Title: Using Social Budgeting Web Software and Natural Resources Software Models to Improve Agricultural Economics Data Collection, Dissemination, and Analysis

Authors: K. P. Boyle*, P. Heilman, R. W. Malone, L. Ma, and R. S. Kanwar

DevTreks Working Paper 01. January, 2012

Kevin P. Boyle is an agricultural economist and President of DevTreks (a nonprofit organization).
Philip Heilman is Research Biologist with the USDA-Agricultural Research Service, Southwest Watershed Research Service, Tucson, AZ.
Robert W. Malone is Agricultural Engineer with the USDA-Agricultural Research Service, National Soil Tilth Laboratory, Ames, IA.
Liwang Ma is a Soil Scientist with the USDA-Agricultural Research Service, Agricultural Systems Unit, Fort Collins, CO.
Ramesh S. Kanwar is Professor and Chair of Agricultural and Biosystems Engineering, Iowa State University, Ames, IA.

*Corresponding author
Kevin P. Boyle
DevTreks
1930 NW Irving St. #402
Portland, OR 97209-1267
503-796-7949
devtrekkers@devtreks.org

Abstract

Agricultural economics needs more and better data. This paper proposes a new approach, social budgeting, for building and analyzing agricultural economic datasets. Social budgeting employs modern information technologies, such as social networks and cloud computing data centers, to collect and analyze economics data over the Internet. DevTreks, an open source, social budgeting web software program, is introduced as an example. A case study involving Midwestern corn soybean production uses DevTreks and the process-based Root Zone Water Quality Model (RZWQM) to illustrate how modern information technology can be used to study economic-pollutant tradeoffs. The paper recommends starting or joining open source, on-line efforts aimed at integrating the physical, chemical, biological, and economic aspects of agriculture.

Keywords: social budgeting, social networks, cloud computing, open source software, economic data, economic services, Nitrogen loading, tile drainage.

1. Introduction - the Emperor has no clothes

Agricultural economics data is hard to collect, disseminate, and use. Most practitioners rely on data that, to varying degrees, limits the full usefulness of their findings. Much of this data suffers from lack of standardization, antiquated software, single user bias, and outdated approaches to information technology (IT) data collection, dissemination, and analysis. Today, most data collected by different farm managers, researchers, or extension staffs cannot be used to support modern computing (the aggregation and analysis of data, by people and machines, in a consistent manner). The data difficulties are compounded by analysts' reliance upon disparate climatic, landscape, demographic, agronomic, and economics data sets.

The problem of inconsistent, incompatible, and missing economics data is not new. In 1991, the Applied and Agricultural Economics Association (AAEA) convened a national task force consisting of 68 economists (the National Task Force on Commodity Costs and Returns Measurement Methods) to deal with aspects of the problem. Their resultant publication, *Commodity Costs and Returns Estimation Handbook* addressed the Task Force's mission to "recommend standardized practices for generating costs and returns estimates ..." (Hallam et al., 1999).

Ten years later, Just and Pope (2001) reviewed the current state of agricultural production data availability and use. They highlighted the data and analytic challenges arising from agriculture's unique dependence on sequential production stages, weather, pests, diseases, lags between the application of inputs and harvest, crop rotations, soils, and farmer abilities. Just and Pope noted that "readily available data tends to consist of highly aggregated public data in which temporal detail (within a growing season) and spatial data (among plots, farms ...) is lost". They concluded, based largely on economists' use of poor quality data, that "the current state of empirical knowledge of agricultural production sums up to little more than an empty box." That conclusion led Just and Pope to the following proposition (p. 721):

> We propose that a significant and complete data base for (…) agricultural production needs to be developed as an investment by/for the agricultural economics profession, … Such a data base could facilitate investigation of many issues identified by this study as blocked by data unavailability. By comparison, the current proliferation of studies with uncommon data bases and incongruent maintained hypotheses has led to endless speculative explanations of differences in results with little comprehensive comparison.

No such complete agricultural economics database has been built in the intervening years. Quotes from four recent articles in agricultural economics publications show that the quality of economics data is still a major issue: "(E)stimates of net farm income derived from (ARMS [Agricultural Resource Management Survey]) (may be) biased downward ... it has been recognized that farm income is generally underreported to the IRS ..." (Key and Roberts, 2009) . "(O)f course, crop budgets and production costs are notoriously difficult to measure" (Goodwin, 2009), "(T)he (economics) evidence is mainly anecdotal ... (Coleman, 2009). "More research and data are needed ... More research and data are also needed ... More research and data are needed ... More research and data are also needed ..." (Bergstrom and Ready, 2009). It should be noted that just about any article in either of the two publications cited, the *American Journal of Agricultural Economics* and the *Review of Agricultural Economics*, could have been quoted to highlight this issue.

Resource and environmental economics analysis also relies on biophysical and economic data and may also be suffering from some "empty boxes". Weyant (2008) notes "(the wide) uncertainty in mitigation cost projections ... is not surprising ... assumptions (must be made over long periods) about productivity growth, fuel prices, technology diffusion, the development of new technologies, and the interest rate/discount rate used in making intertemporal investment decisions". A recent report on climate change (IPCC, 2007) found that "The literature on (climate change) adaptation costs and benefits remains quite limited, and fragmented in terms of sectoral and regional coverage." A contributor to the Ecological Society of America's weblog (Hollister, 2007) wrote: "ecologists and environmental scientists must take a more active role in providing access to both data and the analytical techniques used to analyze those data. As our studies become increasingly broad, our analytical capabilities must also expand and, perhaps more importantly, we should be able to more easily share and reproduce complex analyses."

Regardless of the reasons why individuals manage data poorly, many scientific disciplines, and economics in particular, have not been able to fully exploit modern information technology (IT). The authors find very few examples of the innovative use of modern information technology among economists, who often manage individual datasets in spreadsheets or single files. The reasons why may not be as important as the likelihood for some solutions.

This paper investigates the contributions that modern IT can make to some of these issues by introducing DevTreks, an open source, social budgeting web software program. Social budgeting attempts to combine the advantages of social networking web sites with the analytic power of microeconomics. It relies upon on-line communities to build the site's content. In other words, social budgeting reduces the cost of, and provides a framework for, collaboration. The community of agriculturalists, by pooling their currently fragmented efforts, can then "divide and conquer" the problems associated with incompatible and inconsistent data.

The paper addresses the potential to use IT, social networks, and social budgeting to: 1. improve the collection and dissemination of basic economics data, 2. automate at least some economics analyses, and 3. improve how data is shared between biophysical and economics models. A case study, using two common agricultural science data sets (farm profits and nitrogen (N) loading from tile drained agriculture), and two modern software applications (RZWQM and DevTreks) illustrate the main points. Although the case study focuses on agricultural resource management data, the IT issues are the same for other types of economics and biophysical data.

2 Modern IT - Clothing the Emperor anew

A. What is Social Budgeting

Most definitions of social budgeting focus on how governments can spend money better by including citizens in government budget preparation. UNICEF (2007) describes the process as follows: " In Kenya—where social budgeting was introduced in three districts and at the national level—the initiative seeks to strengthen (government budgeting) practices and make them more effective in delivering social development objectives. Specifically, the objective of the social budgeting exercise was to strengthen the (government budgeting process) in the five areas mentioned in the previous chapter: improving the focus on human development goals, increasing transparency and accountability, promoting wider partnerships, building capacity at the district level and improving the use of data and analysis". Shah (2007) uses a related term, 'participatory budgeting' to describe how: “(Participatory budgeting) offers citizens at large an opportunity to learn about government operations and to deliberate, debate, and influence the allocation of public resources … the enhanced transparency and accountability that participatory budgeting creates can help reduce government inefficiency and curb clientelism, patronage, and corruption”.

In contrast, from DevTreks' perspective, social budgeting can benefit anyone who makes decisions based on budgets, whether governments, individuals, firms, or researchers. DevTreks describes social budgeting on its web site as follows:

“Social budgeting is the on-line sharing of social and economic knowledge among networks and groups of people for the purposes of improving lives and livelihoods. Families finances improve when neighbors show them how to budget their money, or energy use, better. Businesses improve by more easily discovering how their performance, be it profitability or carbon use, compares to their peers. Governments spend money better when project and program managers share information about the benefits, costs, and tradeoffs involved in the delivery of public goods and services.

Social budgeting employs the economics perspective that society, as a whole, is better off when citizens, firms, and governments use resources efficiently. Social budgeting, when done properly, shows people how to conserve scarce resources such as labor, capital, goods, services, know-how, talent, and natural resource assets. “

B. How Should Social Budgeting Be Carried Out?

As with the definition of social budgeting, the nuts and bolts methods for carrying out social budgeting depend on the perspective taken. Most researchers describe the social budgeting process in terms of government accounting and budgeting. Scholz et al. (2000) identify the two major ingredients in social budgeting to be a “Social Accounting System” and a “Social Budget”. The first is a formal financial accounting system implemented using a national statistical reporting system. The second is a formal budget forecasting system that is a mathematical model of the accounting system. Both ingredients differ from traditional government accounting and budgeting systems by their focus on improving social sector spending. The authors describe this system in great detail and recommend that their book be used as a “How to build a Social Budget guidebook”.

In contrast to these financial accounting methods, DevTreks carries out social budgeting using traditional economic data such as input and output prices and operating and capital budgets, which are then aggregated and analyzed using economic analysis methods. The nuts and bolts collection and analysis of this economics data is accomplished through social networks who use online web software applications. How should these web applications deliver social budgeting services?

The authors' experience building DevTreks, and experience collecting agricultural economics production data (25+ years for the lead author), leads them to identify the following characteristics that a modern agricultural economics information system needs to carry out social budgeting:

2.1 Searchable: Data should be easy to find. A 10-10-10 fertilizer price series, an operating budget data set for Midwestern farm rotations, or a capital budget analysis of alternative dairy waste management investment returns, should be one or two clicks away. Google set the standard for easy, relevant searches. *An economics information system should quickly return relevant search results using an interface like Google.*

2.2. Web based: Users should be free to use whatever computer operating system they prefer. The Internet allows them this freedom. An economics information system should be web software that runs in all major browsers. Since cellphones now include web browsers, some economics information should be accessible using mobile web browsers. In addition, the Internet industry has established standards for making web based data accessible to machines. The standards fall under the general category of web services and include the use of Uniform Resource Identifiers (URIs), or web addresses, for every page of data. *An economics information system should be web based, supporting both desktop and mobile device browsers. It should allow users, including machines, to access data using simple Internet protocols, such as URIs.*

2.3 Social: Sites that attract the most visitors are social networks. Sites like "Facebook" or "YouTube" serve as information exchanges that improve how people with common interests can share data. Most of the content on these sites is developed by the users themselves. Niche sites, catering to professional audiences, are evolving from these general purpose social networking sites. Economics data would be easier to collect if thousands of groups of specialists, whether global warming cost estimators, water conservation investment analysts, or grocery store price watchers, joined together in a common social networking site to share their data. *An* *economics information system should be a social networking site that targets professionals who have a need to collect, disseminate, and analyze economics data.*

2.4 Open: Scientists and researchers, and many young computer users, like to tinker with computers and software. Open source software has become an increasingly popular way to allow these tinkerers to collaborate, change, and improve software. The R project is an example of successful, open source statistical analysis project (R Development Core Team, 2009). Statisticians commonly contribute new statistics modules, or "packages", that extend the capabilities of the R Project. Since economics data serves a social purpose -helping people to improve how society, not just individuals, allocate resources- it fits well with open, collaborative software. Open also implies more reproducible results, as other researchers can examine the calculations used in an analysis in detail. *An economics information system should be open-source and managed by competent open-source software managers.*

