Challenges for the Implementation of the Decadal Plan for Long-Term Ecological Research: Land and Water Use Change

Report of a Workshop

Held in San Juan, Puerto Rico
December 16-19, 2008

Sponsored by the LTER Network Office, the National Science Foundation, and the University of Puerto Rico

Steering Committee:
Jess K. Zimmerman (LUQ), Frederick N. Scatena (LUQ), Laura C. Schneider (HBR),
Ted Gragson (CWT), Christopher Boone (CAP),
and J. Morgan Grove (BES)

May 1, 2009

Executive Summary

The Decadal Plan for Long-Term Ecological Research (DP-LTER) in the United States calls for integrated studies of social and ecological systems (SESs) at each the 26 LTER sites. We organized a workshop in San Juan, Puerto Rico in December 2008 to address the implementation of the DP-LTER, focusing on long-term data sets, modeling activities, and cross-site studies that would promote the goals of the DP-LTER. Participants from all 26 LTER sites were invited to the workshop and 21 sites were represented (see List of Participants). Workshop participants met for two days following a half day of seminars given by the Steering Committee and a field trip to the San Juan SES.

The DP-LTER lays out a comprehensive theoretical framework leading to a series of questions linking the human and biophysical components of SESs. Yet, one recent publication suggested that ongoing SES research exhibits such a large diversity of approaches that SES research currently amounts to nothing more than a series of case studies. Given the plans for substantial growth of LTER-SES research over the next decade, there was a perceived need to examine the way biophysical and social scientists sample, analyze, and model SESs and to begin to define the key elements of a unified approach to studying them. This was the goal of the workshop. We focused on the first of the three grand challenges identified in the DP-LTER, land and water use change, with an emphasis given to land use change. While this is a less important issue in marine and remote sites, it was a good starting point to begin to address the issue of implementation.
Prior to the workshop, the Steering Committee decided to organize the meeting around (1) Sampling and IM Strategies; (2) Modeling SESs; and (3) Cross-site Comparative Long-Term Socio-Ecological Research. Workshop participants were then asked to develop three white papers on these topics, which follow this summary. The key findings and recommendations of these white papers are outlined below.

Sampling and IM Strategies. Workshop participants concluded that no single sampling design for collecting long-term socio-ecological data is feasible or is recommendable. While the distinction between “extensive” sampling (measuring pattern and inferring process) and “intensive” sampling (quantification of processes and mechanism) is a useful dichotomy, the when, where, and why of the sampling must be dictated by research questions developed by the site. Thus, the white paper envisions an SES Research Platform that balances local research questions while providing documentation and spatio-temporal referencing to allow the incorporation of results into models and to promote easy comparison with other sites. Specific recommendations include:

Funding should be provided to the LNO to develop an SES data catalog and metadata standards for the compiling of historical and newly collected SES data on the LNO website. This funding should include training of undergraduates, graduate students, staff, and faculty in information management for these purposes.

At the site level, researchers should immediately begin defining the spatial and temporal dimension of their current SES research program. This should include
identifying and developing long-term relationships with key stakeholder groups in their SES.

LTER sites have considerable experience in managing ecological data and there is no inherent reason why existing data management systems cannot be expanded to include SES data. Funding should be provided to sites to augment faculty and IM resources to address fully their SES research and information needs.

As a first step, funding should be provided so that each site can begin to compile historic land cover maps and socioeconomic data for their sites. REU funds and undergraduate workshops would

provide one mechanism to achieve this goal, while instilling in a new generation of researchers the importance of SES research.

Modeling SESs. As with sampling design, the workshop participants found that, while the basic elements of a modeling framework could be identified, there was no single best way to model land use and water change at our LTER sites. Rather, there is a continuum of model types from strictly theoretical models (directed at identifying the emergent properties of systems) to empirical models (for examining site specific historical trends). Commonly used model structures for studying land use change include Cellular Automata models and Agent Based Simulations. Modeling examples from four LTER sites were used to illustrate different modeling approaches. It was emphasized that sites embarking on establishing a land use modeling approach carefully and collaboratively consider the format and scale, spatial extent, spatial and temporal resolution, consistency of data types, and the appropriate way to link data types (i.e. pixels from maps to parcel/household data from census). Specific recommendations include:

As a first step towards developing a collaborative modeling approach, the workshop participants suggested that each interested LTER site assemble two raster land cover maps from at least two points in time. This exercise would help illuminate many issues, such as data compatibility, strengths and weaknesses of existing data, and regional extent of SES research. Most importantly, it would initiate the process of network-level SES analysis and modeling.

If sites can then complement the map analyses with qualitative environmental histories, it would help identify potential mechanisms of change common to all of the sites, or highlight differences.

Interpreted in the context of present day socio-ecological challenges at each LTER site, these results would underscore the ability of this approach address publically-recognized problems and identify subsequent analyses.

Funding should be requested to support these activities, including a meeting of interested PIs and establishing resources in a single laboratory where map analysis and environmental data can be housed.

Note: A supplemental funding request was submitted to NSF to fund this activity after the workshop was completed.

Cross-site Comparative Long-Term Socio-Ecological Research. Cross-site activities are implicitly or explicitly addressed in the other two white papers as an inherent component of long-term SES research. This white paper directly addresses ways in which cross-site SES synthesis can be improved by recognizing different types of cross-site research, the lessons learned from effective cross-site research, and the social-technical nature of cross-site research. Finally, a preliminary effort was made to develop a typology of LTER sites, that is, to identify some inherent structuring of LTER sites that would provide a basis to substructure groups of sites for future cross-site research.

Cross-site SES research is most productively viewed as falling along four widely recognized continua or dimensions, inductive-deductive, generative-verificative , constructive-enumerative, and subjective-objective. Each address, in different ways, the issue of the value of bottom-up site-based research activities versus that of top-down, comprehensive research questions encompassing the entire LTER Network. In the end, the workshop participants suggested a middle-ground approach is best, marshalling site-based interest and resources to address collaboratively the issues raised by the DP-LTER. A long history of collaborative research by LTER scientists informs how this process works best. Key components of effective cross-site research include:

A project champion or champions to guide the effort.

A clear set of goals and collective agreement on what the final product(s) of the research will be.
Clarity of data sharing rules.
Regular face-to-face meetings.

Execution of cross-site research has occurred in several ways (framework provided by an individual person or site, a method, or a series of group meetings), but they all have these key components in common. Using the Grassland Data Integration (GDI) project as a working example, the white paper explores the way in which effective cross-site research involves the interaction of humans and information systems via feedbacks through an iterative process. The socio-technical features revealed by GDI provide an illustrative framework for future SES cross-site studies.

Finally, some first attempts to identify “nested hierarchies” of LTER sites called for in the DP-LTER as a first step to studying land and water use change did not reveal a single best approach to the problem. Regional clustering, examination of the history of cross-site publications, ordinations of environmental and socioeconomic variables, clustering by ecosystem services, and a nominal clustering using a simple matrix of socioeconomic and biome classification all showed some promise in identifying a coherent, attractive structuring of the sites.

As the LTER network has developed and evolved, it has recognized and embraced the notion of SES research. Moreover, the unique attributes of LTER sites and their existing data sets and IM capabilities can allow them to leverage the network’s experience in unique and transformative ways. While the challenges and resource limitations that currently impede the potential of SES research at LTER sites are also recognized, the participants of this workshop were very optimistic that they can be overcome. It is our hope that this collections of papers and recommendations will further that goal.

LIST OF PARTICIPANTS

Karen Baker
PAL
University of California-San Diego
Scripps Institution of Oceanography
2252 Sverdrup Hall
La Jolla, CA 92093-0218 USA
kbaker@ucsd.edu

Christopher G. Boone
CAP
School of Human Evolution & Social Change
PO Box 872402
Tempe, AZ 85287-2402
cgboone@asu.edu

Patrick Bourgeron
NWT
NSTAAR
University of Colorado
Campus Box 450
Boulder, CO 80302-0450 USA
Patrick.Bourgeron@colorado.edu

Nick Brokaw
LUQ
Institute for Tropical Ecosystem Studies
University of Puerto Rico
Rio Piedras, PR 00931
nvbrokaw@uprrp.edu

Dan Childers
CAP
Global Institute of Sustainability
Arizona State University
PO Box 875402
Tempe, AZ 85287 USA
Dan.Childers@asu.edu

Ralf Goericke
Scripps Institution of Oceanography, UCSD
9500 Gilman Drive
La Jolla CA, 92093
CCE
rgoericke@ucsd.edu

Ted L. Gragson
CWT
University of Georgia
Dept. of Anthropology
250 Baldwin Hall
Athens, GA 30602-1619 USA
tgragson@uga.edu

J. Morgan Grove
BES
705 Spear Street
USDA Forest Serice
South Burlington, VT 05403 USA
jmgrove@gmail.com

John A. Harrington, Jr.
KNZ
Department of Geography
254 E Seaton Hall
Manhattan, KS 66506-2904
jharrin@ksu.edu

Nicole Kaplan
SGS
LTER Information Management
Campus Delivery 1499
Colorado State University
Fort Collins, CO   80523-1499
nichole.kaplan@colostate.edu

Gary A. Kofinas
BNZ, ARC
301 Irving Building #1
Institute of Arctic Biology

University of Alaska Fairbanks
Faribanks, AK 99775-7000 USA
ffgpk@uaf.edu

Denise Lach
Department of Sociology
Oregon State University
Corvallis, OR 97331-3703
AND
Denise.Lach@oregonstate.edu

Berry Lyons
MCM
Byrd Polar Research Center
Ohio State University
1090 Carmack Rd, Scott Hall
Columbus, OH 43210-1002 USA
lyons.142@osu.edu

Eda Melendez
LUQ
Institute for Tropical Ecosystem Studies
University of Puerto Rico
Rio Piedras, PR 00931
edacorreo@yahoo.com

John C. Moore
SGS
NREL
Colorado State University
Campus Delivery 1499
Fort Collins , CO 80523-1499
jcmoore@nrel.colostate.edu

Esteban Muldavin
SEV
New Mexico Natural Heritage Program
Biology Department Castetter Hall 167
University of New Mexico
Albuquerque, NM 87131 USA
muldavin@sevilleta.unm.edu

Tischa MuÒoz-Erickson
LUQ
School of Sustainability
P.O. Box 872511

Arizona State University
Tempe, AZ 85287 USA
Tischa.Munoz-Erickson@asu.edu

Jeff Onsted
FCE
Department of Environmental Studies
Florida International University
University Park, ECS 347
Miami, FL 33199 USA
jonsted@fiu.edu

Robert Gilmore Pontius, Jr.
PIE
Clark University
School of Geography
950 Main Street
Worcester MA 01610-1477
rpontius@clarku.edu

Nathan Sayre
JRN
507 McCone Hall
University of California
Berkeley, CA 94720-4740
nsayre@berkeley.edu

Frederick N. Scatena
LUQ
Department of Earth & Environmental Science
University of Pennsylvania
240 S. 33rd Street
Hayden Hall
Philadelphia, PA 19104-6316
fns@sas.upenn.edu

Laura Schneider
HBR
Department of Geography
Rutgers, The State University of New Jersey
Piscataway, New Jersey 08854-8045

laschnei@rci.rutgers.edu

Emily Stanley
NTL
University of Wisconsin
Center for Limnology
680 N. Park St.
Madison, WI 53706 USA
ehstanley@wisc.edu

Stephen K. Swallow
VCR
Woodward Hall - 9 East Alumni Ave.
University of Rhode Island
Kingston, RI 02881-2020
swallow@uri.edu

Jonathon Thompson
HFR
Harvard Forest
324 North Main Street
Petersham, MA 01366
jthomps@fas.harvard.edu

Robert B. Waide
LNO, LUQ
Long Term Ecological Research - Network Office
Department of Biology
University of New Mexico
Albuquerque, NM 87131-1091 USA
rwaide@lternet.edu

Jess K. Zimmerman
LUQ
Institute for Tropical Ecosystem Studies
University of Puerto Rico
Rio Piedras, PR 00931
jkzimmerman@uprrp.edu

1. Sampling and IM Strategies for LTER Socio-Ecologic Studies

FREDERICK N. SCATENA, J. MORGAN GROVE, TISCHA MUNOZ-ERICKSON,
DAN CHILDERS, EMILY STANLEY,
AND EDA MELENDEZ

Introduction

The paper begins with an overview of the goals and initial steps individual LTER sites should consider when designing their sampling procedures and information management (IM) platforms for Socio-Ecologic Studies (SES). The historic and theoretical contexts for existing approaches to social ecological sampling and data prioritization and platforms are then discussed. This is followed by a discussion of IM and technological opportunities that sites are expected to address in the near future. We conclude with recommendations and action items for individual LTER sites and the LNO. While the overall focus of the paper is on SES studies involving land use and water use in the regional landscape of the site, the basic recommendations are considered to pertain to a wide range of SES.

