4. PROJECT DESCRIPTION

4.1 Status and Overview—*The Challenge*

Ecological systems are breathtakingly complex. They encompass the responses and feedbacks of the Earth's biosphere to changes in the geosphere, hydrosphere, and atmosphere across space and time and orders of magnitude. Ecosystems both drive and respond to global change in non-linear ways (Luo 2007; Zhou *et al.* 2008). They involve the behaviors and interactions of the planet's more than 15 million species, from bacteria to humans, with Earth's physical and chemical environment. A grand challenge for the 21st century is understanding ecological systems and forecasting their responses to global change, both natural and anthropogenic, as articulated by the National Research Council (http://www.neoninc.org/about-neon/overview.html), the U.S. Global Change Research Program (www.usgcrp.gov/usgcrp/Library/ocp2009/ocp2009-focus.htm) and other studies. Addressing this challenge is especially critical for grasslands, a vital ecosystem that provides resources and services to human societies worldwide and is the foundation of the life and economy of the Central Plains. As such, grasslands are the focal ecosystem of this proposal.

Ecological forecasting encompasses complex reciprocal impacts among primary drivers and consequences: climate, land-use, and biotic, biogeochemical and hydrological dynamics. Understanding these phenomena is fundamental to sustaining grassland ecosystem services: supplying clean water, recycling essential nutrients, sequestering carbon, preserving biodiversity, and guarding against invasive species and emerging diseases. Natural factors (climate, biotas) and human use of grassland systems for commerce and agriculture have had long-term ecological impacts (Dodds *et al*. 2004), affecting hydrologic systems, soil moisture, terrestrial productivity, biodiversity, species invasions, zoonotic diseases, population structure, and community and ecosystem patterns and processes.

The ability to address the grand challenge of ecological forecasting is severely hampered by the absence of a unified cyberinfrastructure (CI) that mediates the science framework—i.e., data, models, analytics, and narratives—across a CI framework of hardware, software, and collaborative and integration environments (http://www.nsf.gov/pubs/2007/nsf0728/index.jsp?org=NSF). This disparity between the science and CI frameworks keeps powerful CI technologies from fostering cross-domain research on complex ecological phenomena. For the most part, CI tools are advancing research along single, independent pipelines of ecological and biotic data, models, analyses and narrative scenarios, and have yet to fulfill the potential for enabling unified, collaborative research across these pipelines. Essentially, the enormous revolution in CI technologies has not yet been fully harnessed for making complex, interdisciplinary ecological research collaborative for investigators, tractable for science, and beneficial to society. Such harnessing, proposed here for enabling ecological forecasting, has occurred successfully in computational biochemistry, astrophysics and other fields.

During the past two decades, ecology and biodiversity science have evolved from data-limited to data-rich disciplines. Observatory networks with automated, long-term, in-situ measurements, space-borne observatories, and digitized biocollection databases now provide massive streams and stores of data for ecoforecasting. Many of these data lack sufficient resolution and/or geospatial context for modeling at scales amenable to local and regional forecasting of ecological processes. Advanced informatics, networking and computational tools are needed to assimilate, manage, integrate and model climate, land-use/land-cover and biogeochemical data in near real-time. A similar digital data framework is required for deploying biodiversity data for deciphering and modeling species-level biotic phenomena and their impact on ecosystem function. Examples here include: georeferenced spatiotemporal distributions, biology and natural history of native species, invasives, and disease zoonotics; and geospatial biodiversity and epidemiological models that analyze these data to characterize and predict species' spatiotemporal distribution under past, current and future environmental conditions. Advanced analytics and narratives—i.e., driver- and consequence-based scenarios—are needed to integrate and test unified models of ecosystem dynamics.

Despite their individual strengths, no broad-based CI framework or capacity exists in the four partner universities (Kansas–KU; Oklahoma–OU; Kansas State–KSU; Oklahoma State–OSU) for automated sensing, assimilation, modeling, and analysis of massive sets of abiotic and biotic data across ecological parameters. Collectively, the challenge is the development of a CI environment that unites people in a shared, virtual commons for research and education across ecological and biodiversity disciplines.

The Solution. The solution to this challenge is a synthesis of two frameworks: the *science framework* of data, models, analytics and narratives, and the *CI framework* of hardware, software, collaboration environment and integration environment (Fig. 1). Weaving together the CI and science frameworks (Section 4.4 below) will create a 'cyberCommons'—a powerful, integrated environment for knowledge discovery and education in complex environmental phenomena.

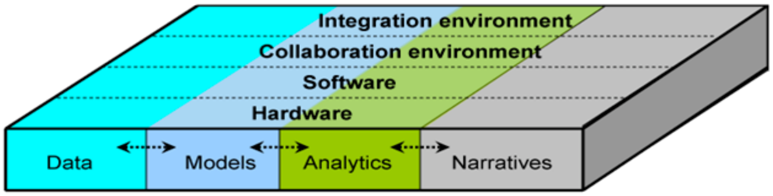


Figure. 1 Integrating the science and CI frameworks.

4.2 Results from Prior Research

Bowman-James, K., PI, L. Krishtalka, W. Dodds, (co-PIs), EPS-0553722, $6.75 M, 4/1/06-3/31/09, Phase V: Building Research Infrastructure to Address Grand Challenge Problems in Ecological Forecasting. *www.k-state.edu/ecoforecasting/*

Researchers at KU, KSU, and Fort Hays State University established a collaboratory for ecoforecasting across the Kansas River Basin to sense, model and predict the effects of climate and land-use on ecosystem function, hydrology and biodiversity, and, in turn, on social systems. The project increased research capacity and expertise with recruitment of four new faculty in biodiversity and ecosystem modeling (KU), and ecophysiology and biogeography (KSU), now sustained by each university. The project also improved critical research infrastructure with the: (1) purchase and installation of ecosystem sensors, 3 carbon flux towers, a large aperture scintillometer, prototype soil moisture network, 10 multi-parameter water-quality sondes, and a mass spectrometer; (2) acquisition of remote sensing data from flyovers and LIDAR images; (3) databasing and georeferencing approximately 300,000 biocollection vouchers of Central Plains animals and plants; and (4) development of an integrated USGS-Kansas statewide soils database for watersheds and ecoregions.

These infrastructure activities increased the spatiotemporal coverage and resolution of essential biotic and abiotic data across the larger Kansas River Basin. Biodiversity data records are served by KU at (http://www.nhm.ku.edu/Hdocs/Collections.html) and can be queried using DiGIR and SPECIFY, an NSF-funded biocollections management program. A new land-cover map of the Kansas River Watershed is served at http://www.kars.ku.edu/projects/ecoforecasting/. GIS scientists programmed the next version of the *Lifemapper* software tool (www.lifemapper.org/index.shtml) for automated modeling and visualization of biodiversity data.

Researchers at KU, KSU and partner institutions integrated these and other tools and data to analyze, model and predict changes in biogeochemical cycles, species populations, and the potential spread of invasive species and emerging diseases, such as West Nile Virus, rabies, and Lyme disease. To date, the research results have been reported in 36 peer-reviewed papers (designated with a † in the References Cited); an additional 19 are in various stages of publication.

Signal research accomplishments include the use of remotely-sensed data to estimate evapotranspiration (Brunsell *et al*. 2008a, 2008b), a framework of scalable parameters for clusters for high-performance computing (Bulatewicz *et al*. 2007), an expanded view of fundamental trophic properties of aquatic ecosystems (Dodds and Cole 2007), ecological economic assessment of wetlands (Gelso *et al*. 2008), improvement in groundwater flow modeling (Jin and Steward 2007, Steward 2007), scaling estimates of ecosystem fluxes to continental scales (Marshall *et al*. 2008), characterization of gas flux rates in grasslands (Tiemann and Billings 2008), a framework for analyzing the biogeography of zoonotic diseases (Peterson 2008), and using biocollections data to monitor and predict the loss of biodiversity (Soberón and Peterson 2008).

The project supported research and training for 33 undergraduates, 38 graduate students, and 4 postdoctoral researchers in a cross-disciplinary collaborative environment. Outreach to state and local government officials increased the potential of the research to inform public policy on health, natural and water resources, and land-use. Importantly, the project created institutional partnerships and united research expertise in ecoforecasting between Kansas and Oklahoma through research symposia and retreats, the result of which is this EPSCoR Track II proposal.

Luo, Yiqi. DEB-0444518, $280,000 2/15/05-2/14/08, Data-model fusion at AmeriFlux sites: Towards predictive understanding of seasonal and interannual variability in net ecosystem exchange.

This research: (1) investigated the application of probabilistic inversion and Markov chain Monte Carlo method to an ecosystem model and isotope data analysis; (2) examined general issues on parameter identification; (3) analyzed seasonal and inter-annual variability in soil respiration and mechanisms underlying biomass growth; (4) and characterized latitudinal patterns of inter-annual variability in net ecosystem production, gross primary production, and ecosystem respiration. The project contributed to development of new software—the prototype of EcoPAD, Ecological Platform for Assimilation of Data, which incorporates the probabilistic inversion approach to data assimilation, and has been adopted by several research groups worldwide and targeted for further development (see Section 4.4.3). To date, the project has resulted in 11 publications (indicated by * in the References Cited) and supported the research of one graduate and two undergraduate students, and one postdoctoral fellow on data-model fusion.

4.3 CI-enabled science: Two critical ecological questions for science and society

Understanding and forecasting ecological states across the Central Plains grasslands encompasses two critical questions identified as grand challenges by the NRC, NSF and other agencies:

1. What are the impacts of changes in land-use/land-cover and climate, both natural and anthropogenic, on biogeochemical cycles and ecosystem function? In turn, what are the feedbacks among these drivers and consequences?

2. What are the impacts of changes in land-use/land-cover and climate, both natural and anthropogenic, on biodiversity—its composition, patterns and dynamics? In turn, how do these changes in biodiversity affect the spread of plant and animal diseases and invasive species, and how do these phenomena influence ecosystem structure, function and services?

The CI challenge in addressing these questions requires advancing four informatics components—data, models, analytics and narratives—for the research community to monitor, diagnose, forecast and synthesize ecological and biodiversity systems under scenarios of changing climate and land-use/land-cover. Section 4.4 (below) describes how the cyberCommons will develop and integrate these informatics components.

Question 1 rationale: Grasslands are highly sensitive to global change phenomena that affect biogeochemical cycles at local, regional and global scales (Albertson and Weaver 1954, Samson and Knopf 1994, Reich *et al.* 2001, Knapp *et al.* 2002, Shaw *et al.* 2002, Briggs *et al.* 2005, Sherry *et al.* 2007). Changes in biogeochemical cycles are both a response to, and driver of, global change (Schlesinger 1997) with feedbacks that affect ecosystem and atmospheric processes, and water quality. Anthropogenic changes in abiotic drivers, including temperature, precipitation, nutrient deposition and CO_2 concentrations, directly influence biogeochemistry and ecosystem structure and function. For example, local and regional grassland productivity is tightly coupled to climate and potential climate change (Knapp *et al*., 2002). Climate change can also alter grassland soil CO_2 flux (Harper *et al*. 2005, Zhou *et al.* 2007). Elevated atmospheric CO_2 concentrations or enhanced regional N deposition rates drive changes in the composition of plant and microbial communities and alter ecosystem processes (Zavaleta *et al*. 2003, Stevens *et al*. 2004, Luo *et al*. 2006, Clark and Tilman 2008).

Changes in local and regional land-use/land-cover also affect biogeochemical processes in the Central Plains. Widespread changes in land-use/land-cover, such as conversion of native grasslands to cropland, and invasion of grasslands by woody species, can substantially alter biogeochemistry, for example, ecosystem-level C and N storage (McKinley and Blair 2008). Likewise, cropland has vastly different rates of N cycling and decreased C sequestration relative to native grasslands (Ajwa *et al.,* 1998), and conversion of cultivated lands back to perennial grasses for the Conservation Reserve Program (CRP) or biofuels leads to long-term changes in N cycling and C sequestration (Baer *et al*., 2002, Baer and Blair 2008, Luo *et al.* 2009). Grassland management also affects soil N availability and trace gas flux (Blair 1997, Johnson and Matchett 2001, Tieman and Billings 2008).

The Central Plains are an ideal laboratory for the study of these changes in ecological states, which are as significant for science as for society. The Central Plains grasslands harbor vast stores of carbon—second only to the Arctic tundra (Schlesinger 1997, Schimel *et al.* 2000)—making it critical to understand the processes that affect net ecosystem C exchange. Along with worldwide grasslands, those of the Central Plains have been altered by anthropogenic activities perhaps more than any other biome on Earth. Land-use changes, such as agricultural conversion, have altered ecosystem structure and function on an unprecedented scale (Samson and Knopf 1994)—changes that are well documented across strong regional precipitation and temperature gradients. Further, rates of land-cover change in the Central Plains, such as conversion of grasslands to woodlands, or to food crop and biofuel production, are likely to increase.

Question 2 rationale: Biodiversity is an excellent indicator of ecosystem function and response to environmental change (Naeem *et al*. 1995, Schlapfer and Schmid 1999, Wardle *et al*. 2000, Loreau *et al*. 2001, Hooper *et al*. 2005, Schmid *et al*. 2002). Biodiversity is a complex phenomenon, including the geographic and ecological distribution of species and their interactions, which are governed by their changing biotic and abiotic landscapes (Soberón and Peterson 2005, Dodds in press). Further, biodiversity patterns and dynamics are intrinsically linked to the spread of invasive species and infectious diseases. Climate change, global connectivity, agricultural demands, and changes in natural and anthropogenic landscapes are increasing species extirpations and the emergence or re-emergence of infectious diseases that threaten humans, biodiversity and ecosystems (Cohen 2000, Bengis *et al*. 2004, Daszak *et al*. 2000, 2007; Jones *et al*. 2008, Dodds 2008). On the Central Plains, biodiversity and ecological processes vary locally and regionally with climate, soils, topography, natural disturbances and land management, which, in turn, are affected by a strong east-west precipitation gradient and a north-south temperature gradient (Dodds 2008).

For example, on the Central Plains, increasing demands for surface and ground water for agriculture and biofuel production is resulting in the extirpation of fish populations. In the U.S., the range of vector-borne infectious diseases (e.g., Lyme disease and West Nile Virus) has expanded rapidly (Chen *et al*. 2006, Diuk-Wasser *et al*. 2006). The spread of West Nile Virus, mediated by birds and mosquitoes over the past decade, has caused declines in several avian species in the U.S. and worldwide (LaDeau *et al*. 2007). Large changes in risk factors (e.g., climate, land-use, social, and economic conditions) are affecting the pathogen-vector-host relationships of infectious diseases, their spatial and temporal dynamics, and their consequent introduction, geographic spread and persistence.

Spatially and temporally explicit biodiversity data are held by many regional, national and international institutions in a variety of forms, including biocollections at museums and herbaria, government agencies, research laboratories, and print publications. Herbaria at the four partnering universities in Kansas and Oklahoma hold more than 1 million vascular plant specimens, roughly 50% of which constitutes the most comprehensive documentation of the floras (and rare and invasive plant species) of the two states and the Central Plains, beginning in the late 1800s. Similar extensive, vouchered biodiversity data document the arthropods (aquatic and terrestrial), fishes, amphibians, reptiles, birds and mammals of the Central Plains.

The four partner Kansas and Oklahoma institutions are among the international leaders in modeling and forecasting the effects of changing climates and land-use/land-cover on biodiversity phenomena, encompassing the potential occurrence, spread and extirpation of mammals, birds, reptiles, fish, insects and plants (Peterson *et al*. 1999, 2002; Soberón 2007, Soberón *et al*. 2007, Soberón and Peterson 2008, Gido *et al*. 2002, Falke and Gido 2006) and the spread of infectious diseases (Peterson 2006, 2007, 2008, Xiao *et al*., 2007, Gilbert *et al*., 2008). The current Kansas EPSCoR RII grant supported digitizing and georeferencing of more than 225,000 biocollections records for Kansas and Central Plains plants, arthropods and vertebrates held at KU and KSU.