2.5 Standards based: A web based data application must meet a combination of professional, data, and information technology standards. Professional standards need to ensure that data is reliable, high quality, and follows 'best-of-class' practices, for calculations and analyses. Examples of these types of standards include the Costs and Returns Estimation Handbook (Hallam et al., 1999) and the American Society of Agricultural and Biological Engineers standards (ASABE, 2009). Information technology standards need to ensure that data can be collected, stored, and made accessible, in the best available manner to the widest possible audience. Examples of information technology standards include the W3C recommendations on XHTML (a web data format, see http://www.w3.org/MarkUp/), the Dublin Core metadata element set (data that describes data, see http://dublincore.org/), and the International Digital Publishing Forum's (IDPF) electronic publishing standards (ebooks, see www.idpf.org). Data standards need to ensure the data can be easily accessed, shared, aggregated and analyzed. Examples of data standards include the UNIFORMATII construction cost classifications (Charette and Marshall, 1999) and the Encyclopedia of Life's biological species classifications (see www.eol.org/). *An economics information system should closely follow the best available professional, data, and information technology standards. When these standards don't exist (i.e., agricultural resource management classification standards), teams should be established to develop the missing standards.*

2.6 Linked: The inventor of the world wide web, Tim Berners-Lee, recently gave a talk explaining his vision of how "linked data" (Internet data that is formally related to other Internet data) is a solution to many of the data shortfalls addressed in this paper (Berners-Lee, 2009). The need for good "linked data" is especially acute for agricultural economics data because of its dependencies upon climatic, landscape, and demographic data. This "linked data", more formally known as the semantic web, is emerging as a new set of Internet technologies for making data more useful (Feigenbaum et al., 2007). It includes schemas for relating data (RDFs, or Resource Description Frameworks), taxonomies for classifying and defining data (ontologies), and inference engines for finding new relations among different ontologies. *An economics information system should return data that has been formally linked and related using semantic web technologies. The resultant data should support new, and better, forms of automated economic analyses.*

2.7 Secure: Once a research team toils for six months building a good data set, the data should not change unless the researchers explicitly grant permission to someone to edit the data. A farmer who agrees to share crop enterprise data, but desires to stay anonymous, should be ensured of their privacy. The Internet was not designed with data security in mind. An economics information system should continually evolve better ways to secure its data. *An economics information system should keep data private if it was provided with an understanding of confidentiality.*

2.8 Long lived: Many popular web sites appear juvenile and temporarily fashionable. In contrast, economics data tends to be serious and have historical importance. An economics web site needs to err on the side of gravitas rather than fashion. More importantly, it needs to be dependable over the long term. *An economics information system should be designed to store data for decades, if not centuries.*

2.9 Easy to use: Data should be easy to enter, navigate through, save, upload, package, calculate, analyze and download. Modern web applications are specifically designed to replace desktop applications. They make full use of modern Internet technologies such as asynchronous data loading (see AJAX, http://en.wikipedia.org/wiki/AJAX) and javascript libraries (see jQuery, http://en.wikipedia.org/wiki/JQuery). These technologies help to overcome common web application annoyances like page flickering, slow data loading, getting lost while navigating, and poor user interaction. *An economics information system should use the latest web technologies to make the web applications as easy to use as many desktop software applications.*

2.10 Rich: Most web sites now commonly support the sharing of photos, videos, and other multimedia. Crop operating budgets would be easier to understand if photos and videos were included of the crop operations. Capital budgets could be assessed more meaningfully if blueprints were available for investment components. An economic analysis might be easier to understand if the author included a video explaining the how and why of the analysis. *An economics system should fully support the use of photos, videos and other multimedia to make economics data richer.*

2.11 Understandable: People with a potential interest in economic data vary from Indian farmers seeking current market price data, extension staffs wondering about the profitability of a prospective new crop introduction, microfinance groups trying to assess how well their group compares to other business organizations, and economic researchers assessing global warming abatement technologies. Practitioners and researchers throughout the world need data that they can understand. This means the data must be international and it must be explained. Data should be accompanied with stories explaining the context of the data. *An economics information system should be international in supporting multiple languages, measurement systems, and production technologies, and all data should be linked to explanatory stories.*

2.12 Extensible: Grameen Bank's App Lab (http://www.grameenfoundation.applab.org/section/index) demonstrates how a software application can be opened to software developers so that they can contribute their own software extensions. Economists have developed thousands of statistical models. Many of these models would make useful software extensions to an economics information system. To name a few, economic extension modules could analyze price indices, determine cost effectiveness, examine marginal costs, find an economic optimal solution, carry out what if scenario planning, and assess risks. *An economics information system should offer incentives to software developers to build "extensions" that extend the economics services offered by the system.*

2.13 Accessible: People with disabilities, including the blind and deaf, should be able to use and analyze the data. The standard row and column nature of price data, in particular, makes this feasible. The Internet industry has recognized this need and developed sound standards for making data accessible (see NIMAS, http://nimas.cast.org/). *An economics information system should follow widely accepted standards for making data accessible to people with disabilities.*

C. How Does DevTreks Carry Out Social Budgeting?

DevTreks delivers social budgeting services through online social networks that focus on a general theme of importance to a social network's members, such as residential construction, agricultural conservation, family budgeting, or government cost benefit analysis. These social networks, which DevTreks refers to as “network groups” are divided further into networks, clubs and members. An agricultural network group might have individual networks devoted to crop production, organic farming, dairy, cattle ranching, or innovative water conservation technologies. Clubs within a cropping network might focus on corn and soybean production in Iowa, or cotton production in Texas. Each club's members might include extension workers, pest control advisers, carbon emission researchers, farmers, or ranchers. Each club offers their content, known as services, to other clubs through service agreements.

DevTreks web software allows clubs to build hierarchical data sets (i.e. DevTreks' services) holding input and output prices, operation and component costs, and operating and capital budgets. The web software includes calculators (i.e. to allocate the cost of agricultural machinery) and analyzers (i.e. mean and standard deviation statistics). It also includes basic story-telling (i.e. case studies, dictionary entries) and multimedia handling (i.e. photo gallery) services. All applications and services have independent search engines. The DevTreks web site includes videos demonstrating how the software works.

A DevTreks club uses standard web addresses, or URIs, to find whatever data needs to be edited or analyzed. These URIs take the form: devtreks.org/agtreks/preview/crops/budget/Corn/12345. The parts of this URI can be interpreted as: host/networkgroup/webaction/network/node/commonname/id. This style of URI, and web content management, also allows machines to access data (this will be explained further in the conclusion). Members of clubs enter data using standard HTML forms which are posted to a DevTreks web site using standard HTTP methods (i.e. GET or POST). The web software uses a model-view-controller architecture to route the action needing to be taken by the server to the proper place. If the action needed is to edit data, web software will edit an XML document associated with the URI and update an enterprise database. Once the edits are completed, the edited XML document is converted to XHTML using a stylesheet and the XHTML response is sent to the web browser that originated the action.

Using the 13 desirable social budgeting traits described above, DevTreks (version beta 0.8.7, January, 2012) can identify many of its current pros and cons:

2.1 Searchable. Each web software application within DevTreks has an independent search engine. These applications include inputs, outputs, operations, components, operating budgets, capital budgets, media resources, linked views, custom documents, networks, members, clubs and services. The search engine includes several ways to filter data including keywords, categories, services and networks. Searches could be enhanced if DevTreks adopted some of the search algorithms employed by modern search engines.

2.2 Web based. All DevTreks content is delivered to all major web browsers using standard XHTML. Clients don't download any software, other than javascript libraries. DevTreks includes a sample project that demonstrates how machines can access all data found in a DevTreks database using standard web services (i.e. using REST, or Representational State Transfer, web services). A mobile phone version of DevTreks is in the planning stage.

2.3 Social. DevTreks' members organize themselves into one or more clubs which can join one or more networks. All content is owned by clubs, not members. Clubs can allow the public, or other clubs, to edit their data. DevTreks recognizes the need to adopt other types of social networking features, such as instant messaging, peer-to-peer feedback, media editing and streaming, and social networking analysis tools.

2.4 Open. DevTreks binaries and database are freely available through two open source software repositories (the DevTreks web site has links to the repositories). The source code is available to anyone planning to make a substantial contribution to the open source project. Although anyone can download and use the web software, DevTreks requests that users make voluntary donations to the project to defray the organization’s expenses. As DevTreks becomes more popular, and gets the early kinks out, it plans to adopt more traditional open-source management practices such as allowing contributors to take responsibility for specific software features and using the source code review and testing services found on open source repository sites.

2.5 Standards based. All DevTreks data is delivered to web browsers using standard XHTML or XML documents. The web software has several calculators available that implement the recommendations made by the Commodity Cost and Returns Task Force (Hallam et al, 1997), including agricultural machinery, irrigation, and net present value calculators. The database has two sample economics data sets that demonstrate the use of classification standards for classifying economics data, including the construction industry's UNIFORMATII (Charette and Marshall, 1999) and an ad-hoc agricultural resource management system that was used in the case study below (section 3.3). DevTreks plans to continue adopting, or developing, improved data and information technology standards, in areas such as schema development, data definition documentation, review of economic calculation and analyses results, application programming interface development, and software testing and documentation.

2.6 Linked. DevTreks does not currently use semantic web technologies that allow Internet data to be formally linked and related. It does, however, have a specific web software application, named "Linked Views", that allow two unrelated pieces of data, such as an agricultural machinery calculator and a farm tractor input, to be easily linked (the semantic web relation between the two pieces of data is currently missing). DevTreks' business plan, which can be downloaded with the open source software, calls for devoting 15% of its time to semantic web data development in the future.

2.7 Secure. Although DevTreks makes use of standard Internet security protocols such as authorization and authentication, its security is still in beta status. It will undergo additional testing before it is fully secure.

2.8 Long lived. DevTreks is in beta testing and can't make any claims about being long lived. The current DevTreks business model, which derives most of its revenues from surcharges associated with clubs paying subscription fees to use other clubs' apps and data is untested. DevTreks plans to continue fine-tuning its business model as successes and failures become clearer.

2.9 Easy to use. The instructional videos on the DevTreks web site can give prospective users some idea about DevTreks' current ease of use. As DevTreks matures, it will adopt standard software industry practices for getting feedback from users about the software's ease of use.

2.10 Rich and 2.11 Understandable. The web software includes one application for managing media resources, such as images and videos. These images and videos can be linked to XML documents, using the "Linked Views" web software application, to produce structured "stories", such as case studies, dictionary items, food recipes, or scientific articles. The "stories" can then be linked to all DevTreks economics content. In addition, all economics content can be localized (i.e. units, currencies, interest rates) by using DevTreks' calculators. Like any Internet company, DevTreks recognizes the need to enhance its media management capabilities by either hiring, or getting contributions from, html designers, graphic artists, video streaming experts, and other media management professionals.

2.12 Extensible. All calculators and analyzers, or "apps", found in DevTreks make use of software patterns (i.e. Microsoft's Managed Extensibility Framework) that keep them completely separated from DevTreks core web software applications. A formal "app store" application programming interface, or API, that developers can follow to add additional apps to DevTreks, is in an early stage of development. Nevertheless, the existing extensions found in DevTreks demonstrate how to add additional apps to DevTreks. DevTreks business model, where clubs subscribe to, and pay for, the apps and data they like, is the only mechanism currently planned for ensuring the quality of these apps.

2.13 Accessible. DevTreks does not currently implement standards for making web data accessible to people with disabilities. This is a recognized shortfall that DevTreks will improve in future releases.