Initial Steps

Developing a SES research program that is well integrated into the existing ecological research efforts will require most sites to *expand their original study* area into the diverse human land uses of the surrounding region. It will also require a *long term commitment* to SES data management that is standardized enough to facilitate cross-site comparisons and allows scaling through data aggregation and disaggregation and the incorporation of new questions and analytical capacities. Participants at the workshop recognized that different sites have different opportunities and constraints and that each site will ultimately be responsible to develop and maintain an SES program that meets their sites needs. However, it was also recognized that the initial steps most sites will be involved with include:

Defining the spatial and temporal dimensions of their SES research area and begin collecting basic demographic data for the region.
Developing environmental histories of their SES research each site and the archiving and digital conversion of historic land use maps and data.
Identifying and developing long-term relationships with stakeholders groups, including public and private Offices of Sustainability, public utilities, indigenous groups, and watershed and health related stakeholders
Development of a network level data catalog and SES meta-data standards that individual sites can use when developing their site level IM system.

Sampling strategies

The original LTER sites were chosen to understand ecological change and process. Therefore, the sampling strategies at most sites were structured around plot based research and monitoring. Historically, the LTER sites have been very good at collecting long-term core environmental measurements and in conducting specific and intensive sampling campaigns that are designed to address specific ecological questions. As the network has developed, there has been a recognized need to expand the focal area of research into adjacent human dominated ecosystems. Consequently, the sampling strategies are evolving into designs that use the original core sites as environmental controls to gauge the change that is occurring in the surrounding SES study area. Therefore a key challenge for LTER sites is to develop strategies that utilize and leverage existing LTER data.

Workshop participants envision that sites will continue both intensive and extensive sampling that is designed to detect social and environmental change and understand their feedbacks and inter-relationships. We define “extensive sampling” to be sampling where the subject matter and geographical coverage is diffuse or widespread. This sampling is typically designed to monitor change over time (Figure 1.1). In contrast, “intensive sampling” typically implies: (a) sampling in a particular area with a dense network of sampling points; (b) sampling on a restricted range of topics by probing on them very deeply with an intricate schedule of questions; and/or (c) sampling occurs over a relatively short time period and is not designed to be continuous. Intensive sampling is commonly used to validate extensive data and analysis and to generate more detailed, process-based and mechanistic understandings. In general, extensive sampling is considered more useful for measuring pattern and inferring process, while intensive sampling is more appropriate for the direct measurement and quantification of process and mechanism. However, the same phenomenon can be measured both extensively and intensively (Figures 1.1, 1.2) and many intensive sampling designs are transitioned into extensive samplings, or vise versa, over time. The types of data that LTER sites need to either collect or acquire from other sources are shown in Figure 1.1. The spatial and temporal range of non-census data s that are already available to LTER sites are shown in Figure 1.2.

Actual sampling will involve combinations of sampling along gradients and in grids and patches. Regardless of how the sample points are selected, a central requirement is that all LTER data must be temporally and spatially referenced and located by *both* the latitude and longitude of parcels *and* their physical addresses. Common gradients will include rural to urban land cover, ethnic composition, and levels of health, income, and pollution levels. Schools, hospitals, community centers, and stakeholders are expected to be involved in all steps of data collection, sharing, and analysis. To facilitate cross-site comparisons, workshop participants recommend that LTER sites use common meta-data conventions and a hierarchy data platform they can build upon to meet their own specific needs (see below).

Figure 1.1. Examples of socio-ecological data types organized by scale and intensity of analysis. Data types marked in green are data that sites must typically acquire, document, and archive. Data types marked in red are typically collected by the sites.

Figure 1.2. Example of non-census data sets with spatial reference to Baltimore City, 1800-2000. Figure developed for Baltimore LTER.

SES Data Platform

Workshop participants agreed that a successful and well integrated SES site and network level research program depends on effective Information Management (IM). Fortunately, data from the natural and social sciences all require similar spatial and temporal identifiers and can be managed in similar ways. Existing LTER sites also have considerable experience in collecting and managing a range of environmental and ecological data. Therefore, *if adequate resources are provided*, there is no inherent reason that existing LTER data management systems cannot be expanded to explicitly include and integrate SES data. In general each site will ultimately need to develop a SES data platform that can provide data for:

Testing and validating models, including what historic and existing models have been most accurate and useful;
Providing the context for future and problem specific studies; including hypothesis generation and testing and determining specific sample intensities
Facilitating network level to global level comparisons.

While each site will collect data to meet their particular needs, all sites will eventually need to develop the capacity to quantify environmental and social change over time. This core methodology and platform will involve sampling georeferenced grid points that can be aggregated by:

Environmental conditions and biophysical patch type
Parcel ownership and social patch type
Modern and historic land use

Core data will also be collected in the following data categories (Table 1.1):

Land cover and use change and the interactions of boundaries and patches
Environmental change (climate, hydrology, biogeochemical, etc)
Resource use (water, energy, biogeochemical, etc)
Demographic change (population size, composition, distribution, etc)
Financial change (employment by sector, GDP, purchasing power, etc)
Ecological and social responses to both natural and anthropogenic disturbances.

Modeling and Cross-Site comparisons

The LTER decadal plan provides core guiding questions for future cross-site SES research of broad scientific and societal relevance. Despite the fundamental importance of these overarching questions, it is clear from this workshop, and from LTER's history, that one size will not fit all. Therefore, it is critical to recognize that questions will be tailored to specific sites, new questions and data needs will emerge, and sets of sites and some data sets may not be relevant or available at all sites. As an extreme example, cross-site comparisons involving ILTER sites have revealed that SES data sets we
consider to be core are not always available in other countries (e.g., parcel data for China). Likewise, understanding dynamics of land use change is a less pressing question

Table 1.1: Basic data categories and core measurements needed for a LTER-SES data platform.

Data Category
Metrics
Sources
Spatial
Scale
Sampling
Frequency
Environmental inputs
Core Climate
(standard LTER core data)
Precip. Temp, Wind, radiation
NOAA
USGS
EPA
Point

Continuous
SES Climate
(health and pollution indices)
Body Heat Index, pollen, index)
NOAA,
EPA,
Site and region
Continuous
Pollution
Air, Water, Soil
Ambient levels,# of Violations
EPA
State
Muncipalities
Point
Continuous
Hydrology
Runoff and extractions
Per capita
Water use
& quality
USGS
Utilities
By sector
Continuous
Landscape Features
Natural Science Maps
Topographic
Geologic
Soils & landforms

Vegetation, NPP
Natural Hazards
Category by area,
DEM's
NRCS, USGS
Regional
Decadal
Social Science Maps
Roads & infrastructure
Housing, commercial
Zoning, easements
Tax & insurance
Hazards (fire maps)
Category by area, per capita
DEM's

Annual +
Utilities
Water, Energy, Waste
Health, Mass transit
Per capita use by location, time
Muncipalities
Parcel to regional
Daily +
Economic inputs and exports
Consumer Spending
Water and Energy
Health care
Durable and consumable
Employment

By sector
Imports and exports
By sector

Per capita use by location
Business
Commerce
Individual to regional
Daily +

Data Category
Metrics
Collaborators
Spatial
Scale

Sampling
Frequency

Demographic and Social Structures
Population
 Income, Occupation
 Education, Religion
 Family structure
by age, race, residency
location
Census

Block, track, zip code
Periodic;
10 yrs
Health
 Mortality and morbidity
 Infrastructure
by age, race,
location
Census
EPA, NIH
Block, track, zip code
Continuous
Periodic
Crime
 Type, violent, non-violent
by age, race,
location
Police
Justice

Daily to Annual
Education
 Levels & standard scores
by age, race,
location
Muncipalities
Census
Block, track, zip code
Annual
Government Jurisdiction
 Local, Municipal,
 Regional, National
Laws and
regulations
Varies

Individual to national

Civil Organizations
 Local, Municipal,
 Regional, National
 International
Watershed groups, NGO's,
Varies
Varies
Annual
Periodic
Recreational Use
 Private and public lands
Visits,
Permits, $ spent
Tourism, Education
Individual site specific
Daily +
History
 Political & social
 Disturbances
archival
Historical
Muncipalities

Archival

for marine and remote sites such as PAL, CCE, or ARC. These examples reflect the larger, ongoing challenge facing the LTER Network of balancing the need to tailor questions and methods to best fit the individual site with the opportunity to work across the network, which is facilitated by standardization of questions, methods, and units.

While cross-site SES research is a new frontier for the LTER network, there have been successful initial explorations and many LTER sites have already embraced SES work and have begun cross-site efforts. The decadal plan also provides guiding questions and a framework for interactions. However, we cannot anticipate the future utility of all of these data sets. We do know there is utility to collecting SES type data without being driven by a specific multi-site based question because the basic SES data will become a
platform and magnet for future studies. Nevertheless we can offer some basic recommendation derived from these initial efforts at LTER sites:

Expertise is developing within the LTER network in merging social and ecological perspectives at both site and cross-site scales. Sites beginning to incorporate social-science data into their existing IM framework should seek expertise from existing sites prior to establishing their SES data management platform.

In some cases, it may be more efficient for an individual or group to do the specialized data acquisition and processing for all sites. It is also anticipated that individual data managers will eventually be responsible for managing specialized cross-site data sets (e.g. Global Climate

Model output, specialized remote sensing data etc). These arrangements will benefit not only from economies of scale and specialization but also they can facilitate cross-site standardization and exchange.

Sharing expertise on *how* to acquire data may be just as useful as sharing information on *what* data is needed. It is thus recommended that the LNO network support a "clearinghouse" type web page where sites can exchange information on data sources and IM techniques.

