic C availability support enhanced microbial action increasing N\textsubscript{2}O emissions. Agricultural management strategies, such as irrigation, drainage, tillage and fallowing can indirectly influence N mineralization and N\textsubscript{2}O emissions (Ibid).

The emissions of N\textsubscript{2}O that result from agricultural N inputs occur through a direct pathway: the microbial processes in the soils where fertilizer is applied, and through two indirect pathways: (1.) following volatilization of NH\textsubscript{3} and NO\textsubscript{x} from managed soils, these gases and their products-- NH\textsubscript{4}+ and NO\textsubscript{3}--can be redeposited to soils and waters and subsequently undergo nitrification/denitrification and (2.) mineral N undergoes nitrification/denitrification in nearby environments after leaching and runoff of N from agricultural fields mainly as NO\textsubscript{3}-- from managed soils (IPCC, 2006, 11.2).

According to the International Fertilizer Industry Association (IFA), worldwide consumption of chemical N fertilizer (in metric tonnes of nitrogen) was 90 million t N for 2004/2005 (IFADATA, 2007). As seen in figure 1, world demand for fertilizer has rapidly increased over the last forty-five years. However growth in fertilizer consumption has not occurred evenly. In developed countries the overall use of N fertilizer has decreased since 1990 and currently remains fairly constant at around 30 million t N/year (figure 2), while the developing world has and continues to experience high growth rates in consumption (figure 3). Based on a continuation of the historical consumption in the developing world depicted in figure 2, and assuming a constant total consumption in the developed world close to today’s, I calculate that world wide consumption of chemical N fertilizers will reach 157 million t N/year by 2050 and grow to 231 million t N/year by 2100. This estimate may be conservative. There is the possibility that demand for chemical N fertilizer in developed countries will increase should they expand their reliance on bio-fuels. Associated with increased fertilizer use are greater N\textsubscript{2}O emissions and a larger contribution to anthropogenic global warming.
Figure 1

World Consumption of Chemical N Fertilizer

Figure 2

Developing World Consumption of Chemical N Fertilizer
Both direct and indirect emissions of N\textsubscript{2}O represent significant sources of anthropogenic greenhouse gas emissions and should both be addressed by regulations seeking to reduce their climate impact. However, indirect N\textsubscript{2}O emissions are difficult to quantify. As was previously mentioned, applied fertilizer is partially lost though volatilization and redeposition or by leaching from the soil by water. While the impacts of N deposition in natural ecosystems, such as eutrophication in aquatic environments, are widely observable, measuring the anthropogenic contribution to N\textsubscript{2}O emissions in these areas is challenging. In contrast, direct emissions from nitrification and denitrification in cropland soil are relatively easily measured, and they can be compared to the amount of N input from fertilizer.

This study consists of a statistical meta-analysis that focuses on analyzing the relationship between direct emissions of N\textsubscript{2}O from agricultural soils and N input by chemical and organic fertilizers. N input from the deposition of manure and urine by domesticated livestock is not included in this study. Data subsets based on crop type and fertilizer type were also included in this study to explore possible differences in N\textsubscript{2}O emission based on management practice. Through use of a multi-level linear regression, this study seeks to create a model for N\textsubscript{2}O emissions that can include data from a large number of studies in different locations with a variety of climatic/environmental conditions. Understanding the influence of these specific variables may help to establish management strategies to more efficiently use N fertilizers, reducing N\textsubscript{2}O emissions and lowering expenses for agricultural producers. This is especially important in the U.S. where the price of natural gas; the primary raw material for chemical N fertilizers, has been extremely volatile (USGAO, 2003). In addition, agricultural practices that increase fertilizer efficiency/decrease usage will also curtail associated indirect emissions of N\textsubscript{2}O.
from the natural environment reducing the negative impact agriculture has on surrounding ecosystems.

**Data/Methods**

**Data Collection**

Data for this study were collected from field experiments reported in peer-reviewed literature. For each literature citation, the data set includes information on: study site location, dates of study, climate/precipitation, soil type, soil texture, soil moisture, pH, organic C content, drainage, N content, sample period length, sample frequency, N fertilizer application rate, fertilizer type, method of application, N\textsubscript{2}O-N emissions over the measurement period in kg/ha (adjusted for control and non-adjusted), and N emissions as a % of fertilizer applied (emission factor). Data taken from these studies span nearly thirty years of research, using methods that have evolved with time. Many studies do not provide complete information for all the variables listed. Soil moisture, often given as water filled pore space (WFPS), is an important factor that influences N\textsubscript{2}O emission by affecting soil O\textsubscript{2} levels. WFPS is inconsistently reported in the literature and could not be included in this study.

The dataset initially contained information for 1201 N\textsubscript{2}O emission measurements from 173 papers. However, upon a qualitative review, the database was reduced to information from 1084 N\textsubscript{2}O measurements taken from 164 studies. Data from experiments using nitrification inhibitors, controlled release fertilizer, or experiments performed on organic soils were omitted, as well as any data that did not list N input or N\textsubscript{2}O-N emissions values.

For more rigorous analysis, the dataset was grouped into subsets by the crop type and the fertilizer type used in each measurement (Tables 1 and 2). These categorical subsets were used to explore differences in N\textsubscript{2}O emission based on the influence of management variables. The published literature reviewed for this project contains disparate data with regards to certain crops and certain N fertilizers. In the interest of maintaining large samples to enable better statistical comparison, some categories contain multiple fertilizers and crops. Fertilizers were organized by the form of nitrogen. Crops were lumped in some cases under *Group Crop Types*, such as in the Cereal group and the Vegetable group which contain a number of similar species. Individual crops that had a large presence in the literature or were worthy of consideration based upon their widespread planting (i.e. wheat, corn and oilseed rape) were further broken into exclusive crop type data subsets. Of the categorical data, only subsets with enough samples to be used in the comparative analysis are listed in bold (table 3)
Table 1. Codes for Nitrogen Fertilizer Type Categories

<table>
<thead>
<tr>
<th>Code</th>
<th>Fertilizers Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Ammonia (aqueous/anhydrous), ammonium chloride, (di)ammonium</td>
</tr>
<tr>
<td></td>
<td>phosphate, ammonium sulfate</td>
</tr>
<tr>
<td>AN</td>
<td>Ammonium nitrate, calcium ammonium nitrate</td>
</tr>
<tr>
<td>M</td>
<td>Manure, slurry, urine, composted municipal waste, organic manure,</td>
</tr>
<tr>
<td></td>
<td>mulch, crop stubble</td>
</tr>
<tr>
<td>N</td>
<td>Nitrate (calcium, potassium and sodium)</td>
</tr>
<tr>
<td>U</td>
<td>Urea</td>
</tr>
<tr>
<td>ANM</td>
<td>Ammonium Nitrate – Manure</td>
</tr>
<tr>
<td>UA</td>
<td>Urea – Ammonia</td>
</tr>
<tr>
<td>UAN</td>
<td>Urea – Ammonium Nitrate</td>
</tr>
<tr>
<td>UM</td>
<td>Urea – Manure</td>
</tr>
<tr>
<td>None</td>
<td>No N fertilizer input used</td>
</tr>
<tr>
<td>Unknown</td>
<td>Not listed</td>
</tr>
</tbody>
</table>