D. Can DevTreks be Used to Deliver Scientific Economics Data?

Those researchers who define social budgeting in terms of better government budgeting, would probably prefer collecting and analyzing social budgeting data using an ideal government accounting and budgeting system (see Andrews (2007), What Would an Ideal Public Finance Management System Look Like?). In contrast, the DevTreks' perspective that anyone can benefit from better budgeting, prefers to use an "Ideal Economics Data Collection, Dissemination, and Analysis System". What would such a system look like?

Both the National Task Force on Commodity Costs and Returns Measurement Methods (Hallam et al, 1997) and Just and Pope (2001) identify the shortfalls in existing economics information systems and make recommendations about how to overcome these shortfalls. Both authors identify the need for national databases of economics production data that has been collected using uniform, consistent, and scientifically valid approaches. They differ in their approach for carrying out the work. The Task Force (page 12-15) recommends a joint government survey-university farm record keeping approach. They present a table contrasting the characteristics of existing data collection systems with the proposed system (page 12-16). Just and Pope believe that neither government surveys nor state university departments are ideal for collecting and maintaining the needed national, or international, data. They think that a non-governmental organization may be a better solution. The Task Force presents a prototype of what such a system would look like (Appendix 13B) and identifies the Internet as being the preferred platform for such a system. Table 2 uses the authors' recommended characteristics for a scientifically valid agricultural economics information system to compare DevTreks with other economics information systems.

Table 2 shows that DevTreks compares favorably with other economics information systems – it allows clubs to: collect data anyway they prefer; collect highly detailed input, output, operating budget and capital budget data; collect longitudinal data; link to structured XML data, such as surveys, natural resource attributes, or firm characteristics; and use established practices for reducing the risk of excess survey exposure. The table points out that DevTreks comes with data sets, including a national data set of crop rotations that demonstrate most of these characteristics. The table shows that DevTreks can probably evolve to meet researchers need for national and international databases of "scientifically valid" agricultural economics production data.

Ultimately, clubs using DevTreks are responsible for delivering "scientifically valid" data. DevTreks provides the web software, the tools, the web site, and some sample data sets, but doesn't deliver economics content itself. The DevTreks business model envisions a market for economics data, where clubs subscribe to, and pay for, data from other clubs. Those clubs that deliver high quality data, and high quality analysis, are expected to generate higher demand for their data, and higher revenues, than other clubs.

Case Study - DevTreks and RZWQM in Iowa corn soybean plot data analysis

Modern IT can help disentangle the complexity inherent in environmental problems in agriculture. Externalities from agricultural production are a significant problem that improved economic data could help address. Tegtmeier and Duffy (2004) estimated that for the U.S. in 2002, externalities, or the cost of agricultural production borne by someone other than producer, were between $5.7 and $16.9 billion. The externality addressed in this case study is N loading from tile-drained agriculture, a critical issue in addressing hypoxia in the Gulf of Mexico (Goolsby et al., 2000; Petrolia and Gowda, 2006). Weersink et al. point out that the textbook approach to agricultural pollution issues of setting a Pigouvian tax based on an optimal residual level may not be feasible given the high cost of quantifying the marginal benefits of reducing externalities. They propose instead that tradeoff curves "may actually be the absolute best method for the purpose of practical decision making" (2002, p. 126). Weersink et al. present a theoretical approach and do not present a case study, nor indeed, any economic or externality data.

The case study presented here illustrates how modern information technology can be used to improve economic data collection, dissemination, and analysis, culminating in the development of a tradeoff curve for N loading. Two software applications (Root Zone Water Quality Model, or RZWQM, and DevTreks) are used to study the economics of nitrate loading on Midwestern, tile-drained, corn soybean production. The characteristics of these software applications highlight common IT shortfalls and advances found in current modeling efforts and information systems development.

The data used in the case study comes from long-term plot experiments carried out by Iowa State University on the Northeast Iowa Experiment Station near Nashua, Iowa (Kanwar, 2006; http://www.ag.iastate.edu/farms/northeast.html, accessed 2/19/2010). The case study reflects a common need in studies of environmental problems caused by agriculture: many factors influence the quantity of pollutants released, such as climate, soils, slope, and crop rotation, tillage system, nutrient application methods, timing, and amount, and ancillary conservation practices. A similar list of factors affect the farmer's net returns, including the producer's knowledge, risk tolerance, equipment stocks, input and output prices, and government programs. The interactions of technical and economic factors jointly determine farm income and agricultural pollutant loading.

3.1 RZWQM - Root Zone Water Quality Model

The Root Zone Water Quality Model, RZWQM (Ahuja et al., 2000), simulates the physical, chemical, and biological processes occurring in one dimension from the bottom of the root zone to the top of the crop plant. RZWQM has been extensively tested, with results published in over 200 papers. Crop yields are model outputs; RZWQM also simulates the flow of water out of the soil into tile drains and the associated concentration of nitrogen, providing a daily estimate of N loading throughout the year.

Measured variables used in the study come from field experiments performed between 1990 and 2003, and include weather records, corn (*Zea mays* L.) and soybean (*Glycine max* (L.) Merr.) yields, tile drainage volume, and N concentration in drainage. The data were collected from 36, 0.4 ha plots, in a randomized complete block design. Due to measurement and soil issues, data from only 30 plots were used for this study. Management treatments were studied in 6 year cycles.

Extensive testing of the RZWQM model with the Nashua dataset is documented in a special issue of *Geoderma* (Vol. 140, No. 3). In particular, Bakhsh et al. (2007) documented the effect of topographic variation on yield. Ma et al. (2007b) evaluated the ability of the model to simulate soil water, the water table, tile flow, and N in tile flow, as well as N uptake and mineralization. Ma et al. (2007a) compared crop yield and N losses for 2 crop rotations and 4 tillage treatments. Malone et al. (2007a) simulated corn yield and N in tile flow from 5 fertilizer and 5 swine manure treatments, as well as long-term effects from varying N application rates, timing, and sources, and the use of winter wheat as a cover crop. Lastly, Malone et al. (2007b) provide an empirical analysis of management effects on N loading in tile flow. Figure 1 summarizes the distribution of 420 RZWQM N loading simulations for 30 plots over 14 years compared to the measured values.

3.2 RZWQM Stand Alone Analysis

The parameterization of RZWQM for the *Geoderma* special issue was the starting point for this case study. Soil hydraulic conductivities and soil water retention curves were determined using soil samples collected in 2001 from a nearby field (Ma et al., 2007b). All soils were assumed to be the same as plot 18 at Nashua.

RZWQM was used to extend the observed results from the Nashua in two ways. First, each management system was simulated in RZWQM for 30 years. With a treatment cycle of 6 years, a two-year crop rotation implies only 3 observations of yield for each crop. As can be seen in Figure 2, weather in any particular year can have a significant effect on both N loading and crop yields, which affect net returns. The very high N loadings in the initial years are probably due to N applications during the very dry 1988 and 1989 years, which led to high annual N loadings during the wet early 1990s, in combination with the use of moldboard plowing (Kanwar, 2006). Weather data (solar radiation and daily rainfall) were derived from an on-site weather station and nearby cities (Saseendran et al., 2007) and a common climate input file from 1964-2003 was used, with the first 10 years of simulation considered a warm-up period. Figure 3 compares N loading simulation results from the 1990-2003 study period at Nashua with the full 30 year period 1974-2003. Average N loading is lower for all management systems for the 30 year simulations.

The second extension of the Nashua dataset is the consideration of management systems that have not been studied at Nashua. A total of 30 management systems were designed to explore combinations of crop rotation, tillage, N application amount, type, and timing, as well as a cover crop (Table 1). Twelve of the simulated management systems (Nashua Treatments) are very similar to those studied at Nashua and in common use in northeastern Iowa. Another 12 management systems (Low Spring) were designed to reduce N loading in tile flow by lowering the N application rate, ensuring N was applied only in the spring, and considering a winter cover crop. Lastly, 6 management systems (High Fall) reflect the fact that it simply is not feasible to apply N fertilizer on all corn acres in the spring, and that fall application rates are higher in anticipation of higher N losses because of the earlier N application. In all management systems N is applied only on the corn crop.

3.3 DevTreks Data Stand Alone Analysis

An early prototype of DevTreks was used to build two economics data sets: 196 rotational crop operating budgets stored as relational data in a database, and 504 budgets stored as whole XML documents in the same database. The 504 budgets represent a complete set of observations for all of the Nashua research plots (36, 1 acre annual plots) for the period 1990 to 2003. These data sets can be found in the database that comes with the DevTreks open source software (although they have been used extensively for software testing). The budgets were built by combining National Agricultural Statistics Service (NASS) input and output prices with the experiment farm's input and output quantity data (USDA - NASS, 2003). The machinery is the experimental farm equipment actually used on the Nashua Farm. All machinery and crop budget calculations derive from guidelines recommended by Hallam et al. (1999).

DevTreks (the prototype had a different name) was deployed in an United States Department of Agriculture web farm (in 2004) and was configured using two web servers, one file system server, and one database server. Each server was deployed on a different physical computer. The prototype allowed two separate groups of researchers in separate states (i.e. Oregon and Arizona) to edit the budgets, at the same time, using two different web browsers (i.e. Internet Explorer and Firefox). Since all of the data delivered to the browsers was standard XHTML, neither group had to make any special adjustments to their browsers (other than allowing JAVASCRIPT to run). Both groups could instantly confirm that the edits were completed satisfactorily. DevTreks social networking features have evolved considerably since the prototype.

The budgets include sections for revenues, operating costs, allocated overhead costs, capital costs, and incentive-adjusted costs. Expenses were calculated for operating costs, such as materials, fuel, and repairs, as well as allocated overhead costs, like machinery capital recovery costs. Net income was derived by subtracting operating and allocated overhead costs from total revenues (crop yield * crop price). These budgets are similar to published budgets for Iowa (Iowa State University, 2006), except the prices are for different time periods, there are some differences in the technologies represented, and there was no charge for land at Nashua.

Statistical comparisons, such as mean and standard deviation operating costs, were made of the crop treatments found in the 504 budget set using a DevTreks statistical analyzer. These analyzers have evolved considerably since the prototype. Current analyzers are built using self-contained software modules, or "extensions", that can be developed independently, added to DevTreks, and then run without having to make any changes to the DevTreks software. These "extensions" are linked to specific content within DevTreks (i.e. an agricultural machinery calculator to a tractor input, a statistical analyzer to a group of crop budgets) using another software application within DevTreks.

The comparisons were possible because the XML nodes of each budget document were tagged using a common tagging system (i.e. LSNT for Late Spring Nitrate Test). This tagging, or data classification, system is completely *ad hoc*: no standard agricultural classification system could be found. This lack of standards to describe economic data remains a fundamental stumbling block to good data management and highlights a significant, current, information technology shortfall. Data can't be aggregated, compared, and fully used for modern computing unless an agricultural classification system exists.

3.4 RZWQM and DevTreks - An integrated analysis

Perhaps the simplest way to address the hypoxia issue would be for farmers to apply less nitrogen, and a typical production economic analysis would hold everything else constant while varying the amount of N applied. An agronomic analysis on the other hand, would emphasize other management variables that can affect the amount of N that enters the tile lines, such as crop rotations, tillage systems, and N application methods and timing. The management systems studied here were designed using both perspectives. Net returns were estimated for each long-term simulation by creating individual budgets for one year's corn and soybean crop under each management system. Long-term returns were calculated on the assumption that the same series of operations were repeated within each crop rotation for 30 years. For simplicity, revenues were calculated as the crop yield for each year times the 2003 price for corn and soybeans.