Preparing for the Future

Individual LTER sites and the LTER network will encounter a number of future challenges and opportunities as they initiate or expand their SES programs. As site-based LTER programs regionalize to incorporate social science and societal relevant activities, the spatial dimensions of their research coverage will grow. In many cases, this will involve regionalizing an LTER site's conceptual view of itself to include the nearest urban areas and human-dominated ecosystems. It is likely that these urban areas will, themselves, be growing. To that end, LTER programs must also be prepared to regularly expand these spatial boundaries in response to the dynamics of their associated urban areas. Likewise, there are some LTER sites that may become multi-site collaboratives because of geographic proximity (e.g. HBR, HFR, and PIE).

Future LTER SES research should also be prepared to take advantage of rapidly expanding sensor and "smart-home" technologies. The NSF is investing heavily in environmental sensor advances through programs such as the National Ecological Observatories Network (NEON). Companion opportunities will certainly be available for the social science research aspects of SES questions. One example is the "SmartHome" technology, in which parcel-level energy and water use will soon be available in real-time through data-streaming wireless and cellular broadcasts. Also, many urban areas are becoming increasingly video-equipped, largely for law enforcement or homeland security surveillance purposes. Access to these "human activity" data may be difficult initially, but will likely become more available—at the least as archived if not-real-time data. In summary, the future of LTER-SES research will almost certainly involve numerous types of sensor-based, real-time-streaming data technologies and data management and analysis systems.

Future data challenges will include the IM of existing and archived data as well. Large bodies of information currently exist in non-digital formats, including public records, legacies of municipal administrations and plans, and oral histories. In some cases, the long-term stability of these data may be in question (e.g. the random filing cabinet that gets discarded during office cleaning). Thus as LTER programs move towards integrated SES research programs, energy should be put into ensuring the long-term stability of these historic documents. This data stability should include digitizing these various forms of information, archiving them with relevant metadata, and organizing them in databases where they are readily accessible. It is also recommended that REU funds and undergraduate workshops be dedicated to archiving SES data.

Finally, the future of LTER SES research will likely include an expansion of the LTER Network itself through the upcoming Urban Long-Term Research Areas (ULTRA) initiative. The ULTRA Program, which is being jointly funded by the U.S. Forest Service and the National Science Foundation, is envisioned to include 7 – 9 new long-term research programs centered on cities. These

ULTRA programs will mimic the current LTER organizational model, but will be focused on fully integrated socio-ecological research. This research will be closely structured around the recommendations of the LTER Decadal Science Plan and the Integrating Science for Society and the Environment Report (ISSE). Existing LTER programs should position themselves to maximize their interactions with and involvement in these future SES-based ULTRA programs. The synergy of these interactions will enhance the ability of the current LTER Network to propel itself into fully integrated and societal-relevant SES research as the network expands to include multi-city urban SES research.

Recommendations

The LNO should initiate the development a SES data catalog and metadata standards following the concepts outlined in this paper (Table 1.1, Figure 1.1). This can be started by compiling existing information from each site and by:

- Maintaining a SES data catalog, metadata standards, and a library of SES survey instruments on the LNO web site.
- Sponsoring network level SES data management training for undergraduates, graduate students, staff, and faculty.

Each site should begin to:

- Define the spatial and temporal dimensions of their SES research area and be prepared to discuss them at the next ASM and at their mid-term reviews.
- Identify and develop long-term relationships with stakeholders groups, including public and private Offices of Sustainability, public utilities, watershed and environmental health related groups.

Each site should identify and/or hire *a scientist* and a *data manager* to:

- Assist in the development of LTER-based SES data standards and catalogs
- Work with existing network expertise to collect basic demographic data and land use maps for their site
- Integrate SES data collection into the site's existing data management system.

Each site should compile historic land cover maps and determine the hydrologic and environmental histories of the site.

- REU funds and undergraduate workshops should be dedicated to archiving SES data

2. A Common Denominator: Modeling Land and Water Use Change in LTER Sites.

LAURA SCHNEIDER, MERRYL ALBER, JOHN MOORE, JEFF ONSTED, ROBERT GILMORE PONTIUS JR., NATHAN SAYRE, AND JONATHAN THOMPSON.

Studying Socio-Ecological Systems (SES): Modeling Land and Water Dynamics

A central challenge for socio-ecological integration is framing research questions that require both social and ecological insights. Although there is recognition that a project can be enriched by the context provided by other disciplines, recent research has shown that socio-ecological patterns and processes are not evident when studied by social and natural scientists separately (Liu et al., 2007). Modeling land and water use change is an ideal nexus for socio-ecological research, with social scientists studying the economic, political, and cultural considerations that influence (or are influenced by) people's use of the land, and ecologists studying how patterns of land and water use are related to ecological variables such as nutrient retention. This type of perspective is particularly important for answering SES questions, which are socially and policy-relevant.

By focusing on land and water use, it is possible for social scientists to study how people make decisions about where they live, and ecologists to study how patterns of land use and cover are related to ecological variables such as nutrient retention. Continuing along this line of logic, in order to be truly SES research, however, it is important to cross this boundary to answer questions such as when changes in nutrient retention might feedback (perhaps through algal blooms and fish populations) to affect choices of why people live or recreate in an area—or about how zoning or decisions to clear wetlands might affect nutrient retention. Land and water use therefore becomes the point of intersection, reflecting drivers and responses in both an ecological and social context. The SES perspective explicitly captures this intersection and recognizes the feedbacks among components within each realm.

There is no single best way to model land use and water systems in LTER sites. Each site has unique challenges and opportunities dictated by site-specific circumstances. In order to do cross site comparisons across the LTER network it is important to think carefully and collaboratively concerning many aspects of format and scale, including causally linked system along analytical scales, spatial extent, spatial and temporal resolution, consistency in data types, and the appropriate way to link data types (i.e. pixels from maps to households/parcel data from census). Though it is not necessary that all of these be exactly the same across sites, greater similarity in data format would allow for more meaningful cross-site studies. A modeling framework to study SES questions might benefit from the following considerations proposed by O'Neill et al. (1986) for any new theory, model, or construct:

The model must be internally consistent.
The model must not be adopted simply because it was successful in another field.
The model must agree with known properties of the system(s) being studied.
The model must be capable of testing new and testable hypotheses.

As a starting point, we suggest that interested LTER sites develop and share maps of their greater study regions showing categories of landscape-use/landscape-cover (LULC) from two or more points in time. Of course, LULC patterns will differ across sites, as will the dominant agent of change (e.g. urban development, wildland fire, invasive species etc.). But, we believe that simply describing changing patterns over time can be a "common denominator approach" that all sites can quickly achieve. In addition to the LULC maps, sites could submit spatially co-registered maps of any independent variables that are believed to influence or be correlated with patterns of past landscape change. This relatively modest goal will "get the ball rolling" and go a long way toward understanding common drivers of landscape change and their social and ecological implications. In some cases, it will be appropriate to simulate the continuation of these changes into the future. In others, it will simply offer a better understanding of the current conditions. Undoubtedly, this exercise will illuminate common themes, foster future collaboration, and begin the long overdue process of centralizing spatial data from all sites.

In the sections that follow we discuss many of the challenges and suggest approaches that might overcome them. We follow with a series of case studies from different LTER sites. We end with a summary in the form of a qualitative assessment of the case studies that leads to a set of recommendations. Note that this discussion is primarily focused on land use modeling, as water use has not been as fully explored. However, we expect that a similar approach could apply to water use as well.

Role of Land Change Models: Diversity of Modeling Approaches

The objective of land change models is to disentangle the complexity of socio-economic and biophysical forces that drive the rate and spatial pattern of land use-cover change and for estimating the socio-ecological impacts of such dynamics. Models are methodological tools to analyze the causes and consequences of landscape dynamics and to support the exploration of future land-use changes under different scenario conditions.

There is a diverse array of approaches to land change modeling. The different approaches lead to remarkably different types of products and insights. A major axis of difference has theoretical models at one end of the spectrum and empirical models at the other. Theory drives the development of theoretical models, so they frequently concern hypothetical landscapes and relationships for which data might not exist. Their purpose is to explore the range of possibilities for how the future might unfold based on hypothesized complex interactions among many entities (e.g. scenarios). They attempt to understand the emergent properties of algorithms. Their uncertainties are difficult to quantify (Messina et al. 2008), since it is challenging, if not impossible, to calibrate and to validate such models. Empirically based models are at the other end of the spectrum. They are closely linked to measured data, so they can be used to examine the implications of historical trends for which data are available at specific sites. They can be used to see whether measurements support any quantifiable theory because the model is a calibration method, and methods of validation are available. Empirical models are not necessarily designed to examine the emergent properties of competing theories and do not necessarily capture complex feedback mechanisms. Castella and Verburg (2007) contrasted these two approaches and found advantages and disadvantages of each. In general, theoretical models tend to take a much longer time to develop, where as empirical models can produce results much more rapidly because they tend to rely on readily available data and to rely on common statistical methods (an exception to this rule are time series analysis).

Two examples of common model structures used to study land change are Cellular Automata (CA) and Agent Based Simulation (ABS). CA is a spatially explicit model where the outcome is the totality of many small and relatively simple micro-simulations. CA models have relatively simple rules; however, these rules, when initialized simultaneously amongst all entities, allow for tremendous complexity in outcome. Nonetheless, CA's simple rules allow for greater transparency and the understanding of results can be rendered probabilistic, rather than deterministic, by allowing for Monte Carlo simulation in its application or other methods of stochastic simulation. This approach yields a range of outcomes over multiple simulations for given conditions, rather than one authoritative prediction. When aggregated, these land use change outcomes can then be expressed geographically, for instance cells that develop more often than others under given conditions can be shown as darker than those areas that develop less often.

In an agent based model, the agent is "a real or abstract entity that is able to act on itself and on its environment; which can, in a multi-agent universe, communicate with other agents; and whose behavior is the result of its observations, its knowledge and its interactions with other agents" (Sanders et al. 1997, as excerpted from Verburg, et al. 2005). Agents are given certain motivations that affect their choices and are then set in an environment, usually a grid representing different land cover types. ABS's, when constructed effectively, can yield fascinating results in the form of emergent properties, i.e., higher level effects that cannot be observed or predicted from watching these agents in isolation.

Modeling examples from LTER

Hubbard Brook Experimental Forest: **Understanding land use legacies: The case of Grafton County, NH.** The objective of modeling land cover change in Grafton County is to determine the role of land-use legacies in the current landscape. The nature and strength of the relation between social and biophysical systems in this region have evolved and changed dramatically since the collapse of agricultural activities in late 1800s. Remarkably a detailed historical record can be exploited to research the relations of human environment systems. Using a series of maps characterizing land cleared for agriculture and forest starting in 1860, and through a logistic spatially explicit model, we quantify past patterns of change to help understand current land cover patterns. The relations between change and socio-ecological independent variables, such infrastructure, demography and forest type, provide the framework to create a business as usual scenario of change in the region for the next 50 years (Schneider et al., in preparation).

Modeling the impact of land-use legacies on vegetation at a regional scale through the use of spatially explicit data it is important because allows for the possibility to link the knowledge of environmental historians and the understanding of ecological place-based research. The results of the model are maps showing potential patterns of land cover considered as critical to estimate carbon storage in the region. The model developed scenarios of future change to estimate potential effect on ecosystem services, such carbon storage and fragmentation effects.