4.4 The CI solution: A cyberCommons to integrate the CI and science frameworks

Meeting the challenge of forecasting future ecological states requires a powerful computational paradigm that unites ideas, people and tools across disciplines and institutions in a seamless, dynamic research architecture. The two overarching science questions (Section 4.3) are themselves dynamic, evolving new complexities as new knowledge is discovered. Addressing them requires a cyber-environment for organizing, connecting and harnessing distributed resources (e.g., data, codes, software libraries, hardware, expertise) as reusable services that can be composed by an emergent organization of research teams to address complex, interdisciplinary questions. Such cyber-environments are the emerging medium for integrative, collaborative research and building virtual research and educational communities.

Our envisioned environment is a cyberCommons. Its architecture integrates the traditional CI framework (hardware, software, research environments) to the scientific process framework of data, models, analytics and narratives. These four informatics components (Table 1, next page) are provided as services though individual files, databases and executables for the research community and for deployment in specific environments that target specific questions.

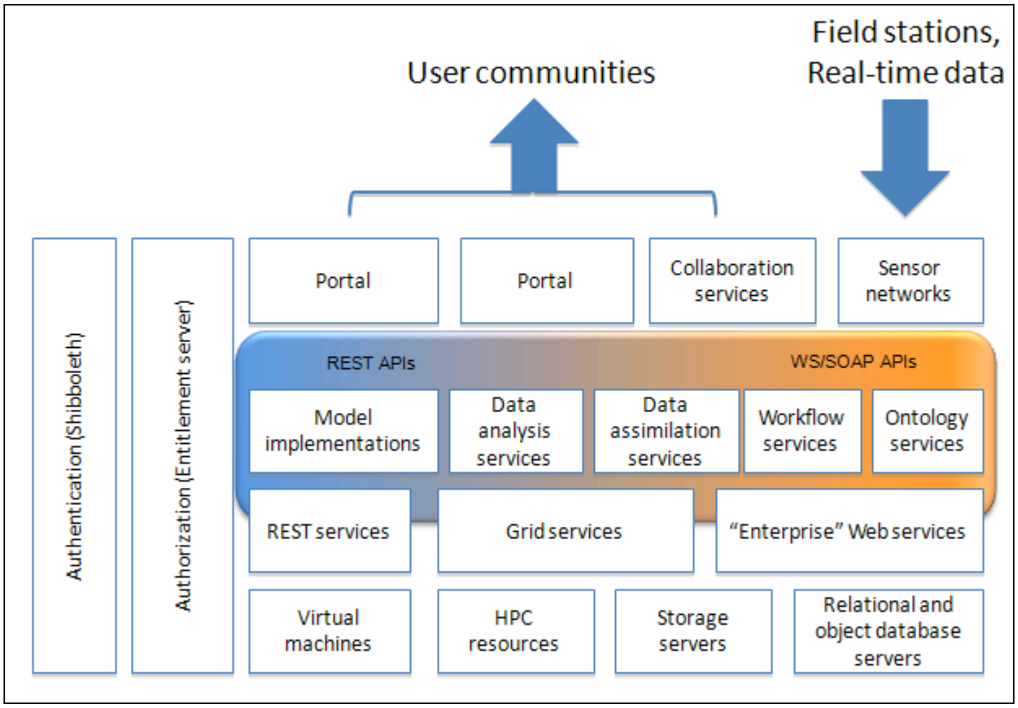


Figure 2. The cyberCommons components

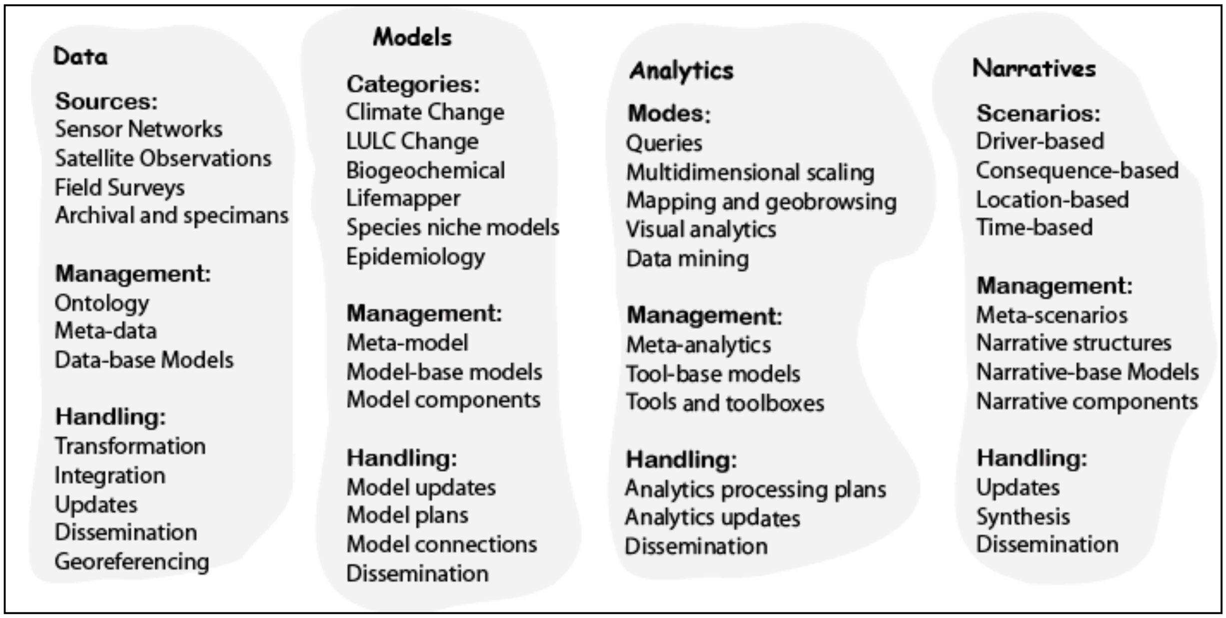
The cyberCommons provides both portal- and service-based functions for investigators to manipulate data, parameterize models, select analytics tools, and formulate narratives (Figure 2). As a web portal, the cyberCommons allows users to browse, manipulate, and download data and models from distributed sites. The network service functions provide mapping, analytics, and computational functions. Light-weight analyses can be served online, such as web GIS mapping and geovisualization. Users will download advanced functions and models to their workstations for analyses. A High Performance Computing (HPC) workbench offers access to supercomputing facilities at OU and other HPC sites.

4.4.1 The science framework—data, models, analytics, narratives

Data sources include sensor measurements, satellite imagery, field surveys and museum vouchered biocollections at different spatial and temporal scales. Data ingestion to mathematical models requires a coherent framework of schemas to transform the data into common georeferenced spatial and temporal values, and common ontological commitments, in turn enabling the development of tools and ecological models for data assimilation and query (Zhang and Gruenwald 2008, Yuan 2007). Common data models such as ArcHydro and OpenMI provide additional tools for bridging ecological data and analytical tools through shared data formats. Building on these approaches to data and metadata, we will develop an object-relational framework for a common ecological data model for the cyberCommons.

The informatics deliverable is: a unified data model using meta-model information, Universal Modeling Language (UML) and Geography Markup Language (GML); ontological schemas defined using the Web Ontology Language (OWL); and the Protégé editor for semantic consistency in the database (http://protege.stanford.edu/overview/protege-owl.html).

Table 1: Informatics components of the cyberCommons science framework



Data

Sources:
Sensor Networks
Satellite Observations
Field Surveys
Archival and specimans

Management:
Ontology
Meta-data
Data-base Models

Handling:
Transformation
Integration
Updates
Dissemination
Georeferencing

Models

Categories:
Climate Change
LULC Change
Biogeochemical
Lifemapper
Species niche models
Epidemiology

Management:
Meta-model
Model-base models
Model components

Handling:
Model updates
Model plans
Model connections
Dissemination

Analytics

Modes:
Queries
Multidimensional scaling
Mapping and geobrowsing
Visual analytics
Data mining

Management:
Meta-analytics
Tool-base models
Tools and toolboxes

Handling:
Analytics processing plans
Analytics updates
Dissemination

Narratives

Scenarios:
Driver-based
Consequence-based
Location-based
Time-based

Management:
Meta-scenarios
Narrative structures
Narrative-base Models
Narrative components

Handling:
Updates
Synthesis
Dissemination

System Models. In parallel with the development of data models, the cyberCommons will integrate system models developed in the four institutions, as well as other models and tools, most of which are now driving ecological research along single pipelines that are not interoperable (see Section 4.4.3 below—Science Cases). Critical here is developing approaches to representing the mathematics of ecological models. System models will be organized into a model base to capture their components and structures, i.e., known real-time and historical ecological processes and interactions among drivers and consequences. Users will select and query models from this model base, and build new models from existing model components.

Analytics. The science of analytics develops computational techniques and tools for exploring potential correlates and embedded patterns in complex, multidimensional data that, ultimately, form scientific hypotheses. The cyberCommons will integrate the collection, management, processing, mining, visualization, and dissemination of data for ecological forecasting into a coherent, comprehensive, accessible analysis environment.

The massive amounts of data from sensor networks, remote sensing platforms, and biocollections databases to be archived and served by the cyberCommons, are prime candidates for data mining protocols to discover new knowledge. However, data transmitted by wireless sensors to servers are more susceptible to losses, delays, and corruption than the wired alternatives, and require techniques to estimate corrupted or missing data. As such, the cyberCommons will expand existing techniques, one for data-stream mining called CFI-Stream (Jiang and Gruenwald 2006), and two for estimating missing data, called WARM and FARM (Gruenwald *et al*. 2007), both developed at OU.

In addition to data mining and other knowledge discovery algorithms, our research will focus on mining of geographic dynamics (Yuan 2009), association rules, scalable decision tree classifiers, and probabilistic decision tree methods (Jiang and Gruenwald 2005; Zhang *et al*. 2007, McGovern *et al*. 2008). These protocols will advance recognition of spatio-temporal patterns among ecological drivers and consequences, e.g., the probability of vegetation phenology responding to shifts in climate patterns.

Visual analytics uses interactive visualization interfaces to facilitate analytical reasoning (Thomas and Cook 2005). Although visual analytics is a young field, useful and general techniques have already been developed into interactive tools, such as Jigsaw (Stasko 2007), for flexible exploration and analysis of both large-scale patterns and fine-grain relationships in multidimensional spatial+temporal+relational data sets. The IMPROVISE environment has been used to develop numerous visual analysis tools by promoting the synthesis of tool interfaces as an integral part of exploration and analysis (Weaver 2004, Weaver *et al.* 2007). Project participants McGovern and Weaver will focus on coupling data mining algorithms with highly interactive visualization techniques, with IMPROVISE as the basis for the interactive analysis aspects of the cyberCommons user interfaces.

Recent advances in web service technologies, such as SOAP and REST, have improved web GIS services. The Open GIS consortium has developed standards for web map, image, and feature services. The cyberCommons will deploy applications currently used at the four institutions (Map Server, ArcIMS, and ArcGIS Server) as well as other web GIS services in a multi-platform approach to maximize the diversity of applications available to meet specific needs.

Narratives are scenarios of change involving drivers and consequences across geospatial and temporal parameters. Each narrative is the model output and analytics outcome resulting from the specific data, models and parameters employed by the investigator to study the relationship among particular drivers and consequences. When narratives are organized in a database framework (a narrative base), they can be searched, compared and explored with narrative analysis tools for consistency and singularity. The cyberCommons will develop a narrative base for ecological research, incorporating representation, management, and process functions to permit examination of narrative ontologies and structures.

4.4.2 The CI framework—hardware, software, collaboration environment, integration environment

Each of the four informatics components in the science framework is mediated through the CI framework of hardware, software, a collaboration environment and an integration environment that compose the CI framework.

Hardware and core CI — The core CI consists of storage, computing and visualization components at each of the four institutions and the campus and regional networks that link them. The hardware will comprise two layers: (1) *dedicated hardware*, i.e., virtual machine (VM) servers, database servers and file servers allocated to researchers at each site, and (2) a *virtual data center*, an integrated consortium gateway to these resources, with virtual organization management, job submission, and monitoring functions.

Core CI deliverables are allocations of existing high performance computing (HPC) and storage resources available to this project from the OU Supercomputing Center for Education Research (www.oscer.ou.edu/education.php), the OU Center for Informatics, the OSU HPC Center, cluster systems at KU and KSU, and systems dedicated to the project participants (see Facilities). This diversity of resources includes compute systems ranging from small workgroup clusters to a large cluster (34.4 teraflops peak) and a large Condor pool (13 teraflops peak), as well as file storage and database servers aggregated from all sites, including an OU 132 terabyte storage system that is scheduled to expand to 400 terabyte in 2009. In addition to the integration of local resources, we will establish virtual machine servers at one or more sites to support the development and management of new services. Portals and services, i.e., software and data components with Web Services Definition Language (WSDL) or REST interfaces, will be provisioned as virtual machines and executed on the consortium's pool of VM servers.

Field sensors, as described below (Section 4.4.3, Science Case 4: Sensor networks for monitoring ecological response) will be deployed at the Konza Prairie Biological Station (an NSF LTER—Long Term Ecological Research—site), the KU Field Station and Ecological Reserves, and OU's Kessler Farm Field Laboratory (KFFL). Oklahoma's Mesonet was funded previously by NSF EPSCoR. Real-time data from the proposed Picarro G1301 WS-CRDS Multi-Species Analyzer and from meteorological stations, stream stations and flux towers included in this part of the CI will be continuously aggregated into a data management system for use by model components.

Networks are a critical piece of the CI, binding the sites and other external resources. Kansas and Oklahoma both have research and education networks, KanREN and OneNet, respectively, and all four institutions are linked by the regional Great Plains Network and the national Internet2 Network. Both OU and OSU are connected to the National LambdaRail (http://www.nlr.net/). Connection speeds for consortium members range from 1Gbps to 10Gbps. Network connectivity from field stations will be provided through KSU, KU and OU network support groups.

Software. Above the core CI hardware level, are the software, middleware and data service components needed to build the cyberCommons. Software CI development will focus on packaging and deploying models and analytical software as composable services using enterprise Web services (WSDL) and Web 2.0 RESTful interfaces. The initial set of services will form the core of a community-driven model system, and will include the following applications both deployed and under development at the member institutions: *Lifemapper* (Ecological Niche Modeling), the plant biodiversity model, the Ecological Platform for Assimilation of Data (EcoPAD), and the Terrestrial Ecosystem Monitoring and Diagnosis System (TEMDIS). Each of these systems is described in Section 4.4.3, Science Cases; the section on Integration Environment (next page) describes approaches to composing and reusing services.

Integration of these models and additional analytical services will provide the CI to support the following activities across the consortium: 1) data analysis, visual analytics and mining algorithms to identify embedded features, patterns, and relationships taxonomically, spatially, and temporally; 2) analysis of relationships among data inputs to determine interactive/reciprocal effects of global change drivers and responses; 3) improving the precision of ecological niche modeling and ecosystem modeling across the Central Plains; 4) data integration, management, and interoperability models to fuse heterogeneous, multiscalar, and multidimensional data; 5) analysis of effects of model input variables, individually and interactively, on model output and model behavior; and 6) refinement of models to address the two science questions discussed above.

