Accessible integration of agriculture, groundwater, and economic models using the Open Modeling Interface (OpenMI): methodology and initial results

T. Bulatewicz¹, X. Yang², J. M. Peterson⁴, S. Staggenborg³, S. M. Welch³, and D. R. Steward²

¹Kansas State University, Dept. of Computing and Information Sciences, Manhattan, Kansas, USA
²Kansas State University, Dept. of Civil Engineering, Manhattan, Kansas, USA
³Kansas State University, Dept. of Agronomy, Manhattan, Kansas, USA
⁴Kansas State University, Dept. of Agricultural Economics, Manhattan, Kansas, USA

Received: 29 October 2009 – Published in Hydrol. Earth Syst. Sci. Discuss.: 25 November 2009
Revised: 6 March 2010 – Accepted: 9 March 2010 – Published: 16 March 2010

Abstract. Policy for water resources impacts not only hydrological processes, but the closely intertwined economic and social processes dependent on them. Understanding these process interactions across domains is an important step in establishing effective and sustainable policy. Multidisciplinary integrated models can provide insight to inform this understanding, though the extent of software development necessary is often prohibitive, particularly for small teams of researchers. Thus there is a need for practical methods for building interdisciplinary integrated models that do not incur a substantial development effort. In this work we adopt the strategy of linking individual domain models together to build a multidisciplinary integrated model. The software development effort is minimized through the reuse of existing models and existing model-linking tools without requiring any changes to the model source codes, and linking these components through the use of the Open Modeling Interface (OpenMI). This was found to be an effective approach to building an agricultural-groundwater-economic integrated model for studying the effects of water policy in irrigated agricultural systems. The construction of the integrated model provided a means to evaluate the impacts of two alternative water-use policies aimed at reducing irrigated water use to sustainable levels in the semi-arid grasslands overlying the Ogallala Aquifer of the Central US. The results show how both the economic impact in terms of yield and revenue and the environmental impact in terms of groundwater level vary spatially throughout the study region for each policy. Accessible integration strategies are necessary if the practice of interdisciplinary integrated simulation is to become widely adopted.

1 Introduction

Policy for large-scale water resource management has broad impacts spanning ecologies, economies, and societies. Such policy must promote sustainable use of water resources to mitigate global trends of increasing water use and uncertain availability. Sustainability requires that competing interests are balanced in an environmentally sound, economically viable, and socially acceptable way. This balance can only be achieved through a clear understanding of the effects of policies on these coupled human-natural systems prior to their adoption. Impact assessment to inform this understanding has traditionally been performed within a particular domain or independently in multiple domains. This disjoint disciplinary science precludes the investigation of cross-discipline effects and feedbacks despite their central importance in many contexts. The need for comprehensive, science-based impact assessment that crosses traditional domain boundaries has motivated a variety of qualitative and quantitative approaches.

Integrated Assessment (IA) is a comprehensive approach to understanding complex environmental processes that combines knowledge and tools across disciplinary domains to guide management decisions (Rotmans and Van Asselt, 1996). The need for integrated assessment in addressing
current global environmental challenges has been well-documented (Parson, 1995; Rotmans and Van Asselt, 1996; Harris, 2002; Pahl-Wostl, 2002; Parker et al., 2002; Jakeman and Letcher, 2003) and water resources are no exception as they are intimately tied to the human and natural systems that depend on them. Parker et al. (2002) identify five dimensions of integration: disciplines, models, stakeholders, scales, and issues, and stated that the most comprehensive form of integrated assessment includes all five. Integrated Assessment and Modeling (IAM) contributes quantitative results along these five dimensions as a supplement to other forms of assessment. The intent of IAM is to combine existing models and methods that were previously employed through isolated disciplinary procedures, approaches, methodologies, and ontologies. While it remains to be seen whether integrated assessment establishes a unifying theory with broadly-accepted methods, it is clear that the transition from traditional reductionist approaches to simulation will require methods and tools for interdisciplinary, integrated modeling.

On this impetus, there has been increasing interest in multidisciplinary integrated modeling, from informal collaborations between researchers within and across institutions (Gomann et al., 2005; Ahrends et al., 2008; Bithell and Brasington, 2009) to international alliances uniting scores of scientists (van Itersum et al., 2008). Governmental recognition for the need of integrated modeling is evident through initiatives such as the European Union’s Water Framework Directive and the US Department of Agriculture’s refactoring of existing models to support integration (Carlson et al., 2009). Despite the success of such work, challenges remain that deter the broad practice of integrated modeling. The greatest of these is arguably the extent of software development necessary, which is particularly prohibitive for small interdisciplinary teams of researchers. Thus, there is a need for effective and agile methods for model integration that require minimal software development yet enable multidisciplinary investigation while building capacity for further collaboration through the exchange of knowledge and tools.