Figure 6 shows histograms comparing the net returns calculated by DevTreks using the measured crop yields with those calculated using RZWQM simulated crop yields. A total of 173 plot years were considered for the 10 systems that were both observed at Nashua and on the list of 30 management systems for the case study (Nashua Treatment). There is an overestimation of crop yields based on the RZWQM simulations, especially for corn, as some of the factors reducing yields, such as hail and insect damage, are not simulated in RZWQM. As a consequence, mean simulated net returns of $277/ha exceeds the mean of measured returns by $37/ha or 15% of the observed net return. Figure 7 provides another view, showing higher net returns for the simulated than the observed for both individual years and averaged returns. For annual values, the r^2 value is 0.56 and the RMSE is $145, while for the results of averages of individual management systems, the r^2 is 0.71 and the RMSE is $64.

If simulated crop yields are too high, one option is to adjust yields by the difference between observed and simulated yields. For simplicity, rather than adjusting net returns downward, for long-term planning, the over-estimation of crop yield was assumed to be offset by the long-term increasing trend in crop yields, which was also ignored in the simulation. Figure 8 shows the long-term corn and soybean yield trend for Chickasaw

County, where Nashua is located. Corn yields are increasing at roughly 2% annually, and soybean yields at almost 1%.

Simulated net returns estimated for the 30 year period are consistently lower than for the 14 year period of observations at Nashua, as was the case for N loading. Interestingly, planting corn in even years as part of a corn soybean rotation seems to better than planting corn in odd years (Fig. 9). As would be expected, when comparing the net return and N loading averages for the 30 year simulations, the Low Spring systems tend to have the lowest N loading values, and the High Fall systems the highest (Fig. 10). All 3 groups of management systems exhibit a broad range of long term average net returns. There is an obvious upper bound on net returns at around $370/ha, with a range of almost 10 to 30 kg/ha of annual average N loading, which contains systems from all 3 groups.

Lastly, Figure 11 shows the tradeoff of N delivered to the edge of the tile system and net returns for the simple case of varying N input amounts in the spirit of Figure 2 from Weersink et al. (2002). To emphasize the value of RZWQM in providing an agronomic foundation for the analysis, the variability in annual results from 5 levels of N input across all 30 management systems is plotted, in addition to lines that show the relationships between N input and N loading, N input and returns, and N Loading and returns. The lines summarizing the tradeoffs were created using lowess relationships that ignored continuous corn systems (shown in red), as those returns were very low compared to rotations of corn and soybeans, at least based on 2003 prices. The same upper limit in net returns around $370/ha as in Figure 10 is visible in the Tradeoff Curve panel, with corresponding annual N loading values in the range of 15 to 30 kg/ha.

Conservationists would want to encourage adoption of the systems at the left of the plateau, just under 15 kg/ha of annual N loading, which are the Low Spring corn soybean rotations with 110 kg/ha of N applied on corn crops. Subsidies would be required for voluntary adoption of the Low Spring cover crop systems that could reduce long term simulated N loading further, to below 10 kg/ha. Although time constraints will require farmers to fall apply some N on corn crops, particularly for producers dedicated to swine production who need to dispose of swine manure, the improved returns to corn soybean rotations should be used to discourage the High Fall corn corn rotations which result in N loadings over 40 kg/ha.

The analysis presented in this case study could be improved in many respects. In fact, every associated profession could suggest useful improvements. A climate scientist might argue that “stationarity is dead” (Milly et al., 2008), so that climate scenarios other than the historical record should be considered. An agricultural engineer might argue for larger machinery or consideration of controlled drainage, at least on flatter areas. A soil conservationist might argue that a forward looking design would consider rotations that include crops for biomass to support cellulosic ethanol production. An economist would want to explore additional input and output price scenarios. Representatives of the Iowa Department of Natural Resources or the Environmental Protection Agency might want to focus on extreme events rather than average annual values, as well as expanding the study to consider a broader range of soils across the state. Farmers would want to know the costs on their particular farms, and they would appreciate a more dynamic decision support framework than the static plot shown in Figure 11. All of these potential improvements underscore the main point of this paper: to address agricultural economic problems systematically and flexibly, modern IT systems are needed to create economic datasets and easily link economic data to agronomic data.

Conclusion and Recommendation

In Hans Christian Andersen's story about the Emperor with no clothes, only a child dares to state the obvious: the Emperor is naked. Economists readily admit that more and better data are needed to further their research. But not as many reach the conclusion drawn by Just and Pope (2001): the profession will produce a meager harvest if it continues to cultivate a data-limited field. Or, in Just and Pope's words, until economic datasets improve, empirical knowledge of agricultural production will remain “little more than an empty box”.

The need for progress in understanding the economics of natural resource issues has never been greater. As the world’s population and income rise, it will be increasingly difficult to provide food at low prices while simultaneously reducing the external costs of agriculture, reducing water use, increasing production of biofuels, conserving natural ecosystems on land and at sea, managing carbon, and adapting to climate change. The potential payoff justifies a substantial effort at improving the datasets available to the profession.

The principle information technology advance identified in this paper is the opportunity that social budgeting offers to collect, disseminate, and analyze, basic economics data. The case study demonstrated that the prototype DevTreks web software was capable of accomplishing this goal. The current software, when mature, will support the automated, online, sharing of economics data that will be so crucial in future information systems that are designed to improve agriculture.

The most serious information technology shortfall identified in this paper is that data, such as agricultural production data, is either not available, or cannot be easily accessed and shared. The case study highlighted this problem. RZWQM can't automatically and easily retrieve the profit data it needs from some other economics information system, such as DevTreks. DevTreks can't get the natural resources data it needs from a natural resources model, such as RZWQM. Can modern information technology alleviate this common problem?

At least three approaches stand out for integrating and automating natural resource and economics data sets. The first is to build either model as an independent "extension", or pluggable software module, to the other (see the contributed packages in the R-Project, explore Grameen Bank's App Lab, or examine the DevTreks.Extensions projects in the open source software). This requires reasonably close collaboration among the model builders on computer platform choices, data models, and data integration interfaces. When both models belong to the same research lab, or Internet company, this is probably the best choice. Closely integrated models can do a very good job of automating data collection, dissemination and analysis and avoid problems associated with data repetition and incompatibility.

A second approach relies on standard Internet protocols, such as the HyperText Transfer Protocol (HTTP) and the Universal Resource Identifier (URI), to improve data sharing. The HTTP protocol defines a way to identify unique resources (URIs), such as images and documents, on the Internet and transfer them to web browsers. The term Representational State Transfer, or REST, web services refer to machines using these protocols to directly share data. The data sent from a URI to a web browser, or machine, "represents" the actual data being stored on a web server. HTTP messages, containing commands such as GET or POST, describe what to do with these representations.

To make this method practical, additional search parameters, sent as HTTP query or request parameters (i.e. "http://mysearchengine.org/search?keyword=corn"), are often needed. The resultant web service "APIs"

(Application Programming Interfaces - instructions describing how data can be mechanically accessed and used) are usually published on the web sites. An example of this type of web service is Amazon's data storage service at http://aws.amazon.com/s3/. Natural resource and economic models could access remote data by writing software instructions that use the web service API. DevTreks includes a sample project that uses Microsoft's Windows Communication Framework and OData data access technology to deliver REST web services.

The third method requires the most time and effort but is probably the best long term solution. Semantic web technologies are used to relate, classify, and analyze the economics and biophysical data. Berners-Lee (2009) suggests three basic rules to follow when first using this method:

1. Use HTTP names to make data available. This is the same as the RESTful web services mentioned in method two.

2. Define content: Formally define the content of the data that is returned from a URI. This can be done using XML schemas or some other data definition language. A particularly useful contribution along these lines would be a way to describe management operations that could be used to simultaneously parametrize a biophysical model and an economic budgeting tool.

3. Return linked data: Data that is returned from a URI should be linked data, not stand alone data. Using RZWQM and DevTreks as an example, links would be made between the farm budgets and the nitrate loading projections. When someone requests data at a standard DevTreks URI they would receive both the budgets and the links to the biophysical data. The W3C, an Internet standards setting organization, has a lot of semantic web technologies available for building these types of links.

The principle advantage to this last approach is that it forces the application developers to focus on the real data problems - data integration is not unique to their models - encompassing the whole fields of agricultural resource management and water quality science. The principle disadvantage is that collaboration may be needed across entire professions for building full blown semantic web features such as data classification and relationship systems. This can be time consuming and expensive, but one can imagine the benefits of being able to pass the same description of management operations to both a biophysical model and an economic model.

An example of this approach in the field of agriculture is The Agricultural Information Management Standards (AIMS) website (FAO, 2010). The stated goal of this web site to “improve coherence among agricultural information systems that will make such systems interoperable. The objectives of AIMS are to create a clearing house for information management standards, and to share and promote the use of common methodologies and tools”. Professional examples from other fields can be found among biologists (Encyclopedia of Life, www.eol.org) and hydrologists (Consortium of Universities for the Advancement of Hydrologic Science - Hydrologic Information System, http://his.cuahsi.org/).

This paper makes three contributions. First, we propose that social budgeting is a feasible approach to building the economic datasets needed to address key agricultural issues and we describe the characteristics needed for social budgeting to be successful. Second, we demonstrate that an example social budgeting web software program, DevTreks, can be used, successfully, to build agricultural economics production datasets. Third, we present a case study linking biophysical and economic data to illustrate current information technology shortfalls (i.e. data couldn't be easily shared), advances (i.e. building online economics data sets), and planned improvements (i.e. using both RESTful web services and semantic web methods to link and share data in the future).

Open source, on-line data collection, dissemination, classification, and analysis efforts have been identified as important contributors to modern information systems development. Contribute to these projects by entering data sets, creating "extensions", and helping to build semantic and ontological systems. Open source efforts may hold the greatest promise for delivering the technical and economic information needed to improve agriculture.

Acknowledgments

We would like to thank Carl Pederson and Ken Pecinovsky of Iowa State University, Kathy Woodard and Caitlin Hall of the University of Arizona, Jim Ayen, Alan Lauver, Barbara Stewart, Steve Brinkman, and Hal Cosby of the NRCS in Iowa, Terry Meade, Bob Jacquis, and Gerardo Armendariz of the Agricultural Research Service.

References

Ahuja, L. R., Rojas, K. W., Hanson, J. D., Shaffer, M. J., Ma, L., (eds)., *The Root Zone Water Quality Model*. Water Resources Publications LLC. (2000), Highlands Ranch, CO.

American Society of Agricultural and Biological Engineers (ASABE), 2009. Standards and Technical Activities. http://www.asabe.org/standards/index.html, accessed 2/19/2010.

Andrews, Matthew. What Would an Ideal Public Finance Management System Look Like? *Budgeting and Budgetary Institutions*. Anwar Shah, editor, (2007), pp.359-383. Washington D.C., U.S.A. : The World Bank

Bakhsh, A., Kanwar, R.S., and Malone, R.W., Role of landscape and hydrologic attributes in developing and interpreting yield clusters. *Geoderma*. **140**(3) 2007), pp. 235-246.

Bergstrom, J. C., Ready, R.C., What have we learned from over 20 years of farmland amenity valuation research in North America? *Rev. Agri. Econ*. **31**(1) (2009), pp. 21-49.

Berners-Lee, T. The next Web of open, linked data. Presentation made to TED2009 Conference. February, 2009. http://www.ted.com/index.php/talks/tim_berners_lee_on_the_next_web.html, accessed 2/19/2010.

Charette, R. P., Marshall, H.E., *UNIFORMAT II Elemental Classification for Building Specifications, Cost Estimating and Cost Analysis*, Department of Commerce, National Institute for Standards and Technology, (1999), NISTIR 6389.

Coleman, A., A model of spatial arbitrage with transport capacity constraints and endogenous transport prices. *Am. J. Agric. Econ*. **91**(1) (2009), pp. 42-56.