Table 2. Codes for Crop Type Categories

<table>
<thead>
<tr>
<th>Code</th>
<th>Crops Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Crop Types</td>
<td></td>
</tr>
<tr>
<td>Cereal</td>
<td>Barley, spring/winter barley, wheat, spring/winter wheat, oats, sorghum</td>
</tr>
<tr>
<td>Grass</td>
<td>Grass, rye grass, grass/clover, pasture, sod, orchard weeds, timothy weeds, weeds</td>
</tr>
<tr>
<td>Legume</td>
<td>Alfalfa, adzuki beans, beans, clover, soybeans</td>
</tr>
<tr>
<td>Tree</td>
<td>Poplar, willow</td>
</tr>
<tr>
<td>Tuber</td>
<td>Potato</td>
</tr>
<tr>
<td>Vegetable</td>
<td>Artichoke, broccoli, cabbage, carrot, cauliflower, celery, oilseed rape (canola), onion, pac choi, peas,</td>
</tr>
<tr>
<td>Other</td>
<td>Banana, miscanthus giganteus, papaya, tobacco, sugar beet, sugar cane, sunflower</td>
</tr>
<tr>
<td>None</td>
<td>Bare, fallow, unplanted</td>
</tr>
<tr>
<td>Unknown</td>
<td>Not listed</td>
</tr>
<tr>
<td>Individual Crop Types</td>
<td></td>
</tr>
<tr>
<td>Bare</td>
<td>Bare soil</td>
</tr>
<tr>
<td>Barley</td>
<td>Barley, winter/spring barley</td>
</tr>
<tr>
<td>Corn</td>
<td>Corn/maize</td>
</tr>
<tr>
<td>Rape</td>
<td>Oilseed rape, canola, rape</td>
</tr>
<tr>
<td>Rice</td>
<td>Flooded and wet/dry rice</td>
</tr>
<tr>
<td>Wheat</td>
<td>Wheat, winter/spring wheat</td>
</tr>
</tbody>
</table>
Table 3. Data Base Summary

<table>
<thead>
<tr>
<th>Crop Type</th>
<th>Measurements</th>
<th>Fertilizer Type</th>
<th>Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cereal</td>
<td>199</td>
<td>A</td>
<td>109</td>
</tr>
<tr>
<td>Grass</td>
<td>315</td>
<td>AN</td>
<td>239</td>
</tr>
<tr>
<td>Legume</td>
<td>40</td>
<td>ANM</td>
<td>30</td>
</tr>
<tr>
<td>Tree</td>
<td>26</td>
<td>M</td>
<td>112</td>
</tr>
<tr>
<td>Tuber</td>
<td>24</td>
<td>N</td>
<td>38</td>
</tr>
<tr>
<td>Vegetable</td>
<td>94</td>
<td>U</td>
<td>126</td>
</tr>
<tr>
<td>Other</td>
<td>41</td>
<td>UA</td>
<td>14</td>
</tr>
<tr>
<td>None</td>
<td>90</td>
<td>UAN</td>
<td>60</td>
</tr>
<tr>
<td>Unknown</td>
<td>29</td>
<td>UM</td>
<td>21</td>
</tr>
<tr>
<td>Bare</td>
<td>50</td>
<td>None</td>
<td>292</td>
</tr>
<tr>
<td>Barley</td>
<td>86</td>
<td>Unknown</td>
<td>43</td>
</tr>
<tr>
<td>Corn</td>
<td>144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rape</td>
<td>51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheat</td>
<td>105</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data Summary

N$_2$O Emissions

The dataset includes measurements of the total N$_2$O emission as kg N$_2$O-N/ha from every study. However, the period over which these measurements were made varies greatly by study (figure 4) from 2 days to 840 days, making comparison of total emissions difficult. To account for this disparity, the average daily emission in each study was calculated by dividing the total N$_2$O emission value by the number of days in its sample period. Average daily emissions are expressed in g N$_2$O-N/ha/day representing a standardized value for each measurement. The resulting distribution of the average daily N$_2$O-N emission values is similar to the distribution of total emission values in that they are both skewed right (figures 5 and 6). Average daily emissions for the crop data subsets and the fertilizer type subsets are summarized in tables 4 and 5.
Figure 4

Sample Period Histogram

Days

Frequency

0 200 400 600 800

Figure 5

Total N2O Emission Histogram

kg N2O-N/ha/study period

Frequency

0 200 400 600
Figure 6

Overall Average Daily N₂O Emissions Histogram

Table 4 Crop Type Average Daily N₂O-N Emissions Summary

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Min (gN/ha/d)</th>
<th>Fert Type</th>
<th>Input (kg N/ha)</th>
<th>Max (gN/ha/d)</th>
<th>Fert Type</th>
<th>Input (kg N/ha)</th>
<th>Median emission</th>
<th>Mean emission</th>
<th>Input Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.00003</td>
<td>None</td>
<td>0</td>
<td>2314.3</td>
<td>N</td>
<td>100</td>
<td>5.75</td>
<td>19.28</td>
<td>0 – 1668</td>
</tr>
<tr>
<td>Bare</td>
<td>0.0533</td>
<td>None</td>
<td>0</td>
<td>429.2</td>
<td>N</td>
<td>100</td>
<td>6.406</td>
<td>21.63</td>
<td>0 – 250</td>
</tr>
<tr>
<td>Barley</td>
<td>0.00006</td>
<td>None</td>
<td>0</td>
<td>160.0</td>
<td>M</td>
<td>1668</td>
<td>5.30</td>
<td>9.80</td>
<td>0-1668</td>
</tr>
<tr>
<td>Cereal</td>
<td>0.0001</td>
<td>None</td>
<td>0</td>
<td>159.5</td>
<td>M</td>
<td>1668</td>
<td>5.357</td>
<td>10.55</td>
<td>0 – 1668</td>
</tr>
<tr>
<td>Corn</td>
<td>0.598</td>
<td>None</td>
<td>0</td>
<td>121.9</td>
<td>U</td>
<td>75</td>
<td>8.348</td>
<td>17.42</td>
<td>0 – 450</td>
</tr>
<tr>
<td>Grass</td>
<td>0.00003</td>
<td>None</td>
<td>0</td>
<td>2314.3</td>
<td>N</td>
<td>100</td>
<td>4.658</td>
<td>30.84</td>
<td>0 – 1230</td>
</tr>
<tr>
<td>Rape</td>
<td>0.0533</td>
<td>None</td>
<td>0</td>
<td>40.57</td>
<td>AN</td>
<td>140</td>
<td>3.585</td>
<td>13.9</td>
<td>0 – 262</td>
</tr>
<tr>
<td>Rice</td>
<td>0.0028</td>
<td>U</td>
<td>122</td>
<td>96.54</td>
<td>M</td>
<td>176</td>
<td>6.849</td>
<td>8.589</td>
<td>0 – 676</td>
</tr>
<tr>
<td>Vegetable</td>
<td>0.0533</td>
<td>None</td>
<td>0</td>
<td>144.4</td>
<td>M</td>
<td>1006</td>
<td>6.849</td>
<td>17.08</td>
<td>0 – 1006</td>
</tr>
<tr>
<td>Wheat</td>
<td>0.0008</td>
<td>None</td>
<td>0</td>
<td>118.5</td>
<td>AN</td>
<td>153</td>
<td>4.767</td>
<td>10.25</td>
<td>0 – 350</td>
</tr>
</tbody>
</table>
Table 5 Fertilizer Type Average Daily N\textsubscript{2}O-N Emissions Summary