Specific data sources to be aggregated and integrated include biodiversity data from vouchered biocollections (e.g. http://vertnet.org/, www.gbif.org/), land-use/land-cover data from Landsat images in USGS archives, MODIS data and other remote sensing sources (http://www.eomf.ou.edu); climate data from local and national sources; carbon flux data from the ORNL AmeriFlux network and other regional carbon flux networks; and biogeochemical data from analytical facilities within the consortium. We will also leverage newer paradigms such as social networking and crowdsourcing for citizen-based data acquisition and sharing. Examples include the georeferenced field photo library at OU for tracking land-use/land-cover and water use (http://www.eomf.ou.edu/photo-browser/), and the NSF GLOBE program for schools. One key project deliverable, enabled by the software layer, will be evaluations of new multidisciplinary models using data from these repositories.

We will also field middleware services for locating and delivering data to models and analytical software, and for linking computational components into workflows. These include ontology-based service repositories, and workflow systems for composing new applications. Portals and associated portlets will provide end-user interfaces tailored to the consortium's varied constituents, including researchers, students, educators, government officials and the general public.

Collaboration environment. A key aim of the cyberCommons is to enable exploration of new, highly integrated models of the interaction and reciprocal feedbacks of climate, land-use/land-cover, biogeochemistry and biodiversity. Its achievement, as well as education and outreach, mandates the formation of interdisciplinary communities of practice (Wenger 1998). Developing such a collaboration environment will encompass: social networking tools such as blogs, wikis, interest profiles, link libraries (e.g. Del.ici.ous) and shared bibliographies (e.g., Connotea); video conferencing and on-line meeting tools such as AccessGrid, H.323, and on-line web meeting services at all sites to support real-time collaboration. The myExperiment site (http://myexperiment.org) exemplifies the type of collaboration environments we will develop to integrate and extend CI for the consortium (DeRoure 2007).

Integration environment. Developing a research environment for multidisciplinary exploration, model building and testing of theory requires the integration of data, models, analytical tools and other components across research teams and physical sites in a flexible manner. In addition to the Web portals as a tool for presenting integrated views of data and models, we will explore and develop models, data and analytical codes as service components, and the composition of service components *as cyberinfrastructure* into new services. As appropriate, these components will be implemented as Grid services (using WSRF and Globus 4), standard 'enterprise' Web services, and as RESTful services that can be combined in mashups with other web-based services such as GoogleMap and Yahoo Pipes. The use of "Web 2.0" approaches to leverage the ongoing creative explosion on the web as a development and production environment for eScience has been extensively described (Pierce 2007; Fox *et al*. 2007). These data and compute services can be organized as workflows using conventional tools such as Kepler and Taverna or, as REST service URLs, can be combined in mashups to create highly interactive web-based tools. Shifting the CI framework from cycles and storage to reusable and extensible services is a novel and cardinal feature of the cyberCommons.

4.4.3 Science Cases

Researchers in the four universities have independently developed research expertise, web-based data portals, web applications, and models dealing with the ecological drivers and consequences encompassed by the two scientific questions: how do changes in climate and land-use/land-cover affect (1) biogeochemical cycles; and (2) biodiversity, invasive species and emerging diseases. Although successful, these efforts—described as Science Cases 1–4 below—are primarily local, incremental and single pipeline, lacking the capability of cross-domain fusion of research across these drivers and consequences. To that end, the cyberCommons will enable transformative research collaboration for modeling and synthesis of ecological complexity.

Science Case 1: Modeling biodiversity patterns and dynamics

The impact of climate change on biodiversity is mediated by ecological processes and their effect on biogeospatial patterns. Global environmental change to natural systems is affecting the composition, distribution and viability of Earth's plants and animal species (Parmesan, 2006, and see many recent exemplary studies on specific organisms: Parmesan *et al.* 1999, Harvell *et al.* 2002, O'Connor *et al.* 2007, Peterson and Shaw 2003, Pounds *et al.* 2006, Peterson *et al.* 2002b, Williams *et al.* 2003).

Two new capabilities, facilitated by computational approaches (simulation models, statistical methods, GIS), have transformed analyses of geospatial patterns of species distributions and their environmental correlates (e.g., Kremen *et al.* 2008): (1) ecological niche modeling uses known occurrence data of a species to estimate its environmental (e.g. climate) constraints and its potential distribution under scenarios of environmental change, particularly climate and land-use/land-cover; (2) synthetic analyses of biological inventory data to examine the geospatial patterns of biodiversity on regional and continental scales. Both approaches reveal how biotas respond to changing climate and land-use/land-cover and are being used in research initiatives by biodiversity, ecological and informatics scientists at KSU (Blair, Dodds, Gido), KU (Beach, Peterson, Soberón, Stewart, Vieglais, Wiley), OSU

(Palmer), and OU (Gruenwald, Weaver, Yuan). Additionally, the emerging field of macroecology is applying computational and quantitative approaches to analyses of large-scale biodiversity spatial patterns (Blackburn *et al*. 2004), for example, a web-services architecture for generating multi-species distribution data grids with user-specified cell sizes and grid extents (Beach and Colwell, in review).

Specifically, KSU researchers focus on: (1) patterns of aquatic biodiversity, primarily fishes, using data on water depletion associated with center-pivot irrigation and construction of impoundments in the Central Plains; and (2) trait-based biodiversity modeling that matches species traits to geographic distributions. KU's biodiversity informatics initiative has pioneered the application of ecological niche modeling to predicting the potential spread and distribution of plant and animal species, including invasives and zoonotic disease vectors/hosts/reservoirs, under scenarios of environmental change (Soberón 2007, Soberón and Peterson 2005, Peterson 2003, 2006, 2007, 2008, Peterson *et al*. 1999, 2005). KU's Biodiversity Research Center has developed (1) *Lifemapper*, a cluster-based, web-services architecture for a global species niche modeling facility (www.lifemapper.org/index.shtml); and (2) the Specify Software Project (www.specifysoftware.org, www.specify6.specifysoftware.org/), a specimen database management system in use by more than 210 biocollections institutions worldwide. OSU has developed quantitative tools for analysis of broad-scale spatial and temporal trends in vascular plant diversity in North America (Palmer 2007, 2006, 2005, Palmer *et al*. 2002). At OU, biodiversity and ecological informatics research focuses on (1) correlating long-term datasets on migrant land-bird populations with disease transmission, and the effects of climate and land-use on avian populations; and (2) use of GIS analysis, spatiotemporal databases, and visual analytics on ecosystem processes (Yuan *et al*. in press).

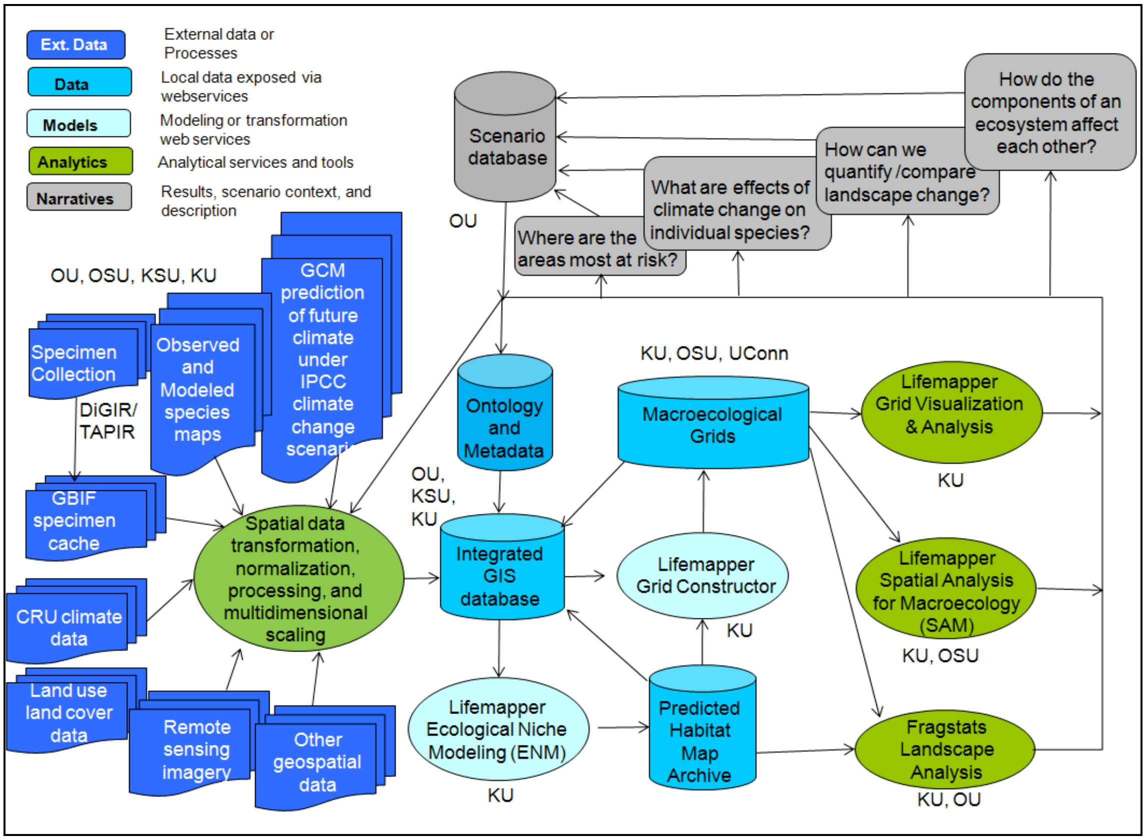


Figure 3. Tasks and workflows associated with environmental niche modeling and Lifemapper, including multidisciplinary data and analysis components

The cyberCommons will be the virtual environment for computational integration of various data sources, modeling agents, and analytical tools to facilitate multidisciplinary, cross-cutting work flows that utilize computational techniques already in place at the four partner universities. It will extend the research on patterns of biodiversity to comparisons of niche vs. trait-based models of species distribution and roles played by ecological vs. evolutionary processes. For example, KSU's tools for mapping geospatial hydrologic conditions and aquatic species distributions should be linked to vegetation models and hydrologic models of phreatophytes to predict the effects of woody expansion on native fish species. OSU's floristic data would benefit from outputs of climatic and land-cover models. OU's web GIS tools for mapping and visual analytics should be integrated with KU's ecological niche modeling tools to address the environmental complexities of species occurrence patterns and predict shifting continental diversity patterns to land-use planning.

For example, Figure 3 (preceding page) illustrates how the cyberCommons can link data, models and analytics associated with single-species environmental niche models (*Lifemapper*) to generate multispecies predictions, macroecological grids and ultimately narratives that address the two critical ecological questions. The environmental data layers used in *Lifemapper* are also deployed by research in Science Case 2 and 3, below. Components and workflows on the left half of Figure 3 exist to varying degrees of automation and integration. Linkages on the right side do not yet exist, and are a development goal of the cyberCommons. For example, the cyberCommons will provide open web portal and service access to point-occurrence biocollections and observational data, land-use/land-cover environmental layers, and expert-derived species distribution maps—all in standard formats (Darwin Core, netCDF, OCG)—to enable data transformation and integration. In the planned model output management system, IPCC (Intergovernmental Panel on Climate Change) climate scenarios will be discovered, transformed and fed into single species ecological niche models; in turn, these models will populate multi-species grids for macroecological models and consequent narrative outputs.

Science Case 2: Forecasting the geography of disease transmission

Emerging infectious diseases have a global impact on biodiversity, human health and national security (Jones *et al*. 2008). Serious health and security threats include: the sudden appearance and spread of West Nile Virus across the Western Hemisphere (Peterson *et al*. 2009); SARS across East Asia; H5N1 avian influenza across the Old World (Gilbert *et al*. 2007, Williams *et al*. 2008); Hantavirus (Sin Nombre Virus) and plague across parts of North America; and the recent anthrax attacks in the US (Broussard 2001). These disease threats reveal the need for a deep and predictive understanding of disease ecology and biogeography, themselves emerging fields of study (Collinge and Ray 2006).

Research on disease and invasive species systems unites expertise, predictive modeling approaches and data at KU (Peterson, Soberón, Martinko, Egbert), KSU (Wisely, Scoglio) and OU (Xiao). Disease geography, like ecoforecasting, is essentially a predictive endeavor. It requires a multiscalar and flexible conceptual framework for anticipating the dynamics of diverse disease systems (Soberón and Peterson 2005, Peterson 2008, Xiao *et al*. 2007, Gilbert *et al*. 2008). In this framework, sustained disease transmission is a function of the interaction among diverse species, i.e., pathogen, host, vector, and where these entities can maintain populations, in turn a function of abiotic, biotic, and mobility parameters (Peterson 2007).

Different factors are dominant at different scales. At moderate scales—i.e., landscape to continental—ecological niche modeling, with biocomputational tools developed at KU, has predicted the distribution of disease component species (Adjemian *et al*. 2006) and the occurrence of disease outbreaks (Gilbert *et al*. 2007). Specifically, ecological niche modeling has been powerful in: (a) describing ecological dimensions of transmission of poorly known diseases (Costa *et al*. 2002); (b) interpolating poorly known spatial distributions to predict additional areas at risk (Peterson *et al*. 2006, Williams *et al*. 2008); (c) anticipating the potential geographic spread of invading species (Peterson 2003); (d) identifying unknown species participating in disease transmission systems (Peterson *et al*. 2002a, Peterson *et al*. 2007); and (e) anticipating spatiotemporal dynamics of vector populations (Peterson *et al*. 2005). Tests against independent data sets confirmed these geospatial predictions, the accuracy of which increases with the integration of remotely sensed data (Gilbert *et al*. 2008).

At finer geospatial scales, the analytical focus is on individuals and their interactions. Network approaches (agent based models) can characterize complex systems involving millions of components interacting in non-linear, heterogeneous fashion. The power of this approach to modeling disease dynamics is its incorporation of individual behavior and spatial heterogeneity. Still-finer scales, such as those involved in an ongoing study of skunk rabies strains in North America, entail modeling of how seasonality, land-use/land-cover and disease etiology affect outbreaks, the adaptive landscape of pathogen transmission, and the epidemiology of different rabies strains.

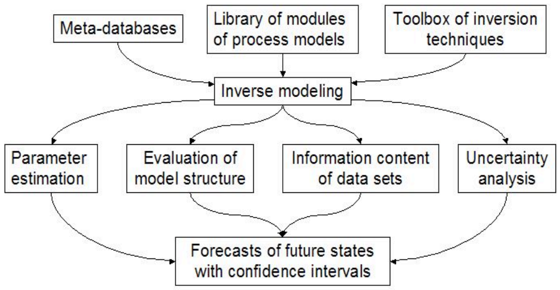
The cyberCommons will enhance the accuracy and degree of complexity in these predictive modeling approaches to disease (and invasive species) systems. It will automate data assimilation from telemetry systems (e.g., for rabies hosts); biodiversity, abiotic, geospatial and remote sensing realms; as well the processing of satellite imagery, and integrate these data pipelines with tools for ecological niche modeling. In so doing, the cyberCommons will unify ongoing research on regional and global disease systems (e.g., rabies, West Nile Virus, Hantavirus, Lyme, H5N1 Avian Influenza, Ebola, Marburg, references cited in this section) across laboratories at KU, KSU and OU and develop their analytical complexity. Ultimately, the modeling results will inform strategies that minimize the risk of human exposure.

Science Case 3: Large-scale biogeochemical cycling

Biogeochemical processes and ecosystem services (e.g., gross and net primary production, net ecosystem production) in the Central Plains are affected by changes in climate, land-use and biodiversity, including invasive species, as well as by natural disturbances. Three research approaches are widely deployed for data collection at KSU, KU and OU: CO_2 eddy flux towers, satellite observations, and in-situ field experiments (e.g., climate warming, altered precipitation and biofuel feed stock harvest at KSU's Konza Prairie and OU's Kessler Farm Field Laboratory. Researchers at OU are developing two major complementary biogeochemical modeling frameworks described below: *EcoPAD* and *TEMDIS*.