Plum Island Ecosystems (PIE). Most of the land change modeling in the PIE has been performed in order to demonstrate general principles concerning rigorous practices concerning empirical models. PIE investigators have followed an approach that is designed to measure the validation accuracy carefully. Therefore, the modeling approach insists on strict separation of calibration and validation information through time. Land cover maps from 1971, 1985, and 1999 are available, so the change between 1971 and 1985 is used to calibrate the model, then the model predicts the change between 1985 and 1999, at which point the predictive power of the model is measured (Schneider and Pontius, 2001). These strict rules concerning modeling practices lead to some severe challenges concerning data availability. Specifically, all of the independent variables must be from 1985 or earlier in order to maintain strict separation of calibration and validation information.

There are very few important variables for which historic data exist. For example, the available maps of roads and protected areas show the contemporary landscape, not the pre-1985 landscape, so it is strictly not legitimate to use them for calibration. But it seems counter-intuitive to make a prediction that does not consider roads and protected areas. This is where the mentality of strict separation of calibration, prediction and validation has introduced challenges. Nevertheless, PIE’s validation results are similar to those of other groups in terms of accuracy. Pontius et al. (2006) examined how the observed levels of accuracy influence the interpretation of the projections into the future. These levels of uncertainty are so large that PIE is now considering a different paradigm of modeling future scenarios. PIE’s proposed future approach will examine the historical data, interpret it in with respect to qualitative aspects of historical processes, and then design scenarios based on contemporary issues that stakeholders face, regardless of historic precedence.

H. J. Andrews: (AND). Two independent scenario planning projects have been completed in western Oregon within the region of the H. J. Andrews LTER. The first, known as the Willamette Futures Project (Hulse et al. 2002; Baker and Landers 2004), was born out of a state initiated program to foster collaboration between scientists and communities toward a scientifically informed vision of growth over the next fifty years. The process started with scientists mapping patterns of past landscape change since Euro-American settlement, which gave context to the current condition. Scientists then mapped the expected patterns of LULC change in each of three future scenarios, which all assumed a doubling of human population over 50-years. Patterns of growth were based on directives from citizen groups and the modeled transitions between categories were based on probabilities derived directly from citizen defined assumptions (Hulse et al. 2004). The future conditions were then evaluated in terms of their effect on water availability, stream condition, and wildlife response.

The other scenario planning project was the Coastal Landscape Analysis and Modeling Study (CLAMS; http://www.fsl.orst.edu/clams/intro.html). CLAMS focused on understanding the aggregate ecological and socio-economic consequences of different forest policies across multiple ownerships of the Coast Range of Oregon (Spies and Johnson 2007). Surveys and interviews of forest landowners were used to determine their expected management intentions under current forest policies. A landscape simulation model was then developed--expressly for the study area and research objectives--to evaluate changes in forest conditions (Bettinger 2001; Bettinger et al. 2005) given expectations of human population growth (Kline et al 2001). The model was then used assess impacts of proposed changes in forest policy on forest structure, wildlife habitat, and timber production (Thompson et al. 2006; Spies et al. 2007).

Potential Model Pilot for Florida Coastal Everglades (FCE). At the FCE site, there is a nascent project underway to address socio-ecological relations through empirical land use modeling. Data is currently being gathered in order to employ a cellular automata (CA) urban and land use change model called SLEUTH (http://www.ncgia.ucsb.edu/projects/gig/). Its name is an acronym which corresponds to the necessary inputs for the model's forecasts: Slope, Land use, Excluded areas, Transportation networks, and Hillshade. All of these geographic phenomena are considered to be highly relevant criteria for past urban growth as well as valuable for the prediction of future urban growth and spread. SLEUTH has been applied in several dozen urban areas across the United States as well as the world at large.

The FCE- LTER site presents an opportunity to utilize SLEUTH in a somewhat novel fashion. First, since the site offers virtually no slopes steep enough to inhibit development, the requisite Digital Elevation Models can then be either used for or replaced by simple elevation points, indicating those areas that are flood-prone and therefore resistant to development. This undertaking also presents an opportunity to use SLEUTH as a starting point for SES pilot modeling exercises. For instance, by creating additional layers, such as demographic characteristics, as well as attitudes and beliefs reflected in geocoded surveys, we can begin to refine SLEUTH's architecture to reflect greater human making decision complexity. This is widely considered to be an important aspect to valid model creation in Urban and Land Use change modeling (Agarwal, et. al., 2000). Second, the model's behavior can be affected by policies either already in place or being explored as future scenarios. The most obvious and justifiable example of such policy integration would involve the use of maps that reflect restrictions of land use (i.e., parks, conservation easements, etc.). However, more ambiguous policies involving incentives or nuanced land use regulation could also possibly be embedded in a cellular automata framework, as long as an accurate map of their currently allocated effects can be created. Zoning as well as planning commission governance, including such small scale phenomena as the willingness to grant land use variances on a property scaling all the way up towards amending a Master Plan, are currently being investigated at the FCE to understand their specific role in land use, land cover change. The satisfactory completion of such a study could result in a more refined suite of inputs into SLEUTH.

A Strategy for Land Use and Water Change Modeling in LTER

The theory of social-ecological systems holds that they are characterized by complex interactions among processes occurring on multiple scales, which can be expected to result in path dependencies, non-linear dynamics, and emergent properties. In view of this, we believe that model development is an on-going, iterative *process* that should incorporate both scientists and stakeholders, thereby contributing to the larger objectives of collaborative, problem-driven inquiry. Models can be used as *tools* to evaluate past change and to understand possible future scenarios. Moreover, model results should be communicated not as rigidly deterministic *predictions* of future outcomes, but as probabilistic *projections*[1]. Below we recommend a strategy for modeling that builds on existing models and datasets, and that capitalizes on the synergies and cross-site opportunities of the LTER network.

A first step is to assemble raster land cover maps from at least two points in time for all interested LTER sites. These data already exist for most sites, and can be subjected to relatively simple analyses involving measurement of change over time and interpretation relative to one or a few basic spatial variables (e.g., slope). This will serve several preliminary purposes: (1) to identify and resolve data compatibility issues across sites; (2) to assess the strengths and weaknesses of existing data; (3) to

---

[1] Maps (Goodchild et al., 1994) that reflect future land use and cover with a "fuzziness", rather than a sharply delineated clarity, that is commensurate with the uncertainty of the prediction as well as the data is one way to geographically express such projections.

identify gaps for further investigation; (4) to define provisionally the regional extent of each site; and (5) to initiate the process of network-level SES analysis and modeling.

In view of the unique temporal scale of LTER, the next step should be to complement the land cover maps with qualitative environmental histories for each LTER region. Understanding land use/land cover/water use change requires baseline information that should reach back in time as far as possible, both for use as possible reference conditions and in recognition of the importance of historical legacies—that is, past disturbance events that have persistent effects on observed conditions. Information should be assembled from existing sources and supplemented with new research based on archival sources, interviews, and oral histories. These findings will aid in identifying potential mechanisms of change to explain the patterns discerned from the comparison of land cover maps. Contacts made in the course of this effort, moreover, will provide initial outreach to various stakeholder groups (e.g., landowners, agency staff, long-time residents, recreationalists, etc.) for potential inclusion in the research program going forward.

The next step is to identify present-day challenges, issues, and threats facing the landscapes in the region of each LTER. These should include both recognized ecological issues (e.g., loss of species, nutrient loss or loading, land use change, habitat fragmentation) and issues raised by interviewees during the environmental history research. Proposed or potential factors that may affect the system (e.g., changes in zoning or regulation) should also be assembled. This effort will serve both to ensure that models address publicly-recognized problems and to identify scenarios for subsequent analysis.

These steps, taken together, will initiate an iterative process of incremental improvement of SES models for each LTER region/site. Improvements may be of several kinds: (1) identification and integration of additional data layers, with which to test hypothesized explanations of variation and change; (2) incorporation of land cover maps from additional points in time; (3) identification of the social-ecological processes that may explain observed change over time; (4) identification of additional data sets for understanding these processes, leading to (5) evaluation of additional scenarios

Synergies among sites will be achieved by periodic comparison of models from each site and analysis of similarities and differences in their scales and in the relative importance of different social-ecological processes for explaining change. This will also be an iterative process, intended to identify gaps for further research and testable hypotheses for explaining social ecological systems at regional and larger scales. Given the state of current knowledge about complex SES, identification of the most important processes and their scales will in itself represent a significant step forward. Cross-site comparisons could be promoted linking research groups from the different sites which deal specifically with modeling land use-cover.

Practical Steps

This section outlines specific steps to make progress on cross-site comparisons with respect to land cover modeling. A reasonable goal is to have some initial descriptive results by the end of summer 2010.

The first step is to identify the LTER sites that are willing and able to participate in this initial cross-site land cover modeling effort[2].. Each site that has GIS data of land cover categories from at least two points in time should be invited to participate in writing a proposal. It would be particularly helpful if each participating site were to have data also on water use, human population, elevation, slope, aspect, political units, and any additional variables that are important to the site's coupled human and natural system.

We will need continued funding to promote this effort; an initial meeting with a representative from each participating site will come with the site's data to discuss procedures for common formats. At the meeting, participants will perform the work to put the data into a common raster GIS platform. Important formatting issues concern a variety of aspects of scale, such as spatial extent, pixel resolution, and category hierarchy. The group must also address the role of metadata in documenting the accuracy of the maps. The group will be lead by the principal investigator who will have oversight and responsibility for the process for a duration equivalent to one month's salary or a one-semester course buyout.

Each participating site will also produce an approximately five page report that describes two issues. First is a historical description of the reasons for the most important land cover changes during the time interval of the land cover maps and how cover classes were determined. Second is a summary of the factors that are likely to influence future land changes. For each future factor, the report will note whether there is historical precedence for the anticipated factor. This information will eventually become important in subsequent phases of modeling because it will indicate whether a model that is calibrated with historical information is likely to be helpful in modeling the future with respect to the factor. For example, if the model will generate a future scenario of a continuation of historical growth of built area, then it would make sense to calibrate the model with historic information. However, if the future scenario were to introduce an entirely new factor, e.g. a new type of agricultural technology, then calibration with past data should be done with caution.

The initial analysis of the submitted data will characterize the land changes between the points in time for which data are available. There are interesting cross-site questions even for the cases where maps from only two points in time are available. The analysis will be designed to examine how various aspects of scale influence the observed land cover changes. We can use the analysis to examine how different processes function at different scales. For example, some macro processes can explain the patterns of land change when there are only two categories (e.g. built versus non-built) over long durations, but other explanations maybe needed when there are several more disaggregated categories, such as agriculture, forest, barren, wetland, etc. Pontius and Malizia (2004) have distilled some important principles that dictate how category aggregation influences measurement of land change. We must also examine the possible influence of error in the maps. Many maps have been created using the best available historical information, so accuracy assessment has neither been done nor ever will be done. This does not mean that we should ignore the possible errors in the maps. Pontius and Lippitt (2006) propose a method whereby we can measure whether suspected levels of error can explain the differences between maps from different points in time.

---

[2] We think that Theresa Valentine, Spatial Information Manager at AND LTER, has already started a list of the spatially-explicit data for different sites.