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Min (gN/ha/d)</th>
<th>Crop Type</th>
<th>N Input (gN/ha/d)</th>
<th>Max (gN/ha/d)</th>
<th>Crop Type</th>
<th>N Input</th>
<th>Median</th>
<th>Mean</th>
<th>Input Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.0015</td>
<td>Other</td>
<td>20</td>
<td>2314.3</td>
<td>Grass</td>
<td>100</td>
<td>5.75</td>
<td>19.28</td>
<td>20 – 1668</td>
</tr>
<tr>
<td>A</td>
<td>0.5205</td>
<td>Legume</td>
<td>40</td>
<td>250</td>
<td>Grass</td>
<td>250</td>
<td>6.13</td>
<td>14.8</td>
<td>20 – 450</td>
</tr>
<tr>
<td>AN</td>
<td>0.25</td>
<td>Tuber</td>
<td>60</td>
<td>209</td>
<td>Grass</td>
<td>300</td>
<td>8.2</td>
<td>17.3</td>
<td>30 – 880</td>
</tr>
<tr>
<td>ANM</td>
<td>1.3</td>
<td>Grass</td>
<td>275</td>
<td>103</td>
<td>Grass</td>
<td>278</td>
<td>17.9</td>
<td>25</td>
<td>92 – 743</td>
</tr>
<tr>
<td>M</td>
<td>0.056</td>
<td>Grass</td>
<td>34</td>
<td>208</td>
<td>Grass</td>
<td>320</td>
<td>13.9</td>
<td>31.6</td>
<td>27 – 1668</td>
</tr>
<tr>
<td>N</td>
<td>0.0333</td>
<td>Grass</td>
<td>100</td>
<td>2314.3</td>
<td>Grass</td>
<td>100</td>
<td>4.453</td>
<td>133.9</td>
<td>57 – 700</td>
</tr>
<tr>
<td>U</td>
<td>0.00015</td>
<td>Other</td>
<td>20</td>
<td>121.9</td>
<td>Corn</td>
<td>75</td>
<td>5.479</td>
<td>13.99</td>
<td>20 – 552</td>
</tr>
<tr>
<td>UAN</td>
<td>2.9</td>
<td>Grass</td>
<td>175</td>
<td>98.9</td>
<td>Vegetable</td>
<td>130</td>
<td>8.055</td>
<td>12.08</td>
<td>35 – 364.5</td>
</tr>
<tr>
<td>UM</td>
<td>0.2887</td>
<td>Rice</td>
<td>192</td>
<td>94.96</td>
<td>Rice</td>
<td>147</td>
<td>14.25</td>
<td>27.08</td>
<td>112 – 676</td>
</tr>
<tr>
<td>None</td>
<td>0.00003</td>
<td>None</td>
<td>0</td>
<td>158.7</td>
<td>Corn</td>
<td>0</td>
<td>2.74</td>
<td>6.33</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 4 shows that nearly all of the lowest N\textsubscript{2}O emission values in the crop type subset were found in plots where no fertilizer was applied. One might expect this because without added inorganic nitrogen for enhanced denitrification/nitrification, N\textsubscript{2}O emissions should be lower. Maximum daily emission values and their associated N input values in the crop type subsets showed no pattern, which suggests that the type of N input and amount of N added may not be as important as other factors influencing emission. The mean and median values show that the distribution of each subset is right skewed like the overall dataset (figure 6) and is demonstrated in histograms as well (see figure appendix). Table 5 shows no obvious patterns in the fertilizer type subset, but like the crop type subsets, the mean and median average daily emissions values indicated that their distributions are right skewed.

**Nitrogen Input**

A histogram of the overall dataset N input (figure 7) shows that the data distribution is skewed right. Crop type and fertilizer type data subsets also show right skewed distributions of N input, but to varying degrees (figure appendix). The distribution of N input values excluding the 292 zero input measurements has a median of 150 kg N/ha and a mean of 194.4 kg N/ha. Initially, plots with no added N were omitted from the study. Non-control emissions values were to be adjusted to reflect fertilizer induced emission by subtracting the natural background N\textsubscript{2}O emissions represented by the control
group emissions values. However, many studies do not include a zero input control making the adjustment of total emissions measurements to fertilizer induced emission error prone. Furthermore, excluding control data would reduce the size of the dataset by more than 25%, and deletes the point where individual regression lines cross the intercept.

Figure 7

Overall Nitrogen Input Histogram

![Histogram of nitrogen input frequencies](image)

Emission Factor (EF)

The IPCC uses default emission factors (EF); the percentage of nitrogen input emitted as N$_2$O, to estimate N$_2$O emissions from managed soils in their tier 1, tier 2 and tier 3 approaches (IPCC, 2006, 11.10). The tier 1 EF is used to estimate N$_2$O emitted from N additions from mineral fertilizers, organic amendments and crop residues and N mineralized from mineral soils as a result of loss of soil carbon. The tier 2 EF is used to estimate the N$_2$O emissions from drained/managed organic soils and the tier 3 EF is used to estimate the amount of N$_2$O emitted from urine and dung deposited by grazing animals (Ibid, 11.10, Table 11.1). The dataset in this study includes measurements of N$_2$O emissions from traditionally managed agricultural soils and is comparable to tier 1. A comparison between the EF found in the collected dataset compared to the IPCC default shows similar values. The tier 1 default EF is 1% of the N applied to soils with an uncertainty range of .3% – 3% (Ibid, Table 11.1). Tables 6 and 7 summarize the overall data set EF and the crop type and fertilizer type subset EFs.
### Table 6 Crop Type Emissions Factor Summary

<table>
<thead>
<tr>
<th>Crop Type</th>
<th>Min EF</th>
<th>Fert Type</th>
<th>Input (kg N/ha)</th>
<th>Max EF</th>
<th>Fert Type</th>
<th>Input (kg N/ha)</th>
<th>Median EF</th>
<th>Mean EF</th>
<th>Input Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.00001</td>
<td>U</td>
<td>122</td>
<td>24.44</td>
<td>M</td>
<td>27</td>
<td>0.9216</td>
<td>1.8</td>
<td>20 – 1668</td>
</tr>
<tr>
<td>Bare</td>
<td>0.018</td>
<td>N</td>
<td>100</td>
<td>16.31</td>
<td>N</td>
<td>100</td>
<td>0.3056</td>
<td>1.313</td>
<td>55 – 250</td>
</tr>
<tr>
<td>Barley</td>
<td>0.0014</td>
<td>AN</td>
<td>285</td>
<td>10</td>
<td>U</td>
<td>25</td>
<td>0.667</td>
<td>1.41</td>
<td>25-1668</td>
</tr>
<tr>
<td>Cereal</td>
<td>0.0014</td>
<td>AN</td>
<td>285</td>
<td>10</td>
<td>U</td>
<td>25</td>
<td>0.819</td>
<td>1.629</td>
<td>25 – 1668</td>
</tr>
<tr>
<td>Corn</td>
<td>0.15</td>
<td>AN</td>
<td>200</td>
<td>9</td>
<td>AN</td>
<td>30</td>
<td>1.099</td>
<td>1.612</td>
<td>30 – 450</td>
</tr>
<tr>
<td>Grass</td>
<td>0.001</td>
<td>N</td>
<td>100</td>
<td>16.2</td>
<td>N</td>
<td>100</td>
<td>0.88</td>
<td>1.712</td>
<td>34 – 1230</td>
</tr>
<tr>
<td>Rape</td>
<td>0.0585</td>
<td>AN</td>
<td>100</td>
<td>6.143</td>
<td>AN</td>
<td>140</td>
<td>1.316</td>
<td>1.581</td>
<td>45 – 262</td>
</tr>
<tr>
<td>Rice</td>
<td>0.00001</td>
<td>U</td>
<td>122</td>
<td>24.44</td>
<td>M</td>
<td>27</td>
<td>.327</td>
<td>1.301</td>
<td>20 – 676</td>
</tr>
<tr>
<td>Vegetable</td>
<td>0.0224</td>
<td>U</td>
<td>250</td>
<td>15.49</td>
<td>AN</td>
<td>51.6</td>
<td>1.296</td>
<td>2.389</td>
<td>45 – 1006</td>
</tr>
<tr>
<td>Wheat</td>
<td>0.0048</td>
<td>AN</td>
<td>86</td>
<td>8.842</td>
<td>ANM</td>
<td>190</td>
<td>0.9</td>
<td>1.525</td>
<td>39 – 350</td>
</tr>
<tr>
<td>Data Set</td>
<td>Min EF</td>
<td>Crop Type</td>
<td>Input (kg N/ha)</td>
<td>Max EF</td>
<td>Crop Type</td>
<td>Input (kg N/ha)</td>
<td>Median</td>
<td>Mean</td>
<td>Input Range</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>-----------</td>
<td>-----------------</td>
<td>--------</td>
<td>-----------</td>
<td>-----------------</td>
<td>--------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Overall</td>
<td>0.00001</td>
<td>Rice</td>
<td>122</td>
<td>24.4</td>
<td>Rice</td>
<td>27</td>
<td>0.9</td>
<td>1.781</td>
<td>20 – 1669</td>
</tr>
<tr>
<td>A</td>
<td>0.025</td>
<td>Grass</td>
<td>100</td>
<td>7.84</td>
<td>Legume</td>
<td>250</td>
<td>0.478</td>
<td>0.965</td>
<td>20 – 450</td>
</tr>
<tr>
<td>AN</td>
<td>0.0014</td>
<td>Cereal</td>
<td>285</td>
<td>15.5</td>
<td>Vegetable</td>
<td>51.6</td>
<td>1.27</td>
<td>1.97</td>
<td>30 – 880</td>
</tr>
<tr>
<td>ANM</td>
<td>0.12</td>
<td>Other</td>
<td>245</td>
<td>11.4</td>
<td>Grass</td>
<td>278</td>
<td>2.1</td>
<td>2.87</td>
<td>92 – 743</td>
</tr>
<tr>
<td>M</td>
<td>0.045</td>
<td>Cereal</td>
<td>450</td>
<td>24.4</td>
<td>Rice</td>
<td>27</td>
<td>0.6</td>
<td>1.79</td>
<td>27 – 1668</td>
</tr>
<tr>
<td>N</td>
<td>0.001</td>
<td>Grass</td>
<td>100</td>
<td>16.31</td>
<td>None</td>
<td>100</td>
<td>0.5455</td>
<td>2.875</td>
<td>57 – 700</td>
</tr>
<tr>
<td>U</td>
<td>0.00001</td>
<td>Rice</td>
<td>122</td>
<td>10</td>
<td>Cereal</td>
<td>25</td>
<td>0.4857</td>
<td>0.9975</td>
<td>20 – 552</td>
</tr>
<tr>
<td>UAN</td>
<td>0.0131</td>
<td>Grass</td>
<td>240</td>
<td>12.4</td>
<td>Tuber</td>
<td>50</td>
<td>2.198</td>
<td>2.916</td>
<td>35 – 364.5</td>
</tr>
<tr>
<td>UM</td>
<td>0.0146</td>
<td>Rice</td>
<td>192</td>
<td>8</td>
<td>Rice</td>
<td>120</td>
<td>0.5325</td>
<td>1.751</td>
<td>112 – 676</td>
</tr>
<tr>
<td>None</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