EcoPAD: Ecological Platform for Assimilation of Data — The accuracy of ecoforecasting increases with the continuous assimilation of data for testing of ecological theory and models (Clark and Gelfand 2006). The cyberCommons will enable development of a data-model fusion tool, the Ecological Platform for Assimilation of Data (EcoPAD), for application to data and models encompassed by the science questions. EcoPAD is supported by meta databases of biogeochemical variables, libraries of modules of process models, and a toolbox of inversion techniques for data assimilation and forward predictions. The functions of EcoPAD's inverse mode will be: (1) parameter estimation, (2) evaluation of different model structures, (3) uncertainty analysis, and (4) quantification of information contained in data sets in diagnostic modeling. EcoPAD's forward mode will incorporate estimated parameter values, their variances, and other sources of information into prognostic models to project future states of ecological systems and associated uncertainties.

EcoPAD's development will require (1) high performance computational power for data assimilation and optimization; (2) more efficient data assimilation algorithms; (3) improved capacity to estimate parameter uncertainty and predicted ecosystem states; (4) improved software workflow; (5) improved integration of EcoPAD's components, (6) conversion of EcoPAD into an operational software system for integration with NEON's planned CI (www.NEONInc.org) and (7) user-friendly interfaces.



EcoPAD schematic

EcoPAD's data assimilation techniques will include: (1) genetic algorithms; (2) simulated annealing;

(3) the Levenberg-Marquardt minimization method in combination with a quasi-Monte Carlo algorithm; (4) Markov chain Monte Carlo sampling with Metropolis-Hasting (Metropolis et al. 1953, Hasting 1970); (5) ensemble Kalman Filter (Kalman 1960); and (6) an adjoining function of ecosystem C processes for inverse analysis. This work builds on success in data assimilation using the CASA model (Cooperative Adaptive Sensing of the Atmosphere) for net primary productivity in combination with a layered soil carbon dynamic model (Zhou and Luo 2008). Previous work, using the Terrestrial Ecosystem Model includes development and application of six assimilation techniques to multiple heterogeneous datasets from Duke FACE, eddy-flux sites, and spatially distributed measurements of plant and soil carbon pools and fluxes (Luo *et al*. 2003, Xu *et al*. 2006; White *et al*. 2002, 2005, 2006).

TEMDIS, the Terrestrial Ecosystem Monitoring and Diagnosis System, is a framework in development at OU for attribution analysis of CO_2 fluxes of terrestrial ecosystems in a changing world. It is composed of several satellite-based diagnostic models, and process-based prognostic models for the assimilation of remotely sensed data, including the Vegetation Photosynthesis Model (VPM) and the PROSAIL-2 radiative transfer model. VPM is built upon the concepts of leaf chlorophyll content, light absorption by leaf chlorophyll, and light use efficiency, and is driven by satellite images and daily climate data. The VPM model has been evaluated at more than 30 CO_2 eddy flux tower sites, ranging from boreal forests, deciduous forests, grassland, alpine meadow, and croplands (winter wheat, maize), to tropical forests (Xiao *et al*., 2005, Mahadevan *et al*., 2008).

PROSAIL-2 radiative transfer model (RTM), which builds on advances in the canopy-level SAIL (Braswell *et al*. 1996) and leaf-level PROSPECT models, is used to retrieve simultaneously canopy properties and leaf properties from daily MODIS images (Zhang *et al*. 2005; Zhang *et al*. 2006). Inversion of PROSAIL-2 for one MODIS pixel (500-m spatial resolution) takes approximately two days of computation on a server (Zhang *et al*., 2005) as it provides the probability distribution of individual variables as estimates of uncertainty through the Markov Chain Monte Carlo simulations (a modified Metropolis algorithm, Metropolis *et al*. 1953). Landscape- and regional-level simulations (thousands to millions of MODIS pixels) by PROSAIL-2 require access to supercomputers and grid and cloud computing resources. The cyberCommons will provide the ready access to high performance computing and the geospatial data repository required to run PROSAIL-2, in turn advancing ecological, weather and climate forecasts (e.g., the NSF-funded Linked Environments for Atmospheric Discovery (LEAD) program). The TEMDIS models will be advanced with better geospatial data (e.g., satellite imagery, weather, climate), data visualization, work flow, resolution and statistical inference.

Remote sensing data on changes in land-use/land-cover is critical to advancing the modeling of biogeochemical cycles. The Kansas Applied Remote Sensing Program (KARS) at KU uses satellite imagery and aerial photography datasets to map land-cover, model land use dynamics, and analyze landscape characteristics. Pre-processed imagery for the Central Plains includes: AVHRR time-series NDVI imagery since 1989; MODIS time-series NDVI imagery from both the Terra (since 2000) and Aqua (since 2002) sensors; and, for Kansas, Landsat TM and ETM+ multi-seasonal (spring, summer, fall) imagery for 1991-1994 and 2000-2005, and statewide digital land cover maps derived from Landsat imagery (1990, 1994, and 2005). This coverage complements the nationwide National Land Cover Database developed by the USGS.

Access to these datasets through the Kansas Data Access and Support Center (DASC, the Kansas GIS data clearinghouse) requires time-consuming processing steps. The cyberCommons will create an interactive portal for researchers to identify and access specific data stacks, e.g., by watershed, ecoregion, or political boundary. It will also provide training for researchers who need to create specialized datasets such as local area maps of woody plant invasion and spatial patterns and landscape impacts of CRP lands, crop expansion, e.g. corn for ethanol, and suburban growth.

Science Case 4: **Sensor networks for monitoring ecological response**

Forecasting biogeochemical and biodiversity responses to environmental change requires intensive data inputs from experimental and observational studies to inform modeling and scaling activities.

Unfortunately, data from different regions are often collected with different methods, hampering data compatibility and integration for analysis and forecasting. Oklahoma benefits from having statewide standard installations of Mesonet sensor networks. Kansas requires comparable sensor networks and wireless communication capabilities in order to increase the scope and scale of ecosystem measures for advancing the accuracy and resolution of ecological forecasting across the Central Plains. As part of the cyberCommons, these CI enhancements will occur at focal field study sites—the KSU Konza Prairie Biological Station (KPBS), the KU Field Station and Ecological Reserves (KSR), and the Kessler Farm Field Laboratory (OU). The resulting near real-time assimilation of data from field experiments and sensor networks will yield rapid data analysis and integration of modeling results with EcoPAD and TEMDIS, which will scale up through multi-resolution satellite sensors (Landsat, Hyperion, MODIS).

The 3,487-ha KPBS field site (a NEON candidate core site) includes 60 experimental watersheds with varied land-use (prescribed fire and grazing, agriculture) and land-cover (native prairie, restored prairie, gallery forest, shrubland), an intensively monitored stream system, and long-term plot-scale experiments. Three eddy covariance (flux) towers continuously monitor carbon, water and energy flux over native grasslands and sites undergoing woody plant encroachment, a critical land-cover change in the Central Plains. A critical CI enhancement for research will be the installation of a tower network for wireless communication, data transmission, and automated animal tracking across KPBS. New wireless sensor networks will measure environmental heterogeneity and abiotic drivers (i.e., wind speed, solar radiation, surface temperature, soil moisture at three depths, relative humidity) in the footprint of the flux towers and in other targeted areas. Soil moisture is a critical, but poorly characterized, driver of ecological processes in grasslands (Knapp *et al*. 2002, 2008). A new network of automated instruments will provide real-time access to unsaturated soil infiltration rates at multiple scales. The deployment of spatially optimized, vadose zone monitoring networks across watersheds will advance understanding of plant and microbial responses to environmental heterogeneity. Each node will monitor soil water content, temperature, and soil water potential near upper and lower soil profile boundaries. These data are critical for modeling water and energy fluxes in unsaturated soil. The cyberCommons will integrate each of these data streams and data from a long-term climate change experiment at KPBS (Fay *et al*. 2000, Knapp *et al*. 2002) to feed EcoPAD and TEMDIS, being developed at OU (Science Case 3, above).

The KSR site includes long-term experiments on land-cover change and an eddy flux tower that extends our sensing network to a site with higher rainfall and aggrading forests. Two new networks will be deployed in the KSR flux tower footprint: (1) four soil respiration chambers (Li-8100) and sensors to monitor abiotic variables (soil temperature, moisture, etc); and (2) a wireless remote network of microclimate stations, similar to those at KPBS, to measure environmental heterogeneity. Common sensors and data collection methods will ensure compatible data across vegetation types, climates, and soils. Addition of a Picarro G1301 WS-CRDS Multi-Species Analyzer to the KSR flux tower will simultaneously monitor fluxes of carbon, water and methane. Addition of a Large Aperture Scintillometer (LAS), to complement one at KPBS, will allow calculation of evapotranspiration across large areas by residual measure of the surface energy balance, ultimately for upscaling measured fluxes from the towers and validation of surface energy fluxes from satellite data.

Central Plains ecosystems will likely be strongly affected by the expanding biofuel industry. The Kessler Farm Field Laboratory is conducting two climate warming and biofuel feedstock experiments as test beds for CI development for ecological forecasting (Luo *et al*. 2001, 2009). The experiments manipulate three environmental factors: climate warming, altered precipitation, and biofuel feedstock harvesting. In addition, the research focus of the new OK NSF EPSCoR RII Track 1 award is the development of ethanol from switchgrass and related crops. The cyberCommons will augment the existing sensors in the experiments with wireless networks, which will automate measurements of temperature, soil moisture, wind speed, precipitation, and other environmental and biological variables and feed the data into models and data assimilation software.

The cyberCommons will assimilate data from sensor networks via a unified data model to feed ecosystem process models (e.g., EcoPAD, TEMDIS) for scaling activities, for comparing measured carbon, energy, and water fluxes with model outputs, and for narrative runs under multiple climate scenarios predicted by downscaled IPCC climate models. The multiscale, biometeorological modeling framework uses different models at their native spatial and temporal resolutions to develop scaling relationships, and to assess responses and feedbacks to climate change. Using IPCC AR4 results (4th Assessment Report) as forcing fields, the Weather Research and Forecasting (WRF) model will be coupled with ecological models (EcoPAD, TEMDIS) to assess regional-scale processes at annual-synoptic timescales. Ecosystem process models will assess the ecological implications of these regional climate forcings. Large eddy simulations will assess the effects of changes in land-use and surface fields on atmospheric transfer of mass and energy fluxes. This framework will allow the scaling of high frequency sensor measurements (20 hz) to validate the Large Aperture Scintillometer and eddy covariance data, and to scale results up to the region to assess impacts across the Central Plains. Remote sensing Soil-Vegetation-Atmosphere-Transfer (SVAT) model outputs will validate model runs. Field-collected data will validate each step prior to conducting the narrative AR4 output scenarios.

A second component of the sensor network at KPBS focuses on automated tracking of animal movements to support studies of animal-plant interactions (Science Case 1, above) and epidemiological modeling (Science Case 2, above), both of which are also NEON goals. Patterns of host/vector/reservoir animal movements and contact rates are key to understanding the emergence and spread of infectious diseases and maintenance of species diversity in ecological systems. For example, in the Central Plains, skunks are an important vector of rabies. An Automated Radio Telemetry System (ARTS) can determine patterns of vector movement in relation to land cover, and assess how overwintering in communal dens affects host-to-host contact and the probability of rabies transmission. Such animal tracking is critical for developing models of the spread of Hantavirus (*Peromyscus* vector), Avian Influenza and other zoonotic diseases. Data on animal movement and activity (e.g., interactions, mortality, dispersal) will be integrated with climate and landscape features into a model of disease transmission in rural and urban environments. Model outputs will increase understanding of the role of grassland ecosystem services in regulating the spread of zoonotic disease.

4.4.4 The cyberCommons Deliverables and Timeline

The cyberCommons will be developed in four stages. ***Stage 1*** **(0-6 months)** will involve recruitment of personnel, acquisition of hardware and software, and establishment of virtual data centers at each site. ***Stage 2*** **(6-18 months)** will focus on development of the cyberCommons web portal configured on a virtual machine as gateways to all four data centers. Initially, the web portal will be populated with existing data sets and tools from the four institutions, and with customized web GIS applications for mapping and browsing. Middleware with management capabilities and applications will be developed for users to access the OU Supercomputer via the web portal. Concurrently, major research effort will be devoted to analyses of: (1) data and function requirements; (2) workflow; and (3) user interface needs. The results will specify development of an ontological system, a unified data model, and narrative structures for ecological and biodiversity data integration and information management, which, in turn, will specify the development of tools for visual analytics and data mining. At the end of Stage 2, project researchers will test and evaluate the web portal and provide feedback for revision and refinement.

Stage 3 **(18-27 months)** will prototype the cyberCommons by along two paths. First, the project's biologists and information scientists will continue iteratively testing and refining the web portal and tools developed during Stage 2, i.e., the ontology, unified data model, workflows, visual analytics, data mining tools, and user interface. Concurrently, we will develop network service capabilities, including APIs and SOAP applications to provide geospatial analytical functions and tools for visual statistics, mapping, graphing, and web mining. Forums, blogs, and other virtual social networking tools will be developed for communication among and feedback from users.

Stage 4 **(27-36 months)** will complete implementation of the cyberCommons by evolving the prototype to an operational CI system, including all specified functions, user registration and accounting, security, coupling with other CI systems, and capabilities for expansion.

Table 2. cyberCommons deliverables

Category	Component	Lead and (team)	Month
People	Recruitment of graduate students and software engineers	All	12
	Develop communication plans and reporting strategies	KU/OU (All)	3
Hardware	Establish virtual data center nodes at each site	KU (All)	6
	Allocate compute and storage at each site	KU (All)	6
	Connect local HPC and storage and virtual data center nodes	OU (All)	9
	Sensor networks at KPBS and KU Ecological Reserves	KSU (KU)	
	1. specify and purchase sensors		12
	2. specify, purchase and install radio towers		12
	3. install sensor stations and field network components		24
	4. implement animal collars/ triangulation		36
	5. hydrologic sensor test network		36
	6. field network/data rate upgrade		12
	7. data models/sharing processes for assimilation into eco-models		24
	8. Data products for biological and geochemical models		36
Software	HPC gateways set up (management process, queues, storage allocations)	OU	6
	Virtual Machine provisioning and management process	KU	12
	Software development of models (EcoPAD, TEMDIS)	OU (All)	36
Collaboration environment	Top-level portal and integration of existing resources	KU (All)	12
	VO collaboration support services (blogs, wiki, discussion groups, video and audio conferencing, etc.)	KU/OU	6
	Resource catalogs of APIs, service addresses, and data dictionaries for databases, models, analytical codes and other services	KU/OU	18
	Portal, portlet and mashup development standards	OU/KU	18
	VO authentication and authorization services (Shibboleth, Missouri Entitlement Server software)	KU	12
Integration environment	EcoPAD: Ecological Platform for Assimilation of Data	OU (All)	
	1. core software		12
	2. Web interface, and		24
	3. integration with sensor networks at KPBS, KSR, KFFL, AmeriFlux		24
	4. forecasts based on land-use/land-cover and climate change		36
	5. policy making and outreach applications		36
	TEMDIS including PROSAIL-2 RMT and VPM	OU	36
	Ontology server and management processes	OU	27
	Data-assimilation and uncertainty analysis algorithms using comprehensive ecosystem and ecological forecasting models (e.g. NCAR Land Surface Model)	OU	
	Climate change/Biofuel experiments at KFFL	OU (All)	12
	cyberCommons integration environment	All	
	1. cyberCommons web portal for existing data and tools	KU/OU	18
	2. Geospatial data management system and query tools	OU/KU/OSU	18
	3. Ecological ontology and data modeling	OU/KU (All)	18
	4. Develop data mining tools and algorithms	OU/OSU	24
	5. Develop visual analytics tools and algorithms	OU/OSU	24
	6. Develop narrative elements and structures	OU	24
	7. Network services (SOAP and REST APIs)	KU	24
	8. Prototype data models, tools, and narratives	OU	27
	9. Prototype network services applications	KU	27
	10. Testing and evaluation	KU/OU (All)	30
	11. Make cyberCommons operational	KU/OU (All)	36

Education, outreach and training	CI-mediated for-credit seminars for graduate and senior undergraduate students in aspects of computational ecology and biodiversity science, including biodiversity- eco- and geoinformatics, and ecological niche and biogeochemical process modeling.	All	36
	Graduate collaborative research program in ecological modeling using cyberCommons data and tools	All	36
	Portals to introduce ecological concepts, research and tools to P-16 students and teachers, and the general public, including interactive modules for K-12 students and informal science education settings.	All	36
	Implement informatics tools, such as Lifemapper II, for use in KU and OU (and other) natural history museums for development of next generation informal science education programming in ecological forecasting.	KU/OU (All)	
	Graduate student training in developing integrated research environments for ecological modeling.	All	36

4.5 Education

Cyberinfrastructure creates environments and opportunities to expand education beyond the teacher-student relationship and the classroom-laboratory. CI-mediated education transcends disciplinary, geographic, and institutional boundaries and affords broader participation by diverse populations. A cyberCommons can enlarge the range of materials and modes of access and delivery for learners to explore, for example, ecological field observations, data analysis, or team computational modeling of the reciprocal effects of climate and land-use/land-cover changes on ecosystem structure and function on the Central Plains. Web modeling tools for students, teachers and the general public project can enrich both formal and informal science education by enabling inquiry-based analysis of ecological phenomena—for example, "backyard" observations, or testing model outputs with local data.