As a result of decades of research in the field of modeling and simulation, there is a wealth of established and trusted models across nearly all scientific disciplines, and as a result of increased interest in integrated modeling over the past decade, there are a variety of methods and tools for linking models. By leveraging this existing work, the extent of software development necessary for model integration is minimized. In addition, through the use of standards-based linking techniques, integration efforts contribute to building capacity for further studies and collaboration between scientists and across institutions. Thus, model linking using existing models and tools is an accessible means to integrated modeling for small interdisciplinary teams of researchers enabling the simulation of human-natural systems.

Irrigated agricultural systems (Fig. 1) possess close interactions between biological, hydrologic, and economic processes. In such a system, policy that regulates the consumption of water has important implications for agricultural production and in turn the economic viability of the farmer and the rural society as a whole. Through the integration of agricultural, groundwater, and economic models, these cross-domain impacts can be simulated. In this paper, we describe how we used the linked-model approach to build an integrated model for the purposes of studying process interactions in irrigated agricultural systems in semi-arid grasslands. We present the integration methodology next, followed by a description of the models and how we created reusable modeling components from them. The efficacy of this integrated approach is illustrated by studying the feedbacks and interactions across the agricultural system.

2 Methods

2.1 Model linking method

Model linking, that is, enabling individual models to collectively conduct a simulation, is typically a challenging task due to the inherent differences in how models are designed and used in different disciplines (Bulatiewicz and Cuny, 2005). Differences between models vary from their implementation (e.g. programming languages) to their conceptualization (e.g. spatial and temporal discretization). A fundamental challenge in linking models is resolving these incompatibilities. Furthermore, high-frequency interaction between models demands automated methods that allow the linked model to execute autonomously. The need for model linking (also called model coupling) has motivated considerable work toward methods and tools (Larson et al., 2001; David et al., 2002; Buis et al., 2006; Bulatiewicz and Cuny, 2006; Ford et al., 2006; Joppich and Kurschner, 2006; Gregersen et al., 2007).

These tools provide a mechanism for model programs to exchange meaningful data, enabling the modeled quantities of one simulation to be incorporated into the simulation of another model as the simulations are being carried out, thus enabling the modeling of process interactions. Extreme approaches are to either create a specialized tool that communicates and translates quantities between a particular set of models, or enable the models to interact through a general-purpose tool using a standard protocol.

Specialized tools are appropriate in situations where a model is only intended to be linked to a specific set of models in a fixed configuration or when general-purpose tools prove too restrictive or impractical. In cases where a model may be linked to a variety of other models in different configurations, it is advantageous to employ a standard protocol that supports experimentation through rapid reconfiguration of the participating models and the interactions between them without the need for programming. The use of standard protocols also
promotes collaboration between modelers across institutions and the establishment of a common repository of reusable model components.

For these reasons, we have utilized the Open Modeling Interface (OpenMI) (Moore and Tindall, 2005; Gregersen et al., 2007) which describes a standard way for models to exchange data with each other. The interface defines a set of capabilities that a model must possess in order for it to be linked to other models. These capabilities are both descriptive, to support the task of specifying model interactions at the domain level, and functional, to support the execution of a linked model. To fulfill the descriptive requirements, a model must be capable of providing a list of the domain quantities that it simulates and those that it uses as input, along with the units and spatial distribution of each. These are called output exchange items and input exchange items. To fulfill the functional requirements, a model program must be capable of sending and receiving data (that corresponds to the exchange items) during execution. A model that meets the requirements of the interface is called a linkable component and can be linked to other models in a highly automated way. Using visual software tools (OATC, 2009), a scientist chooses a set of linkable components of interest, interactively specifies the quantities to be exchanged between them, and then executes the linked model. Components can be easily added and removed and the links between them reconfigured, facilitating experimentation and rapid prototyping of linked models.

The extent of software development that is required to conform an existing model to the interface varies and is most influenced by whether or not the model itself is modified in the process. Changing a model source code to meet the requirements of the interface permits complete control in how the capabilities are supported by the model, but is not always possible or practical. In cases where a model source code is not available or is too complex to modify (or re-verification of the modified model is not feasible), the model can be left unchanged and treated as a black box to which inputs are given and from which outputs are read. This reuse strategy has been referred to as full model reuse (Robinson et al., 2004) and it can substantially reduce the software development effort involved in the process, but may impose limits in the operation of the linkable component, for example, in the frequency at which it can exchange data with other components. When a model program is treated as a black box, then the greatest frequency that it may exchange data is the minimum time span of a single execution of the model. Whether the time span is acceptable or not depends on the particular model and the needs of the study. For example, consider a groundwater model and surface water model that are to be linked at time scales appropriate for each model (hours to seasons for surface water and hours to decades for groundwater). If linkable components are created for these models using the black-box approach, then the models would not be able to exchange information at intervals that are shorter than the simulation period of a single run of either model. So if
a single run of the groundwater model simulates a time period on the order of months or years, then data could not be exchanged at a sufficient rate to capture daily interactions between the groundwater and surface water. In such a case, the models would require modification to enable them to exchange data after each internal time step. In addition, a model must not rely on an interactive graphical user interface for input and/or output, since this would prevent a wrapper program from accessing the input and/or output. In this work, we have adopted the black box approach to conforming the relevant models to the interface. We use the term black box to refer only to the model program; knowledge of the model itself, its theoretical basis, concepts, and assumptions is essential to the sound application of any model. We present the models next, followed by the design of the linkable components.