The mean EF of the overall dataset is 1.781%, higher than the IPCC estimate, but within the range of uncertainty given. A histogram of the dataset’s EF values explains this discrepancy (figure 8). The right skewed distribution of EF values shows that while most of the data are close to the IPCC default, there are a few large positive values in the distribution. These extreme values affect the dataset’s mean. Use of the median as a more robust measure of center is appropriate. The dataset median of 0.9% is very close to the IPCC default. While the measures of center values of EF are close to the IPCC default, tables 6 and 7 indicate that the minimum and maximum EF values of the overall data set and the subsets are far from the range of uncertainty. These discrepancies from the IPCC EF value suggest that the default may not accurately depict the amount of N₂O emitted from a given fertilizer N input.
The Multi-level Linear Model

As was previously mentioned, the usable N$_2$O database consists of measurements summarized from 164 separate papers. Each study was performed under different environmental conditions with different management variables. Even studies carried out at the same research station are subject to variation in soil properties, such as organic carbon, pH and minerizable nitrogen and fluctuating climatic conditions like temperature and rainfall, which influence the processes of nitrification and denitrification (Bouwman et al., 2002). These differences presumably account for the extreme values of EF found for some studies in the data set. The high variability in conditions present at the experiment sites prevents the use of a simple linear regression applied to the entire dataset. To accurately model the relationship between N input and N$_2$O emissions the use of a multi-level linear model (MLM), a model that takes the variation between studies into account, is necessary. The MLM utilized in this study can be thought of in two ways: (1.) a linear regression in which intercepts and slopes are allowed to vary by group or (2.) a regression that includes a categorical input variable representing group membership (Gelman and Hill, 2006 p. 247).

The general MLM with varying intercepts and varying slopes is represented as:

$$ y[i] \sim N(\alpha_j[i]+\beta_j[i]x_i,\sigma^2_y), \text{ for } i=1,...,n $$

With the i’s representing the individual and j[i] representing the group association of the i$^{th}$ observation.
The “group” could be one of a number of categorical variables taken from the N₂O database, including individual study (data reference), fertilizer type, crop type and soil type. To prevent confusion in the following explanation of the model, the categorical variable by which the regression coefficients will vary is the individual study.

Multi-level regression is a method of performing a data analysis between the two extremes of excluding a categorical predictor from a model (complete pooling) or estimating separate models within each level of the categorical predictor (no pooling) (Ibid, 248). Consider a model seeking to estimate the relationship between fertilizer nitrogen input and N₂O emissions. In this case, a complete-pooling model would estimate the linear regression across all studies (Ibid, 249). Different studies are treated as replications. The group indicator of the study is therefore not included. Complete pooling ignores the inherent variation among studies that produces different N₂O emissions and different relationships between N input and emission as seen in figure 9. The black line is the pooled regression model and the colored lines represent regression models for individual studies. This pooled regression model of the N₂O data does not represent the relationship seen in many of the individual studies and would not provide an accurate tool to estimate N₂O emissions in any particular study.

Figure 9

![Pooled and Un-pooled Regression](image)

A no-pooling model (colored lines) overstates the study level variation by providing separate regression models fit for each individual study (Ibid, 249). The validity of these models is limited by the small sample of N₂O measurements provided in each study (Ibid, 252). A linear regression performed on studies with fewer
measurements produces more variable estimates with higher standard errors (Ibid, 252). The model cannot accurately describe the relationship between N input and N₂O emissions based on the small number of measurements. Some studies in the database provide only a single N₂O emission measurement making use of a regression model impossible. Also, unpredictable weather and environmental conditions or human error may yield N₂O emissions measurements that are atypical of the measurement area. Emission predictions based upon one study period for an area could give results that are not representative.

Even using a study with a large number of measurements, the no-pooling regression model is not able to estimate N₂O emissions at other locations because of the variable site conditions. A site-specific model no matter how accurate is limited to estimation at that site. This model, like the complete-pooling model is not a useful tool to estimate emissions of N₂O. The natural variation of conditions over time and between studies would not be taken into account, leading to problems similar to the complete pooling model.

The MLM is known as a partial-pooling model because it is compromise between the complete-pooling and no-pooling models. The multi-level estimate for a given study \( j \) can be approximated as a weighted average of the mean of the N₂O emissions measurements in the study (the no-pooling regression) and the mean of the overall emissions measurements (the complete pooling regression) (Ibid, 249). The weighted average applied to each study depends on information in the individual study and the average of all the studies. Estimates from studies with smaller samples are weighted more heavily (stronger pooling) by the MLM and pull the regression estimate closer to the overall emissions average. Estimates from studies with larger samples are weighted less, and the regression estimate is closer to the individual study regression model (weaker pooling) (Ibid, 250, 253).

The MLM creates a weighted average regression for the dataset and individual group level regression lines for each study. This enables the MLM to make two different predictions based upon the dataset; 1.) the MLM can predict the N₂O emission for a new experiment within an existing study in the database and 2.) the MLM can predict the N₂O emission for a new experiment in a new study (Ibid, 267-268). The prediction capability of the MLM makes it a valuable tool for anyone trying to calculate N₂O emissions.