The four universities will develop common, CI-mediated for-credit seminars for graduate and senior undergraduate students in aspects of computational ecology and biodiversity science, including biodiversity- eco- and geoinformatics, and ecological niche and biogeochemical process modeling. For example, KU is developing a curriculum enhancement to prepare the next generation of researchers to use the critical research tools of computational science and high performance computing. The new curriculum will provide targeted, real-world learning opportunities and workforce development for students at the cyberCommons by drawing on specific modeling and data analysis activities.

Beyond fostering individual graduate research, the CyberCommons will be a platform for enabling true interdisciplinary graduate research, wherein teams of students from multiple disciplines, departments and institutions bring the natural and social sciences to bear on a single project—for example, the social and ecological consequences of expanding biofuels production across the Central Plains. Further, student research projects in cyber disciplines will be able to bring technological and methodological innovations to the proposed CI. These innovative resources, through open portals, will also be available to new audiences in the states' numerous community and four-year colleges, and state agencies. As products are completed, the cyberCommons will have consortium-wide training on their use, e.g., how to select data layers, use portals, compose new models and run new narratives. Broader audiences of faculty and students will be invited to attend the training sessions. This extends OU's ongoing NSF-funded CI education initiative, *Supercomputing in Plain English* (http://www.oscer.ou.edu/education.php; Neeman *et al*. 2002, 2006, 2008a, b) for students, faculty and staff at dozens of institutions. In 2010 and 2011, building on OU's ongoing series of Supercomputing Symposia, the cyberCommons will host a weeklong summer ecoinformatics workshop in Kansas and Oklahoma for undergraduate, graduate students, research scientists and faculty.

In addition to higher education and workforce development, the CI environment is a logical conduit to introduce ecological concepts, research and tools to a broader community, including interactive modules for K-12 students and portals for the general public. For example, Kansas will be able to incorporate and extend Oklahoma's current suite of outreach activities that include Girl Scouts geospatial badges, GIS and GPS activities for high schools, summer workshops for public school teachers, and GIS workshops for state and local government officials. In turn, Oklahoma will be able to incorporate and extend KU's public educational outreach on NSF's Assembling the Tree of Life, and KSU's broad LTER-based educational programs affiliated with the Konza Prairie Biological Station, including online ecological databases generated by participants, student-run research (http://www.k-state.edu/konza/keep/education/index.htm), the Konza Environmental Education Program (KEEP) for K-12 students and the public (www.ksu.edu/konza/keep), the Konza Prairie Schoolyard LTER program, and the statewide Konza SLTER program of data collection at satellite prairie sites (Prairies Across Kansas).

Finally, two of the participating universities (KU, OU) have distinguished natural history museums with extensive outreach programs in informal science education to K-12 and the public. The CyberCommons will bring the powerful informatics tools, such as Lifemapper, to the museums for development of next generation informal science education programming in ecological forecasting.

4.6 Diversity Plan

During the past 15 years, Kansas, Oklahoma and regional state partners have developed high-speed research and education networks (i.e., the collaborative Great Plains Network, Oklahoma's One Net, KanED, and KANREN) to link higher education institutions, K-12 schools, libraries, and government agencies to the Internet and its increasing collection of resources. As a result, the geographically distributed population of teachers and learners representing the gender, ethnic and institutional diversity of the two states will be able to access and benefit from the products and resources of the CyberCommons. Use of networks to reach diverse populations is particularly timely, as the minority populations of both states are increasing, and now exceeds 50% of the population in four of the 18 largest cities in Kansas.

For example, Oklahoma's OneNet will enable science portals to be adopted and used by the 23 smaller universities and colleges distributed throughout the state mostly in rural areas and smaller towns. Three 2-year colleges and three smaller rural 4-year colleges, all of whom are committed to incorporating ecological forecasting models into their science curricula, will lead the effort. In Kansas, science portals will enrich the curriculum at the state's six universities, 19 community colleges, and 18 independent colleges and universities. Both states will introduce CyberCommons resources to teachers and Native American students in Oklahoma and Kansas to enable them to participate in data gathering and analysis as an enhancement of ongoing programs in the natural and social sciences. Among the resources are portals that the project will tailor to environmental issues across Native American lands.

In similar fashion, Kansas and Oklahoma have developed a rich array of successful diversity programs, including those fueled by EPSCoR RII awards. For example, each year Emporia State University in Kansas hosts Expanding Your Horizons, a program where more than 200 middle-school girls from across the state to explore careers in STEM disciplines with women scientists as mentors, as does KSU's GROW (Girl's Researching Our World: www.k-state.edu/grow/) program. Oklahoma EPSCoR supports Supplemental Instruction (SI) and GRE preparation at the historically black Langston University (LU), resulting in significantly increased GPA scores for participants. Using existing state networks and the CI resources and architectures of this project, the KS-OK consortium will build on, enrich and extend these programs for women and minorities through science portals where students and teachers can explore and participate in 21st Century ecological science using the data, analysis and modeling products of this project.

The consortium will invite participation at its annual symposia, targeting faculty from geographically remote areas of the states, HBCUs and tribal colleges, and institutions serving large numbers of women and minorities. Stipends for travel, in the state's matching funds, will assist and encourage attendance.

4.7 Dissemination and Communication Plan

The project will implement a unified dissemination and communication plan across the two states to communicate the value and benefit of the consortium's work and to improve science literacy among its citizens. Fundamental to the plan will be identification of constituent groups, development of targeted messages, and selection of communications channels and formats for dissemination. Expected constituent groups include: journalists, state-level policymakers (governors, legislators, agency heads and key staff), the federal delegation, science and cyber faculty at area colleges and universities, K-12 science and cyber educators, symposium audiences, environmental organizations, and other groups with potential interest in the goals of this proposal. The audiences will receive targeted announcements of and invitations to project events, and descriptions of accomplishments and results using a range of communication channels, formats and venues, including websites, web portals, press releases, e-newsletters, brochures, presentations and media interviews, posters, e-mail, and annual symposia. News releases will be the most immediate and frequent public communications, initially to introduce the project and describe its importance to the region and thereafter to announce events and highlight accomplishments and impacts. Project staff will leverage distribution of e-documents by engaging other organizations to share their membership lists when appropriate.

The cyberCommons web site will describe the purpose, goals and expected impact of the project, and provide links to web resources and research and educational portals that result from project activities. Staff will advertise the website to constituents and recruit the nation's other state EPSCoR programs to avail their constituents of the resources provided by the cyberCommons. In order to put a "local face" on the cyberCommons, staff will arrange for project scientists to make informal presentations to local educational and civic groups in each state (Rotary, Lions, Chamber of Commerce, associations of science teachers), journalists, and various media, which also will be posted on the project website.

Semi-annual symposia for all project participants will be the cornerstone event for both internal and external communications. Each year, the symposia will alternate between KS and OK and include demonstrations of the developing infrastructure. In Year 1, the symposia will target local and state agency policymakers from departments such as agriculture, health and environment, wildlife and parks, and water resources. The objective will be to raise their awareness of the project and discuss with them the potential utility of project results to their respective missions. In Year 2, the symposia will target higher education institutions, including community colleges, historically black colleges and universities, tribal colleges, and private colleges and universities in the two states. Objectives will be to increase institutional diversity, introduce this constituency to the research and education applications of the developing cyberinfrastructure, and enlist their participation as users of emerging products at project portals and websites. To facilitate attendance among this constituency, the project will offer travel stipends on the basis of distance and need. The Year 3 symposia will target elected officials and their staff members to showcase the collaboration between states and the value of the cyberinfrastructure and modeling products as tools to enhance research and education and inform public policy.

Symposia will feature presentations from leaders in CI innovation, and be scheduled to maximize attendance by members of the AAAS Assessment Team, the External Advisory Team, the State EPSCoR Committees and targeted groups. Programs will feature presentations, demonstrations, and posters that illustrate the power of the cyberCommons for advancing knowledge discovery of environmental phenomena for science and society. Press releases will announce each symposium, with space reserved for journalists who wish to attend.

4.8 Evaluation and Assessment

The consortium will use both formative and summative evaluation to assess, monitor, and improve project activities. Two groups of external experts will contribute to this process: (1) the American Association for the Advancement of Science (AAAS); and (2) an External Science Advisory Committee (ESAC). The consortium will define an internal process to gather and prepare data for the AAAS and ESAC teams to conduct their assessment and for the project's Leadership Team to monitor on-going activities.

AAAS Panel. The AAAS will assemble a five-member panel, including one AAAS senior staff member and four experts in the CI areas proposed for improvement in this proposal. Early in the award period, the AAAS panel will meet project participants, review implementation plans, and provide supplemental or alternative strategies for successful implementation where warranted. The panel will review the benchmarks and metrics of the program and advise on the best methods for tracking progress. Thus, the panel will be providing formative project evaluation and program guidance on implementation.

Near the beginning of the second year, the AAAS panel will assess progress and suggest areas for improvement and alternate strategies, if needed, for completing the project goals. In the final six months of the award period, the panel will begin a summative evaluation and address the sustainability of the project. Each of the panel meetings will occur on-site and alternate between Kansas and Oklahoma. At the conclusion of each visit, the panel will prepare a report of its findings for dissemination to all project stakeholders.

External Advisory Committee. An External Advisory Committee (EAC) will provide scientific, CI and management review and counsel to the project teams throughout the award period. The EAC will attend the semi-annual project symposia, meet with project leaders and participants, and prepare reports describing the strengths and weakness of the work underway and suggesting course corrections for improvement. Members of the EAC will be available to advise project teams on an informal basis as needs arise. To date, four distinguished scientists have agreed to serve on the EAC: Peter Arzberger, Chair, PRAGMA—Pacific Rim Applications and Grid Middleware Assembly (and, at UC San Diego, Director, National Biomedical Computation Resource); Geoffrey Fox, Professor, Computer Science, University of Indiana (web-based integration environments); Deanna Pennington, Informatics Research and CI Outreach, LTER Network Office, University of New Mexico; and David Schimel, CEO of NEONInc (and Senior Scientist, NCAR).

Metrics. Table 3 lists examples of the quantitative and qualitative metrics that will be used to evaluate and assess the project's progress and outcomes.

Table 3. Sample Evaluation and Assessment Metrics

Indicator	Quantitative Metrics	Qualitative Metrics
Research and implementation	• Publications in ecological, biodiversity, GIS and computer science • Presentations at professional venues • New datasets, tools, models, etc. served though web portals • Specification, acquisition, installation of cyberCommons hardware (see Table 2 Deliverables) • Implementation of software, collaboration environment and integration environment layers (see Table 2 Deliverables)	• Interdisciplinarity and collaboration in the dissemination of project results • Data, model fusion and disciplinary fusion in generating research results • New research areas and synergies • Sustainable research clusters • Major research success stories • Invited participation at professional conferences, panels, and boards • Professional awards and honors
Education, training and mentoring	• Interdisciplinary doctoral and masters projects • New courses, seminars, etc. fusing science and informatics • Junior faculty supported and mentored through cyberCommons activities • Training in HPC and related CI for faculty and students • Ecoinformatics seminars, workshops, symposia	• Students and faculty working in interdisciplinary teams • Integration of research, education, and innovation through courses, seminars, and program development • Students prepared for STEM workforce in ecoinformatics and cyberinformatics • Enhanced junior faculty research productivity and competitiveness

Diversity, outreach and communication	• Outreach activities to and participants from K-12, regional community colleges, HBCUs, etc. • Participants recruited from women and underrepresented groups • Public relations and communication	• Establishment of research and educational relationships with targeted constituents and participating institutions • Integration of portal-based research and educational resources into institutional curricula • Public recognition of cyberCommons
Sustainability	• cyberCommons-generated research and education proposals • Additional state and institutional investments in cyberCommons-related research and education	• cyberCommons resources enable projects within and beyond consortium and ecoforecasting • Expansion of supercomputing networks and CI • Contribution to KS and OK S&T development • Interaction, collaboration and partnerships with LEAD and other successful external CI groups

4.9 Sustainability Plan

The consortium will sustain the cyber-enabled activities and impacts in several ways.

Eight new faculty. Both states have demonstrated long-term commitment to the science to be enabled by the cyberCommons by creating eight new full-time tenure-track faculty positions in ecological and biodiversity forecasting. Specifically, at OU, two positions are devoted to ecological research and two to software engineering and ecoinformatics of data and models. Two positions at KU cover global biodiversity and ecological modeling; at KSU, the two recent faculty hires study ecogeography and ecophysiology. These eight faculty augment one of the strongest research groups nationally and internationally in ecological and biodiversity science and informatics.

Building on existing programs. The cyberCommons will increase research competitiveness by integrating the consortium's existing people, ideas and tools in a new CI framework for cross-domain investigation of complex environmental phenomena. Specifically, the cyberCommons will leverage and combine the strengths and expertise of: at OU—the Supercomputing Center for Education and Research, the Center for Informatics, the Center for Spatial Analysis, the Earth Observation and Modeling Center, and the Center for Ecological Forecasting and Data Assimilation; at KU—the Biodiversity Research Center, the Natural History Museum, the Kansas Biological Survey, the Kansas Applied Remote Sensing program, and the Kansas Ecological Reserves, a candidate NEON satellite site; at KSU—the Konza Prairie Biological Station, an LTER site and candidate NEON core site.

Further, the cyberCommons complements and builds upon existing NSF EPSCoR programs in both states. In Kansas, the current Track 1 RII award is funding the *science* infrastructure for ecological forecasting; in Oklahoma, the previous Track 1 award supported the cross-disciplinary area of plant virus biodiversity and ecology. Its recent Track 1 award, which focuses on ethanol production from switchgrass and related crops, as well as the pending Track 1 proposal from Kansas will strengthen the research infrastructure for understanding environmental phenomena related to changing climates and land-use/land-cover.

