DISSERTATION

REGIONAL AND NATIONAL-SCALE ANALYSIS OF CROPLAND CARBON CYCLING

Submitted by

Erandathie Lokupitiya

Graduate Degree Program in Ecology

In partial fulfillment of the requirements
for the Degree of Doctor Of Philosophy
Colorado State University
Fort Collins, Colorado
Summer 2006

COLORADO STATE UNIVERSITY

July 06, 2006

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY ERANDATHIE LOKUPITIYA HERE ENTITLED ANAYLYSIS OF CARBON DYNAMICS OF US AGRICULTURAL SOILS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Richard Conant

Scott Denning

Michael Lefsky

Adviser Keith Paustian

Director Indy Burke/ Bill Laurenroth

ABSTRACT OF DISSERTATION

REGIONAL AND NATIONAL-SCALE ANALYSIS OF CROPLAND CARBON CYCLING

Increased global greenhouse gas (GHG) emissions, including carbon dioxide (CO_2), are known to contribute to global warming. Previous research has found carbon (C) sequestration in agricultural soils as a potential way of mitigating atmospheric CO_2 emissions.

In the first part of this study, I evaluated the methods used by Annex 1 (developed) countries in inventorying the sources and sinks of agricultural soil GHG emissions. In the second part, I assessed cropland soil C balance and C storage, considering residue C inputs and CO_2 output from soils, at regional and national scale. One of the main components in this study was estimating the crop residue C inputs, using available county-level yield and area data for major US crops during the period 1982-1997. Since the existing annual data reported by the National Agricultural Statistics Service (NASS) have a large number of gaps (missing data), I filled those gaps by using regression analyses with the data from the Census of Agriculture, and a suite of linear mixed effect models that incorporated county level environmental and economic variables. Then these comprehensive crop yield and area databases were used to estimate C inputs from crop residues (and cropland NPP).

Interannual and spatial variability of residue C inputs were studied in relation to changes in production, weather and climate. I also evaluated the potential use of Advanced Very High Resolution Radiometer (AVHRR) Normalized Difference Vegetation Index (NDVI) data in estimating the crop aboveground biomass and residue C inputs. Finally I estimated the soil C stocks and annual stock changes in cropland soils (and CO_2 loss due to decomposition), by using the Introductory C Balance Model (ICBM), and evaluated the C storage in soil and overall C balance over the US cropland. During the 16-year period, we found an average cropland NPP and residue C input rate of 504 and 312 Tg C yr^{-1}, respectively. The interannual variability of C stocks ranged mostly within 20 Tg, and the overall C stocks increased by 14 Tg towards the end of the study period, implying the importance of cropland C dynamics in the overall C cycle.

Erandathie Lokupitiya
Graduate Degree Program in Ecology
Colorado State University
Fort Collins, Colorado 80523
Summer 2006

ACKNOWLEDGEMENT

I want to thank several people who have helped me reach this far. First I want to express my gratitude to Dr. Ted Elliott (deceased), for all the support, interest and willingness he had for accepting me as one of his students. Although I was not fortunate enough to see him, I am so grateful to him for everything he did during his last days in life, to secure my future as a student at CSU.

I would like to extend my deep gratitude to my advisor, Dr. Keith Paustian, for all his advice, guidance, efforts, and support for my success as a Ph.D. student under him. He has been a great advisor and a mentor, who has always helped to get the best out of me. He has been a great inspiration, and I am so grateful to him for directing me in the right path during the last few years. I would also like to acknowledge my graduate committee, Drs. Rich Conant, Scott Denning, and Michael Lefsky for all the advice, support and time they have spent on me, and leading me through to achieve certain targets and goals as a Ph.D. student. I am so grateful for their honest efforts in training me to develop as a scientist.

My research took close to four years of time, and there are so many people who extended their help and support during that period. I would like to extend my heartfelt thanks to Dr. Jay Breidt for his advice in statistics on certain major parts of my research. I am also grateful to Mark Easter, Steve Williams, John Norman, Sunil Kumar, and Kay Dudek for their help and support. I would like to extend my sincere thanks to everybody in GDPE, NREL and the Dept. of Soil and Crop Sciences, who have given me the opportunity to pursue my studies in such a pleasant and friendly environment.

I am grateful to Drs. O.A. Andren, T. Kätterer, S.R. Koontz, Sian Mooney, and all the people including the inventory specialists and officers at the national greenhouse offices in a large number of countries, and officials at the USDA (NASS, ARS, and ERS), who helped me by providing necessary information, help and sources for my research.
Thanks to CASMGS for funding my research.

Last, but not least, I want to thank my husband Ravi, all my friends, and my family back home for all their love and support.

WITH HEARTIEST GRATITUDE

TO

TED ELLIOTT

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Global warming due to increased carbon dioxide and other greenhouse gas (GHG) emissions has become a major environmental concern. Carbon dioxide concentration in the atmosphere has increased from 280 ppmv in 1800 to 368 ppmv in 2000, largely due to human activities (IPCC, 2001a). While carbon dioxide from fossil fuel combustion constitutes the dominant source of anthropogenic emissions, land use and agriculture are also significant greenhouse gas sources (IPCC, 2001a), for CO_2 as well as other gases (i.e. nitrous oxide and methane). However, terrestrial ecosystems, including agricultural systems, can be both sources and sinks for CO_2. Soils constitute the single largest carbon stock in the terrestrial biosphere and hence improved knowledge of soil carbon dynamics is crucial for a clearer understanding of human perturbations of the global C cycle.

Croplands contain the most intensively managed soils, which globally are estimated at around 170 Pg of carbon in the upper 1 m (Cole et al. 1996). With increasing emphasis on the need for mitigating global GHG emissions, considerable research has focused on the potential usefulness of cropland as a sink for CO_2 (e.g. Cole et al., 1996; Lal et al., 1998; Paustian, 1997). The process of increasing the amount of C stored in soils (as well as in long-lived biomass such as trees) is often termed 'carbon sequestration'. Because the C entering the soil is derived from CO_2 fixed by plants, carbon sequestration constitutes a net removal of CO_2 from the atmosphere. Soil carbon sequestration can be achieved by adopting management practices that increase the amount of C added to soils and/or reduce the C emitted (e.g. from microbial decomposition) from soils. Various studies have found that management practices such as no till or reduced tillage, crop rotation, fertilization and organic amendments, etc., can help significantly increase the carbon storage within agricultural soils (Burke et al., 1995; Paustian et al., 1995, 1997). At global scale, potential C sequestration in agricultural soils through changes in land use management practices has been estimated at 600- 900 Tg yr^{-1} (Cole et al., 1996).

My research has focused on improving estimates of soil C dynamics in agricultural ecosystems at regional and national scales, in the United States. A main objective of the work was to provide better quantification and understanding of the role of cropland soils in GHG emissions and in GHG mitigation through C sequestration.

International cooperation for mitigating GHG emissions was strengthened by the adoption of the United Nations Framework Convention on Climate Change (UNFCCC) in 1992. Member countries of the UNFCCC now prepare national inventories of GHGs to evaluate the emissions by sources and removal by sinks. Thus, the first part of my research (*Chapter 2*) involved a review and analysis of methods currently used by UNFCCC member states to inventory GHG emissions from agricultural soils.

The availability of spatially-explicit data for estimating C inputs is essential in developing accurate inventories and estimating accurate C balances for cropland ecosystems. Currently in the US, there is extensive statistical data on annual agricultural production dating back several decades, which represents a very rich data source. However, these data were not specifically intended to address C balance questions and other limitations exist that need to be dealt with to apply the data to analyzing ecosystem C dynamics. The main sources of agricultural statistics in the US are annual surveys (including yields and crop areas) conducted by the National Agricultural Statistics Service (NASS). A more comprehensive data set, but collected only every five years, is provided by the US Census of Agriculture (Ag Census).

As with virtually all large databases, gaps in both spatial and temporal coverage, limit the utility of the extant databases. Over the past 2-3 decades, satellite-based sensors have been deployed with capabilities to estimate vegetation attributes, including productivity (Campbell, 2002**;** Doraiswamy et al., 2001, 2003, 2004, 2005; Hill and Donald, 2002). Thus, remote sensing represents an alternative or complementary data source to fill in gaps in ground-based surveys. As part of my research *(Chapter 3)*, I investigated the potential of remote sensing, using a vegetation index based on the Advanced Very High Resolution Radiometer (AVHRR), to estimate crop production in an agriculturally dominated state, Iowa.

Empirical modeling approaches can also be used to fill data gaps and process agricultural statistics for use in C cycle analysis. In the third part of my dissertation *(Chapter 4)*, I derived and tested a variety of statistical models to combine NASS and Ag Census data and to fill gaps in yield and crop area estimates. The new synthetic database was then used, with empirical relationships relating crop biomass components to yield measurements, to analyze the spatial and temporal patterns of crop net primary production and crop residue inputs to soil for the whole US (*Chapter 5*), for a 16 year period (1982-1997).

The change over time in soil C stocks, and hence the net exchange between the vegetation-soil system and the atmosphere, is determined by the balance of C inputs and losses of C (predominantly through decomposition). Currently there is no national measurement network in the US to directly measure soil C stock changes over time. Hence, soil C balance estimates at regional or national scales use model-based approaches (Paustian et al., 2002; Eve et al., 2001, Ogle et al. 2003). I used a simple dynamics simulation model, the Introductory Carbon Balance Model (ICBM) (Andren and Kätterer, 1997, Andren and Kätterer, 2001, Andren et al., 2004), to analyze spatial and temporal trends in US cropland soil C (*Chapter 6*), where the model was driven, in part, by the estimates of C inputs described above. A particular focus of the study was to analyze the interannual variability of projected changes in soil C stocks and the implications of this variability for assessing compliance with mitigation policies.
An overview of the major findings and conclusions of these studies is provided below.

National inventory methods by the Annex 1 countries in estimating agricultural soil greenhouse gas emissions (Chapter 2).

So far, the developed countries have been responsible for the majority of global GHG emissions; these countries are referred to as 'Annex 1' countries under the United

Nations Framework Convention on Climate Change (UNFCCC). According to the provisions of the UNFCCC, parties to the convention need to prepare national inventories of GHGs (excepting those controlled by the Montreal Protocol), including the emissions by sources and removal by sinks, and submit those to the Conference of Parties (COP), the convention's principal policy making body. These inventories need to follow the revised 1996 guidelines (IPCC, 1997a,1997b,1997c) provided by the Intergovernmental Panel on Climate Change (IPCC), as elaborated by the Good Practice Guidance (IPCC-GPG, 2000, 2003; IPCC, 2001b), and the UNFCCC reporting guidelines on annual inventories.

I studied the methods used by Annex 1 countries (a total of 39 countries, except for Turkey who recently ratified the convention) under the UNFCCC in estimating the GHG emissions/sinks in agricultural soils, for reporting in their annual national GHG inventories. The IPCC guidelines (IPCC-GL) outline methods structured according to three tiers, with increasing complexity and data requirements for higher tiers: 1) Tier 1 consists of simple equations and global default emission factors provided in the IPCC-GL (and IPCC-GPG); 2) Tier 2 uses the IPCC-GL default equations but requires country-specific parameters that better account for local climate, soil, management and other conditions; and 3) Tier 3 methods are based on more complex models and inventory systems, typically using more disaggregated activity data that better capture variability in local conditions (IPCC, 2000, 2003). I assessed to what extent Annex 1 countries have proceeded to develop county-specific Tier 3 methods in estimating agricultural soil GHG emissions for their national inventories. I found that 82% of the countries report nitrous oxide (N_2O) emissions in the inventories, and about 56% of the countries still use the default method recommended by the IPCC. Less than 35% of the countries report CO_2 emissions, and only a handful of countries have developed advanced Tier 3 methods. These Tier 3 methods mostly incorporate complex, process-based models. Lack of spatially explicit activity data is a major constraint for implementing higher tiered, country-specific methods for inventory purposes.

Studying the C dynamics of US croplands (Chapters 3-6).

Chapters 3-6 of my dissertation focused on the regional and national scale C dynamics of US cropland over a 16-year period (1982-1997). This study period was chosen based on the availability of digital data in both National Agricultural Statistics Service (NASS) and the Census of Agriculture (Ag Census), the two main national agricultural statistical databases. In Chapters 3-5, I mostly focused on estimating and evaluating the spatial and temporal variation of crop residue C inputs to cropland soils by filling the gaps in available county-level annual crop statistics. In the final chapter (*Chapter 6*) I focused on studying the overall C dynamics in the US cropland, considering the balance between the annual C inputs and outputs from the US cropland soils, with particular emphasis on the interannual variability and implications on short-term C dynamics in the cropland ecosystem.

In order to achieving the above research targets, I decided to use the annual crop statistics reported by NASS as a basis for a new synthetic database of crop NPP and residue C inputs, since it conducts and reports annual surveys of production and area for all major crop species in the US. To develop a comprehensive database (without gaps), I explored the potential for augmenting the data with information from remote sensing (*Chapter 3*) as well as statistical modeling to fill gaps and derive estimates of residue C inputs (*Chapters 4 and 5*).

I evaluated the potential for using of remotely sensed information in estimating aboveground biomass and residue C inputs, during the years when NASS has not reported yield data. In several past studies, remotely sensed information has been used in estimating crop yields and aboveground biomass (Hill and Donald, 2002; Hansen and Schjoerring, 2003; Doraiswamy et al., 2001, 2003, 2004, 2005). As described in Chapter 3, I used canonical correlation analyses between Normalized Difference Vegetation Index (NDVI) and biomass variables, followed by best subset multiple regression analyses, to develop model relationships between NDVI and aboveground biomass. Biweekly AVHRR NDVI (1 km resolution) and crop biomass derived using allometric equations relating yields to crop biomass (Williams and Paustian, submitted) were used as the input datasets. Canonical correlation creates a new set of canonical variates from the original NDVI in which the information from highly correlated, temporally close NDVI are combined. This approach was chosen as the best approach since it removed any effects from the multicollinearity among temporally close NDVI.

In the canonical correlation analyses, a cross-validation approach was taken, using the biomass data (estimated from the yields reported by NASS) for Iowa under three scenarios, in which 10, 20, and 40% of the data were made randomly missing.

Regression models were derived between the crop biomass and the canonical variates of the training (non-missing) data. The ability of using the NDVI data from one, two and three years in estimating the missing biomass data within the same period and a different period that fall outside the time period used in model derivation, were evaluated, as detailed in Chapter 3.

NDVI was positively correlated with crop biomass and original NDVI pixel values during the period from end June to end August; however, this correlation varied in value among the crops and different years, ranging from 0.1 to 0.84. It was found that combining data from all three years is the best way in predicting the missing data in any of the participating years. The estimated values under the all three scenarios were close to the observed values with less than 5% relative error, and the approach we used was found to be very effective in predicting biomass from NDVI, especially in the presence of missing data.

Several past studies had estimated the C sequestration potential in the US (Eve et al., 2001; Sperow et al. 2003; Ogle et al., 2003). Due to the large range of variation in the estimates by these studies, we felt the need of using a comprehensive crop database for having more accurate estimates of C stock changes in cropland soils. In this effort, I assumed that any influence in management would be reflected in the observed crop productivity over the study period that we chose (i.e. 1982-1997).

Thus, to estimate C inputs and soil C stocks over the entire study period, I developed comprehensive yield and crop area databases for major US crops, as detailed in Chapter 4, by filling the gaps in the county-level crop areas and yields reported by NASS during 1982- 1997. I evaluated several statistical approaches to fill in the gaps in NASS crop yield and acreage data over this period. Leave-one-out and leave-k-out procedures were used to find out the most suitable statistical method (Lokupitiya et al., 2006), and regression analyses between NASS and Ag Census crop yield data and multiple imputation technique in SAS (version 8.2) were found to be the best methods. I filled some of the gaps in county-level yields and crop areas by using regression analyses with the data from Census of Agriculture (reported every 5 years). The remaining gaps were filled using linear mixed effect models incorporating county-level environmental and economic variables. These models were run at Land Resource Region (LRR) level. The environmental variables used in filling the the models for gaps in yield data included mean monthly summer temperature, annual precipitation, precipitation/potential evapotranspiration ratio, and irrigated/total crop area ratio, and the economic and environmental variables used for filling the gaps in crop area data included crop price, fertilizer cost (unit cost of anhydrous ammonia), diesel cost, and precipitation of the preceding year. This way I developed comprehensive county-level yield and crop area databases for major US crops by filling all the gaps in the data reported by NASS. A thorough quality assurance and quality control procedure was followed on the final data created, and the crop data were compared against the data from the National Resource Inventory (NRI).

Then using those databases, I estimated and studied the interannual variability of residue C inputs in relation to trends in cropland productivity, including NPP, and variation in weather over the 16-year period (*Chapter 5*). Residue C inputs were estimated by using the dry biomass in yields, and allometric relationships incorporating information from harves indecs and root: shoot ratios for specific crops, in deriving residue C inputs from yield dry matter (Williams and Paustian, submitted). The analyses on the temporal and spatial variation of estimated residue C inputs were performed for different Crop Production regions (CPRs) and the country as a whole.

I found an overall increase in residue C inputs among the different CPRs over the period despite the interannual variability caused by weather changes. The residue C inputs showed an inverse relationship with increase in growing season temperature over the study period. Among the different crop types, corn had the largest C inputs, followed by soybean, wheat, and hay crops. The average cropland NPP over the study period was 504 Tg C yr^{-1}, and the highest NPP was observed in 1994 (i.e. 570 Tg C yr^{-1}), which was about 40% of the CO_2-C emissions from fossil fuel burning by the country during that year. The estimates were close to the NPP estimates by other recent studies (Lobell et al., 2002). The findings of the current study implicated the importance of the US cropland, especially the North Central region and Central and Northern Plains that had the largest crop areas, in absorbing atmospheric CO_2-C and the overall US C cycle.

In the final part of the study (*Chapter 6*), I modeled the decomposition dynamics of the residue C inputs I estimated, using the Introductory C Balance Model (ICBM; Andren and Kätterer, 1997, Andren et al., 2004). ICBM includes two soil organic matter pools and four parameters other than residue C inputs. The parameters include two decomposition coefficients for 'young' and 'old' soil C pools (k_Y *and* k_O), a humification coefficient (h), and an external influence coefficient (r_e) to represent any impact from changes in weather including temperature and soil water balance. The daily variation in soil moisture, temperature, and potential and actual evapotranspiration information were used in estimation of r_e considering the differences in crop type, soil type, and practice (irrigated vs rain fed crops); the daily values were averaged to represent the county-level annual environmental variation which represented the environmental influence on the decomposition of crop residues.

I found that the ICBM-predicted C stocks were within the ranges observed for soil pedon data reported by the National Soil Information System (NASIS, USDA), based on soil samples taken from different parts of the country. The estimated soil C stocks followed both the variation in cropland area and the observed changes in weather over the 16-year period. Since total cropland C stocks showed drastic variation due to changes in crop areas over time, I considered only the C stocks and stock changes in the permanent cropland areas (i.e. the minimum crop area observed for a county over the 16-year period was considered as the permanent cropland), in the analyses. There was an overall increase in C stocks in the different CPRs over the 16-year period. According to the model results, the total C stocks in the irrigated and rain fed cropland at national scale were 45 Mg ha^{-1} and 36 Mg ha^{-1}, respectively, and the irrigated cropland was only about 10% of the total cropland over the study period.

The largest C stocks were observed in the cropland areas in the Far West and North Central regions. C stocks in the US permanent croplands increased by 14 Tg over the 16 year period (on the average 0.9 Tg C yr^{-1}), which is slightly lower than the estimate by Ogle et al. (2003) for the same period (1982-1997). The interannual variability was still significant with gains and losses in C stocks due to the observed interannual weather variation. The trend clearly showed the impact from climate events such as El Nino. Overall net gains of C stocks were high after 1993, and significant increases (over 10 Tg C yr^{-1}) were found especially in the years following those that had high productivity. Significant losses could be found in the years 1983, 1984, 1987, 1989, and 1994, which were preceded by years that had lower than average productivity due to extremes in weather. The estimated annual 'apparent' cropland NEP (i.e. the balance of NPP left after reducing the CO_2 released in microbial respiration and any grain exports from the US) ranged from 14 to 50 Tg, during the period 1994-1997. Therefore further improvement of cropland productivity and increased use of land under management practices such as no till (which is currently only about 20% of the US cropland (CTIC, 2002)) would further reduce net CO_2 emissions from the US croplands, resulting in increases in C stocks and NEP, which would have significant implications on the overall C cycle and CO_2-C mitigation by the country.

REFERENCES

Andrén, O., and T. Kätterer. 1997. ICBM- the Introductory Carbon Balance Model for exploration of soil carbon balances. Ecol. Applications 7(4),1226-1236.

Andrén, O., and T. Kätterer. 2001. Basic principles of soil carbon sequestration and calculating dynamic country-level balances including future scenarios. p 495-511. *In* R. Lal, J.M. Kimble, R.M. Follett, and B.A. Stewart. (ed.) Assessment Methods for Soil Carbon. Lewis Publishers, Boca Raton, FL.

Andrén, O., T. Kätterer, and T. Karlsson. 2004. ICBM regional model for estimations of dynamics of agricultural soil carbon pools. Nutr. Cycl. Agroecosys. 70:231-239.

Angers, D. A., and M. R. Carter. 1996. Aggregation and organic matter storage in cool, humid agricultural soils. pp. 193-211. IN M. R. Carter and B. A. Stewart (eds.). Structure and Organic Matter Storage in Agricultural Soils. CRC Press, Inc., Boca Raton, Florida.

Burke, I.C., E.T. Elliott, and C.V. Cole 1995. Influence of macroclimate, landscape position, and management on soil organic matter in agroecosystems. Ecol. Applications 5(1), 124- 131

Buyanovsky, G.A., and G.H. Wagner. 1986. Post-harvest residue input to cropland. Plant Soil 93, 57- 65

Campbell, C.A., and R. de Jong. 2001. Root-to-straw ratios- influence of moisture and rate of N fertilizer. Can. J. of Soil Sci. 81(1), 39- 43.

CAST, 2004. Climate change and greenhouse gas mitigation: challenges and opportunities for agriculture. Council for Agricultural Science and Technology (CAST). Ames, IA.

Cole, V., C. Cerri, K. Minami, A. Mosier, N. Rosenberg, D. Sauerbeck, J. Dumanski, J. Duxbury, J. Freney, R. Gupta, O. Heinemeyer, T. Kolchugina, J. Lee, K. Paustian, D. Powlson, N. Sampson, H. Tiessen, M. Van Noordwijk and Q. Zhao. 1996. Agricultural options for mitigation of GHGemissions. p 745-771. *In* R.T. Watson, M.C. Zinyowera, R.H. Moss, and D.J. Dokken, (Ed.) Climate Change 1995. Impacts, adaptations and mitigation of climate change: Scientific-technical analyses. IPCC Working Group II. Cambridge University Press, Cambridge.

Conservation Technology Information Center (CTIC). 2002. National Crop Residue Management Survey News Release. [Online]. Available at http://www.ctic.purdue.edu/Core4/CT/ctsurvey/2002/2002NewsRelease.html. CTIC, West Lafayette, IN.

Doraiswamy, P., S. Hollinger, T. Sinclair, A. Stern, and B. Akhmedov. 2001. Evaluation of Modis Products for Applications in Regional Crop Condition and Yield Assessment. Remote Sensing For Agriculture Ecosystems And Hydrology

Doraiswamy, P.C., S. Moulin, P. W. Cook, and A. Stern. 2003. Crop yield assessment from remote sensing. Photogramm. Eng. Rem. S. 69(6), 665- 674

Doraiswamy, P.C., J.L. Hartfield, T.J. Jackson, B. Akhmedov, J. Prueger, and A. Stern. 2004. Crop condition and yield simulations using Landsat and MODIS. Remote. Sens. Environ. 92, 548-559

Doraiswamy, P. C., T. R. Sinclair, S. Hollinger, B. Akhmedov, A. Stem, and J. Prueger. (2005). Application of MODIS derived parameters for regional crop yield assessment. Remote. Sens. Environ. 92, 192-202

Eve, M.D., M. Sperow, K. Paustian, and R.F. Follett. 2001. National-scale estimation of changes in soil C stocks on agricultural lands, Environ. Pollut. 116, 431- 438

Hansen, P.M., and J.K. Schjoerring. 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crop using normalized difference vegetation indices and partial least squares regression. Remote. Sens. Environ. 86, 542-553.

Hay, R.K.M. 1995. Harvest index: a review of its use in plant breeding and crop physiology. Ann. Appl. Biol. (1995), 126, 197-216.

Hill, M.J., and G.E. Donald. 2003. Estimating spatio-temporal patterns of agricultural productivity in fragmented landscapes using AVHRR NDVI time series. Remote. Sens. Environ. 84, 367- 384

IPCC 1997a. Revised 1996 IPCC guidelines for national GHGinventories reporting instructions (Volume1). Intergovernmental Panel on Climate Change

IPCC 1997b. Revised 1996 IPCC guidelines for national GHGinventories workbook (Volume 2). Intergovernmental Panel on Climate Change

IPCC 1997c. Revised 1996 IPCC guidelines for national GHGinventories reference manual (Volume 3). Intergovernmental Panel on Climate Change

Intergovernmental Panel on Climate Change (IPCC). 2000. J. Penman, D. Kruger, I. Galbally, T. Hiraishi, B. Nyenzi, S. Emmanuel, L. Buendia, R. Hoppaus, T. Martinsen, J. Meijer, K. Miwa, and K. Tanabe (ed.) Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories. IPCC/IGES, Hayama, Japan.

IPCC 2001a. R.T. Watson et al. (eds) Climate Change 2001: Synthesis report. Cambridge University Press.

IPCC 2001b. Good practice guidance and uncertainty management in national GHG inventories. Intergovernmental Panel on Climate Change

Jenkinson, D.S. 1971. Studies on the decomposition of ^{14}C-labelled organic matter in soil. Soil Science III, 64-70

Lal, R., J. Kimble, R.F. Follet, and C. Cole 1998. The potential for U.S. cropland to sequester carbon and mitigate greenhouse effect. Ann Arbor Press, Chelsea, Michigon.

Lobell, D.B., J.A. Hicke, G.P, Asner, C.B. Field, C.J. Tucker, C.J., and O. Los. 2002. Satellite estimates of productivity and light use efficiency in United States agriculture, 1982-98. Glob. Change Biol. 8, 722-735

Lokupitiya, E., and K. Paustian. 2005. Agricultural soil GHGemissions: A review of national inventory methods. J. Environ. Qual.. *In Press.*

Lokupitiya, R.S., E. Lokupitiya, E., and K. Paustian. 2006. Comparison of missing value imputation methods for crop yield data. Environmetrics 17(4), 339- 349.

Ogle, S.M., J. Breidt, M.D. Eve, and K. Paustian. 2003. Uncertainty in estimating land use and management impacts on soil organic carbon storage for US agricultural lands between 1982 and 1997. Glob. Change Biol. 9:1521-1542.

Paustian, K., E.T. Elliott, H.P. Collins, C.V. Cole, and E.A. Paul. 1995. Use of a network of long-term experiments for analysis of soil carbon dynamics and global change: the North American model. Aust. J. Exp. Agr. 35: 929-939

Paustian, K., H.P. Collins, and E.A. Paul. 1997a. Management controls on soil C. p. 15- 49. *In* E.A. Paul, E.T. Elliott, K. Paustian, and C.V. Cole (ed.) Soil Organic Matter in Temperate Agroecosystems. CRC Press, New York.

Paustian, K., J. Brenner, K. Killian, J. Cipra, S. Williams,S., E.T. Elliott, M.D. Eve, T. Kautza, and G. Bluhm. 2002. State-level analyses of C sequestration in agricultural soils. p 193- 203. *In* R. Lal, J.M. Kimble, and R.F. Follett (ed.) Agricultural practices and policies for C sequestration in soil, New York, Lewis Publishers.

Sinclair, T.R. 1998. Historical changes in harvest index and crop nitrogen accumulation. Crop Science 3(38), 638- 643.

Sperow, M., M.D. Eve, and K. Paustian. 2003. Potential soil C sequestration on U.S. agricultural soils. Climatic Change 57:319-339.

CHAPTER 2

AGRICULTURAL SOIL GREENHOUSE GAS EMISSIONS: A REVIEW OF NATIONAL INVENTORY METHODS

ABSTRACT

Parties to the UNFCCC are required to submit national GHG inventories, together with information on methods used in estimating their emissions. Currently agricultural activities contribute a significant portion (ca. 20 %) of global anthropogenic GHG emissions, and agricultural soils have been identified as one of the main GHG source categories within the Agricultural Sector. However, compared to many other GHG sources, inventory methods for soils are relatively more complex and have been implemented only to varying degrees among member countries.

This review summarizes and evaluates the methods used by Annex 1 countries in estimating CO_2 and N_2O emissions in agricultural soils. While most countries utilize the IPCC default methodology, several Annex1 countries are developing more advanced methods that are tailored for specific country circumstances. Based on the latest national inventory reporting, about 56% of the Annex 1 countries use IPCC Tier 1 method, about 26% use Tier 2 methods, and about 18% do not estimate or report N_2O emissions from agricultural soils. More than 65% of the countries do not report CO_2 emissions from the

cultivation of mineral soils, organic soils or liming, and only a handful of countries have used country specific, Tier 3 methods. Tier 3 methods usually involve process-based models and detailed, geographically specific activity data. Such methods can provide more robust, accurate estimates of emissions and removals but require greater diligence in documentation, transparency and uncertainty assessment to ensure comparability between countries. Availability of detailed, spatially explicit activity data is a major constraint to implementing higher tiered methods in many countries.

INTRODUCTION

Efforts to mitigate increasing greenhouse gas (GHG) concentrations through international cooperation have accelerated since the beginning of the 1990's, with the adoption of the United Nations Framework Convention on Climate Change (UNFCCC) in 1992 and the Kyoto Protocol in 1997. The UNFCCC entered into force in 1994 and so far over 180 countries have ratified it. The prime objective of the UNFCCC is to stabilize atmospheric GHG emissions at a level that would not further damage the climate system from human interference.

Under the UNFCCC, member countries are expected to submit national GHG inventories. Estimating the sources and sinks of GHG emissions at a national level is needed to quantify the sources and sinks from individual countries and to assess compliance with international agreements to reduce emissions. Accurate inventories also improve understanding of the relative importance of different sinks and sources and their spatial distribution. Developed countries that are listed in Annex I of the UNFCCC (referred to as ‘Annex I countries’) are expected to submit an annual national inventory report (NIR) containing detailed and complete information on their emissions for all years from a base year (which is 1990 for the majority of the countries) to the current year for the annual inventory submission (UNFCCC, 1999).

Annex 1 countries have submitted national GHG inventories since 1994, and since 1996 they have made annual submissions. The quality, information content, and transparency differ substantially among the reporting countries (Acosta et al., 2002) and only some countries have published detailed method descriptions. According to Acosta et al. (2002), methodological problems associated with GHG inventories include the use of outdated emission factors, insufficiently robust assumptions, calculation errors, and use of methods, parameters, and emission factors that deviate substantially from IPCC default method and values.

Inventory methods for land use and agriculture related activities, in general, and for soils in particular, are arguably among the more complex and least developed inventory sectors, yet soils are a significant sink/source category for many countries. This paper reviews the methodologies used by member parties to the UNFCCC to prepare national GHG inventories of soil emissions and sinks. Our objective is to assess the current state, recent developments, and future trends in soil emission inventory methods. Since many non-Annex 1 countries have not yet reported soil emission inventories, our focus is on methodologies used by Annex 1 countries.

IMPORTANCE OF AGRICULTURAL SOILS AS A SOURCE AND SINK OF GHG EMISSIONS

Currently, agricultural activities contribute a significant portion (ca. 20%) of global GHG emissions (Figure 1), making it imperative for many countries to inventory GHG emissions from agricultural soils. Nitrous oxide (N_2O) is the predominant GHG added to the atmosphere from agricultural soils, in terms of global warming potential. N_2O is produced through both nitrification and denitrification, which are microbial-driven soil processes that naturally occur in soils; these processes are stimulated by increased N inputs. Currently about 60% of total N_2O emissions in Annex 1 countries are derived from agriculture (Figure 2). Anthropogenic increases in N_2O emissions in agricultural soils occur mainly as a result of fertilizer and manure use. Total N_2O emissions from all Annex 1 countries (Table 1), as well as for the US alone, remained fairly constant during the 1990s (Figure 3).

Soils are the major land surface carbon (C) reservoir, containing roughly 1500 Pg C, about three times the amount of C in terrestrial vegetation (Schlesinger, 1997). Agricultural soils contain around 170 Pg of C in the upper 1m (Cole et al., 1996). Emissions of CO_2 from soil are significant, particularly for land use change, often involving the conversion of native ecosystems to agricultural uses (Figure 1). Agricultural soils can also contribute to the removal and storage of atmospheric carbon dioxide (CO_2) into soil organic matter. C storage (or C sequestration) in agricultural soils can be increased by management practices such as no-till, organic amendments, conservation reserves, and improved crop rotation and fertilization practices (Paustian et al., 1997, 2000 ; Smith, 2004; Ogle et al., 2005). The potential C sequestration in global agricultural soils through changes in such management practices has been estimated at 600- 900 Tg yr^{-1} (Cole et al., 1996). According to IPCC (2000a), potential C stock changes due to land use and management improvements could be 125 Tg yr^{-1} on annual croplands and as much as 800 Tg on all agricultural land (including grazing lands and agroforestry), by 2010, with concerted efforts to adopt best management practices.

Although methane (CH_4) emissions from agricultural soils, specifically in paddy rice, are an important GHG source globally, CH_4 from rice cultivation is of minor importance for Annex 1 countries (Figure 4). Hence, only methods for estimation of CO_2 emissions and removals and N_2O emissions from agricultural soils are dealt with in this paper.

METHODOLOGIES FOR ESTIMATING GHG EMISSIONS AND REMOVALS IN AGRICULTURAL SOILS

According to Articles 4 and 12 of the UNFCCC, national inventories need to use comparable methodologies agreed upon by the Conference of Parties (COP), the UNFCCC's principal policy-making body. Currently, parties to the UNFCCC use the 1996 Revised IPCC Guidelines (IPCC 1997 a, b, c), as elaborated by the IPCC Good Practice Guidance (IPCC 2000b, 2003), and the UNFCCC Reporting Guidelines on Annual Inventories (FCCC/CP/1999/7). The IPCC Guidelines provide general guidance, default calculation methods, and reporting formats for national inventories in order to promote transparency and comparability among countries. The IPCC Guidelines (referred to as 'IPCC-GL') and Good Practice Guidance (IPCC-GPG) provide a default methodology (Tier 1) for estimating GHG emissions as well as guidance in developing and using country-specific data and methods (referred to as Tier 2 and Tier 3 methods – see below). Some Annex 1 countries have begun to develop and implement these country-specific 'higher-tiered' methodologies. The methodologies currently used by different member countries to the UNFCCC in estimation of CO_2 and N_2O are summarized in Table 2. A description of the 1996 IPCC default methodology and methods developed by some Annex 1 countries for agricultural soils follow.

IPCC inventory methodology

As described in the IPCC-GPG, methodologies for agricultural soils could fall under three main tiers, with increasing complexity and data demands occurring at the higher tiers: 1) Tier 1 consists of simple equations and default emission factors provided in the IPCC-GL (and IPCC-GPG); 2) Tier 2 uses the IPCC-GL default equations but requires country-specific parameters that better account for local climate, soil, management and other conditions; and 3) Tier 3 methods are based on more complex models and inventory systems, typically using more disaggregated activity data that better capture variability in local conditions (IPCC 2000b, 2003). Member countries are encouraged to provide uncertainty estimates as well as mean values (IPCC, 1997a, b, c).

N_2O emissions from agricultural soils

The IPCC-GL and IPCC-GPG provides methodology for estimating: 1) direct N_2O emissions from soils; 2) indirect emissions of N_2O (i.e. due to losses from N volatilization and leaching that are subsequently deposited in non-agricultural environments and are subject to loss as N_2O); and 3) direct soil emissions of N_2O from animal production (during waste storage for confined livestock and from livestock grazing). Direct N_2O emissions from soil are estimated using an equation (Eq. [1]) that incorporates emissions from all major N inputs and does not discriminate among different N sources (IPCC 1997b, c).

$$N_2O_{\text{ DIRECT}} (\text{kg N yr}^{-1}) = [F_{SN}+F_{AW}+F_{CR}+F_{BN}]*EF_1 + F_{OS}*EF_2 + F_{PRP}*EF_3 \qquad [1]$$

Where:

F_{SN} = N input from synthetic fertilizer use (kg N yr^{-1})

F_{AW} = N from livestock manure applied to soil (kg N yr^{-1})

F_{BN} = total N input in N-fixing crops (kg N yr^{-1})

F_{CR} = N input from crop residues (kg N yr^{-1})

F_{PRP} = N input from animal excretion on pasture/range/paddock (kg N yr^{-1})

F_{OS} = area of cultivated organic soils (ha)

EF_1= emission factor for direct N application (kg N_2O-N kg^{-1} N added)

EF_2 = emission factor for cultivated organic (e.g. peat) soils (kg N_2O-N ha^{-1} yr^{-1})

EF_3= emission factor for excretion on pasture/range/paddock (kg N_2O-N kg^{-1} N added)

Indirect soil N_2O emissions are defined as deriving from the volatilization (as NH_3 and NO_X) and subsequent deposition of previously applied N, as well as emissions from applied N that has been transported to riparian or aquatic environments through leaching and runoff. Thus the actual emissions may be occurring in non-agricultural environments, but the original source of the applied N was from agricultural soils. Other N_2O emissions from animal waste management systems are included in the agricultural sector but are not related to soil management.

According to the latest National Inventory Report (NIR) submissions, a majority of reporting countries use the IPCC default method (about 70% of the countries) and default emission factors (about 75% of the countries) in estimating N_2O emissions from agricultural soils.

Emissions/ removals of CO_2 from agricultural soils

Depending on the management practices being used, and their relative effect on C inputs from residues vs. C losses from decomposition, agricultural soils can be either a net source or a net sink for C (Paustian et al., 1997, 2000; Lal, 2004; Smith, 2004). The IPCC methodology estimates net CO_2 emissions (sinks and sources) from: 1) changes in C stocks of mineral soils due to changes in land-use practices; 2) CO_2 emissions from organic soils converted to agriculture or plantation forestry; and 3) liming of agricultural soils (IPCC, 1997b, c, 2003).

For CO_2 emission from mineral soils, the net change in soil organic C is estimated for lands under different categories of land use and management, stratified by climate and soil type, over a specified time period; the default time period is 20 years. Estimates of soil C stocks are for the top 30 cm of the soil profile, where the impacts of changes in land use and management are greatest and where most field measurements have been reported (Ogle et al., 2003a). Hence, changes in soil C stocks that may occur deeper in the profile are not captured by the method.

The default methodology uses a set of coefficients (stock change factors) based on soil type, climate, disturbance history, productivity, and management practices. Climate is divided into nine categories based on average annual temperature and precipitation. Soils are defined by taxonomic characteristics based on broadly defined soil properties, including texture, clay mineralogy, morphology, and drainage, that influence the ability of a soil to store organic matter. Default values for reference C stocks and stock change factors are stratified according to climate and soil type. Reference C stocks represent values found under native, unmanaged ecosystems.

The basic method combines the reference C stock, stock change factors, and activity data for land use and management changes over time. Mineral soil C stock change is estimated as shown in Eq. [2] (IPCC 1997c, 2003).

$$\Delta SC = [(SC_0 - SC_{(0-T)})]/D \qquad [2]$$

Where,

$SC = \Sigma_c \Sigma_s \Sigma_i SC_R * F_{LU} * F_{MG} * F_I * A$

Where:

ΔSC = annual soil carbon stock change (Mg C yr^{-1})

SC_0 = soil organic carbon stock in the inventory period end year, for current land use and management (Mg C ha^{-1})

$SC_{(0-T)}$ = soil organic carbon stock T years prior to the end year of the inventory period (Mg C ha^{-1})

c represents climate zone, *s* soil type and *i* is the set of land management systems defined for the country.

A= land area of each parcel (ha)

SC_R = the reference carbon stock (Mg C ha^{-1})

F_{LU} = stock change factor for land use type (dimensionless)

F_{MG} = stock change factor for management/disturbance regime (dimensionless)

F_I = stock change factor for carbon input level (dimensionless)

D= time period for transition between equilibrium C stocks, as represented by stock change factors (default is 20 years).

While the formulation is designed to be generic for all soils, the interpretation and values of these factors vary according to the type of ecosystem (i.e. cropland, grassland, forest) and the changes in land use and management that are being represented. For example, the land use factor provides a baseline level for C stocks in permanent cropland, short- ($\leq$20 yr) and long- term (>20 yr) cropland set-aside (to perennial grasses or trees), shifting cultivation, and managed grassland and forest, relative to stock levels in native unmanaged ecosystems (where F_{LU}=1). For cropland, the stock change factor for management regime (F_{MG}) specifies relative C stock values for different tillage regimes. For cropland and managed grassland (pasture, hayland), the input factor (F_I) relates to management practices that affect the relative amount of C returned to the soil (as plant-derived residues or exogenous additions like animal manure) for a particular land use type. Hence for cropland, it is dependent on the type of crops grown, whether residues are removed or retained, and whether manures are added. For grassland, F_I depends on management practices that influence primary productivity, such as fertilization, species improvement, and grazing regime. More detailed definitions and default values for stock change factors are given in the Good Practice Guidance for LULUCF (IPCC, 2003).

The default inventory period, to which land use and management activity are applied, is 20 years, i.e., default values for relative stock change factors are estimated for a 20-year period (IPCC, 2003). However, the method can be applied for an inventory period of 20 years or less, using the default factors. In calculating inventories for year '*t*', land areas are stratified by climate, soil type, and "initial" land use and management for year *t*-T (where T is the length of the inventory period), and then land use and management conditions are specified for the same areas in year *t*. Changes in C stocks from one land use/management system to another are assumed to be linear over time. Implicit in the method is that if there are no changes in land use or management (i.e. F_{LU}, F_{MG} and F_I are unchanged) for a given land area over the inventory period, then soil C stocks remain constant and there is no net emission or removal of CO_2.

CO_2 emissions from cultivated organic soils and from liming of soils are handled differently from C stock changes in mineral soils. Both are considered as only sources of CO_2 and are estimated using simple emission factors (i.e. annual loss per unit area) multiplied by the areas of each activity. Net C loss from organic soils is calculated using the land area and annual loss rates that vary by broad climate divisions and land-use. The default values for annual loss rates of C given in this method are derived from a global survey on published literature. Emissions from liming are estimated assuming that the carbonate-C added (i.e. limestone or dolomite) is emitted as CO_2 in the year of application.

CORINAIR methodology

Countries of the European Union (EU) prepare their national inventories according to the EU emission inventory program known as CORINAIR (CORe INventory of AIR emissions in Europe). It was initiated in 1985 to assist in the development of consistent, comparable, and transparent national inventories for “conventional” air pollutants such as SO_x, NO_x, and VOC. This system has evolved over time and guidelines for preparing atmospheric emission inventories by the EU member countries are included in the EMEP (i.e. Co-operative Programme for Monitoring and Evaluation of the Long-Range Transmission of Air Pollutants in Europe) / CORINAIR Atmospheric Emission Inventory Guidebook, first published in 1996.   The source categories covered under CORINAIR 1990 include eleven main source sectors, including agriculture.

CORINAIR methodology involves more disaggregated source categories and spatial detail, and CORINAIR-based estimates can be transformed appropriately for different reporting purposes, including the IPCC format. EMEP/CORINAIR (2002, 2004), has classified GHG emissions from agricultural soils (in accordance with Common Reporting Format (CRF)/ IPCC classification) under the categories: 1) *Cultures with fertilizers* (where emissions of NH_3, N_2O and NO_x, CO_2, CH_4, and non- CH_4 volatile organic compounds (NMVOCs) are estimated); and 2) *Cultures without fertilizers* (where emissions of NH_3, N_2O, NO_x and VOCs are estimated). The methodology suggested for estimating N_2O emissions follows the IPCC default methodology; IPCC Direct and Indirect N_2O subsource categories in agricultural soils have been reported as CORINAIR sub-sectors for Cultures with/without fertilizers (EMEP/CORINAIR, 2004).

In addition to the simpler methodology based on the IPCC default method, an improved methodology for N_2O emissions is given; methods have been developed based on multivariate regression analyses that incorporate important factors such as climate, weather, and soil conditions that control N_2O emissions. Mechanistic simulation models such as DNDC (Li, 2000) have also been applied at regional scale to inventory N- trace gas emissions (EMEP/ CORINAIR,2004).

No alternative methodology from the IPCC has been suggested for estimating CO_2 emissions and removals under *Cultures with fertilizers*. IPCC land use changes considered under *Cultures with fertilizers* include: a) conversion of woodland to grassland and cropland, b) conversion of grassland to cropland and *vice versa*, and c) other land use change activities that include drainage of wetlands and cultivation of organic soils. Although this latter category of land use changes is a significant source of CO_2 in certain countries (especially in northern Europe), such CO_2 emissions are rarely reported in national inventories, as no default methodology has been provided under the CORINAIR methodology. However, higher emission factors compared to IPCC for CO_2 released from cultivated organic soils have been suggested based on recent measurements made in Europe (EMEP/CORINAIR, 2004).

National GHG accounting systems developed by certain Annex 1 countries

Currently the IPCC default methodology dominates among the methods used by Annex 1 countries in estimating national GHG emissions in agricultural soils. In recognition of the limitations inherent in using global and regional default values, the IPCC-GL and IPCC-GPG encourage countries to develop methodologies more appropriate for their national circumstances. However, development of such methods requires considerable time and resources, including testing and validation prior to implementation. Consequently, relatively few country-specific systems have been fully implemented to date. In this section, country-specific methods developed by certain Annex1 countries to estimate agricultural soil CO_2 and/ or N_2O emissions and removals are discussed. In each of these countries, emissions and/or removals from soils are a major component of the total impact of GHG in the agricultural sector (Fig. 4). Some of these methods are still in the development process and some are not fully utilized for agricultural soils at the national level and have been used and validated mostly at regional or project-level. An overview of the methods used by all Annex 1 countries in estimating CO_2 and N_2O emissions from agricultural soils is given in Table 2.

Australia

Australia has about 21% forest area and two-thirds of the land is in agricultural and pastoral use, of which about 90% is used for grazing livestock (Australian Greenhouse Office, 2002a). Currently about 19% of the total CO_2-equivalent GHG emissions in Australia comes from the agriculture sector and about 5% come from other land use. Emissions from agricultural soils in 2002 show an increase of 29% compared to 1990, mainly due to an increase in area under cultivation, increased rates of fertilizer application, and increased number of livestock (UNFCCC/NIS, 2004).