Figure 10 shows an example of a successful varying slope, varying intercept MLM using the same data seen in figure 9. The weighted average model (black line) is compared to 5 selected group level models (colored lines) from the dataset. The partial pooling has successfully taken into account conditional variation between each site and has demonstrated that the relationship between N input and N₂O emission has a constant slope. Variation in the intercept indicates that different sites have naturally different background emissions.
Results

Multilevel Linear Regression

The multi-level regression run in this study uses the varying intercepts and varying slopes model, with N input as the predictor variable, the average daily $\text{N}_2\text{O}$-N emissions as the response variable and the studies as the categorical group level predictor. Because both the $\text{N}_2\text{O}$ emission and N fertilizer application data are highly skewed (figures 3, 4 and 5 and in the figure appendix), the natural logarithm transformation is applied to both variables to reduce the leverage of large values. Total emission values from the database that were zero or negative were replaced with the arbitrary value of 0.00001 kg N/ha/period before being converted to an average daily emission value to enable the use of log transformation. The average daily $\text{N}_2\text{O}$-N emission measurements were transformed with the natural log, $\ln(g \text{ N}_2\text{O-N}/\text{ha/day})$ to produce a more normal distribution (figure 11, see figure appendix for subset histograms). Though there are still outliers in the data, they are not as distant.

In some of the crop type and fertilizer type data subsets, the skew of N input is not as severe and might prevent one from using a natural log transformation. However, to keep continuity among the response and predictor variables and make the regression equation easier to explain, all N inputs were log transformed. N input values were transformed with the formula; $\ln(N \text{ input/ha} + 1)$ because the natural log of zero yields an error term. The addition of one unit to each input value does not impact the data, because
the majority of input values are in the hundreds. A histogram of \( \ln(N \text{ input} + 1) \) shows a more normal distribution for the data and a cluster of \( N \) inputs of 1 kg N/ha associated with zero values (figure 12 see figure appendix for subsets).

Figure 11

![Overall Log(Average Daily Emissions) Histogram](image)

Figure 12

![Overall Log(N Input+1) Histogram](image)
Overall Model

A regression analysis of the entire dataset using the multi-level linear model to take into account conditional variation among study sites was performed to establish if the amount of N input directly effects N$_2$O emission independent of all other variables. The varying slope/varying intercept model yields the estimated regression line averaging over all groups: \( \ln(y) = 0.48 + 0.34(\ln(x+1)) \) with a standard error of 0.221 on the intercept coefficient and 0.038 on the slope coefficient. A model is also estimated within each individual group (study). Note: I will refer to the weighted average MLM given for the entire dataset as the “average model” and the weighted individual study MLMs as “group level models”. Figure 13, is a plot of the model’s random effects or the group level errors in both the slope and intercept. The coefficients of the average model are represented by the vertical lines and the deviation of the group model coefficients are represented by the horizontal lines. This indicates how much the intercept and slope in group level models have shifted up or down compared to the estimated coefficients of the average model.

Figure 13

![Graph showing the model's random effects for intercept and slope. The graph includes vertical lines indicating the average model coefficients and horizontal lines indicating the deviation of the group model coefficients.](image)
The intercept represents average log emission when input is zero. The very different intercept values or \( \text{N}_2\text{O} \) emissions at zero input, indicates that there are naturally large emission differences between studies without applied N input. The random effect plot also shows that the group level model’s slope coefficients are mostly positive and increase as the intercept decreases. Positive slope coefficients imply that as input increases \( \text{N}_2\text{O} \) emission will also increase. This is true for all of the study groups except for the three left-most, whose difference between the overall slope coefficients and their own is less than -0.344, which indicates that their individual group model slopes are negative.

A standard normal quantile plot of the Overall MLM’s residuals reveals that the residual distribution is not well described by the normal distribution. The deviation from the straight line suggests that the residual distribution has heavier tails than a normal distribution, a typical behavior from a cross-sectional data. The effect of heavy tails is that we may overestimate the standard error of the estimated model coefficients, thereby reducing the statistical power. For this study, a slight overestimation of coefficient standard error will lead to a more conservative statistical inference. In other words, if a slope is significant with heavy tail residual distribution, the actual significance level is higher.

Figure 14
Crop Type Models

Crop specific MLMs were used to account for categorical differences in crop type that might prevent measurements from being grouped into one model. A study done by Dobbie et al. (1999) found that the $\text{N}_2\text{O}$ emissions from fertilized small-grain cereals (wheat and barley) were consistently lower than from fertilized grassland, potatoes and broccoli, which had similar emissions. Others have found emissions from wheat to be lower than from corn (Bronson and Mosier, 1993). Separating the data by crop type also reflects differences in various management and environmental factors. The environment and climate suitable for certain crops are unsuitable for others. Crops also differ in their growing periods, and associated land-use, planting, fertilization, irrigation and harvesting practices. Crop based MLMs have a practical application in that a successful model would provide a valuable tool for comparative analysis and better management of $\text{N}_2\text{O}$ emissions. The findings of the above studies are not necessarily representative of all available data. Site-specific environmental and climatic conditions may be responsible for the apparent variation in $\text{N}_2\text{O}$ emission from different crop types. By sub-setting the data by crop type and using the multi-level linear model to take into account conditional variation among study sites, this analysis indicates if the type of crop planted influences $\text{N}_2\text{O}$ emission as a function of N input. The regression equation and standard error of each model are summarized in table 8.

<table>
<thead>
<tr>
<th>Crop Type</th>
<th>Regression Equation</th>
<th>SE Intercept</th>
<th>SE Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare</td>
<td>$\ln(y) = 1.26 + 0.14\ln(x+1)$</td>
<td>0.43</td>
<td>0.06</td>
</tr>
<tr>
<td>Barley</td>
<td>$\ln(y) = 0.99 + 0.11\ln(x+1)$</td>
<td>0.95</td>
<td>0.22</td>
</tr>
<tr>
<td>Cereal</td>
<td>$\ln(y) = 0.43 + 0.28\ln(x+1)$</td>
<td>0.39</td>
<td>0.09</td>
</tr>
<tr>
<td>Corn</td>
<td>$\ln(y) = 1.24 + 0.23\ln(x+1)$</td>
<td>0.25</td>
<td>0.04</td>
</tr>
<tr>
<td>Grass</td>
<td>$\ln(y) = -0.15 + 0.46\ln(x+1)$</td>
<td>0.43</td>
<td>0.08</td>
</tr>
<tr>
<td>Rape</td>
<td>$\ln(y) = -0.14 + 0.43\ln(x+1)$</td>
<td>0.93</td>
<td>0.18</td>
</tr>
<tr>
<td>Rice</td>
<td>$\ln(y) = 0.64 + 0.22\ln(x+1)$</td>
<td>0.58</td>
<td>0.08</td>
</tr>
<tr>
<td>Vegetable</td>
<td>$\ln(y) = 0.24 + 0.35\ln(x+1)$</td>
<td>0.48</td>
<td>0.09</td>
</tr>
<tr>
<td>Wheat</td>
<td>$\ln(y) = 0.13 + 0.32\ln(x+1)$</td>
<td>0.57</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Bare MLM

The random effects plot of the Bare MLM indicates that there essentially no differences between the slopes of the group level models and the average model (figure 15). The extremely small variation suggests that measurements on bare plots will be consistently similar independent of site location. A standard normal quantile plot shows the Bare MLM’s residuals to be close to normally distributed indicating that the model accurately describes the relationship between input and emission (figure 16).
Figure 15

Figure 16

Bare MLM
Cereal, Barley and Wheat MLMs

A plot of the Cereal MLM’s random effects indicates that there are relatively large differences in slope between the average model (0.28) and the individual group models which range from -0.075 to 0.844 (Figure 17). The Cereal dataset was broken down into two subsets of consisting of the most represented individual crop types; wheat and barley. The Wheat MLM has a slope of 0.32, while the Barley MLM has a slope of 0.11, which explains some of the variation between the average model slope and the group level model slopes. However plots of both the Wheat MLM’s and Barley MLM’s random effects (figures 18 and 19) also indicate large variation between the average model and the group level models.

Figure 17
Figure 18

Figure 19
Corn MLM

A plot of the Corn MLM random effects (figure 20) indicates that differences in slope between the average model (0.23) and the group level models (0.11 – 0.34), are significant.