Building for future programs. The cyberCommons pre-adapts researchers in both states to participate fully in NSF's planned installation of NEON (www.neoninc.org) as a continental-scale research platform to sense the environment for forecasting ecological states. Similarly, The Department of Homeland Security's recommended placement of the National Bio and Agro Defense Facility in Manhattan, Kansas (www.dhs.gov/xres/labs/gc_1188509623607.shtm) will generate enormous opportunity for research partnerships that integrate ecological and biodiversity science with the genetics and epidemiology of zoonotic disease systems.

Finally, sustainability also entails sustaining knowledge discovery for society—a fundamental educational deliverable. The research and educational programs fostered by the cyberCommons and its integrative, collaborative environment will prepare a new generation of students, trained across biocomputational disciplines, to increase our knowledge of complex environmental systems. Further, CyberCommons educational portals will sustain public understanding and recognition of these systems, their application to environmental management, and the states' investments in these S&T enterprises.

4.10 Management and Coordination Plan

The KS-OK CI consortium will use a stable structure of teams and CI tools to manage and coordinate the activities of the cyberCommons through shared tasks and accountabilities. Cross-state and cross-institution teams will have specific tasks, responsibilities and performance measures; cross-team membership, in combination with the leadership team, will ensure coordination and synchrony among component areas. By encouraging shared leadership rather than hierarchy, the project will foster the capacity, ethic and work environment to share information and ideas freely—a cardinal property of creative and productive organizations.

Our primary coordination principle is a continuous focus on clarity—of project goals and milestones, of communication mechanisms and expectations, and of personal and group expectations and responsibilities. In practice, appropriate cyber tools will link people and activities, e.g., list-serves, blogs and wikis, dedicated websites, data and model sharing and visualization software, and science portals designed to address the needs of the project's constituents. Table 4 below lists the project management entities and their roles and mechanisms; Table 5 below lists the project participants.

Leadership Team. The two PIs and four co-PIs from OK and KS will form the Leadership Team: PIs Risser (OK) and Bowman-James (KS), and co-PIs Krishtalka and Dodds (KS) and Luo and May Yuan (OK). This team will have ultimate responsibility for managing and guiding the project and ensuring its success. Paul Risser has extensive EPSCoR experience, previously serving as chair of the OK State EPSCoR Committee and in scientific and academic leadership roles nationally and internationally, e.g., President of Oregon State University, OK Chancellor for Higher Education, Interim Director of the Smithsonian's National Museum of Natural History. Kristin Bowman-James, a distinguished professor of chemistry, is project director for Kansas NSF EPSCoR and PI of the current RII grant. Yiqi Luo is an internationally recognized scientist with experience in multi-investigator projects in ecology and informatics. Leonard Krishtalka, professor of ecology and evolutionary biology, is director of the KU Biodiversity Institute and serves on national and international advisory bodies for biodiversity science and informatics. He and Walter Dodds are co-PIs on a multidisciplinary, multi-institutional ecological forecasting project in Kansas. Dodds, professor of biology at KSU, serves as co-PI for the Konza Prairie LTER site. Supported by the OK/KS Project Coordinators, the Leadership Team will maintain regular e-mail and telephone communication, meet monthly by tele- and videoconference, quarterly by videoconference and semi-annually at the project symposia.

State EPSCoR Committees. The Leadership Team will report regularly to the State EPSCoR Committees, which will provide general project oversight and represent the accomplishments of the cyberCommons to respective state constituencies.

Project Directors Team. Four co-PIs, Krishtalka, Dodds, Luo and Yuan, will have primary responsibility for working with the two project coordinators and the project teams to: (1) monitor and guide the accomplishment of project deliverables and milestones across both states; (2) provide interstate and inter-institutional coordination of the science and CI components of the project plan.

The CI Team unites CI and science expertise across the collaborating universities in ecological and biodiversity informatics: James Beach and Rick McMullen, (KU); Dan Andresen (KSU); Luo, Yuan and Henry Neeman (OU). This six-person team, reporting to the Project Directors Team, is responsible for planning and implementing the hardware, software, collaborative environment and integration environment of the cyberCommons according to the scheduled project deliverables and milestones. The CI team will consult regularly and iteratively with project scientists to ensure that the CI solutions are meeting the science needs.

Ad hoc Teams will be appointed by the Project Directors Team and/or the CI Team to address new issues, ideas, technologies and new user constituencies that become important to the success of the project. These ad hoc teams could draw members from within or outside the project's current consortium, such as government agencies, scientific societies, educational institutions, and industry.

Project Coordinators Barbara Paschke (KS) and Lynda Snake (OK) will provide support and logistical services for the project teams, such as scheduling meetings, travel and symposia, coordinating hardware and software purchases, and fostering public relations and educational outreach. Paschke served as Associate Director of Kansas NSF EPSCoR, has coordinated statewide and interstate multidisciplinary projects, and has conducted public relations and communications for agencies and programs. Linda Snake has been the research grant manager for five Botany/Microbiology department faculty, with responsibilities for financial reports, personnel actions, travel, and coordination of meetings and special projects.

External Evaluation Panel. Through a contractual services arrangement, the American Association for the Advancement of Sciences (AAAS) will evaluate the project and report their findings to the State EPSCoR committees, the State EPSCoR directors, the Leadership Team, and the project participants. The AAAS panel's protocols are described in Section 4.8 Evaluation and Assessment.

External Advisory Committee. As described in Section 4.8, an external advisory committee will help the consortium implement the cyberCommons.

Table 4. Management and Coordination Structure, Roles, and Mechanisms

Who	What: roles	How: management/coordination mechanisms
State EPSCoR Committees	• Procure state matching funds • Provide state-level oversight • Provide public and government advocacy	• Promote project and state investment (match) • Review progress reports semi-annually • Meet with public and governmental bodies
Leadership Team	• Responsible for entire project • Fiscal oversight • Timely completion of project deliverables • Implementation of AAAS and EAC review recommendations • Submission of annual and final reports • External relations • Ensure grant compliance	• Hold monthly video conferences to monitor activities and progress, resolve issues, and discuss emerging opportunities • Communicate regularly with team members by e-mail and conference calls • Serve as spokespersons for the project to external audiences • Organize and compose documents for external constituents • Review annual and final report documents
Project Directors Team	• Daily project management • Monitor progress toward deliverables and milestones • Assess need for ad hoc teams • Monitor budget and review requested expenditures • Coordinate project activities • Provide evaluation materials to AAAS, EAC • Disseminate project results	• Serve as members of the Leadership Team • Hold monthly videoconferences with CI Team, quarterly videoconferences with all project personnel • Appoint and charge ad hoc teams • Organize and hold annual project symposia • Communicate regularly with all project personnel • Organize and compose project documents for external evaluators • Present project results to professional and public audiences (e.g., publications, society meetings)
CI Team	• Develop detailed implementation plan for the cyberCommons • Manage, schedule and adjust tasks to build the cyberCommons • Disseminate project results	• Develop and supervise detailed tasks and timelines for implementation of hardware and software deliverables • Review implementation plan and progress with Project Directors Team, AAAS panel, EAC and project scientists; emend plan based on feedback • Deploy communication technologies (wikis, blogs, listserves, websites, etc.) Present results at project, professional and public meetings

Project Coordinators	• Provide logistical and support services to the Leadership, Project Directors and CI teams • Coordinate activities and communications across teams, institutions and states • Assist teams with dissemination and document preparation communications • Foster educational and public outreach	• Attend team meetings; prepare agendas, minutes, budget reports, etc. • Maintain records of project personnel, correspondence, reports, etc. • Arrange and schedule meetings, workshops, videoconferences, symposia and other events • Distribute/solicit information • Prepare project reports, documents, Web materials, press releases, brochures, newsletter • Assist project teams in serving as liaison to educational organizations
Ad hoc Teams	• Address new and/or unanticipated issues and opportunities requiring specific expertise	• To be determined by the assignment and charge from the Project Directors or CI teams.
AAAS Panel	• Provide formative and summative evaluation	• Annual on-site reviews with project personnel • Annual reports to State EPSCoR Committees, Leadership Team and project participants
External Advisory Committee	• Provide periodic assessment of and recommendations for the success of the project, primarily the CI and science components	• Attend annual project symposia, meet with project teams and review project documents • Participate in videoconferences as needed • Prepare reports for the Leadership Team

Table 5. Project participants

Name	Univ	Department	Expertise
D. Andresen	KSU	Computer & Information Science	Sensor networks; HPC systems
F. Ballantyne	KU	Kansas Biological Survey; Ecology & Evolutionary Biology	Ecosystem modeling
J. Beach, co-PI	KU	Biodiversity Research Center	Research management; biodiversity informatics; CI education and training
S. Billings	KU	Kansas Biological Survey; Ecology & Evolutionary Biology	Biogeochemistry
J. Blair	KSU	Biology; Konza Prairie	Biogeochemistry
K. Bowman-James, PI	KU	KS NSF EPSCoR; Chemistry;	Research management
J. Briggs	KSU	Konza Prairie	Remote sensing and environment
N. Brunsell	KU	Geography	Biogeochemistry
D. Brunson	OSU	HPC center	Physical CI
D. Caragea	KSU	Computer & Information Science	Data mining; information visualization
D. Chandler	KSU	Civil Engineering	Ecohydrology
W. Dodds, co-PI	KSU	Biology; Konza Prairie	Biogeochemistry
S. Egbert	KU	Kansas Applied Remote Sensing; Geography;	Remote sensing data
C. Ferguson	KSU	Biology; Herbarium	Biodiversity data and informatics
C. Freeman	KU	Biodiversity Research Center; Kansas Biological Survey	Biodiversity data and informatics; training
K. Gido	KSU	Biology	Biodiversity informatics
C. J. Grady	KU	Biodiversity Research Center	Geospatial data and software engineering
J. Grissom	OU	Center for Informatics	Software engineering/development
L. Gruenwald, co-PI	OU	Computer Science	Databases, data models, data mining

J. Kelly	OU	Oklahoma Biological Survey	Animal biodiversity, infectious disease modeling
L. Krishtalka, co-PI	KU	Biodiversity Institute; Ecology & Evolutionary Biology	Research management; biodiversity data and informatics
S. Lakshmivarahan	OU	Computer Science	Data assimilation
Y. Luo, co-PI	OU	Botany/Microbiology	Biogeochemical modeling, data/ model assimilation
E. Martinko	KU	Kansas Biological Survey; Ecology & Evolutionary Biology	Research management; biodiversity; remote sensing technologies
A. McGovern	OU	Computer Science	Data mining/knowledge discovery, machine learning
R. McMullen, co-PI	KU	Center for Research	CI architecture; Research computing
H. Neeman	OU	OSCER	Physical CI, HPC software development; CI education and training
J. Nippert	KSU	Biology	Ecophysiology
M. Palmer, co-PI	OSU	Botany	Plant biodiversity, biostatistics
B. Paschke	KU	Kansas Biological Survey	Project coordination, outreach, education
D. Peterson	KU	Kansas Biological Survey	Remote sensing databases; land-cover mapping
A. T. Peterson	KU	Biodiversity Research Center; Ecology & Evolutionary Biology	Biodiversity informatics, ecological niche modeling, disease modeling;
C. Rice	KSU	Agronomy	Biogeochemistry, soils
P. Risser, PI	OU	Research Cabinet	Research management; Ecology
C. Scoglio	KSU	Electrical and Computer Engineering	Overlay and virtual networks; modeling
L. Snake	OU	Botany/Microbiology	Project coordination and management
J. Soberon	KU	Biodiversity Research Center; Ecology & Evolutionary Biology	Biodiversity informatics and modeling; Global biodiversity policy
A. Stewart	KU	Biodiversity Research Center	Geospatial data and software engineering
D. Vieglais	KU	Biodiversity Research Center	Biodiversity informatics, data mining and modeling, software engineering; CI training
C. Weaver	OU	Computer Science	Human computer interaction, visualization, visual analytics
E. Wiley	KU	Biodiversity Research Center; Ecology & Evolutionary Biology	Biodiversity, informatics, ecological niche modeling
S. Wisely	KSU	Biology	Ecology and disease epidemiology
X. Xiao	OU	Botany/Microbiology	Remote sensing, biogeochemical modeling, ecology & disease modeling
M. Yuan, Co-PI	OU	Atmospheric & Geographic Science	Temporal GIS, data modeling, web mapping; science education
G. Zolnerowich	KSU	Biology	Biodiversity informatics

E. REFERENCES

Adjemian, J., E. Girvetz, L. Beckett and J. Foley. 2006. Analysis of Genetic Algorithm for Rule-Set Production (GARP) modeling approach for predicting distributions of fleas implicated as vectors of plague, *Yersinia pestis*, in California. *Journal of Medical Entomology,* 43:93:103.

Ajwa, H. A., C. W. Rice, and D. Sotomayor. 1998. Carbon and nitrogen mineralization in tallgrass prairie and agricultural soil profiles. *Soil Science Society of America Journal* 62:942-951.

Albertson, F. W. and J. E. Weaver. 1954. Nature and degree of recovery of grassland from the great drought of 1933 to 1940. *Ecological Monographs,* 14:395-479.

†Andresen, D., J. Hatcliff, G. Singh and S. Warren. 2007. Observations and directions for high-confidence sensor networks, *Proceedings of the Composable and Systems Technology for High Confidence Cyber-P*, 1.

*Arnone III, J. A., P. S. J. Verburg, D. W. Johnson, J. D. Larsen, R. L. Jasoni, A. J. Lucchesi, C. M. Batts, C. von Nagy, W. G. Coulombe, D. E. Schorran, P. E. Buck, R. H. Braswell, J. S. Coleman, R. A. Sherry, L. L. Wallace, Y. Q. Luo, D. S. Schimel. 2008. Prolonged Suppression of Ecosystem Carbon Dioxide Uptake After an Anomalously Warm Year. *Nature*, 455:383-386.

Baer, S. G., D. J. Kitchen, J. M. Blair and C. W. Rice. 2002. Changes in ecosystem structure and function along a chronosequence of restored grasslands. *Ecological Applications,* 12:1688-1701.

Baer, S. G. and J. M. Blair. 2008. Grassland establishment under varying resource availability: A test of positive and negative feedback. *Ecology* 89:1859-1871.

Baret, F. and T. Fourty. 1997. Radiometric estimates of nitrogen status in leaves and canopies in G. Lemaire (Ed.), *Diagnosis of the Nitrogen Status in Crops.* Berlin:Springer, 201-227.

†Behmer, S.T. and A. Joern. 2008. Coexisting generalist herbivores occupy unique nutritional feeding niches, *Proceedings of the National Academy of Sciences* (USA), 105, 1977-1982.

Bengis, R. H., F. A., Leighton, J. R., Fischer, M. Artois, T. Morner and C. M. Tate. 2004. The role of wildlife in emerging and re-emerging zoonoses. *Revue Scientifique Et Technique De L Office International Des Epizooties,* 23:497-511.

†Bengtson, J. R., M. A. Evans-White and K. B. Gido. 2008. Effects of grazing minnows and crayfish on stream ecosystem structure and function, *Journal of the North American Benthological Society*, 27, 772-778.

Blair, J. M. 1997. Fire, N availability and plant response in grasslands: A test of the transient maxima hypothesis. *Ecology* 78:2539-2368.

†Branson, D., H. Branson, A. Joern and G. A. Sword. 2006. Sustainable management of insect herbivores in grassland ecosystems: new perspectives in grasshopper control, *BioScience*, 56, 743-755.