Australia has developed a National C Accounting System (NCAS) based on resource inventories, field studies, modeling, and remote sensing. NCAS involves a verified model-based accounting system operating at highly disaggregated spatial and temporal scales (25m, monthly time steps; Australian Greenhouse Office, 2002b). Several submodels comprise the Full C Accounting Model (FullCAM) for estimating land use change emissions. FullCAM has components that incorporate C exchange between the atmosphere and agriculture- and forestry-related activities. Emissions and removals are estimated for biomass as well as soil C pools.

The five sub-models of the FullCAM include a physiological growth model for forests (CAMFor), a C accounting model for cropping and grazing system (CAMAg), a residue decomposition model (GENDEC), and the Rothamstead soil C model (Roth-C) (Brack and Richards, 2001; Australian Greenhouse Office, 2002b). The FullCAM model provides a linkage between these sub-models.

CAMAg reflects the management impacts on C accumulation and allocates crop biomass to various plant product pools and to decomposable and resistant organic residues. Change in agricultural soil C is estimated using Roth-C model (Coleman and Jenkinson, 1995), based on soil type, land use and management history, and residue inputs from different cropping systems. This model has been calibrated against long-term field measurements and verified using paired sites (undisturbed vs cleared sites) in areas of major land use change; although further measurements from long-term field experiments are needed to refine the model (G. Richards, personal communication, 2004). Since CAMFor and CAMAg are both included within the FullCAM, transitional activities such as deforestation, afforestation, reforestation, and mixed systems such as agroforestry can be represented (Australian Greenhouse Office, 2002b). In addition, a mathematical framework has been developed to incorporate the ability to estimate non-CO_2 GHGs within the FullCAM model. This will enable the estimation of non-CO_2 emissions from forestry (i.e. CH_4 from decomposition and burning, N from decomposition, burning, fertilization, and soil preparation) and agriculture (i.e. N from fertilizer application, animal excrement, soil management, decomposition, and burning; CH_4 from decomposition and burning) (Australian Greenhouse Office, 2002b).

Austria

In Austria, agricultural area is about 41%, and forests occupy about 46% of the total land area. Agricultural area includes arable land, grassland, as well as vineyards and orchards. Natural reforestation in former agricultural land, and aforestation activities have contributed to the further extention of the country's forest area within the last few decades (UNFCCC/NIS, 2004).

The "Australian Carbon Balance Model" (ACBM) is being developed for full C accounting of C stocks in Austria, considering 1990 as the baseline. The overall model covers five national subsystems; agriculture, forestry, energy, production, and waste. For each module, C stocks, flows, processes and control variables have been identified. The model is formulated to estimate the full national C balance, including intersystem C flows to show the impact of one action on the other components of the model, and subsequent net flux to the atmosphere (Orthofer et al., 2000; Gerzabeck et al., 2003).

Agricultural soil is one main component within the module for agriculture. Soil C dynamics are modeled using a simple approach that divides organic matter into three pools based on residence times. Factors for simulation and calculation of C stocks are derived using official soil survey data (Orthofer et al., 2000). Net emissions (expressed in CO_2 equivalents) from agricultural soil estimated using the ACBM were 13% lower than emission estimates made using the IPCC-GL. This difference was mainly due to the effect of C sequestration in agricultural products and forest soils and of the net effect from C import and export (Gerzabeck et al., 2003).

Soil C estimates from ACBM have not been so far reported in the National inventory submission, although C stock changes in Austrian forest soils during 1990- 2010 were estimated using this modeling approach (Weiss and Schlamadinger, 2000).

Canada

In Canada about 42% of the country is covered by the forest, and about 7% of the land is under agriculture. About two-thirds of the farmland is used for crops and improved pasture (Environment Canada, 2001).

At present, Canada is using the IPCC Tier 1 method to estimate agricultural soil N_2O emissions, and CO_2 emissions and removals from agricultural soils are estimated using the Century model that has been calibrated for Canadian conditions (Smith et al., 1997; C. Liang, personal communication, 2004). However, a new National C and Greenhouse-gas emission Accounting and Verification System for agriculture (NCGAVS) is being developed to estimate soil C change and direct N_2O emissions from agricultural soils (McConkey et al., 2003). It is a model-based system that uses integrated databases of information on land, land management, and land use change.

NCGAVS will provide a more detailed country-specific methodology for estimating N_2O and CO_2 emissions/removals from Canadian agricultural soils. The basic geographic units used in NCGAVS are the Soil Landscapes of Canada (SLC) polygons, which are mapped at a scale of 1:1,000,000, yielding around 5000 polygons for the entire country. Soil C changes and N_2O emissions will be estimated for the components within a SLC, so that important attributes such as soil organic matter, texture, and topographic effects, etc. will be addressed. Inclusion of topographic effects is an improvement compared to the existing inventory methodology.

The initial version of NCGAVS will concentrate on estimating sources/sinks of soil C and direct N_2O emissions from land-use change, crop- and grazing-land management, and revegetation under primary agriculture over a 5-year inventory period. Since CH_4 emissions from agricultural soils are not a significant source in Canada, it will not be included initially. It is expected that NCGAVS will be used for the inventory submission in 2006 (C. Liang, personal communication, 2005). Future versions are planned to have more extensive coverage, accounting for all GHG sources and sinks in Canadian agriculture. Since N_2O represents the largest GHG source for Canadian soils, further efforts to quantify N_2O emissions will be a major focus (McConkey et al., 2003).

Germany

In Germany, about 54% of land was under agriculture in 1997, with two-thirds of the area used for annual crops. About 3% of agricultural land is organically farmed. Fertilizer use and livestock are the major sources of agricultural sector GHG emissions of N_2O and CH_4 (BMU, 2002).

Germany has developed a system using two mechanistic models, Denitrification and Decomposition (DNDC) (Li et al., 2000) and Photosynthesis and Evapotranspiration-Nitrification- Denitrification and Decomposition (PnET-N-DNDC; based on the PnET model by Aber et al., 1996) for estimating N_2O emissions from agricultural and forest soils, respectively (Butterbach-Bahl et al., 2001, 2002, 2004). These process-based models integrate complex interactions among primary drivers, soil environmental factors, and biogeochemical reactions. Input parameters include daily climate data, soil properties (organic matter, texture, pH, bulk density), vegetation (crop/ forest type, age), and land management practices.

This approach has been tested regionally, and so far modeled N_2O emissions have been comparable with the estimates derived from using IPCC guidelines. DNDC-based estimates at the regional scale were slightly higher (about 10%) compared to the estimates based on IPCC default method (Butterbach-Bahl et al., 2002). Model validation is difficult however, due to the scarcity of field measurements over entire years or entire regions, as well as the high spatial variability of N_2O emissions (Butterbach-Bahl, 2004). Although the IPCC methodology has been used in the current and past inventory submissions, a national inventory of N_2O emissions using the simulation modeling approach will be used in future German national inventory submissions (K. Butterbach-Bahl, personal communication, 2004).

Methodology for estimating CO_2 emissions from German agricultural soils is still in development. CO_2 emission from cultivated organic soils is a key category. Soil C stocks for annual and perennial cropland, vineyards, grassland, and fallow land have been estimated using remotely-sensed data and soil data; emission factors have been derived using a review of about 200 national and international studies (A. Gensior, personal communication, 2004).

New Zealand (NZ)

Agriculture is the principal industry in NZ, predominantly as intensive and extensive pastoral systems. Only about 1% of the land area of NZ is devoted to annual cropland. 49% of total GHG emissions in 2002 were from agriculture, and N_2O emissions from agricultural soils accounted for 34% of agricultural sector emissions (NZ MfE, 2001; UNFCCC/NIS, 2004).

Although soil C stock changes are not currently reported (UNFCCC/NIS, 2004), a soil Carbon Monitoring System (CMS) is being developed to account for changes in soil C stocks due to land cover changes occurring in the recent past (i.e. conversion of grazing land to plantation forestry and to native woody vegetation). The soil CMS of NZ is based on a simple empirical model, similar in concept to the IPCC Tier 1 approach. This accounting system uses three data layers (i.e. soil, climate, and land use), for which steady-state soil C stocks are assigned using geo-referenced soil C measurements. The six IPCC recommended soil classes, supplemented by a separate class for podzol soils, are defined and a more detailed breakdown of climatic zones compared to the IPCC approach is used (Scott et al., 2002; Tate et al., 2003). In order to incorporate erosion impacts on soil C, an erosivity index (i.e. the product of slope and mean annual precipitation) has been included with soil type, climate and land use as the major determinants of soil C stocks. Soil C values for different land cover/land use categories have been estimated for each of 18 soil-climate classes (Scott et al., 2002; Tate et al., 2003).

Advantages over the IPCC default method include a NZ-specific representation of soil and climate conditions (e.g. Table 3), measured soil C stocks for climate-soil-land use conditions in NZ, and factoring in effects of erosion, which is high in part of the country. CMS-based estimates for soil C values were slightly higher for high clay-activity soils in cold temperate dry climates and slightly lower for high clay-activity soils in cold temperate moist soils, compared to the default IPCC soil C values for native vegetation in similar soil/climate categories (Table 3). For estimation of soil C changes, periodic update of national land cover/land use data is essential. Currently, uncertainty about changes in areas under different land cover and limited data on the effects of key land-use changes on soil C stocks are the major sources of uncertainties in estimating soil C emissions/removals.

A recent research-based effort (employing methods different from those being implemented for official national reporting), using models, remote sensing, and field data, suggests that soil C stocks in New Zealand are roughly in balance, with some accumulation occurring in grassland and scrub ecosystems and net losses associated with soil erosion (Tate et al., 2000; Trotter et al., 2004).

Sweden

In Sweden, more than two-thirds of the country is covered by forests, and agricultural land comprises 8% of land area, concentrated in the southern part of the country. A significant proportion (9%) of annual cropland is on cultivated peat soils (histosols) (Swedish Ministry of the Environment, 2001; UNFCCC/NIS, 2004).

For estimating soil C budgets, Sweden is developing a simulation-based approach, using the Introductory Carbon Balance Model (ICBM) (Andrén and Kätterer, 1997, 2001). This two-pool model has been calibrated using long-term field data and has been incorporated into a regional framework to enable estimates of soil C emissions/removals for national reporting (Andrén et al., 2004). The model is conceptually simple and, with suitable input data (i.e. annual agricultural statistics, daily weather data, climate region, soil type, and crop type, etc.), it can be run and optimized in a conventional spreadsheet program (Andrén et al., 2003, 2004). This model approach is still in the testing phase, and currently only the emissions from organic soils are reported in the NIR.

United Kingdom (UK)

In the UK, about 47% of the land area is used for intensive crop and pasture production while about 30% comprises less intensively managed grazing systems (Defra, 2001).

Soil C changes are estimated using a matrix of land use change (derived from land surveys), linked to a dynamic empirical model of C gain or loss. The model is conceptually similar to the IPCC approach, with the important difference that changes in C stocks over time are modeled as non-linear, using an exponential function. Soil C changes with time, for a particular land use transition, are estimated as shown in Eq. [3].

$$C_t = C_f - (C_f - C_0)\, e^{-kt} \qquad [3]$$

Where:

C_t = C stock at time t (Mg C ha^{-1})

C_0 = initial C stock (Mg C ha^{-1})

C_f = equilibrium C stock under the new land use (Mg C ha^{-1})

k = specific rate of C change ($year^{-1}$)

t = time period (year).

For example, if the inventory year is 1990 and a base year of 1930 is chosen to represent the initial C_0 stock, then the total soil C loss or gain from 1930-1990 is estimated using Eq. [4].

$$X_{1990} = \sum_{T=1930}^{t=1990} A_T (C_0 - C_f)(1 - e^{-k(1990-T)}) \qquad [4]$$

Where:

A_T= area under the particular land use transition (ha)

Negative values of X_{1990} indicate removals (gains) in C, while positive values represent C loss. Similarly, the calculation can be made over the interval 1930-1989, as shown in Eq. [5].

$$X_{1989} = \sum_{1930}^{1989} A_T (C_0 - C_f)(1 - e^{-k(1989-T)}) \qquad [5]$$

The net change in C in 1990 is the difference between the estimated values from above Eqs. [4] and [5] (i.e. $F_{1990} = X_{1990} - X_{1989}$).

To apply the model, data is required to estimate the change in equilibrium C stocks from the initial to the final land use during a transition. These are calculated for each land use category, as area-weighted averages by major soil types, by countries (i.e. Scotland, England, and Wales) within the UK. Mean changes in equilibrium soil C stocks are calculated as shown in Eq. [6].

$$\overline{C_{ijc}} = \frac{\sum_{s=1}^{6} (C_{sijc} L_{sijc})}{\sum_{s=1}^{6} L_{sijc}} \qquad [6]$$

Where:

i= initial land use; j= new land use; c= country; s= soil group

C_{sijc} = change in equilibrium soil C for a specific land use transition

L_{sijc} = land area by soil type for a specific land use transition.

The rate of C change depends on the type of land use transition. For a transition where C is lost, a 'fast' specific rate constant (i.e. 'k') is applied, and for transitions where C is gained, a 'slow' specific rate is applied; time ranges relevant to complete transitions were selected using a literature search on measured rates of soil C, and expert judgment (UNFCCC/NIS, 2004; R. Milne, personal communication, 2004). Land use change matrices for the periods 1947-1980 and 1984-1990 have been used in applications to date (UNFCCC/NIS, 2004). The C stock changes reported for the NIR (UNFCCC/NIS, 2004) using this methodology, include means and estimates of uncertainty based on a Monte-Carlo approach, computed separately for England, Scotland, and Wales. For Northern Ireland, C stock estimates have been made using an IPCC-based method, as currently no land-use change matrix is available for the country (UNFCCC/NIS, 2004).

The DNDC model has also been used to estimate N_2O emissions from UK agricultural soils. UK-specific, county-level soil characteristics, daily climate, crops, livestock, and farming practices had been used as input data. Model validation had been done using available, but limited, field data. To be consistent with the IPCC approach, emission factors calculated using model estimates, had been used along with the activity data for different source categories to estimate N_2O emissions. DNDC- based N_2O estimates from indirect emissions, and agricultural practices (excluding animal waste during storage), were about 40% lower than the estimates made using the IPCC default method (Brown et al., 2002).

United States (US)

In the US, arable land covers 19% of total land area with an additional 6% in intensive pasture and 21% in rangeland. Although cropland area has remained relatively stable over the last century (Lal et al., 1998), recent trends show a 12% decline in cropland over the past 20 years and other significant changes in land use and management continue to occur (NRI, 2002). Important trends affecting agricultural soils include set-aside of marginal lands in conservation reserves, reductions in tillage intensity, and increases in cropping intensity (Ogle et al., 2003a).

Currently, the US estimates soil C stock changes using a modified version of the IPCC default methodology with US-specific reference C stocks and stock change factors (Ogle et al., 2003a). Activity data are stratified by IPCC-defined climate and soil types. A comprehensive national database, the National Resources Inventory (NRI), is the primary source of land use and management data. NRI records land use, crop type, and other information (e.g. irrigation, pasture improvement, soil type) on more than 400,000 permanent inventory points on agricultural land. Surveys have been conducted on 5-year intervals (1982-1997), although currently the NRI is transitioning to an annual collection of data on a subset of inventory points. Supplemental data including county-level tillage practices (CTIC, 1998), fertilizer use (ERS, 2003), and manure production (Edmonds et al., 2003), are included in the inventory. A Monte-Carlo approach is used to estimate 95% confidence intervals of stock changes for each climate-soil combination (Ogle et al. 2003 a, b).

A more advanced simulation approach using the Century model is being developed to estimate soil C emissions/removals. Annual changes are computed dynamically as a function of inputs of C to soil (e.g., crop residues, manure) and C emissions from organic matter decomposition, which are governed by climate and soil factors as well as management practices. The model simulates all major field crops (maize, wheat and other small grains, soybean, sorghum, cotton) as well as hay and pasture (grass, alfalfa, clover). The same sources of input data as in the IPCC-based methodology are used for management variables, included tillage, fertilization, irrigation, drainage, and manure addition. Preliminary results predict that cropland mineral soils are a net sink of about 21 Tg yr^{-1}, which is higher than the estimates (11 Tg yr^{-1}) using the IPCC approach (US EPA, 2004). Both methods attribute C gains to conservation set-aside and reduced tillage, but the simulation approach also accounts for the long-term trend of increasing crop productivity, which is not captured by the IPCC method.

Soil N_2O emissions were previously estimated using the IPCC Tier 1 methodology, with activity data derived from county-level databases on mineral fertilizer use, animal manure use, crop residues and N-fixing crops, sewage sludge application, and grazing animals (US EPA, 2004). However, the US has developed a simulation-based approach using the DAYCENT model (Parton et al., 1998; DelGrosso et al., 2001) to estimate N_2O emissions from agricultural soils. This approach has advantages over the empirical IPCC method in that it can better capture the interaction between different management conditions, including fertilization and manuring practices, soils, and varying climate. A major challenge is deriving activity data (such as synthetic fertilizer and manure nitrogen inputs) at a suitable spatial scale, since existing fertilizer use databases are aggregated at the country-level. Preliminary estimates using the dynamic method are 10-15% lower than with the IPCC method and with greater interannual variability, due to the inclusion of weather effects in the simulation approach (S. DelGrosso, personal communication, 2005). The largest factor accounting for the difference is lower emissions from N-fixing crops estimated by DAYCENT compared with the IPCC default method.

DISCUSSION AND CONCLUSION

Agricultural soils are a significant source of N_2O emissions for all Annex 1 countries and an important CO_2 source/sink category for many of them. However, compared to some other sectors of country GHG reporting, reporting of soil emissions and removals is, at present, generally less comprehensive and less uniformly applied across countries. Even for Annex1 countries, Tier 1 methods and default emission factors are the most widely used alternatives (Figures 5 and 6).

The relatively large percentage of countries not reporting soil emissions/removals and the predominance of Tier 1 approaches among countries that do report are likely due to two main factors. First, the activity data that are required, even for Tier 1 methods, are more difficult to obtain compared to other statistics such as those on energy consumption. For example, while statistics on country-level mineral fertilizer use are generally available, estimating N addition (and resultant N_2O flux) from sources such as animal manure and sewage sludge are more uncertain. Secondly the methods themselves, particularly for soil CO_2 flux, are arguably more complex and require information from several sources. For estimating CO_2 emissions from mineral soils, the Tier 1 IPCC method requires some level of stratification of land area according to climate and soil type, as well as additional information on land use and management changes over time. An additional reason which could account for the low reporting of soil CO_2 emissions/removals (Figure 6) is that many Annex1 countries may not consider this to be a 'key source' category. However, according to IPCC recommendations, determination of key sources should be based on an initial (e.g. Tier 1) inventory estimate. Hence as more countries strive to adopt Good Practice Guidance, the frequency of reporting for all soil-related categories should increase.

Despite these challenges, a number of countries have successfully implemented soil emission inventories and several are in the process of developing advanced, computational and data intensive methods, tailored to national circumstances (Tier 3).

The IPCC default approach has the advantages of having a relatively simple structure as well as providing default emission and stock change factors, so that the main requirement for individual countries is to obtain suitable activity data. To facilitate this, the methods were designed to work with globally available data sets, at a minimum. However, with simplification there are tradeoffs in the form of increased uncertainty, particularly with the application of global default values. Inherently, globally averaged emission factors will be subject to error when applied to a particular country or region having conditions different than the global mean. In addition, global defaults for soil processes are likely biased in that most of the data used in their derivation are from temperate locations, where the preponderance of field research has been done. Hence, tropical conditions are often underrepresented. Finally, the use of highly aggregated data (e.g. national totals for soil N input sources) results in a loss of information about sub-regional (within country) differences in sources and sinks of greenhouse gases.

For both soil N_2O and CO_2, the process controlling emissions and removals are highly influenced by spatially and temporally varying conditions such as temperature, soil moisture, and soil chemical and physical properties. These differences are unlikely to be adequately captured when national-level aggregate data are used. Additional limitations to the IPCC method for mineral soil C changes include the lack of inclusion of soil erosion and transport and restriction of C stock changes to the top 30 cm of soil. For organic soils, the data available for estimation of emission factors is quite limited and hence uncertainty is higher than for the analogous stock change factors for mineral soils (IPCC, 2003).

Country-specific approaches are being rapidly developed to overcome some of these deficiencies and improve estimates of soil-derived emissions. For soil C estimation, elaboration of the IPCC method using country-specific stock change and reference C stocks, along with other modifications, have been implemented in NZ and the US. The UK approach is functionally similar, but includes a non-linear change with time from an initial to a new equilibrium C stock (whereas the IPCC default assumes a linear transition). Fully dynamic approaches employing simulation models and detailed activity data are being developed in Australia, Austria, Canada, Sweden, and the US.

For N_2O, nearly all countries continue to utilize the IPCC base methodology, although dynamic simulation approaches have been implemented in Germany and the US, and are under development in Australia (Australian Greenhouse Office, 2002b) and Canada.

At present, few countries have estimated inventories using both advanced Tier 3 methods and simpler (Tier 1 or Tier 2) IPCC-default methods. Hence it is difficult to make a general assessment of the adequacy of the more simple IPCC approach compared to more sophisticated inventories. Austria reported a difference of 13% in soil C emission/removals between simple and advanced inventory methods, and N_2O emissions in the US inventory differed by 10-15% comparing IPCC default and simulation model-based approaches. Soil C values in New Zealand estimated using CMS were slightly higher or lower compared to IPCC default C values, for different soil/climate categories. In the UK, DNDC- based N_2O estimates were 40% lower compared to the estimates from

IPCC default method. Although advanced methodologies have been developed for inventorying soil C stock changes in Australia, Canada, Sweden, and UK, so far no study has been done on how those estimates compare with the estimates from the IPCC method.

The higher Tier methods developed by certain Annex 1 countries are designed to be more representative of country-specific soil, climate, and management conditions, and in most cases, these methods are applied at a finer spatial scale than is used for a Tier 1 approach, Hence, the inventory results obtained are expected to be more accurate and with lesser uncertainty than when using the default method. However, at present there have been few instances of rigorous uncertainty assessments applied to either IPCC default (Tier 1) inventories or Tier 3 country-specific methods. Few countries have measurement networks, which can provide independent validation inventory estimates (Ogle and Paustian, 2005). Hence a 'head-to-head' comparison of different inventory methods and Tier levels, with regard to accuracy and uncertainty, is not currently possible. Hopefully as more countries develop and implement more advanced soil GHG emission/removal inventories, supplementary estimates will be made using IPCC methods to provide an evaluation of the reliability of continuing the more general IPCC approaches, which are likely to remain the dominant method for most developing countries, at least for the near future.

Acknowledgements

We are grateful for assistance provided by the inventory specialists and national greenhouse offices from many countries in preparing this review. Support for this work was provided by the Consortium for Agricultural Soil Mitigation of Greenhouse Gases (CASMGS) funded by Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture, under Agreement No. 2001-38700-1109.

REFERENCES

Aber, J.D., P.B. Reich, and M.L. Goulden. 1996. Extrapolating leaf CO_2 exchange to the canopy: a generalized model of forest photosynthesis compared with measurements by eddy correlation. Oecologia 106:257-165.

Acosta, R., S. Pesmajoglou , J. Grabert, and R. Lichte. 2002. Technical review of national greenhouse gas inventories, paper presented at the 11th International Emission Inventory Conference, "Emission inventories- Partnering for the future", Atlanta, Georgia, USA. 15-18 April 2002.

Andrén, O., and T. Kätterer. 1997. ICBM- the Introductory Carbon Balance Model for exploration of soil carbon balances. Ecol. Applications 7(4):1226-1236.

Andrén, O., and T. Kätterer. 2001. Basic principles of soil carbon sequestration and calculating dynamic country-level balances including future scenarios. p 495-511. *In* R. Lal, J.M. Kimble, R.M. Follett, and B.A. Stewart. (ed.) Assessment Methods for Soil Carbon. Lewis Publishers, Boca Raton, FL.

Andrén, O., T. Kätterer, and T. Karlsson. 2003. Carbon balances in Swedish Agricultural soils: Improving IPCC methodology with limited resources. p 127-132. *In* C.A.S. Smith (ed.) Soil Organic Carbon and Agriculture: Developing Indicators for Policy Analyses. Proceedings of and OECD expert meeting, Ottawa, Canada. Agriculture and Agri- Food Canada, Ottawa, and Organization for Economic Co-operation and Development, Paris.

Andrén, O., T. Kätterer, and T. Karlsson. 2004. ICBM regional model for estimations of dynamics of agricultural soil carbon pools. Nutr. Cycl. Agroecosys. 70:231-239.

Australian Greenhouse Office, 2002a. Australia's third national communication on climate change: A report under the United Nations Framework Convention on Climate Change. Australian Greenhouse Office. Canberra.

Australian Greenhouse Office, 2002b. Greenhouse gas emissions from land use change in Australia: An integrated application of the national carbon accounting system,

National Carbon Accounting System Technical Report. Australian Greenhouse Office. Canberra.
BMU. 2002. Third Report by the Government of the Federal Republic of Germany in accordance with the Framework Convention of the United Nations. 2002. Bundesministerium für Umwelt, naturschutz und Reactorsicherheit. Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, Germany.
Brack, C.L. and G. Richards. 2001. FullCAM: An integrated model for predicting carbon stocks at a landscape and national scales. Proc IUFRO Conference Forest Modeling for Ecosystem Management, Forest Certification, and Sustainable Management., Vancouver, British Columbia., Canada12-18 August 2001.
Brown, L., B. Syed, S.C. Jarvis, R.W. Sneath, V.R. Phillips, K.W.T. Goulding, and C. Li. 2002. Development and application of a mechanistic model to estimate emission of nitrous oxide from UK agriculture. Atmos. Environ. 36:917-928.
Butterbach-Bahl, K., F. Stange, and H. Papen. 2001. Regional inventory of nitric oxide and nitrous oxide emissions for forest soils of southeast Germany using the biogeochemical model PnET-N-DNDC. J. Geophys. Res. 106(D24):34155-34166.
Butterbach-Bahl, K., M. Kesik, P. Miehle, and H. Papen. 2002. Calculation of regional budgets of N-trace gas emissions from soils using mechanistic models: Results from a case study for Saxony, Germany. p 119-124. *In* Van Ham, Baede, Guicherit, and Williams-Jacobse (Ed.) Non-CO_2 Greenhouse Gases. Mill press, Rotterdam.
Butterbach-Bahl, K., M. Kesik, P. Miehle, and H. Papen, C. Li. 2004. Quantifying the regional source strength of N-trace gases across agricultural and forest ecosystems with process based models. Plant Soil 260:311-329.
Cole, V., C. Cerri, K. Minami, A. Mosier, N. Rosenberg, D. Sauerbeck, J. Dumanski, J. Duxbury, J. Freney, R. Gupta, O. Heinemeyer, T. Kolchugina, J. Lee, K. Paustian, D. Powlson, N. Sampson, H. Tiessen, M. Van Noordwijk and Q. Zhao. 1996. Agricultural options for mitigation of greenhouse gas emissions. p 745-771. *In* R.T. Watson, M.C. Zinyowera, R.H. Moss, and D.J. Dokken, (Ed.) Climate Change 1995. Impacts, adaptations and mitigation of climate change: Scientific-technical analyses. IPCC Working Group II. Cambridge University Press, Cambridge.
Coleman, K., and D.S. Jenkinson. 1995. RothC-26.3 – a model for the turnover of carbon in soil: Model description and users guide. Lawes Agricultural Trust, Harpenden, UK.
CTIC. 1998. National Crop Residue Management Survey. Conservation Technology Information Center [Online]. Available at www.ctic.purdue.edu/CTIC/CTIC.html (accessed 25 July, 2004; verified 20 April, 2005).
Defra. 2001. The UK's Third National Communication under the United Nations Framework Convention on Climate Change. Department for Environment, Food and Rural Affairs, UK.
Del Grosso, S.J., W.J. Parton, A.R. Mosier, M.D. Hartman, J. Brenner, D.S. Ojima, and D.S. Schimel. 2001. Simulated interaction of carbon dynamics and nitrogen trace gas fluxes using the DAYCENT Model. p 303-330. *In* M.J. Schaffer, L. Ma, S.

Hanson (Ed.) Modeling Carbon and Nitrogen Dynamics for Soil Management. Lewis Publishers, Boca Raton, FL.
Edmonds, L., R. L. Kellogg, B. Kintzer, L. Knight, C. Lander, J. Lemunyon, D. Meyer, D.C. Moffitt, and J. Schaefer. 2003. Costs associated with development and implementation of Comprehensive Nutrient Management Plans. Part I – Nutrient management, land treatment, manure and wastewater handling and storage, and recordkeeping. [Online]. Available at www.nrcs.usda.gov/technical/land/pubs/cnmp1.html (accessed 25 July, 2004; verified 20 April, 2005). Natural Resources Conservation Service, U. S. Department of Agriculture.
EMEP/CORINAIR. 2002. EMEP/CORINAIR Atmospheric Emissions Inventory Guidebook. Technical Report No. 30. European Environment Agency, Copenhagen.
EMEP/CORINAIR. 2004. EMEP/CORINAIR Emissions Inventory Guidebook. Technical Report No. 30. European Environment Agency, Copenhagen.
Environment Canada. 2001. Canada's Third National Report on Climate Change: actions to meet commitments under the United Nations Framework Convention on Climate Change. Government of Canada.
ERS. 2003. Ag Chemicals and Production Technology. Economic Research Service. [Online]. Available at www. ers.usda.gov (accessed 25 July, 2004; verified 20 April, 2005) USDA.
Gerzabek, M.H., F. Strebl, M. Tulipan, M., and S. Schwarz. 2003. Quantification of carbon pools in agriculturally used soils of Austria by use of a soil information system as basis for the Austrian Carbon Balance Model. p 73-78. *In* C.A.S. Smith (ed.) Soil Organic Carbon and Agriculture: Developing Indicators for Policy Analyses. Proceedings of and OECD expert meeting,Ottawa, Canada. Agriculture and Agri- Food Canada, Ottawa, and Organization for Economic Co-operation and Development, Paris.
Intergovernmental Panel on Climate Change (IPCC). 1996. R. Watson, M.C. Zinyowera, R.S. Moss, and D.J. Dokken (ed.) Climate Change 1995. Impacts, Adaptations, and Mitigation of Climate Change: Scientific –Technical Analyses. Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. Cambridge, UK.
Intergovernmental Panel on Climate Change (IPCC). 1997a. J.T. Houghton, L.G. Meira Filho, B. Lim, K. Tréanton, I. Mamaty, Y. Bonduki, D.J. Griggs, and B.A. Callander (ed.) Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories: Reporting Instructions (Volume1). IPCC, Meteorological Office, Bracknell, UK.
Intergovernmental Panel on Climate Change (IPCC). 1997b. J.T. Houghton, L.G. Meira Filho, B. Lim, K. Tréanton, I. Mamaty, Y. Bonduki, D.J. Griggs, and B.A. Callander (ed.) Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories: Workbook (Volume 2). IPCC, Meteorological Office, Bracknell, UK.
Intergovernmental Panel on Climate Change (IPCC). 1997c. J.T. Houghton, L.G. Meira Filho, B. Lim, K. Tréanton, I. Mamaty, Y. Bonduki, D.J. Griggs, and B.A. Callander (ed.) Revised 1996 IPCC Guidelines for National Greenhouse Gas

Inventories: Reference Manual (Volume 3). IPCC, Meteorological Office, Bracknell, UK.

Intergovernmental Panel on Climate Change (IPCC). 2000a. R. Watson, I.R. Noble, B. Bolin, N.H. Ravindranath, D.J. Verardo, and D.J. Dokken (ed.) Land use, Land-use Change, and Forestry: A Special Report. Cambridge University Press. Cambridge, UK.

Intergovernmental Panel on Climate Change (IPCC). 2000b. J. Penman, D. Kruger, I. Galbally, T. Hiraishi, B. Nyenzi, S. Emmanuel, L. Buendia, R. Hoppaus, T. Martinsen, J. Meijer, K. Miwa, and K. Tanabe (ed.) Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories. IPCC/IGES, Hayama, Japan.

Intergovernmental Panel on Climate Change (IPCC). 2003. J. Penman, M. Gytarsky, T. Hiraishi, T. Krug, D. Kruger, R. Pipatti, L. Buendia, K. Miwa, T. Ngara, K. Tanabe, and F. Wagner (ed.), Good Practice Gudinance for Land Use, Land-Use Change and Forestry. IPCC/IGES, Hayama, Japan.

Lal, R., J. Kimble, R.F. Follet, and C.V. Cole. 1998. The potential for U.S. cropland to sequester carbon and mitigate greenhouse effect. Ann Arbor Press, Chelsea, Michigan.

Lal, R. 2004. Agricultural activities and the global carbon cycle. Nutrient Cycling in Agro-ecosystems 70:103-116.

Li, C.S. 2000. Modeling trace gas emissions from agricultural ecosystems. Nutrient Cycling in Agro-ecosystems 58:259-276.

Li C.S., J. Aber, F. Stange, K. Butterbach-Bahl, and H. Papen. 2000. A process-oriented model of N_2O and NO emissions from forest soils: 1. Model development. J. Geophys. Res.-Atmos. 105 (D4):4369-4384.

McConkey, B.G., C.M. Monreal, T. Huffman, G.T. Patterson, J.A. Brierley, R. L. Desjardins, P. Rochette, M.R. Carter, M. Bentham, and S. Gameda. 2003. National Carbon and Greenhouse-gas emission Accounting and Verification System for Agriculture (NCGAVS). P. 79-92. *In* C.A.S. Smith (ed.) Soil Organic Carbon and Agriculture: Developing Indicators for Policy Analyses. Proceedings of and OECD expert meeting,Ottawa, Canada. Agriculture and Agri- Food Canada, Ottawa, and Organization for Economic Co-operation and Development, Paris.

National Resources Inventory (NRI). 2002. [Online]. Available at www.nrcs.usda.gov/technical/land/nri02/nri02lu.html. (acessed 25 July, 2004; verified 20 April, 2005). United States Department of Agriculture – Natural Resource Conservation Service.

NZ MfE. 2001. New Zealand's Third National Communication under the Framework Convention on Climate Change. Ministry for the Environment, Wellington, New Zealand.

Ogle, S.M., J. Breidt, M.D. Eve, and K. Paustian. 2003a. Uncertainty in estimating land use and management impacts on soil organic carbon storage for US agricultural lands between 1982 and 1997. Glob. Change Biol. 9:1521-1542.

Ogle, S.M., J. Breidt, M.D. Eve, and K. Paustian. 2003b. Uncertainty in soil organic carbon storage for US agroecosystems due to land use and management between 1990 and 1997. p. 133-143. *In* C.A.S. Smith (ed.) Soil Organic Carbon and

Agriculture: Developing Indicators for Policy Analyses. Proceedings of OECD expert meeting, Ottawa, Canada. Agriculture and Agri- Food Canada, Ottawa, and Organization for Economic Co-operation and Development, Paris.

Ogle, S.M., F.J. Breidt, and K. Paustian. 2005. Agricultural management impacts on soil organic carbon storage under moist and dry climatic conditions of temperate and tropical regions. Biogeochemistry, *in press*.

Ogle, S.M and K. Paustian. 2005. Soil organic carbon as an indicator of environmental quality at the national scale: Inventory monitoring methods and policy relevance. Canadian Journal of Soil Science (in press).

Orthofer, R., E. Gebetsroither, F. Strebl, H. Schwaiger, G. Jungmaier, J. Spitzer, L. Canella, G. Hintermeier, A. Windsperger, and F. Schneider. 2000. The Austrian Carbon Balance Model (ACBM) Final Project Report. Austrian Research Centers, Austria.

Parton, W.J., M. Hartman, D. Ojima, and D. Schimel. 1998. DAYCENT and its land surface submodel: description and testing. Global and Planetary Change 19:35-48.

Paustian, K., H.P. Collins, and E.A. Paul. 1997. Management controls on soil carbon. p. 15- 49. *In* E.A. Paul, E.T. Elliott, K. Paustian, and C.V. Cole (ed.) Soil Organic Matter in Temperate Agroecosystems. CRC Press, New York.

Paustian, K., J. Six, E. Elliott, and H.W. Hunt. 2000. Management options for reducing CO_2 emissions from agricultural soils. Biogeochemistry 48:147-163.

RIVM. 2002. RIVM Report 773201 006. Greenhouse Gas Emissions in the Netherlands 1990- 2000. National Inventory Report 2002. RIVM. National Institute of Public Health and the Environment, Bilthoven, the Netherlands.

Schlesinger, W.H. 1997. Biogeochemistry: an analysis of global change. Academic Press, San Diego, California.

Scott, N.A., K.A. Tate, D.J. Giltrap, C.T. Smith, R.H. Wilde, P.F.J. Newsome, and M.R. Davis. 2002. Monitoring land-use change effects on soil carbon in New Zealand: quantifying baseline soil carbon stocks. Environ. Pollut. 116:S167-S186.

Smith, P. 2004. Carbon sequestration in croplands: the potential in Europe and the global context. Eur. J. Agron. 20:229-236.

Smith, W.N., P. Rochette, C. Monreal, R.L. Desjardins, E. Pattey, and A. Jaques. 1997. The rate of carbon change in agricultural soils in Canada at the landscape level. Can. J. Soil Sci. 77:219-229.

Swedish Ministry of the Environment. 2002. Sweden's third national communication on Climate Change under the United Nations Framework Convention on Climate Change. Ministry of the Environment, Sweden.

Tate, K.R., N.A. Scott, A. Parshotam, L. Brown, R.H. Wilde, D.J. Giltrap, N.A. Trustrum, B. Gomez, and D.J. Ross. 2000. A multi-scale analysis of a terrestrial carbon budget. Is New Zealand a source or sink of carbon? Agr. Ecosyst. Environ. 82:229-246.

Tate, K.R., N.A. Scott, S. Saggar, D.J. Giltrap, W.T. Baisden, P.F. Newsome, C.M. Trotter and R.H. Wilde. 2003. Land-use change alters New Zealand's terrestrial carbon budget: uncertainties associated with estimates of soil carbon change between 1990- 2000. Tellus Series B-Chemical and Physical Meteorology 55 (2):364-377.

Trotter, C.M., K.R. Tate, S. Saggar, N.A. Scott, and M.A. Sutherland. 2004. A multi-scale analysis of a national terrestrial carbon budget and the effects of land-use change. P 311-341. *In* M. Shiyomi et al. (ed.) Global Environmental Change in the Ocean and lands. TERRAPUB, Tokyo.

UNFCCC. 1999. FCCC/CP/1999/7. UNFCCC documents [Online]. Available at www. unfccc.int/resource/docs/cop5/07.pdf. (accessed 25 July, 2004; verified 20 April, 2005). UNFCCC Secretariat, Bonn, Germany.

UNFCCC/NIS. 2003. National Inventory Submissions 2003. [Online]. Available at www. unfccc.int/national_reports/annex_i_ghg_inventories/national_inventories_submissions/items/618.php. (acessed 25 July, 2004; verified 20 April, 2005). UNFCCC Secretariat, Bonn, Germany.

UNFCCC/NIS. 2004. National Inventory Submissions 2004. [Online]. Available at www. unfccc.int/program/mis/ghg/submis2004.html. (accessed 25 July, 2004; verified 20 April, 2005). UNFCCC Secretariat, Bonn, Germany.

US EPA, 2004. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2002. [Online]. Available at www. epa.gov/globalwarming/publications/emissions. (acessed 25 July, 2004; verified 20 April, 2005). US Environmental Protection Agency. Washington, D.C.

Weiss, P., and B. Schlamadinger. 2000. Research activities related to the role of forests and forestry in climate change mitigation in Austria. Biotechnol. Agron. Soc. Environ. 4(4): 252-258.

Table 1. N_2O emissions (Gg of CO_2 equivalent) from Agricultural soils in Annex 1 countries (Compiled using the data from the UNFCCC Greenhouse Gas Inventory Database). NR = Not reported.

Country	1990	2000		
	Agricultural Soils	Agricultural soils	National total	% Agricultural soils vs national total
Australia	14,669	18,077	31,906	56.66
Austria	1,024	987	2,515	39.24
Belarus	NR	NR	NR	NR
Belgium	4,910	4,891	13,422	36.44
Bulgaria*	16,712	NR	NR	NR
Canada	27,364	33,654	53,938	62.39
Croatia*	2,361	NR	NR	NR
Czech Republic	7,568	4,732	8,175	57.88
Denmark	9,797	7,853	9,083	86.46
Estonia	952	366	414	88.41
European Community	198,043	189,726	338,111	56.11
Finland	4,373	3,496	7,183	48.67
France	51,975	50,571	76,891	65.77
Germany	30,926	27,351	60,080	45.52
Greece	6,501	6,370	11,009	57.86
Hungary	1,414	11,339	12,698	89.30
Iceland	70	69	124	55.65
Ireland	6,552	6,666	9,725	68.54
Italy	20,337	20,554	43,176	47.61
Japan	9,607	8,055	36,870	21.85
Latvia	2,998	952	1,288	73.91
Liechtenstein	NR	NR	NR	NR
Lithuania	NR	NR	NR	NR
Luxembourg*	147	NR	94	NR
Monaco	NR	NR	NR	NR
Netherlands	6,650	7,352	16,980	43.30
New Zealand	11,454	12,100	12,654	95.62
Norway	2,623	2,535	5,154	49.19
Poland	13,358	10,712	23,896	44.83
Portugal	4,791	4,634	8,258	56.12
Romania*	7,766	NR	NR	NR
Russian Federation	NR	NR	NR	NR
Slovakia	4,154	2,181	3,085	70.70
Slovenia*	1,416	NR	NR	NR
Spain	16,023	18,570	30,497	60.89
Sweden	3,792	3,603	6,916	52.10
Switzerland	2,404	2,165	3,619	59.82
Ukraine	NR	NR	NR	NR
United Kingdom	30,353	26,829	43,878	61.14
USA	267,088	297,561	425,345	69.96
Total	761,770	783,951	1,296,892	60.45

*- excluded in calculating the total of the columns, as one or more estimates are missing for these countries

Table 2. Methodologies used by Annex 1 countries for agricultural soil emission estimation according to the latest submissions of National Inventory Report (NIR) and Common Reporting Format (CRF), in summary. For N_2O emissions, methods and emission factors have been classified based on the major methodology being used. For CO_2 emissions, methods and emission factors are given in the order: mineral soil/ organic soil/ liming.

Country	N_2O emissions		CO_2 emissions/ removals		Additional Remarks
	Methods	Emission Factor	Methods	Emission Factor	
Australia	T2	CS	T3/NE/NE	CS/NE/NE	N_2O – Australian-specific methods and emission factors. Methodology incorporates increased emissions due to soil disturbance caused by cropping practices; soil disturbance is a country-specific category that combines emissions from several IPCC categories. CO_2- Only stock change due to pasture improvement and minimum tillage (with methods and emission factor given as country-specific) given in the CRF under removals from mineral soils ((UNFCCC/NIS, 2004).
Austria	T1	D	NE	NE	No official reporting of soil C stock changes; estimates based on Tier 3 methodology are planned for future (UNFCCC/NIS, 2004).
Belarus	NR	NR	NR	NR	N_2O Estimates given although no details on specific methods or emission factors provided. No information on estimates of CO_2 from agricultural soils provided in either NIR or CRF (UNFCCC/NIS, 2004).
Belgium	T1	D	NE	NE	N_2O- calculated using IPCC methodology with country- or region- specific data. A national study going on regarding CO_2 estimation (UNFCCC/NIS, 2004)
Bulgaria	T1	D	NE	NE	CO_2- not estimated as country specific data are lacking and certain activities not taking place (UNFCCC/NIS, 2004).
Canada	T1	D	T3/T1/T1	CS/D/D	C stocks- Century model calibrated to Canadian conditions used (Smith et al. 1997). CO_2 emission from liming estimated using stoichiometric relationships relevant to the breakdown of dolomite and limestone into CO_2 and other minerals (UNFCCC/NIS, 2004).
Croatia	NR	NR	NR	NR	N_2O estimates are given, but no sufficient details on the specific method or emission factor available in either NIR or CRF (UNFCCC/NIS, 2004). CO_2- not reported; data for stock changes from cultivation of mineral soil not well documented (UNFCCC/NIS, 2004).
Czech Republic	T1	D	NE	NE	N_2O- A complex methodology that incorporates agricultural soils and other sub sectors, based on recent studies. Methods for CO_2 under GPG are still being developed (UNFCCC/NIS, 2004).
Denmark	T2	CS	NE	NE	N_2O- Methods include IPCC T1b (e.g. emission from N-fixing crops) and models (e.g. emission from N leaching and runoff) relevant to Danish conditions (UNFCCC/NIS, 2004) .