2.2 Domain models

Both human and natural processes interact at the parcel level in irrigated agricultural systems (Fig. 1). Agricultural decision making, including crop choice and irrigation practice, is influenced by external factors such as crop value and water-use policy as well as internal factors such as water availability and the cost of pumping. The yield that results from crop growth is limited by water availability through irrigation and precipitation and other weather variables such as solar radiation and relative humidity.

2.2.1 Agricultural model

We used the Erosion Productivity Impact Calculator (EPIC) model to serve as the agricultural model. EPIC (Williams et al., 1990) estimates crop growth (biomass) on a daily time step based on intercepted radiation estimates and radiation use efficiency. Biomass accumulation duration is determined by user input or when minimum temperature limits occur. Grain yields are estimated as a function of crop harvest indices. Crop water use is estimated based on potential evapotranspiration and estimated light interception.

Plant water supplies are determined from soil profiles which are controlled by user input. Up to ten soil layers can be defined to depths appropriate for each crop. For corn, soil profile depth was set to 2 m. Input values for soil profiles that are important for estimating soil water content on each daily time step include: layer depth, soil texture, volumetric water content at −0.033 and −1.5 MPa, and bulk density. System inflows include precipitation and irrigation, with soil water outflows are crop water use, soil evaporation, and drainage. EPIC includes an automatic irrigation feature in which irrigation dates and amounts are a function of plant stress. Soil losses (Bhuyan et al., 2002), groundwater recharge (Laroche et al., 1999), climate change (Brown and Rosenberg, 1997; Dhakhwa et al., 1997; Tan and Shibasaki, 2003), soil C sequestration, soil water content (Lee et al., 1996; Costantini et al., 2002; Huang et al., 2006), and cropping systems management have all been evaluated with EPIC (Warner et al., 1997).

2.2.2 Groundwater model

Groundwater provides a reliable source of water for irrigated agriculture to supplement precipitation in crop production. Groundwater stores are replenished by recharge, which is the portion of precipitation and irrigation that filters downward past the root zone of plants to the top of the groundwater table where pores between granular aquifer media are fully filled with water. Groundwater moves from regions with higher groundwater elevation to lower elevation and eventually emerges to the land surface through withdrawal in wells, baseflow to streams, or uptake by phreatic vegetation with roots that tap the top of the groundwater table. The groundwater model predicts the elevation of groundwater, the velocity of groundwater movement, and the impact of changes in natural or anthropogenic forcings, such as climate change or changes in pumping practices.

The computation model adopted for this study is founded in the Analytic Element Method (Strack, 1989). The model utilizes mathematical expressions that exactly satisfy the continuity of flow equation as well as Darcy’s Law, which is a constitutive relation between groundwater flow and changes in groundwater elevation. Equations have been developed to model recharge and irrigation in pumping wells (Steward et al., 2009), changes in bedrock elevation (Steward, 2007; Steward et al., 2008), and groundwater ecohydrologic interactions with plants (Steward and Ahring, 2008). The data to support this study is readily available in GIS format (Steward and Bernard, 2006b; Yang et al., 2009).

2.2.3 Economic model

The economic model predicts irrigators land-use choices on irrigated fields each year. A common modeling approach in the groundwater economics literature is an optimization procedure (e.g., mathematical programming) structured to reproduce a land managers annual decision process (Young et al., 1986; Bernardo et al., 1993; Burness and Brill, 2001). Such models typically maximize profits subject to land- and water-use constraints. Here, we follow a common method from the literature on the economics of land use (Lichtenberg, 1989; Hardie and Parks, 1997; Wu et al., 2004) in which land use choices are predicted by probabilistic equations that were econometrically estimated from land managers observed choices. The two methods are conceptually equivalent, as the form of the probabilistic equations is derived under the assumption of individual optimization. The econometric approach was selected for our study because of the field-specific nature of the land use data, where the observed decisions are typically the corner solutions of planting.
all available land to a single crop. These data are not well suited to the available techniques for calibrating optimization models (e.g., Howitt, 1995).