DevTreks. *Social budgeting that improves lives and livelihoods.* http://sourceforge.net/projects/devtreks or http://devtreks.org/, accessed 01/09/2012.

Food and Agriculture Organization of the United Nations. Agricultural Management Information Standards (AIMS) http://aims.fao.org/, accessed 11/29/2010.

Feigenbaum L., Herman, I., Hongsermeirer, T., Neumann, E., Stephens, S., The semantic web in action. *Scientific American*. **297**(6)(2007), pp. 90-97.

Goodwin, Barry K., Payment limitations and acreage decisions under risk aversion: A simulation approach. *Am. J. Agric. Econ*. **91**(1) (2009), pp.19-41.

Goolsby, D.A., Battaglin, W.A., Aulenbach. B.T., Hooper, R.P., Nitrogen flux and sources in the Mississippi River Basin. *The Science of the Total Environment.* **248 (**2000), pp.75-86.

Government of Kenya/UNICEF Programme of Cooperation. Social Budgeting, Investments in Kenya's Future, Working Paper No 2, 2007 http://www.unicef.org/socialpolicy/files/Social_budgeting_investing_in_Kenyas_future.pdf last accessed 11/23/2010

Hallam, A., Eidman, V. E., Morehart, M., Klonsky, K., editors, 1999. Commodity Costs and Returns Estimation Handbook: A Report of the AAEA Task Force on Commodity Costs and Returns, Staff General Research Papers 1315, Iowa State University, Department of Economics.

Hollister, J.W., 2007. Moving forward with ecological informatics and reproducibility. EcoTone, the Ecological Society of America Blog, http://www.esa.org/esablog/?p=62, accessed 2/19/2010.

Intergovernmental Panel on Climate Change (IPCC) 2007, Climate Change 2007: Impacts, Adaptation and Vulnerability. Contributions of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

Just, R. E., Pope, R.D., The agricultural producer: Theory and statistical measurement. *Handbook of Agricultural Economics, Volume I*. Bruce Gardner and Gordon Rausser, eds., (2001), pp.629-741. Amsterdam, The Netherlands: Elsevier Science B.V.

Kaspar, T.C., Kladivko, E.J., Singer, J.W. 2005. Chapter 10 "Potential and Limitations of Cover Crops, Living Mulches, and Perennials to Reduce Nutrient Losses to Water Sources from Agricultural Fields" in Proceedings of the Gulf Hypoxia and Local Water Quality Concerns held September 26-28, 2005 in Ames, Iowa, p. 132.

Iowa State University, 2006. Estimated Costs of Crop Production in Iowa. http://www.econ.iastate.edu/faculty/duffy/pages/fm1712_2006.pdf, accessed 10/1/2009.

Kanwar, R.S., Effects of cropping systems on NO_3N losses to tile drain. *J. Amer. Water Res. Assoc.* **42**(6) (2006), pp. 1493-1501.

Key, N., Roberts, M.J., Nonpecuniary benefits to farming: Implications for supply response to decoupled payments. *Am. J. Agric. Econ*. **91**(1) (2009), pp.1-18.

Ma, L., Malone, R.W., Heilman, P., Jaynes, D.B., Ahuja, L., Saseendran, S.A., Kanwar, R.S., Ascough, J.C., RZWQM simulated effects of crop rotation, tillage, and controlled drainage on crop yield and nitrate-N loss in drain flow. *Geoderma*. **140**(3) (2007a), pp. 260-271.

Ma, L., Malone, R.W., Heilman,P., Karlen, D.L., Kanwar, R.S., Cambardella, C.A., Saseendran, S.A., Ahuja, L., RZWQM simulation of long-term crop production, water and nitrogen balances in Northeast Iowa. *Geoderma*. **140**(3) (2007b.), pp. 247-259.

Malone, R.W., Ma, L., Heilman, P., Karlen, D.L., Kanwar, R.S., Hatfield, J., Simulated N management effects on corn yield and tile-drainage nitrate loss. *Geoderma*. **140**(3) (2007a), pp. 272-283.

Malone, R.W., Ma, L., Karlen, D.L., Meade, T., Meek, D., Heilman, P., Kanwar, R.S., Hatfield, J.L., Empirical analysis and prediction of nitrate loading and crop yield for corn-soybean rotations. *Geoderma*. **140**(3) (2007b), pp. 223-234.

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J**.**, Stationarity is dead: Whither water management? *Science*, **319**(5863) (2008), pp. 573-574.

Petrolia D.R., Gowda, P.H., Missing the boat: Midwest farm drainage and Gulf of Mexico hypoxia, *Review of Agricultural Economics*, **28**(2) (2006), pp. 240-253.

R Development Core Team, 2006. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN:3-900051-07-0.

Saseendran, S.A., Ma, L., Malone, R., Heilman, P., Ahuja, L., Kanwar, R.S., Karlen, D.L., Hoogenboom, G., Simulating management effects on crop production, tile drainage, and water quality using RZWQM-DSSAT. Geoderma. **140**(3) (2007), pp. 297-309.

Scholz, W.; Chicon, M.; Hagemejer, K. *Social Budgeting. Quantitative Methods in Social Protection Series*. Geneva, International Labor Office/ International Social Security Association, 2000

Shah, Anwar. Overview. *Participatory Budgeting*. Anwar Shah, editor, (2007), pp.1-18. Washington D.C., U.S.A. : The World Bank

Tegtmeier, E.M., Duffy, M.D., External costs of agricultural production in the United States *International Journal of Agricultural Sustainability*. **2**(1) (2004), pp. 1-20.

United States Department of Agriculture, Economics Research Service and National Agricultural Statistics Service. Agricultural Resources Management Survey. Washington D.C. U.S.A. http://www.ers.usda.gov/Briefing/ARMS/, accessed 2/19/2010.

United States Department of Agriculture, National Agricultural Statistics Service. *Agricultural Statistics 2003*. United States Government Printing Office, Washington D.C.

Weyant, J.P., A critique of the Stern Review's Mitigation Cost Analyses and Integrated Assessment. *Review of Environmental Economics and Policy*. **2**(1) (2008), pp. 77–93.

Weersink, A., Jeffrey, S., Pannell, D., Farm-level modeling for bigger issues. *Rev. Agri. Econ.* **24**(1) (2002), pp. 123-140.

Figure Captions

Figure 1. Histograms of RZWQM simulated and measured annual N loadings into tileflow for 30 plots over 14 years.

Figure 2. Box plots of measured N loading and net returns based on measured yields at Nashua for the 1990-2003 period.

Figure 3. Comparison of 14 year average simulation results by management system compared to 30 year results for N loading. CC is continuous corn; CS is corn in even years; SC is soybeans in even years.

Figure 4. Histograms of RZWQM simulated and measured annual net returns for 30 plots over 14 years.

Figure 5. Scatterplot of net returns based on simulated and observed crop yields showing both annual values and averaged net returns for 10 management systems. See Table 1 for management system descriptions and number of years compared for each system.

Figure 6. Time series plot showing increasing crop yield trends for Chickasaw County, IA.

Figure 7. Comparison of 14 year average simulation results by management system compared to 30 year results for net returns. CC is continuous corn; CS is corn in even years; SC is soybeans in even years.

Figure 8. A scatterplot comparing the net returns and N loading from 30 year simulation results for 3 groups of management systems.

Figure 9. Plot summarizing the tradeoff approach proposed by Weersink et al. (2002) for 30 year simulation results of 30 management options on N loading and net returns. CC is continuous corn; CS is corn in even years; SC is soybeans in even years. Locally smoothed lines ignore CC values.

Table 1. Management systems used in long-term simulations. CP=”Chisel Plow”, NT=”No Till”; CC=”Continuous Corn”, CS=”Corn-Soybean”, SC=”Soybean-Corn”; SM=”Swine Manure”, UAN=” Urea Ammonium Nitrate”.

	Tillage	**Rotation**	**N Amount kg N/ha**	**N Type**	**Season**	**Cover Crop**	**Plot Years**
Nashua Treatments	CP	CC	150	SM	Fall	No	18
	CP	CS	150	SM	Fall	No	20
	CP	SC	150	SM	Fall	No	22
	NT	CC	150	SM	Spring	No	
	NT	CS	150	SM	Spring	No	8
	NT	SC	150	SM	Spring	No	6
	CP	CC	150	UAN	Spring	No	18
	CP	CS	150	UAN	Spring	No	30
	CP	SC	150	UAN	Spring	No	27
	NT	CC	150	UAN	Spring	No	
	NT	CS	150	UAN	Spring	No	12
	NT	SC	150	UAN	Spring	No	12
Low Spring	CP	CC	135	UAN	Spring	No	
	CP	CC	135	UAN	Spring	Yes	
	NT	CC	135	UAN	Spring	No	
	NT	CC	135	UAN	Spring	Yes	
	CP	CS	110	UAN	Spring	No	
	CP	CS	110	UAN	Spring	Yes	
	NT	CS	110	UAN	Spring	No	
	NT	CS	110	UAN	Spring	Yes	
	CP	SC	110	UAN	Spring	No	
	CP	SC	110	UAN	Spring	Yes	
	NT	SC	110	UAN	Spring	No	
	NT	SC	110	UAN	Spring	Yes	
High Fall	CP	CC	200	Anhydrous	Fall	No	
	NT	CC	200	Anhydrous	Fall	No	
	CP	CS	168	Anhydrous	Fall	No	
	NT	CS	168	Anhydrous	Fall	No	
	CP	SC	168	Anhydrous	Fall	No	
	NT	SC	168	Anhydrous	Fall	No	

Table 2. Characteristics of Alternative Data Collection Systems

Characteristics of Data	Current USDA Survey	Current University Farm Record System	Integrated USDA and University System	DevTreks NGO System
1. Based on probability sampling (1)	YES	NO	YES	YES. Clubs can collect data anyway they prefer, including from probability surveys.
2. Consistent procedures used across states (1)	YES	NO	YES	YES (see #3)
3. Data accuracy (1) Accurate reporting Detailed information Close local scrutiny	MODERATE	MODERATE	MODERATE	YES. DevTreks includes a national data set of crop rotations for the USA, demonstrating how consistent, accurate, and detailed data can be collected using DevTreks. The data was collected from cooperative extension offices throughout the USA, who collected the base data using a great deal of local scrutiny.
4. Longitudinal data(1)	NO	YES	YES	YES. DevTreks includes a sample data set containing 14 years of experimental plot data for corn soybean rotations demonstrating how longitudinal data can be collected.
5. Cost (1)	HIGH	HIGH	HIGH	LESS HIGH. The social networking features of DevTreks may be able to decrease this cost substantially
6. Data availability (1)	Limited	Varies by State	YES	Fee-based or Free. DevTreks business model allows clubs to charge subscription fees for their data. Data will be available to wider audiences if clubs charge either low, or no, fee.
7. Highly detailed microeconomics data (2)	NO (input use by crop and application rate missing)	YES	MODERATE	YES (see #3 and #4)
8. Highly detailed investment data (2)	NO	Varies by State	MODERATE	YES. DevTreks includes a capital budgeting application with data that is as detailed as the operating

				budget data (i.e. inputs and outputs).
9. Other data (i.e. firm characteristics, soil quality) (2)	YES	YES	YES	NOT DIRECTLY. DevTreks allows economics data to be linked to structured XML data, such as surveys, natural resource attributes, or firm characteristics.
10. Excessive survey exposure and right to privacy exposure (2)	YES	NO	POSSIBLY	POSSIBLY. Clubs who collect their data using surveys risk the problem of excess survey exposure. All clubs collecting firm level data face right-to-privacy issues. These risks and issues have been mitigated by techniques developed by cooperative extension economists, such as collecting data from groups of farmers rather than individual farmers.