Estonia	T1	D	T1/NE/NE	D/NE/NE	CO_2- Stock changes in mineral soils only; T1a method and default emission factors used (UNFCCC/NIS, 2004).
Finland	T1	D	T1/T1/T1	D/D,CS/D	Estimated based on activity data from annual agricultural statistics, publications, databases and agricultural experts. Both IPCC default and national values for emission factors and other parameters used (UNFCCC/NIS, 2004).
France	C/ T2	D,CS	T2/NE/T1	CS/NE/D	Methodology for N_2O emissions available in a special document, i.e. CITEPA ("Méthodologie utilisée pour les inventaires de NH_3 et de N_2O provenant des activités agricoles : évolution et perspectives") (UNFCCC/NIS, 2003). French methodology does not consider C emission/ uptake in relation to the nature of different soils; thus CO_2 reported for 'all soil types' under mineral soils (UNFCCC/NIS, 2004).
Germany	C/T2	D,CS	NE/NE/T2	NE/NE/CS	CO_2 from liming estimated using a country-specific emission factor (UNFCCC/NIS, 2004).
Greece	T1	D	T2/NO/NO	CS/NO/NO	Estimates of mineral soil C stock changes are not officially accepted due to lack of readily available data (estimates of stock changes in mineral soils for afforestation of agricultural soils given). CO_2 from cultivation of organic soil and liming are given as not occurring. Good Practice Guidance (GPG) followed, but currently not fully implemented for LULUCF sector (UNFCCC/NIS, 2004).
Hungary	T2	CS	T1/NO/T2	D/NO/CS	N_2O- IPCC method with activity data and emission factors derived using agricultural statistics; CO_2 from mineral soils and liming estimated; liming involves limestone and a lower emission factor compared to the IPCC default has been used (UNFCCC/NIS, 2004).
Iceland	T1	D	NR	NR	
Ireland	T1	D	NR/NR/T1	NR/NR/D	N_2O- GPG and a country-specific method used in estimating N_2O emissions from agricultural soils; Tier 1b used to estimate N inputs from N fixing crops and crop residues returned to the soil; Development of T2 methods for N_2O in progress. CO_2- Emission from mineral soils and organic soils not reported, pending the results of major research in this area (UNFCCC/NIS, 2004).
Italy	T1	D	T1/NR/NR	D/NR/NR	CO_2 for mineral soils estimated, with default IPCC emission factors for base factor, tillage and input factors. National expert evaluations used for the amount of organic carbon in soil (UNFCCC/NIS, 2004).
Japan	T2	CS	NE	NE	
Latvia	T1	D	NE/T1/T1	NE/D/D	N_2O- T1 method with IPCC default, and national emission factors and parameters; activity data from agricultural statistics and agricultural experts. CO_2- No estimation for mineral soils, as data are not available.(UNFCCC/NIS, 2004).
Liechtenstein					No submission.
Lithuania	T1	D	NE/T1/NE	NE/D/NE	CO_2 from organic soil (upland crops) estimated (UNFCCC/NIS, 2004).
Luxembourg	NR	NR	NR	NR	
Monaco	NO	NO	NO	NO	

Netherlands	T2	CS	NE	NE	N_2O- Based on methods described in the N_2O background document by Kroeze (1994, as cited in RIVM, 2002). Under indirect N_2O emissions, 'Background agricultural soils' reported that include N_2O emissions from cultivation of histosols and emissions from manure and fertilizer applications done in the past, to reflect the long term emission effects of past agricultural practices (UNFCCC/NIS, 2003, 2004).
New Zealand	T1	D,CS	NR/NR/T2	NR/NR/CS	N_2O- A country-specific emission factor has been used for emissions from animal production. CO_2 emission from liming estimated, with a lower C conversion factor compared to the default. (NRI and UNFCCC/NIS, 2004).
Norway	T1	D	NR/NR/T2	NR/NR/CS	CO_2 from liming estimated using data from Norwegian Agricultural Inspection service, and a country-specific emission factor (which is equivalent to the default) (UNFCCC/NIS, 2004).
Poland	T2	CS	T1/NR/NR	D/NR/NR	Follows the revised IPCC Guidelines with application of specific (Tier 2/3), and simple methods (in few cases) (UNFCCC/NIS, 2003, ,2004).
Portugal	T1	D	NE	NE	
Romania	T1	D	NE	NE	CO_2 not estimated due to the lack of data (UNFCCC/NIS, 2004).
Russian Federation					No submission.
Slovakia	T1	D	NO	NO	
Slovenia	T1	D	T1/NE/T1	D/NE/D	CO_2 from only mineral soils and liming estimated (UNFCCC/NIS, 2004).
Spain	T1	D	NE/NO/NO	NE/NO/NO	N_2O- IPCC guidelines followed; activity parameters and variables based on national data, and default emission factors. CO_2 not estimated due to lack of reliable data (UNFCCC/NIS, 2004).
Sweden	T2	CS	NR/T2/T1	NR/CS/D	N_2O - Activity data are mainly based on official Swedish statistics; default and national emission factors used for different sub sources in Direct N_2O emissions, and emissions from animal production. CO_2- Removal from mineral soils not reported; for emissions from organic soils, a national emission factor used. IPCC methodology and default emission factors used for emission from liming (UNFCCC/NIS, 2004).
Switzerland	T2	CS	NR/T2/NR	NR/CS/NR	N_2O- Estimated using IULIA, an IPCC-derived national method that uses default emission factors, but makes adjustments in the source categories, appropriating the conditions in Switzerland; both default and country specific emission factors used for estimating emission from sub source categories. CO_2- cultivated organic soils only (UNFCCC/NIS, 2004).
Ukraine					
United Kingdom	T1	D	T3/T2/T2	CS/CS/CS	CO_2 from application of limestone, chalk and dolomite to agricultural soils estimated using emission factors based on the stoichiometry of the reaction, assuming pure limestone and dolomite (UNFCCC/NIS, 2004).

United States of America	T1	D	T2/T2/T1	CS/CS/D	N_2O- Revised IPCC methodology and national methodologies. CO_2 from all three source categories (mineral soils, organic soils and liming) estimated. IPCC methodology with continued improvements being used to estimate C stocks (UNFCCC/NIS, 2004).
Non-Annex 1 countries					Emissions (sources or sinks) from agricultural soils rarely recorded; IPCC default method is mostly used in the recorded instances.

Note: Under Methods: T1- IPCC Tier1, T2- IPCC Tier2 (i.e. IPCC default method and country specific emission factor/s), T3- Country specific, C- CORINAIR;

under emission factors: D- IPCC Default, CS- Country Specific;

NO- Not Occurring; NE- Not Estimated; NR- Not Reported;

(Sources: Latest National Inventory Reports (NIR)s/ Common Reporting Format (CRFs) and information obtained by contacting responsible country offices)

Table 3. Organic C values derived for arable soils in New Zealand using the soil Carbon Monitoring System (CMS), compared with IPCC GPG/default soil C values (Mg C ha^{-1}, 0-30cm) for native vegetation in similar soil/climate categories.

IPCC soil/ climate	**IPCC**	**CMS**
High Clay Activity/ Cold temp dry	50	67
High Clay Activity / Cold temp moist	95	84
Aquic	87	82
Volcanic	70-130	99

(Source: Scott et al., 2002 ; IPCC GPG, 2003)

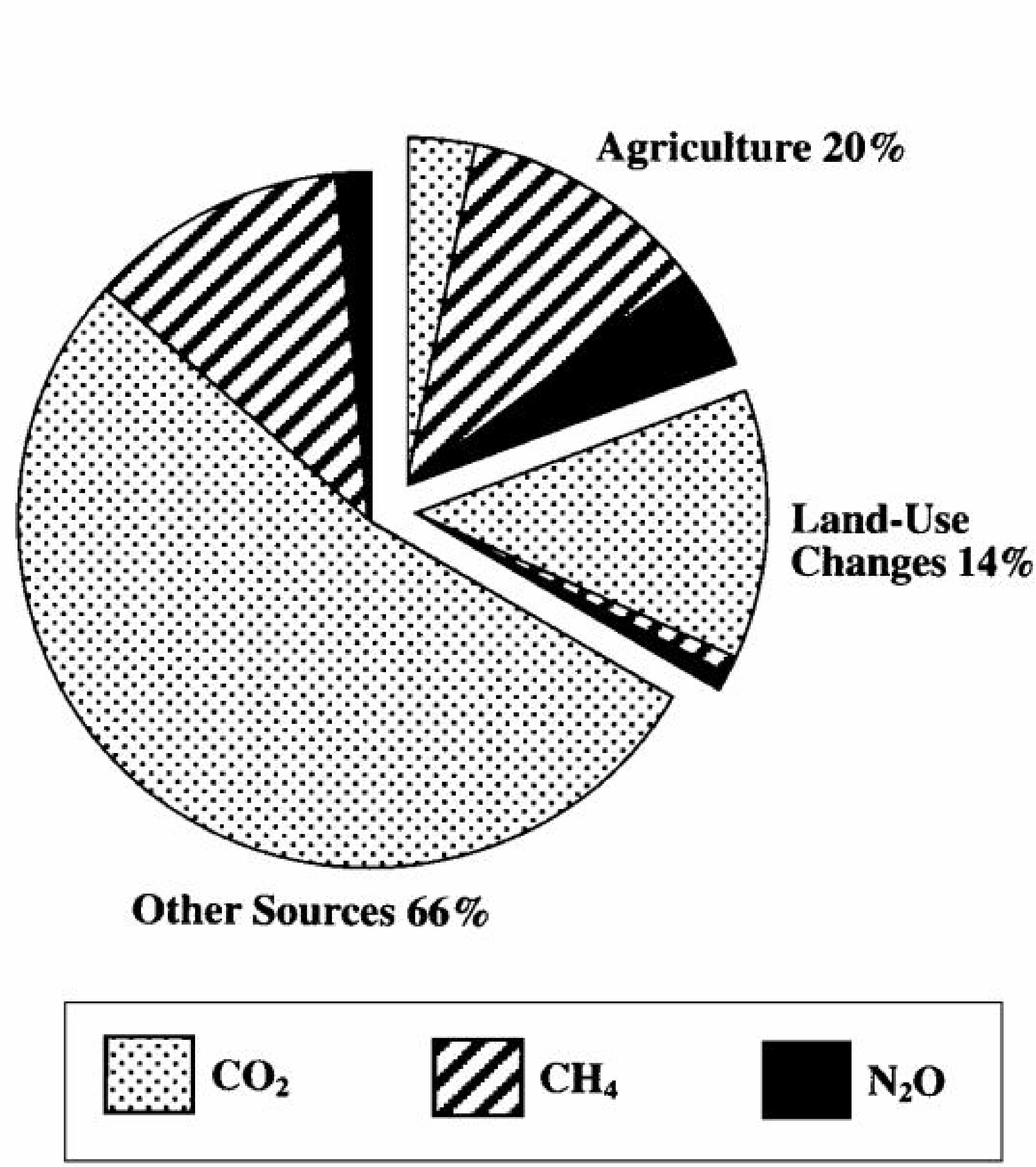


Figure 1. Percent contribution of agricultural sector greenhouse gas emissions to increased radiative forcing attributed to anthropogenic enhancement of the greenhouse effect (IPCC, 1996).

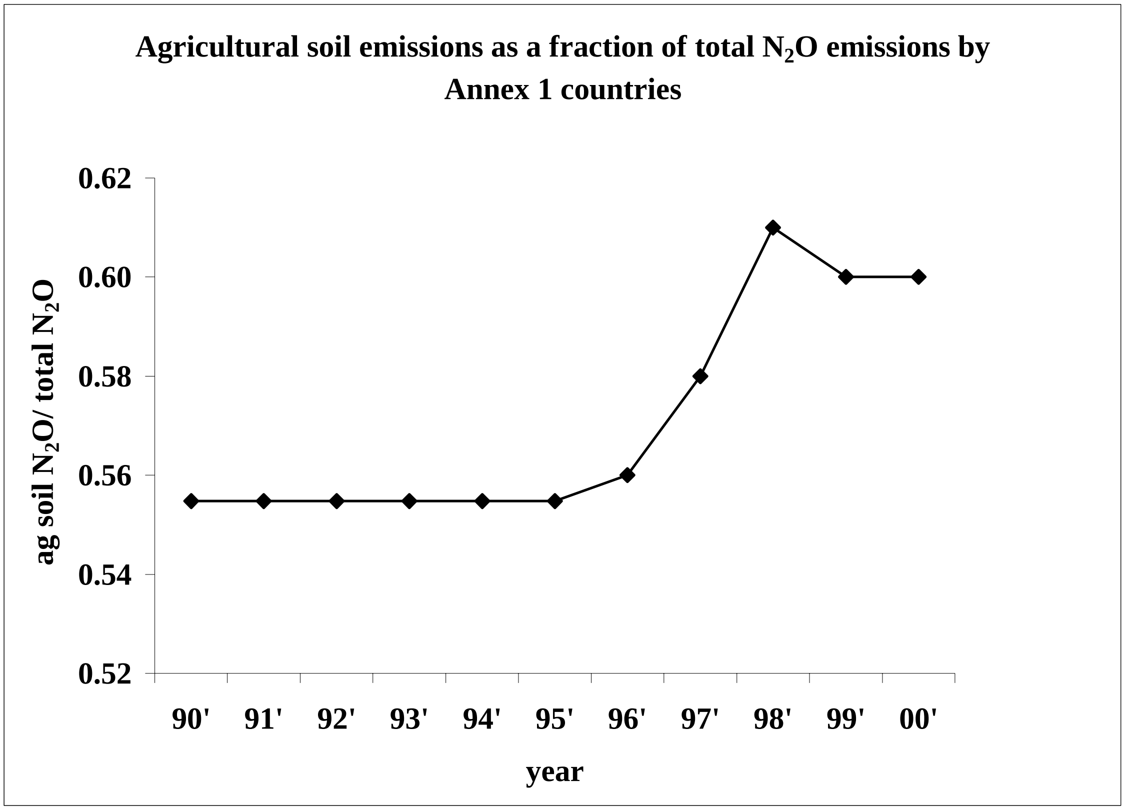


Figure 2. N_2O emission from agricultural soils as a fraction of cumulative N_2O emissions by Annex 1 countries from 1990- 2000.

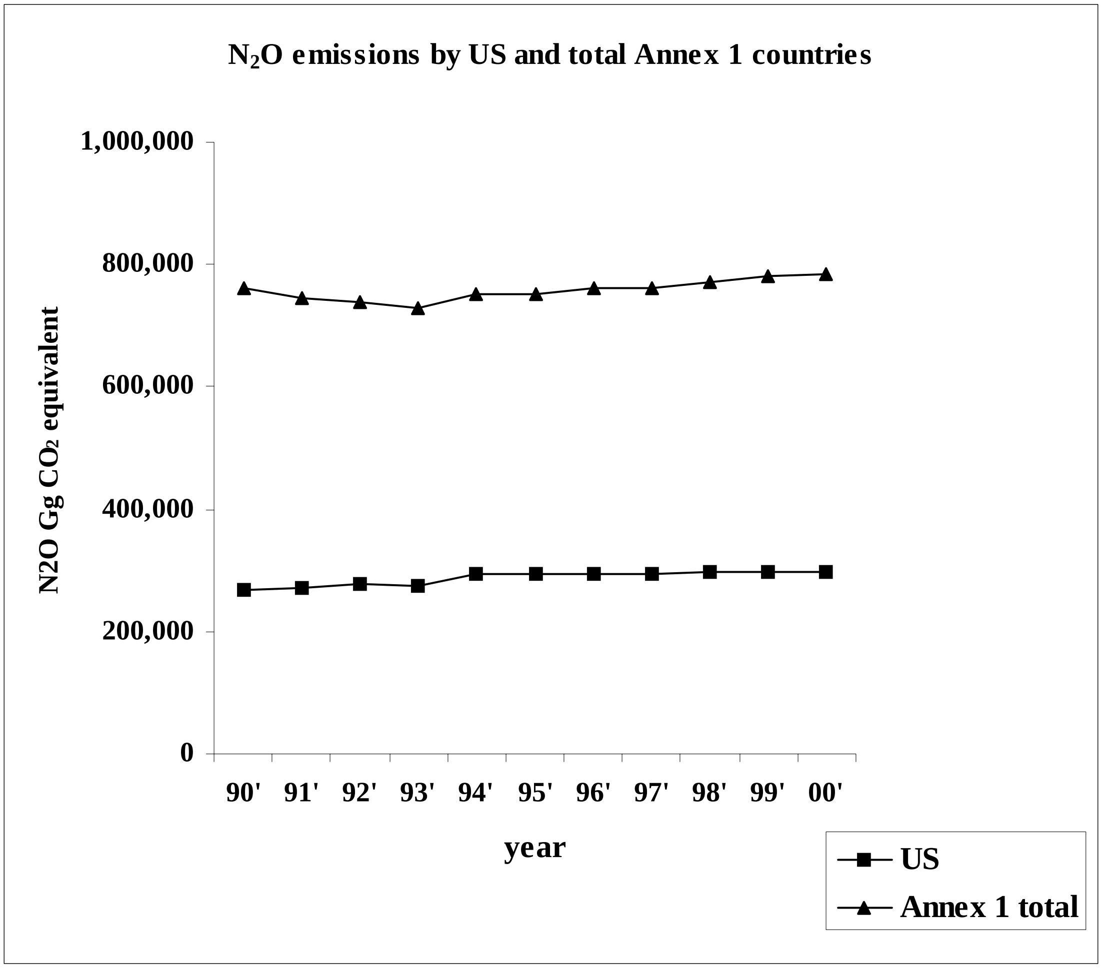


Figure 3. N_2O emission from agricultural soils by US and all Annex 1 countries from 1990- 2000.

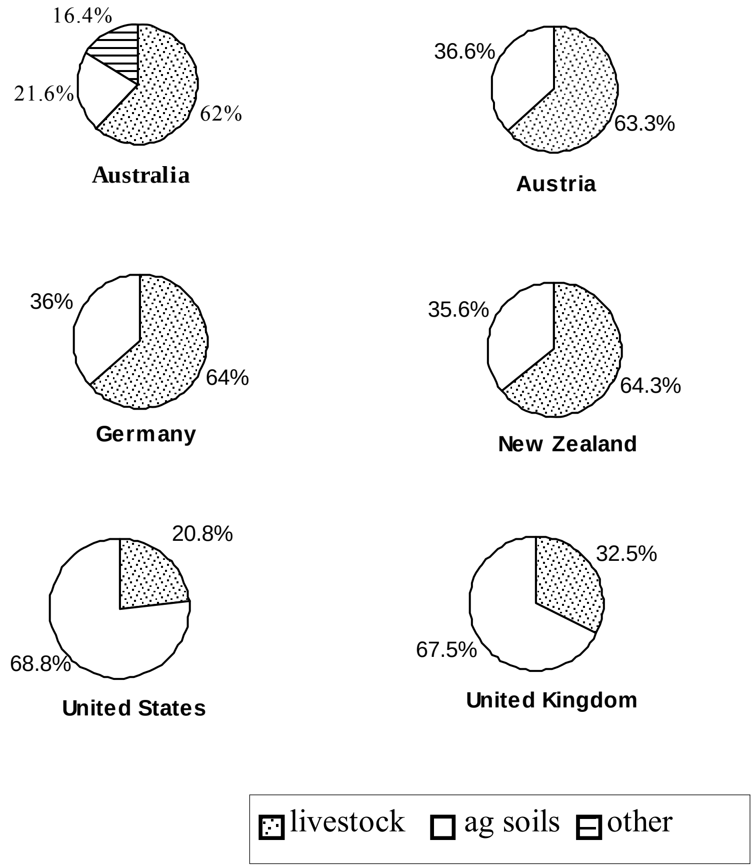


Figure 4. Relative impact of agricultural soils to the overall agricultural sector, in terms of Global Warming Potential, (computed as absolute values for both sinks and sources combined for CO_2, N_2O and CH_4), as a percentage of the total agricultural sector emissions for some selected Annex 1 countries. 'Other' emissions represent combined N_2O and CH_4 emissions from burning of savannas and crop residues and CH_4 from rice.

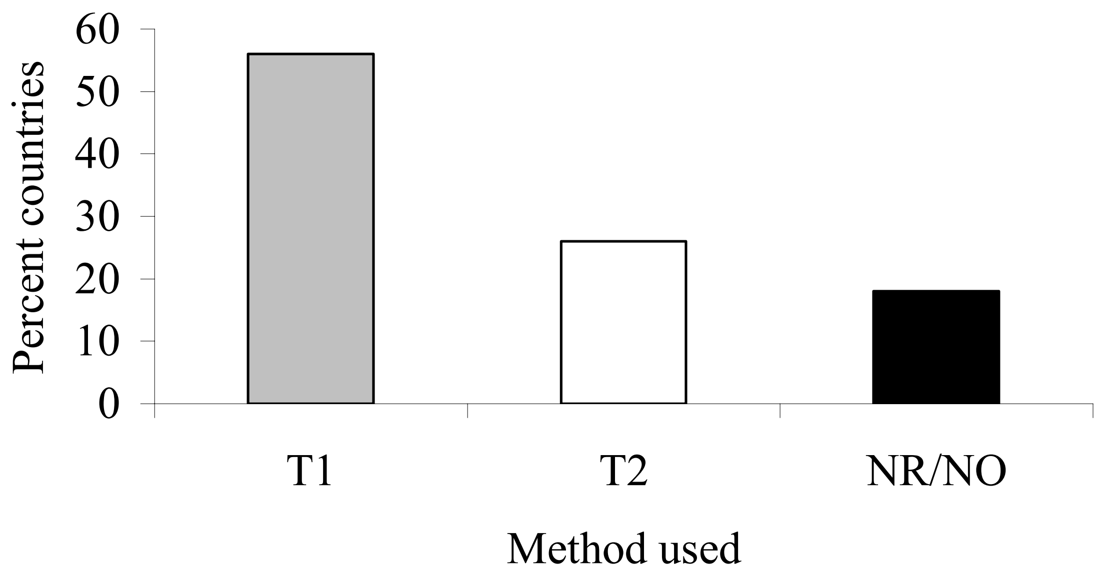


Figure 5. Methods used for estimating N_2O emissions by Annex 1 countries based on the NIRs and CRFs submitted in 2004 (T1- Tier1; T2- Tier2, NR/NO- Not Reported or Not Occurring).

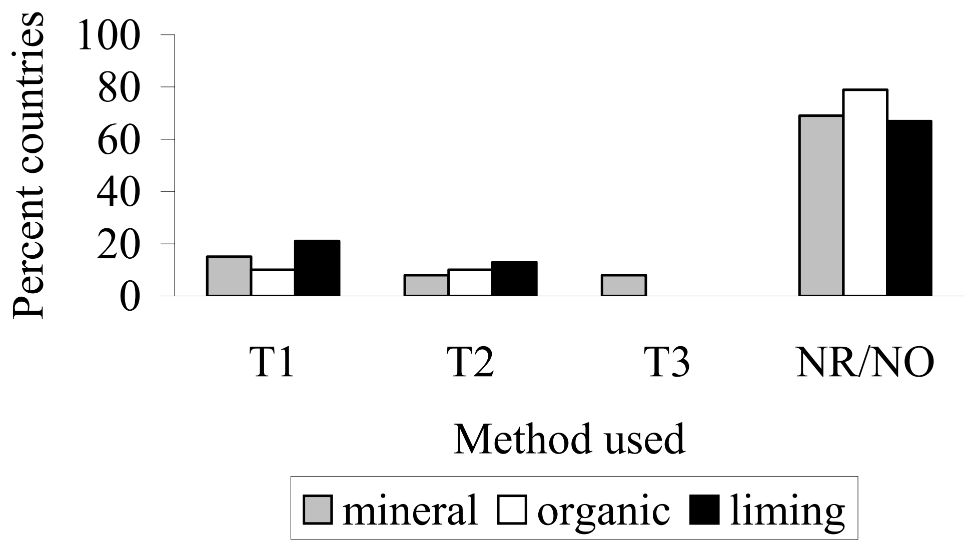


Figure 6. Methods used for estimating CO_2 emissions by Annex 1 countries based on the NIRs and CRFs submitted in 2004. (mineral- stock changes in C from mineral soils, organic- emissions from cultivated organic soils, liming- emissions due to liming, T1- Tier1; T2- Tier2, T2- Tier 3, NR/NO- Not Reported or Not Occurring).

CHAPTER 3

USE OF AVHRR TIME SERIES FOR ESTIMATING MISSING CROP BIOMASS VALUES DERIVED FROM THE NATIONAL AGRICULTURAL STATISTICS SURVEY

ABSTRACT

Carbon input from crop residues is an important determinant of the carbon balance in agricultural soil. Crop residue production can be estimated from biomass and crop yield data from ground-based surveys. However, survey data may be unavailable for certain time periods and/or locations. Remotely sensed data is collected on a regular schedule and may also provide more spatially resolved data compared with crop yield surveys. We analyzed the relationship between composited biweekly AVHRR NDVI and crop aboveground biomass, using biomass estimated from county-level yield data reported by NASS (National Agricultural Statistics Survey) for three crops (corn, soybean, and oats), during the years 1992, 1997, and 2002. The aim of the study was exploring relationships between NDVI and crop biomass for potential future use in filling the gaps in counties where no NASS-reported yields are available. Aboveground biomass was estimated from the Pathfinder biweekly dataset of NDVI values, using canonical correlation analyses followed by best subset multiple regression incorporating canonical variates from NDVI

variables. Cross validation of the model estimates was done by randomly splitting the full dataset into training and application subsets, simulating a 10% to 40% range of percent missing values. NDVI and biomass values of the major crops in Iowa during a given year were well correlated, and the estimated values were very close to the observed values, with < 5% relative error and R^2 values > 0.8 in most cases. Using the available (training) data from a single year or a combination of years to derive models for filling the missing (i.e. validation) data within the same time period, yielded mean estimated biomass values with less than 1% relative error and bias. However, applying the models derived using the data from any single year (or a combination of years), on a different year with missing data was less appropriate since it yielded low R^2 values for the relationships between biomass and NDVI, although the mean residuals were low. Thus multiple regression analyses using biomass and NDVI canonical variates were found to be a suitable approach in predicting missing biomass for subsequent estimation of crop residue carbon inputs, when the training data and validation data are from the same time period.

INTRODUCTION

The use of remote sensing for crop forecasting goes back to the early 1970's (reviewed in MacDonald & Hall, 1980). Since then, agricultural agencies in various countries (e.g., Canada, Hungary, and US) have been using sensors such as the National Oceanic and Atmospheric Administration Advanced Very High Resolution Radiometer (NOAA

AVHRR) and Landsat imagery to forecast crop yields and crop conditions (Allen et al., 2002; Csornai et al., 2002; Reichert & Cassey, 2002)

Several studies have used multispectral and hyperspectral data for spatially explicit crop forecasting and yield estimation. Many use the Normalized Difference Vegetation Index (NDVI) to estimate biomass (Lozano-Garcia et al., 1991; Hansen and Schjoerring, 2003) and crop yields (Tucker et al., 1983,1985; Quarmby, 1993; Senay et al., 2000; Yang et al., 2000; Doraiswamy et al., 2001, 2003, 2004, 2005; Hill and Donald, 2003; Knudby, 2004); these studies were mostly conducted at field scale. Certain studies have used integrated NDVI over the crop growth period too correlate NDVI and biomass (e.g., Tucker et al., 1983, 1985; Quarmby, 1993). NDVI is a vegetation index that ranges between -1 and +1, and is the difference between near infrared and red channels normalized by their sum (i.e. NDVI= (NIR- R)/(NIR+ R). Increasing positive values indicate increasing green vegetation, and negative values indicate non-vegetated surface features such as water, ice, snow, clouds, etc. The relationship between vegetation indices such as NDVI and biomass depends on the relationship between the vegetation index and Leaf Area Index (LAI) and the relationship between LAI and biomass. According to Curran (1981), NDVI is directly related to biomass when biomass is linearly correlated with LAI. In comparing several vegetation indices for plants in salt marshes, Modenese et al. (2005) found NDVI to give the highest correlation with aboveground biomass. Currently, the National Agricultural Statistics Service (NASS) of the US Department of Agriculture (USDA) uses biweekly composite NDVI images from AVHRR to monitor crop condition, and for crop forecasting, while Landsat imagery is mainly used to estimate crop acreage at the county level, under its Cropland Data Layer Program.

Our study was carried out as part of an effort to assess the carbon (C) dynamics of agricultural soils in the conterminous U.S. Crop residues are the main component of C inputs to agricultural soils, and include about 50-60% of the total aboveground crop biomass. Crop yields can be used to estimate above and below ground biomass (and hence residue C inputs) based on allometric relationships for different crops (Buvanovsky and Wagner 1986; Prince et al., 2001; Campbell and Jong, 2003; Williams and Paustian, submitted). In our main study we considered a 16- year period from 1982 to 1997 during which digitial databases were available from both NASS (annual data) and Census of Agriculture (Ag Census; data reported every 5 years), although Ag Census data were not used in this study. One limitation of using the yield data reported by NASSis missing data for certain years in certain counties. Therefore we explored the potential use of remote sensing in estimating crop aboveground biomass in those counties where there is no yield information, using county-level crop production in Iowa as a test bed.

MATERIALS AND METHODS

Remotely sensed data used in this study included biweekly AVHRR NDVI images from 1992,1997, and 2002 (corrected for cloud-contaminated pixels), and the 1992 National Land Cover Dataset (NLCD) produced by the US Geological Survey (USGS) using Landsat images. Annual yield data reported by NASS for 1992,1997, and 2002 were used as the ground data for estimating aboveground biomass; these years were chosen as the most recent years that both NASS and Ag Census have reported yields. The state of Iowa was chosen for the study, since it is one of the crop states in the US where NASS has excellent reporting of yields and acreage without missing data, thus making it a good place to observe relationships between crop aboveground biomass and NDVI.

NOAA AVHRR biweekly NDVI images (1 km resolution) of the conterminous U.S. were obtained from the Pathfinder dataset for the years 1992, 1997, and 2002 (available at ftp://data.nodc.noaa.gov/pub/data.nodc/pathfinder). For each of the three years, images collected during the growing season (beginning of April to end of October) were combined as bands within a single multi temporal image. A county map of the U.S. was used to get the county boundaries and extract composite NDVI images for Iowa. The NLCD coverage for Iowa was recoded to exclude non-crop areas and mask the composite NDVI images from the three years. The cropland areas selected consisted of the NLCD categories for small grains (i.e. oats in Iowa) and row crops (corn and soybean). Since the NLCD coverage with crop layers had 30 m resolution, pixels from NLCD crop layers were aggregated to 1000m resolution for masking the NDVI images. If cropland area in the combined NLCD pixels were >75% annual cropland, the entire pixel was classified as annual cropland, otherwise the combined pixel was classified as non-cropland. Average biweekly NDVI pixel values (from the separate biweekly layers of the composite image) were then calculated for each county. Image processing was performed using Erdas Imagine 8.6 (Leica Geosystems) and ArcGIS 8.1 (ESRI).

NASS reports annual county-level yields by extrapolating yield data collected from a representative sample of farms in each country. Crop yields and areas for the years 1992, 1997, and 2002 reported by NASS (available at http://www.nass.usda.gov/Data_and_Statistics/index.asp) were used to estimate percent crop area and aboveground biomass. In Iowa, the major crops in all years (1992, 1997, and 2002) were corn (*Zea mays* L.) and soybean (*Glycine max* L.); oats (*Avena sativa* L.) was also considered in this study as a third major crop. For each county, the percentage of the county's total area occupied by any of the three crops, and the percentage of each crop within the total annual crop area were considered.

Aboveground crop biomass was calculated using the county-level yield data reported by NASS, and allometric equations relating grain yield to biomass for each crop (Williams and Paustian, submitted). Crop yields were corrected for moisture content, and converted to biomass of yield dry matter, before applying the crop allometric equations to estimate total aboveground biomass for comparing with NDVI values.

The range of dependent and independent variables used and the basis for selection of the methodology

Since the aim of the study was to derive general relationships between remotely sensed data and crop biomass estimated from yield statistics, the following crop variables were considered as the dependent variables:

a. Mean aboveground biomass per hectare of each crop ([C,S,O]AGBM kg ha^{-1} ,where C,S,O stand for corn, soybean and oats)

b. Area-weighted biomass (AWBM kg ha^{-1}; the sum of the aboveground biomass of the crops, weighted by their area fraction)

b. i.e. AWBM= $\sum([C,S,O]AGBM * [C,S,O]AF)$, where *AF* is the area of the particular crop as a fraction of the total crop area

Area-weighted biomass from the three crops was included assuming it might match better with the NDVI signal at pixel level, as it represents the mixed-crop biomass per hectare. Biomass data for the above variables were estimated for three different years (i.e., 1992, 1997, and 2002).

The pixel values for two-week periods during the growing season of 1992, 1997, and 2002 (Table 1) were considered as independent variables. Since there were slight differences in the beginning and end dates of the biweekly time intervals in the three years, biweekly periods of the growing season in 1997 and 2002 were matched with the corresponding 1992 periods that had the greatest date overlap for the analyses and interpretation of the results.

Initial analyses of the data showed significant correlations between temporally adjacent NDVI values, such that the values from different time periods cannot be treated as independent (i.e. they exhibit multicollinearity). To address the problem of multicollinearity, we used canonical correlation analyses to model crop biomass as a function of NDVI.

Canonical correlation analyses

Canonical correlation analysis (CCA) is a statistical approach that summarizes multiple variables from two datasets as pairs of canonical variates. Although CCA treats both sets of variables identically, it is convenient to label one dataset independent and the other dependant, in this case these are the remotely sensing NDVI values and crop biomass values, respectively. Pairs of canonical variates are created as linear combinations of the original variables in each datasets. CCA maximizes the correlation between linear combinations of variables from one set with linear combinations of variables from another set. The advantage of CCA is that it quantifies the redundancy in each set of variables. This, in turn, allows us to analyze both X and Y variables in terms of their relationships to other variables within their own dataset and to variables in the other dataset.

In this CCA analysis, monthly NDVI values were considered as one set of variables, and county-level corn aboveground biomass (CAGBM), soybean aboveground biomass (SAGBM), oats aboveground biomass (OAGBM), and area weighted biomass (AWBM) were considered as a second set of variables. One advantage of CCA is it eliminates the multicollinearity associated with both the biweekly NDVI values and the crop biomass values. It also provides more interpretable results, as the patterns of correlation within and between datasets are reduced to a smaller number of variates that are ranked by their importance in explaining variance in each dataset. As an example, we could have used separate multiple regression analyses to predict aboveground biomass for corn and soybean from NDVI values. However, doing so would obscure the fact that corn and soybean aboveground biomass co-vary due to two effects: the limitation on total area of cropland in each county, and the fact that good years for corn are generally good years for soybean as well. In addition, the coefficients for the resulting regression equations would be difficult to interpret, as they would combine numerous effects into a single linear combination of the independent variables, whereas CCA separates out effects into separate canonical variates.

Best subset multiple regression analyses were performed to estimate each dependant variable from the canonical variates derived from the NDVI dataset. To obtain the model with the highest predictive power, the best subset of the entire set of independent variables was chosen based on Mallows' C_p value (a measure of model fit) and R^2 values for each subset of the independent variables; i.e., the subset that gave the lowest C_p (with C_p value less than or equal to 1, when p= number of parameters), and highest R^2 was chosen as the best.

These analyses were performed using datasets in which 10, 20, or 40 percent of the training data were randomly removed to represent missing data (the validation dataset), and the remaining data were used as a training data set to create regression equations. To obtain results that were insensitive to the particular selection of missing variables, canonical correlations and best subset multiple regression analyses were performed iteratively 100 times for cross-validation. Analyses were performed in the IDL software package (Research Systems, 2005). The following scenarios were evaluated:

a. Analyses were carried out separately for each year: 10, 20, and 40 percent of the data from that year were removed from the dataset for each year and treated as missing data. Regression equations developed using the training dataset were then applied to the reserved counties. In addition, to determine the year-to-year consistency of the equations, these same equations were extrapolated to estimate aboveground biomass values for the two other years,

b. Combined data from two years were used together and separated into training and validation sets, with the unused year's data used to check the ability to extrapolate beyond those years.

c. Data from all three years were combined and 10, 20, and 40 percent of the missing data from the same set were considered as the validation data set for the models obtained using training data.

RESULTS

The final composite NDVI images for the crop layers in 1992, 1997, and 2002 are shown in Figure 1. When the temporal variation in NDVI was studied for each year, we found that NDVI increased from the beginning of April, and remained high from mid June through early October in 1992 and 2002, and mid June through mid September in 1997. Highest NDVI in all three years were observed from the end of July or early August until early September. In general, NDVI increased during this period, the lowest being in 1992 and highest in 2002. The highest county-averaged NDVI observed for Iowa cropland were 0.63, 0.71, and 0.79 in 1992, 1997, and 2002, respectively. The increasing trend over the years was more conspicuous during the period of high NDVI, but towards the end of the growing season, NDVI in 1997 was lower in certain biweekly periods, compared to the corresponding periods in 1992 (Figure 2).

According to the crop area information reported by NASS, about 80% of the counties (i.e. 78 out of 99) in Iowa had more than 50% land cover with these crops, and the average crop area in the counties were 89%. The percentage of the area occupied by corn within the cropland of a county ranged from 52 to 88% in 1992, 43 to 79% in 1997, and

38 to 70% in 2002. The percent area for soybean ranged from 3 to 48% in 1992, 14 to 55% in 1997, 24 to 59% in 2002, and that for oats ranged from 0 to 12% in 1992, 4 to 9% in 1997, and 0 to 6% in 2002. In 1992, the average percent crop area for corn, soybean and oats were 61%, 37%, and 2%, respectively; in both 1997 and 2002, the average percent annual crop areas for corn, soybean and oats were 53, 46 and 1%, respectively (Figure 3). The county-level average yields reported by NASS for these three crops were slightly different between the three years, and the highest yields and biomass values were found in 2002; average corn yields were 9, 8.5, and 10 Mg ha^{-1}, soybean yields were 2.9, 3, and 3.2 Mg ha^{-1}, and oats yields were 2.4, 2.6, and 2.7 Mg ha^{-1}in 1992, 1997, and 2002, respectively. Thus the yields for the crops increased over time, as reflected in the increased NDVI, too (Figure 2).

Canonical correlation analyses

For any of the years or combination of the years, the highest correlation between the first NDVI and crop biomass canonical variates was 0.92; the corresponding p-value of <0.0001 rejected the null hypothesis that all the canonical correlations are zero. The results for multivariate statistical tests also confirmed the significance of the canonical correlations obtained from the analyses.

Of the four canonical variates that contributed towards the observed model relationships with crop biomass (Tables 2 and 3), the first canonical variate (CV1) had the highest loadings from the original NDVI dataset; CV1 had positive loadings from NDVI pixel values of the biweekly periods in early April to mid/end June, and early September to end October in all three years (Figures 4 and 5). In all the models, CV1 was negatively correlated with crop biomass (Tables 2 and 3). Since CV1 has the highest (positive) loadings from the NDVI at the early and end phases of the crop growth cycle, the negative coefficient of CV1 in all the models with biomass indicates that NDVI is negatively or less correlated with biomass during the early and final phase of the crop growth cycle. This was confirmed by the negative correlation between biomass variables and the original NDVI pixel values from individual biweekly periods of the same time intervals. In addition, there was a positive correlation between crop biomass and original NDVI pixel values during the period from end June to end August; however, this correlation varied in value among the crops and different years, and ranged from 0.1 to 0.84). The loadings from original NDVI pixel values on the second, third, and fourth canonical variates were very low, except for the relatively high loadings on the second canonical variate in 1997 (Figure 4). However, all four canonical variates seemed to follow the same pattern of variation, with positive loadings from NDVI pixel values towards the end phases of crop growth. All four canonical variates contributed towards the model relationships with the biomass of each individual crop and area weighted total biomass.

Model relationships between NDVI derived canonical variates and crop biomass in relation to the extent of missing data

Model relationships were obtained for 10, 20, and 40 percent missing data in biomass under the scenarios a) data from a single year, b) two years of data combined, and c) all three years of data combined. In general, they had relatively high R^2 values for the models obtained from training and validation data sets (e.g. table2, figure 6). However, when the same model relationships were extrapolated to a different year or a combination of years, the same model relationships yielded relatively low R^2values (Figures 6 and 7). The extent of the difference between the mean estimated and observed values varied depending on the observed values in the year or the two years combined in the training dataset. For instance under scenario 2, when the models for CAGBM from combined data in 1992/1997 were applied on the extrapolated data in 2002, the mean estimated values were 12% lower than the observed values (i.e. relative error 12%), and when models from 1992/2002 combination was applied on 1997 data, the mean estimated values were 3% higher than the observed values (relative error –3%). The ratio of root mean square error/ mean estimated value (RMSE/ MPRED) ranged between 0.1 to 0.2 for CAGBM, 0.05- 0.2 for soybean, 0.1 to 0.3 for oats, and 0.1 to 0.2 for AWBM for the values in the extrapolated datasets under the first and second scenarios. Under the third scenario when the data from all three years were combined for the analyses, the model relationships estimated biomass values with less than 1% relative error (i.e. (observed-estimated)/ observed) and less than 0.05 of RMSE/ MPRED ratio for both training and validation data sets when 10, 20, and 40 percent data were missing. Thus the mean estimated values were very close (within 4% across all the biomass variables) to the mean observed values for each biomass variable (Table 4). However, the R^2values were slightly lower than those obtained for the training and validation data sets under the first two scenarios (i.e. the single year scenario and with combined data for two years); the observed R^2 for the validation data ranged for the third scenario ranged between 45-50% for AWBM, 60-70% for CAGBM, 50-60% for SAGBM, and 20-30% for OAGBM.

We analyzed the average bias (i.e. average residuals) for all three scenarios. Soybean always had the lowest bias (mostly within 5 kg ha^{-1}; Figure 9). However, oats being a minor crop with minimum crop area, showed the highest bias (still within 20 kg ha^{-1}) in relation to the mean observed biomass values. Corn had very low biases (less than 7 kg ha^{-1}) in 1992 and 1997, but the bias was slightly higher (close to 15 kg ha^{-1}) in 2002, in a year when the average corn biomass was much higher compared to the other two years; but still this bias was negligible since the average observed corn biomass in 2002 was 18353 kg ha^{-1}).

DISCUSSION

The current study tested the feasibility of using remotely sensed AVHRR NDVI information to predict crop aboveground biomass and estimate the residue carbon inputs from major crops in Iowa. Using original biweekly NDVI pixel values as independent variables in developing model relationships with biomass was not advisable, due to the presence of multicollinearity among the biweekly NDVI pixel values in certain time periods, especially during the early crop growth. Multicollinearity problems were avoided by using canonical correlation analyses that combined information from closely related, county- averaged biweekly NDVI pixel values during crop growth period into separate canonical covariates. Aboveground crop biomass variables derived using the annual yield data reported by NASS, served as the dependent variables.

In analyzing the correlations between original NDVI and individual crop biomass variables, biomass of all three crops was positively correlated with NDVI from end June to end August when the NDVI was at a maximum. Since corn and soybean had the highest crop area (mostly over 90%) and biomass, The largest contribution to NDVI must have come from these two crops . The usual harvest date for corn and soybean is October, and harvest dates for oats normally fall in July in Iowa. After the end of August, NDVI was negatively correlated with biomass, during a period when highest biomass should be found in the crops, especially for soybean and corn, due to maturation and end phase of grain-filling. The grain-filling period, where 40-50% of the biomass is allocated to grain, usually falls within the last 50-60 days of growth cycle in corn plants. Thus the maximum NDVI was observed within the beginning of the grain filling period of corn plants. The negative coefficient of the first canonical variate with biomass (and the observed correlation between biomass and original NDVI) denotes that biomass was negatively correlated with NDVI when the crop is close to harvest (in September-October). These results are in accordance with a study by Curran (1981), in which a negative correlation was found with high biomass in vegetation and NDVI. Canopy opening, and similar reflectance from drier or senescing vegetation towards the end of the crop cycle and soil in NIR and visible range (Todd et al., 1998; Campbell, 2002), might have led to low NDVI and negative correlation with biomass.

The coefficients obtained for the first canonical varate for different crops and different years varied, indicating the differences among the crop growth cycles and the differences in the biweekly periods in terms of crop phenology in different years. Both NDVI and biomass depends on external environmental factors such as precipitation and temperature; crop biomass depends on the number of growing degree-days and the temperature during the grain filling, etc. Therefore when such environmental factors vary between different years, it makes it less accurate to use the model derived from one year or a combination of two years on a different year. Our study confirmed this by showing low R^2 values when the models from scenarios 1 and 2 were extended to a different year, although the mean residual values were low. Overall, the best results were obtained when the models from training data were applied for the missing data in the same year or the combination of years. According to the results of the analyses, the mean estimated values or R^2 values were not much dependent on the extent of the missing data considered in the study; the results obtained for cross-validation using all three cases of missing-ness (i.e. 10, 20, and 40% of missing data), were very close, implying that our method could be used when more than 40% of the values are missing (i.e as the basis for crop biomass estimates with a smaller quantity of training data).

Although the first canonical variate was the most useful canonical variate in predicting and interpreting the NDVI-biomass relationship, the purpose of the current study was to come up with model relationships between NDVI and crop biomass. Thus the other canonical covariates were also considered in the best subsets multiple linear regression analyses, to come up with the subset of the canonical variates that would give the best model fit between the NDVI and biomass of each individual crop. The current study showed that canonical correlation analysis followed by best subset multiple regression analyses, is an improved approach in predicting aboveground biomass using NDVI values, especially as a means of predicting biomass when there are missing data in the reported crop statistics.

The mean estimated values for training data and validation data from 100 runs under each case of missing data for all three scenarios were very close, within 5% relative error. This confirms that when models are derived from the available data within the same time period (within the same year or the combination of the years that are relevant to the missing data), this approach is highly successful; The observed R-sq value gradually decreased when we increased the time period for choosing training and validation data from 1-, 2- or 3-years However, using the data from all three years was the best approach in predicting the missing data in any of the participating years, than using one or two years' data to predict the missing data in a different year, since the model extrapolation to a different year resulted in very low R-sq values.

CONCLUSION

Overall, the methodological approach used in the current study yielded model relationships between NDVI canonical variates and biomass variables with high R^2

values, and estimated values with low relative errors (and RMSE/MPRED ratios). Canonical correlation analyses between NDVI pixel values and biomass data, and subsequent best subset regressions incorporating canonical variates, was used as a methodology for avoiding effect from multicollinearity among adjacent biweekly NDVI pixel values. The results showed that the model relationships derived from this approach can be valid in predicting biomass values for up to 40 percent of the missing data. However, the missing data should be filled only with the models derived from the available data pertaining to the same time period, to better account for the specific phenological changes over the corresponding time period. Canonical correlation analyses revealed that NDVI and crop biomass are well correlated during the middle of the crop growth from mid June to end August, and using all the canonical variates from original biweekly NDVI pixel values in subsequent best subsets multiple regression analyses was needed in determining model relationships for biomass of individual crops. Overall, we found this approach suitable for filling missing biomass data at county-level, to be used in estimating residue carbon inputs or similar purpose. Since it incorporates low resolution AVHRR NDVI data and available county-level yield data as the input data for model derivation, we find this as a better approach for regional or national scale studies, than for field scale studies. This approach could be further enhanced by using MODIS NDVI data that have higher spatial, spectral, and radiometric resolution.

Acknowledgements

This research was supported by the Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture, under Agreement No. 2001-38700-11092 to the Consortium for Agricultural Soils Mitigation of Greenhouse Gases (CASMGS). We thank Sunil Kumar and Kay Dudek from Colorado State University, and anonymous reviewers.

REFERENCES

Allen, R., G. Hanuschak, and M. Craig. 2002. History of Remote Sensing for Crop Acreage in USDA's National Agricultural Statistics. Available at http://www.usda.gov/nass/nassinfo/remotehistory.htm

Campbell, J.B. 2002. Introduction to remote sensing. (pp. 450- 485). New York: Guilford Press.

Curran, P.J. 1981. Multispectral remote sensing for estimating biomass and productivity. Proc. 1st int. Symp. Brtiish Photobiology. Soc., 5-8

Csornai, G., Cs. Wirnhardt, Zs. Suba, et al. 2002. Remote sensing based crop monitoring in Hungary. Available at http://www.fomi.hu/internet/magyar/Projektek/remsensmonit.htm.

Doraiswamy, P., S. Hollinger, T. Sinclair, A. Stern, and B. Akhmedov. 2001. Evaluation of Modis Products for Applications in Regional Crop Condition and Yield Assessment. Remote Sensing For Agriculture Ecosystems And Hydrology

Doraiswamy, P.C., S. Moulin, P. W. Cook, and A. Stern. 2003. Crop yield assessment from remote sensing. Photogramm. Eng. Rem. S. 69(6), 665- 674

Doraiswamy, P.C., J.L. Hartfield, T.J. Jackson, B. Akhmedov, J. Prueger, and A. Stern. 2004. Crop condition and yield simulations using Landsat and MODIS. Remote. Sens. Environ. 92, 548-559

Doraiswamy, P. C., T. R. Sinclair, S. Hollinger, B. Akhmedov, A. Stem, and J. Prueger. (2005). Application of MODIS derived parameters for regional crop yield assessment. Remote. Sens. Environ. 92, 192-202

Hansen, P.M., and J.K. Schjoerring. 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crop using normalized difference vegetation indices and partial least squares regression. Remote. Sens. Environ. 86, 542- 553.