The modeling framework is described by Hendricks (2007) and Steward et al. (2009). Irrigators are assumed to choose the crop on each field that maximizes expected income i.e., expected crop revenue less production costs. Revenue of each crop is determined by its expected price and its expected yield, which in turn depend on site-specific factors such as soil type. Variation in production costs of a particular crop can be explained by variation in the application cost of water, which depends on variation in energy prices and in pumping lifts. The probability of a particular crop being planted is assumed to be an increasing function of its expected income. The multinomial logit model (Maddala, 1983) was applied to estimate functions that relate the probability of each crop being planted to the underlying causal factors (crop prices, energy prices, pumping lift, and other site-specific variables). Data to estimate the model consisted of 1956 field-level observations from the study region spanning the period 1991–2004. This dataset contains records on irrigated fields in the county that were planted to the four most common irrigated crops: alfalfa, corn, sorghum, and soybeans. Once estimated, the logit equations constitute the simulation model to predict land use choices in the study region. When invoked for a particular simulation year, the model randomly assigns one of the four crops to each irrigated parcel, where the logit equations and field-level data govern the probability that each crop is selected.

2.2.4 Related work

Within the disciplinary context of integrating agricultural, groundwater, and economics models, there has been considerable research dating back to the 1960s (Burt, 1964, 1966). This literature is reviewed by Koundouri (2004), Brouwer and Hofkes (2008), and Harou et al. (2009). As Harou et al. point out, one of the criticisms of previous integrated models is that they have drastically simplified certain model components and spatially aggregated the base data. For instance, the hydrologic relationships are often reduced to simple mass balance equations that update the water stock between periods over a large region. This is particularly true when models are integrated in a holistic (Braat and van Lierop, 1987) manner in which a single model simulates all disciplinary processes endogenously (i.e. economic and hydrologic equations are solved simultaneously). Explicit hydrologic models can predict spatially detailed water level changes that are often relevant for management decisions at a regional scale. Such models are better accommodated by modular (or compartmental Braat and van Lierop, 1987) integration in which independent models transfer data. Data is either transferred in one direction from one model to another, or there is a bidirectional exchange.

Data transfer between models has been realized in different ways with varying levels of automation. Draper et al. (2003) used an economic model to estimate economic benefit and loss functions that were then used as input to a water resource optimization model. In Volk et al. (2008), ecological and economic models exchanged data through an intermediary GIS that stored model inputs and outputs. Ahrends et al. (2008) coupled a groundwater model to an economic model employing a custom program that executed the models and translated input/output files in an automated fashion. The IWRAM DSS (Jakeman and Letcher, 2003) employed a variety of modular approaches, from creating component models in a common programming language or software tool allowing them to be executed together, to creating an integrating engine that executed customized models written in different languages (Cuddy et al., 2005).

Two of the fundamental challenges of the modular approach are (1) identifying the appropriate data exchanges and (2) enabling the models to transfer data (Cai et al., 2003; Brouwer and Hofkes, 2008). Our work is one of the first studies to utilize a standardized, modular linking interface that provides automated execution and data exchange between multidisciplinary models without requiring them to be modified.

2.3 Creating linkable components

To make a model into a linkable component, it must be given the capabilities required by the Open Modeling Interface. Since we did not wish to modify the models, the capabilities were implemented in a helper program called a wrapper that acts as a broker between the model and other linkable components. The wrapper meets the descriptive and functional requirements of the interface and invokes the model program for specific time spans as necessary. The basic functional requirements of a linkable component are that it can (1) accept input quantities that are incorporated into its simulation (set values), (2) advance the simulation forward in time, and (3) provide output quantities to other models (get values) (although some components may only support (1) or (3) and not both). Note that a linkable component only advances its simulation time in response to a request for a quantity at a later point in simulation time. We used a common design (Fig. 2) for the three model wrappers in which the wrapper generates the model input files for the next time span to simulate, including any input quantities that it received from other linkable components, executes the model, and then reads the output files from which it extracts the output quantities that can be provided to other linkable components. After each execution of the model, the wrapper advances its internal clock accordingly. To enable the component to exchange quantities with other components at the highest frequency possible, the wrapper executes the model for the shortest time span supported by the model, which is 1 year in the case of these.
models. The software development effort involved in creating the wrappers is minimal since the wrapper need only keep track of the current simulation time and perform simple file input and output. The wrapper is written in the C# programming language and the models are written in FORTRAN or Scilab. Such differences in the modeling tools customary in different disciplines are common and are easily accommodated by the wrapper design.

The quantities that we chose to expose in each component are given in Table 1. Most of the primary inputs and outputs of each model are exposed in the component. The more inputs and outputs that are exposed, the greater the component’s potential for reuse. During execution, when a component provides a quantity to another component or accepts a quantity from another component for a specific point in time, a set of values is exchanged where each value represents the quantity at a different geo-referenced location. These locations can be points, line segments, or polygons, and each quantity is associated with a group of these geographic elements, called an element set. The processes simulated by these three components exchange quantities that are associated with pumping wells, so the element set associated with these quantities is the set of 2-D points at which pumping wells are located throughout the study region. Thus, each value exchanged between the components is associated with a specific 2-D point (one of the elements in the set) where a well is located. The groundwater and economic models employed in this study are spatially distributed, but the agricultural model is not. For this reason, the wrapper for the agricultural model is slightly different than the other wrappers in that it executes the model once for each point in the element set, enabling the linkable component to perform a spatially distributed simulation even though the model itself is a lumped model.