(1) Taken from Table 12.1, page 12-16, in Hallam et al, 1997

(2) Derived from Just and Pope (2001)

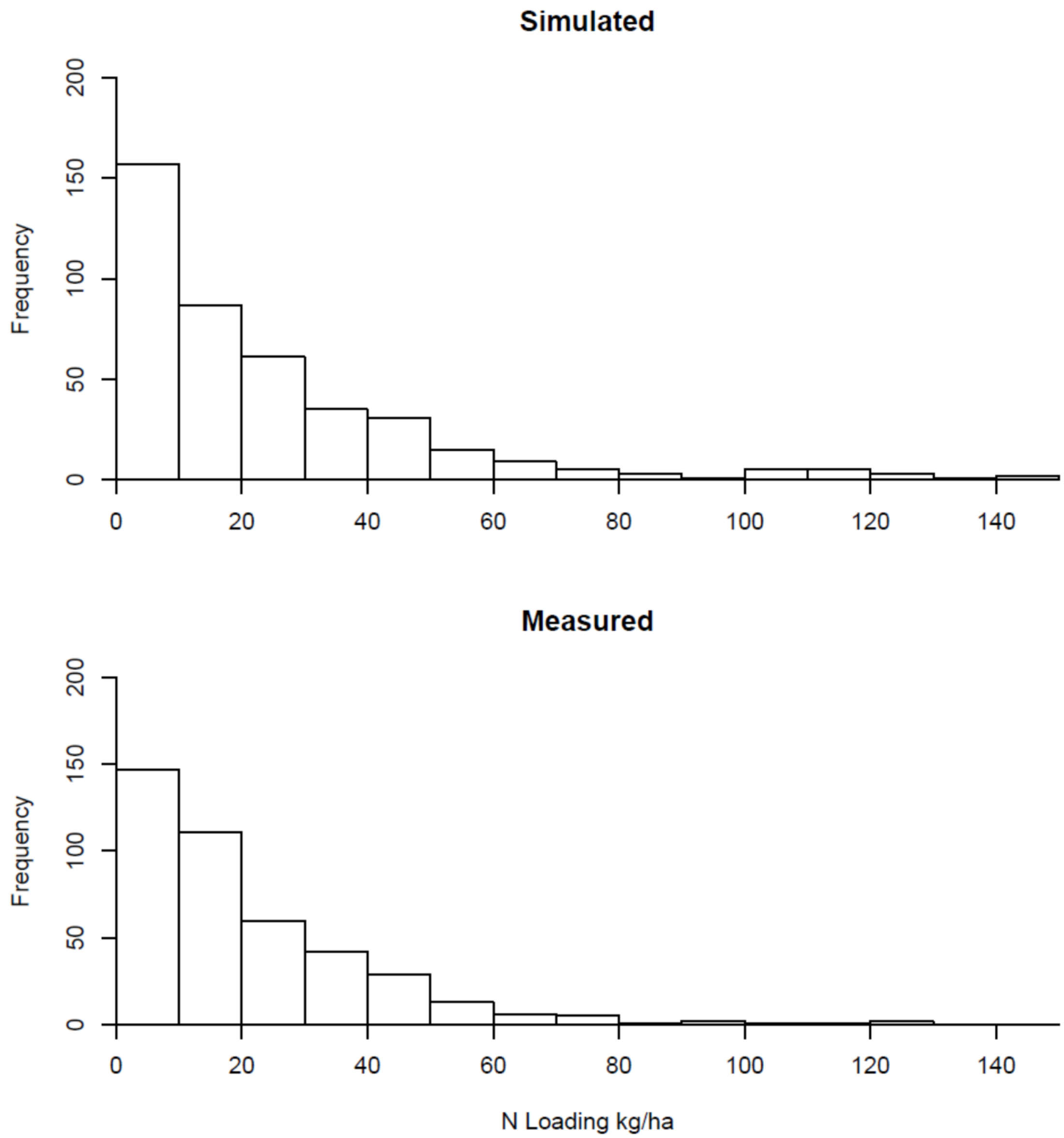

Simulated
Frequency
0
50
100
150
200
0
20
40
60
80
100
120
140
Measured
Frequency
0
50
100
150
200
0
20
40
60
80
100
120
140
N Loading kg/ha

Figure 1

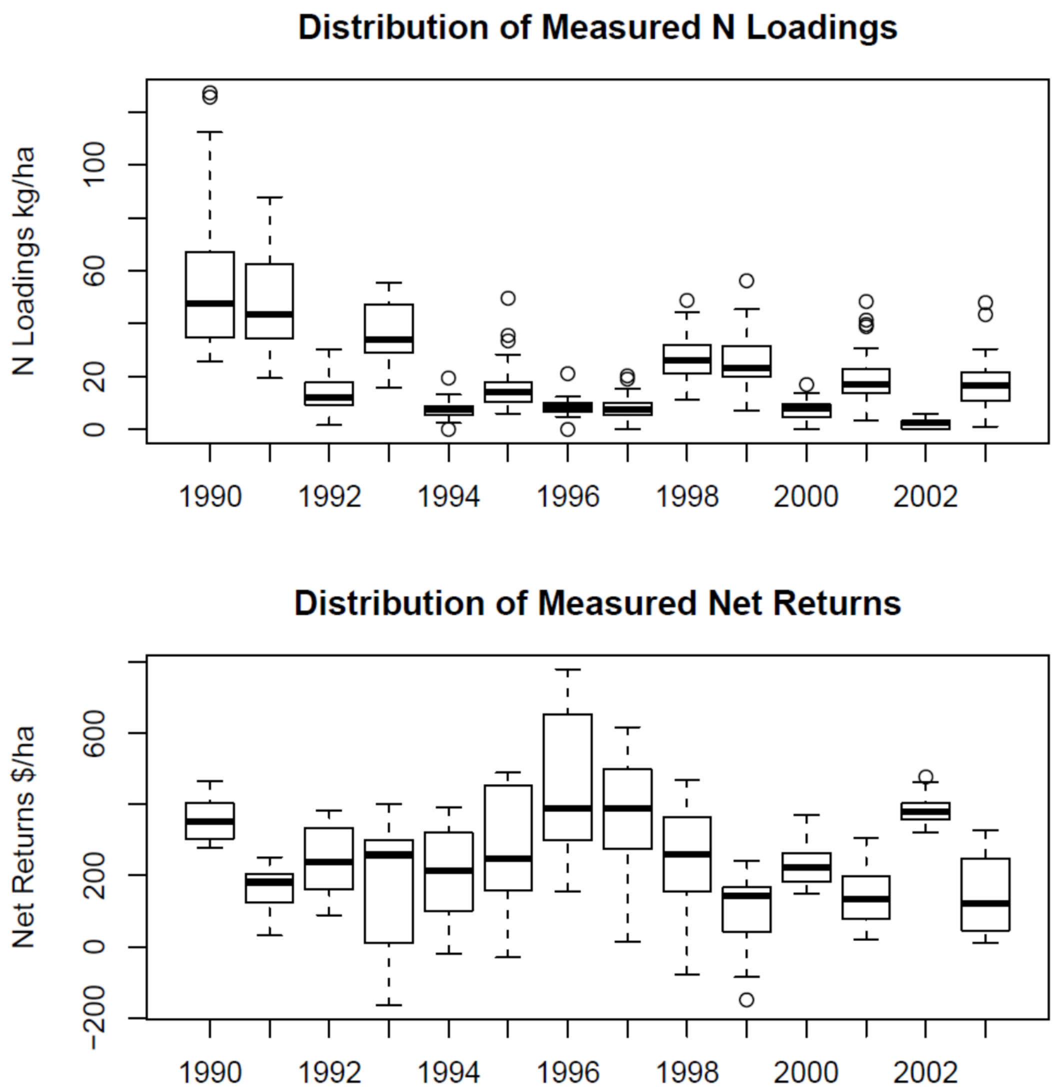

Distribution of Measured N Loadings
N Loadings kg/ha
0
20
60
100
1990
1992
1994
1996
1998
2000
2002
Distribution of Measured Net Returns
Net Returns $/ha
−200
0
200
600
1990
1992
1994
1996
1998
2000
2002