Hill, M.J., and G.E. Donald. 2003. Estimating spatio-temporal patterns of agricultural productivity in fragmented landscapes using AVHRR NDVI time series. Remote. Sens. Environ. 84, 367- 384

Knudby, A. 2004. An AVHRR-based model of groundnut yields in the Peanut Basin of Senegal. Int. J. Remote. Sens. 25(16), 3161-3175

Lozano-Garcia, D.F., R.N. Fernandez, and C.J. Johannsen. 1991. Assessment of regional biomass-soil relationships using vegetation indexes. IEEE T. Geosci. Remote. 29(2), 331-339

MacDonald, R.B., and F.G. Hall. 1980. Global crop forecasting. Science 208(4445), 670-679

Modenese, L., M. Camuffo, E. Belluco, A. Marani, and M. Marani. 2005. Relations between biomass and vegetation indices of halophytic plants in Venice salt marshes. Geophysical Research Abstracts, 7, 09448.

Moran, M.S., Y. Inoue, and E.M. Barnes. 1997. Opportunities and Limitations for Image-Based Remote Sensing in Precision Crop Management. Remote. Sens. Environ. 61, 319-346.

Reichert, C. and D. Cassey. 2002. A Reliable Crop Condition Assessment Program (CCAP) Incorporating NOAA AVHRR Data, a Geographic Information System and the Internet. Avialable at http://gis.esri.com/library/userconf/proc02/pap0111/p0111.htm

Senay, G.B., J.G. Lyon, A.D. Ward, and S.E. Nokes. 2000. Using high spatial resolution multispectral data to classify corn and soybean crops. Photogram. Eng. Rem. S. 66(3), 319-327

Smith, R.C.G., J. Adams, D.J. Stephens, and P.T. Hick. 1995. Forecasting wheat yield in a Mediterranean-type environment from the NOAA satellite. Aust. J. Agr. Res., 46, 113-25.

Thenkabail, P.S. 1992. LandsatThematic Mapper indices for evaluating management and growth characteristics of soybean and corn. Trans ASAE (Am Soc Agric Eng) 35, 1441-1448

Todd, S.W., R.M. Hoffer, and D.G. Milchunas. 1998. Biomass estimation on grazed and ungrazed rangelands using spectral indices. Int. J. Remote. Sens. 19, 427-438

Tucker, C.J., C. Vanpraet, E. Boerwinkel, and A. Gaston. 1983. Satellite remote sensing of total dry matter production in the Senegalese Sahel. Remote. Sens. Environ. 13, 461-474

Tucker, C.J., C.L. Vanpraet, M.J. Sharman, and G. Van Ittersum. 1985. Satellite remote sensing of total herbaceous biomass production in the Senegalese Sahel: 1980-1984. Remote. Sens. Environ. 17, 233-249

Wannebo, A., and C. Rosenzweig. 2003. Remote sensing of US cornbelt areas sensitive to the El Nino-Southern Oscillation. Int. J. Remote. Sens., 24 (10), 2055-2067

Yang, C., J.H. Everitt, J.M. Bradford, and D.E. Escobar. 2000. Mapping grain sorghum growth and yield variations using airborne multispectral digital imagery. Trans ASAE (Am Soc Agric Eng), 43(6), 1927- 1938

Yang, C., J.H. Everitt, and J.M. Bradford. 2004. Airborne hyperspectral imagery and yield monitor data for estimating grain sorghum yield variability. Trans ASAE (Am Soc Agric Eng), 47 (3), 915-924

Yang, C., J.H. Everitt, and J.M. Bradford. 2004. Airborne hyperspectral imagery and yield monitor data for estimating grain sorghum yield variability. Trans ASAE (Am Soc Agric Eng), 47 (3), 915-924

Table 1. Time periods relevant to the biweekly composite NDVI images of 1992 and 1997 that encompass the crop growth cycles.

1992	**1997**	**2002**
Apr 03- Apr 16	Mar 28- Apr 10	Apr05- Apr18
April 17- Apr 30	Apr 11- Apr 24	Apr19- May02
May 01- May 14	Apr 25- May 08	May03- May16
May 15- May 28	May 09- May 22	May17-May30
May 29- Jun 11	May 23- Jun 05	Jun1- Jun13
Jun 12- Jun 25	Jun 06- Jun 19	Jun14- Jun27
Jun 26- Jul 09	Jun 20-Jul 03	Jun28- Jul11
Jul 10- Jul 23	Jul 04- Jul 17	Jul12- Jul25
Jul 24- Aug 06	Jul 18- Jul 31	Jul26- Aug08
Aug 07- Aug 20	Aug 01- Aug 14	Aug09- Aug22
Aug 21- Sep 03	Aug 15- Aug 28	Aug23- Sep05
Sep 04- Sep 17	Aug29-Sep11	Sep06- Sep19
Sep 18- Oct 01	Sep 12- Sep 25	Sep20- Oct03
Oct 02- Oct 15	Sep 26- Oct 09	Oct04- Oct17
Oct 16- Oct 29	Oct 10- Oct 23	Oct18- Oct31
	Oct 24- Nov 06	Apr05- Apr18

Table 2. Best subsets multiple regression models between the canonical variates from the relevant NDVI pixel bands and aboveground biomass in 1992 data when 10, 20, and 40% data were missing.

40% data missing	**R^2**
CAGBM=25941.3691 - 354.713*CV1- 144.198*CV2 - 174.365*CV3 - 132.656*CV4	0.83
SAGBM= 15324.9 - 72.8236*CV1 - 27.4222*CV2 - 8.2546*CV3 - 27.0204*CV4	0.62
OAGBM=-12382.4 - 244.101*CV1 - 22.7414*CV2 - 22.0027*CV3 + 4.3636*CV4	0.92
AWBM= 25422.55 - 141.741*CV1- 15.1164*CV2 - 192.989*CV3 - 172.344*CV4	0.61
20% data missing	
CAGBM= 29598.17 - 324.767*CV1- 114.266*CV2 - 220.329*CV3 - 187.053*CV4	0.81
SAGBM=17267.1484 -66.6498*CV1-22.5345*CV2 -1.8776*CV3 - 39.3428*CV4	0.59
OAGBM= -12223.1 - 222.983*CV1+18.528*CV2 +52.6297*CV3 - 12.4709*CV4	0.92
AGBM= 28438.03 - 127.991*CV1-20.9689*CV2 - 236.678*CV3 - 190.907*CV4	0.59
10% data missing	
CAGBM= 29069.4844 -313.1467*CV1-97.3937*CV2 -221.8544*CV3-00.4418*CV4	0.81
SAGBM=17136.9297 -64.5785*CV1-16.1217*CV2+ 0.0085*CV3-38.7376*CV4	0.58
OAGBM= -12885.21-215.338*CV1+10.6914*CV2 +60.9894*CV3+9.0038*CV4	0.92
AWBM= 28509.6855 -124.4203*CV1-26.0491*CV2-251.147*CV3 - 228.7913*CV4	0.58

Table 3. Best subsets multiple regression models between the canonical variates from NDVI pixel bands and aboveground biomass in 1997 and 2002 for training data with 10% data missing.

1997	R^2
CAGBM= -16066.978- 1372.9088CV1-186.2685*CV2-124.6874*CV3+196.6339*CV4	83%
SAGBM=-4869.7495- 375.6839*CV1+171.8032*CV2+ 194.6913*CV3+ 39.2563*CV4	78%
OAGBM=-9096.083- 557.1525*CV1- 79.682*CV2+ 9.4436*CV3-191.1776*CV4	55%
AWBM= 25422.55 - 141.741*CV1- 15.1164*CV2 - 192.989*CV3 - 172.344*CV4	84%
2002	
CAGBM=-100639.0703- 824.3716*CV1- 29.4531*CV2+ 45.33*CV3+ 4.2968*CV4	85%
SAGBM= -43935.6406- 221.3674*CV1- 35.015*CV2- 22.1413*CV3-1.2646*CV4	83%
OAGBM=29324.2031- 216.8267*CV1- 23.1073*CV2- 22.9806*CV3+ 17.8367*CV4	24%
AWBM= -84203.3672 - 652.5745*CV1- 41.6448*CV2+ 30.9411*CV3-0.6962*CV4	86%

Table 4. Mean values from the observed and estimated values for training- and validation biomass data when the data from all three years are combined for developing model relationships through canonical correlation analyses. CAGBM- Corn Aboveground Biomass; SAGBM- Soybean Aboveground Biomass; OAGBM- Oats Aboveground Biomass; AWBM- Area Weighted Biomass.

Variable	Mean observed Biomass for all data (Kg ha^{-1})	Percent missing data	Mean Estimated values for training data (Kg ha^{-1})	Mean estimated biomass for validation data (Kg ha^{-1})
CAGBM	17281.18 ± 1879.95	10	16909.66 ± 32.31	16912.39 ± 262.43
		20	16915.74 ± 51.0	16893.77 ± 168.08
		40	16916.11 ± 94.42	16887.97 ± 132.35
SAGBM	6933.0 ± 522.68	10	6945.13 ± 8.28	6949.42 ± 65.46
		20	6947.58 ± 13.69	6942.95 ± 45.77
		40	6947.72 ± 24.84	6941.78 ± 33.29
OAGBM	5084.6 ± 1231.38	10	5090.45 ± 20.88	5105.31 ± 96.41
		20	5095.09 ± 35.98	5096.08 ± 85.01
		40	5090.66 ± 52.05	5086.91 ± 75.26
AWBM	12645.67 ± 1292.17	10	12407.28 ± 25.98	12405.55 ± 169.31
		20	12411.1 ± 43.22	12399.53 ± 109.56
		40	12412.41 ± 67.31	12387.13 ± 93.61

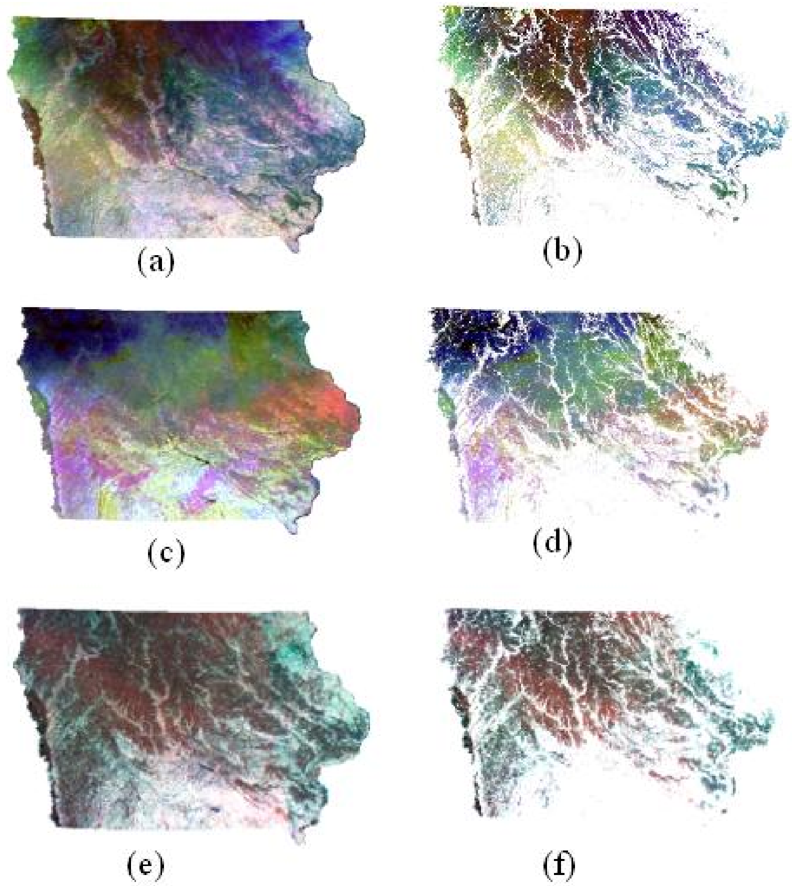


Figure 1. Images used in the data analyses: (a), (c) and (e) represent the false-color composites for NDVI images during the crop season in1992, 1997, and 2002, respectively, and (b), (d), and (f) are the same images after being masked using the relevant crop layers of the NLCD Landsat image. (Note: image layers in (a) and (b) correspond to the biweekly NDVI from April 17- 30 (blue), May 01- 14 (green), and May 15- 28 (red) in 1992, (c) and (d) correspond to the biweekly NDVI from April 11- 24 (blue), April 25- May 08 (green), and May 09- May 22 (red) in 1997, and (e) and (f) correspond to the biweekly NDVI from April 19- May 02 (blue), May 03- May 16 (green), and May 17- May 30 (red) in 2002).

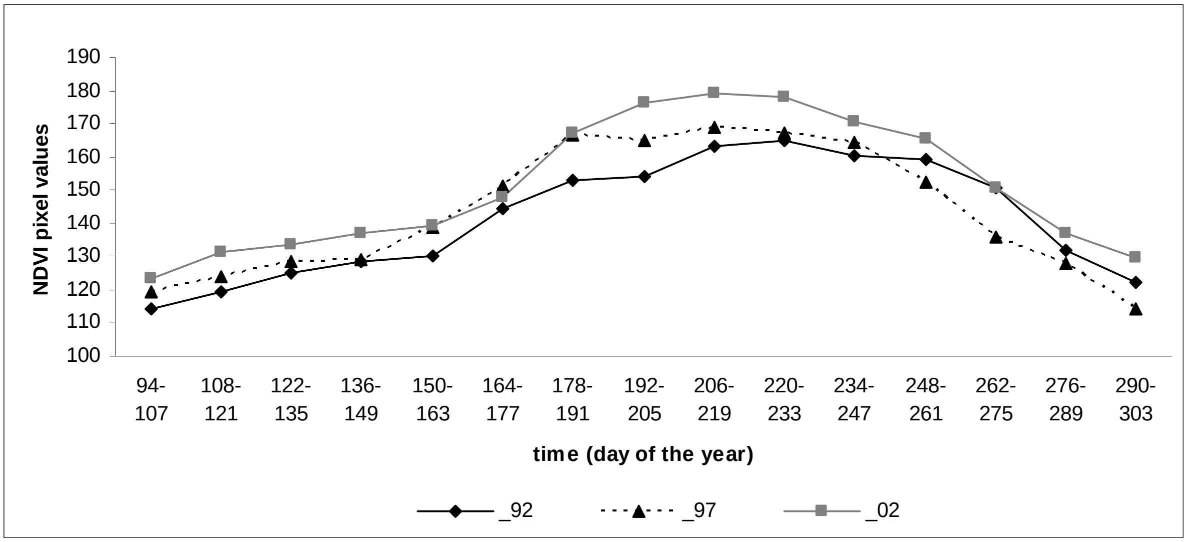


Figure 2. Variation of NDVI averaged for the whole state of Iowa with biweekly time periods in 1992 with corresponding time periods in 1997 and 2002. The same trend was observed at individual county level.

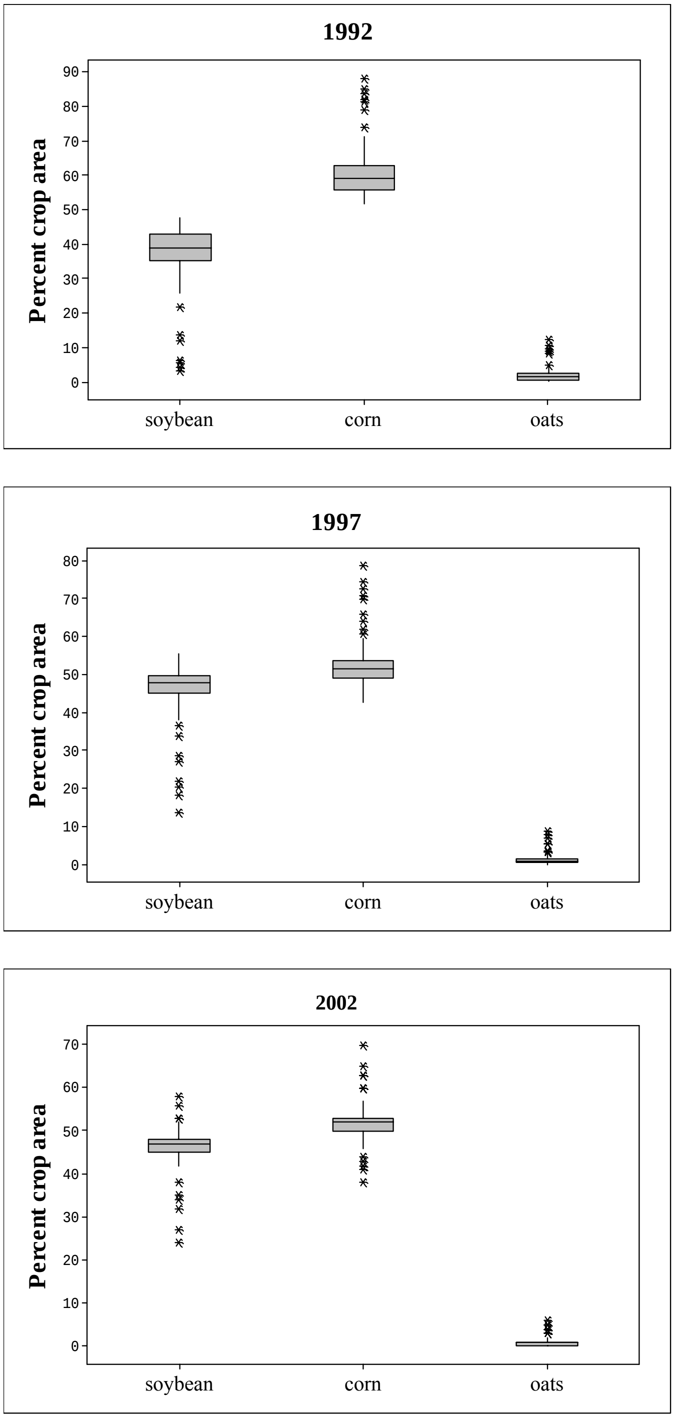


Figure 3. Box plots for crop area occupied by each crop as a percentage of total crop area in 1992,1997, and 2002.

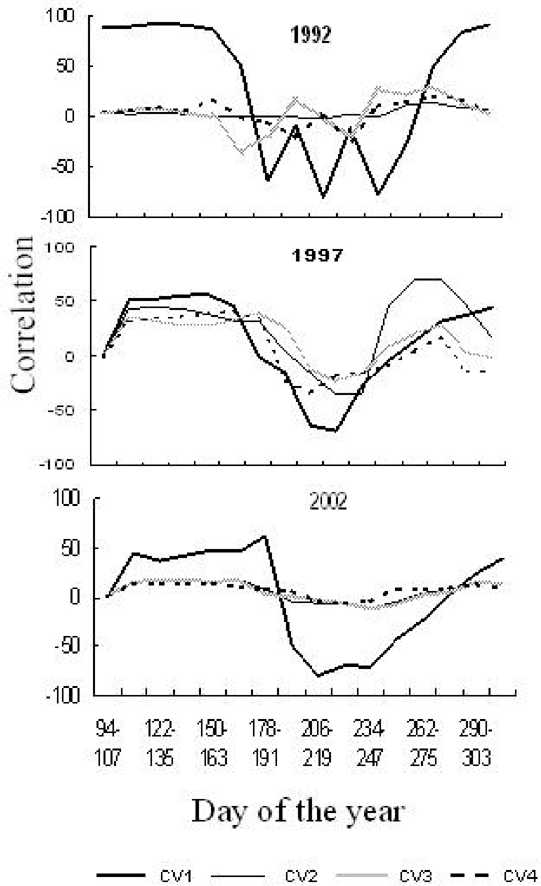


Figure 4. The correlation between the original NDVI pixel bands and canonical variates with 90% data in the training set

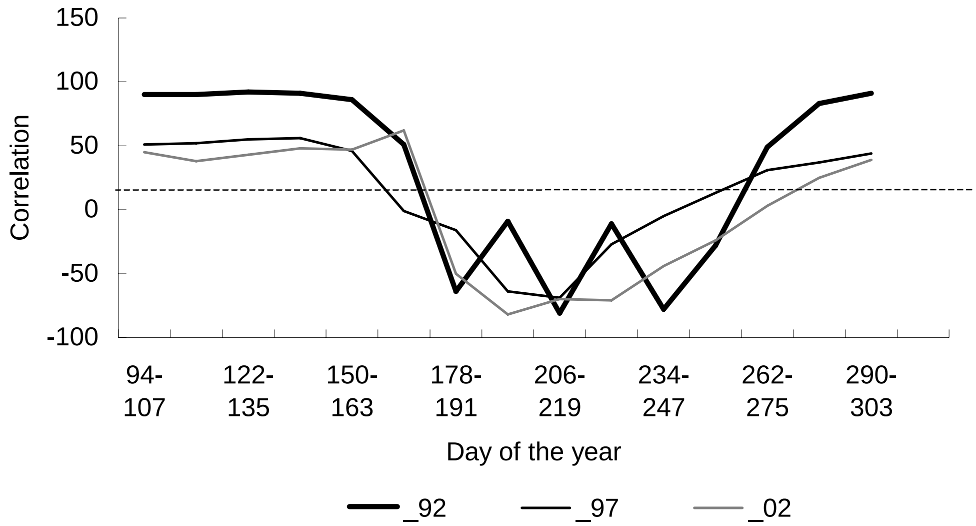


Figure 5. Correlation between original pixel bands vs first canonical variate with 90% data in the training dataset (10% missing)

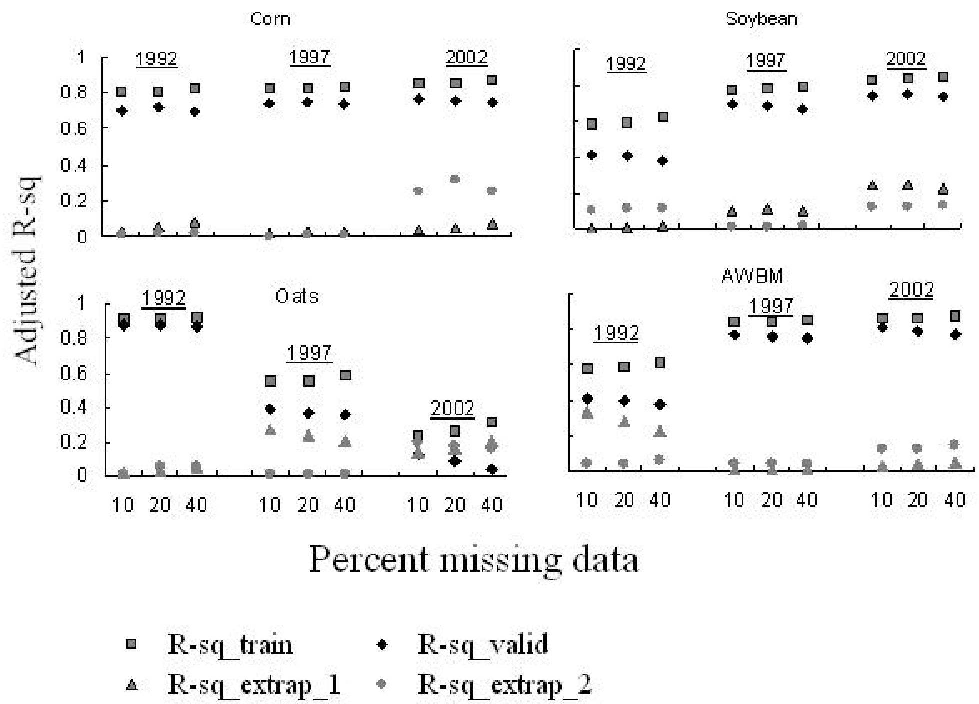


Figure 6. Adjusted R-sq values for the estimated relationships between NDVI canonical variates and corn, soybean, oats and area weighted biomass (AWBM) under single year data scenario. R-SQ_train: adjusted R-sq for the training data set from each year when 10%, 20%, and 40% of the data were missing; R-SQ_valid: adjusted R-sq when the model obtained from the training data is applied on validation data; R-SQ_extrap_1 and R-sq_extrap_2 : adjusted R-sq when the model obtained from the training data from each year is applied on the data from each of the remaining years that were not included in deriving the model relationships; extrap_1 refers to the year that is lower in value and the extrap_2 refers to the year that is higher in value (for instance, for the training data were from 1992, extrap_1 corresponds to 1997, and extrap_2 corresponds to 2002).

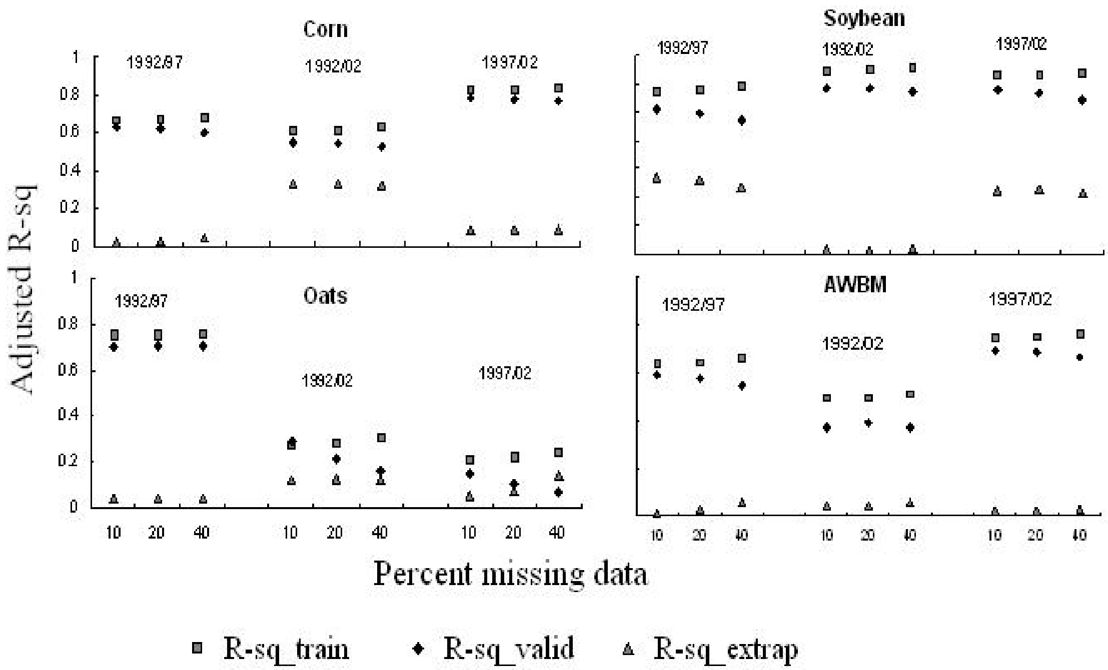


Figure 7. Adjusted R-sq values for the estimated model relationships between canonical variates from NDVI and corn, soybean, oats, and area weighted biomass (AWBM) data combined from different years (i.e. 1992/97, 1992/02 and 1997/02). R-sq_train: R-sq from the models for training data set from each two-year combination when 10%, 20%, and 40% of the data were missing; R-sq_valid: adjusted R-sq when the model obtained from the training data is applied on a data set with 10-, 20-, and 40% data missing randomly; R-sq_extrap: adjusted R-sq when the model obtained from the training data from each two-year combination is applied on the data from the remaining year that was not included in the data combination.

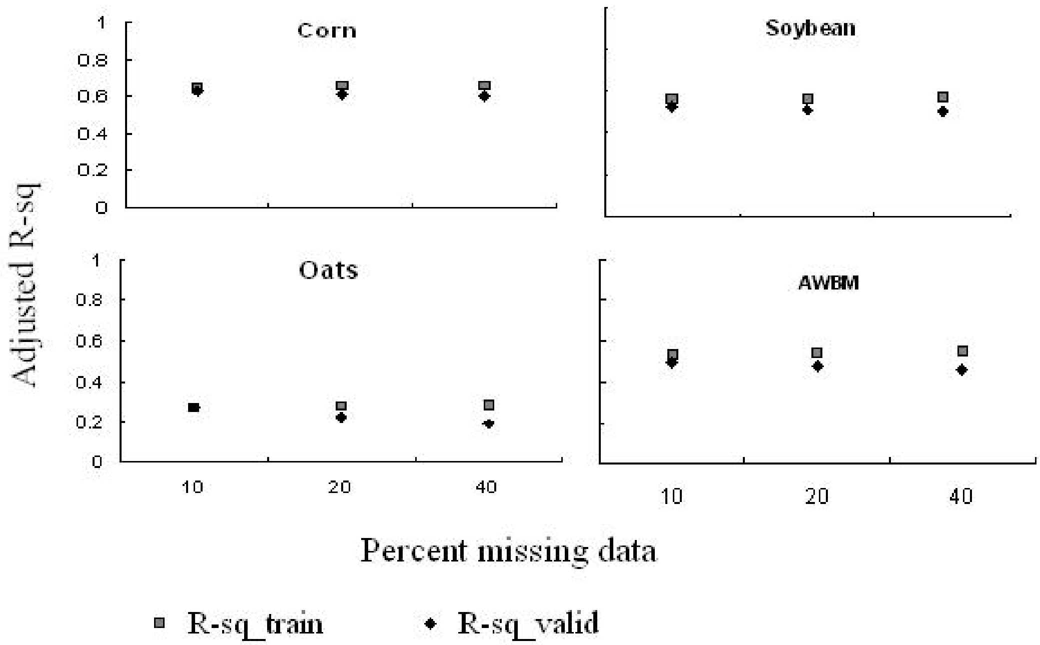


Figure 8. Adjusted R-sq values for training and validation data when the data from all the years are combined; R-sq_train: R_sq for the models derived using training data; R_sq_valid: R_square when the models from training data are applied on the validation data set

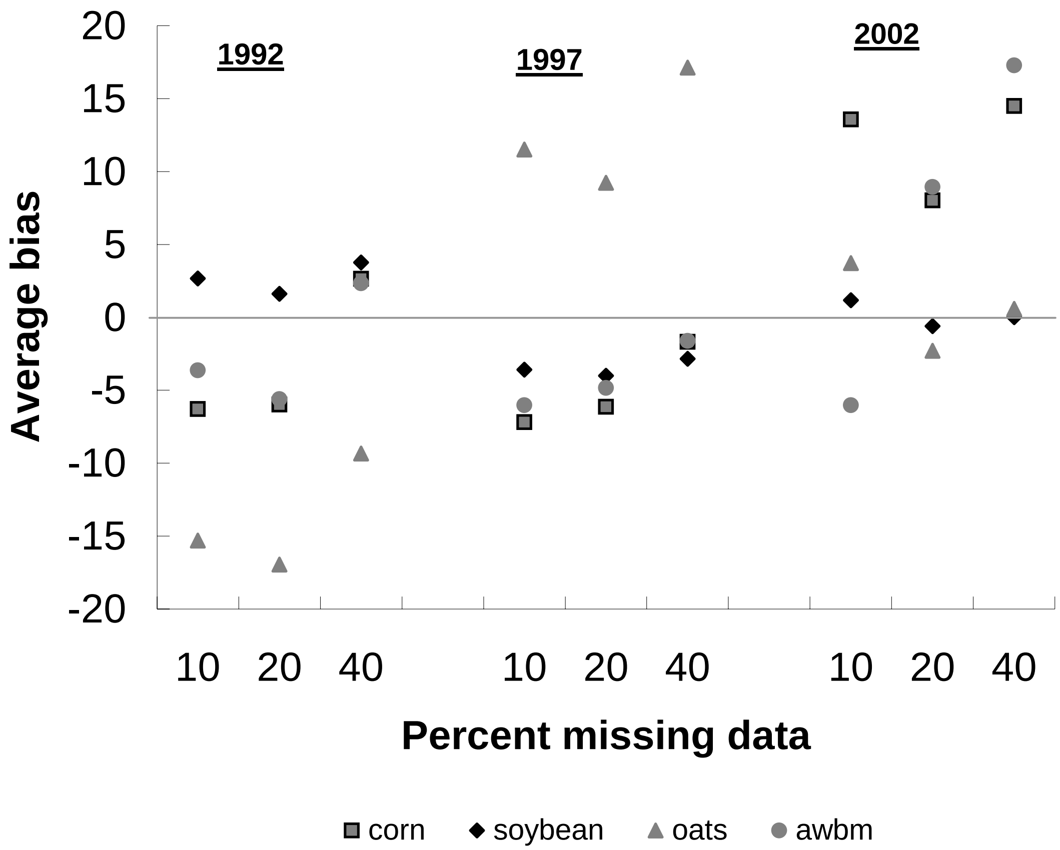


Figure 9. Average bias (i.e. average residuals) of the estimated values from 100 iterations when 10, 20, and 40% data were missing within a single year

CHAPTER 4

DERIVING COMPREHENSIVE COUNTY-LEVEL CROP YIELD AND AREA DATA FOR ESTIMATING CARBON DYNAMICS IN US CROPLAND

ABSTRACT

Carbon (C) sequestration in agricultural soils has been proposed as a way to help mitigate carbon dioxide (CO_2) buildup in the atmosphere. Crop production data, collected across the US over many decades, provides a unique resource for analyzing the impacts of carbon inputs from crop residues on spatial and temporal trends in soil C at regional and continental scales. However, significant gaps in reported crop yields and area need to be filled to accurately assess soil C changes.

We created comprehensive county-level databases for nine major crops of the US for a 16-year period, by filling the gaps in existing data reported by National Agricultural Statistics Service (NASS). We used a combination of regression analyses with data reported by NASS and the Census of Agriculture and linear mixed-effect models incorporating county-level environmental, management and economic variables pertaining to different agro-ecozones. Predicted yield and crop area data were very close to the data reported by NASS, within 10% relative errors. Using the linear mixed-effect model approach gave the best results in filling 84% of the total gaps in yields and 83% of

the gaps in crop areas of all the crops. Regression analyses with Ag Census data filled 16% of the gaps in yields and crop areas of the major crops reported by NASS.

Crop yields and crop area data, along with information on harvest index and root:shoot ratios, can be effectively used in estimating county-level crop residue C inputs for the entire country, to model C dynamics and determine the potential contribution from US agricultural soils in mitigating greenhouse gas emissions.

INTRODUCTION

Agricultural soils have been identified as a potential sink for increasing amounts of atmospheric greenhouse gases (GHGs) that lead to global warming, especially CO_2 (Cole et al., 1996; Lal et al., 1998; Paustian et al., 1997). Carbon sequestration in agricultural soils can be achieved by increasing C additions to soil and/or reducing decomposition of organic matter. Potential C sequestration in global agricultural soils through changes in land use management practices has been estimated as 600- 900 Tg yr^{-1} (Cole et al., 1996). According to Sperow et al. (2003), the potential C sequestration in US agricultural soils through such practices may be 80 to 100 Tg yr^{-1}.

Decomposition and soil C storage rates are highly dependant on the type and the amount of biomass added to the soils. Crop yields can be used to estimate aboveground residue biomass using relationships between the crop yields and aboveground biomass for different crops (i.e. harvest indices). Similarly, production of belowground residues can be estimated using equations derived from studies quantifying plant C partitioning, by developing root: shoot ratios of production for different crops (e.g. Buvanovsky and Wagner 1986; Prince et al., 2001; Campbell and Jong, 2003). Combining per ha estimates of C inputs together with cropland area provides an estimate of the CO_2-C fixed by agricultural lands and the amount of C that can be returned to the soil as residues.

Currently county-level crop yields in the US are reported by two main agricultural databases maintained by the US Department of Agriculture (USDA): the database of the National Agricultural Statistics Service (NASS), and the Census of Agriculture (Ag Census). The NASS crop yield data are produced annually using a survey done on selected farms which is extrapolated statistically to estimate county-level crop yields. Ag Census crop yield estimates are produced every 5 years during the years ending in "2" and "7"; Ag Census contacts every farmer within a county by mailing report forms to collect data and thus Ag Census data are more complete and comprehensive (Pawel and Fesco, 1988; USDA, 1998; R. Korkosh, personal communication, 2004).

The county-level data reported by NASS are used extensively for designing government policies, supporting research and other purposes. However, missing data (i.e. gaps) in certain parts of the country and reporting of yield and crop area information only at state-level for certain states, presents limitations in using the data for comprehensive analyses of the C balance of agricultural lands. Such data can be used as input to other models to analyze spatial patterns of regional scale C dynamics and/or for validation purposes, for example satellite-derived information on crop production and area extents. The aim of our study was to derive complete county-level crop yield and crop area databases, by filling the gaps in the yield and crop area data reported by NASS over the period 1982-1997, using Ag Census data and statistical models incorporating appropriate county-level environmental, management, and economic variables.

Studies done so far to estimate crop yields include models incorporating various agro-meteorological variables (e.g. Berka et al., 2003) or combinations of agro-meteorological, hydrological, management and economic variables, such as the EPIC model (e.g. Cavero et al., 2001; Tan and Shibasaki, 2003). Some studies have used combinations of ground based- and satellite-based information (Rudorff and Batista, 1990; 1991; Reynolds et al., 2000; Lobell et al., 2003; Doraiswamy et al., 2003, 2004, 2005; Yang et al., 2004; Tao et al., 2005) to estimate yields. In some of these studies, yields have been estimated through combining agro-meteorological variables, with remotely sensed information in statistical models (e.g. Rudorff and Batista, 1990, 1991; Smith, 1995; E. Lokupitiya, M. Lefsky, and K. Paustian, unpublished data) or simulation models based on remotely sensed information, such as Carnegie Ames Stanford Approach (CASA) and Global Production Efficiency Model Version 2.0 (GLO-PEM2; e.g. Tao et al., 2005). In certain other studies, remotely sensed information has been combined with crop models such as EPIC (e.g. Doraiswamy et al., 2003; Yang et al., 2004) and FAO Crop Specific Water Balance Model (CSWB; Reynolds et al., 2000), to estimate yields. These models have been mostly used in field- or regional- scale estimation of crop yields.

In studies for crop area estimation, either remotely sensed information (Bauer et al., 1978; Hixson et al., 1981; MacDonald and Hall, 1980; Csorni et al., 2002) or purely statistical models (Griffith, 1999) have been used. Remotely sensed information has also been used in improving the precision of ground-sampled data for area estimates (Gonzalez-alonso, 1997; Allen et al., 2002).

In this study, we first evaluate the existing national crop yield and area databases, their characteristics and compatibility between NASS and Ag Census. We describe the methods used for the imputation of missing data for crop yields and area and evaluate the appropriateness of these methods in imputing long-term data gaps in national crop statistics.

METHODS

Evaluation of the available national crop statistics for major crops in the US

Yields and crop area of alfalfa (*Medicago sativa* L.) hay , barley (*Hordeum vulgaris* L.), corn (*Zea mays* L.) for grain, corn for silage and green chop, oats (*Avena sativa* L.), other hay (hay other than alfalfa; i.e. tame hay, small grain hay, wild hay), sorghum (*Sorghum bicolor* L.), soybean (*Glycine max* L.), and wheat (*Triticum aestivum* L.) were considered in this study. NASS has reported crop statistics at the state level for more than 100 years and at the county level for over 70 years, for most of the country. Data reported in NASS include planted and harvested crop area, yield and total production and management practices (e.g. irrigation, summer-fallowing). The data from NASS (available at http://www.usda.gov/nass/pubs/histdata.htm) were reorganized into separate databases (using Microsoft Access 2000) of yields and crop areas for each crop with county FIPs, state, and yields (or crop areas) for each year of the 16- year period, 1982- 1997. Similarly, the data reported in Ag Census (by the Department of Commerce Bureau of the Census and USDA) were reorganized into separate databases of yields and crop areas for each crop with county FIPs, state, and yields (or crop area) for 1982, 1987, 1992, and 1997.

Compatibility of the crop yield and crop area estimates by NASS and Ag Census were evaluated by mapping (in Arc GIS version 8.2) the number of years the Ag Census and NASS have so far reported data under each crop for the period 1982-1997, and by mapping the absolute differences and percent differences (e.g. the difference in NASS crop yield as a percentage of the yield reported by Ag Census) in the crop yields and crop area reported by NASS and Ag Census for each crop. Percent differences were used to find the distribution of any "outliers" or data representing extreme differences between the NASS and Ag Census databases. Thus any discrepancies among the crop yields and crop areas reported by the existing 2 databases were evaluated taking the differences in the reported values and level of reporting (county- versus state-level) into consideration. In doing this we did a thorough study on the survey strategies used for data reporting in the two databases, and found out the counties that fall within different ranges of percent differences, and any states and counties that are not reported by either of the databases, using data queries in Microsoft Access 2000, and maps in Arg GIS (version 8.2).

Synthesis of comprehensive crop yield and area databases

Following the preliminary analyses of discrepancies between the data reported by NASS and Ag Census, a step-wise procedure was used to fill gaps and derive complete county-level databases of crop yields and areas, and estimate residue C inputs (Figure 1). Because NASS data is collected each year, it was chosen as the foundation database and data from Ag Census served in adjusting and filling missing data, as described below. The synthetic database produced is referred to as 'NASSus'.

Filling gaps in crop yields and areas reported by NASS where Ag Census data were available

Leave-one-out and leave-k-out procedures were used to find out the most suitable statistical method for imputing NASS data (Lokupitiya et al., 2006). Regression analyses between NASS and Ag Census crop yield data and multiple imputation technique in SAS (version 8.2) were found to be the best methods. Spatial statistical analyses such as the Kernel regression and kriging were found as the least suitable methods to be considered. Therefore regression analyses between NASS and Ag Census yield data were performed to replace extreme data or 'outliers' in NASS yields, and fill in the gaps. In order to detect outliers, a criterion based on the lower quartile (Q1; 25%th percentile), upper quartile (Q2; 75%th percentile), and interquartile distance (IQ) was used; any value lower than (Q1 - 3*IQ) and any value greater than (Q2 + 3*IQ) were removed. Regression analyses performed separately between Ag Census and NASS yields for each crop and each of the years 1982,1987,1992 and 1997, were used to replace outliers and fill the gaps in NASS yield data. Crop areas reported in AgCensus that were missing for corresponding years and counties in the NASS data were added to the NASSus database without adjustment.

Using environmental variables to fill remaining gaps in crop yields

Linear mixed-effect models (Littell et al. 1996) were used for filling remaining gaps in the yields in NASSus data, utilizing environmental and management factors such as irrigation, to predict yields.

County-level weather and irrigation data were chosen as the covariates, with yields as the dependent variable, in the mixed models. Mean monthly summer temperature (MST), annual precipitation (P), precipitation/potential evapotranspiration ratio (P/PET), irrigated/total crop area ratio (ITA) were the variables that represented fixed effects in the models. County FIPs code was the only variable representing any random effects; the random effect represents variation among counties due to factors other than the fixed effects.

Preparation of environmental data for linear mixed effect model runs.

Annual potential evapotranspiration was calculated for 1982-1997, using the method by Thornthwaite (1948), using weather data from the gridded PRISM (Daly et al., 1994) dataset (http://www.ics.orst.edu/prism) for the conterminous US (Figure 3).

The USA has a total of 20 Land Resource Regions (LRRs), which delimit contiguous areas with similar geographical, climate and land use conditions (Figure 3). County-averaged yield data from different major crops and county weather variables were grouped by LRRs over the entire time series and sorted by year and county. In addition, ITA for a county was used as a conditional variable to determine whether moisture-limitations would be included in the model. If a county had an ITA greater than 0.5 (i.e. majority of cropland is irrigated), then P and P/PET were not used in the prediction of yield.

NASS typically reports separate explicit categories for irrigated and non-irrigated cropland where both are present as significant land area fractions. In some arid counties, the entire area for a particular crop is likely to be irrigated, hence NASS may only report total area for that crop. Similarly, in many eastern US counties, where irrigation is minimal, NASS may only report total crop area. In such instances, crops were designated as primarily irrigated or non-irrrigated based on location, type of crop and long-term climate averages.

Selection criteria for the best-linear mixed effect models and quality control measures for the predicted crop yields.

Eq. [1] gives the basic model used in the linear mixed-effect models for crop yields:

$$Y = X.\beta + Z.u + \varepsilon \qquad [1]$$

Where:

Y= yield

X= design matrix of covariates (or fixed effects, including the intercept, P, P/PET, ITA, MST, and interactions between these variables)

β= vector of coefficients corresponding to the fixed effects

Z= design matrix of '0's and '1's for the random effects; for each fips, with '1' in column j indicating that observation is from county j

u= vector of coefficients corresponding to the random effects for each fips

ε= error vector (which may be autocorrelated with time).

The linear mixed-effect models were run separately for each LRR with autoregressive order 1 (AR1) covariance structure, with time as repeated measures, to fill in the remaining yields in counties. Thus for each crop, several models with different combinations of the above covariates and the crop yield as the response variable, were run on each LRR in SAS (version 9.1). The model with the lowest Akaike Information Criterion (AIC) was chosen as the best model for each LRR; the chosen model was used in filling the county-level yield gaps in different LRRs under different crops. Altogether 10 models were attempted on each LRR per crop; if convergence criteria were not met, or the final Hessian was not positive definite, either AR(1) or random effects had to be dropped to have the convergence criteria met.

The predicted yields from the linear mixed-effect models were compared against a default method that used the mean of the observed values within a county as the predicted values. A few counties had only one observed (reported) yield or crop area value across the entire time series in an LRR under certain crops, while the other counties had several observations with relatively large variation in crop area or yields over the time series. Hence, no single standard statistical method could be adopted to screen for outliers. Therefore the following procedure was adopted for identifying outliers. First, where predicted values fell outside plus/minus three times the mean of the observed values, they were given an initial designation as potential outliers. Where potential outliers (in predicted values) also fell outside the range of the observed values for the particular counties, they were given a final designation as outliers, and were replaced with the mean of the observed values (i.e. the default option). Final data screening was done considering the occurrence of the crop at county-level, i.e., imputed data were removed from counties that had no reported occurrence of the crop.

Using environmental and economic variables in gap-filling in crop area data

The mixed models for crop area data included economic and weather variables from the previous year as fixed effect predictor variables: precipitation, crop price, fertilizer cost (unit cost of anhydrous ammonia), and diesel cost. County area and cropland area set aside in the Conservation Reserve Program (CRP) for the current year were also used as fixed effect variables. County FIPs served as the only variable for random effects.