A model need only be made into a linkable component once and can be reused thereafter. We found that it required approximately one week for a skilled software developer to create working components from the three models. Since all components share the same interface, any component can be linked to any other (although it is up to the scientist to ensure that the linkages are meaningful). In this way, utilizing standard interfaces for linking models builds capacity for further work and collaboration, and contributes a common repository for reusable model components.

### 2.4 Creating the integrated model

The system conceptualization in Fig. 1 provides a blueprint for the design of the linked model. The direct process interactions include the extraction of water from the aquifer for irrigation and the economic choice of the crop that is grown, which is based in part on the water depth. These interactions translate into three quantity exchanges between the linkable components as shown in Fig. 3. Recharge to the aquifer from the agricultural model was not explicitly represented in the linkages as recharge past the root zone to the aquifer is buffered through a very deep vadose zone that dampens surficial signals. Instead, the groundwater model uses a long-term average recharge rate.

At the start of the first simulation year, the groundwater component provides the observed saturated thickness to the economic component. The economic component then selects the crops to be grown for that year for each well and informs the agricultural component of its selection. The agricultural component simulates the growth of the crops and provides the groundwater component with the year-end total water pumped by each well. The groundwater component then simulates the water level at the end of the year and provides the saturated thickness to the economic component as the starting thickness for the following year. The circular execution of the components continues until the end of the simulation period is reached.

Using the available software tools, we interactively specified the quantities to be exchanged between the models and then executed the linked model. The execution of a linked

### Table 1. Input and output exchange quantities exposed by the linkable components.

<table>
<thead>
<tr>
<th>Model</th>
<th>Direction</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>Input</td>
<td>Crop choice</td>
</tr>
<tr>
<td>Agriculture</td>
<td>Output</td>
<td>Irrigated water use, Yield, ET, PET, Runoff, Percolation</td>
</tr>
<tr>
<td>Groundwater</td>
<td>Input</td>
<td>Pumping rate, Recharge</td>
</tr>
<tr>
<td>Groundwater</td>
<td>Output</td>
<td>Groundwater head, Saturated thickness</td>
</tr>
<tr>
<td>Economic</td>
<td>Input</td>
<td>Saturated thickness</td>
</tr>
<tr>
<td>Economic</td>
<td>Output</td>
<td>Crop choice, Irrigated water use</td>
</tr>
</tbody>
</table>

![Fig. 2. Model wrapper software design. Dashed lines indicate the flow of data.](image-url)
Fig. 3. Circular pattern of data exchange between the models.

model begins with the component that is marked with a trigger, which is the groundwater component in this case (since it is the last of the components to execute on a given year). The trigger asks the groundwater component to perform its simulation. Before the groundwater component executes a time step, it must obtain the current water use from the agricultural component by requesting it. When the agricultural component receives a request, it advances its simulation time to the requested time.

Similarly, before the agricultural component executes a time step, it must obtain the current crop choice from the economic component by requesting it, which causes the economic component to advance its simulation time to the requested time. After the economic component advances its simulation time, the crop choice is provided to the agricultural component which can then advance its simulation time and provide the resulting water use to the groundwater component, which can in turn advance its simulation time. These dependencies are automatically identified and the order in which the components advance their simulation times and exchange data is fully automated.

3 Study: water policy and agricultural production

By linking together agricultural, groundwater, and economic components, simulations can be performed that capture the dynamics between these processes in irrigated agricultural systems. Policy scenarios can be investigated by re-parameterizing one of the components, and then observing the effects in the other components. In this study we demonstrate how the linked model can be utilized to evaluate the human and natural impacts of two sustainable water use policies.

3.1 Study region: semi-arid grasslands of the Central Plains

The Ogallala Aquifer has supported irrigated agriculture, and in turn the rural societies, throughout the semi-arid grasslands of the Central Plains of the US for 50 years. This has resulted in groundwater declines of up to 30 m (Custodio, 2002) in some areas. The transition to sustainable use of the aquifer will impact ecologies, economics, demographics, and the landscape, all of which must be considered in the development of policy to guide this process. In this retrospective study, we show what the resulting effect of two policies would have been in 2004 for a county in western Kansas had they been adopted in 1991. We chose a study region of one county (Sheridan County, KS, approximately 50 km · 50 km) as this is the typical aggregation level for economic data such as prices and yields. The data sources for each model are given in Table 2.