There are many types of interesting analyses that can be performed simply by characterizing the historical changes, without stepping into the realm of predictive modeling. Even for cases where there are maps from only two points in time, Pontius et al. (2004) offer a technique to detect systematic transitions among three or more categories. Alo and Pontius (2008) expand those methods to test whether the systematic processes at one site are different than the systematic processes at a different site. An obvious next step is to expand the methods to see whether the processes at one time interval continue into a subsequent time interval. The example below illustrates what can be done when there are maps from three points in time (thus two time intervals) and there is a single independent variable, which is slope in this case. Figure 2.1 shows an overlay of three maps of built versus non-built for 1951, 1971, and 1999 for Worcester and the surrounding towns in Central Massachusetts. We compare this map to an independent variable, slope (Chen and Pontius, in review). Figure 2.2 shows how the growth of built was concentrated on the flat slopes until 1971, but slope was not an important factor for the gain of built after 1971. So, if we were to calibrate a model with the relationship observed during pre-1971 in order to predict post-1971, then the model would have low predictive accuracy. This case illustrates the usefulness of strategic examination of the data, while not yet entering the realm of predictive modeling. Effective methods to visualize quantitatively and to explain qualitatively the historic changes should be an initial goal of our cross-site modeling effort.

These recommended initial steps have been informed by a two multi-year, multi-laboratory studies concerning cross-site comparisons of land change. The first is the Human Environment Regional Observatory (HERO) program, which compared four sites in Massachusetts, Pennsylvania, Kansas, and Arizona. We found that it was helpful simply to examine a two-map overlay of historic land cover categories from each site. The second is an effort connected to the international LUCC community to examine the validation of predictive land change models (Pontius et al. 2008). The comparison of 13 cases studied revealed that decisions concerning formatting of the data can frequently have a larger influence on the modeling results than the design of the simulation algorithm. Furthermore, many of the participants reported that they suspected that there were substantial errors in the data, so the output of the model contains some errors that are associated with the data and others that are associated with the model. It is difficult to disentangle these types of errors when metadata that describe the accuracy of the data do not exist. Therefore, much can be learned by simply thinking carefully about the design and examination of historical data, even before enter the realm of predictive modeling. Furthermore, if a model extrapolates past trends and is found to be erroneous, then it simply means that the process of the simulated land transformation was not stationary over time, but this is a result that we could see simply by a careful examination of the historic data. Though the challenges involved in attempting SES integrated modeling are daunting, we believe the first steps outlined above are modest enough for us to gain solid footing and keep our balance on this new ground but still ambitious enough to move us forward towards a greater understanding of social environmental systems across the LTER sites.

Figure 2.1. (a) time intervals of change in built based on an overlay of land cover maps of 1951, 1971, and 1999 on the left; (b) slope on the right.

Figure 2.2. (a) land cover changes areas in bins of slope showing that most land has flat slopes in upper figure; (b) land cover change intensities in bins of slope showing that newly built land was concentrated on flat slopes during pre-1971 intervals, but not during 1971-1999 in lower figure.

References Cited

Agarwal, C., G. L. Green, M. Grove, T. Evans, and C. Schweik (2000). "A Review and Assessment of Land-Use Change Models" *Dynamics of Space, Time, and Human Choice*. Bloomington, IN: Center for the Study of Institutions Population, and Environmental Change, Indiana University.

Alo, Clement and Robert Gilmore Pontius Jr. 2008. Identifying systematic land cover transitions using remote sensing and GIS: The fate of forests inside and outside protected areas of Southwestern Ghana. Environment and Planning B 35: 280-295.

Baker, J. Landers H. 2004. Alternative futures analysis for the Willamette River Basin, Oregon. Ecological Applications 14:311-312.

Bettinger, P. 2001. Challenges and opportunities for linking the modeling of forest vegetation dynamics with landscape planning models. Landscape and Urban Planning. 56:107-124.

Bettinger, P., M. Lennette, K.N. Johnson, and T.A. Spies. 2005. A hierarchical spatial framework for forest landscape planning. Ecological Modeling. 182: 25-48.

Castella, Jean-Christophe and Peter H. Verburg. 2007. Combination of process-oriented and pattern-oriented models of land-use change in mountain area of Vietnam. Ecological Modeling 202: 410-420.

Chen, Hao and Robert Gilmore Pontius Jr. in review. Diagnostic tools to evaluate a land change prediction along a gradient. Computers, Environment and Urban Systems.

Goodchild, M. F., Buttenfield, B., & Wood, J. (1994). Visualizing fuzzy maps. In H. M. Hearnshaw & D. J. Unwin (Eds.), *Visualization in geographical information systems* (pp. 141–149). Chichester, UK: Wiley & Sons.

Hulse, D., S. Gregory, and J. Baker. 2002. Willamette River Basin Planning Atlas: Trajectories of Environmental and Ecological Change. Oregon State University Press.

Kline, J. D., A. Moses, and R. J. Alig. 2001. Integrating urbanization into landscape-level ecological assessments. Ecosystems, 4:3-18.

Liu, J, Thomas Dietz, Stephen R. Carpenter, Marina Alberti, Carl Folke, Emilio Moran, Alice N. Pell, Peter Deadman, Timothy Kratz, Jane Lubchenco, Elinor Ostrom, Zhiyun Ouyang, William Provencher, Charles L. Redman, Stephen H. Schneider, William W. Taylor. 2007. Complexity of coupled human and natural systems. Science 317:1513-1516.

Messina, J.P., T.P. Evans, S.M. Manson, A.M. Shortridge, P.J. Deadman, and P.H. Verburg. 2008. Complex systems models and the management of error and uncertainty. Journal of Land Use Science 3: 11-25.

Pontius Jr, Robert Gilmore, Wideke Boersma, Jean-Christophe Castella, Keith Clarke, Ton de Nijs, Charles Dietzel, Zengqiang Duan, Eric Fotsing, Noah Goldstein, Kasper Kok, Eric Koomen, Christopher D. Lippitt, William McConnell, Alias Mohd Sood, Bryan Pijanowski, Snehal Pithadia, Sean Sweeney, Tran Ngoc Trung, A. Tom Veldkamp, and Peter H. Verburg. 2008. Comparing the input, output, and validation maps for several models of land change. Annals of Regional Science 42: 11-47.

Pontius Jr, Robert Gilmore and Christopher D Lippitt. 2006. Can error explain map differences over time? Cartography and Geographic Information Science 33(2): 159-171.

Pontius Jr, Robert Gilmore, Anna J Versluis and Nicholas R Malizia. 2006. Visualizing certainty of extrapolations from models of land change. Landscape Ecology 21(7): 1151-1166.

Pontius Jr, Robert Gilmore and Nicholas R Malizia. 2004. Effect of category aggregation on map comparison. Lecture Notes in Computer Science 3234: 251-268. in M J Egenhofer, C Freksa, and H J Miller (eds): GIScience2004.

Pontius Jr, Robert Gilmore, Emily Shusas and Menzie McEachern. 2004. Detecting important categorical land changes while accounting for persistence. Agriculture, Ecosystems & Environment 101: 251-268.

Schneider L.C. and R. G. Pontius Jr. 2001. Modeling land-use change in the Ipswich watershed, Massachusetts, USA. Agriculture, Ecosystems & Environment 85: 83-94.

Spies, T.A. and K.N. Johnson. 2007. Projecting Forest Policy and Management Effects across Ownerships in Coastal Oregon. Ecological Applications. 17:3-4.

Spies, T.A., K.N. Johnson, K.M. Burnett, J.L. Ohmann, B.C. McComb, G.H Reeves, P. Bettinger, J.D. Kline, and B. Garber-Yonts. 2007. Cumulative Ecological and Socioeconomic Effects of Forest Policies in Coastal Oregon. Ecological Applications. 17:5-17.

Thompson J.R., K.N. Johnson, M. Lennette, T. Spies, P. Bettenger. 2006. Historical disturbance regimes as a reference for forest policy in a multiowner province: a simulation experiment. Canadian Journal of Forest Research 36: 401-417.

3. Cross-Site Comparative Long-Term Socio-Ecological Research

CHRIS BOONE, PATRICK BOURGERON, TED GRAGSON, JOHN HARRINGTON,
NICOLE KAPLAN, GARY KOFINAS, DENISE LACH, BERRY LYONS, STEPHEN
SWALLOW, BOB WAIDE,
AND JESS ZIMMERMAN

Rationale for Cross-Site Comparative Research

There are currently several LTER sites working together on cross-site socioecological science (SES) issues that could represent the beginnings of increased, long-term site socioecological activities. This effort should be fostered for a number of compelling reasons. Cross-site activities can take advantage of the heterogeneity - social, economic, political, ecological, geophysical, etc. - among and between the 26 LTER sites and create controlled experiments or leverage natural experiments that allow comparisons over multiple temporal and spatial scales. Cross-site socioecological research does carry the challenge of linking studies both geographically and across disciplines. This requires careful planning to generate questions and experiments, and develop research designs so that empirical testing results in generalized understanding of socioecological dynamics that goes beyond current theoretical ideas (i.e., Panarchy). One goal of these cross-site activities should be to develop stylized facts and rules of thumb that are common to all socioecological research, acknowledging that there is always tension between broad generalities and context-specific considerations.

Integration is not achieved by gluing parts together to assemble a data set. Rather, the focus and design of cross-site activities must be jointly defined by social and biophysical scientists. The questions and the problems must be compelling to all parties, and address the challenges presented when integrating diverse data types (or data from diverse sources). It is furthermore crucial to identify the range of issues and/or questions that require a transdisciplinary approach. This will help avoid the problems of parallel play and products that are nothing more than stapled interdisciplinary efforts (Miller 1982). Fully integrated interdisciplinary efforts will lead scientists from each discipline to identify new questions or approaches that were not apparent from the perspective of a single discipline.

Types of Cross-Site Research

Cross-site research can minimally consists of case study comparisons, the integrative testing of ideas/hypotheses using data from a hierarchical cluster of sites, and long-term or longitudinal assessments. Case studies carried out at each site could create an informative and detailed local understanding in a narrative format that analyzed across sites could identify common themes versus place-to-place differences. Integrative cross-site research is typically the product of individual or collaborative hypotheses that are assessed with data from multiple sites, e.g., LTER sites in desert locations could test the extent to which local water resource managers are aware of within-site cultural differences in environmentally-related water use knowledge. With long-term data on SES conditions and connections at a number of sites it becomes possible to address place-to-place differences in temporal variation and examine, for example, relative differences in coupled SES variation over time.

Comparison is an indispensable scientific technique -- no analytic statement about empirical observations can be made without at least one comparison providing the contrast that permits either inductive generalization or deductive proof. However, drawing a sharp distinction between quantitative and qualitative studies is artificial – the first type does not mean it derives from a discipline classified as a "science" any more than the second must result from one of the "humanities." It is more productive to consider research and in particular comparative, cross-site research, as a problem in data reduction and control of variation (Preissle & LeCompte 1981, Bollen, Entwisle & Alderson 1993). Figure 3.1 portrays the significant contrasts in research as relative characteristics of selected analytic strategies (broken lines indicate that the strategy varies on a dimension; solid lines indicate invariance). Each of the four dimensional characteristics of research design may be conceptualized as a continuum, since most studies may be placed in locations between the extremes.

Figure 3.1. Relative characteristics of selected analytic strategies (after Preissle & LeCompte 1981).

The **inductive-deductive dimension** refers to the place of theory in a research study. Purely inductive research would begin with collection of data, then builds theoretical categories and propositions from relationships discovered among the data. Purely deductive research begins with a theoretical system, develops operational definitions of the propositions and concepts of the theory, and matches them empirically to some body of data.