Braswell, B. H., D. S. Schimel, J. L. Privette, B. Moore, W. J. Emery, E. W. Sulzman and A. T. Hudak. 1996. Extracting ecological and biophysical information from AVHRR optical data: An integrated algorithm based on inverse modeling. *Journal of Geophysical Research-Atmosphere,* 101:23335-23348.

Briggs, J. M., A. K. Knapp, J. M. Blair, J. L. Heisler, G. A. Hoch, M. S. Lett and J. K. McCarron. 2005. An ecosystem in transition: causes and consequences of the conversion of mesic grassland to shrubland. *BioScience,* 55:243-254.

Broussard, L. A. 2001. Biological agents: Weapons of warfare and bioterrorism. *Molecular Diagnosis,* 6:323-333.

†Brunsell, N. A., J. M. Ham and C. E. Owensby. 2008a. Assessing the multi-resolution information content of remotely sensed variables and elevation for evapotranspiration in a tallgrass prairie environment, *Remote Sensing of Environment*, 112, 2977-2987.

†Brunsell, N. A., J. M. Ham and C. E. Owensby, 2008b, Year 3: Assessing the multi-resolution information content of remotely sensed variables and elevation for evapotranspiration in a tall-grass prairie environment, *Remote Sensing of Environment*, 112, 2977-2987.

†Bulatewicz, T., D. Andresen, S. Welch, W. Jin, S. Das and M. Miller, 2007. A Framework for Scalable Parameter Estimation on Clusters, *Proc. of the 8th LCI International Conf. on High-Performance Clustered Computing*, 1.

Chen, H. Y., H. H. Stratton, T. B. Caraco and D. J. White. 2006. Spatiotemporal Bayesian analysis of Lyme disease in New York State, 1990-2000. *Journal of Medical Entomology,* 43:777-784.

Clark, J. S. and A. E. Gelfand. 2006. A future for models and data in environmental science. *Trends in Ecology and Evolution,* 21(7):375-80.

Cohen, M. L. 2000. Changing patterns of infectious disease. *Nature,* 406:762-767.

Colwell, R., P. Epstein, D. Gubler, M. Hall, P. Reiter, J. Shukla, W. Sprigg, E. Takafuji, and J. Trtanj. 1998. Global climate change and infectious diseases. *Emerging Infectious Diseases,* 4: 451-452.

Collinge, S. K. and C. Ray, editors. 2006. *Disease Ecology: Community Structure and Pathogen Dynamics.* Oxford University Press, Oxford.

Costa, J., A. T. Peterson and C. B. Beard. 2002. Ecological niche modeling and differentiation of populations of Triatoma Basiliensis Neiva, 1911, the most important Changas disease vector in northeastern Barzil (Hemiptera, Reduviidae, Triatominae). *American Journal of Tropical Medicine & Hygiene* 67:516-520.

†Craine, J. M., E. G. Towne, A. Joern and R. G. Hamilton, in press, Consequences of climate variabiltiy for the performance of bison in tallgrass prairie, *Global Change Biology*.

Daszak, P., A. A. Cunningham and A. D. Hyatt. 2000. Wildlife ecology - Emerging infectious diseases of wildlife - Threats to biodiversity and human health. *Science,* 287:443-449.

Daszak, P., K. Jones, M. Levy, J. Gittleman, N. Patel, V. Mara, S. F. Nakielny, J. Chan and M. N. Labo. 2007. Global trends in emerging infectious diseases. *American Journal of Tropical Medicine and Hygiene,* 77:284-285.

DeRoure, D. and C. Goble. 2007. myExperiment - A Web 2.0 Virtual Research Environment. In: *International Workshop on Virtual Research Environments and Collaborative Work Environments*, May 2007, Edinburgh, UK. http://eprints.ecs.soton.ac.uk/13961/

Diuk-Wasser, M. A., H. E. Brown, T. G. Andreadis and D. Fish. 2006. Modeling the spatial distribution of mosquito vectors for West Nile virus in Connecticut, USA. *Vector-Borne and Zoonotic Diseases,* 6:283-295.

†Dodds, W. K. and J. J. Cole. 2007. Expanding the concept of trophic state in aquatic ecosystems: It's not just the autotrophs, *Aquatic Sciences*, 69, 427-439.

†Dodds, W. K. 2008. Trophic state, eutrophication, and nutrient criteria in streams, *Trends in Ecology and Evolution*, 22, 670-676.

Dodds, W. K. 2008. *Humanity's Footprint. Momentum, Impact and Our Global Environment.* University of Columbia Press. 270 pp.

†Dodds, W. K. and R. M. Oakes. 2008. Headwater influences on downstream water quality, *Environmental Management*, 41, 367-377.

Dodds, W. K. In press. *Laws, Theories and Patterns in Ecology*. University of California Press.

Falke, J. A. and K. B. Gido. 2006. Effects of reservoir connectivity on stream fish assemblages in the Great Plains. *Canadian Journal of Fisheries and Aquatic Sciences,* 63:480-493.

Fay, P. A., J. D. Carlisle, A. K. Knapp, J. M. Blair and S. L. Collins. 2000. Altering rainfall timing and quantity in a mesic grassland ecosystem: Design and performance of rainfall manipulation shelters. *Ecosystems* 3:308-319.

Fox, Geoffrey C., R. Guha, D. F. McMullen, A. F. Mustacoglu, M. E. Pierce, A. E. Topcu and D. J. Wild. 2007. Web 2.0 for Grids and e-Science INGRID 2007 - Instrumenting the Grid 2nd International *Workshop on Distributed Cooperative Laboratories* - S.Margherita Ligure Portofino, ITALY.

†Gelso, B. R., J. A. Fox, and J. M. Peterson. 2008. The Cost of Wetlands: Effects of Size, Hydration, and Dispersion, *American Journal of Agricultural Economics*, 90, 172-185.

Gido, K. B., C. S. Guy, T. R. Strakosh, R. J. Bernot, K. J. Hase and M. A. Shaw. 2002. Long-term changes in the fish assemblages of the Big Blue River Basin 40 years after the construction of Tuttle Creek Reservoir. *Transactions of the Kansas Academy of Science* 105:193-208.

†Gido, K. B., J. F. Schaefer and J. A. Falke. In revision. Convergence of littoral zone fish communities in reservoirs, *Freshwater Biology*.

Gilbert, M., X. Xiao, P. Chaitaweesub, W. Kalpravidh, S. Premashthira, S. Boles and J. Sligenbergh. 2007. Avian influenza, domestic ducks and rice agriculture in Thailand. *Agriculture, Ecosystems & Environment* 119:409-415.

Gruenwald, L., H. Chok, H. and M. Aboukhamis. 2007. Using Data Mining to Estimate Missing Sensor Data, *The IEEE ICDM Workshop on Optimization-Based Data Mining Techniques with Applications*.

Harper, C. W., J. M. Blair, P. A. Fay, A. K. Knapp, and J. D. Carlisle. 2005. Increased rainfall variability and reduced rainfall amount decreases soil CO2 flux in a grassland ecosystem. *Global Change Biology* 11:322-344.

†Harrington, J. 2006. Climate Change and Persistence in Western Kansas: A Reanalysis. *Papers of Applied Geography Conferences*, 29, 49-55.

Harvell, C. D., C. E. Mitchell, J. R. Ward, S. Altizer, A. P. Dobson, R. S. Ostfeld, M. D. Samuel. 2002. Climate Warming and Disease Risks for Terrestrial and Marine Biota. *Science* 296(5576): 2158–2162. DOI: 10.1126/science.1063699.

Hasting, W. K. 1970. Monte Carlo sampling methods using Markov chain and their applications. *Biometrika,* 57:97-109.

†Hoeinghaus, D. J. and S. C. Zeug. 2008, Can stable isotope ratios provide for community-wide measures of trophic structure? Comment, *Ecology* 89:2353-2357.

†Hoeinghaus, D. J., K. O. Winemiller and A.A. Agostinho. 2008. Hydrogeomorphology and river impoundment affect food-chain length in diverse Neotropical food webs. *Oikos*, 117:984-995.

Hooper U., F. S. Chapin III, J. J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J. H. Lawton, D. M. Lodge, M. Loreau, S. Naeem, B. Schmid, H. Setälä, A. J. Symstad, J. Vandermeer, D. A. Wardle. 2005. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs* 75:1 pp. 3-35.

Jiang, X. R. and L. Gruenwald, 2005. Microarray Gene Expression Data Association Rules Mining Based On BSC-Tree and FIS-Tree, *Data and Knowledge Engineering* 53:1.

Jiang, N. and L. Gruenwald. 2006. CFI-Stream: Mining Closed Frequent Itemsets in Data Streams, the *Proceedings of the ACM International Conference on Data and Knowledge Discovery* (KDD).

†Jin, W. S. Staggenborg, S. Lauwo, M. Miller, S. Das, D. Andresen, J. Peterson, D. Steward and S. M. Welch. 2006. Coupling Crop and Hydrologic Models for Decision Support in the High Plains Aquifer: Preliminary Steps, *Proceedings of the 5th International Conference on the Analytic Element Method*, 1.

†Jin, W. and D. R. Steward. 2007. The transition of flow patterns through critical stagnation points in two-dimensional groundwater flow. *Advances in Water Resources*, 30, 16-28.

Johnson, L. C. and J. R. Matchett. 2001. Fire and grazing regulate belowground processes in tallgrass prairie. *Ecology,* 82:3377-3389.

†Jonas, J. L. and A. Joern. In press. Host plant quality alters grass:forb consumption by a mixed-feeding insect herbivore, Melanoplus bivittatus, *Ecological Entomology*.

Jones, K. E., N.G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L.Gittleman, and P. Daszak. 2008. Global trends in emerging infectious diseases. *Nature* 451:990-993.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82(1):35–45.

†Kettle, N., L. Harrington, J. Harrington, Jr. 2007. Groundwater Depletion and Agricultural Land Use Change in the High Plains: A Case Study from Wichita County, Kansas. *The Professional Geographer* 59:221-235.

Knapp, A. K., P. A. Fay, S. L. Collins, M. D. Smith, J. D. Carlisle, C. W. Harper, B.T. Danner, M. S. Lett and J. K. McCarron. 2002. Rainfall variability, carbon cycling, and plant species diversity in a mesic grassland. *Science* 298:2202-2205.

Knapp, A. K., C. Beier, D. D. Briske, A. T. Classen, Y. Luo, M. Reichstein, M. D. Smith, S. D. Smith, J. E. Bell, P. A. Fay, J. L. Heisler, S. W Leavitt, R. Sherry, B. Smith and E. Weng. 2008. Consequences of more extreme precipitation regimes for terrestrial ecosystems. *BioScience* 58:811-821.

Kremen, C., A. Cameron, A. Moilanen, S. J. Phillips, C. D. Thomas, H. Beentje, J. Dransfield, B. L. Fisher, F. Glaw, T. C. Good, G. J. Harper, R. J. Hijmans, D. C. Lees, E. Louis, Jr., R. A. Nussbaum, C. J. Raxworthy, A. Razafimpahanana, G. E. Schatz, M. Vences, D. R. Vieites, P. C. Wright, M. L. Zjhra. 2008. Aligning Conservation Priorities Across Taxa in Madagascar with High-Resolution Planning Tools. *Science* 320(5873): 222–226. DOI: 10.1126/science.1155193.

LaDeau, S. L., A. M. Kilpatrick and P. P. Marra. 2007. West Nile virus emergence and large-scale declines of North American bird populations. *Nature* 447:710-U713.

Loreau M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, M. A. Huston, D. Raffaelli,8 B. Schmid, D. Tilman, D.A. Wardle. 2001. Biodiversity and Ecosystem Functioning: Current Knowledge and Future Challenges. *Science* 26:294:5543.pp. 804-808.

*Luo, Y. 2007. Terrestrial Carbon-Cycle Feedback to Climate Warming. *Annual Review of Ecology, Evolution, and Systematics* 38:683-712.

Luo, Y. Q., L. H. Wu, J. A. Andrews, L. White, R. Matamala, K. V. R. Schafer and W. H. Schlesinger. 2001. Elevated CO2 differentiates ecosystem carbon processes: Deconvolution analysis of Duke Forest FACE data. *Ecological Monographs,* 71:357-376.

Luo, Y. Q., L. W. White, J. G. Canadell, E. H. DeLucia, D. S. Ellsworth, A. Finzi, J. Lichter, and W. H. Schlesinger. 2003. Sustainability of terrestrial carbon sequestration: A case study in Duke Forest with inversion approach. *Global Biogeochemical Cycles,* 17(1)1021, doi:10.1029/2002GB001923.

Luo, Y., D. Hui, D. Zhang. 2006. Elevated Carbon Dioxide Stimulates Net Accumulations of Carbon and Nitrogen in Terrestrial Ecosystems: A Meta-Analysis. *Ecology* 87:53-63.

Luo YQ, R. Sherry, XH Zhou, SQ Wan. In press. Terrestrial Carbon-Cycle Feedback to Climate Warming: Experimental Evidence on Plant Regulation and Impacts of Biofuel Feedstock Harvest. *Global Change Biology.*

*Luo, Y. Q., E. S. Weng, X. W. Wu, C. Gao, X. H. Zhou and L. Zhang. In press. Parameter Identifiability, Constraint, and Equifinality in Data Assimilation with Ecosystem Models. *Ecological Applications*.

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y., Gottlieb, E. 2008. A satellite-based biosphere parameterization for net ecosystem CO2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), *Global Biogeochemical Cycles*, 22, doi:10.1029/2006GB002735.

†Marshall, J. D., J. M. Blair, D. Peters, G. Okin, A. Rango and M. Williams. 2008. Predicting and Understanding Ecosystem Responses to Climate Change at Continental Scales, *Frontiers in Ecology and the Environment*, 6, 273-280.

McGovern, Amy and Nathan Hiers, and Matthew Collier, and David J. Gagne II, and Rodger A. Brown. 2008. Spatiotemporal Relational Probability Trees. *Proceedings of the 2008 IEEE International Conference on Data Mining.*

McKinley, D. C. and J. M. Blair. 2008. Woody plant encroachment by Juniperus virginiana in a mesic native grassland promotes rapid carbon and nitrogen accrual. *Ecosystems* 11:454-468.

†McLauchlan, K. K. 2008. Reconstructing nutrient and energy flows through terrestrial ecosystems of the Holocene. *Palaios* 23:721-723.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller. 1953. Equations of state calculations by fast computing machines. *Journal of Chemical Physics* 21:1087-1092.

†Morse, C. A., C. C. Freeman and R.L. McGregor. 2007. New, corrected, and interesting records for the Kansas vascular flora. *J. Bot. Res. Inst. Texas*, 1:753-761.

†Murdock, J. N., W. K. Dodds and D. l. Wetzel. 2008. Subcellular chemical imaging of localized benthic algal nutritional content via HgCdTe array FT-IR. *Vibrational Spectroscopy* 49:179-188.

Naeem, Shahid, Lindsey J. Thompson, Sharon P. Lawler, John H. Lawton, 1995. Empirical Evidence that Declining Species Diversity May Alter the Performance of Terrestrial Ecosystems, Volume 347:1321, pp. 249-262.

Neeman, H., J. Mullen, L. Lee and G. Newman, 2002: Supercomputing in Plain English: Teaching High Performance Computing to Inexperienced Programmers. *Proceedings of the 3rd LCI International Conference on Linux Clusters: The HPC Revolution 2002.*

Neeman, H., L. Lee, J. Mullen and G. Newman. 2006: Analogies for Teaching Parallel Computing to Inexperienced Programmers. *inroads: SIGCSE Bulletin* 38(4):64-67.

Neeman, H., H. Severini and D. Wu. 2008: Supercomputing in Plain English: Teaching Cyberinfrastructure to Computing Novices. *inroads: SIGCSE Bulletin* 40(2):27-30.