3.2 Verification of the models

Each model was individually calibrated to the study region in earlier works. The calibration of the agricultural model identified values for 10 model parameters that most closely reproduced observed water-use and yield over the period 1990–2001 (Bulatewicz et al., 2009). Applying the calibrated model to the time period of this study (1991–2004), the RMSE for irrigated water use at the well-level was $64.5 \times 10^3$ m$^3$ and the RMSE for county-level yield was 1.9 t averaged over the four crops included in the calibration. The economic model and groundwater models were calibrated to the time period 1991–2004 in Steward et al. (2009). The economic model predicted the correct crop for 4560 of 9767 observations (47%, single run). Many of these incorrect predictions, however, were a result of the observed crop choice not being included in the model. The prediction accuracy was 74% (4560/6161) within the subset of observations that reported crops that were included in the model.

The groundwater model was calibrated to match the groundwater elevation at observation wells. The residual difference between the simulated and observed groundwater elevation at 70 observation wells throughout the study region was 4.69 m and is within the yearly groundwater elevation fluctuations.

3.3 Verification of the linked model

In linked models, each model must not only be verified when executed independently, but also when executed collectively. When a model is executed in a linked fashion, some of its input data are replaced with simulated data that are provided by other models. If the simulated data differ from the original input data, this may induce greater (or lesser) error in the model results than when the model is executed independently. In linked configurations that include feedback loops, this error can be compounded. Thus, when a model
Table 2. Model data sources.

<table>
<thead>
<tr>
<th>Agricultural Model Data</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water Use</td>
<td>Wilson et al. (2005)</td>
</tr>
<tr>
<td>Irrigation Type</td>
<td>Wilson et al. (2005)</td>
</tr>
<tr>
<td>Crop</td>
<td>Wilson et al. (2005)</td>
</tr>
<tr>
<td>Soil</td>
<td>USDA (2006)</td>
</tr>
<tr>
<td>Weather</td>
<td>High Plains Regional Climate Center (2009)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Economic Model Data</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrigated Acres</td>
<td>Wilson et al. (2005)</td>
</tr>
<tr>
<td>Irrigation Type</td>
<td>Wilson et al. (2005)</td>
</tr>
<tr>
<td>Crop Type</td>
<td>Wilson et al. (2005)</td>
</tr>
<tr>
<td>Weather</td>
<td>Kansas Weather Data Library (2008)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Groundwater Model Data</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-development Saturated Thickness</td>
<td>Gutentag et al. (1984)</td>
</tr>
<tr>
<td>Hydraulic Conductivity</td>
<td>Cederstrand and Becker (1998b)</td>
</tr>
<tr>
<td>Specific Yield</td>
<td>Cederstrand and Becker (1998a)</td>
</tr>
<tr>
<td>Recharge</td>
<td>Hansen (1991)</td>
</tr>
<tr>
<td>Historical Pumping Rate</td>
<td>Wilson et al. (2005)</td>
</tr>
<tr>
<td>Historical Groundwater Level</td>
<td>Hausberger et al. (1998)</td>
</tr>
</tbody>
</table>

is executed in a linked fashion, there is greater importance placed on the accuracy of the outputs that are provided to other models than may be placed on these outputs when executed independently.

This sensitivity to the accuracy of exchanged quantities was observed during the initial trial runs of the linked model. The error in the water level simulated by the groundwater model systematically increased after each year of simulation resulting in a significantly higher water level at the end of the simulation period than was observed. This was a result of the agricultural model’s consistently low prediction of county-wide annual water use. Although the individually-calibrated agricultural model reported in the earlier study achieved low absolute relative error in county-wide annual water use (13% for corn), the total simulated annual water use was considerably lower than observed on each year of the simulation (totaling $89.6 \times 10^6 \text{m}^3$ less than the observed $1388.4 \times 10^6 \text{m}^3$ over all years). This bias toward low water use in the total county-wide annual water use error of the agricultural model made it unsuitable for linking to the groundwater model. With respect to the verification of the groundwater model, the error in the water use simulated by the agricultural model must not only be small, but must also not be biased high or low, so that the error in total water use over the entire simulation period is minimized. This highlights the challenge of ensuring that the quantities provided by a model meet the specific accuracy requirements of the model that consumes them, which may vary from one level of spatial or temporal aggregation to another.

The earlier work in the calibration of the individual agricultural model identified 250 different parameterizations of the model and reported the parameter set with the lowest error. Some of these parameter sets had comparably low error (as they were the result of the same optimization process) yet did not exhibit the same bias toward low simulated water use. We chose one of these parameter sets that (when executed independently) has water use and crop yield error that is equivalent to the reported parameter set and that closely matched the total observed water use over the entire simulation period. These are given in Table 3.