Figure 20
Grass MLM

A plot of the Grass MLM random effects (figure 21) indicates that differences in slope between the average model (0.46) and the group level models (-0.285 – 1.758), are significant.

Figure 21

Rice MLM

The random effects plot of the Rice MLM displays little variation between the slope of average model and the slopes of group level models (figure 22). A standard normal quantile plot of the model’s residuals for the most part indicates normality and suggests that the model accurately describes the relationship between N input and N₂O emission (figure 23).
Figure 22

Figure 23
Vegetable and Oilseed Rape

The random effects plot of the Veg MLM displays significant variation between the slope of average model and the slopes of group level models (figure 24). The variation could be due to the large number of different individual crops contained within the Veg dataset. However, the Rape MLM; a model based on oilseed rape measurements, which make up more than half of the Veg dataset and Veg MLM, also displays significant variation between the average model slope and the group level model slopes (figure 25).

**Figure 24**
Fertilizer Type Models

Based on the relative contribution of denitrification and nitrification to N$_2$O emission and the different conditions under which each process is favored, it is logical to assume that the type of fertilizer applied may influence emissions. There are several studies which indicate that fertilizer type significantly affects N$_2$O emission. Under wet conditions, Clayton et al. (1997) found that N$_2$O emissions from NO$_3^-$ containing fertilizers were much higher than that from NH$_4^+$ containing fertilizers on grassland plots. The higher emissions were attributed to higher denitrification rates in the NO$_3^-$ amended soil. This study also found that N$_2$O emissions were higher from NH$_4^+$ and urea amended plots than from NO$_3^-$ fertilizers under dry conditions (Ibid). Relatively low N$_2$O emissions from cattle slurry compared to NO$_3^-$ and NH$_4^+$ fertilizers applied to grassland have been found (Tilsner, 2003 and Velthof et al., 1997). Plots amended with multiple forms of N-fertilizer may also produce different emissions. Combined application of manure and NO$_3^-$ fertilizer may enhance N$_2$O production, because the addition of available C with manure may increase denitrification of the NO$_3^-$ applied with the fertilizer (Clayton et al., 1997, Duxbury et al. 1982 and Velthof et al., 1996). The findings of the above studies may in some cases be conditional and are not representative of all available data. Other factors may be responsible for the apparent variation in N$_2$O emission from different fertilizer types. By separating the data by fertilizer type and using the multi-level linear model to take into account conditional variation among study sites, this analysis indicates if the form of N-fertilizer influences N$_2$O emissions as a function
of N input. The regression equation and standard error of each model are summarized in table 9.

Table 9

<table>
<thead>
<tr>
<th>Fertilizer Type</th>
<th>Regression Equation</th>
<th>SE Intercept</th>
<th>SE Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$\ln(y) = -0.70 + 0.52(\ln(x+1))$</td>
<td>0.81</td>
<td>0.16</td>
</tr>
<tr>
<td>AN</td>
<td>$\ln(y) = -0.038 + 0.50(\ln(x+1))$</td>
<td>0.61</td>
<td>0.12</td>
</tr>
<tr>
<td>ANM</td>
<td>$\ln(y) = 0.07 + 0.49(\ln(x+1))$</td>
<td>2.15</td>
<td>0.39</td>
</tr>
<tr>
<td>M</td>
<td>$\ln(y) = -1.60 + 0.80(\ln(x+1))$</td>
<td>0.93</td>
<td>0.17</td>
</tr>
<tr>
<td>N</td>
<td>$\ln(y) = 4.19 - 0.35(\ln(x+1))$</td>
<td>3.10</td>
<td>0.61</td>
</tr>
<tr>
<td>U</td>
<td>$\ln(y) = -5.52 + 1.43(\ln(x+1))$</td>
<td>1.83</td>
<td>0.36</td>
</tr>
<tr>
<td>UAN</td>
<td>$\ln(y) = 1.60 + 0.13(\ln(x+1))$</td>
<td>0.76</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Ammonia MLM

The random effects plot of the Ammonia MLM displays little variation between the slope of average model and the slopes of group level models (figure 26). A standard normal quantile plot of the model’s residuals is approximately normal which indicates that the model accurately describes the relationship between N input and N$_2$O emission (figure 27).
Ammonium Nitrate MLM

The random effects plot of the AN MLM displays little variation between the slope of average model and the slopes of group level models (figure 27). A standard normal quantile plot of the model’s residuals displays normality which indicates that the model accurately describes the relationship between N input and N\textsubscript{2}O emission (figure 29).
Figure 28

Figure 29
Ammonium Nitrate – Manure MLM

The random effects plot of the ANM MLM displays significant variation between the slope of average model (0.49) and the slopes of group level models (0.36 – 0.62) (figure 30). A standard normal quantile plot of the model’s residuals is approximately normal which indicates that the model describes the relationship between N input and N\textsubscript{2}O emission (figure 31).

Figure 30
Manure MLM

The random effects plot of the Manure MLM displays no significant variation between the slope of average model and the slopes of group level models (figure 32). A standard normal quantile plot of the model’s residuals displays normality which indicates that the model accurately describes the relationship between N input and N₂O emission (figure 33).
Nitrate MLM

The random effects plot of the Nitrate MLM displays no significant variation between the slope of average model and the slopes of group level models (figure 34). A standard normal quantile plot of the model’s residuals is approximately normal which indicates that the model accurately describes the relationship between N input and N$_2$O emission (figure 35).

Figure 34

Figure 35
Urea MLM

The random effects plot of the Urea MLM displays no significant variation between the slope of average model and the slopes of group level models (figure 36). However, a standard normal quantile plot of the model’s residuals indicates a degree of non-normality which indicates that the model may not accurately describe the relationship between N input and N\textsubscript{2}O emission (figure 37).

Figure 36
Urea – Ammonium Nitrate MLM

The random effects plot of the UAN MLM displays no significant variation between the slope of average model and the slopes of group level models (figure 38). A standard normal quantile plot of the model’s residuals is approximately normal which indicates that the model accurately describes the relationship between N input and N\textsubscript{2}O emission (figure 39).
Figure 38

Figure 39
Analysis/Discussion

Overall Model

In the Overall MLM, three group level models with negative slope coefficients (1, 109 and 139) did not fit the pattern of the rest of the data. These studies correspond to a high level of natural emissions and have the following equations of the regression line:

- Group 1: \( \ln(y) = 5.4663 - 0.18(\ln(x+1)) \)
- Group 109: \( \ln(y) = 4.2882 - 0.057(\ln(x+1)) \)
- Group 139: \( \ln(y) = 3.772 - 0.042(\ln(x+1)) \)

The group 1 study (Abbasi and Adams, 2000) was done on grass and compared three plots fertilized with 100 kg N/ha potassium nitrate (KN) to three plots fertilized with 250 kg N/ha diammonium phosphate (DAP). The soil properties of pH, % organic carbon and % nitrogen available were the same for all plots. Measurements on the KN plots were performed for 7 days while measurements on the DAP plots were performed for 28. Total emissions from the KN plots were 0.69, 7.85 and 16.2 kg N\(_2\)O-N/ha corresponding to WFPS values of 63%, 71% and 84%. Total emission from the DAP plots were 0.375, 1 and 7 kg N\(_2\)O-N/ha corresponding to WFPS values of 63%, 71% and 84%. Aside from the obvious effect WFPS had on total N\(_2\)O emissions, the main point of interest is that plots fertilized with KN had much higher emissions than plots fertilized with DAP even over a shorter measurement period. The high emission values of the KN plots corresponding to the relatively lower N input resulted in the negative slope given by the MLM. The difference in emissions suggests that the type of fertilizer used may be influential on N\(_2\)O emissions and should perhaps be separated when modeling the relationship of N input to N\(_2\)O emissions.