Neeman, H., H. Severini, D. Wu and K. Kantardjieff. 2008: Teaching Supercomputing via Videoconferencing. *Proceedings of TeraGrid 2008.* Awarded Best Education Paper.

O'Connor, M. I., J. F. Bruno, S. D. Gaines, B. S. Halpern, S. E. Lester, B. P. Kinlan and J. M. Weiss. 2007. Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proceedings of the National Academy of Sciences* 104:1266-1271. DOI:10.1073/pnas.0603422104.

Palmer, M. W., P. Earls, B. W. Hoagland, P. S. White and T. M. Wohlgemuth. 2002. Quantitative tools for perfecting species lists. *Environmetrics*. 13:121-137.

Palmer, M. W. 2005. Temporal trends of exotic species richness in North American floras: an overview. *Ecoscience* 12:386-390.

Palmer, M. W. 2006. Scale dependence of native and alien species richness in North American floras. *Preslia* 78:427-436.

Palmer, M. W. 2007. Species-area curves and the geometry of nature. Pp. 15-31 *in* D. Storch, P. A. Marquet and J. H. Brown, editors. *Scaling Biodiversity*. Cambridge University Press, Cambridge.

Parmesan, C., N. Ryrholm, C. Stefanescu, J. K. Hill, C. D. Thomas, H. Descimon, B. Huntley, L. Kaila, J. Kullberg, T. Tammaru, W. J. Tennent, J.A. Thomas and M. Warren. 1999. Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature* 399:579-583. DOI:10.1038/21181.

Parmesan, C. 2006. Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics* 37: 637-669.

*Pataki, D.E., T. Xu, Y. Luo and J. R. Ehleringer. 2007. Inferring biogenic and anthropogenic CO_2 sources across an urban to rural gradient. *Oecologia*. 152:307-322.

Peterson, A. T. 2003. Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology* 78:419-433.

Peterson, A. T. 2006. Ecological niche modeling and spatial patterns of disease transmission. *Emerging Infectious Diseases* 12:1822-1826. Peterson, A. T. 2007. Ecological niche modelling and understanding the geography of disease transmission. *Veterinaria Italiana* 43:393-400.

Peterson, A. T. 2008. Biogeography of diseases: A framework for analysis. *Naturwissenschaften*.

Peterson, A. T., J. Soberón, and V. Sánchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. *Science* 285:1265-1267.

Peterson, A. T., V. Sanchez-Cordero, C. B. Beard and J. M. Ramsey. 2002a. Ecologic niche modeling and potential reservoirs for Changas disease, Mexico. *Emerging Infectious Diseases* 8:662-667.

Peterson, A. T., M. A. Ortega-Huerta, J. Bartley, V. Sánchez-Cordero, J. Soberón, R. H. Buddemeier and D. R. B.Stockwell. 2002b. Future projections for Mexican faunas under global climate change scenarios. *Nature* 416:626-629.

Peterson, A. T. and J. Shaw. 2003. Lutzomyia vectors for cutaneous leishmaniasis in Southern Brazil: ecological niche models, predicted geographic distributions, and climate change effects. *International Journal for Parasitology* 33(9):919-931.

Peterson, A. T., C. Martínez-Campos, Y. Nakazawa, and E. Martínez-Mayer. 2005. Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 99:647-655.

Peterson, A. T., R. R. Lasdh, D. S. Carroll, and K. M. Johnson. 2006. Geographic potential for outbreaks of Marburg hemorrhagic fever. *American Journal of Tropical Medicine & Hygiene* 75:9-15.

Peterson, A. T., M. Papes, D. S. Carroll, H. Leirs and K. M. Johnson. 2007. Mammal taxa constituting potential coevolved reservoirs of filoviruses. *Journal of Mammalogy* 88:1544-1554.

†Peterson, J. M. and K. Schoengold. 2008. Using Numerical Methods to Address Water Supply and Reliability Issues: Discussion. *American Journal of Agricultural Economics*, 90, 1350-51.

Peterson, A. T., A. Robbins, R. Restifo, J. Howell and R. Nasci. 2009. Predictable ecology and geography of West Nile virus transmission in the central United States. *Journal of Vector Ecology.*

Pierce, Marlon E., G. C. Fox, J. Y. Choi, Z. Guo, X. Gao and Y. Ma. 2007. Using Web 2.0 for Scientific Applications and Scientific Communities. To appear in: *Concurrency and Computation: Practice and Experience*. Special Issue for 3rd International Conference on Semantics, Knowledge and Grid SKG2007, Xian China October 28-30.

Pounds, J. A., M. R. Bustamante, L. A. Coloma, J. A. Consuegra, M. P. L. Fogden, P. N. Foster, E. La Marca, K.L. Masters, A. Merino-Viteri and R. Puschendorf. 2006. Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* 439:161-167.

Reich, P. B., J. Knops, D. Tilman, J. Craine, D. Ellsworth, M. Tjoelker, T. Lee, S. Naeem, D. Wedin, D. Bahauddin, G. Hendrey, S. Jose. K. Wrage, J. Goth and W. Bengston. 2001. Plant diversity enhances ecosystem responses to elevated CO2 and nitrogen deposition. *Nature* 410:809-812.

†Saak, A. E. and J. M. Peterson. 2007. Groundwater Use under Incomplete Information, *Journal of Environmental Economics and Management* 54:214-228.

Samson, F. and F. Knopf. 1994. Prairie conservation in North America. *BioScience* 44:418-421.

Schlapfer, F. and B. Schmid. 1999. Ecosystem effects of biodiversity: a classification of hypotheses and exploration of empirical results. *Ecological Applications* 9(3):893-912.

Schmid B, J. Joshi, and F. Schläpfer. 2002. Empirical evidence for biodiversity-ecosystem functioning relationships. In: *Functional Consequences of Biodiversity: Experimental Progress and Theoretical Extensions*-Kinzig A, Tilman D, Pacala S, eds. Princeton, NJ: Princeton University Press.

Schimel D., J. Melillo, H. Q. Tian, A. D. McGuire, D. Kicklighter, T. Kittel, N. Rosenbloom, S. Running, P. Thornton, D. Ojima, W. Parton, R. Kelly, M. Sykes, R. Neilson and B. Rizzo. 2000. Contribution of increasing CO2 and climate to carbon storage by ecosystems in the United States. *Science* 287:2004-2006.

Schlesinger, W. H. 1997. *Biogeochemistry - An Analysis of Global Change.* Academic Press, San Diego, CA.

Shaw, M. R., E. S. Zavaleta, N. R. Chiariello, E. E. Cleland, H. A. Mooney, C. B. Field. 2002. Grassland responses to global environmental changes suppressed by elevated CO2. *Science* 298:1987-1990.

Sherry, A., X. Zhou, S. Gu, J. A. Amone III, D. S. Schimel, P. S. Verburg, L. L. Wallace, and Y. Luo. 2007. Divergence of reproductive phenology under climate warming. *PNA,* 104:198-202.

*Sherry, R. A., E. Weng, J. J. Arnone III, D. W. Johnson, D. S. Schimel, P. S. Verburg, L. L Wallace and Y. Luo. 2008. Lagged Effects of Experimental Warming and Doubled Precipitation on Annual and Seasonal Aboveground Biomass Production in a Tallgrass Prairie. *Global Change Biology* 14:2923-2936.

†Smith, C. M., J. M. Peterson and J. C. Leatherman, 2007, Attitudes of Great Plains Producers About Best Management Practices, Conservation Programs, and Water Quality, *Journal of Soil and Water Conservation* 62, 97A-103A.

Soberón, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters* 10:1115-1123.

Soberón, J. and A. T. Peterson. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics* 2:1-10.

Soberón, J., R. Jiménez, J. Golubov and P. Koleff, P. 2007. Assessing completeness of biodiversity databases at different spatial scales. *Ecography* 30:152-160.

Soberón, J. and A. T. Peterson, A.T. 2008. Monitoring biodiversity loss with primary species-occurrence data: Toward national-level indicators for the 2010 Target of the Convention on Biological Diversity. *AMBIO.*

Stasko, J. (2007), Carsten Goerg, Zhicheng Liu, and Kanuoriya Singhal. Jigsaw: Supporting Investigative Analysis through Interactive Visualization. *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST)*, 131-138, IEEE.

†Steward, D. R. 2007. Groundwater response to changing water-use practices in sloping aquifers, *Water Resources Research* 43, W05408.

†Steward, D. R. and E. A. Bernard. 2007. Integrated engineering and landscape architecture approaches to address groundwater declines in the High Plains Aquifer. *AEESP* 121-134.

†Steward, D. R., P. Le Grand, I. Jankovic and O. D. L. Strack. 2008. Cauchy integrals for boundary segments with curvilinear geometry. *Proceedings of The Royal Society of London.* Series A, 464, 223-248.

†Steward, D. R. and T. Ahring. In press. An analytic solution for groundwater uptake by phreatophytes spanning spatial scales from plant to field to regional. *Journal of Engineering Mathematics*.

†Steward, D. R., X. Yang and S. Chacon. In press. Groundwater response to changing water-use practices in sloping aquifers using convolution of transient response functions. *Water Resources Research*.

Thomas, J. and K. A. Cook. 2005. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society.

†Tiemann, L. and S. A. Billings. 2008. Carbon controls on nitrous oxide production with changes in substrate availability in a North American grassland. *Soil Science* 17:332-341.

†Upendram, S. and J. M. Peterson. 2007. Irrigation Technology and Water Conservation in the High Plains Region. *Journal of Contemporary Water Research and Education* 137, 40-46.

Wardle, David A., Karen I. Bonner, Gary M. Barker, 2000. Stability of ecosystem properties in response to above-ground functional group richness and composition. *Oikos* 89(1):11-23.

Weaver, C. 2004. Building Highly-Coordinated Visualizations in Improvise. *Proceedings of the IEEE Symposium on Information Visualization*, 159-166, IEEE Computer Society.

Weaver, C., David Fyfe, Anthony Robinson, Deryck W. Holdsworth, Donna J. Peuquet and Alan M. MacEachren. 2007. Visual Exploration and Analysis of Historic Hotel Visits. *Information Visualization* 6:1, 89-103, Palgrave-Macmillan.

Wenger, Etienne. 1998. *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.

White, L., and Y. Luo. 2002. Inverse analysis for estimating carbon transfer coefficients in Duke Forest. *Applied Mathematics and Computation* 130:101-120.

*White, L. and Y. Luo. 2005. Model-based CO_2 data assessment for terrestrial carbon processes: Implications for sampling strategy in FACE experiments. *Applied Mathematics and Computation* 167:419-434.

White, L. W., Y. Luo and T. Xu. 2005. Carbon sequestration: inversion of FACE data and prediction. *Applied Mathematics and Computation* 163: 783-800.

*White, L., F. White, T. Xu and Y. Luo. 2006. Estimation of Parameters in Carbon Sequestration Models from Net Ecosystem Exchange Data. *Applied Mathematics and Computation* 181:864-879.

White, L., F. White, Y. Luo and T. Xu. 2006. Estimation of parameters in carbon sequestration models from net carbon exchange data. *Applied Mathematics and Computation* 181:864-879.

*White, L. and Y. Q. Luo. 2008. Modeling and inversion of net ecological exchange data using an Ito stochastic differential equation approach. *Applied Mathematics and Computation* 196:686-704.

Williams, S., E. E. Bolitho and S. Fox. 2003. Climate change in Australian tropical rainforests: an impending environmental catastrophe. *Proceedings of The Royal Society: Biological Sciences* 270: 1887.

Williams, R., F. O. Fasina and A. T. Peterson. 2008. Predictable ecology and geography7 of avian influenza (H5N1) transmission in Nigeria and East Africa *Transactions of the Royal Society of Tropical Medicine and Hygiene* 102:471-479.

†Williams, J.R., S. Mooney and J. M. Peterson. In press. What is the Carbon Market: Is There a Final Answer? *Journal of Soil and Water Conservation*.

Xiao, X., Q. Zhang, D. Hollinger, J. Aber, J., and B. Moore, 2005. Modeling seasonal dynamics of gross primary production of evergreen needleleaf forest using MODIS images and climate data, *Ecological Applications* 15(3):954-969.

Xiao, S., Gilbert, M., Slingenbergh, J., Lei, F., and Boles, S., 2007. Remote sensing ecological variables and wild bird migration related to outbreaks of highly pathogenic H5N1 bird flue, *Journal of Wildlife Diseases* 43(3, Supplement):S40-S46.

*Xu, T., L. White, D. Hui and Y. Luo. 2007. Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction, *Global Biogeochemical Cycles* 20, GB2007, doi:10.1029/2005GB002468.

Yuan, M. 2007. Adding Time. *Handbook of Geographic Information Science*, J. Wilson and S. Fotheringham, editors, published by Blackwell Publishing. P. 169-184.

*Yuan, W., Y. Luo, A. D. Richardson, R. Oren, S. Luyssaert, I. A. Janssens, R. Ceulemans, S. Zhou, T. Grünwald, M. Aubinet, C. Berhofer, D. D. Baldocchi, J. Chen, A. L. Dunn, J. Deforest, D. Dragoni, A. H. Goldstein, E. Moors, J. W. Munger, R. K. Monson, A. E. Suyker, G. Starr, R. L. Scott, J. Tenhunen, S. B. Verma, T. Vesala and S. C. Wofsy. 2009. Latitudinal Patterns of Magnitude and Interannual Variability in Net Ecosystem Exchange Regulated by Biological and Environmental Variables. *Global Change Biology.* In press.

Yuan, M. 2009. Knowledge discovery of geographic dynamics in spatiotemporal data. In *Geographic Data Mining and Knowledge Discovery*(2ed) edited by H. Miller and J. Han. To be published by CRC/Taylor and Francis.

Zhang, Q., X. Xiao, B. Braswell, E. Linder, E., F. Baret, and M. Berrien. 2005. Estimating light absorption by chlorophyll, leaf and canopy in a deciduous broadleaf forest using MODIS data and a radiative transfer model. *Remote Sensing of Environment 99*, 357-371.

Zhang, Q., X. Xiao, B. Braswell,, E. Linder, S. Ollinger, M. L. Smith, J. P. Jenkins, F. Baret, A. Richardson, A. and B. Moore, B. 2006. Characterization of seasonal variation of forest canopy in a temperate deciduous broadleaf forest, using daily MODIS data. *Remote Sensing of Environment,* In Press.

Zhang, J., W. Liu, and L. Gruenwald. 2007. A Successive Decision Tree Approach to Mining Remotely Sensed Image Data. In, *Knowledge Discovery and Data Mining: Challenges and Realities with Real World Data*, X. Zhu and I. Davidson (eds.), Idea Group Publishing, Inc., ISBN 1-59904-252-5.

Zhang, J. and L. Gruenwald. 2008. Embedding and Extending GIS for Exploratory Analysis of Large-Scale Species Distribution Data, *The ACM International Conference on Advances in GIS Systems*.

Zhou, T., and Y. Luo. 2008. Spatial Patterns of Ecosystem Carbon Residence Time and NPP-Driven Carbon Uptake in the Conterminous USA. *Global Biogeochemical Cycle*. 22, GB3032, doi:10.1029/2007GB002939.

*Zhou, X., S. Wan and Y. Luo. 2007. Source components and interannual variability of soil CO_2 efflux under experimental warming and clipping in a grassland ecosystem. *Global Change Biology*. 13, 761–775, doi: 10.1111/j.1365-2486.2007.01333.x

Zhou, X., E.Weng, Y. Luo. 2008. Modeling patterns of nonlinearity in ecosystem responses to temperature, CO_2, and precipitation changes. *Ecological Applications*. 18: 453-466