Figure 2

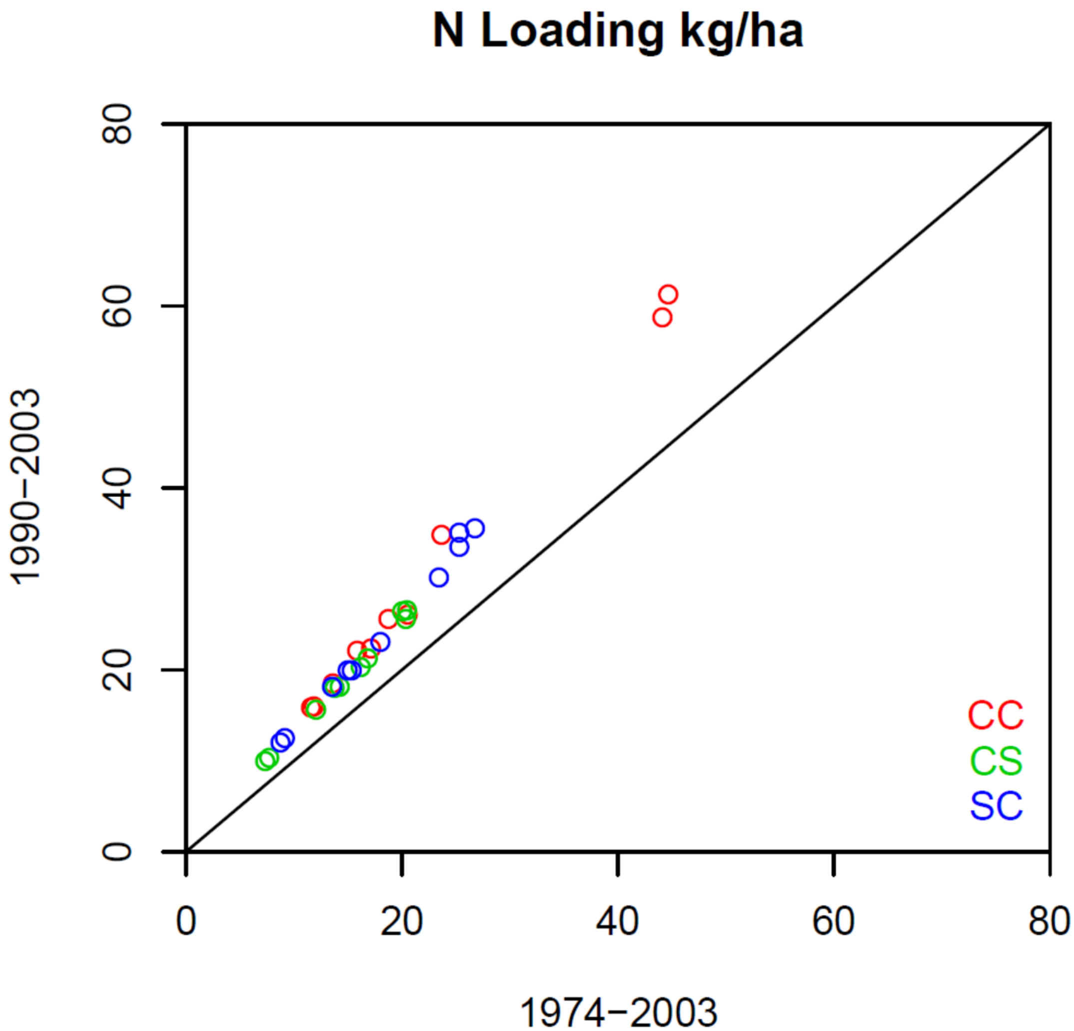


Figure 3

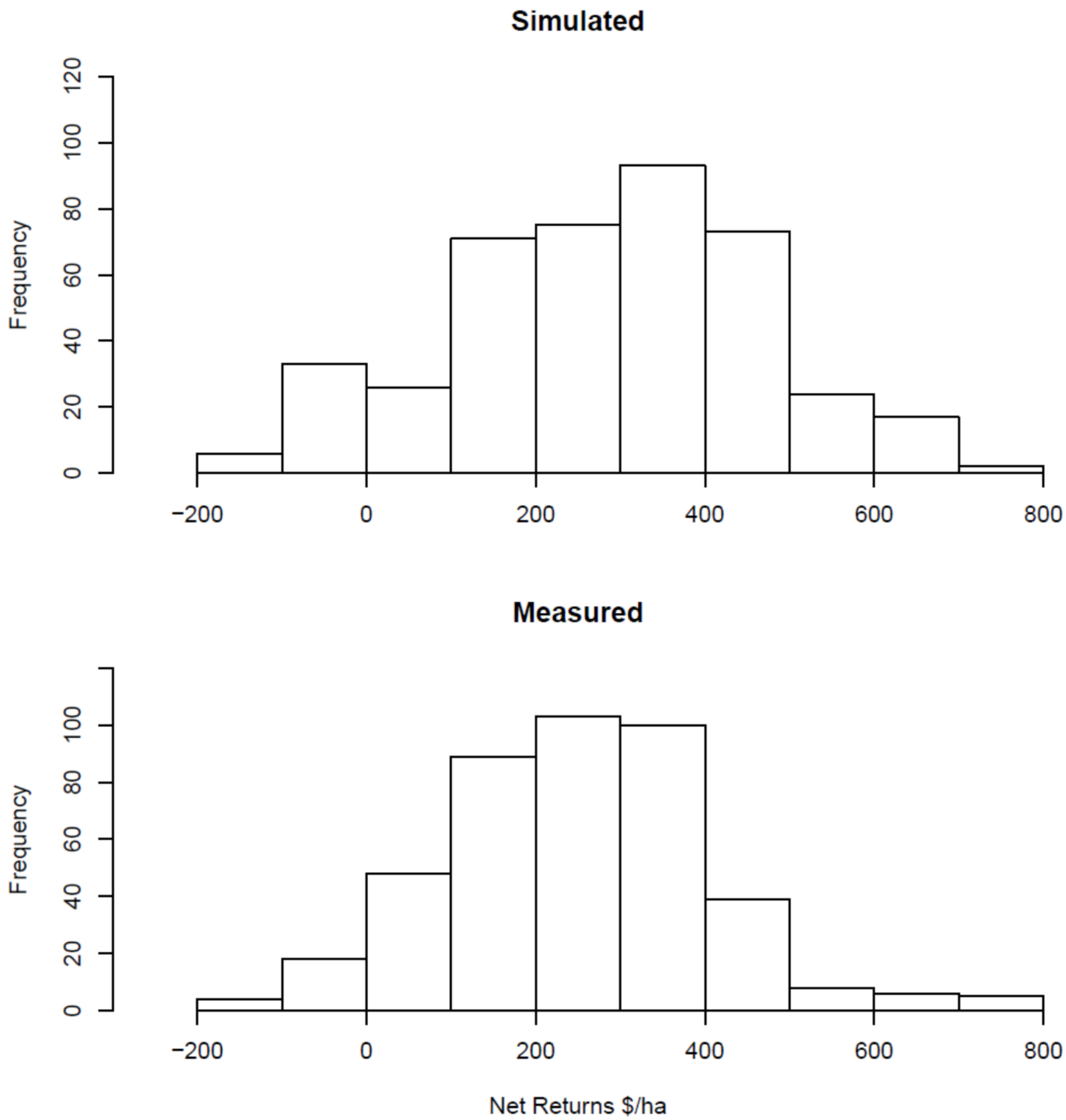
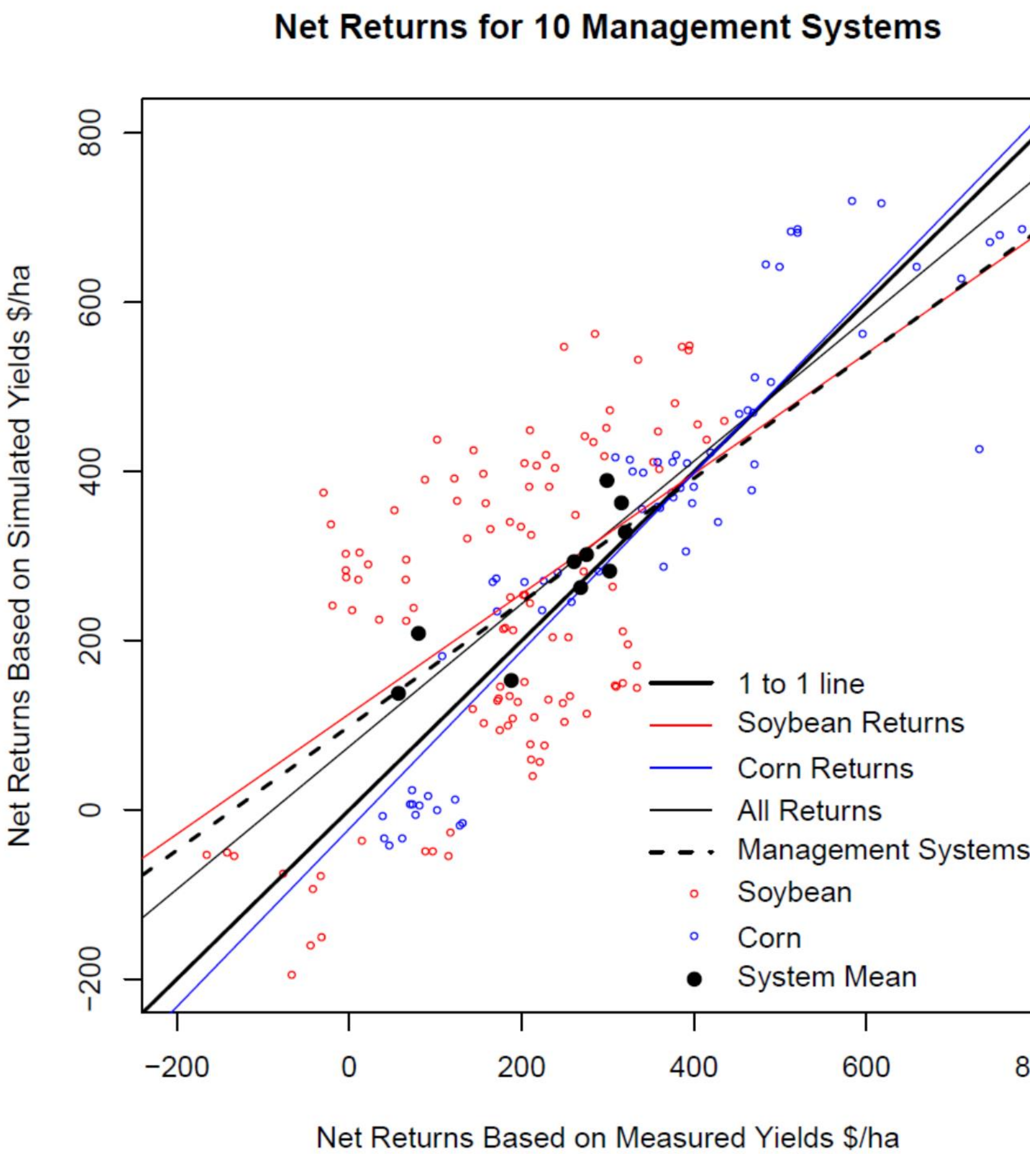

Simulated
Frequency
0
20
40
60
80
100
120
−200
0
200
400
600
800
Measured
Frequency
0
20
40
60
80
100
−200
0
200
400
600
800
Net Returns $/ha

Figure 4



805
806 Figure 5
807

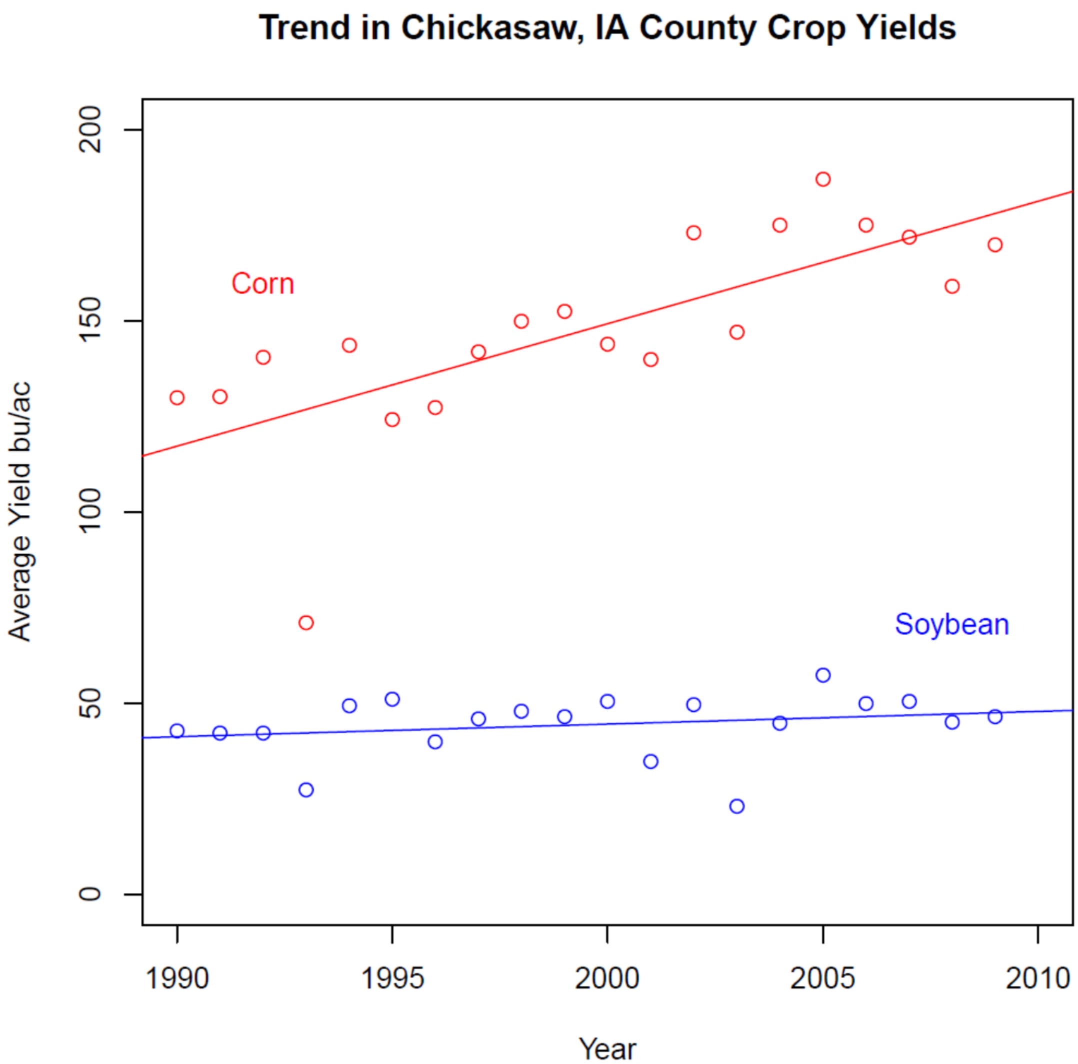

Trend in Chickasaw, IA County Crop Yields
Corn
Soybean
Average Yield bu/ac
0
50
100
150
200
1990
1995
2000
2005
2010
Year