The **generative-verificative dimension** denotes the position of evidence within a study, as well as the generalizability attempted in the study. Verificative research verifies or tests propositions developed elsewhere, and also commonly attempts to generalize. The goal is to establish not only the extent to which a proposition is 'true', but also the universe of populations to which it applies. Generative research, in contrast, seeks to discover constructs and propositions using one or more data bases as the source of evidence. Generative research is often inductive, while verificative research is frequently deductive. Nevertheless, generative research may be informed by theory just as verificative research may have no theoretical framework.

The **constructive-enumerative dimension** of an investigation refers to the ways in which the units of analysis of a study are formulated and delineated. A constructive strategy seeks to derive analytic categories by a process of abstraction in which units of analysis are developed or discovered. The process of enumeration relies on previously derived or defined units of analysis that are then subjected to systematic counting procedures.

Finally, research designs may be characterized along a **subjective-objective continuum**. When patterns are viewed from the perspective of the object, entity or group under investigation then the appropriate strategies are those that elicit and analyze subjective data in order to reveal how the research subject conceptualizes their own experiences and world-view. In an objective approach, conceptual categories and explanatory relationships visible to external observers are applied by the investigator to the analysis of the object of inquiry.

Many of the questions and sub-questions in the ISSE report section on land and water (especially Questions 2, 3, 4, & 5) can be fashioned as comparative or integrative research efforts. For example, Q2b (the relative influence of human and biophysical drivers of change) can be studied at each site, but in the context of the LTER network there is an advantage to comparing findings across two or more sites or even across the spectrum of ecosystems in the network. This level of analysis offers the possibility of generalizing findings and even generating new theories about human-environment connections.

Another subset of questions for which cross-site studies are relevant are those that examine changing relationships between humans and landscapes over time. Q2a, for example, asks how historical trajectories of land change and ecosystem management constrain future trajectories. Using information from multiple LTER sites would allow us both to examine land change over time andassess the relative impact of ecosystem constraints on historical and current decisions. For example, rural interactions arise at local, regional and continental scales, often mitigated by policies and markets across these scales. Urban centers may drive demands for food and fiber or renewable energy production, while leaving the ecosystem impacts as remote, ignored afterthoughts. However, nearby urban centers might constrain management of ecosystems through, for example, policies designed to protect air quality that limit use of prescribed fire. This meso-scale, indirect effect then alters fire frequency and distribution in ecosystem processes, potentially affecting future fire risks or post-fire erosion after extreme events. Understanding alternative human responses to these interactions would be a unique area of research, with the topographic and human landscape context across sites revealing different principles and limitations on generalize-able knowledge.

Following are additional examples of cross-site Social-Ecological Science (SES) that falls within four (2-5) of the main themes for Land Use and Water Change in the ISSE report.

Q2: *By what mechanisms do humans directly and indirectly drive the dynamics of working and urban systems?* This question suggests examining land and water use in relation to the boom and bust cycles affecting the economy and motivating migration around various sites. Studies could explore the cyclical dynamics of land and water use change through economic cycles or cycles of ecosystem or climate change and their interactions. Mechanisms may include, but not be limited to, individual choices allocating land or water to various purposes from residential to cultivation to inactive management, and reach levels of collective choice. Cultural norms at different sites ranging from communities of indigenous to non-indigenous peoples could be examined to see how they may lead to different ethical or preference choices, may facilitate the adoption of alternative political-regulatory responses, or result in voluntary community actions. Social group heterogeneity may interact differently with ecosystem change in coastal, mountain, plains, marine or polar systems. Questions could also be developed at the continental scale regarding how near or distant urban centers, particularly in their demands for food and fiber, affect relatively remote ecosystems.

Q3: *What are the causes of human activities that are linked to working and urban system change and how does feedback from ecosystem change influence future causes?* This question suggests study of the processes of human mitigation and adaptation to climate change at regional and local scale, and how these efforts affect rate of ecosystem change and trajectories of SES change. Studies could be designed around heterogeneity in the ecological landscape and its interactions with heterogeneity in human preference and cultural structures to create different types of land-use and land cover changes that might include creation of different types of human communities. Working lands are often modified (in the past as well as the present) by the human pursuit of provisioning ecosystem services, including food, fiber and water. Such actions create spatial and temporal trajectories of land use and land cover types that may alter (enhance or degrade) production of these services under future conditions. However, as economic and social cycles develop, some ecosystems also attract human use in response to aesthetic or cultural ecosystem services. Furthermore, human behavior and choice may be constrained by land use legacies that influence biological systems through time. Cross-site studies may draw on how these factors play out in coastal, mountain, plains, marine or polar sites at different stages of transition to or from a human focus on provisioning ecosystem services or cultural-aesthetic services.

Q4: *How does urban and working system change influence ecosystem structure and function and the delivery of ecosystem services at local, regional and continental scales?* To answer this question, a model of land-use change that was generic in structure and function could be developed to project tradeoffs in land value as land use and land cover change. Such generalizeable land use models would enable multiple sites to stylized facts or rules-of-thumb about how socio-ecological systems change, while also identifying when context modifies or over-rides generalized (modeled) connections between human and biological systems. Application of a general model in coastal, mountain, plains or polar sites (and their regions) would help identify when social decisions need to recognize specific regional or local contexts. Models could be designed to incorporate interactions or feedbacks between changes in aesthetic qualities of ecosystems or in productivity for provisioning services, and how these changes interact with human behavior and choice to alter land and water allocations over time in different systems.

Q5: *How can public policy and private management decisions be informed by the knowledge of how human settlement and management affect ecosystem performance?* Cross-site studies could help address this question by identifying the interactions between individual and group responses to ecological and economic change in land cover. These responses may be defined within alternative local norms in relevant communities, from use of municipal or state regulatory processes to local norms for group action, or shared responsibility to create new economic or market institutions. Study could involve the responsiveness of groups to science-based information about ecological change through resource management, their ways of addressing issues of uncertainty in knowledge, and the extent to which principles of adaptive management can be implemented in various SES contexts. For example, what are the social and ecological factors affecting performance of SES governance?

Lessons Learned from Effective Cross-Site Research

Generalization and integration are characteristics of science, and a network of sites can assist with gathering and integrating local data for cross-site comparisons. Such comparative analyses facilitate the generation of new theory and the testing and refinement of existing theory. In addition, the synergy of colleagues working together across sites can lead to the cross-fertilization of existing ideas and the development of new questions. The value of cross-site, comparative socioecological research in the context, for example, of the LTER Network is clear. How, then, do we promote cross-site research. Are there lessons to learn from past experience for how to proceed with this type of research that go beyond "asking the right question" to "collaborating effectively"?

The first step is simply identifying long-term research projects that are case studies for the "lived experience" of collaborative research. Information derived from cross-site projects that have succeeded is an important step for identifying the lessons learned or best practices of cross-site, collaborative research. This can provide a source of ideas and inspiration to others for what conditions and processes helped advance collaboration versus those that impeded collaboration. There is value in identifying false starts and barriers, but ultimately a focus on the lessons learned can help avoid the mistakes of the past and advance the cause of ISSE.

Table 3.1 presents a list of multi-year, multi-site research projects, some within and some beyond the LTER Network. Some are explicitly socioecological and

Table 3.1: Multi-year, multi-site research projects thought by participants to be emblematic of successful cross-site collaborations. An asterisk indicates a project that was explicitly socioecological and designed, in principal, to be transdisciplinary. Names in italics are discussed in more detail elsewhere in the text.

Project Name	**PI/Leaders**
Ecosystem Services *	Chapin – BNZ / Kinzig – CAP
LIDET	Harmon – AND
Biodiversity & Ecosystem Function	Waide & Willig – LUQ
Grasslands Data Integration (GDI)	Cushing – Evergreen (Knapp/Kaplan – SGS)
Magnuson Variability Study	Magnuson – NTL / VandeCastle – LNO
Working Lands *	Swinton – KBS; recent start-up
Land Use Scenarios *	Foster – HFR; recent start-up
ILTER – Ecosystem Services and the Application of ISSE *	Bourgeron – NWT
EcoTrends - Socio economic *	Boone – CAP/Gragson – CWT/Grove – BES

EcoTrends – Biophysical	Peters – JOR
Center for Tropical Forest Science (CTFS)	Davies – Smithsonian/Harvard (Contact: Zimmerman – LUQ)
LINX	Mulholland – Oak Ridge
Agricultural Landscapes in Transition (AgTrans) *	Redman – CAP/Foster – HFR
Ison/Isof	Magnuson – NTL
NPP	Knapp – SGS
Lakes in the Landscape	Kratz – NTL
Land Fragmentation	Boone – CAP/Harrington – KNZ
Dissolved Organic Carbon	McKnight – FCE
Climate Variability	Goodwin – KNZ
Urban/Suburban Lawns	Grove – BES; recent start-up
HERO	Harrington – KNZ
Global Change	Harrington – KNZ

transdiciplinary in nature, while others are largely disciplinary or interdisciplinary collaborations within a single domain that in most cases is biological or ecological. This is not an exhaustive or representative list. It is simply a list recalled by the participants in the Puerto Rico Workshop (that was nevertheless significant). In reflecting on these projects, preconditions and conditions for success were identified.

Preconditions. It is clear that ultimate project success also depended in part on the existence of certain preconditions that facilitated establishment of a long-term collaborative research relation. These include the following:

Origins of idea
Have a productive / seductive brainstorming session formulating focus
Everyone resonates with the idea
Cross-cutting and compelling question(s)
Workshops can spawn projects
Concurrent grant activity / RFP
Available and comparable data
Agreement to a shared mindset, language, focus/problem, time scale, framework?
Leadership / champion / convener
Realistic objectives
Common relationship / experience of LTER PIs
Discussions and agreement/expectation on data sharing policy; system of data that allows for local control

Leadership. Beyond the preconditions that make possible the initiation of collaboration, there are characteristic activities that contributed to the success in meeting project goals. Of central importance are an effective leader and a project champion who are associated with a project for its duration, which do not have to be the same person. Some individuals are necessary to bring the group together and guide the group in an initial brainstorming and idea generation phase. Interdisciplinary research and interdisciplinary teams will benefit when the principal investigator/leader has prior experience working with these kinds of projects and can help the group develop communication bridges between group members who use different language and jargon, sometimes to express very similar thoughts.

The Champion. The champion, on the other hand, is the individual who is able to shepherd the project through to completion. In some cases, this has been a post-doc responsible for gathering data sets, making the data sets comparable, arranging teleconferences, initiating papers, developing workshops, or authoring related proposals. The essential feature of this individual is their ability to devote considerable time to the project, whereas an effective leader may be better used to motivate development of new projects. For instance, the AgTrans project had an effective leader whose skills were in great demand so that he lacked the time to actually champion the project, which was done by a post-doctoral researcher who collected and analyzed the long-term agricultural census data that was the foundation of the research. The combination of an effective and experienced leader with the post-doc's determination and dedication resulted in the successful completion of an edited volume, *Agrarian Landscapes in Transition: Comparisons of Long-Term Ecological & Cultural Change* (2008).