Data preparation. Crop price data for the previous year were obtained from NASS (available at http://www.usda.gov/nass/pubs/histdata.htm). Available state-level prices of the crops were extracted for the period 1981- 1996. Where state-level price data were not available, the mean price of the multi-state crop production region was used. Since no price data were available for corn for silage in NASS, corn for silage price per ton was estimated by multiplying the per bushel price of corn grain by 9 (Barkley, 2002). All crop prices were adjusted for inflation using the Gross Domestic Product-Implicit Price Deflator (GDP-IPD; S. R. Koontz, personal communication, 2005).

Fertilizer and diesel price data of the previous year were extracted from the USDA Agricultural Prices Annual Summary reports for the period 1981- 1996 (available at http://usda.mannlib.cornell.edu/reports/nassr/price/), and adjusted for inflation using the GDP-IPD. Previous year's precipitation was extracted from the PRISM data grid described above. Cropland area enrolled in CRP since the beginning of the program in 1986 to 1997 was obtained from the ERS (2005) and FSA (2005). County-level crop area data and the data from all the predictor variables over the entire time series for the period 1982-1997 were then compiled and organized in a similar structure as detailed above for yield data.

Selection criteria for the best linear mixed-effect models and quality control measures for the predicted crop area. Linear mixed-effect model analyses were performed with auto regressive order 1 (AR1) covariance structure with crop area as the dependent variable, county FIPs as the random effect, and different combinations of the following variables as the fixed effects: previous year's precipitation, previous year's crop price, previous year's fertilizer price, previous year's diesel price, CRP crop area, and county area. The analyses were performed under the following model options:

a. Crop area as the response (y) variable, and diesel price, fertilizer price, crop price, CRP crop area, and county area as regression (x) variables

b. Crop area/ county area as the y variable, and the rest of the variables as x

c. Crop area/ county area as the y variable, with rest of the variables standardized (by dividing the value of each variable by the std. deviation of each variable), as x variables

d. Taking all the variables standardized including the y variable in c above

e. Log-transformed crop area as the y variable, log-transformed county-area as an x variable, and all the other x variables standardized (by dividing by the std. deviation); predicted y values were exponentiated back to get the predicted crop area values from the models.

Linear mixed effect models were run on each different crop at LRR level, considering the entire time series. The model with the lowest AIC was selected as the best model for filling the gaps in crop area data.

Detection of any outliers and quality control of the predicted crop area were carried out in the same way as for the yields. The total of the crop area aggregated at state-level was compared against the state-level cropland crop area based on the information collected by National Resources Inventory (NRI), as an additional quality control measure.

Estimation of the carbon inputs from crop residues

Under improved agricultural management practices in the recent decades, about 60% of the crop (stover and roots) is left on- and below-ground, with the remainder being removed by harvest; this becomes added to soils as crop residue. Dry biomass of the aboveground crop residue can be estimated by subtracting the dry grain biomass (i.e. removed in harvest), from the total aboveground dry biomass. Crop yields were converted to biomass of yield dry matter by correcting for moisture content. Then the biomass of aboveground residue dry matter for each crop was estimated from crop-specific harvest indicies (Williams and Paustian, submitted). Then the biomass of the total belowground residue dry matter was derived using crop-specific shoot: root ratios corrected for rhizodeposition based on analyses of published biomass partitioning studies (Williams and Paustian, submitted). County-level totals do not account for removal of aboveground residue that is used for bedding, fuel or other purposes (e.g. grazing).

RESULTS

Preliminary analysis of data

Except for a few outlying values, Ag Census and NASS crop data were very similar, although some significant deviations were found. When all the crops were considered together (Figures 4 and 5), 80-99% of counties reported crop yields within 30% or less difference between NASS and Ag Census and 15-70% of counties had differences of 5% or less in reported crop yields (Table 1).

The number of outliers was relatively small (less than 1- 5%) for each crop within each year. For instance, Figure 4 shows the deviation of NASS wheat yields from those reported by Ag Census for 1997. Since NASS collects information from a sample of farmers within a county and extrapolates that information to the entire county, it may create occasional anomalous values that were obvious during the years when both NASS and Ag Census data are reported. It was found that the extreme differences between Ag Census and NASS for the same crop were very low even when the absolute differences between the yields were compared (Figure 5). According to Figure 5, corn yields reported by NASS and Ag Census for a majority of the counties had a difference of only about 10 bushels per acre (i.e. 0.6 tonnes per hectare).

Existing gaps in the crop yields reported by NASS. Since NASS reports crop yield data annually, while Ag Census reports every 5 years, we used NASS as the underlying database for developing complete and comprehensive databases for the major crops in the US. Although NASS reports annual data, NASS does not report crop yields in certain states and counties that are known to contain particular crops, particularly hay crops. For instance, NASS does not report alfalfa crop yields in the counties of 21 states in the conterminous US, while Ag Census reports data for those states (Figure 6, Table 2). Similarly, Ag Census has not reported county-level yields and/or crop area of barley, corn for grain, corn for silage, oats, other hay, sorghum, soy, and wheat for some counties where NASS has reported data (Table 2). Even in the years that both Ag Census and NASS have reported data, we found that still there are missing values in certain counties.

Synthesis of comprehensive databases of crop yields and crop area

Filling initial gaps in NASS data using Ag Census

The regression analyses between NASS and Ag Census yields for 1982, 1987, 1992, and 1997, gave high R^2 values (e.g. Table 3). For alfalfa hay, barley, corn for grain, oats, sorghum, soybean and wheat, close to 90% of the variation in the data reported by NASS were explained by Ag Census data. However, compared to other crops, corn for silage and other hay showed a weaker relationship between NASS and Ag Census (lower R^2 values). For all the crops, the slope of the regression was very close to 1 indicating a good relationship between the NASS and Ag Census data. This was further evident since the intercept was close to zero in the majority of the crops (except for corn for silage and green chop, oats and sorghum, which could be due to the under representation of the data in either of the databases under counties in certain states). These regression models were used to replace the outliers (or the data that are extremely different compared to Ag Census) in NASS data, and fill in the gaps during the above four years when both NASS and Ag Census have reported crop yield and area. Using this process, 16% of the gaps in NASS yields and crop areas were filled.

Using environmental variables as covariates in linear mixed-effect models for filling the remaining gaps in the county-level crop yields

The best linear mixed-effect models with different combinations of environmental variables for predicting missing yields of the crops in the counties of different LRRs are summarized in Table 4. The best predictions (the lowest AIC) for crop yields in the western LRRs (i.e. A, B, C, D, and E) were obtained mostly with mean monthly summer temperature (MST); for LRR E, both MST and ITA seemed to contribute equally in predicting the yields in different crops (Table 4). The best models for the rest of the country tended to be those that included the fixed effect variables P, MST, ITA, and P/PET. For alfalfa hay in LRRs C, O, U, other hay in LRRs C, and U, and corn for silage in LRR O, convergence criteria were met in the models, but the final Hessian was not positive definite. This was found when the observed data for the crop in the particular LRR was very few and sparse. In such cases, convergence criteria were satisfactorily met, when the models were run with all the fixed effect variables and random variable, but no autoregression, or when the time correlation was removed from the model. Most of the time, the AIC for the best model and the AICs for the rest of the models differed by more than 2 (i.e. the value normally used in as the best-fit criterion under AIC).

Quality control/ quality assurance of the final yields. The number of 'outliers' in the predicted data was extremely low (Table 5) and only three crops (i.e. alfalfa hay, barley and corn) contained them; these very few outliers were replaced with the mean of the observed values for the particular counties (or FIPs). All the predicted yields had relative errors less than 1. Over the 16-year period the majority of the missing data in NASS were for alfalfa hay (21% of the total gaps), other hay (25%), and corn for silage (11%); the percentage missing data in the rest of the crops ranged between 5-10% of the total gaps in the yields reported by NASS. Eighty-four percent of the total gaps in NASS (and 99.998% of the gaps in the NASS and Ag Census combined data) were filled using the mixed models. Only 0.02% of the gaps in NASS were filled with the default option (i.e. where the imputed values were designated as outliers

Figure 7 shows the alfalfa yields from initial NASS database, NASS and Ag Census combined, and the completed alfalfa yields, with the imputed values (NASSus database). About 21% of the gaps in alfalfa yields reported by NASS were filled using the Ag Census information, and about 78% of the gaps were filled using the linear mixed effect models; the remaining 1% of the gaps were filled with the default option. Figure 8 illustrates the yield trend over time with inclusion of the predicted values for corn in two counties, depicting the compatibility of the predicted (for the missing years) and the observed yields.

Using environmental and economic variables in filling the remaining gaps in crop area data

Out of the mixed model options attempted, the final option (i.e. log-transformed crop area as y variable and all the other x variables standardized with log-transformed county-area as an x variable) was found to give the best model results, with the lowest relative errors and convergence criteria well met. The best models for filling the crop area gaps in counties of each LRR were chosen based on the lowest AIC values (e.g. Table 6), and those models are summarized in Table 7. No remarkable trends could be observed in the response of the crop area to the fixed effect variables used in the best models, but the crop area in the majority of the LRRs under each crop seem to have the best models with the combinations of the three variables, diesel price, fertilizer price, and crop price of the previous year. When the importance of each single fixed effect variable is concerned, the diesel price of the previous year seems to be the most important predictor, being the sole predictor (other than the county area) for at least one LRR in a majority (5 out of 9) of the crops. CRP area, either alone, or in combination with the other variables, seems to be more powerful in predicting the crop area when the area under CRP is high. This was particularly obvious for LRR G, having close to 30% of the cropland in CRP during most of the time in the 16-year period, and more than 95% counties in the LRR containing land in CRP. However, other LRRS (e.g. B, F, and H) that had even a greater percentage of the counties with CRP, did not have CRP area as a predictor in most of the best models for any of the crops. The reason could be that CRP area was about 20% or less of the total cropland area of these LRRs.

Quality control/ quality assurance of the final crop area. The relative errors produced from the mixed models were much lower than the default option (i.e. if the gaps were simply filled with the mean of the observed crop area; Table 8). Occasionally, using the best mixed model yielded outliers, especially if only one or very few observed data were present over the entire time series; this trend was obvious for certain FIPs in LRR 'U' for alfalfa hay, LRR 'O' for corn for silage, and LRR 'K' for sorghum. However, the number of outliers in each crop was less than 1% of the total predicted values (Table 5).

The gaps in the few counties that had outliers were filled using the default method, by using the mean of the observed values. Overall, 83% of the total gaps in the crop areas reported by NASS (and 98.5% of the remaining gaps in NASSus database) were filled with the linear mixed effect model approach and only 1% of crop area gaps in NASS were filled using the default method (Figure 9). Table 9 provides a summary of the gaps filled in the crop area of each crop.

In comparing the final crop area with the cropland area reported by NRI points at state level, we found that out of the 48 states in the US, 13 states had total crop area (all the crops combined together) that slightly exceeded the NRI cropland area. However, the final total cropland area aggregated at state level was very close to the state-level cropland area according to the NRI, with an R^2 exceeding 99%. This was observed for

1982, 1992, and 1997, during which years NRI had also reported data (Figure 10). As evidence of the overall goodness-of-fit of the linear mixed effect model approach, the trend in the total cropland area after inclusion of the predicted values for county-level missing crop areas, followed a trend very similar to the observed (reported) crop area trend over time (Figure 11).

Estimation of the crop residue carbon inputs

Intensity of C inputs estimated using the crop yield and crop area data largely reflects the geographical distribution of precipitation/potential evapotranspiration. The highest C inputs on cropland (county-weighted; kg ha^{-1}) occur in areas with high precipitation or irrigation. Thus the Corn Belt region had the highest county-weighted carbon inputs and was closely followed by the Central Valley region of California where crops are grown essentially under irrigation (Figure 12). The geographic pattern of county-weighted C inputs was similar during dry years and wet years, but on the average, during a wet year the amount of inputs added per ha was about 20% higher (Figure 12).

By filling gaps in missing data, total estimated C inputs from the whole US over the 16-year period were increased by 6%. While significant at the national-level, the consequence of gap filling for estimating C inputs was even greater for certain crops and local areas. Gap filling accounted for about 45% of the total carbon inputs from alfalfa hay, 30% from other hay, 10% from corn for silage and green chop, and less than 5% from each of the other crops. In about 20% of the counties (585 out of the 3044 counties where crops were grown), the gap filling alone added more than 50% of the total carbon inputs (tonnes) that occurred over the 16-year period. Gap-filling added 100% of the carbon inputs over the 16-year period in about 1.7% of the counties (51 out of 3044) that did not have yields or crop area reported by NASS, less than 10% of the carbon inputs in about 55% of the counties (1692 out of 3044), and 10-50% of the carbon inputs in about 26% of the counties (803 out of 3044).

DISCUSSION

In studying the C balance in agricultural soils, both CO_2 output from the decomposition process and C inputs from crop residues are important. Carbon inputs and net primary production can be estimated from observed crop yields and crop-specific allometric functions. One drawback in using the available US national crop statistics is missing yields and crop area information at county level. Thus, more comprehensive crop yield and area data can be used in a variety of analyses of regional C balance studies.

Out of the 3111 counties in the conterminous US, we found that only 67 counties had no crops reported. In initial evaluation of the existing discrepancies between the two main crop statistical databases, NASS and Ag Census, we found that the most of the data are very close and comparable (Table 1). However, the difference in the survey methods yielded occasional, extremely different values in NASS, compared to Ag Census. Certain differences in the reported county-level crop statistics could also be attributed to the way NASS and Ag Census reports the farm (or farmer) information; if a farmland extends over several counties, NASS surveys use the location of the headquarters as the location of the farmland. However, for Ag Census the county in which the operator earns most of his income is reported. This discrepancy becomes visible in those counties with little agriculture (R. Korkosh, personal communication, 2004). For certain crops, NASS does not report county-level data for certain states where the crop is present, and overall, NASS has a significant number of missing data at county level. NASS reporting is also restricted by Title 13 of the US code that stipulates that data are not to be published if it would disclose the operations of a single farm within a county, but it is permitted to release the 'number of farms' information observed for a county (Griffith, 1999). This is another reason for county-level missing data in NASS. Initially, we filled the gaps in NASS using the data from Ag Census; however, less than 20% of the total gaps in NASS data could be filled using the Ag Census information.

Using linear mixed-effect models with environmental, management and economic variables to impute missing data yielded lower relative errors compared to a default method of simply using the mean of the observed values for a county. Overall, the linear mixed effect model approach filled more than 80% of the total gaps in NASS data. In a few instances, where county data were very sparse, models needed to be modified by dropping either autoregression or random effects to meet the convergence criteria. Less than 1% of the missing or imputed values were filled using the county-level time series mean as a default method. Availability of a very low number of observations has been problematic in certain other studies as well. According to Tao et al. (2005), the CASA model overestimated yields in areas with few observations, while it performed better in areas with dense crop coverage.

Using the AR(1) covariance structure yielded predicted values that showed a good time correlation/trend in the predicted yields (Figure 8), and overall the mixed model approach seemed to perform well; only a handful of 'outliers' had to be replaced with the means of the observed values. Compared to other complex simulation models, our approach was straightforward, with fewer parameters. It also incorporated essential environmental and economic factors, in addition to spatial and temporal autocorrelation effects.

Incorporation of the county area was essential in the models for predicting the crop areas. In our study we incorporated log-transformed county area as a predictor variable, while Griffith (1999) had considered the density of area (by dividing by the county area) to incorporate any effect from the size of a county. Griffith (1999) had taken the relationship between an agricultural commodity and the number of farms producing that commodity, along with the spatial autocorrelation in the statistical models used in small area estimation in Michigan and Tennessee. With the model options having log-transformed county area, we got better relative errors compared to having the area density as the dependent variable.

Since no other ground-based database is available (except for NASS and Ag Census) to compare the final results at county-level, we aggregated the predicted crop area at state-level, and compared those with the state-level cropland area based on the NRI. The final crop area for different crops at state-level was very close, although there were slight differences in certain crops at state-level. This could be mostly due to differences in the reporting by NRI and NASS or Ag Census, especially in terms of the differences of small grain crops (due to differences in sampling time), and differences in reporting hay crop categories. The total cropland from all the major crops according to our final crop area (after filling all gaps) were very close to the cropland area from NRI estimates (R^2 =0.99).

Carbon inputs showed a geographical variation in the level of addition, being highest in the areas with high precipitation and in arid/semiarid areas with irrigated cropland. The same trend could be found during both dry years and wet years, although the amount of C was higher during the wet years (Figure 12). The filling of gaps made it possible to estimate C inputs for a significant number of counties where the data were missing during all or certain number of years during the period concerned. Gap filling added a significant proportion of C inputs both county wise and crop wise. For instance, gap filling added more than 50% of the total C inputs in about 20% of the 3044 counties where the crops were grown; 10-50% of the total C inputs were added in about 26% of the counties. Gap filling in yields and crop areas added close to half of the total C inputs from alfalfa hay, and close to one third of the C inputs from other hay.

CONCLUSION

A majority of the missing data in yields and crop areas reported by NASS during the 16-year period were for alfalfa hay (21% of the total gaps), other hay (25%), and corn for silage (11%); missing data in the remaining crops were less than 10% of the total gaps. The effect of gap-filling was greater for certain counties and certain crops, especially for hay crops in certain states where NASS does not report county-level data, and certain counties with small crop areas. The use of environmental, economic and management variables in linear mixed models while taking the spatial and temporal correlation into account, allowed filling the largest proportion of the data gaps; regression analyses with Ag Census also helped fill a significant portion of the gaps during 1982, 1987, 1992 and 1997. Overall, the methodological approaches we used in this study enabled us reach the goal of estimating the total county-level C inputs in residues from major crops in the US cropland, while creating complete county-level yield and acreage datasets for those crops. The next step is to use the information from the current study in modeling the C dynamics in the US agricultural soils, and its potential to contribute towards CO_2 mitigation.

Acknowledgements

We are grateful for the assistance and support given by Mark Easter, John Norman, S.R. Koontz of Colorado State University, Roy Korkosh of NASS/ USDA, Hari Eswaran of NRCS/ USDA, Paul Hesse of National Energy Information Center (NEIC), Sian Mooney of University of Wyoming, Shawn Bucholtz and Alexander Barbarika of FSA/ USDA, Wen Huang, Tim Parker and Daniel Hellerstein of ERS/ USDA. Support for this work was provided by the Consortium for Agricultural Soil Mitigation of Greenhouse Gases (CASMGS) funded by Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture, under Agreement No. 2001-38700-1109.

REFERENCES

Allen, R., G. Hanuschak, and M. Craig. 2002. History of Remote Sensing for Crop area in USDA's National Agricultural Statistics. [On line]. Available at http://www.usda.gov/nass/nassinfo/remotehistory.htm (accessed 25 Janurary, 2006; verified 20 February, 2006).

Barkley, M. 2002. Pricing corn silage. Penn State Cooperative Extention Office in Bedford County. Bedford, Pennsylvania.

Bauer, M.E., M.M. Hixson, B.J. Davis B.J., and J.B. Etheridge. 1978. Area estimation of crops by digital analysis of landsat data. Photogramm. Eng. Rem. S. 44 (8): 1033-1043.

Berka, L. M. S., B.F.T. Rudorff, and Y.E. Shimabukuro. 2003. Soybean yield estimation by an agrometeorological model in a GIS. Sci. agric. (Piracicaba, Braz.) 60 (3): 433-440.

Burke, I.C., E.T. Elliott, and C.V. Cole. 1995. Influence of macroclimate, landscape position, and management on soil organic matter in agroecosystems. Ecol. Appl. 5(1): 124- 131.

Buyanovsky, G.A., and G.H. Wagner. 1986. Post-harvest residue input to cropland. Plant Soil 93: 57- 65.

Campbell, C.A., and R. de Jong. 2001. Root-to-straw ratios- influence of moisture and rate of N fertilizer. Can. J. of Soil Sci. 81(1): 39- 43.

Cavero, J., E. Playan, N. Zapata, and J.M. Faci. 2001. Simulation of maize grain yield variability within a surface-irrigated field. Agron. J. 93 (4): 773-782.

Cole, V., C. Cerri, K. Minami, A. Mosier, N. Rosenberg, D. Sauerbeck, J. Dumanski, J. Duxbury, J. Freney, R. Gupta, O. Heinemeyer, T. Kolchugina, J. Lee, K. Paustian, D. Powlson, N. Sampson, H. Tiessen, M. Van Noordwijk and Q. Zhao. 1996. Agricultural options for mitigation of greenhouse gas emissions. p 745-771. *In* R.T. Watson, M.C. Zinyowera, R.H. Moss, and D.J. Dokken, (ed.) Climate Change 1995. Impacts, adaptations and mitigation of climate change: Scientific-technical analyses. IPCC Working Group II. Cambridge University Press, Cambridge.

Csornai, G., Cs. Wirnhardt, Zs. Suba, P. Somogyi, G. Nádor, dr.L. Martinovich, L. Tikász, A. Kocsis, B. Tarcsai, and Gy. Zelei. 2002. Remote sensing based crop monitoring in Hungary. Available at

http://www.fomi.hu/internet/magyar/Projektek/remsensmonit.htm (accessed 10 Janurary, 2005; verified 20 February, 2006).
Daly, C., R.P. Neilson, and D.L. Phillips. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. Journal of Applied Meteorology 33: 140- 158.
Doraiswamy, P.C., S. Moulin, P.W. Cook, and A. Stern. 2003. Crop yield assessment from remote sensing. Photogramm. Eng. Rem. S. 69(6): 665- 674.
Doraiswamy, P.C., J.L. Hartfield, T.J. Jackson, B. Akhmedov, J. Prueger, and A. Stern. 2004. Crop condition and yield simulations using Landsat and MODIS. Remote Sens. Environ. 92: 548-559.
Doraiswamy, P. C., T.R. Sinclair, S. Hollinger, B. Akhmedov, A. Stem, and J. Prueger. 2005. Application of MODIS derived parameters for regional crop yield assessment. Remote Sens Environ 92: 192-202.
ERS, 2005. Economic Research Service. USDA. Conservation Reserve Program Database. Available at http://usda.mannlib.cornell.edu/ (accessed 15 May, 2005; verified 20 February, 2006).
FSA, 2005. Farm Service Agency. USDA. Washington, D.C.
Gonzalez-alonso, F., J.M. Cuevas, R. Arbiol, X. Baulies. 1997. Remote sensing and agricultural statistics: crop area estimation in north-eastern Spain through diachronic Landsat TM and ground sample data. Int. J. Remote Sens. 18(2): 467 – 470.
Griffith, D.A., 1999. A methodology for small area estimation with special reference to one-number agricultural census and confidentiality: results for selectedmajor crops and states. RD Research Report No. RD-99-04. National Agricultural Statistics Service. United States Department of Agriculture.
Hixson, M.M., B.J. Davis, and M.E. Bauer. 1981. Sampling landsat classifications for crop area estimation. Photogramm. Eng. Rem. S. 47: 1343-1348.
Lal, R., J. Kimble, R.F. Follet, and C.V. Cole. 1998. The potential for U.S. cropland to sequester carbon and mitigate greenhouse effect. Ann Arbor Press, Chelsea, Michigan.
Littell, R.C., G.A. Milliken, W.W. Stroup, and R.D. Wolfinger. 1996. SAS system for mixed models. SAS Institute, Inc., North Carolina.
Lobell, D.B., G.P. Asner, J.I. Ortiz-Monasterio, and T.L. Benning. 2003. Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. Agric. Eocsyst. Environ. 94: 205- 220.
Lokupitiya, R.S., E. Lokupitiya, K. Paustian. 2006. Comparison of missing value imputation methods for crop yield data. Environmetrics. *In press*
MacDonald, R.B. and F.G. Hall. 1980. Global crop forecasting. Science 208(4445): 670- 679
Paustian, K., E.T. Elliott, H.P. Collins, C.V. Cole, and E.A. Paul. 1995. Use of a network of long-term experiments for analysis of soil carbon dynamics and global change: the North American model. Aust. J. Exp. Agr. 35: 929-39
Paustian, K., H.P. Collins, and E.A. Paul. 1997. Management controls on soil carbon. p. 15- 49. *In* E.A. Paul, E.T. Elliott, K. Paustian, and C.V. Cole (ed.), Soil Organic Matter in Temperate Agroecosystems. CRC Press, New York.

Pawel, D., R. Fecso. 1988. On the use of correlations in crop yields. Proceedings of the Survey Research Methods Section, American Statistical Association (1988).

Prince, S. D., J. Haskett, M. Steninger, H. Strand, and R. Wright. 2001. Net primary production of US Midwest croplands from agricultural harvest yield data. Ecol. Appl. 11: 1194- 1205.

Reynolds, C.A., M. Yitayew, D.C. Slack, C.F. Hutchinson, A. Huete, and M.S. Petersen. 2000. Estimating crop yields and production by integrating the FAO crop specific water balance model with real-time satellite data and ground-based ancillary data. Int. J. Remote Sens. 21(18): 3487- 3508.

Rudorff, B. F. T. and G.T. Batista. 1990**.** Spectral response of wheat and its relationship to agronomic variables in the tropical region. Remote Sens. Environ. 31 (1): 53-63.

Rudorff, B. F. T., and G.T. Batista. 1991. Wheat yield estimation at the farm level using TM-Landsat and agrometeorological data. Int. J. Remote Sens. 12(12): 2477-2484.

Smith, R.C.G., J. Adams, D.J. Stephens, and P.T. Hick. 1995. Forecasting wheat yield in a Mediterranean-type environment from the NOAA satellite. Aust. J. Agric. Res. 46: 113-25.

Sperow, M., M.D. Eve and K. Paustian. 2003. Potential soil C sequestration on U.S. agricultural soils. Climatic Change 57: 319-339.

Tan, G., and R. Shibasaki. 2003. Glolbal estimation of crop productivity and the impacts of global warming by GIS and EPIC integration. Ecol. Model. 168: 357-370.

Tao, F., M. Yokozawa, Z. Zhang, Y. Xu, and Y. Hayashi. 2005. Remote sensing of crop production in China by production efficiency models: model comparisons, estimates and uncertainties. Ecol. Model. 183: 385-396.

Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. Geogr. Rev. 38(1): 55- 94.

USDA, 1982 to 1997. Agricultural Prices. Annual Summary 1981- Annual Summary 1996.
USDA. Washington, D.C. available at
http://usda.mannlib.cornell.edu/reports/nassr/price/ (accessed 05 May, 2005; verified 20 February, 2006).

USDA, 1998. USDA's National Agricultural Statistics Service: The Fact Finders of Agriculture. USDA. NASS.

Wilson, W.W., and J.L. Sebaugh. 1981. Established criteria and selected methods for evaluating crop yield models in the AgRISTARS program. Proceedings of the Survey Research Methods Section, American Statistical Association (1981).

Yang, P., G.X. Tan, Y. Zha, and R. Shibasaki. 2004. Integrating Remotely Sensed Data with an Ecosystem Model to Estimate Crop Yield in North China. Paper presented at the Geo-Imagery Bridging Continents XXth ISPRS Congress, 12-23 July 2004 Istanbul, Turkey.
http://www.isprs.org/istanbul2004/comm7/papers/29.pdf (accessed 25 Janurary, 2006; verified 20 February, 2006).

Table 1. Percentage of counties having < 30% and <5% difference, respectively, in NASS yield data compared to Ag Census

crop	**1982**		**1987**		**1992**		**1997**	
	<30%	<5%	<30%	<5%	<30%	<5%	<30%	<5%
Barley	92	37	93	40	91	23	99	60
Wheat	94	50	93	29	93	26	98	42
Alfalfa hay	82	19	80	15	78	13	92	37
Corn	95	37	96	46	95	35	99	71
Corn for silage	94	46	91	35	88	24	96	48
Sorghum	89	45	91	35	84	17	91	39
Other hay	77	24	81	19	80	17	88	36
Soybean	97	59	97	60	97	47	99	77
Oats	95	50	93	51	92	23	98	40

Table 2. States with reported data for county-level crop information (i.e.yields and crop area) by Ag Census and NASS during the study period (1982- 1997).

Crop	State
Alfalfa hay	AL[2], AR[2], AZ, CA[2], CO, CT[2], DE[2], FL[2], GA[2], IA, ID, IL, IN[2], KS, KY, LA[2], MA[2], MD[2], ME[2], MI[2], MN, MO[2], MS[2], MT, NC[2], ND, NE, NH[2], NJ, NM, NV, NY, OH[2], OK, OR, PA, RI[2], SC[2], SD, TN, TX, UT, VA, VT[2], WA, WI, WV, WY
Barley	AZ, CA, CO, DE, ID, KS[1], KY[1], MD, MI[1], MN, MT, NC[1], ND, NE[1], NJ[1], NM[1], NV, OK[1], OR, PA[1], SC[1], SD, TX[1], UT, VA[1], WA, WI[1], WV[1], WY
Corn	AL, AR[1], AZ, CA[1], CO, CT, DE, FL, GA, IA, ID[1], IL, IN, KS, KY, LA[1], MA, MD, MI, MN, MO, MS, MT[1], NC, ND, NE, NH, NJ, NM, NY, OH, OK[1], OR[1], PA, RI[2], SC, SD, TN, TX, UT[1], VA, VT, WA, WI, WV, WY[1]
Corn for silage	AR[1], AZ[1], CA[1], CO, DE[1], IA, ID[1], IL, IN, KS, KY[2], LA[1], MA[2], MD[1], MI[1], MN[1], MO[2], MS[1], MT, NC[1], ND, NE, NH[2], NJ[2], NM[1], NY, OH[2], PA, RI[2],, SD, UT, VA[1], VT[2], WA[1], WI, WV, WY
Oats	AL[1], AR[1], CA[1], CO[1], GA[1], IA, ID[1], IL[1], IN, KS, MD[1], ME[2], MI[1], MN, MO[1], MT, NC[1], ND, NE, NY, OH, OK[1], OR, PA, SC[1], SD, TX[1], UT, VA[1], VT[2], WA[1], WI, WV, WY[1]
Other hay	AL, AR[2], AZ, CA[2], CO, CT[2], DE[2], FL[2], GA[2], IA[2], ID[2], IL, IN[2], KS, KY, LA[2], MA[2], MD[2], ME[2], MI[2], MN, MO[2], MS[2], MT, NC[2], ND, NE, NH[2], NJ, NM, NV, NY, OH[2], OK, OR, PA, RI[2], SC[2], SD, TN, TX, UT, VA, VT[2], WA, WI, WV, WY
Sorghum	AL[1], AR, AZ[1], CA[1], CO, GA[1], IA[1], IL[1], IN[1], KS, KY[1], LA, MO, MS, NC[1], NE, NM, OK, SC[1], SD[1], TN[1], TX, VA[1]
Soy bean	AL, AR, DE, FL, GA, IA, IL, IN, KS, KY, LA, MD, MI, MN, MO, MS, NC, ND[1], NE, NJ, OH, OK, PA[1], SC, SD[1], TN, TX, VA, WI
Wheat	AL, AR, AZ, CA, CO, DE, GA, IA, ID, IL, IN, KS, KY, LA, MD, MI, MN, MO, MS, MT, NC, ND, NE, NJ[1], NM, NV, NY[1], OH, OK, OR, PA, SC, SD, TN, TX, UT, VA, WA, WI[1], WV, WY

([1]- states reported by NASS, but not by Ag Census; [2]- states reported by Ag Census, but not by NASS)

Table 3. Regression models obtained for Ag Census and NASS crop yields for 1997.

Crop	Regression model	R^2
Alfalfa hay	NASS= 0.2497+ 1.006 Ag Census	0.86
Barley	NASS=- 0.1863+ 1.030 Ag Census	0.97
Corn for grain	NASS= 0.9819+ 1.009 Ag Census	0.97
Corn for silage & green chop	NASS= 9.7973+ 0.865 Ag Census	0.73
Oats	NASS= 4.7180+ 0.974 Ag Census	0.90
Other hay	NASS= 0.3297+ 0.907 Ag Census	0.71
Sorghum	NASS= 5.8482+ 0.967 Ag Census	0.92
Soy bean	NASS= 0.2880+ 1.018 Ag Census	0.98
Wheat	NASS= 0.3128+ 1.051 Ag Census	0.96

Table 4. Summary of the models used in filling the gaps in crop yields in different Land Resource Regions (LRRs) of the US. (MST=mean summer monthly temperature; P=annual precipitation; PET= annual potential evapotranspiration; ITA= irrigated/total crop area).

Model	alfalfa hay	barley	corn	Corn for silage	Oats	Other hay	sorghum	Soy	wheat
◘ MST	D		A, B, C, D, E	A, B, C, D, E	A, B, C, D, H	A, D, E	C		A
◘ P/PET*ITA	A	H					L		
◘ P, (P/PET*ITA)		J				R			
◘ ITA, P									
◘ no autoregression	C, O, U			O		C, U			
◘ no fixed effects				J					
◘ P, MST				R		L			O
◘ P*MST, P, MST		F	J			F, J, O	I	O	
◘ ITA	B, G	A, C, D, I							B, C, D, E
◘ MST, (P/PET*ITA), (P/PET*ITA*MST)						H	F	F, G	
◘ (P/PET*ITA), MST, (P/PET*ITA*MST), P	H, I, J, K, L, M, S	G, K, L, M, N, S, T	G, H, I, K, L, M, O, P, R, T, U	H, I, K, L, M, T	F, G, H I, J, L, N, O, P, R, S, T	G, K, M, O	G, K, O, T	H, I, K, L, M, N, P, U	G J, K, L, M, N, P, T
◘ ITA, MST	E, T	E, B		G	E	B, S	D, E, S		H
◘ P*MST, P, MST, (P/PET*ITA)	F, N, P, R	P, R	F, N, S	F, N, P, S	K, M	I, N, P, T	H, J, M, N, P	J, R, S, T	F, I, R, S

Table 5. Percentages of “outliers” in the total predicted values (including those predicted for both missing and the observed values) from mixed models, for yields and crop areas.

Crop	% “outliers” in predicted crop area	% “outliers” in predicted yields
Alfalfa hay	0.7	0.04
Barley	0.6	0.05
Corn for grain	0.2	0.004
Corn for silage	0.2	No outliers
Oats	0.72	No outliers
Other hay	0.01	No outliers
Sorghum	1	0.06
Soy	0.55	0.002
Wheat	0.002	No outliers

Table 6. AIC values from different model options for wheat area in some LRRs. AIC values of the best models are shown in bold. Diesel$=diesel price of the previous year; fert$=fertilizer price of the previous year; crop$= crop price of the previous year; CRP=crop area set aside under the Conservation Reserve Program during the current year; county_area= area of the county; P= precipitation of the previous year.

Model	LRR A	LRR B	LRR C	LRR D	LRR E
◘ All variables	419.38	316.36	539.72	3408.63	2506.07
◘ Diesel$, county_area	435.48	412.98	567.88	**3404.77**	2506.39
◘ Fert$, county_area	445.97	418.20	567.67	3409.26	**2502.34**
◘ Crop$, county_area	420.17	324.38	542.40	3414.19	2509.87
◘ P, county_area	445.06	420.50	566.93	3426.76	2510.64
◘ CRP, county_area	445.50	414.99	566.84	3428.41	2514.13
◘ Diesel$, fert$, crop$, county_area	417.47	**313.08**	536.78	3406.57	2505.17
◘ Diesel$, fert$, diesel$*fert$, crop$, county_area	419.05	313.72	**536.05**	3408.50	2506.90
◘ Crop$, P, CRP, county_area	421.95	328.38	545.46	3415.79	2510.45
◘ Diesel$, fert$, crop$, CRP, county_area	419.42	314.49	537.80	3408.34	2505.24
◘ Diesel$, fert$, crop$, P, county_area	**417.42**	314.96	538.70	3406.89	2505.87

Table 7. Summary of the models used in filling the gaps in crop areas in different Land Resource Regions (LRRs) of the US crop area. Pl. see the description under the title of Table 6 for abbreviations.

Model	alfalfa hay	barley	Corn	Corn for silage	Oats	Other hay	sorghum	Soy	wheat
◘ Diesel$, fert$, crop$, P, CRP, county_area	M, R, H	G, H, I, P, S	C, G, H, K, L, N, P	F, M, P, R, S	H, L, M, P, T, S, N	T, N	T	J, M, N, P	F, H, M, N, P
◘ Diesel$, county_area		C, D, M	D	C, E, G, L, N	D		D		D
◘ Fert$, county_area				D		D			E
◘ Crop$, county_area			B	G	A, E				C, G
◘ P, county_area	C, F, T	J, S		T		H, O			
◘ CRP, county_area	G	G, L			C			K	
◘ Diesel$, fert$, crop$, county_area		E	A, I			B, J	K		O
◘ Diesel$, fert$, diesel$*fert$, crop$, county_area	I, L, P, N	F, N, R, T	E, O	J	B, F, G, R	A, D, E, H, J, K, M, R, S	G, I, M, N, O	F, G, O	B, I, K
◘ Crop$, P, CRP, county_area	D		U, S		I, J	G, I, L	E, L	S, U, R, L	L
◘ Diesel$, fert$, crop$, CRP, county_area		B, H			O	P		I	
◘ Diesel$, fert$, crop$, P, county_area	B, A, E, J, S	A	M, T, F	B, C, E, K, L, N, T	I	F	F, H, J	T	A, R, S, T, J
◘ Diesel$, crop$, CRP, county_area		K	R, J	A, H, R	K		C, P		
◘ Default (no random effects or no AR(1))	C, O, U			O		C, U			

Table 8. Percent relative errors (i.e., (observed-predicted)/observed) associated with the predicted values of crop area from the mixed model runs for different crops, and those if the predicted values were the default (i.e. mean of all the observed values).

Crop	Relative error with the mixed model	Relative error with the default option
Alfalfa hay	-2.64	-10.35
Barley	-13.48	-41.76
Corn for grain	-7.28	-26.8
Corn for silage	-8.19	-21.14
Oats	-10	-41.2
Other hay	-3.39	16.1
Sorghum	-9.93	-52.7
Soy	6.33	-33.47
Wheat	-8.9	-30.34

Table 9. Percentage gaps[1] filled through the different methodologhical approaches.

Crop	Gaps in NASS	% gaps[1] filled from NASS to combined NASS+Ag Census database	% gaps in NASS filled by mixed models	% gaps in NASS filled by the default option[2]
Alfalfa hay	23017	20.77	77.81	1.42
Barley	7049	3.18	94.95	1.87
Corn	7822	10.92	87.78	1.3
Corn for silage	12417	18.56	80.98	0.46
Oats	9626	7.45	89.86	2.69
Other hay	27554	24.84	75.13	0.03
Sorghum	8938	2.52	94.63	2.85
Soy	5854	7.79	89.02	3.19
Wheat	7371	10.46	89.43	0.11

[1]- gaps correspond to the number of missing data in different years during the 16-year period in all the counties where each crop is grown.
[2]- only the "outliers" of the predicted values from the mixed models were filled with the default option (i.e. means of the observed values)

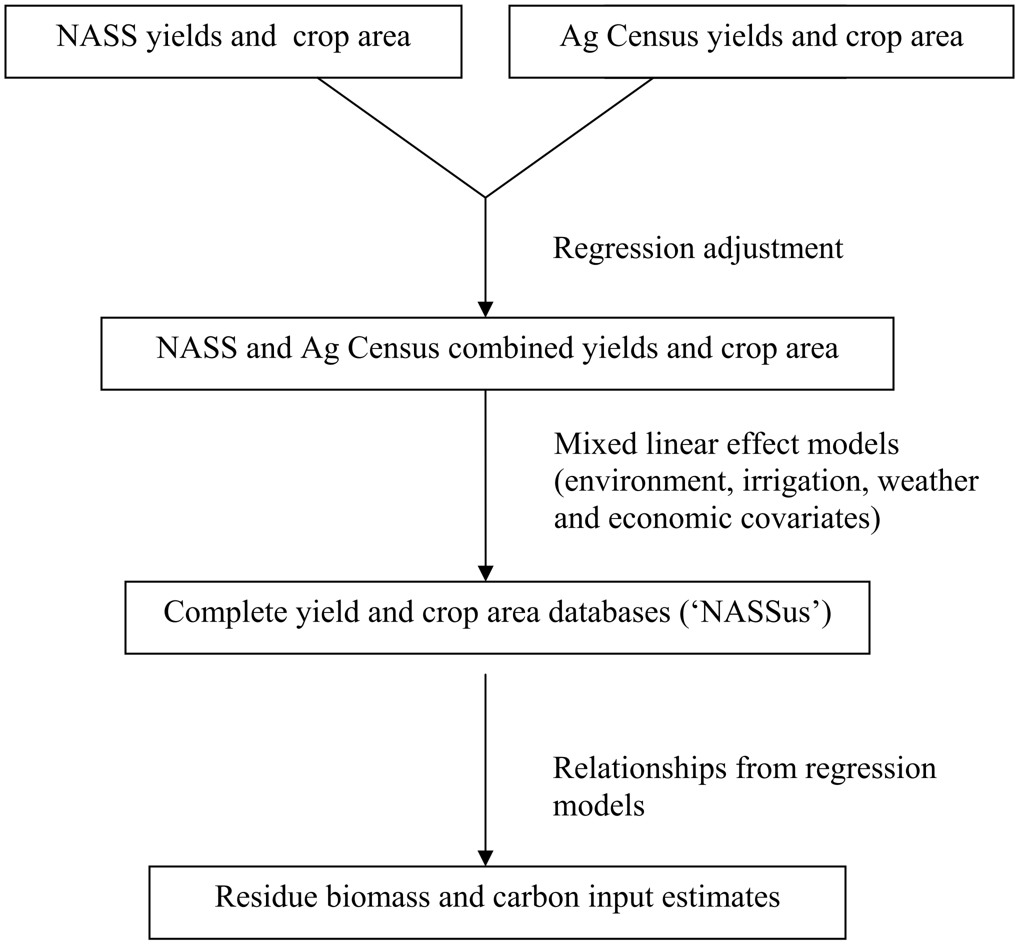


Figure 1. Methodological approach

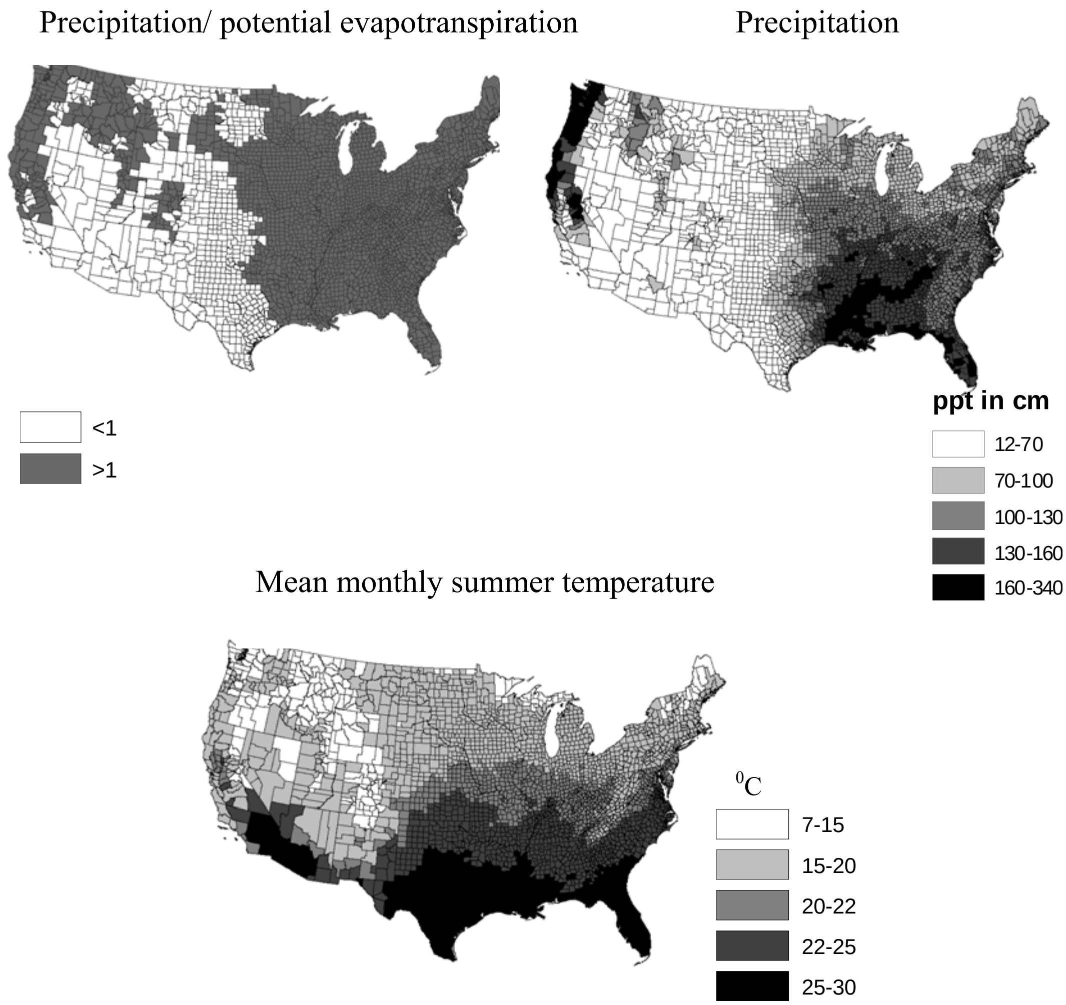


Figure 2. Maps of the main environmental variables used in the mixed models for filling the gaps in yields (Only the data from 1982 are shown here).