Figure 6

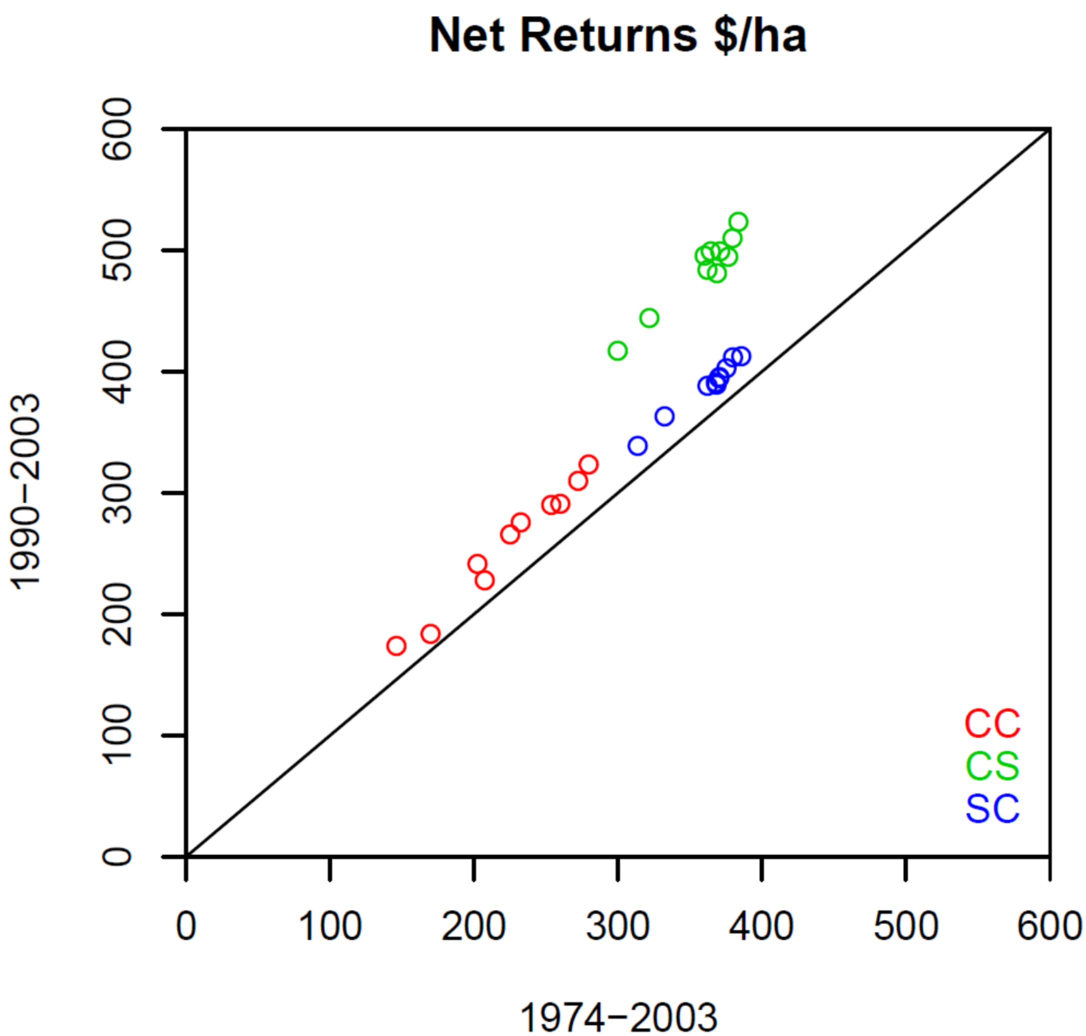
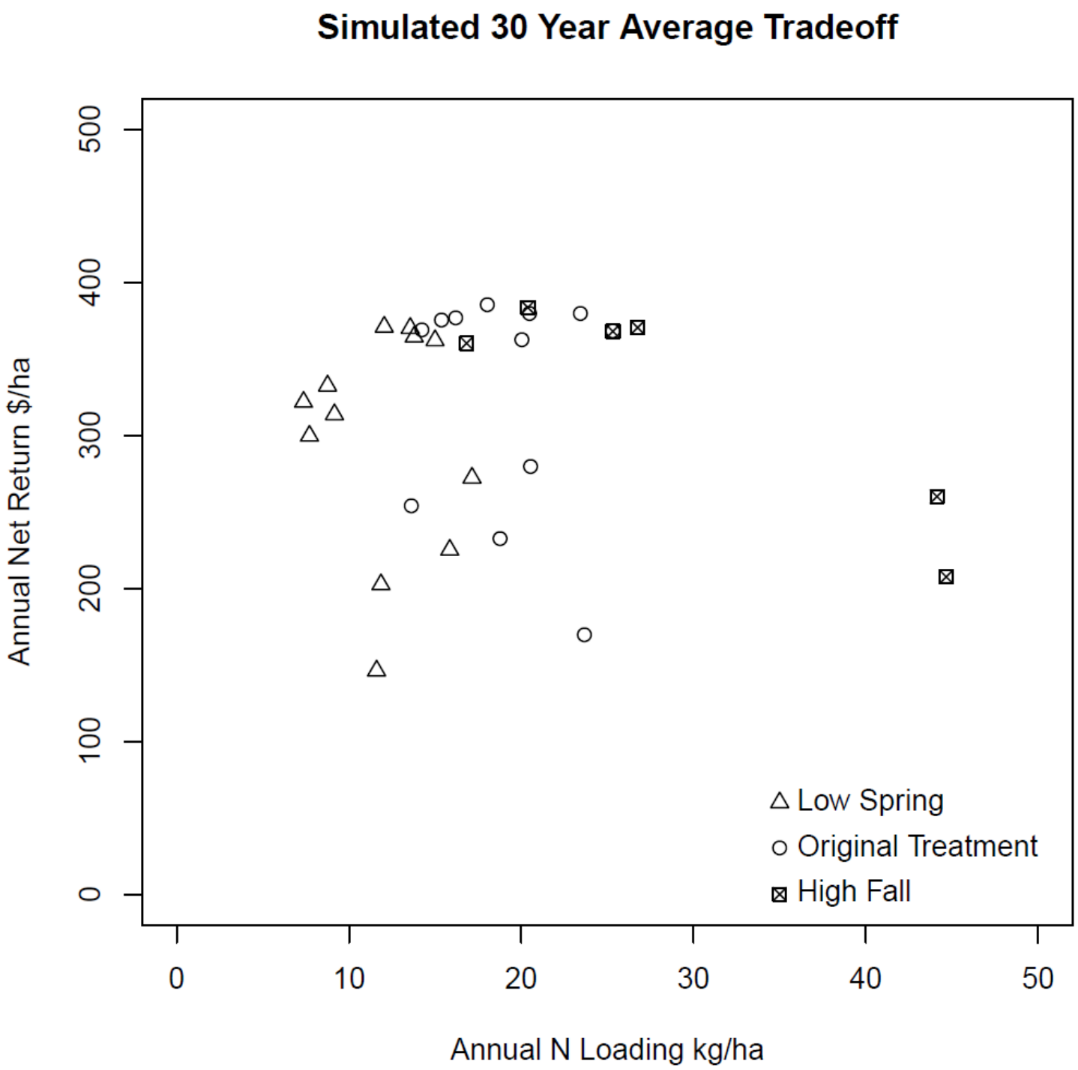
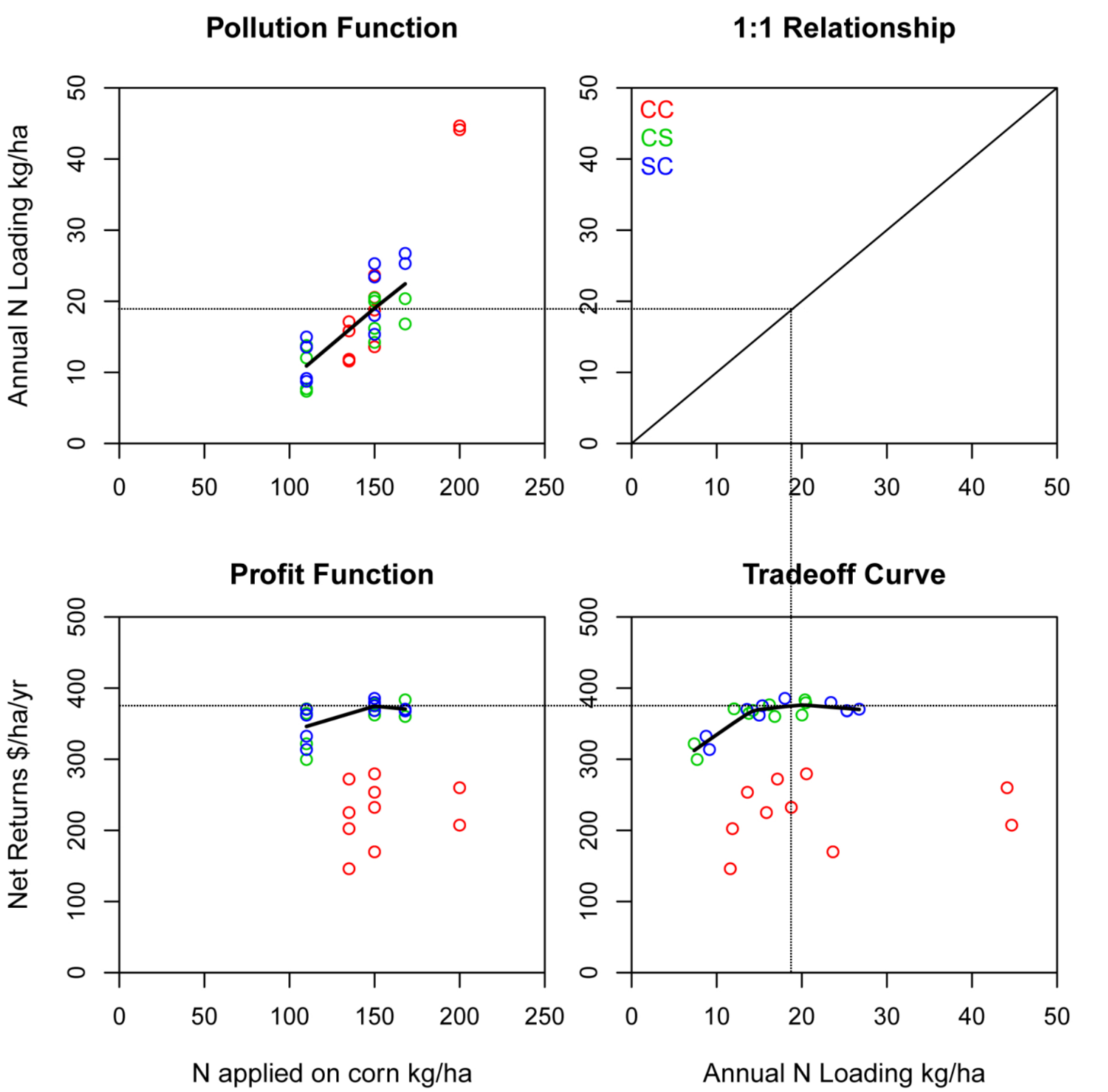

Net Returns $/ha
1990−2003
0
100
200
300
400
500
600
1974−2003
CC
CS
SC

Figure 7



814
815 Figure 8
816



817
818 Figure 9
819