The group 109 study (Nyborg et al., 1997) was done on bare soil and compared two control plots to two plots with input from KN. The studies were performed on the same soil type and the soil properties of pH, % organic carbon and % nitrogen available were the same. The only difference was that one experiment was performed in 1989 and one experiment was performed in 1990. The 1989 experiment yielded total emissions of 1.34 kg N\(_2\)O-N/ha from the control and 3.5 kg N\(_2\)O-N/ha from a plot fertilizer with 57 kg N/ha over 33 days of measurement. The 1990 experiment yielded total emissions of 6.03 kg N\(_2\)O-N/ha from the control and 16.31 kg N\(_2\)O-N/ha from a plot fertilizer with 100 kg N/ha over 38 days of measurement. While one might expect climatic conditions at a site to be fairly stable, the large differences between the 1989 and 1990 measurements on the same soil suggests that there may have been an environmental factor not reported such as weather that may have influenced N\(_2\)O emissions. The high emissions from the control plot in 1990 caused the MLM to estimate a negative slope for this group. Yearly differences in environmental factors may need to be accounted for in the model by breaking studies down into smaller groups based upon the year they were performed. This would however reduce the sample size of many groups causing the model to more completely pool the data and take into account less variation between groups.
The group 139 study (van Cleemput et al., 1994) was done on grass and compared a plot fertilized with 200 kg N/ha as KN and a plot fertilized with 295 kg N/ha as AN to a control. The studies were performed on the same soil type and the soil properties of pH, % organic carbon and % nitrogen available were the same. The KN plot yielded a total emissions of 6.4 kg N$_2$O-N/ha, the AN plot yielded a total emissions of 8.4 kg N$_2$O-N/ha and the control plot yielded a total emissions of 4.5 kg N$_2$O-N/ha. The only difference was the days over which they were measured. The KN plot and the control were measured over 50 days while the AN plot was measured over 280 days. When calculating the average daily emission to be used in the MLM the emissions from the AN plot is made lower than the KN and control plot due to the larger number of days it was measured. The large difference in measurement periods and the N inputs of the AN and KN plots compared to their total emissions suggests that KN fertilizer produces relatively more N$_2$O over a shorter period of time and with less total N input. The existence of such a relationship supports the idea that N$_2$O emissions models should be split categorically by fertilizer type.

The significant deviation in slope coefficients seen the Overall MLM random effects plot (figure 13) and the non-normal distribution of the Overall MLM’s residual in the standard normal quantile plot (figure 14), indicates that the relationship between N input and N$_2$O emissions is not independent of all variation between studies. In reaction to the failure of the Overall MLM, the data were broken into previously mentioned subsets by crop type and fertilizer type in an effort to isolate any systematic variation caused by these variables that would affect the analysis.

Crop Type Models

With two exceptions; the Bare MLM and the Rice MLM, the crop type datasets displayed large variations in slope between their average model and their individual group models. The large variation in slopes indicates that breaking the Overall dataset into subsets based on crop type does not minimize the slope variation seen in the Overall MLM. This indicates that crop type differences probably do not influence N$_2$O emissions. The large variation in slopes prevents one from accurately estimating N$_2$O emissions based upon these models. An examination of the two successful crop type models may indicate what caused them to display no significant variation in slope coefficients.

The Bare MLM consists of 51 observations contained within 13 groups (studies). The studies were performed in a variety of regions with different environmental conditions, including the U.S., Australia, Finland, Canada, Spain, France, Sweden, Germany and the UK. Likewise, the Rice MLM which consists of 85 observations contained within 13 groups includes data from both flooded rice and upland rice studies performed in a multitude of places (U.S., China, India, Australia, Japan, the Philippines and Indonesia). Measurements in both datasets were taken in an assortment of climatic zones and under diverse management conditions. It is therefore unlikely that environmental similarities produced the slope similarities. The only consistency among the studies in each dataset was the in the predominant use of one fertilizer treatment. Sixty-five percent of the measurements in the Rice data set were made on plots that had been fertilized with urea or urea in combination with another fertilizer. The over-
representation of urea fertilizer compared to others may have produced the extremely low variation between the slope of the average model and the individual group model slopes. The Bare dataset also displays some consistency of the type of fertilizer used. Over one third of the measurements in the Bare dataset were made on unfertilized plots and 27% of the measurements were taken on plots fertilized with ammonia accounting for 62% of the observations in combination. However, 53% of the observations in the Grass MLM which displayed the most slope variation are made up of only two fertilizer forms. It seems unlikely that the slope similarity seen in the Bare MLM was caused by fertilizer choice.

**Fertilizer Type Models**

With the exception of the ANM model, the random effects plots for each fertilizer type show no significant variation between the average model slope coefficient and the group level model coefficient. The similarity of slope in each model over a variety of studies and conditions indicates that the form of N fertilizer used heavily influences N\textsubscript{2}O emissions. Because the analysis was done after transforming the data with the natural log, the increase in emissions for a given increase in N input is multiplicative not additive.

The model regression equation:

\[ \ln(y) = \beta_0 + \beta_1 \ln(x+1) \]

is the same as:

\[ y = e^{(\beta_0)\ln(x+1)}^{\beta_1} \]

for a baseline input of \( x_0 \), the emission is:

\[ y_0 = e^{(\beta_0)\ln(x_0+1)}^{\beta_1} \]

This equation indicates that the relationship between average daily emissions (\( y \)) and N input (\( x \)) is that for every 1 percent increase in \( x \) from the baseline value (\( x_0 \)) \( y \) increases by \( \beta_1 \) percent:

\[ y_1 = e^{(\beta_0)\ln[(x_0+1)(1.01)]}\beta_1 = y_0 * (1.01)^{\beta_1} = y_0 * (1.01)^{0.52} \]

An example using the A MLM equation helps illustrate this point. The baseline is calculated on an input of 1 (\( x_0 = x+1 \)). With baseline equation; \( y_0 = e^{(-0.70)\ln(1+1)}^{0.52} \), we get 0.497 g N/ha/day as our baseline emission value. For a 1% increase in \( x \) from this baseline value, \( y \) increases by ~0.52%. This yields us a value of 0.500 g N/ha/day.

\[ y_1 = y_0 * (1.01)^{0.52} \]
Although the intercepts ($\beta_0$) of the fertilizer type models are different resulting in very different baseline emissions for each model, the slope coefficient ($\beta_1$) indicates how much additional N$_2$O emissions we can expect for a given increase in N input. The regression equations in table 10 show that the A, AN, and ANM models have similar slope coefficients and thus will produce a similar increase in emissions for a given N input. The analogous relationships across fertilizer types contradict the findings of previously mentioned studies, which indicated that the N$_2$O emissions from AN should be larger than A and that N$_2$O emissions from ANM should be larger than both. The MLM (manure) slope contradicts the finding that plots fertilized with manure produced relatively lower emissions than plots amended with chemical fertilizers. The inconsistency between the models and the study findings does not indicate that the studies are incorrect, only that the relationships found in each study may be caused by factors other than the form of N used in fertilization.

**ANM (Ammonium Nitrate – Manure) MLM**

Differences in the slope coefficient between the average model and the 10 group level models in the ANM dataset ranged from (-0.125 to 0.133). The group level slope coefficients are all positive which indicates that the relationship between input and emission is positive. The high variability between slope coefficients could be caused by; (1) the inconsistent composition of manure across studies and/or (2) the low number of studies included in the analysis. Manures are a mixture of mineral N, organic N and C compounds, salts and water. Manures and crop residue are known to stimulate the denitrification process by increasing organic C availability (Aulakh et al., 1991 and Beauchamp et al., 1989). Animal manures have large variability in mineral composition due to differences in animal species source and nutrition (Chadwick et al., 2000). Although the inorganic N input was standardized to a kg N/ha value in each study, differences in mineralizable C could be responsible for the variability in slope. However, the manure MLM does not exhibit a large variation in model slope coefficients, which implies that differences in organic C are unlikely to be the source of discrepancy. Slope variation caused by the small sample of studies can be addressed in future work by adding data.