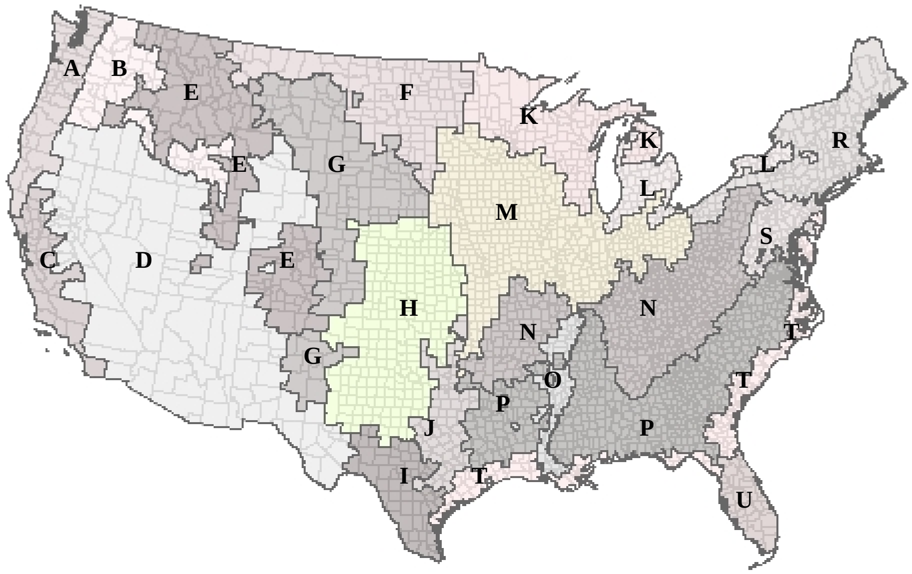


Figure 3. Map of the US Land Resource Regions (LRRs)

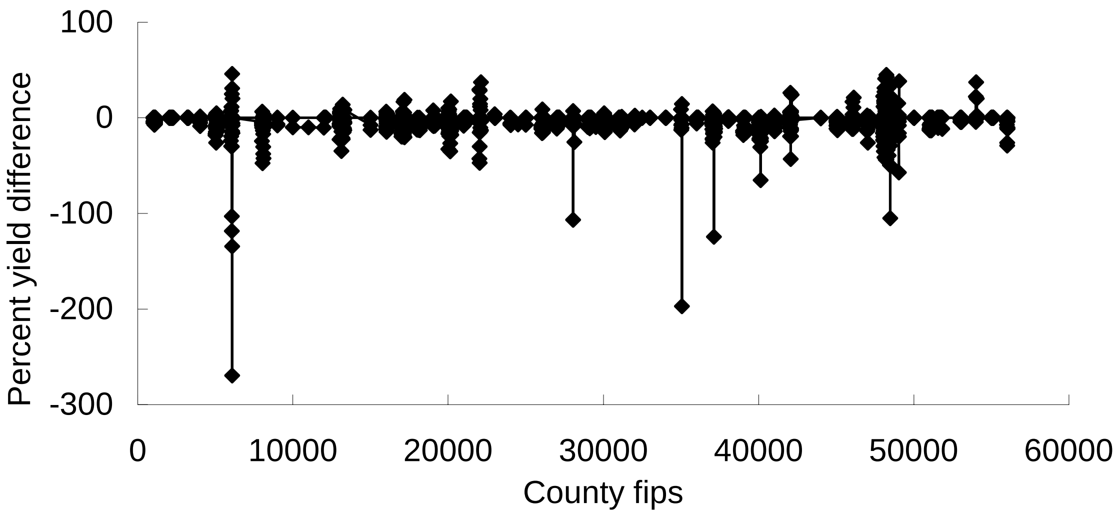


Figure 4. Difference in NASS wheat yield as a percentage of Ag Census data 1997

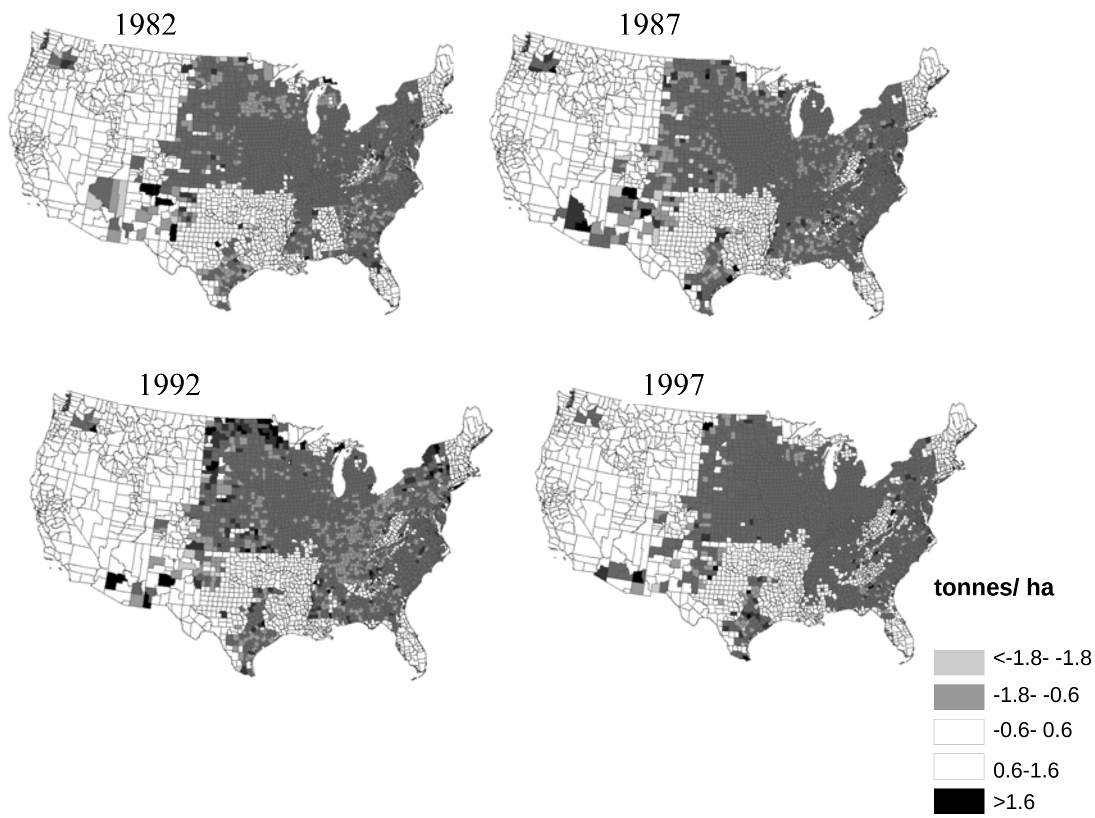


Figure 5. Difference in corn yields- Ag Census versus NASS

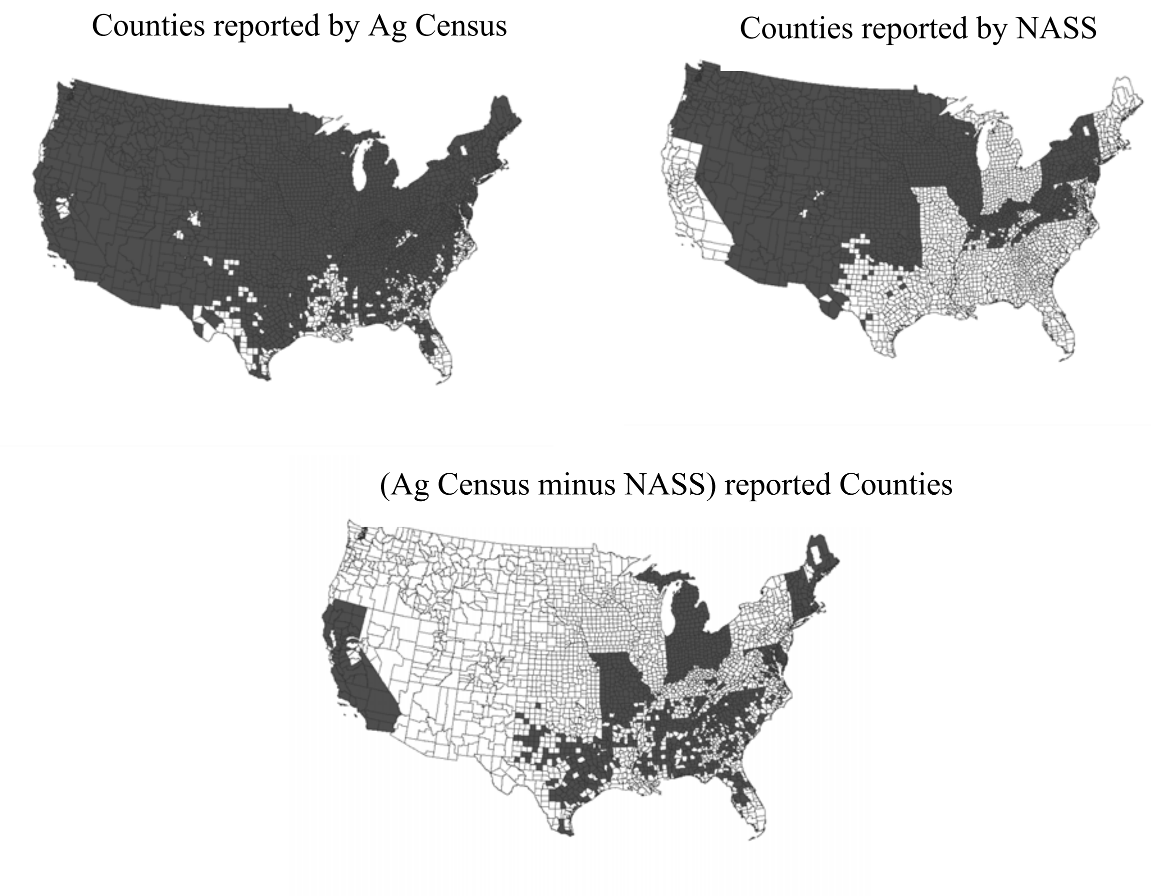


Figure 6. NASS and Ag Census differences in the reported counties – Alfalfa hay

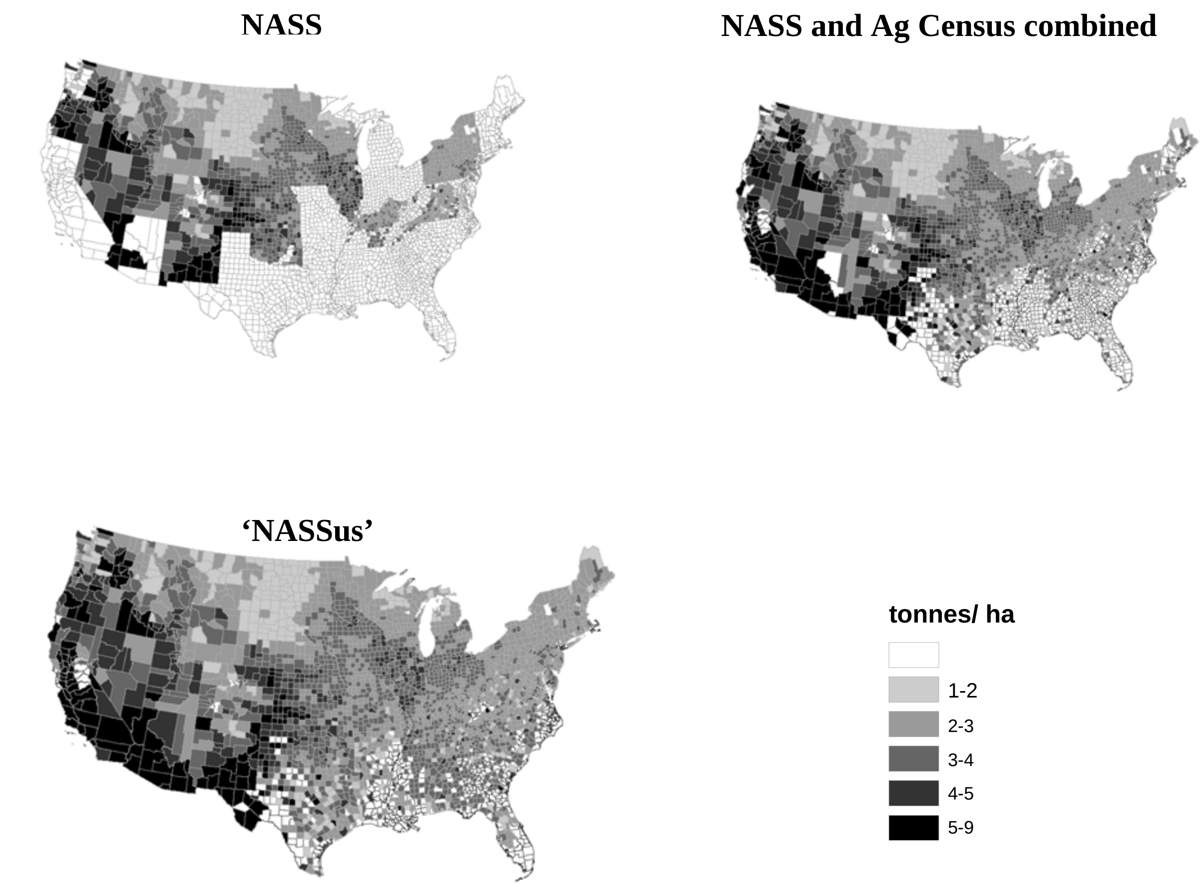


Figure 7. Original NASS, NASS- Ag Census combined, and final gap-filled database (i.e. 'NASSus') of crop yields in 1997- Alfalfa hay

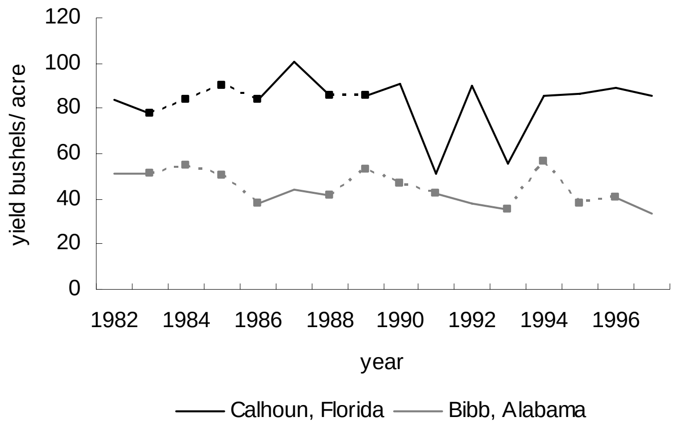


Figure 8. Trend in corn yields with observed (continuous line) and predicted (broken line and filled squares for the values; for years with no data) values, for two counties in two different states during the 16-year period.

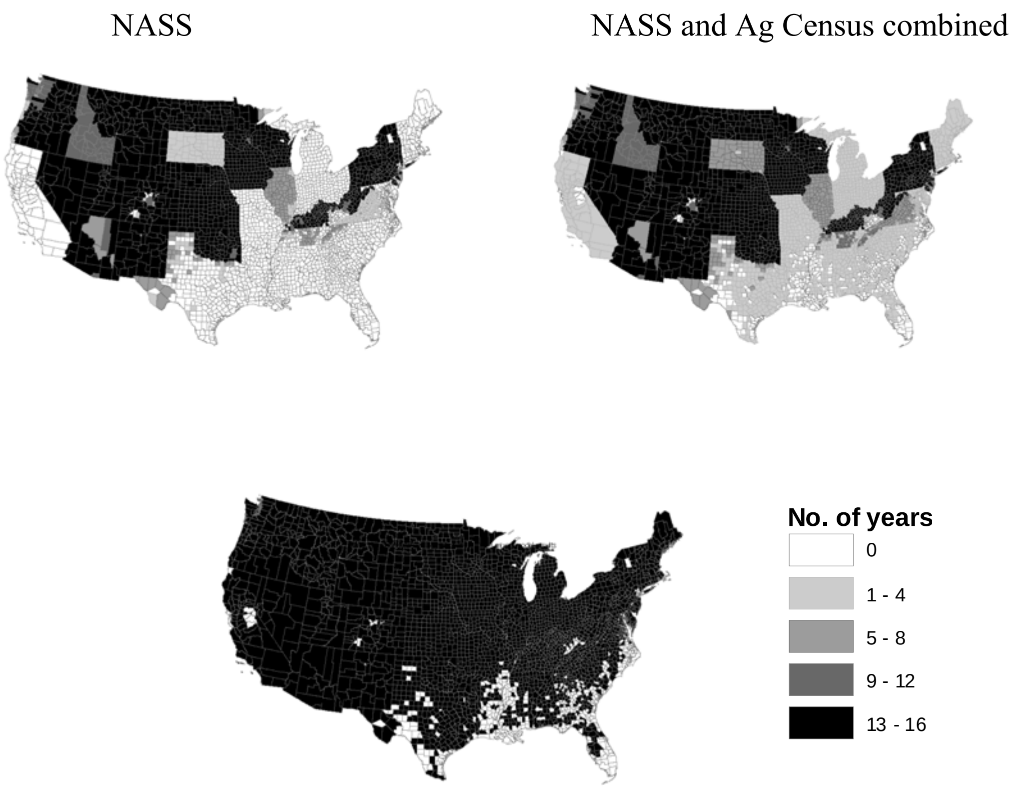


Figure 9. NASS, NASS and Ag Census combined, and 'NASSus' counts of years with alfalfa hay yields and crop area.

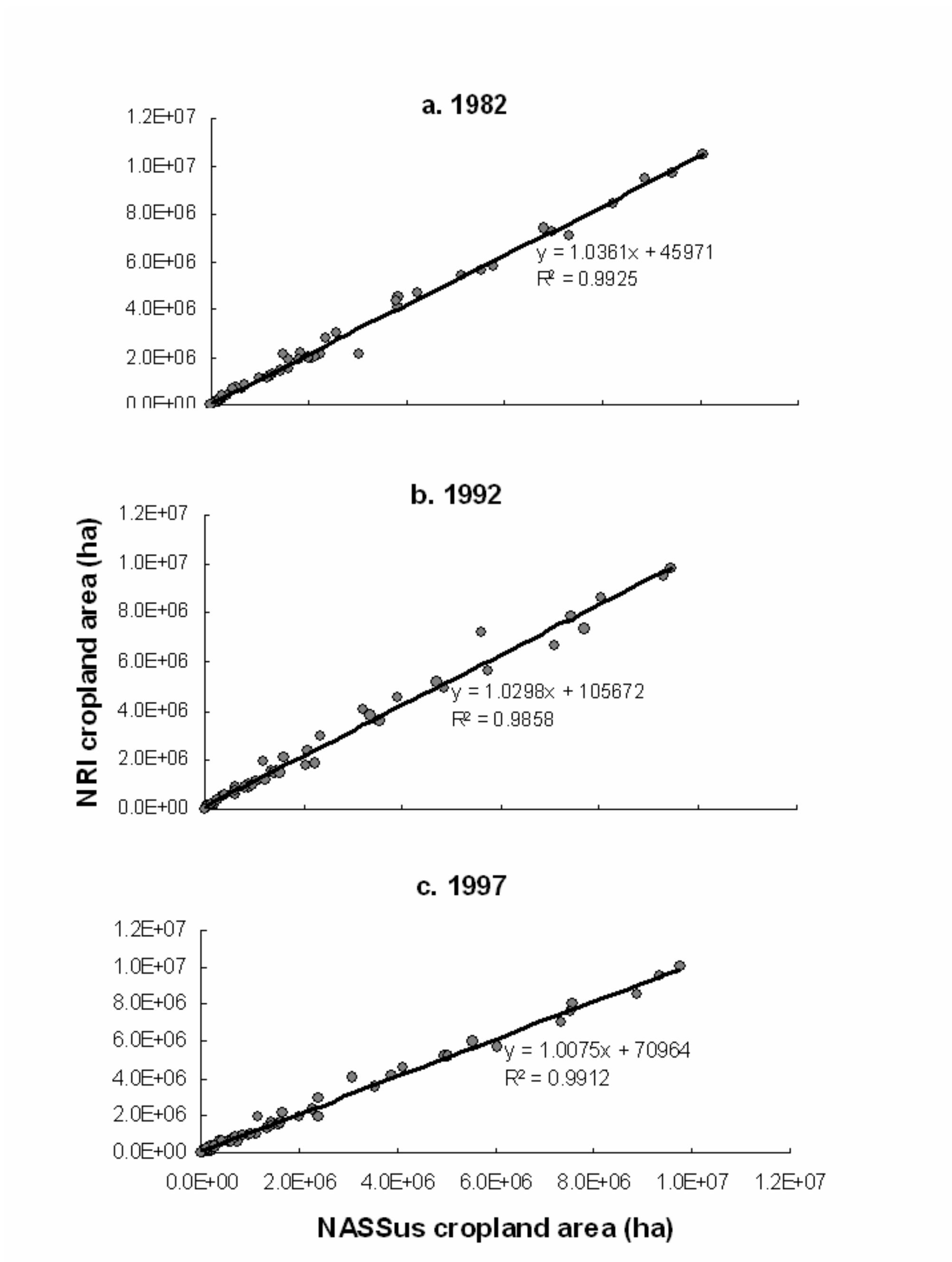


Figure 10. Final cropland area ('NASSus') of all the major crops for 1982 (top), 1992 (middle), and 1997 (lower), aggregated at state-level plotted against the state-level US total cropland area according to the National Resources Inventory (NRI).

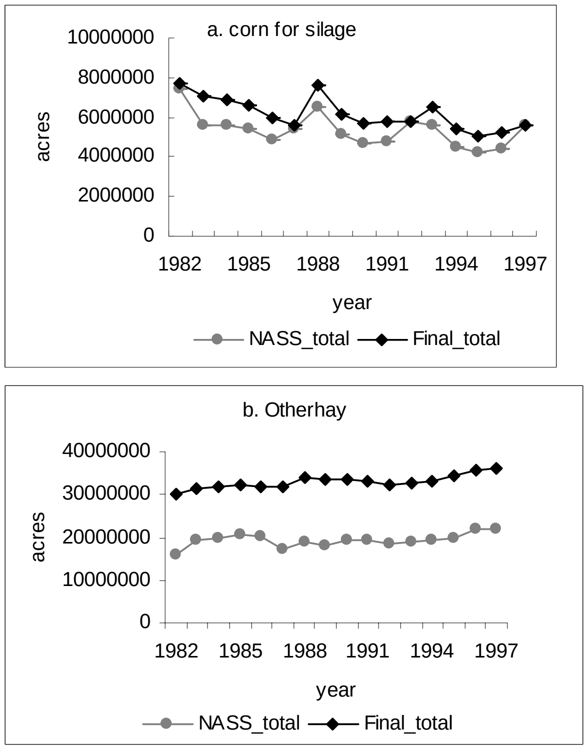


Figure 11. Corn for silage and other hay crop area reported by NASS and the final crop area ('NASSus') after gap filling

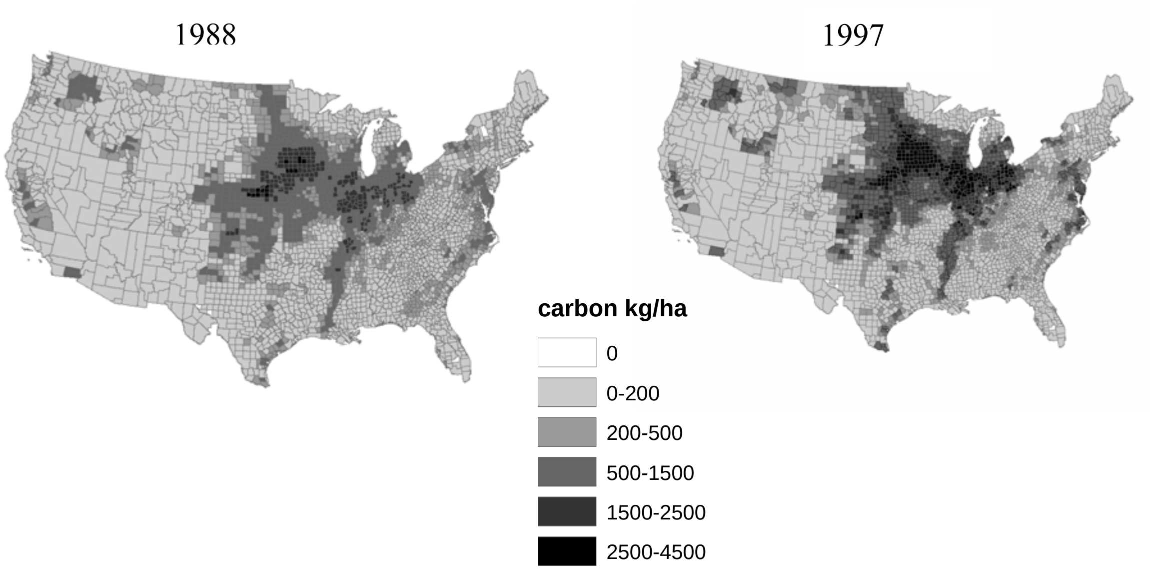


Figure 12. Total carbon inputs from the major crops of the US weighted by county area kg/ha- during a dry year (i.e. 1988; left) and a wet year (i.e. 1997; right)

CHAPTER 5

TEMPORAL AND SPATIAL VARIABILITY OF RESIDUE C INPUTS TO US AGRICULTURAL SOILS: IMPLICATIONS FROM THE TRENDS IN CROP PRODUCTION, CLIMATE AND WEATHER

ABSTRACT

Carbon (C) dynamics in agricultural soils depend directly on C inputs from crop residues. Since crop residues account for about 50-60% of the biomass, trends in residue C inputs also reflect any trends in the cropland net primary production (NPP). We analyzed the temporal and spatial variation of crop residue C inputs in US cropland soils over a 16-year period (i.e., 1982-1997) and their relation to climate variables, to interpret the interannual variability and temporal trends in the C balance of US agricultural soils.

Out of the seven US Crop Production Regions (CPRs), the North Central and Far West regions had the highest yields, NPP and C input rates. However, total annual residue C inputs were highest in the North Central and Central and Northern Plains regions that had the largest proportion (about 70%) of US cropland area. North East and Delta States, which encompass about 12% of US cropland, had the lowest total C inputs per ha. Average C input rates ranged from 1.8 ± 0.1 Mg ha^{-1} yr^{-1} in Delta States to 3.0 ± 0.3 Mg ha^{-1} yr^{-1} in North Central region; average NPP per unit cropland area ranged from 3.1 ± 0.2 Mg ha^{-1} yr^{-1} in Delta States to 5.4 ± 0.2 Mg ha^{-1} yr^{-1} in Far West region.

The interannual variability of C inputs was correlated with climate variables, especially the mean growing season temperature. The lowest yields and C input rates were observed in 1988, the year with the highest mean growing season temperature during the study period. Total residue C inputs and C input rates for US cropland were inversely proportional to the mean growing season temperature over the study period. This trend was also observed when comparing the spatial differences among the CPRs, as well. The Delta States and Southern Plains had the observed highest mean growing season temperatures and the lowest C input rates over the study period. A quadratic relationship incorporating total growing season precipitation and mean growing season temperature closely predicted the observed annual variation in residue C input rates at the CPR level, depicting the crop response to prominent weather changes, including severe droughts.

Total net C uptake (NPP) by the major US crops varied from a minimum of 379 Tg C yr^{-1} in 1988 to a maximum of 570 Tg C yr^{-1} in 1994, averaging 504 Tg C yr^{-1} over the study period. The observed maximum cropland NPP in 1994 was equivalent to about 40% of the total CO_2-C emitted from US fossil fuel combustion. Interannual variability (between adjacent years) in the total residue C returned to the US agricultural soils was as high as 96 Tg, and averaged 47 Tg. Variations of these magnitudes have major implications for estimating the C balance of US croplands over short time scales.

INTRODUCTION

The agricultural sector is among the major sources of increased greenhouse gas (GHG) emissions, both globally and in the US. Cropland occupies about one-fifth of the area of the US and it has had a significant impact on the country's economy and the environment. Total cropland of the US extends over seven major Crop Production Regions (CPRs) that vary in geography and climate: Far West, Central and Northern Plains, Southern Plains, North Central, Delta States, Northeast, Southeast. The number of farms in the US has decreased since the mid twentieth century, from about 5.5 million to less than 2 million in 2000, but the total cropland area under production has remained fairly constant (Ray et al., 2003).

The role of agriculture in carbon (C) cycling in the US and the potential for C sequestration and GHG mitigation by the agricultural sector has been widely examined over the past decade (Cole et al., 1995; Eve et al., 2002; Follett, 2001; Lal et al., 1998, 2004; Ogle et al., 2003; Paustian et al., 1995, 1996, 1997a, 1997b, 2001, 2002; Sperow et al., 2003; West and Marland, 2002a,b). These studies discuss how improved management practices that add and/or store more residue C, such as no-till, crop rotation, reduced summer fallow, and conservation buffers, etc., could help sequester C and reduce carbon dioxide (CO_2) emissions in agricultural soils.

However, the amount of C fixed by agroecosystems, and hence the C removed in products (i.e. grain, forage) or returned as residues to the soil has a dominant role in the agricultural C balance. Long-term trends in productivity and residue inputs to soil, driven by technological and management changes, also impact the magnitude and direction of change in soil C storage (Allmaras, 1999; Paustian et al., 1995, 1997). Moreover, because of the high rates of annual primary productivity of many agricultural crops, relative to non-cropland vegetation, agricultural systems exert a disproportionate influence on short-term C cycle fluxes. Thus, the high seasonal and interannual variability in CO_2 uptake and C additions to soil, due to climate variability, in croplands has significant implications for short-term changes (i.e. over a few years) in C inventories (Marland et al., 2003) as well as atmospheric-based estimates of regional C cycling (Denning et al., 1995; Tian et al., 1999).

In the majority of the countries in the world, lack of comprehensive agricultural statistics has been a major constraint in developing accurate national inventories of cropland emissions. The National Agricultural Statistics Service (NASS) of the US reports annual crop statistics for the major US crops. Some studies (Prince et al., 2001; Lobell et al., 2003; Hicke and Lobell, 2004) have used the crop statistics reported by NASS for estimation of productivity and NPP at regional scale. Lobell et al. (2002) used NASS yield data from 1992 to validate the national cropland NPP estimates using remotely sensed information. However, there are significant gaps or missing data in the county-level yields and crop areas reported by NASS. For instance, NASS does not report annual county-level yield and crop area data for 21 states where alfalfa hay and other hay crops are grown. Thus using the existing NASS data alone for estimating national level crop productivity over a significant time period (e.g. Hicke et al., 2002) might underrepresent the true picture of the cropland productivity.

We used a new database, derived from the crop survey data reported by NASS and the US Agricultural Census, to quantify and interpret interannual variability in cropland NPP and residue C additions in the US, for a 16-year period. In earlier work, (Lokupitiya et al., submitted), we developed comprehensive county-level crop yield and area databases for major crops in the US, using a suite of statistical models with climate and economic data to fill data gaps in the crop statistics reported by NASS and Ag Census. The databases include the dominant crop types in the US, i.e., alfalfa (*Medicago sativa* L.) hay, barley (*Hordeum vulgaris* L.), corn (*Zea mays* L.) for grain, corn for silage and green chop, oats (*Avena sativa* L.), other hay (hay other than alfalfa; i.e. tame hay, small grain hay, wild hay), sorghum (*Sorghum bicolor*), soybean (*Glycine max* L.), and wheat (*Triticum aestivum* L.) that make up over 90% of the total US harvested cropland area. In this second phase of the study, we used completed crop yield and area databases to estimate crop residue C (both aboveground and belowground) inputs in the US agricultural soils to study their temporal and spatial variation over a 16-year period, 1982-1997. We analyzed spatial and temporal variation in crop C input in relation to observed variation in crop yields, area extent, and climate over the study period.

MATERIALS AND METHODS

In earlier work (Lokupitiya et al., submitted), we combined and reconciled data on the annual crop areas and yields reported by the National Agricultural Statistics Service (NASS) and Census of Agriculture (Ag Census) to develop complete county-level yield and crop area databases. Missing data in yields and crop areas were filled by using regression analyses of primary NASS and Ag Census data and linear mixed-effect models incorporating several environmental and economic variables, run at Land Resource Region (LRR; Figure 1b) level. The gap filled databases were subject to thorough

Quality Assurance/ Quality Control (QA/QC) and validated against the cropland data from National Resources Inventory (NRI).

The completed yield and crop area databases were then used to estimate residue C inputs for alfalfa hay, barley, corn for grain, corn for silage, oats, other hay, sorghum, soybean, and wheat over the 16- year period, 1982-1997. Crop yields were corrected for moisture content and equations containing crop-specific harvest indices and root:shoot ratios and C concentrations (Buvanovsky and Wagner 1986; Campbell and Jong, 2003; Williams and Paustian, submitted; Prince et al., 2001) were used to calculate total above- and belowground biomass production (NPP) and residue C inputs from different crops. Crop area information was used to estimate total inputs over the cropland area (ha) in each county. Monthly precipitation and temperature data for the 16-year period were obtained from the PRISM database (Daly et al., 1994), which consists of gridded (4 km^2) values for the conterminous US, and aggregated to county-level. County-level monthly potential evapotranspiration was estimated for the whole US for the study period using the method by Thornthwaite (1948). Correlation between each weather variable and C input rates (Mg ha^{-1} yr^{-1}) for each CPR were determined using the Pearson correlation; county-level data were aggregated to the whole US to test correlations between residue C inputs and weather variables at national level. To determine what combination of weather variables could predict the exact observed temporal pattern of C input rates, county-level data for annual, monthly, or mean growing season (i.e. monthly data averaged for the period April to September) weather variables were used as independent variables in a suite of linear mixed- effect models. These models were run (SAS version

9.1) for each CPR using combinations of the county-level data for above weather variables (including squared and interaction terms) and C input rates, to investigate the interannual variability of C inputs with variation in weather. The model with the lowest Akaike Information Criterion (AIC) was chosen as the best model. For summary interpretations, model estimated values from the county-level data for C inputs and other variables were aggregated by CPRs (Figure 1). Interannual variability in C input rates in different CPRs were analyzed using a 5-year moving average across the 16-year period by studying the variation of the slope of the observed curve over time.

RESULTS

Spatial variation of residue C inputs

The geographic distributions of major cropland species vary substantially (Figure 2). Hay, wheat, and corn occur throughout most of the conterminous US, while the other crops, particularly oats, barley, and sorghum have more limited distributions.

Spatial variation in C inputs depends on the relative dominance of different crop species, their production potential and biomass allocation, and climatic and other environmental factors. Most counties in the western US had very low residue C inputs per total county area, although the C inputs per ha of cropland were high in those counties where crops are grown (Figure 3), due to the prevalence of irrigation. Per county C input rates were highest in the counties of the Corn Belt region, which falls within North Central CPR (Figure 1).

Total residue C inputs over the entire period was highest in the North Central region followed by the Central and Northern Plains regions (Figure 4 and Table 1), which had the largest percentage of total cropland area (on the average 40% in North Central and 30% in Central and Northern Plains). Delta States and North East regions that had the lowest percentage of the total cropland area in any given year together encompassed about 10% of the total annual C inputs in the US cropland. The latter two regions had about 10-15% cropland area compared to the North Central and Central and Northern Plains. Although corn was the main crop during the period, Far West and Delta States in general had the lowest corn crop area, and the North Central region had the highest corn area.

Carbon input rates over the study period were closely related to yields. Highest yields and C input rates were observed in Far West and North Central regions for the majority of the crops. Delta States, Southern Plains and Central and Northern Plains regions in general had low yields and C input rates, except for relatively high yields and C input rates for small grain crops (except for barley) in Delta States, alfalfa hay and corn in Southern Plains, and soybean and corn for grain in Central and Northern Plains.

Small grain crops (i.e. barley, oats and wheat) in general had the highest yields in the Far West region, followed by eastern regions, North Central and Delta States. Relatively low yields and C input rates for small grain crops were found in Southern Plains and Central and Northern Plains.

Hay crops had high yields and C input rates in North Central, Far West and South East regions, and low yields and C input rates in Central and Northern Plains and North East regions. Southern Plains had relatively high yields and C input rates from hay other than alfalfa.

Row crops corn and sorghum had the highest yields and C input rates in the Far West region (Table 2). The yields and C input rates of corn, sorghum and soybean were high in the North Central region, too. Sorghum and soybean yields were low in Southern Plains and Delta States, and South East region in general had low yields for all three row crops.

Crop residue C inputs clearly showed some variation due to climatic effects, especially due to variation in mean growing season temperature. Mean growing season temperature differed by a maximum of 6 ^{0}C between production regions. Southern Plains (24.6 ± 0.45 ^{0}C) and Delta States (24.8 ± 0.56 ^{0}C) had the highest mean growing season temperatures closely followed by South East (23.2 ± 0.6 ^{0}C) region. Both Far West and North Central regions, which had high C inputs, had moderate temperatures (averaging between 17- 20^{0} C) during the mean growing season. Annual C input rates from all the regions (combined) were negatively correlated with the range of mean growing season temperatures observed across the regions ($r= -0.64$; $p<0.05$).

Precipitation/potential evapotranspiration (P/PET) ratio, which is related to moisture availability for plant growth, varied greatly across the different crop production regions. The percentage of counties with P/PET ratio less than 1 (i.e. low moisture availability) was high for the CPRs Central and Northern Plains (ranged 34- 91% of the total number of counties within the region), Far West (34- 70% of the total), and Southern Plains (47-91% of the total), and much lower for the other regions (data not shown). Majority of the counties in Far West and Southern Plains had a large proportion of irrigated crop area (>50% of the total cropland area).

Although the PET varied among the CPRs, within a CPR it varied temporally within a narrow range. Highest PET in any given year was observed in the Southern Plains, followed by the Delta States, South East, North Central, Far West, North East, and Central and Northern Plains, respectively.

The Far West region had the lowest mean annual precipitation (50.7 ± 9.9 cm) across the time series, followed by the Central and Northern Plains (57.5 ± 6.8 cm) and Southern Plains (79.5± 8.7 cm). Delta States, North East, and South East regions had precipitation over 100 cm, and Delta States had the highest average of the annual precipitation over the period (139.8 ± 15.5 cm). The North Central region, which had the highest total C inputs, had an intermediate average annual precipitation (90.7 ± 11.6 cm). There was no statistically significant correlation between C inputs and P/PET when the data from all the regions were combined.

Comparison with cropland NPP

Residue C input rates (Mg ha^{-1} yr^{-1}) closely followed the trends in annual cropland net primary production (Mg ha^{-1} yr^{-1}). Highest NPP and C input rates were observed in the

Far West and North Central regions, followed by Central and Northern Plains. The average regional annual cropland NPP over the 16-year period ranged from 3.1 ± 0.25 Mg ha^{-1} yr^{-1} in Delta States to 5.36 ± 0.22 Mg ha^{-1} yr^{-1} in the Far West. Although the average NPP in crop areas was high in Far West, the crop areas were relatively a very small percentage of the total county areas in the region as a whole (only 5% crop area). At national scale, the US cropland NPP over the 16-year period ranged from 3.61 Mg ha^{-1} yr^{-1} in 1988 to 5.07 Mg ha^{-1} yr^{-1} in 1994. Total annual C uptake in NPP by the US cropland averaged 504.4 ± 57.7 Tg yr^{-1} (Table 3), and total US residue C inputs was 61-62% of the total US cropland NPP.

Temporal variation in residue C inputs

The lowest between-year variation for C input rates (19% difference between the minimum and maximum) over the 16-year period was observed for Delta states, and the highest variation (40%) was observed for the North Central region.

Variation of C inputs is directly related to both temporal and spatial variation of the yields and the area extent of different crops. Over the 16-year period, the US area (harvested) under crops such as alfalfa hay, barley, and oats in general had decreased over time, while the total area under corn and soybean had increased towards the end of the period (Figure 5). Average yields of barley, corn, oats, sorghum, soybean, and wheat showed an overall increasing trend over the 16-year period, although there was considerable variation from year to year, due to variation in weather. Average yields of hay crops and corn for silage remained relatively stable (Figure 6). The largest proportion of the total cropland area over the 16-year period was occupied by corn (24%), wheat (22%), and soybean (22%). These crops were followed by hay (other hay and alfalfa hay; jointly about 21% of the cropland area), sorghum, barley, oats, and corn for silage, each of which occupied less than 5% of the total cropland area.

In general, yields and C input rates in different CPRs were relatively low in 1983, 1988, 1991, 1993 (except for Southern Plains in 1993), and 1995. Total residue C inputs were the lowest in 1983 and 1988 and the majority of CPRs had the lowest annual C input rates during 1983, and/or 1988. For the most part, yields and C inputs were relatively high across the CPRs during the years 1992, 1994, 1996 (except for the low C input rates in Southern Plains in 1996), and 1997. During 1982, 1984, 1986, 1987, 1989, 1990, more mixed results were shown; yields from certain crops were high in certain regions, while they were low in the other regions during the same year.

Based on the trend in the moving averages, average C input rates (from all the crops combined) increased from 1982 to 1997, by 0.22 Mg ha^{-1} in the Central and Northern Plains, 0.23 Mg ha^{-1} in the Delta States, 0.42 Mg ha^{-1} in the Far West, 0.25 Mg ha^{-1} in North Central, 0.08 Mg ha^{-1} in North East, 0.23 Mg ha^{-1} in South East. In the Southern Plains, because of the lowest C input rates observed in 1996, the overall trend was neither increasing nor decreasing over the study period.

Carbon inputs varied with changing weather over the study period. The lowest C inputs to US cropland soils were observed in 1988, a prominent drought year with the lowest average annual precipitation (82 ± 35.4 cm), and the second highest mean growing season temperature (21.6 ± 3. ^{0}C) during the 16-year period (Figure 7). Within the different regions, there was no statistically significant relationship between temporal variation of C input rates and P/PET alone. Statistically significant relationships could be observed for the CPRs Central and Northern Plains, Far West, Southern Plains and Delta States when C input rates were regressed with the county-level data for mean growing season temperature and P/PET ratio.

Area-weighted annual residue C input rates (Mg ha^{-1} yr^{-1}) for the whole US were negatively correlated with the mean growing season temperature over the study period (i.e. $r= -0.65$; $p<0.05$). For individual CPRs, area-weighted annual C input rates were also negatively correlated with the mean growing season temperatures over the time series, in Delta States ($r= -0.51$; $p<0.05$), North Central ($r= -0.62$; $p<0.05$), North East ($r= -0.62$; $p<0.05$), and South East ($r= -0.74$; $p<0.05$) regions. The highest mean growing season temperatures were observed in 1987 and 1988, and the lowest C inputs were observed in 1988 (Figure 7b). Similarly, during other years that had above average growing season temperatures, the crop C inputs tended to be lower (Figure 7b).

The observed temporal trend in C input rates was well predicted by a quadratic relationship (i.e. the model with the lowest AIC in the mixed model runs; Table 4) that incorporated mean growing season temperature, squared mean growing season temperature, mean growing season precipitation, squared mean growing season precipitation, and interaction between the mean growing season temperature and mean growing season precipitation. The model coefficient for the quadratic term for mean growing season temperature was negative for all the CPRs, and the weather variables well predicted the very low C input rates observed in 1983, 1988, and 1993 in the Central and Northern Plains and North Central regions (Figure 8). The model correctly predicted the observed trend for the C input rates in the other CPRs, as well.

DISCUSSION

Net primary production and yields in croplands vary spatially and temporally due to spatial and temporal differences in climate, management and other factors. Regional patterns in crop production are largely driven by differences in the productivity potential of different crop species, which are largely governed by climate, but modified by management (e.g. irrigation). Temporal trends in production are influenced by developing technology (i.e. crop genetics, fertilization) and changes in management and agricultural policies, and by long-term trends in climate and CO_2 concentrations. Embedded within these longer trends, is the interannual variability of production, largely driven by climate variability and climate cycles (e.g. El Nino, La Nina), the effects of which also vary spatially across the continental US. Spatial variability is also related to the differences in soil quality, moisture and nutrient availability that are directly or indirectly related to climatic differences.

Crop NPP and residue C inputs showed significant interannual variability over the 16 year period. The total annual residue C inputs to the US agricultural soils over the 16-year period ranged from a minimum of 234 Tg C yr^{-1} in 1988 to 352 Tg C yr^{-1} in 1994. Total NPP ranged from 379 Tg C yr^{-1} in 1988 to 570 Tg C yr^{-1} in 1994. Interannual (year to year) variation in total residue C inputs was significant over the study period, ranging from 9-95 Tg C yr^{-1} ; interannaul variation in total C in annual NPP ranged from 15-151 Tg C yr^{-1}.

Our estimates are slightly lower than the total cropland NPP for the US estimated by Lobell et al. (2002) using satellite data (in CASA model) for the period 1982-1998. Their NPP estimates based on the satellite estimates in 1992 had been validated against the NPP estimates from NASS data in 1992. They estimated an NPP value of 620 Tg yr^{1} for the US cropland, which is 60 Tg higher than our estimate for 1992. Validation of estimates based on remotely sensed data for a drought year (i.e. giving NPP estimates for both ends of the range) might have better helped understand the degree to which the NPP derived from the two methods are compatible. Lobell et al. (2002) also incorporated a few more crops, compared to ours, in their study. Our study is the first study to estimate the C inputs to soils over the period, and hence no comparison could be done against the estimated crop-specific and total residue C inputs or annual C input rates over the period.

Yields and C input rates in different regions showed an overall increase over the study period. In the analysis of variation of the curve derived from the moving average, we found that the average residue C input rates in a majority of the CPRs increased by 0.22-

0.25 Mg ha^{-1} from the beginning to the end of the period. The long-term trend of increasing yields for most major crops, mainly since the 1930's is well recognized and is largely attributed to an array of technology and management developments, including increased fertilizer use, crop genetics, and pest management (Reilly and Fuglie, 1998). Although genetic improvements, particularly development of shorter stature varieties for small grains (e.g. wheat, barley) in the 1960s-70s, led to an increase in harvest index, characteristic harvest indices have remained relatively constant in recent decades (Williams and Paustian, submitted). Hence, the trend in yield increases during our study has driven a commensurate increase in net CO_2 assimilation and C inputs to soil.

Total US CO_2-C emissions from fossil fuel burning in 1994 were 1384 Tg C (US EPA, 1999). Thus the total cropland NPP was equivalent to 40% of the fossil fuel CO_2-C emissions and residue C inputs were equivalent to about 25% of the total fossil fuel CO_2-C emissions. This evidences the importance of cropland C dynamics in the overall C cycle of the US, especially in the regions like North East and Central and Northern Plains (where the largest crop areas are found compared to the other regions). In contrast to other anthropogenic processes in the C cycle such as CO_2 emissions from energy-use, C fluxes associated with CO_2 assimilation, harvest export and the soil C balance on croplands, exhibit a much higher degree of interannual variability.

Interannual variability of total C inputs and per ha C input rates directly depended on the yields and crop area extent. During the study period corn, wheat, soybean and hay crops occupied over 90% of the US cropland, and oats and barley occupied less than 5% of the total cropland. Over time, area under soybean and corn increased, while the area under alfalfa hay, barley, oats, and sorghum decreased; the yields of majority of the crops increased over time. Depending on the yields and area extent, annually the largest amount of C inputs were added from corn, followed by soybean, wheat and hay crops. But on a per ha basis the highest C input rates were from corn, sorghum, alfalfa hay, wheat, and soybean, respectively, although the area extent of sorghum was relatively very low (less than 5%).

The mean growing season temperature among the different CPRs during this period ranged between 16.7 and 25.8 ^{0}C, and the regions that had the highest mean growing season temperature had the lowest C input rates. The C input rates were inversely proportional to the mean growing season temperature over the 16-year period. Although we observed this trend for the country as a whole (and in few regions), regional differences existed in the influence of variation in temperature on yields. For instance, although we observed the lowest mean growing temperature and precipitation for the whole US in 1988, all the regions except for the North Central region and Central and Northern Plains, experienced less different or lower mean growing season temperatures compared to the previous year, but with lowered precipitation. Lowering of temperature and precipitation in Delta states (that normally has the highest precipitation among the regions) seem to have caused a slight increase in the C input rates compared to the neighboring years. Thus the overall trend we observed for the whole country seem to reflect mostly the effects in the North Central region and Central and Northern Plains, that had the largest crop areas, lowest precipitation and highest temperature in 1988.