Clear Goals. Another key ingredient is to have a clear set of goals and a collective agreement among the members of a project as to what the final product(s) will be. For any collaborative research, agreement on goals and products is a critical first step, but for cross-site research, agreement up-front will help to allay future problems that may arise from questions about data sharing, joint publications, and distribution of resources. Effective cross-site research can and should begin with a clear set of rules. For example, the CTFS has been a successful cross-site research project spanning all continents. It is an effective network because it adopts a common methodology with a common goal to measure and monitor biodiversity of tree species in tropical forests. All plots have similar dimensions that are prescriptively measured using the same methods.

Managing Data. While prescriptive measurement contributes to the longevity of the CTFS network, the network could be improved by the formulation of clear data sharing rules. Some sites immediately publish their findings on web sites while others do not release their data until their scientists have had the opportunity to publish from it. In some cases, data are shared between investigators without other sites knowing that the data have been shared. The lack of data sharing rules can easily break down trust in a network, and trust is critical to the functioning of all cross-site research.

Allocating Resources. Clarity of data rules relates directly to another key feature of collaboration: the need for clarity from the beginning of a project as to what incentives for cooperation or participation will be. How fiscal, physical and personnel resources will be distributed needs to be clearly tied to effective member participation and meeting project goals. By outlining specific goals and *deadlines* for those goals, and explicitly stating the rewards or penalties to result from meeting or not meeting those obligations, cross-site research can avoid the difficult situation of supporting unproductive or uncooperative researchers to the end of the project. While unproductive researchers on a team are typically not invited to participate in future proposals and cross-site proposals, it more seriously undermines the spirit and purpose of cross-site or team research as it is happening. This can be avoided, in some cases, by carefully articulated research plans and clearly outlined consequences of meeting or not meeting project milestones.

Communication. Regular meetings, either face-to-face or by phone or video conference, move the project forward and also provide peer pressure on members to be prepared throughout the life of a project. Cross-site research can also benefit by being tied to a graduate or undergraduate course, where obligations are clear, meetings are regular, and performance quickly evaluated. This is another opportunity for the project champion to ensure follow-through with the research goals. For example, an effective champion might serve as a moderator for a threaded discussion board that keeps investigators engaged and produces an archive of ideas.

Face-to-Face Interaction. Despite some clear advantages of using distance technologies to keep participants connected in long-term collaborative projects, face-to-face meetings remain critical especially near the end of the project when the final products are being created. Beyond fulfilling the goals of the project itself, it is at the conclusion of one project where the opportunities to generate spin-off projects and keep the momentum going on professional relationships - these are hard earned investments that are extremely valuable for further collaborations. The most effective face-to-face meetings are those that occur in distant, neutral places away from the day-to-day responsibilities of work places. This is a proven formula enforced by the National Center for Ecological Analysis and Synthesis (NCEAS). Other NCEAS principles include sharing data, coming to meetings with data sets prepared for analysis, limiting group sizes to 15, mandating a diversity of participants that includes graduate students, and providing ample time for informal discussions. These principles could be effectively emulated in LTER cross-site research to good purpose.

Number of Sites. Closely related to the effective number of participants in a collaborative project is the question of how many sites are necessary for the research to be classified as cross-site. Simply prescribing the number of sites is artificial and forces relations that can easily failed for many reasons. It is much more effective to say the number of sites should depend on the number necessary to answer the question. For example, what is the acceptable power of a test and how many sites would it take to reach it? Alternatively, the sites to be included in a study could be determined on the basis of one or more typologies or site clusters (see below).

Strategy. Finally, cross-site research can be successfully executed in several ways. One approach is for a single individual or small group of individuals to collect data from multiple sites and then by themselves analyze those data. An example of this approach is the Magnuson Variability Study (see Table 3.1). Another approach is for a single individual or site to provide a framework that other sites use to perform their research either individually or in concert with other sites. The strict sampling method of the CTFS, which has been effective at reaching project goals, is an example of this approach. A third approach, which appears to be the most common, is for multiple sites to gather at the inception of an idea and to design the project through a series of negotiations and meetings. The initial number of sites might be reduced or augmented depending on the process of formulating the research question, searching for available data, and identifying key individuals and points of contact from each of the sites. This was the approach followed in the AgTrans project.

In summary, success in cross-site collaborative research not only depends on having a great question. It also depends on recognizing and establishing the basis for productive, long-term relations between the members of the research team. From the preceding issues, the conditions for success identified empirically from the lived experience in several projects include:

Issues of Onset:

A leader with a strong disciplinary/interdisciplinary/transdisciplinary perspective
A well articulated strategy or plan of attack
A vision of a final product(s)

Activities related to rewards
Having a core group with prior experience helps
Follow-through based on peer pressure and prior defined expectations is a must
Adequate labor at all levels (i.e., students, post docs, technicians, etc.)
Data organized, quality-controlled, and accessible to the point that meaningful analysis is possible while original data are archived unchanged, i.e., data management strategy in place.

Issues of Participation:
Are all sites needed? No.
Don't be prescriptive - there should be enough sites to answer the question with power
There is more than one way to execute cross-site research:

- X collects your data and does it
- Y gives framework and people work within it
- Group generates framework and does analysis collectively

Issues of Communication:
Generate the initial idea, then have members go away to think but not for too long
Identify a small number of papers that help ground the origin of the idea
Once a month over the life of a project people share readings to have shared scholarship
Face-to-face meetings are important, especially at the end
Empirical question: can cross-site research be conducted exclusively in a virtual setting?
The "distant neutral place" is important for achieving focus in a group
Schedule regular meetings – Don't let time pass
Have enforceable costs
Human relationships matter – talk about the Red Sox, not just about the research

These examples above include project-based studies and program-based studies. The ISSE framework is flexible enough to handle project-based studies that may use funds in traditional short or mid-term cycles, but also embed such studies in a long-term program of cross-site research on a broad, focal theme. ISSE is designed to foster cross-site studies of questions requiring long-term data drawn from sites or long-term experiments implemented about new questions at individual sites.

Cross-Site Collaborative Research as a Social-Technical Activity

Collaborative research is not merely a technical, but a socio-technical activity in which positive social relations induced by the proper mix of incentives and disincentives along with the necessary resources (fiscal, physical and personnel) are brought together productively over an extended period of time. The Grassland Data Integration (GDI) project (source: Nicole Kaplan) is presented as a case study of how social and technical aspects are merged productively in a research context. GDI is a joint effort by ecologists interested in annual aboveground net primary production (ANPP), and Long Term Ecological Research (LTER) information managers and computer scientists interested in data integration, semantics, and pragmatics (Cushing et al 2008).

The case study illustrates how humans and information systems interact to manage a legacy of data from a natural system and emphasizes the need to apply good stewardship (Baker & Chandler 2008), and communicate expectations and feedback through an iterative process as an integration system is designed. In general outline, the socio-technical features of GDI as a cross-site data integration project include:

Data integration is iterative – it is more than data compilation as it involves integrating the understanding data derived from individual cases, i.e., synthesis!

Reaching the objectives depends on participants using a common language to conceptualize the system; this includes both in developing the experimental design as well as the standardization of species codes.

Expectations about the analysis were clearly articulated: this was true both for the intentions of the ecologists as well as the system design for facilitating action; this translated into high-validity of the results of statistical analysis.

Technology was viewed as a service to the project, rather than the project being a service to technology. This approach leads to prioritizing local human needs: Does the system satisfy the local human's intentions? Is the system flexible and robust enough to answer cross-site research questions? Can technology be expanded to incorporate new data and answer new questions? Are there opportunities for feedback from the system and evaluation?

There were discussions leading to agreement and expectation about sharing data, codified into a Data Sharing Policy. The data system agreement allows for local control. In a similar vein, there were clear rewards for documenting data during project analysis and the drafting of manuscripts. When an investigator does not feel responsible they are less likely to carry through on their responsibilities.

In the GDI project, ecologists, information managers and computer scientists developed a conceptual design to integrate long-term ANPP data from across LTER sites in the grassland region. The database was developed to facilitate dynamic standardization and integration of data and metadata, and to perform queries and analysis of cross-site production at different scales. The project objectives were to make the integration process more efficient, enable cross-site analysis, conserve fine levels of data granularity, and eventually accommodate ANPP data from other grassland sites as well as sites outside the grassland biome.

During the project participants learned that in the past, the challenges to integrating cross-site data had been underestimated and thus much of the data that were to be incorporated into the database had been collected without standard protocols or metadata. The reliability and utility of an integrated data product depends on documenting the data as they are loaded, determining a statistically valid level of comparison, and transforming the data into a standardized format. However, maintaining a sustainable data warehouse over the long term is not feasible without semi-automated tools for data insertion, integration, documentation, and validation. Even more importantly, to undertake this project the information managers and computer scientists required the advice of ecologists familiar with the data to be incorporated. It was the collaboration of ecologists (responsible for experimental design and data analysis), information managers (accountable for data access), and computer scientists (responsible for producing technical solutions) at the onset of the project that led to the ultimate success.

Differences in data granularity (whether data were collected by species or growth habit) and experimental design had to be resolved in order to integrate site-level data. For example, at some sites ANPP is measured directly by harvesting total standing crop biomass whereas at other sites it is estimated from species-specific regression relationships between biomass and plant volume or coverage (McNaughton et al. 1996). In addition, ANPP data had been collected at different spatial (e.g., one-quarter square meter vs. hectare), temporal (e.g., seasonal or annual) and biological (e.g., species or life form) scales. Although such differences are common in biotic data, they were further complicated by differences in the language used by ecologists to describe data from the field, by information managers to describe data in the information system, and by computer scientists to describe technology applied.

It became important that scientists communicate the experimental design at each site along with how the designs were conceptualized as well as their intention to perform population or community dynamics analysis. The interaction of ecologists, computer scientists, and information managers was crucial to ensure that the analysis of the data were statistically valid and would support production of scientific papers. More specifically, the interaction led to the development of a common language to describe the experimental design between sites and that was captured in the design of the GDI database. The development, agreement and enactment of this language took several iterations with information managers and ecologists from each site, with the computers scientists. It was necessary for the language to be built into the data model to support successful insertions to and queries from the database, and has been essential to ensuring data are being analyzed on a comparable level.

While each site participating in the GDI project bases ANPP on field measurements, sites have idiosyncratic semantics for describing their experimental design and the experimental units where they collect data. For example, he smallest unit at which data were collected was termed the “experimental unit” whereas the unit appropriate for analysis to which experimental unit data were aggregated was termed the “sampling unit”. Developing a common semantics to reduce site-level idiosyncrasies in how research effort served to capture the socio-technical relationships between the data collected in the field, the database design, and requirements of ecologists to perform useful analyses.

The iterative database creation process allowed for feedback from the system to the ecologists and brought data errors to light. The finished database contained fewer errors than the source data as normalization highlighted errors of absence: data missing from certain plots or certain years and blank species or mass data. Integration highlighted errors of context: species with entries in the data but not the species tables, data from mislabeled or nonexistent plots, or plots with bad coding information. In addition, some basic validation checks removed observations with negative or zero weight, and observations for years outside the known span of the experiments. Questionable data were resolved by conversations with data providers.