**N (Nitrate) MLM**

The Nitrate MLM was the only model to produce a negative slope coefficient. The random effects plot of the model shows no significant variation between the average and group level slopes indicating that extreme outliers are not affecting the average model. Plots of average daily emission and nitrate-N input both in the natural log scale and the original scale indicate why the slope is negative (figure 40 and 41).
Figure 40

$log (\text{Av Daily Emission}) \text{ vs } log (\text{N Input}+1)$

Figure 41

$\text{Av Daily Emission vs N Input}$
Average daily N₂O-N emissions values increase to a peak at 80-100 kg N input then drastically decrease, but continue a positively increasing slope at higher input values. The Nitrate dataset consists of 18 separate studies, including additional data in future analysis may smooth out the large peak. There is a possibility that the use of the average daily value exacerbated differences in emissions for studies with high total N₂O-N emissions and short study periods. Abbasi and Adams (2000) measured emissions of 0.69, 7.85 and 16.2 kg N₂O-N/ha over seven days giving average daily emission values of 98.57, 1121.43 and 2314.27 g N₂O-N/ha/day respectively. However, a plot of total study period emissions and N input indicates a spike in the distribution around input values of 100 kg N/ha existed before the conversion to average daily emission (figure 42).

Additional work on a Nitrate Model may require splitting the dataset into two N input subsets; 1-100 kg N/ha and >100 kg N/ha to correctly describe the relationship between nitrate N input and N₂O emissions.

**Figure 42**
UAN (Urea – Ammonium Nitrate) MLM

The low slope of the UAN model compared to the slopes of U and AN models; the two fertilizers included as a combined application in the UAN dataset, warrants a more in-depth examination of the data to search for a possible explanation. The random effects plot and standard normal quantile plot (figures 38 and 39) show no indication that the model is inaccurate. Plots of average daily emission and nitrate N input both in the natural log scale and the original scale indicate why the slope is relatively low (figures 43 and 44).

Figure 43
Average daily N\textsubscript{2}O emission values corresponding to N inputs >200 kg N/ha are generally lower than the emission values corresponding to N inputs <200 kg N/ha. The extremely high emission value at an input of 130 kg N/ha also helps flatten the slope. There is a possibility that the use of the average daily emission value artificially lowered N\textsubscript{2}O-N emissions at high N input values with long sample periods or artificially increased N\textsubscript{2}O-N emission values corresponding to lower N inputs with shorter sample periods. A plot of total study period emissions and N input does show that the high value at 130 kg N/ha input was made more extreme by calculating the average daily emission (figure 45). However, the same general relationship exists as depicted in the two previous plots. The UAN MLM included only 15 separate studies with varying combinations of urea – ammonium nitrate fertilizer inputs. Perhaps the dataset should be further broken down based on relative contribution of each fertilizer type. Before future work on the model can be pursued additional data is needed.
Manure (M) and Urea (U) MLMs

The M MLM and U MLM had the highest slope coefficients; 0.80 and 1.43 respectively. The large difference between the slopes of these two MLMs and the others, although significant, is explainable. As was previously mentioned, manure fertilizers contain organic C compounds in addition to inorganic N which has been shown to enhance N₂O emissions. The availability of C supports higher levels of activity in denitrifying microbes and subsequently causes microsite anaerobiosis which encourages denitrification as the dominant microbial process, producing more N₂O emissions than a site experiencing higher levels of nitrification (Azam et al., 2002). Urea also includes organic C in its chemical formula (NH₂)₂CO which could produce the same augmenting effect on emissions.
Conclusion

Based upon the multi-level linear analysis of the dataset, this study has shown that the relationship between N input and N₂O emissions is not independent of all conditional variation between studies. While the categorical analysis of differences in crop type did not have a significant influence on N₂O emissions, the categorical analysis of N fertilizer forms usually showed a significant influence on emissions.

The A, AN, M and U models appear to accurately describe the relationship between different forms of N input and N₂O emissions. These models are useful to estimate the relative amount of emission between different fertilizer forms and could aid agricultural best management practices.

The ANM model, which displayed significant slope variations, and the UAN and N models, which depicted significant relationships that contrasted with scientific knowledge, were derived from relatively small samples. It is possible that the UAN and N models accurately describe the relationship between their respective forms of N input and N₂O emissions and that the variation in the ANM model may be caused by inconsistencies in manure composition. However, before drawing any conclusions, additional work with a greater number of measurements is necessary.

The successful MLMs have some drawbacks. Transformation of the data by the natural log, while necessary to perform a linear regression analysis, causes the relationship between N input and average daily N₂O emission to be multiplicative, not additive. The models do yield important comparative increases in emission relative to the baseline. However, soil baseline N₂O emissions need to be known in order to utilize the equations.

The regression response variable; average daily N₂O emissions (g N₂O-N/ha/day), lacks definite estimation ability. An average daily emission value is successful in comparing emissions across fertilizer types, but without knowing the number of days N₂O emissions will be influenced by the presence of fertilizer N, the average daily emission value cannot be used to estimate a total N₂O emission value from a plot. The total value is important for anyone trying to compile a greenhouse gas inventory or calculate greenhouse gas credits that might be used under a cap and trade carbon credit system.

Despite its faults, the analysis performed in this study yielded important insight into the factors influencing N₂O emissions from agricultural soils. These findings can be used to guide both management strategies and further research dealing with an increasingly important topic.
Reference:


Van Cleemput O, Malkanti RM, d’Ydewalle Y and Baert L (1990) Denitrification influenced by incorporated harvest residues


http://www.epa.gov/climatechange/emissions/downloads06/07Agriculture.pdf

http://www.grida.no/climate/ipcc_tar/wg1/127.htm

US GAO United States General Accounting Office (September, 2003) Natural Gas: Domestic Nitrogen Fertilizer Production Depends on Natural Gas Availability and Prices, Report to the Ranking Democratic Member, Committee on Agriculture Nutrition and Forestry, U.S. Senate, GAO-03-1148: 1-32

Data:


Models:

Database Reference:

Abassi, MK and Adams WA (2000) Gaseous N emission during simultaneous nitrification-denitrification associated with mineral N fertilization to a grassland soil under field conditions, *Soil Biology and Biochemistry*, **32**: 1251-1259


Breitenbeck GA, Bremner JM (1986) Effects of rate and depth of fertilizer application on emission of nitrous oxide from soil fertilized with anhydrous ammonia, *Biology and Fertility of Soils*, 2:201-204


Christensen, S (1983) Nitrous oxide emission from a soil under permanent grass: seasonal and diurnal fluctuations as induced by manuring and fertilization, Soil Biology and Biochemistry, 15: 531-536


Clough TJ, Sherlock RR, Cameron KC and Ledgard SF (1996) Fate of urine nitrogen on mineral and peat soils in New Zealand, Plant and Soil, 178: 141-152


Thornton FC and Valente RJ (1996) Soil Emissions of Nitric Oxide and Nitrous Oxide from No-till Corn, Soil Science Society of America Journal, 60 (4): 1127-1133


Van der Weerden TJ, Sherlock RR, Williams PH and Cameron KC (1999) Nitrous oxide emissions and methane oxidation by soil following cultivation of two different leguminous pastures, Biology and Fertility of Soils, 30: 52-60


Velthof GL, Brader AB and Oenema O (1996a) Seasonal variations in nitrous oxide losses from managed grasslands in the Netherlands, Plant and Soil, 181: 263-274


Weiske A, Benckiser G and Ottow JCG (2001) Effect of the new nitrification inhibitor DMPP in comparison to DCD on nitrous oxide (N2O) emissions and methane (CH4) oxidation during 3 years of repeated application in field experiments, *Nutrient Cycling in Agroecosystems*, **60**: 57-64


Data taken directly from


Hutchinson et al., (1993)


Parton et al., (1988)

Ruser et al., (2000)


Teepe et al., (2000)
Van der Weerden et al., (2000)
Watanabe et al., (2000)
Whalen et al., (2000)
Williams et al., (1999)