The inverse relationship with increasing growing season temperature was quite prominent for the C input rates observed in the North Central region. For instance, for a 1^0C increase in the growing season temperature (and a decrease in precipitation of 8 cm compared to the previous year), a decrease of 0.76 Mg ha^{-1} in C input rate was observed in 1983. The same change in magnitude (i.e. a decrease by 0.76 Mg ha^{-1}) was observed in 1993, when the growing season temperature in the North Central increased by 1^0C (compared to1992), with the highest growing season precipitation observed for the region during the 16-year period. This might imply a potential reduction in productivity in the region with increases in growing season temperature, despite any associated variation in precipitation. Lobell and Asner (2003) observed decreasing trends for soybean and corn yields with increase in the growing season temperature and they also could not observe a relationship between the yields and precipitation alone (Lobell and Asner, 2003). However, inclusion of growing season precipitation in a quadratic model with growing season temperauture in our study, helped predicting the exact observed pattern of C input rates over the 16-year period (Figure 8).

As described above, the temporal variation we observed in C input rates were mostly related to changes in weather driven by climatic events such as El Nino. The drought in 1988 has been attributed to variation in the sea surface temperature in the tropical Pacific, and mostly restricted to the northern plains, west coast and South East (Rasmusson and Wallace 1983, Trenberth et al., 1988, Palmer and Branković, 1989). Drought in the mid west region occurred from April to June 1988, and by July of the same year, 43% of the area of the country was in severe drought (Trenberth et al., 1988). Drought can affect the critical stages of crop growth and yields. For instance, corn yield can be severely affected by the water stress at tasseling; four days of visible wilting before tasseling could reduce corn yield by 10-25% (Rosenzweig et al., 2000). Also, drought is the most important factor affecting soybean yields in the South Eastern US. (Hansen et al., 1997). Soybean is sensitive to water stress and drought at planting and from flowering to pod-fill.

Low C input rates were also observed in 1983, 1991 and 1993. On the average, C input rates were moderate to high during the years 1982, 1985, 1992, 1994, 1996 and 1997. All of these years encompassed El Nino events, some of which seem to have caused drastic effects, due to extreme weather, mostly during the second half (of the episode). For instance, during 1992-1993 episode, high yields (and thus residue C input rates) were observed in 1992 in the North Central region when the growing season temperature was relatively low. During 1993, growing season temperature was higher than the previous year and the floods in the Mississippi river valley occurred due to very high precipitation. Fungal epidemics in corn, soybean and alfalfa caused by high moisture, and outbreak of soybean sudden death syndrome in the US Midwest in 1993 (Rosenzweig et al., 2000), may have led to the (low yields and) low residue C inputs observed. During the same year, severe summer drought occurred in the South East region causing yield losses (Lott, 1994), thus reduced residue C inputs as we observed in our study. Thus the same El Nino event could yield differential responses in different CPRs.

According to our analyses, the major impacts on yields and residue C input rates were observed during the El Nino events 1982-1983, 1987-1988, and 1992-1993, the effects of which were reflected mostly in the relatively high yields and C inputs during the first part of the episode, and reduced yields and C inputs during the second half of the episode. Although such interannual variability could be observed, the overall trend in NPP and residue C inputs was increasing during the period. This implies that management and technological improvements could still help increase C uptake in the US croplands through increased NPP and higher productivity that could lead to increased soil C storage, despite certain short-term impacts from the extreme weather events.

CONCLUSION

Total C inputs over the 16- year period varied depending on the crop area extent, yields, crop type, weather, and climate. Largest amounts of C inputs were added from the crops that had the greatest area extent. Region wise, the North Central region, which had a significant crop area, moderate mean growing season temperatures, and precipitation/ potential evapotranspiration ratio >1, had the highest yields and residue C inputs. The Far West region had a very low crop area (about 5%), but the C input rates (ha^{-1}) within the cropland area was relatively high. This region had the lowest mean growing season temperatures and precipitation lower than potential evapotranspiration; thus most of the crops were irrigated. Highest C input rates were observed for corn, followed by sorghum, alfalfa hay, wheat, and soybean. But since the area of sorghum was very low, the total C inputs added were much more significant for the other crops that had relatively high crop areas.

Within the temperature range observed for the period, C inputs in the whole US and certain Crop Production regions showed a negative correlation with the mean growing season temperature. In the US, the yields and C input rates were low during the years that had relatively high average mean growing season temperatures. Also the regions that had relatively high mean growing season temperatures (e.g. Delta States and Southern Plains) had relatively low production and residue C input rates. Although no statistically significant relationship could not be observed for precipitation, the interaction between the temperature and precipitation in a quadratic relationship in linear mixed effect models, proved the impact of precipitation especially during the prominent drought years. The annual cropland NPP estimated for the period ranged from a minimum of 378.8 Tg C yr^{-1} in 1988 to a maximum of 569.5 Tg C yr^{-1} in 1994. Cropland is a significant component of the C cycle for the continental US; for example, cropland NPP was equivalent to about 40% of the CO_2-C emissions from fossil fuel burning by the country. Despite the interannual variability associated with weather changes, overall C inputs to the US cropland soils increased over time. However, the interannual variability in the amount of C added to soil, was generally greater than the magnitude of the total mean increase over the 16-year period. This degree of variability in C inputs has implications for estimating short-term changes in the net C balance of cropland soils.

Acknowledgements

We are grateful for the assistance given by Mark Easter and John Norman. Support for this work was provided by the Consortium for Agricultural Soil Mitigation of Greenhouse Gases (CASMGS) funded by Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture, under Agreement No. 2001-38700-1109.

REFERENCES

Allmaras, R.R., H.H. Schomberg, C.L. Douglas, Jr., T.H.Dao. 1999. Conservation tillage's unforeseen advantage. Resource: Engineering & Technology for a Sustainable World 6, 7-8.

Buyanovsky, G.A., and G.H. Wagner. 1986. Post-harvest residue input to cropland. Plant Soil 93, 57- 65

Campbell, C.A., and R. de. Jong. 2001. Root-to-straw ratios- influence of moisture and rate of N fertilizer. Can. J. of Soil Sci. 81(1), 39- 43.

Cole, V., C. Cerri, K. Minami, A. Mosier, N. Rosenberg, D. Sauerbeck, J. Dumanski, J. Duxbury, J. Freney, R. Gupta, O. Heinemeyer, T. Kolchugina, J. Lee, K. Paustian, D. Powlson, N. Sampson, H. Tiessen, M. Van Noordwijk and Q. Zhao. 1996. Agricultural options for mitigation of greenhouse gas emissions. p 745-771. *In* R.T. Watson, M.C. Zinyowera, R.H. Moss, and D.J. Dokken, (ed.) Climate Change 1995. Impacts, adaptations and mitigation of climate change: Scientific-technical analyses. IPCC Working Group II. Cambridge University Press, Cambridge.

Conservation Technology Information Center (CTIC). 1998. National Crop Residue Management Survey News Release. [Online]. Available at http://www.ctic.purdue.edu/CTIC/nr.html (accessed 05 December, 2005; verified 09 January, 2006). CTIC, West Lafayette, IN.

Conservation Technology Information Center (CTIC). 2002. National Crop Residue Management Survey News Release. [Online]. Available at http://www.ctic.purdue.edu/Core4/CT/ctsurvey/2002/2002NewsRelease.html. CTIC, West Lafayette, IN.

Daly, C., R.P. Neilson, and D.L. Phillips. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. Journal of Applied Meteorology 33, 140- 158.

Denning, A.S., I.Y. Fung, and D.A. Randall. 1995. Latitudinal gradient of atmospheric CO_2 due to seasonal exchange with land biota. Nature 376, 240-243.
Edmeades, G.O., J. Bolanos, S.C. Chapman, H.R. Lafitte, and M. Banzinger. 1999. Selection improves drought tolerance in tropical maize populations: I. Gains in biomass, grain yield, and harvest index. Crop Sci 39, 1306-1315.
Eve, M.D., K. Paustian, R.F. Follett, and E.T. Elliott. 2001a. A national inventory of changes in soil C from national resources inventory data. p 593- 609. *In* R. Lal, J.M.Kimble, R.F. Follett, and B.A.Stewart (ed.) Assessment methods for soil C. CRC Press, Boca Raton, Fla. ,
Eve, M.D., M. Sperow, K. Paustian, and R.F. Follett. 2001b. National-scale estimation of changes in soil C stocks on agricultural lands, Environ. Pollut. 116, 431- 438
Eve, M.D., M. Sperow, K. Howerton, K. Paustian, and R.F. Follett. 2002. Predicted impact of management changes on soil C storage for each cropland region of the conterminous United States. J. Soil Water Conserv. 57(4),196-204
Follett, R.F. 2001. Soil management concepts and carbon sequestration zin cropland soils. Soil Till. Res. 61, 77- 92.
Hansen J.W., A.W. Hodges, and J.W. Jones. 1998. ENSO influences on agriculture in southeastern United States. J Climate11 (3): 404-411
Hicke, J.A., D.B. Lobell, and G. Asner. 2004. Cropland area and net primary production computed from 30 years of USDA agricultural harvest data. Earth Interactions (8), 1-20
Kamara, A.Y., A. Menkir, B. Badu-Apraku, and O. Ibikune. 2003. The influence of drought stress on growth, yield and yield components of selected maize genotypes. J. Agr. Sci. 141, 43-50
Legler DM, K.J. Bryant, and J.J. O'Brien. 1999. Impact of ENSO- related climate anomalies on crop yields in the US. Climatic Change 42 (2), 351-375
Lobell, D.B., J.A. Hicke, G.P, Asner, C.B. Field, C.J. Tucker, C.J., and O. Los. 2002. Satellite estimates of productivity and light use efficiency in United States agriculture, 1982-98. Glob. Change Biol. 8, 722-735
Lobell, D.B., and G.P. Asner. 2003. Climate and management contributions to recent trends in US agricultural yields. Science 299, 14 February 2003.
Ogle, S.M., J. Breidt, M.D. Eve, and K. Paustian. 2003. Uncertainty in soil organic carbon storage for US agroecosystems due to land use and management between 1990 and 1997. p. 133-143. *In* C.A.S. Smith (ed.) Soil Organic Carbon and Agriculture: Developing Indicators for Policy Analyses. Proceedings of OECD expert meeting, Ottawa, Canada. Agriculture and Agri- Food Canada, Ottawa, and Organization for Economic Co-operation and Development, Paris.
Palmer, T.N. and ĉ. Branković. 1989. The 1988 US drought linked to anomalous sea surface temperature. Nature 338, 54 – 57.
Paustian, K., E.T. Elliott, H.P. Collins, C.V. Cole, and E.A. Paul. 1995. Use of a network of long-term experiments for analysis of soil carbon dynamics and global change: the North American model. Aust. J. Exp. Agr. 35: 929-39
Paustian, K., E.T. Elliott, G.A. Peterson, and K. Killian. 1996. Modeling climate, CO_2, and management impacts on soil C in semi-arid ecosystems. Plant Soil 187, 351-365.

Paustian, K., H.P. Collins, and E.A. Paul. 1997a. Management controls on soil C. p.15-49. *In* E.A. Paul, E.T. Elliott, K. Paustian, and C.V. Cole (ed.) Soil Organic Matter in Temperate Agroecosystems. CRC Press, New York.

Paustian, K., E. Levine, W.M. Post, and I.M. Ryzhova. 1997b. The use of models to integrate information and understanding of soil C at the regional scale. Geoderma 79, 227- 260.

Paustian, K., E.T. Elliott, M.R. Carter. 1998. Tillage and crop management impacts on soil C storage: use of long-term experimental data. Soil Till. Res. 47, vii-xii

Paustian, K., K. Killian, J. Cipra, G. Bluhm, J.L. Smith, E.T. Elliott. 2001. Modelling and regional assessment of soil C: A case study of the Conservation Reserve Program. Soil Sci. Soc. Am. J. 57, 207- 225.

Paustian, K., J. Brenner, K. Killian, J. Cipra, S. Williams,S., E.T. Elliott, M.D. Eve, T. Kautza, and G. Bluhm. 2002. State-level analyses of C sequestration in agricultural soils. p 193- 203. *In* R. Lal, J.M. Kimble, and R.F. Follett (ed.) Agricultural practices and policies for C sequestration in soil, New York, Lewis Publishers.

Prince, S. D., J. Haskett, M. Steninger, H. Strand, and R. Wright. 2001. Net primary production of US Midwest croplands from agricultural harvest yield data. Ecol. Appl. 11: 1194- 1205.

Ray, D. , D. De La Torre Ugarte, and T. Kelly. 2003. Rethinking US agricultural policy: changing course to secure farmer livelihoods worldwide. Agricultural Policy Analysis Center. The University of Tennessee.

Reilly, J.M., and K.O. Fuglie. 1998. Future yield growth in field crops: what evidence exists? Soil & Till. Res. 47, 275-290

Schertz, D. 1986. Conservation tillage: an analysis of crop area projections in the United States. J. Soil Water Conserv. 33: 256- 258.

Sperow, M., M.D. Eve and K. Paustian. 2003. Potential soil C sequestration on U.S. agricultural soils. Climatic Change 57:319-339.

Thornthwaite, C.W. 1948. An approach toward a rational classification of climate. Geogr. Rev. 38(1), 55- 94.

Trenberth, K. E., G.W. Branstator, and P.A. Arkin. Science 242, 1640−1645

Uri, N.D. 2000. Global climate change and the effect of consevation practices in US agriculture. Environ. Geol. 40 (1-2): 41- 52.

US EPA. 1999. Inventory of U.S. greenhouse gas emissions and sinks: 1990- 1997. US EPA. Washington, D.C.

West, T.O. and G. Marland. 2002a. A synthesis of carbon sequestration, carbon emissions, and net carbon flux in agriculture: comparing tillage practices in the United States. Agr. Ecosyst. Environ. 91, 217- 232.

West, T.O. and G. Marland. 2002b. Net carbon flux from agricultural ecosystems: methodology for full carbon cycle analyses. Environmen. Pollut. 116, 439- 444.

Table 1. Annual C inputs Tg over the 16-year period in the major Crop Production Regions of the US.

Year	CNP	DS	FW	NC	NE	SE	SP	US total
1982	82.8	13.1	21.1	164.0	11.8	27.7	23.4	344.0
1983	66.2	10.5	20.7	104.8	9.7	18.0	18.3	248.3
1984	80.1	12.6	21.9	151.2	12.3	25.6	20.5	324.1
1985	85.6	11.3	20.8	168.7	12.8	24.7	22.8	346.6
1986	84.7	9.4	20.1	157.8	11.2	18.0	19.2	320.6
1987	80.6	8.9	18.4	141.9	10.4	17.2	16.6	293.9
1988	62.1	9.9	18.1	101.6	8.9	16.9	16.3	234.0
1989	70.5	9.1	19.1	149.7	10.3	20.8	17.1	296.6
1990	86.2	9.6	20.5	155.8	11.0	18.1	18.5	319.6
1991	82.9	8.1	18.1	143.0	9.5	17.5	17.2	296.2
1992	91.9	10.2	17.9	170.9	10.9	21.0	22.3	345.2
1993	79.9	8.6	20.3	127.0	9.5	16.5	18.9	280.8
1994	94.6	9.9	19.1	179.0	10.9	20.7	18.1	352.4
1995	79.5	8.7	21.0	146.5	9.8	18.0	16.4	299.9
1996	96.9	12.0	22.1	160.0	11.0	20.3	17.6	339.9
1997	97.8	10.9	21.6	168.4	10.1	19.0	21.8	349.6

Table 2. Ranges (minimum-maximum) for the area weighted annual C input rates during the period 1982- 1997 (Mg C per ha per year) for major US crops in the US observed in different CPRs

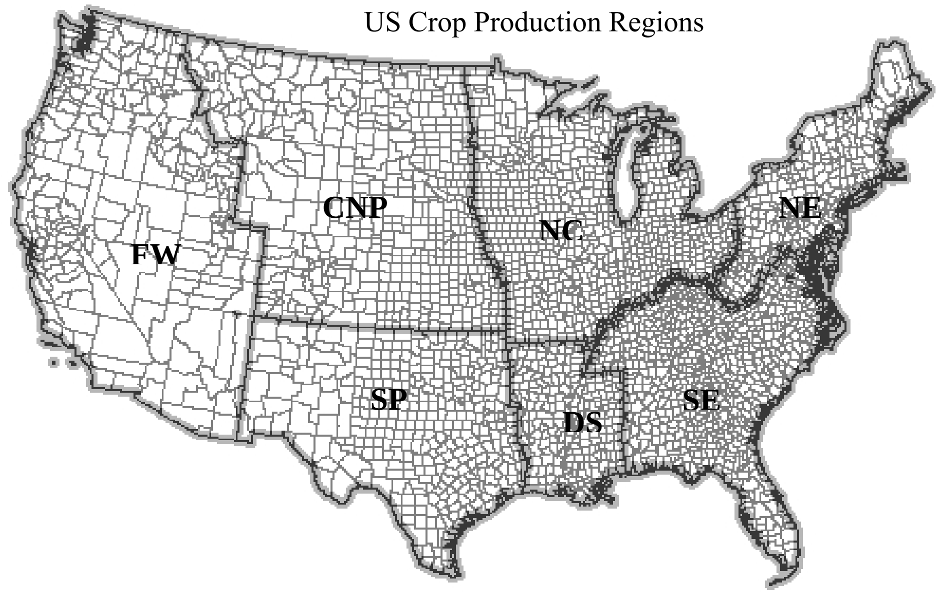
CPR	Range of annual C input rates (Mg ha^{-1} yr^{-1})									
	Alfafa hay	Barley	Corn for grain	Corn for silage	Oats	Other hay	Sorghum	Soybean	Wheat	Overall
CNP	1.88- 2.32	1.06- 2.04	3.39- 4.99	0.92- 1.18	0.92- 1.18	0.80- 1.13	2.13- 3.32	1.62- 2.36	1.55- 2.14	2.06-2.70
DS	2.10- 2.58	No crop	3.04- 4.34	No crop	No crop	1.14- 1.35	2.44- 3.01	1.46- 2.05	1.39- 2.99	1.80-2.48
FW	3.79- 4.17	1.95- 2.57	5.06- 6.35	1.67- 1.90	1.67- 1.90	1.24- 1.44	3.21- 3.59	No crop	3.14- 3.96	3.07-3.72
NC	1.77- 2.87	1.26- 2.41	3.08- 5.46	0.92- 1.24	0.92- 1.24	1.04- 1.27	2.52- 3.91	1.83- 2.56	2.16- 3.01	2.37-3.95
NE	2.12- 2.47	1.92- 2.52	2.82- 4.55	1.00- 1.32	1.00- 1.32	1.05- 1.16	No crop	1.68- 2.25	2.36- 3.49	1.95-2.56
SE	2.23- 2.91	1.57- 2.42	2.39- 4.21	0.97- 1.33	0.97- 1.33	1.16- 1.37	2.10- 2.91	1.45- 2.06	1.91- 3.01	1.76-2.57
SP	3.02- 3.58	1.54- 1.96	3.61- 5.18	1.46- 1.65	1.46- 1.65	0.92- 1.20	2.18- 2.64	1.54- 2.01	1.41- 2.09	1.96-2.33

Table 3. C in total cropland NPP (Tg C) over the period in different CPRs and the US

year	CNP	DS	FW	NC	NE	SE	SP	US total
1982	131.4	19.1	36.0	263.1	22.9	43.7	36.6	552.8
1983	105.0	15.4	35.2	168.0	19.4	28.8	29.5	401.3
1984	126.8	18.6	37.4	241.6	23.9	41.1	32.2	521.6
1985	135.3	16.9	35.9	269.8	24.6	40.2	36.6	559.4
1986	134.7	14.1	35.1	252.8	21.9	29.9	31.1	519.6
1987	127.8	13.3	32.2	227.0	20.5	28.7	27.1	476.5
1988	98.8	14.8	31.7	161.2	17.6	28.2	26.5	378.8
1989	111.7	13.7	33.5	238.1	20.0	34.7	28.0	479.7
1990	135.7	14.5	36.0	248.2	21.4	30.5	29.7	515.8
1991	131.6	12.4	32.4	227.5	18.5	29.8	28.3	480.6
1992	145.3	15.8	31.3	272.1	21.0	35.4	36.2	556.9
1993	126.5	13.1	34.9	201.9	18.5	28.3	30.7	453.9
1994	149.7	15.2	33.4	285.1	21.0	35.3	29.8	569.5
1995	126.3	13.3	37.2	232.3	18.8	31.1	27.2	486.2
1996	153.8	18.4	38.1	253.7	21.0	34.7	29.4	549.2
1997	155.5	16.8	38.0	267.0	19.6	32.3	35.9	565.2
Average	131.0	15.3	34.9	238.1	20.7	33.3	30.9	504.2

Table 4. AIC values from several different model options attempted for weather and C input relationships. The quadratic relationship in the last row was chosen as the best model to predict the observed trends in C input rates. ppt= precipitation; tmean=mean monthly temperature; ratio= precipitation/ potential evapotranspiration; Prefix ‘S’ stands for “summer” to indicate the growing season and ‘Ann’ stands for “Annual”.

Model	CNP	DS	FW	NC	NE	SE	SP
Sppt Stmean	8482	-750	2186	16500	658	2863	13654
Sppt Stmean Sppt*Stmean	8466	-757	2172	16181	601	2748	13645
Stmean Annppt	8445	-695	2231	16499	636	2810	13862
Stmean Sratio Stmean*Sratio	8462	-769	2176	16021	574	2742	13679
Stmean Annppt Stmean*Annppt	8446	-694	2194	16460	627	2700	13808
Stmean Annratio	8462	-704	2231	16484	687	2833	13927
Stmean Sratio	8466	-757	2182	16498	684	2865	13732
Stmean	8486	-681	2229	16498	775	2881	13967
Sppt Stmean Sppt*Stmean $Sppt^2$ $Stmean^2$	7885	-791	2150	13208	400	2626	13545



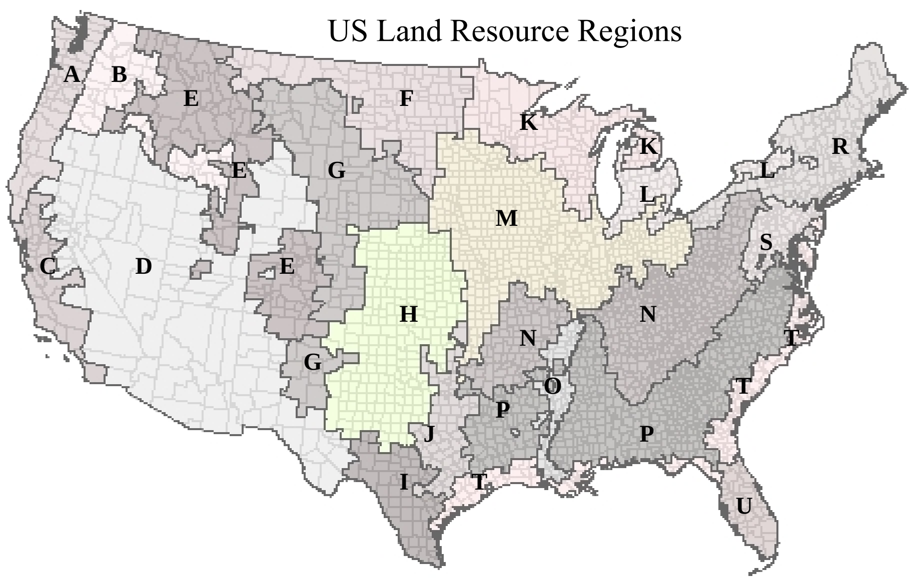


Figure 1. Counties categorized under different Crop Production Regions (CPRs) and Land Resource Regions (LRRs) of the US. The description of the CPRs is as follows: CNP- Central and Northern Plains; DS- Delta States; FW- Far West; NC-North Central; NE- North East; SE- South East; SP- Southern Plains

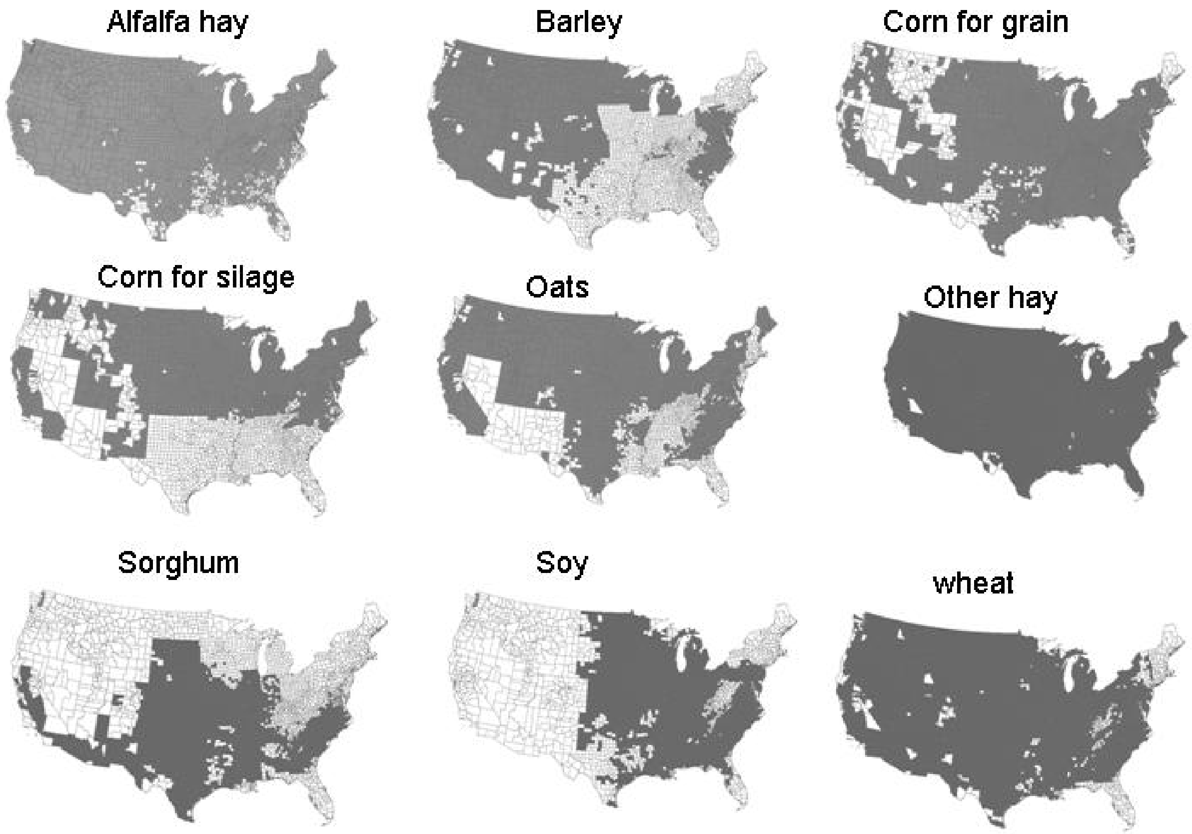


Figure 2. Counties where major crops of the US were present during the 16 year period, 1982- 1997. The data are from the comprehensive databases we developed in the first part of this study.

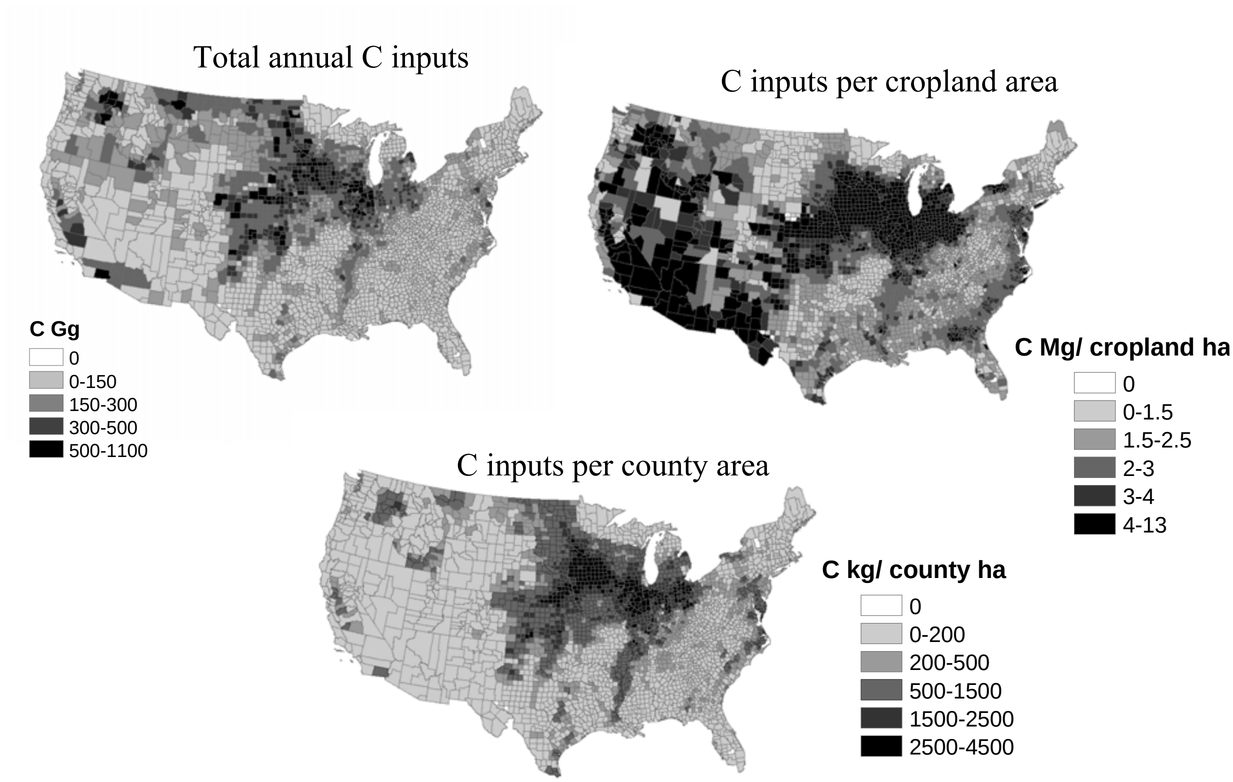
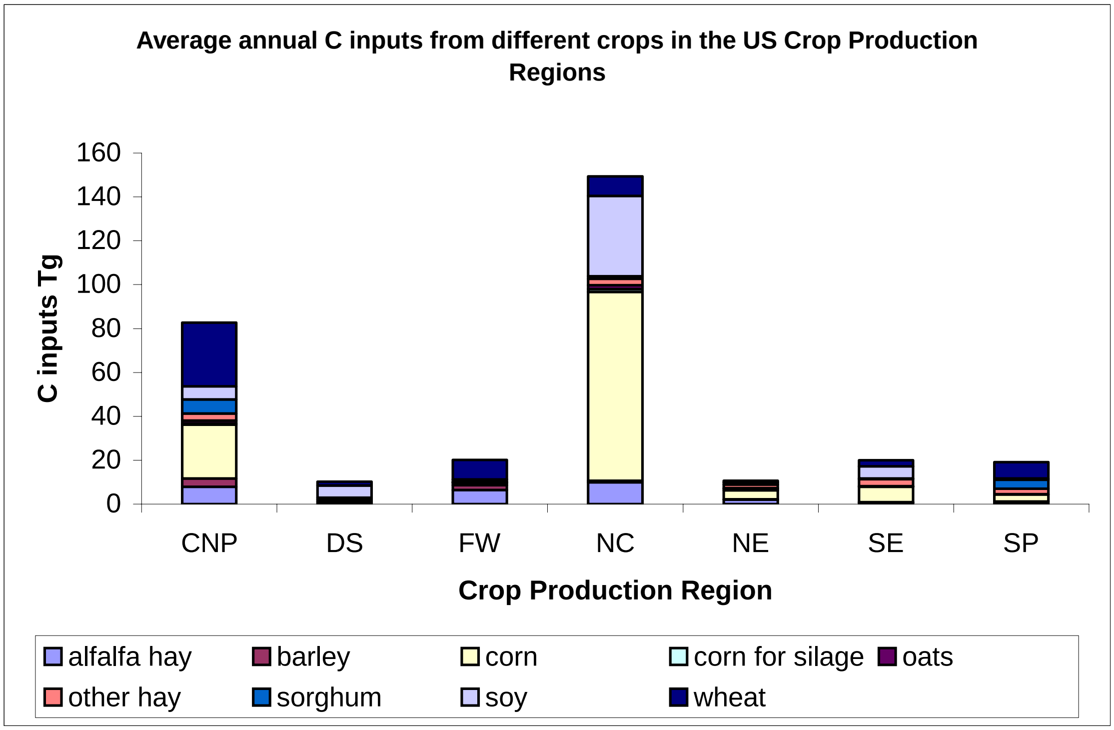


Figure 3. County-level total C inputs (Gg; top left), C inputs per cropland area (Mg ha^{-1}), and C inputs per county area (kg ha^{-1}) in the US during 1997.



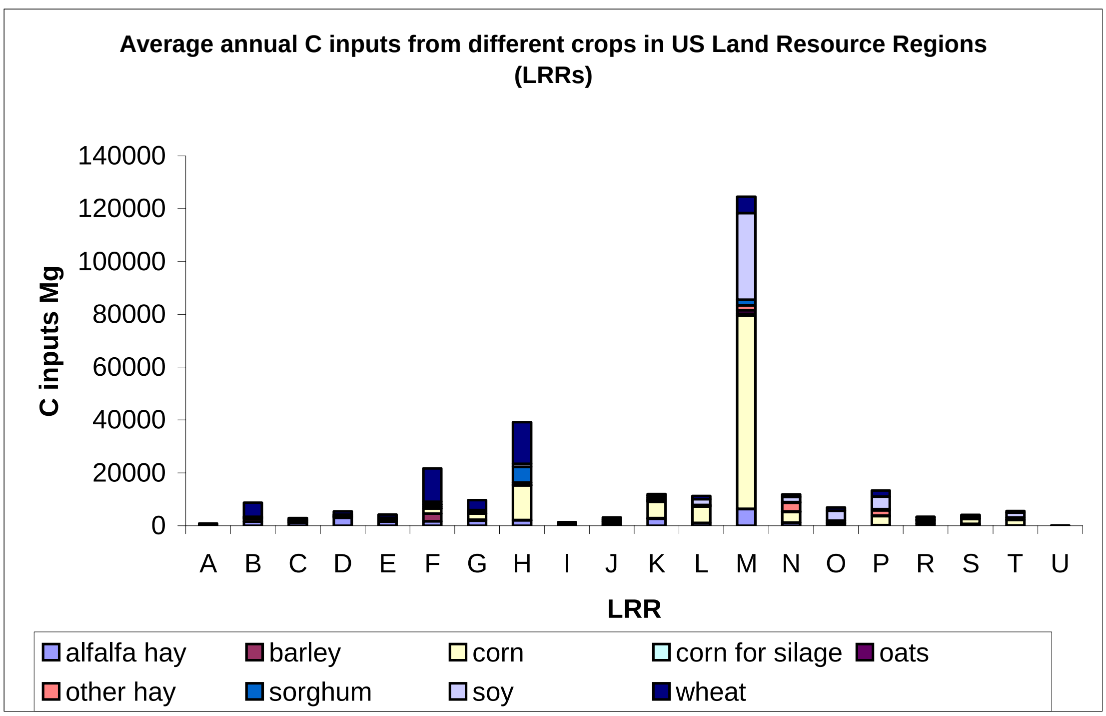


Figure 4. Spatial variability in average annual residue C inputs over the period 1982-1997, as shown for different a) Crop Production regions and b) Land Resource Regions (LRRs). See caption Fig. 1 for the description of the CPRs.

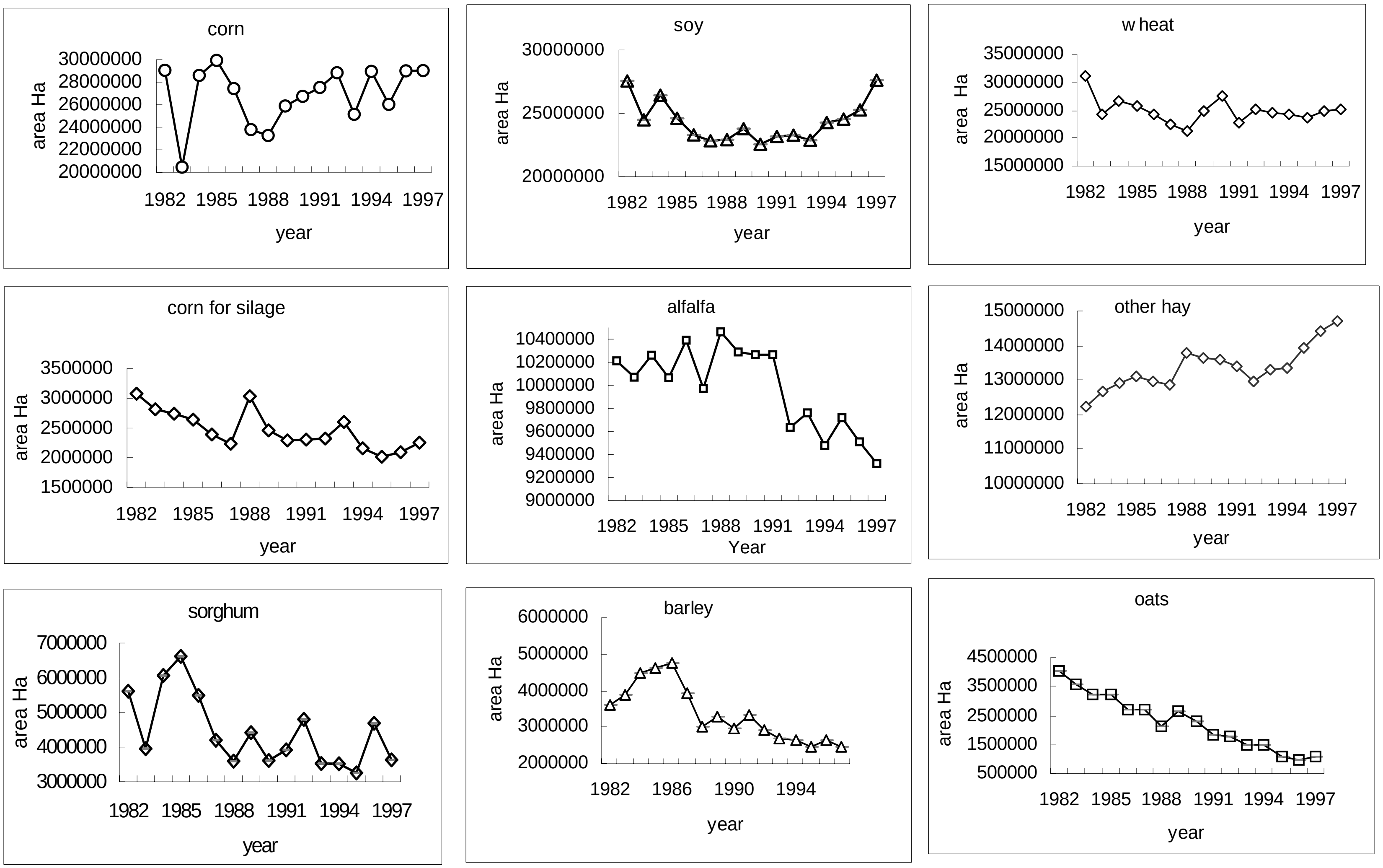


Figure 5. Temporal variation of total crop area under different crops (after filling the gaps)

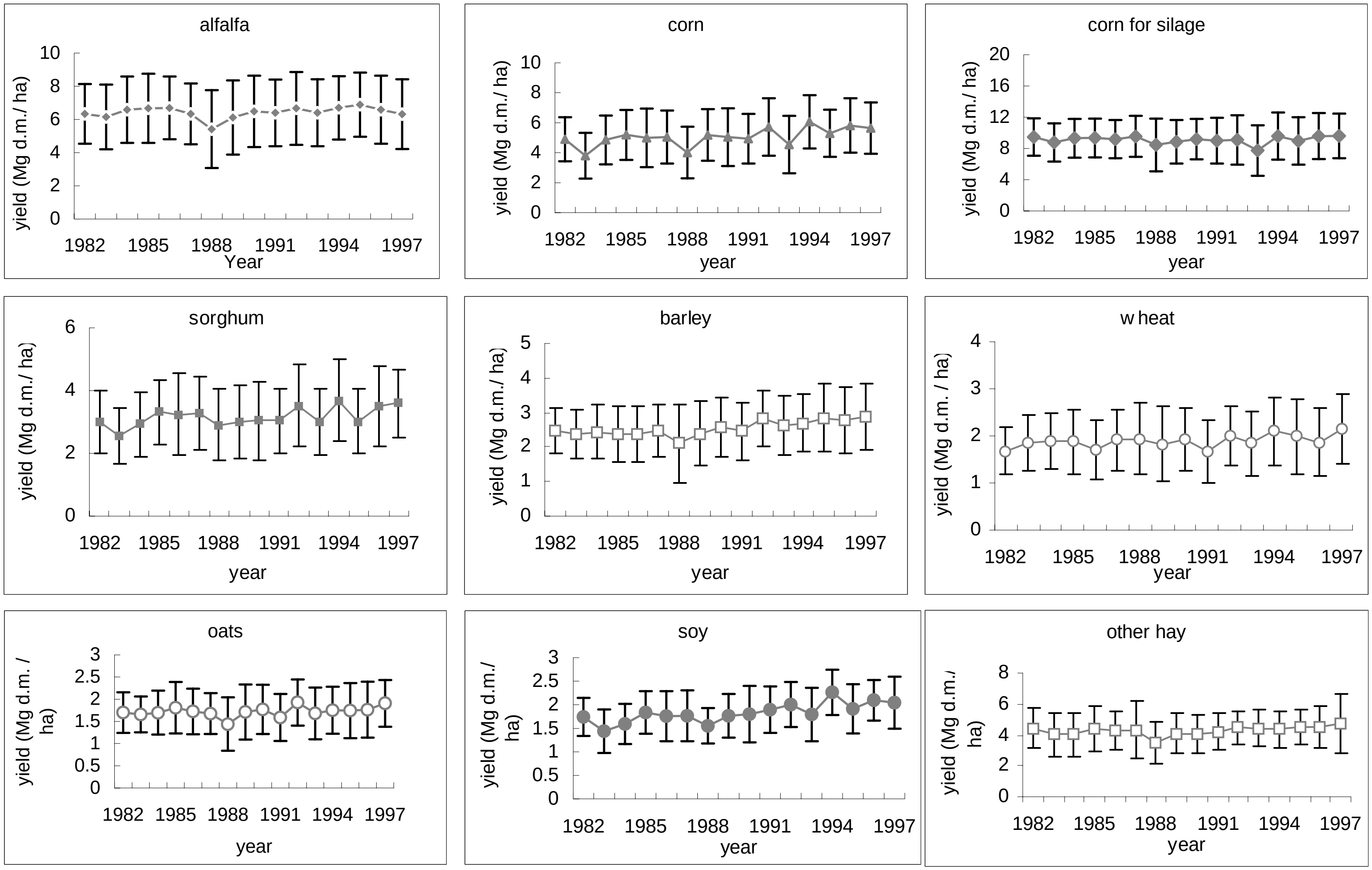


Figure 6. Variation of average yield dry matter (estimated using the county-level yield data) of alfalfa hay, barley, corn, oats, other hay, sorghum, soybean, and wheat over the 16-year period. The error bars represent the standard deviation.

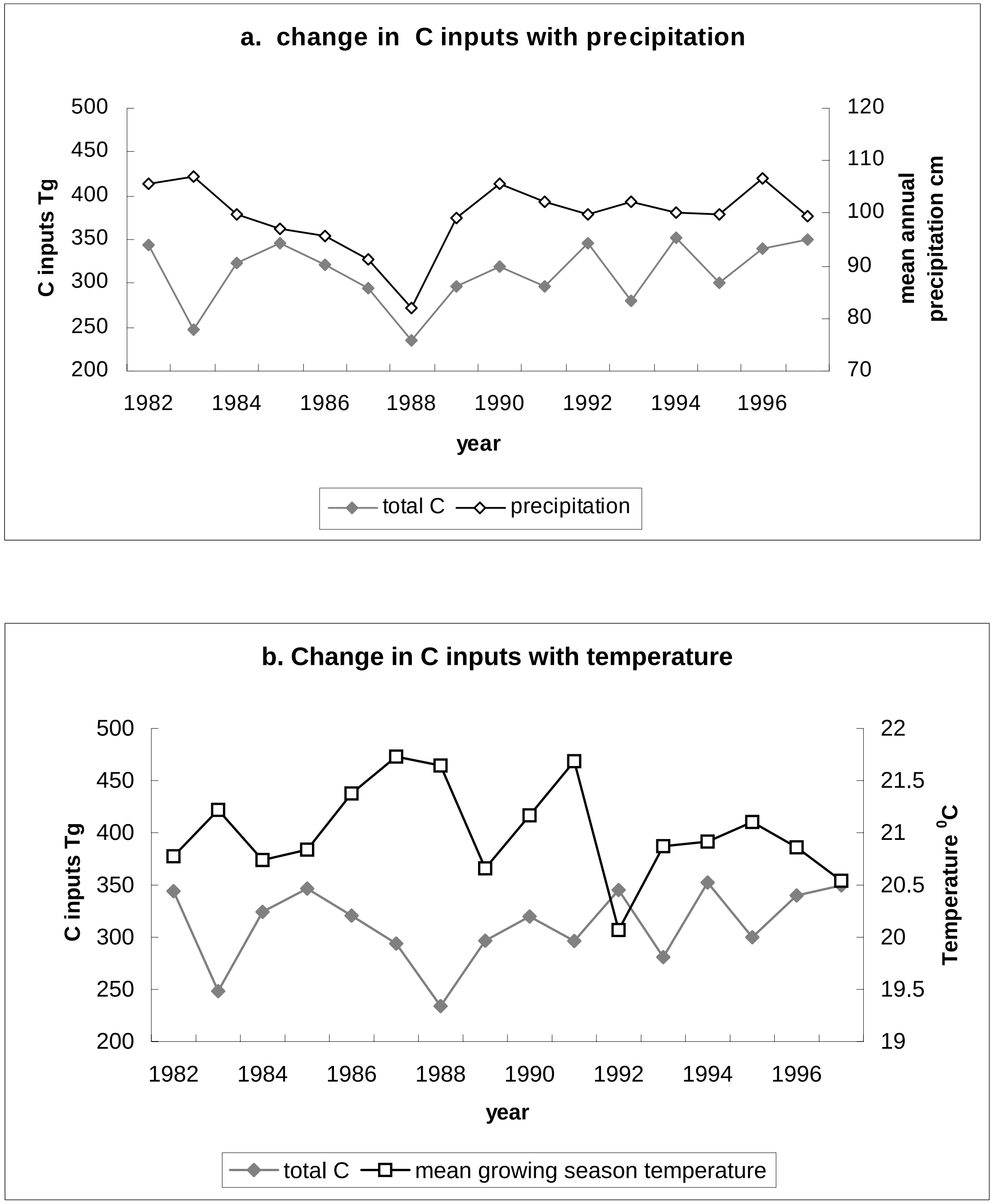


Figure 7. Temporal variation of total annual C inputs with a) mean annual precipitation and b) mean growing season temperature in the US agricultural soils. Similar variations were observed for carbon input rates (Tg per ha; Total C inputs for the whole US weighted by the crop areas aggregated from all the counties), too.