When the agricultural model was executed independently with the alternative parameter set, the simulated total county water use was $31.7 \times 10^6 \text{m}^3$ lower than observed ($947.1 \times 10^6 \text{m}^3$). The RMSE of water use per well was $61.6 \times 10^3 \text{m}^3$ and 1.9 t for yield at the county-level, averaged over the four crops. When executed in a linked fashion, the
Table 3. EPIC parameter values used for the simulation of corn.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRI</td>
<td>Minimum application interval for automatic irrigation</td>
<td>11 d</td>
</tr>
<tr>
<td>BIR</td>
<td>Water stress factor to trigger automatic irrigation</td>
<td>0.83</td>
</tr>
<tr>
<td>ARMN</td>
<td>Minimum volume allowed for automatic irrigation per application</td>
<td>9.3 mm</td>
</tr>
<tr>
<td>ARMX</td>
<td>Maximum volume allowed for automatic irrigation per application</td>
<td>32.8 mm</td>
</tr>
<tr>
<td>WA</td>
<td>Biomass to energy ratio</td>
<td>49.0 t ha⁻¹ MJ⁻¹</td>
</tr>
<tr>
<td>TB</td>
<td>Optimum temperature for plant growth</td>
<td>34.5 °C</td>
</tr>
<tr>
<td>TG</td>
<td>Minimum temperature for plant growth</td>
<td>8.38 °C</td>
</tr>
<tr>
<td>DLAI</td>
<td>Fraction of growing season completed when the ratio of leaf area to ground area begins to decline</td>
<td>0.83%</td>
</tr>
<tr>
<td>RLAD</td>
<td>Leaf area decline rate</td>
<td>5.45</td>
</tr>
<tr>
<td>RBMD</td>
<td>Biomass-energy decline rate</td>
<td>2.84</td>
</tr>
</tbody>
</table>

total water-use was only $2.5 \times 10^6$ m$^3$ lower than observed ($1388.4 \times 10^6$ m$^3$) with similar error to the original parameter set (water use RMSE per well of $69.2 \times 10^3$ m$^3$ and average 1.7 t for yield at the county-level). Using the alternate parameter set for the agricultural model, the groundwater model water level error was 4.76 m, comparable to the error when executed alone. The accuracy of the economic model was minimally affected, resulting in an error of 46% (4524 correct of 9780 observations) and 73% (4524/6170) when the observation was one of the modeled crops. The results of the linked model for the historical scenario (in which existing policies were used) is shown in Fig. 4a. The net change in the saturated thickness (left) between the predevelopment groundwater table and the end of the simulation period is shown, and the size of the wells are scaled based on the total water used over the simulation period. The total revenue of each parcel over the simulation period is also shown (center), where each well is associated with the PLSS land unit in which it resides. This was estimated by multiplying the annual yield of the crop grown on each parcel by the average price of the crop that year. The crop that was chosen by the economic model most often over the 14-year simulation period is shown in the map on the right, and was predominantly corn. The 13 parcels marked as “multiple” indicates that more than one crop was associated with a parcel.

These results show that the area facing the most intensive irrigation results in the greatest agricultural production and the largest decline in groundwater level. It is clear that this level of agricultural production cannot be sustained through irrigated water use, but it is not clear what effects policy for reduced water use may have on production levels. To investigate this question, we consider two policies.

### 3.4 Water policy scenarios

We considered two policies for limiting water use to sustainable levels adopted from Steward et al. (2009): regulating water use, and offering an incentive-based water-right buyback program. The regulation policy strictly enforces the prior appropriation doctrine to limit groundwater consumption to match the natural recharge rates in a specific high-priority area (outlined in Fig. 4b). Within the high-priority area, the total water use would have to be reduced by 90% to match natural recharge, based on the five-year average water use of each well prior to 1991. This resulted in the removal of 173 wells from the high-priority area leaving the 23 wells with the most senior rights. These wells were removed from the agricultural model and all other parameters of the three models remained the same as in the historical scenario.

The results in Fig. 4b indicate that the policy results in less groundwater decline in the high-priority area with a maximum water level decline of 9.8 m as compared to 22.1 m in the historical scenario. The resulting economic impact is also considerable, reducing the total county revenue by 30% from $323.2 \times 10^6$ to $229.0 \times 10^6$ over the study period (Table 4).

In the incentive policy, the same net water savings is achieved as in the regulation policy, but is done so through a county-wide voluntary water-right retirement program. It is assumed that each landowner would be offered a government payment just equal or slightly above his losses in income from the water right. This payment would vary from landowner to landowner. A reduction of 34% in total water use over the entire county is necessary to match natural recharge (again, based on the 5-year average water use of each well and the average recharge over the region). Since
Fig. 4. Linked model results for three water use policies.
participation in the water-right buy-back program is assumed to be random, we performed 30 trials of this simulation, each with a different random sampling of participating wells. On average, each trial included 542 of the 729 wells in the county (st. dev. 10). These simulations were performed on a computing cluster due to the significant runtime (2 days) of the linked model which allowed the 30 trials to be performed concurrently.