The database allows comparison of ANPP between LTER sites and vegetation types, but the discipline of data integration and preliminary scientific analysis also led to improving data quality for subsequent analysis. For example, an early statistical comparison of three LTER sites (i.e., JRN, SGS and SEV) erroneously suggested that Jornada was significantly more productive than similar grassland sites despite the fact that it is the warmest and driest. As a consequence, the Jornada site updated its species' regressions and emphasized that the prevalence of a single species (i.e., *Yucca elata*) at one site can influence cross-site analysis. Finally, differences in site-level compliance with the data sharing policy impacted the robustness of the system. Ecologists required data from overlapping years to perform meaningful analysis, however, some sites were reluctant to provide data based on quality issues, concerns over data being re-used inappropriately, and sharing data prior to publishing the results themselves. Bringing the ecologists together to address these socio-technical concerns helped information managers and computer scientists create policies, functions and content in the system to satisfy each local need.

Interdisciplinary collaboration and teamwork during the data integration design process are keys to success. Ecologists, information managers, and computer scientists, all had to identify and address the socio-technical challenges to make the GDI a successful data integration project. It required time and attention to data quality issues by local information managers and ecologists during data transformation and integration, and a fairly deep understanding of how the data would eventually be analyzed. A team approach with defined roles and expectations fostered a partnership of ecologists, information managers and computer scientists that continues today in an attempt to now integrate data from other LTER sites. The GDI project has helped information managers become more aware of LTER synthesis projects, helped ecologists become aware of system requirements.

Typologies to Describe the LTER Network

The Decadal Plan calls for research on land- and water-use change to be “structured as nested hierarchies, with regional teams nested within network-wide teams” (p. 12). Intuitively, cross-site studies of long-term geospatial changes in land and water use are more feasibly developed among smaller groups of “similar” sites. But, how similarity is to be defined is not presently clear. Sites could be grouped by region (i.e., Northeast, Southeast, etc.). Alternatively, sites could be grouped by ecosystem type (e.g., forest, grassland, coastal, etc.). In both instances, particular sites fail to fall neatly in to clearly identifiable regions or ecosystem types. Following is a first attempt to use objective methods to develop typologies that sort the LTER sites into hierarchically structured groups.

Geographic, environmental, and historical factors offer the potential to facilitate interactions among LTER sites and between the ecological and social science disciplines. The structure of the interaction is expected to depend on the relative strength of ties between this set of factors as well as such drivers as the working relationships between individuals or the movement of individual researchers among institutions. The relevance of typologies of the 26 LTER sites based on geography, climate, history and socioeconomic conditions can be tested, for example, against research relationships among sites based on jointly authored publications. An ongoing investigation of social networking among LTER sites (Christian et al. in prep.) provides a description of these research relationships. In a similar vein, a clustering of sites based on research themes and biomes might also be useful.

The location of LTER sites (Figure 3.2) suggests that seven clusters of sites might be linked by geographic proximity, in this case defined as separation by only a few hours driving distance. One pair of sites (i.e., Bonanza Creek and Arctic) is separated by about six hours’ drive, but is included since access to the Arctic site by road is through Fairbanks, near Bonanza Creek. Geographic proximity facilitates scientific interaction because of existing local or regional connections among institutions, because of common research questions and management issues, and because of the relatively low cost of moving between sites. Moreover, geographic proximity is often linked to environmental similarity, but this is not always the case.

Of the seven sets of sites in close geographic proximity, five sets are also in the same climate domain as defined by the National Ecological Observatory Network (NEON; Figure 3.3). Two pairs of sites (i.e., Bonanza Creek-Arctic and Niwot Ridge-Short Grass Steppe) are in different climate domains. In addition, two sites (i.e., Florida Coastal Everglades-Luquillo) are in the same climate domain although separated by 1700 miles.

Figure 3.2. Grouping LTER sites by geographic proximity.

Figure 3.3. LTER sites placed within NEON climate domains.

Ordination of the 26 LTER sites (Figure 3.4) by nine environmental variables reveals several clusters of sites that are substantially removed from the main body of sites: tropical sites (FCE-LUQ), Antarctic sites (MCM-PAL), high latitude and high mountain sites (NWT-BNZ-ARC). Half the sites form a dense cluster around the mid-point of the only significant ordination axis, which correlates with temperature, length of the growing season, and potential evapotranspiration.

Although similar environments may stimulate comparison, more often than not ecologists examine trends along environmental gradients. Thus, a group of four sites arrayed from desert to short-grass steppe (CAP-JRN-SEV-SGS) have a strong history of interaction, although this history is also related to long-term personal and professional relationships among investigators. While there are a large number of other sites with similar environmental conditions, three in particular have been included in gradient studies involving the four sites mentioned above. These sites (KNZ, KBS, and CDR) are dominated by short-stature vegetation, as are the four sites along the desert-short-grass gradient.

Figure 3.4. An ordination (principle components analysis) of LTER sites using biophysical variables, mostly climate related. Labels on the axes of the ordination reflect the dominant influence of the original variables.

Ordination of 22 LTER sites by 6 socioeconomic variables appear to contrast urban versus rural sites with little useful clustering evident (22 rather than 26 sites were used because there is no socioeconomic data for the Antarctic sites, the blue water marine sites or the French Polynesia site, Figure 3.5). (Data: http://coweeta.ecology.uga.edu/ trends/catalog_trends_base2.php) The six variables analyzed were population, urban population, farm workers, service employees, population density for the year 2000 of counties where LTER sites are located and population growth for 1980-2000. A single significant axis distinguished areas of low vs. high population and was highly positively correlated with numbers of farm workers, service employees, and urban population. The second (non-significant) axis did not distinguish sites with high numbers of farm workers vs. high number of service employees or low vs. high urban population, as might be expected. Rather, it distinguished sites by high vs. low population growth (BES low) and population density (BES high).

Previous efforts at integration among LTER sites (Waide et al. 1999, Mittelbach et al. 2001) suggest that similarity in ecosystem structure and in the core questions driving research efforts also can stimulate comparison. Hence, in the LTER Network, comparisons among forested sites and among coastal sites would be expected. In addition, urban (CAP, BES) and augmented (NTL, CWT) sites have social science components that could facilitate synthesis among these sites. An analysis of social networking among LTER sites based on jointly-authored publications (Christian et al. in prep.) documents the development of cooperative projects within the LTER Network.

Figure 3.5. An ordination (principle components analysis) of LTER sites using socioeconomic variables, all population related. Labels on the axes of the ordination reflect the dominant influence of the original variables.

Figure 3.6 shows three points of this analysis. In 1983, five of the existing eight sites are involved in joint projects. In 1995, the pattern of relationships between sites suggests a dynamic situation in which partnerships between pairs of sites dictate the structure. The pairing of sites reflects geographic proximity, habitat similarity, and shared personnel. By 2005, a complex web of interactions exists, with a core of strong interactions among the older sites and an outer ring of more recent sites with still-developing connections.

In 2007, an LTER working group created a typology based on the most important ecosystem services provided by 26 sites (Figure 3.7). The patterns formed in this analysis are difficult to interpret and may reflect to some degree different values attributed to critical services at different sites. However, in general it is clear that freshwater provisioning and climate regulation are services that distinguish LTER sites from each other.

Figure 3.6. The history of the funding of LTER sites and diagrams of social interactions among LTER sites based on co-authored publications in 1986, 1995, and 2003.

Figure 3.7. Depiction of sites sharing 4 or 5 common critical services, with 'networks' of sites having all strong linkages (3, 4, or 5 common critical services) shown in similar colors. Ecological similarities among sites are indicated, along with the most common critical services within the network.

We attempted a nominal cluster analysis of sites (Figure 3.8) based on a simple matrix of dominant biomes, research themes, or conditions in the LTER sties and in the regions immediately surrounding each site. Three socioeconomic codings (urban, ex-urban, working systems) were combined with seven biome/habitat codings to produce a hierarchical classification of sites with three large clusters and several smaller clusters. Some of the small clusters appeared to make logical groupings (boreal/temperate forest sites, predominantly coastal sites, and grasslands) while others did not (e.g., KBS falls in with PAL, MCR, and CCE). Greater scrutiny of the variables used to code sites and a more careful consideration of site classification might resolve these problems.

Figure 3.8. A cluster analysis of LTER sites based on coding sites by biome and socioeconomic classifications. Assigned classifications are shown in the presence/absence matrix, with the clustering of sites shown on the left. The upper dendrogram shows how variables clustered in terms of their ability to distinguish sites.

Each of the approaches developed here has strengths and weaknesses and the results in many cases (e.g., ordinations, nominal clustering) depend strongly on the extent and accuracy of the underlying data. Interestingly, the history of the development of the LTER Network, more than might be expected at first, seems to be a more useful topology than other approaches, in part because it strongly reflects the strength of ongoing cross-site activity.

Organizational Considerations

The programmatic implementation of cross-site LTER studies raises a number of organizational questions. The ISSE document provides a common framework and at set of research questions to guide SES studies at the site and cross-site levels. To what extent if at all should these efforts be coordinated? Is cross-site research best achieved through a mostly bottom-up process of identifying research questions and organization of research groups? Is there a need for a level of cross-site study to which all sites contribute and in which they participate?

The best approach is a middle ground. This would include an institutionalized set of incentives that encourage and develop the capacity of sites to engage in SES research and participate on cross-site studies. Incentives can take the form of seed funding to initiate projects as well as meaningful acknowledgements at site reviews for special efforts in SES cross-site study efforts. Special funding for network-level SES synthesis is worthy of consideration as well. Including a social scientist as a member of review teams is one part of this process that currently happens in some cases, but not all. In addition there is a need for a formal committee to provide guidance, coordination, and leadership to the LTER community on this enterprise. The purpose of this group is not to direct SES research at or across sites, but to help in identifying common interests and building an SES community-of-practice as articulated in the Decadal Plan.

In the effort to build the capacity of LTER sites to engage in cross-site SES studies, it is important to recognize that sites are at different stages of development. Some only conduct site-level analysis, others are working at or moving toward regional-scale of analysis, and a small number have well developed SES study programs and are already engaged cross-site research projects. In building the capacity of sites for network-level SES research it would be helpful to explore how those with more experience can help those just exploring the integration of a social dimension in to their program. Without the peer support as well as the participation of key individuals such as NSF program officers the LTER Network may lead to ever-increasing disparity between sites engaged in SES studies relative to those that are not engaged.

References Cited

Baker, Karen S. and Cynthia L. Chandler. 2008. Enabling long-term oceanographic research: Changing data practices, information management strategies and informatics. Deep-Sea Research II 55: 2132– 2142.

Bollen, Kenneth A., Barbara Entwisle, and Arthur S. Alderson Source. 1993. Macrocomparative Research Methods. Annual Review of Sociology 19: 321-351.

McNaughton, S. J., D. G. Milchunas, and D. A. Frank. 1996. How can net primary productivity be measured in grazing ecosystems? Ecology 77: 974-977.

Miller, Raymond. 1982. Varieties of interdisciplinary approaches in the social sciences. Issues in integrative studies 1: 1-37.

Mittelbach, G.G. , C. S. Steiner, K. L. Gross, H. Reynolds, S. M. Scheiner, R. B. Waide, M. R. Willig, and S. I. Dodson. 2001. What is the observed relationship between species richness and productivity? Ecology 82: 2381-2396.

Preissle, Judith and Margaret D. LeCompte. 1981. Ethnographic Research and the Problem of Data Reduction. Anthropology & Education Quarterly 12: 51-70.

Waide, R. B., M. R. Willig, G. Mittelbach, C. Steiner, L. Gough, S. Dodson, G. P. Juday, and R. R. Parmenter. 1999. The Relationship Between Primary Productivity and Species Richness. Annual Review of Ecology and Systematics 30: 257-300.