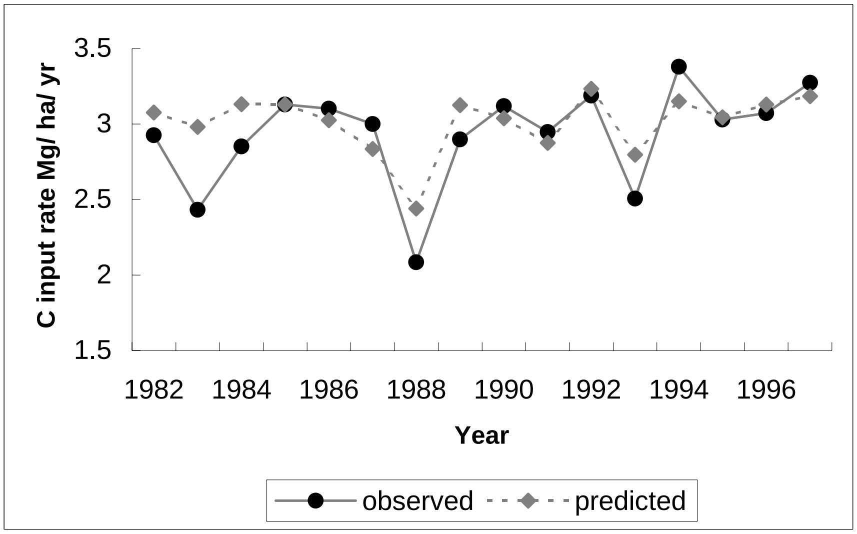


Figure 8. Interannual variability of observed C input rates and those predicted using the weather variables in a mixed model, for North Central region.

CHAPTER 6

C DYNAMICS IN US AGRICULTURAL SOILS ESTIMATED USING THE INTRODUTORY CARBON BALANCE MODEL (ICBM)

ABSTRACT

Carbon dynamics in agricultural soils depends mostly on the C inputs added from crop residues and the C output to the atmosphere from microbial decomposition of the added inputs. We studied C balance in US agricultural soils, using the Introductory Carbon Balance Model (ICBM), with C inputs estimated from county-level yield data as input to the model. We studied the spatial and temporal variability of soil C stocks in relation to changes in C inputs that mostly reflected the variation in production and weather over the study period.

Estimated soil C stocks in the permanent cropland ranged between 3.07 to 3.11 Pg (to a depth of 0-20 cm), which was mostly influenced by the interannual varation in the mix of crop species and residue C inputs. From the beginning to the end of the study period (i.e. 1982-1997), total soil C stocks increased by only 14 Tg, and the interannual variation observed for total C stocks was mainly dependent on the amount of residue C added from the previous year. The interannual variation in soil C stocks was much less than the

variation observed for the residue C inputs. The interannual (year to year) variation in total C stocks ranged within 20 Tg over the study period. Higher C stock rates (45.4 ± 0.1 Mg ha^{-1} yr^{-1}) were observed for irrigated soils (~ 10% of total cropland) compared to rainfed soils (36.2 ± 0.1 Mg ha^{-1} yr^{-1}) at national scale; however in certain regions (e.g. in the Far West and North Central regions) the opposite could be observed. We found that apparent Net Ecosystem Productivity (i.e. the sum of soil C stock change and the exported harvest) in the US cropland ranged between 14 and 50 Tg yr^{-1} during the period 1994-1997. Our study did not consider impacts from addition of any organic amendments or changes in tillage practices. The magnitude of the contribution from cropland soils in mitigating CO_2-C emissions could be even higher once we include these factors in the analyses.

INTRODUCTION

Carbon balance in terrestrial ecosystems has large implications on the variation in atmospheric carbon dioxide (CO_2) emissions. Atmospheric CO_2 has increased from 280 ppmv in 1850 to 374 in ppmv in 2003, mostly as a consequence of human disruption of the balance in the global carbon cycle, including land use activities. US agricultural soils is currently responsible for about 7% of the country's annual CO_2-C emissions (US EPA, 1999), and there is a significant potential for sequestration of C, especially in the temperate soils, by adopting appropriate improved management practices that help mitigate atmospheric CO_2 emissions (Lal et al., 1998; Paustian et al., 1997, 2002; Follett, R.F., 2001).

Plant litter including residues from croplands, and other organic amendments to soils undergo decomposition and release CO_2 through mediation of microbial activity, having significant influence on the C cycle at ecosystem level. The focus on global warming and its implications on overall carbon cycle has prompted more and more research for studying the changes in terrestrial ecosystem C cycling. Although there are means of direct estimation of soil C changes (e.g. measurements based on field sampling and lab analyses, isotope studies) and CO_2 emissions (Eddy covariance measurements), there is still not enough observed data to evaluate temporally and spatially varied C dynamics (i.e. changes in soil C stocks) at regional or national scale. Therefore most of the studies that have estimated such large scale C dynamics, have mostly used complex simulation modeling approaches (either top-down or bottom-up approaches) incorporating a range of input variables including statistics related to vegetation and/or remote sensing (Denning et al., 1995, Lobell et al., 2002, Potter et al., 1993, Potter, 1999, Parton et al., 1987, Paustian et al., 1995, 2002).

Modeling soil C dynamics and decomposition extends far back to few decades ago (e.g. Olson, 1963), and with recent emphasis on the potential role of soils in mitigating atmospheric CO_2, further development and refinement of models with special emphasis on soil C dynamics has increased (Parton et al., 1987, Jenkinson, 1990, Jenkinson et al., 1991, Andren and Kätterer, 1997, Andren et al., 2004, Del Grosso et al., 2005). These modeling approaches have been especially helpful for inventory purposes, when countries report annual total emissions and sinks under different sectoral categories, including agricultural soils. The short-term C dynamics and interannual variability of C stock changes have been a major concern in developing inventories and studying overall cropland contribution to the C cycle, and our aim of this study was to evaluate the US cropland C dynamics using the estimated comprehensive residue C input data in a simple simulation modeling appraoch.

We used the Introductory Carbon Balance Model (ICBM), originally developed for Swedish agricultural soils, in estimating C dynamics (i.e. overall C stock changes and interannual variability in soil C) of the US agricultural soils for a 16-year period, 1982-1997. ICBM is a simple 2-pool soil C model run with 5 parameters (i.e. annual carbon input rate, decomposition constants for the two carbon pools, an external influence coefficient, and a humification constant). The model estimates for Swedish agricultural C stocks have been validated against observed data from long-term experiments (Andren and Kätterer, 1997, Andren and Kätterer, 2001, Andren et al., 2004). This model has been applied for regions in Sweden (ICBMregion model; Andren and Kätterer, 2004), Canada and Africa (O.A. Andren, 2006, personal communication; Andren and Kätterer, submitted). The current study is the first study attempting the use of ICBM to study national scale C dynamics in the US.

MATERIALS AND METHODS

ICBM region (ICBMr) model description

The ICBMregion (ICBMr) model has two main soil C pools: a 'young' C pool (Y) and an 'old' C pool (O). The model incorporates annual carbon inputs *i* (Mg ha^{-1} yr^{-1}) and four other parameters: decomposition constant for Y (i.e. fraction of Y that decomposes in a year) k_Y (yr^{-1}), decomposition constant for O (i.e. fraction of O that decomposes in a year) k_O (yr^{-1}), external influence coefficient to represent influence from soil climate r_e (dimensionless), and humification coefficient (fraction that enters from Y to O, after subtracting the CO_2-C outflow) h (dimensionless). Since we considered the change in C stocks over a relatively short period, and since the aim of the study was mostly to study the interannual variability using the estimated residue C inputs, we assumed that the C stocks are at steady state at the beginning of the study period. Initial, steady state C stocks were calculated considering the mean residue C input rate over the 16-year period as the initial equilibrium C inputs, and the average r_e value for the period as the initial equilibrium r_e at the equilibrium state; the defaults values were used for k_Y, k_O, and h. The initial stocks were calculated using the steady state equations shown in Figure 1. Equations 1 and 2 were the model equations used in calculation of the young and old soil C stocks (Andren and Kätterer, 2004).

$$Y_t = (Y_{t-1} + i_{t-1})e^{-k_Y r_e} \qquad [1]$$

$$O_t = \left(O_{t-1} - h\frac{k_Y(Y_{t-1} + i_{t-1})}{k_O - k_Y} \right) e^{-k_O r_e} + h\frac{k_Y(Y_{t-1} + i_{t-1})}{k_O - k_Y} e^{-k_Y r_e} \qquad [2]$$

Y = Young C pool (Mg ha^{-1}); O= old C pool (Mg ha^{-1}); t= time (year); i= C input (Mg ha^{-1}); h = humification coefficient (dimensionless); k_Y= decomposition constant for Y ($year^{-1}$); k_O = decomposition constant for O ($year^{-1}$); r_e= external influence coefficient.

Model inputs and parameter estimation

County-level annual crop residue C inputs (Mg ha^{-1} yr^{-1}) for the US were estimated using a database of crop residue and yields for major crops in the US cropland for 1982-1997 (Lokupitiya et al., submitted). Crops included alfalfa (*Medicago sativa* L.) hay, barley (*Hordeum vulgaris* L.), corn (*Zea mays* L.) for grain, corn for silage and green chop, oats (*Avena sativa* L.), other hay (hay other than alfalfa; tame hay, small grain hay, wild hay), sorghum (*Sorghum bicolor* L.), soybean (*Glycine max* L.), and wheat (both winter wheat and spring wheat; *Triticum aestivum* L.).

The default values used as the decomposition coefficients in ICBMr were used for the current study, as well ($k_Y = 0.8$, $k_O = 0.006$), and $h = 0.13$; O. Andren, 2006, personal communication). The external influence coefficient, r_e was estimated using the weather to soil climate module *(W2r_e)* as detailed in Andren and Kätterer (2004). Parameter r_e incorporates three parameters representing soil water balance (r_w), soil temperature (r_t), and a cultivation factor (r_c) as shown in eq 3.

$$r_e = r_W * r_T * r_c. \qquad [3]$$

Estimation of r_c factor in ICBMr model is still in the testing phase (O.Andren, 2006, personal communication), and thus a default value of $r_c = 1$ was assumed for the current study. The r_e values using the other two components of (i.e. r_W and r_T) were estimated for different crop types (9) and soil types (11) of the US at county level. Soil data used (i.e. volumetric water content at field capacity and wilting point by soil type), and the percentage of each soil type at county level for the US were obtained from the STATSGO (http://www.ncgc.nrcs.usda.gov/products/datasets/statsgo/) database.

Estimation of r_W and r_T.

Soil water balance, r_W was estimated based on the concepts and equations derived from FAO guidelines for computing crop water requirements (Allen et al., 1998, Andren and Kätterer, 2004). The daily water balance in the soil was calculated using daily precipitation (*PPT*), actual evapotranspiration (actual evaporation from soil and actual transpiration from crops), crop interception, run off and percolation (Andren and Kätterer, 2004).

We calculated soil potential evapotranspiration (PET) using the Hargreaves ET_o equation (Hargreaves and Samani, 1985, Allen et al; 1998, Samani, 2000). The daily precipitation, minimum and maximum temperatures, solar radiation (W m^{-2}) at county level were extracted using the Daymet database (http://www.daymet.org/) for the PET calculation. For the counties with both irrigated and rainfed agriculture, irrigation water amounts for each crop were extracted from the Census of Agriculture, and added to precipitation, to estimate actual total water input to the irrigated cropland soil.

Crop coefficients (K_c, K_{cb}, and K_e; Allen et al., 1998) based on the growth cycle of individual crops were used in the estimation of interception. In calculating these coefficients, crop start date and end date information was compiled using the crop calendars published by USDA (1997); for perennial alfalfa, the start dates of the growth period within a year was compiled based on the last frost date in the spring.

Volumetric water content at field capacity and wilting point for different soil types in the US were used in calculating the relative water content and r_W. Air temperature from the Daymet database was adjusted to soil temperature using a simple empirical equation, and soil temperature was used in a quadratic relationship to derive r_T (Andren and Kätterer, 2004).

Model runs

Daily r_e values at county level were calculated for each crop by soil type, for a topsoil depth of 20cm, separately for irrigated crop areas and rainfed areas over a 17-year period from 1981-1997, and averaged to get annual r_e values for each category. In this study, we ran ICBM for combined C inputs from all the crops; thus the final annual r_e values from all crop and soil types were aggregated to two main r_e values: irrigated vs rainfed. Carbon stocks from the irrigated, rainfed, and total cropland in the whole US were estimated for the period 1982-1997, using the equations 1 and 2.

The results from the model runs were analyzed for the total C stock changes in relation to variation in the stocks in young and old pools. The estimated C stocks were validated against the C stocks reported in the pedon data (NRCS, 1997) from soil samples collected in different states; each pedon sample had specific information on the ecosystem type (e.g. forest vs cropland), soil series, and, other relevant information including the soil depth. To avoid any impact from the large variation seen in the crop areas over the 16-year period, the interannual and spatial variability of C stocks were estimated for the permanent cropland (i.e. the lowest cropland area observed for a given county during the entire time series, which was constant for the entire time series). Estimated C stocks during the 16-year period were analyzed considering the variation associated with weather changes, C inputs, and the practice (i.e. irrigated vs rainfed). Spatial variation was mostly analyzed at regional level (at Crop Production Region (CPR) level).

We also estimated the "apparent" Net Ecosystem Productivity (NEP; i.e., the sum of soil C stock change and the exported harvest) for the US cropland during the period 1994-1997. In this analysis we included the annual Net Primary Production (NPP), CO_2-C lost in decomposition, and the total C in the US grain export over the period. We assumed that the consumed grain within the country adds no C to agricultural soils, and exported grain would avoid a certain amount of C emission that would have otherwise occurred within the country. Therefore we considered that both soil C stock changes and the C in exported grain contribute in mitigating CO_2 emissions from the croplands. In this study we considered the sum of C in grain export and soil stock changes as the apparent NEP of the cropland ecosystem.

RESULTS

C stocks in different pools

The observed r_e values reflected the temperature and precipitation influences on decomposition, with the highest values in the South East region and irrigated crop areas in the west (Figure 2). The drought years (i.e. 1983, 1988) had the lowest r_e values during the period, reflecting the reduced decomposition during those periods, and the highest r_e values were observed during the years 1992, 1995, and 1997. The young pool had a relatively very low concentration of C compared to the old pool (Figure 3), and the C stocks in the young pool ranged from 86-109 Tg over the 16-year period; C stocks over 100 Tg could be observed during the years 1982, 1983, 1993, 1995, and 1997, despite the higher rates of decomposition as reflected in the higher r_e values observed in 1995 and 1997. The old pool had 25-35 times higher C stocks compared to the young pool, over the 16-year period. The variation in the old pool fell within a very narrow range in the permanent cropland, increasing only by 19 Tg from the beginning to the end of the study period. The overall variation in C stocks mostly reflected the variation in the young pool, despite the size of the pool.

Temporal and spatial variation in C stocks

Over the 16-year period, C stocks in the total US permanent cropland ranged between a minimum of 3073 Tg in 1984 to a maximum of 3105 Tg in 1997; this seemed to be mostly influenced by changes in the mix of crops and impacts from the interannual variation in weather (Figure 3). C stocks increased by 14 Tg C from the beginning to the end of the period in the permanent cropland with slight interannual variability.

Interannual variability of C stocks was relatively low compared to that in the residue C inputs (Figure 4). An alternating pattern in C inputs and C stocks could be observed in relation to changes in weather. For instance, the lowest C inputs were observed in 1988; however, the corresponding decrease in C stocks was observed in 1989. The year-to-year change in C stocks mostly fell within 20 Tg, however, that of C inputs showed significant dampening with a year-to-year variation that ranged between 2-80 Tg over the study period (Figure 4). Although there was significant variation between the C inputs and C stocks at crop area level (Figure 5), at entire county level, a similar pattern of spatial variation was observed for both C inputs and soil C stocks (Mg C ha^{-1} yr^{-1}; Figure 6).

Average C stocks in US irrigated cropland (45.4 ± 0.1 Mg C ha^{-1} yr^{-1}) were 9 Mg C higher than the rainfed cropland over the study period (36.2 ± 0.1 Mg C ha^{-1} yr^{-1}) (Figure 7). The irrigated crop area in the US permanent cropland was 74 Mha less than the rainfed crop area, making up 10% of the total permanent cropland area over the study period.

Significant spatial variation in rainfed and irrigated C stocks could be observed among the different Crop Production Regions (CPRs) in the US. Irrigated cropland in the CPRs were relatively very low except for Far West (irrigated land about 57% of the total cropland), Central and Northern Plains (about 16%), Southern Plains (about 13%), and Delta States (12%). Irrigated areas in North Central (0.4%) and North Eastern (<0.001%) regions were very low. Although the overall country average showed higher C stock values for irrigated land, region wise the same trend could not be observed (Table 1). Overall, at CPR level, highest C inputs and stocks were observed in the North Central, Far West, and Central and Northern Plains regions (Table 1 and figures 5 and 6).

Comparison of the estimated C stocks with C in pedon samples

C stocks in Pedon data (NRCS, 1997) for a number of samples available for a 20 cm depth from 15 states were compared against the ranges estimated C stocks for those states (Table 4). The estimated C stocks at county level showed a relatively large variation. Thus the range estimated for the counties in a state was compared against the range observed for the number of soil samples in each state with pedon data. The estimated range most of the time fell within the observed range (Table 2).

Overall cropland NEP and C stock change

Net Ecosystem Productivity for the US cropland includes the C uptake in NPP minus any C that would exit the system. In ICBM, a fraction the previous year's input (i.e. the fraction left after microbial respiration in decomposition) becomes incorporated as part of the current years observed C stocks. In this study we estimated the apparent NEP (i.e. NPP- sum of the C stock changes and grain exports) in the permaneent cropland area in the US (Table 3). The apparent NEP estimated for the total cropland ranged between 14 to 50Tg yr^{-1} during the period 1994-1997.

DISCUSSION

In this study we predicted the total C stocks at county-level for the whole US. The aim of the study was to detect how much the soil C stocks in the US cropland would change due to addition of residue C. We did not consider any organic (manure) amendments or changes in tillage practices in this study. Overall model estimated C stocks were compared with the C stocks from soil samples reported in the pedon database by the NRCS, and we studied the spatial and temporal variation of C stocks in the US cropland soils, in comparison to variation in the residue C inputs added from the major crops.

Average C stocks observed for the total permanent cropland (89.4 Mha) in the US were about 3085 Tg during the 16-year period. The observed interannual variation in C stocks were mostly due to variation in the young C pool, the soil pool mostly impacted by the residue C inputs added in the previous year. Thus any change in the residue C inputs added during a given year, was mostly reflected in the young soil C pool in the following year. However, this trend was masked in the total C stocks in the total cropland in a given year, since the variation in total cropland mostly reflected the area under the cropland during that year. Therefore, for all the analyses performed in this study, we considered the permanent cropland of the US to avoid any variation in crop area.

In our analyses, we found that there was a lag between the variation in C inputs and C stocks; high residue C inputs in a given year was reflected in the high C stock in the following year (Figure 4). Thus high C stocks (over 100 Tg) could be observed during the years (i.e.1983, 1993, 1995, and 1997) that were preceded by high productivity and residue C inputs to agricultural soils (as described in the preceding chapter). This trend also reflected the impact from the increased C inputs due to increased production towards the end of the period. Similarly the years with low C inputs (e.g. 1983, 1988, and 1993) were followed by years that had lower C stocks. The results confirmed the validity of the external influence coefficient ($r_{e)}$ values used in predicting any impact on decomposition from soil climate that was driven by daily variation in weather.

Another noticeable observation for the interannual variability of C inputs and C stocks was the less variation in C stocks compared to the interannual variability observed for the residue C inputs (Figure 4). Except for a high parallel increase shown in 1993 for the high residue C input added in 1992, the year-to-year variation in C stocks was about half or less in maginitude, compared to the year-to-year variation observed for C inputs in most of the years. C stocks in the permanent showed a slow but gradual increase, with 14 Tg higher C stocks towards the end of the period (compared to the beginning).

The estimated C stocks clearly predicted the expected outcome for the spatial variation in the C stocks among the different CPRs. For instance, in the Far West region, where about 57% of the cropland in irrigated, C stocks were relatively higher compared to the other regions; however, in the same region, higher C stocks could be observed for the rainfed cropland (although the inputs were higher in the irrigated cropland; Table1). This was observed especially for the Northern Pacific Coast Range, Foothills, and valleys in Washington and Oregon, where crops are grown under substantial soil moisture.

Similarly, in the North Central region where most of the cropland is rainfed (only about 0.3% is irrigated), irrigated land had much lower C stocks compared to the rainfed cropland. Thus, although the overall country had higher C stock rates (Mg ha^{-1} yr^{-1}) in irrigated cropland, region wise the opposite trend could be observed for certain regions; however this did not seem to have much effect at national scale, due to very low cropland under irrigation, in general.

Due to high variability observed among the county-level C stocks within a given state, and lack of sufficient soil samples within the pedon database, the estimated C stocks could be validated marginally.   Overall, the ranges estimated seemed to tally with the observed (measured) low C stocks in certain states in the North East and South East regions (e.g, KY, MD, and NJ). Although relatively high C stocks were found in the North Central region, the estimated range from ICBM was much narrower (lower) compared to the observed range (Table 2). Due to high range observed for C stocks among the pedon samples, no proper spatial variation could be detected, other than the relatively low C stocks observed for certain states in South East and North East CPRs, and high stock values for the North Central (as described above) and Central and Northern Plains. Lack of substantial representation of all the CPRs in the soil samples used, also restricted any proper comparison for spatial variability of C stocks over the country.

The permanent cropland of the country ranged between 70- 85% of the total cropland during the study period. According to DOE/EIA (2005), the mineral soils in US cropland were accumulating 52.4-51.7 Tg CO_2 (i.e. 14.3-14.1 CO_2-C) during 1990- 1997. Our estimates for the same period ranged between 10-11 CO_2-C Tg, which was only slightly lower than the values reported by DOE/EIA (2005), as the agricultural mineral soil C sink (2004). Therefore, overall, the model seems to perform faily well, and it seems to correctly predict the expected interannual- and spatial variability in C stocks.

We found that the NEP, or the apparent NEP in the total US cropland during the period 1994- 1997 was about 1-4 percent of the estimated CO_2-C emission from fossil fuel burning (USEPA, 2002; table 3). Soil C stock change was about 20-30% of the apparent NEP (Table 3). Thus our study found the impact of both soil C stocks and the overall cropland productivity (i.e. implications from high NPP and C inputs), on overall C cycling in the US cropland.

Estimated CO_2-C release due to residue decomposition in our study (220- 240 Tg) in the permanent cropland) was only slightly higher than the range illustrated by the USDA (2003; i.e. 200- 220 during the period 1990- 1995). Since we did not take into account any other organic amendments to soils, and any influence from different crop management practices, the estimates from the current study could be considered as the baseline estimates for the apparent NEP and cropland C stocks for the US.

CONCLUSION

The ICBM model seemed to well predict the interannual and spatial variability in C stocks in the US cropland soils during the study period, although the C stocks predicted seem to be slightly lower compared to the estimates by the DOE/EIA (2005). The total soil stock predicted for the 16-year study period was about 3.1 Pg C in a depth of 20 cm in the cropland soils. The interannual variability in soil C stocks based on permanent cropland estimates ranged within 20 Tg over the study period, and it was much lower compared to the interannual variation observed for crop residue C inputs. Based on our study, apparent NEP estimates calculated for the US cropland ranged between 14-50 Tg during the period 1994- 1997, implying the importance of C dynamics in the cropland soils, in the CO_2 mitigation potential of the country, and in overall C cycle. Although the interannual variability in C stocks occurred within a narrow range, the stocks clearly reflected the influences of the production (i.e. variation in the C inputs added), and associated changes in weather. Our estimates are slightly lower than the studies that have predicted the “potential” C sequestration in the US agricultural soils. We considered a constant crop area (only the permanent cropland) for this analysis, and did not take into account any influence from specific management practices or the impacts in the land use changes that occurred in the croplands (e.g. CRP) within the same period of time. Thus findings of our study reflect more of “detrended”, baseline C stock estimates, that would reflect the existing short-term C dynamics in croplands that are useful in climate change policy and inventory purposes.

Acknowledgments

This research was supported by the Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture, under Agreement No. 2001-38700-11092 to the Consortium for Agricultural Soils Mitigation of Greenhouse Gases (CASMGS). We would like to thank Mark Easter and Steve Williams for their support.

REFERENCES

Allen, R.G., L.S. Pereira, D. Raes, and M. Smith, M. 1998. Crop evapotranspiration – guidelines for computing crop water requirements. FAO irrigation and drainage paper 56. FAO, Rome, Italy.

Andrén, O., and T. Kätterer. 1997. ICBM- the Introductory Carbon Balance Model for exploration of soil carbon balances. Ecol. Applications 7(4),1226-1236.

Andrén, O., and T. Kätterer. 2001. Basic principles of soil carbon sequestration and calculating dynamic country-level balances including future scenarios. p 495-511. *In* R. Lal, J.M. Kimble, R.M. Follett, and B.A. Stewart. (ed.) Assessment Methods for Soil Carbon. Lewis Publishers, Boca Raton, FL.

Andrén, O., T. Kätterer, and T. Karlsson. 2004. ICBM regional model for estimations of dynamics of agricultural soil carbon pools. Nutr. Cycl. Agroecosys. 70:231-239.

Del Grosso, S.J., Parton, W.J., Mosier, A.R., Holland, E.A., Pendall, E., Schimel, D.S., and D.S. Ojima. 2005. Modeling soil CO_2 emissions from ecosystems. 2005. Biochemistry 73, 71-91.

Denning, A.S., I.Y. Fung, and D.A. Randall. 1995. Latitudinal gradient of atmospheric CO_2 due to seasonal exchange with land biota. Nature 376, 240-243.

DOE/EIA. 2005. Emissions of greenhouse gases in the United States 2004. Department of Energy/ Energy Information Administration, Washington, D.C.

Follett, R.F. 2001. Soil management concepts and carbon sequestration in cropland soils. Soil & Tillage Research 61, 77- 92.

Hargreaves, G.H., and Z.A. Samani. 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric. 1(2), 96–99

Jenkinson, D.S. 1990. The turnover of organic carbon and nitrogen in soil. Philosophical Transactions of the Royal Society of London. Series B, Biological Sicences 329, 361-368.

Jenkinson, D.S., D.E. Adams, and A. Wild. 1991. Model estimates of CO_2 emissions from soil in response to global warming. Nature 351, 304-306.

Lal, R., Kimble, J., Follet, R.F., Cole, C.V., 1998. The potential for U.S. cropland to sequester carbon and mitigate greenhouse effect. Ann Arbor Press, Chelsea, Michigan.
Lobell, D.B., J.A. Hicke, G.P, Asner, C.B. Field, C.J. Tucker, C.J., and O. Los. 2002. Satellite estimates of productivity and light use efficiency in United States agriculture, 1982-98. Glob. Change Biol. 8, 722-735
NRCS, 1997. National Survey Laboratory Characterization Data. Digital Data. USDA-NRCS, Lincoln, NE. Available at http://nasis.usda.gov/downloads/home.shtml. (accessed 25 June, 2006; verified 02 July, 2006). Natural Resources Conservation Service, U. S. Department of Agriculture.
Olson, J.S. 1963. Energy storage and the balance of producers and decomposers in ecological systems. Ecology 44(2): 322-331.
Parton, W.J., D.S. Schimel, C.V. Cole, and D.S. Ojima. 1987. Analysis of factors controlling soil organic matter levels in Great Plains Grasslands. Soil Sci. Soc. Am. J. 51, 1173- 1179.
Paustian, K., E.T. Elliott, H.P. Collins, C.V. Cole, and E.A. Paul. 1995. Use of a network of long-term experiments for analysis of soil carbon dynamics and global change: the North American model. Aust. J. Exp. Agr. 35: 929-39
Paustian, K., H.P. Collins, and E.A. Paul. 1997. Management controls on soil carbon. p. 15- 49. *In* E.A. Paul, E.T. Elliott, K. Paustian, and C.V. Cole (ed.), Soil Organic Matter in Temperate Agroecosystems. CRC Press, New York.
Paustian, K., J. Brenner, K. Killian, J. Cipra, S. Williams,S., E.T. Elliott, M.D. Eve, T. Kautza, and G. Bluhm. 2002. State-level analyses of C sequestration in agricultural soils. p 193- 203. *In* R. Lal, J.M. Kimble, and R.F. Follett (ed.) Agricultural practices and policies for C sequestration in soil, New York, Lewis Publishers.
Potter, C.S. 1999. Terrestrial biomass and effects of deforestation on the global carbon cycle. BioScience 9(10): 769-778.
Potter, C. S., J. T. Randerson, C. B. Field, P. A. Matson, P. M. Vitousek, H. A. Mooney, and S. A. Klooster, 1993. Terrestrial ecosystem production: A process model based on global satellite and surface data. Global Biogeochem. Cycle. 7, 811–841
Samani, Z. 2000. Estimating Solar Radiation and Evapotranspiration Using Minimum Climatological Data. J. Irrig. and Drain. Engrg. 126(4), 265-267
USDA, 1997. Usual planting and harvesting dates for field crops. USDA.
USEPA. 2002. In brief: The US greenhouse gas inventory. U.S. Environmental Protection Agency. Washington, D.C.
USDA. 2003. U.S. Agriculture and Forestry Greenhouse Gas Inventory: 1990-2001. United States Department of Agriculture.

Table 1. Average C inputs and stocks (Mg ha^{-1}) in Crop Production Regions (CPRs) in the US cropland (i.e. all crops)

	CNP		DS		FW		NC		NE		SE		SP	
	Input	Stock	Input	Stock	Input	Stock	Input	Stock	Input	Stock	Input	Stock	Input	Stock
Rainfed	1.8 ± 0.2	35 ± 0.1	1.8 ± 0.2	15 ± 0.1	2.2 ± 0.2	75 ± 0.3	3 ± 0.4	43 ± 0.2	2 ± 0.2	29 ± 0.2	1.9 ± 0.2	19 ± 0.2	1.6 ± 0.1	19 ± 1.6
Irrigated	3.6 ± 0.3	52 ± 0.7	2 ± 0.1	15.2	3.5 ± 0.2	44 ± 0.2	0.9 ± 0.2	10 ± 0.1	0.9 ± 0.2	12 ± 0.1	1.6 ± 0.2	19 ± 0.2	3.5 ± 0.2	39 ± 3.5
Overall	2.1 ± 0.2	37 ± 0.2	2 ± 0.2	15 ± 0.1	2.9 ± 0.2	58 ± 0.2	3 ± 0.2	43 ± 0.2	2 ± 0.2	29 ± 0.2	1.9 ± 0.2	13 ± 0.2	1.6 ± 0.1	21 ± 1.9

Table 2. Estimated and observed (field measured) ranges of state-level soil C stocks (Mg ha^{-1}) for a depth of 20 cm.

State	CPR	Estimated Range	Observed range	No. of samples
IA	NC	26- 55	44- 78	4
IL	NC	18- 62	20- 119	55
IN	NC	19- 52	32- 46	3
KS	CNP	14- 52	14- 48	16
KY	SE	10- 33	26- 42	3
MD	NE	20- 34	15- 44	8
MI	NC	21- 62	61- 104	4
MN	NC	17- 53	41- 149	6
MO	NC	10- 36	24- 44	11
NE	CNP	12- 72	18- 66	21
NH	NE	20- 29	69- 102	3
NJ	NE	10- 36	16- 30	4
OH	NC	16- 50	42- 102	8
VT	SE	21- 26	34- 66	10
WI	NC	21- 51	17- 87	16

Table 3. Values of the variables relevant to in NEP calculation for **permanent cropland** only.
NPP: Net Primary Production; grain CO_2-C: CO_2 respired from consumed grain; FF: Fossil fuel; NEP: Net Ecosystem Productivity.

Year	NPP Tg	C inputs Tg	Grain C	Grain Export Tg	grain C consumed	CO2_C from soil respiration Tg	total C_stocks Tg	CO2 from FF burning Tg	Apparent NEP Tg (C stock change + export)
1994	456.8	207.4	249.4	27.4	222.0	220.4	3099.8	1353.6	14.4
1995	400.9	250.7	150.2	38.8	111.4	240.5	3086.8	1366.1	49.1
1996	430.7	219.3	211.4	34.1	177.2	222.3	3097.1	1416.8	31.2
1997	437.9	235.7	202.2	29.8	172.3	224.6	3094.1	1435.4	40.9

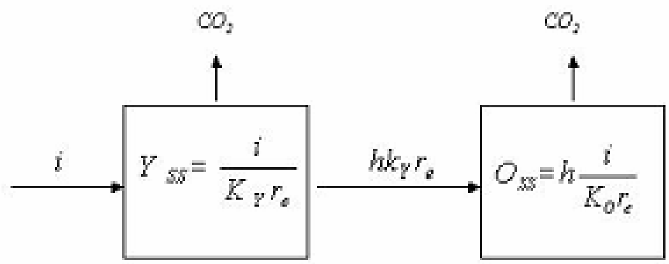


Figure 1. Model structure of the Introductory Carbon Balance Model (ICBM). i = annual carbon inputs (Mg ha^{-1}), k_Y = decomposition constant for 'young' organic matter, Y (i.e. fraction of Y that decomposes in a year; yr^{-1}), k_O = decomposition constant for 'old' organic matter, O (i.e. fraction of O that decomposes in a year; yr^{-1}), r_e = external (i.e. climate, management) influence coefficient (dimensionless), and h = humification coefficient (dimensionless). Y_{ss} and O_{ss} denote the steady state condition for Young and Old soil C pools (Source: Andren and Kätterer, 2001, 2004).

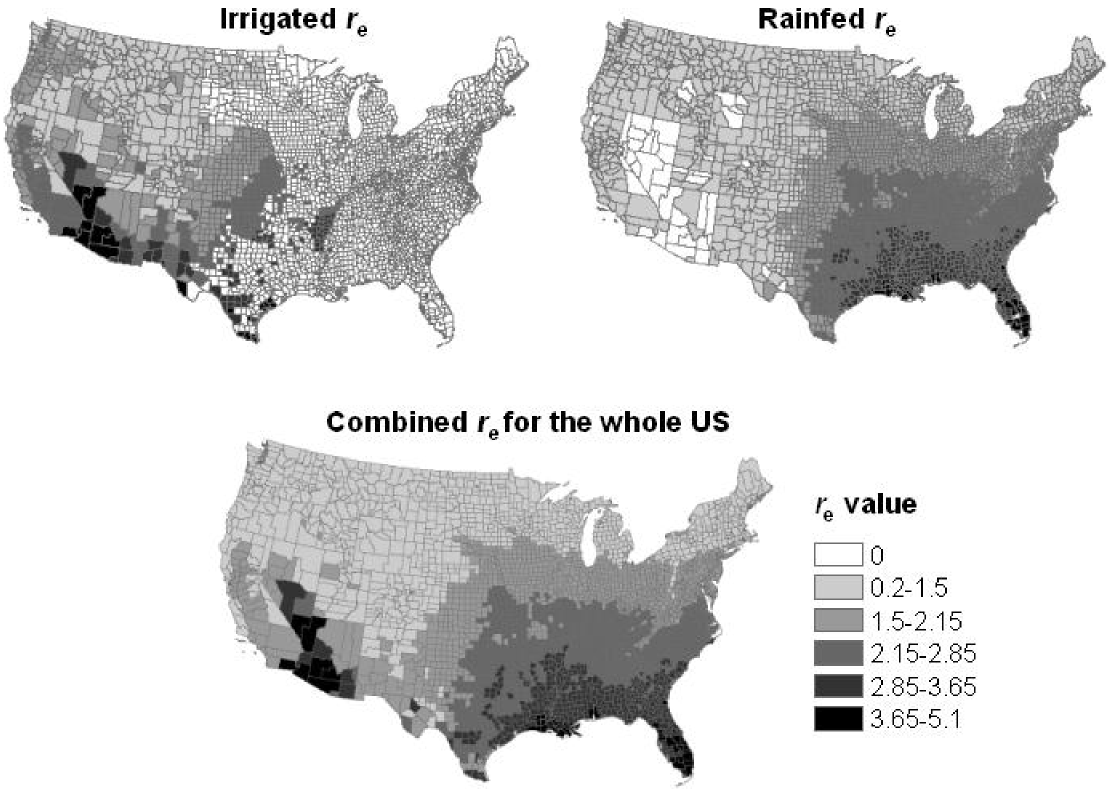


Figure 2. External influence coefficient (r_e) estimated from irrigated (top left), rainfed (top right), and area weighted average (using combined rainfed and irrigated areas; bottom).

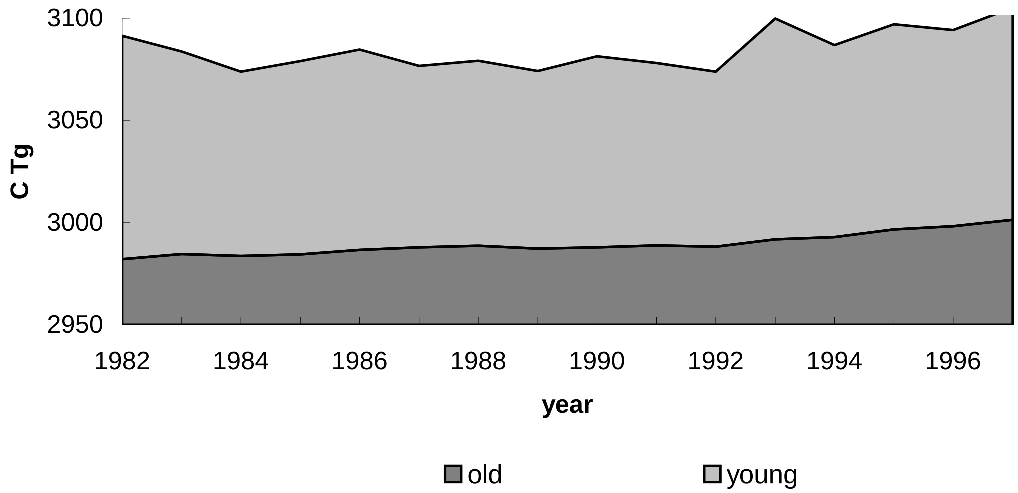


Figure 3. Variation in young, old and total C pools in the US permanent cropland. The variation in the C stocks in the permanent cropland reflected the variation in the young pool.

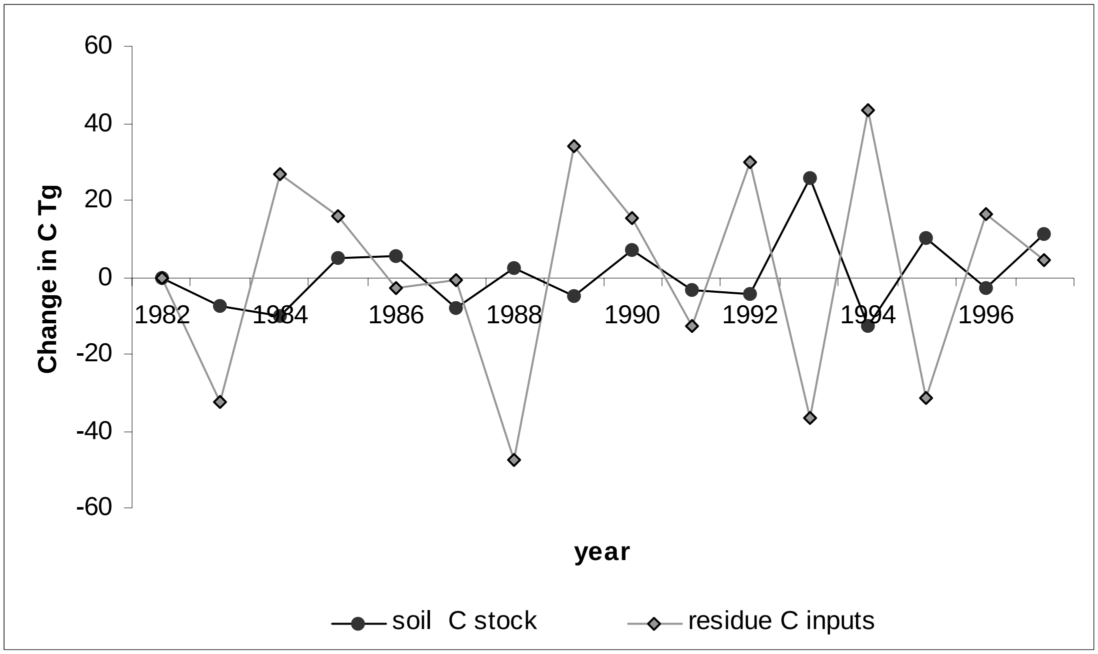


Figure 4. Interannual (year to year) variation in C inputs and soil C stock in the permanent cropland in the US

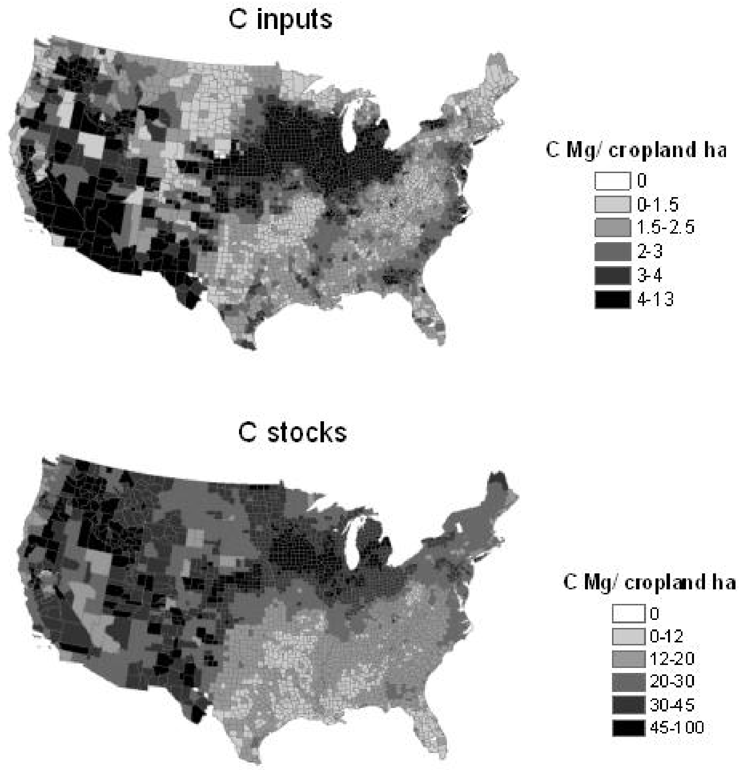


Figure 5. The spatial variation and magnitude of C inputs and stocks per cropland ha in 1997

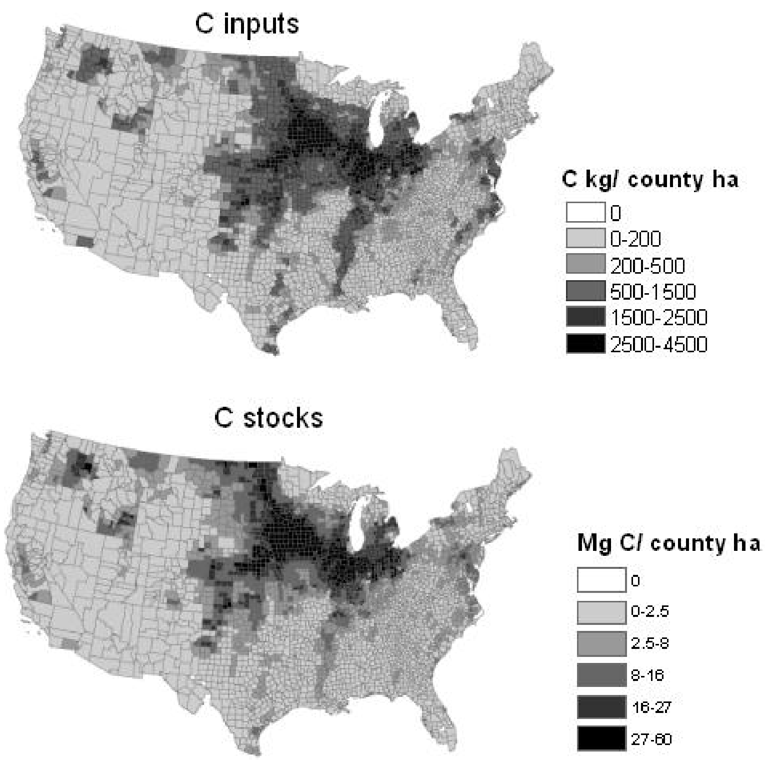


Figure 6. The magnitude of C inputs and stocks per county ha in 1997.

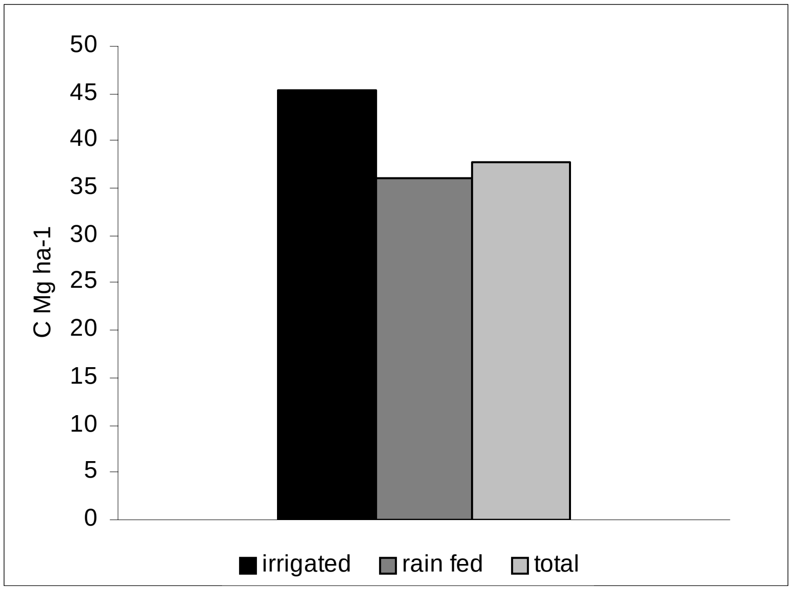


Figure 7. C stocks (Mg ha^{-1}) in the US irrigated, rainfed and total cropland.