The results of this policy in terms of maximum groundwater level decline, total revenue, and crop selection are illustrated in Fig. 4c. The groundwater decline indicated in the figure is an average of the maximum groundwater level decline at the end of the simulation period over all 30 trials. The minimum water level was 14.9 m, as compared to the 22.1 m in the historical scenario. The revenue per parcel (center) is the 30-trial average of the total revenue produced by each parcel over the simulation period. The crop choice indicated for each parcel (right) is the crop that was chosen most often over all trial-years (30 trials-14 years). The 34% reduction in groundwater consumption resulted in a 30% reduction in both yields and revenue throughout the county.  

1 Under a voluntary incentive policy, the revenue losses would be offset by the government payments to buy the water rights, so that the net reduction in local revenue would be less drastic than

The variation of the results of the 30 trials is shown in Fig. 5 in terms of standard deviations for per-parcel revenue (center) and groundwater use and water level change (left). The greatest variation in pumping rates and the resulting groundwater level is in the area of the most intensive irrigation. The variation in crop choice (right) is presented as the number of number of different crops chosen for a parcel over the 30 trials. In most cases, the crop chosen for a parcel was the same in all 30 trials (crops only differed between trials for 44 parcels).

4 Conclusions

In this work we have presented an approach to constructing multidisciplinary integrated models through the linking of individual domain models. The software development effort was minimized through the reuse of existing models and model-linking tools. Although the utility of model reuse has been debated (Paul and Taylor, 2002; Pidd, 2002; Robinson et al., 2004), we have found the black box approach to model reuse to be an effective way to simplify the task of creating reusable modeling components from existing models. The flexible design of the OpenMI accommodates this approach for both hydrological models, for which the interface was designed, as well as models from other domains, thus providing a suitable medium to bring together traditionally isolated modeling disciplines. Utilizing an open standard in integrated modeling studies means that the components developed for a particular study can be reused in other linked studies shown in Fig. 4c. However, the figure accurately reflects overall income losses because the government payments simply transfer the economic burden from local landowners to taxpayers.
models without the need for additional programming, and thus contributes to a common repository of model components.

We employed this approach to construct an integrated agricultural-groundwater-economic model. The reusable model components were designed as lightweight wrappers around the model programs (Fig. 2). The agricultural component was designed to support spatially-distributed simulation by repeated application of the underlying field-scale model program. The linked model was configured such that the relevant quantities were exchanged between pairs of components in a circular fashion (Fig. 3).

A case study demonstrating the utility of the coupled model in the context of irrigated agricultural systems was performed. Data were compiled for a study region in the semi-arid grasslands overlying the Ogallala Aquifer in the Central Plains of the US. The models were verified both individually and collectively as the integrated model to reproduce historical irrigated water use, crop choices and yields, and groundwater decline. The initial parameter set selected for the agricultural model was found to induce error in the groundwater model that compounded on each year of simulation. Through the use of an alternative parameter set, the error was reduced and the overall consistency of the linked model was maintained. This highlights the challenge of ensuring that the quantities provided by a model meet the specific accuracy requirements of the model that consumes them, and that the error is not accumulated or amplified beyond acceptable limits on successive iterations of the linked model.

The results of the integrated model using historical data from 1991 to 2004 are presented in Fig 4. The results of the integrated model for the retrospective application of two water-use policies that achieve sustainable levels of groundwater consumption were presented in terms of economic productivity and change in groundwater use/levels both spatially (Fig. 4) and over the complete study region (Table 4).

The presented case study is an example of the kinds of comprehensive simulations that are possible through integrated modeling. The results illustrate the spatial patterns of changes in water-use, production, and revenue that result from two alternative policies aimed at sustainable water-use. Although both policies produce the same county-wide effect, they do so with differing impacts throughout the region, illustrating the effect of policy that targets specific geographic areas. The integrated model provides a means for further studies that investigate the impacts of policy across multiple domains and is currently being used in a study to evaluate policies for limited irrigation.

Multidisciplinary integrated models play an important role in the integrated assessment of policy that impacts societies and ecologies. Such quantitative methods provide unique insight into the broad effects of water resource policy in coupled human-natural systems and address the needs of agencies and stakeholders for integrated assessment in the development of policy. For integrated modeling to become common practice, practical methods for building such models are necessary, and we believe that this work contributes to that need.

Acknowledgements. This work was supported in part by the National Science Foundation (grants NS9889;GECE001505 and EPS0553722), the US Department of Agriculture/Agriculture Research Service (Cooperative Agreement 58-6209-3-018), and the Provost Office’s Targeted Excellence Program at Kansas State University. Any findings, opinions, conclusions, or recommendations expressed herein are those of the authors and do not necessarily reflect the views of any funding units.

Edited by: P. van der Zaag

References


Hendricks, N. P.: Estimating irrigation water demand with a multinominal logit selectivity model, Department of Agricultural Economics, Kansas State University, Manhattan, Kansas, 134 pp., 2007.


Lichtenberg, E.: Land quality, irrigation development, and cropping patterns in the Northern High Plains, Am. J. Agr. Econ., 71(1),


