Please submit your completed report in MS Word format to cbalaban@ufl.edu by October 18, 2013. *Don't spend a lot of time on formatting*. Each report will be reformatted to be incorporated into the final document, so there is no need for complex formatting. It would also be very helpful if you name your report **[University]-[PI last name].docx**. For instance: UF-Balaban.docx.

FESC FINAL PROJECT REPORT TEMPLATE **Due October 18, 2013**

Title Page: Project Title, PI Name, Institution, Partners, Project Time Period, Date

Executive Summary (Should be 2 to 8 pages in length. The Executive Summary should be a stand-alone document and completely describe the project and give the results. Items that should be included are accomplishments, benefits to the state and how funds were leveraged. The Executive Summary will be used as the project description and results for the FESC Annual Report. The last sentence of the Executive Summary will refer the reader to the full report on the FESC website which will be entered by FESC office.)

The full report format is given below

Table of Contents

Title: Database infrastructure for integrative carbon science research

PI: Sabine Grunwald **Co-PI's:** Timothy A. Martin

Students (name/degree sought): C.W. Ross (M.S. completed 2011); X. Xiong (Ph.D., completed summer 2013);

Staff members: Brandon Hoover and Risa Patarasuk

Summary

We developed the TerraC system (Terrestrial Carbon Information System) [http://terraC.ifas.ufl.edu] that synthesizes terrestrial carbon and environmental data into an online database structure and facilitates data sharing and synthesis analysis. This project helped

to leverage a $20 million NIFA-USDA funded project (PINEMAP: "Integrating Research, Education and Extension for Enhancing Southern Pine Climate Change") that has adopted TerraC as the centralized data infrastructure for the project.

The overall goal of this project was to develop a carbon data information system for the carbon science community focused on Florida and the southeastern U.S. The specific objectives were to develop a web-based database infrastructure and toolsets that integrate carbon and environmental data and facilitate collaboration and interactions among scientists and others interested in carbon and global climate change sciences. Terrestrial carbon is a key indicator in the global carbon cycle that interrelates with the global climate and other biogeochemical cycles. Energy consumption and fossil fuel emissions have impacted severely the carbon cycle and lead to warming effects in the atmosphere. It is critical to integrate data, information, and knowledge related to carbon and other environmental properties to enhance our knowledge and develop adaptation and mitigation strategies to counteract negative impacts in the future.

The TerraC system provides support to synthesize and integrate carbon and environmental data to assess impacts of global climate and land use change, investigate adaptability to natural and anthropogenic induced forcings and disturbances, vulnerability and risk assessment related to the terrestrial carbon balance, ecosystem services (e.g. carbon sequestration), and more.

Goals and Objectives

The overall goal of this project was to develop a carbon data information system for the carbon science community focused on Florida and the southeastern U.S. The specific objectives were to develop a web-based database infrastructure and toolsets that integrates carbon and environmental data and facilitates collaboration and interactions among scientists and others interested in carbon and global climate change sciences.

1. Motivation for Project

Rising CO_2 emissions in the atmosphere and effects on global climate change have been well documented, and future impacts are uncertain but potentially devastating. In the southeastern U.S., specifically Florida, natural and agro-forest ecosystems have much potential to sequester carbon in biomass and soils due to unique climatic and landscape conditions. However, research gaps exist to accurately assess terrestrial carbon content/density, carbon pools, and carbon fluxes at coarse scales, ranging from county to the region and larger. This project was motivated to address these obstacles by creating a database infrastructure for the carbon science community, focused on ecosystems in Florida and the southeastern United States. The TerraC system provides support to synthesize and integrate carbon and environmental data to assess impacts of global climate and land use change, investigate adaptability to natural and anthropogenic induced forcings and disturbances, vulnerability and risk assessment related to the terrestrial carbon balance.

2. TerraC Architecture

We developed a SQL-based (SQL = structured query language) web-accessible data infrastructure to compile carbon and environmental data into a centralized tool. TerraC manages user accounts at a variety of permission levels, provides tools to store, view, query, access, and download carbon data. A new .NET application was built to handle the interaction between the client and the database to accommodate fast streaming of large data files. With the growth and demand of modern datasets, this technology is a requisite for tools such as TerraC.

3. Description of the TerraC Information System

TerraC provides a data engine which allows managing, archiving, sharing, editing, modifying, and querying carbon and associated environmental data. These data are derived from various projects and sources; thus, provide a wide array of different carbon measurements, in various ecosystems and geographic regions, and spatial and temporal scales. The TerraC data engine facilitates synthesis and modeling to gain better insight into carbon cycling from micro, plot, field, watershed, basin, large region, and global scales.

Data Sharing and Usage Policy

Data users submitting data to or use data from the TerraC Information System agree to abide by the terms and conditions explained in this document. Data users may be held responsible for any misuse that is caused or encouraged by failure to abide by this agreement.

Definitions

Project: Set of one or more datasets that contain carbon (and related) environmental data.

Dataset: Set of data comprised of one or more data fields that contain carbon (and related) data that is part of a single project.

Roles of users

Project owner (leader): Principal Investigator or person with similar credentials responsible for collecting and managing the original, quality controlled data generated by a specific project. The project leader needs to initiate a project before a dataset can be submitted to TerraC and is responsible for the quality of all datasets under his/her projects. The project leader controls the levels of data sharing and can assign one or more data managers to each of his/her projects.

Data contributor (or manager): User that has read/write access to a dataset in TerraC. The data manager has privileges to submit a new dataset to a project and access and modify existing ones in part or as a whole. The project leader needs to assign a user manager status before he/she can submit a new dataset or modify an existing one in a project.

Data user: User that can view a dataset in TerraC. The data user can read public datasets and also private datasets as long as he/she has been granted access to them by the project leader. The data user cannot submit a new dataset or modify existing ones unless he/she receives manager status from the project leader to a project.

Data sharing: Data stored in TerraC can be shared at three access levels. The access levels are chosen by the project leader to control access to their projects by different users. Different access levels can be assigned to different users, the level being project- and user-specific. Levels 1 and 2 mirror the roles of data user and data manager, respectively. Level 3 in the most restricted access level.

Levels of data sharing:

Level 1 – Public with read-only access: Access to the data is open to all TerraC users. Any person that has a TerraC user account (i.e. data users) can view the data, but not modify it directly from the TerraC database. Only the project leader can modify/edit data.

Level 2 – Private read/write access: Access to the data is open to data managers who were assigned (approved) by the project leader to have permissions to view and modify/edit data directly from TerraC. *Private read/write access* is password-protected.

Level 3 – Private read-only access: Access to the data is restricted to the project leader and users selected by the project leader. Users can only view the data, but not modify it directly from TerraC.

The project leader controls the sharing of data in TerraC. He/she provides leadership for collaboration with new partners on behalf of the project teams. The project leader can switch sharing levels from Level 3 to 2 and 1, but not vice versa, meaning if the data are released to other users or the general public this right cannot be reversed.
Data users who are interested in to gain access to a specific protected dataset can contact the project leader and negotiate agreement of data use of a specific project. The project leader may agree to share data with the data user to collaborate on a joint project, work on a co-authored research publication, or use them for other purposes.

Data usage: Data users are expected to use data obtained from TerraC to the highest level of professional integrity and ethics. Data users must abide by the following guidelines when distributing or publishing data obtained from TerraC:

Data sharing and usage in TerraC is governed by the **Attribution Non-Commercial Share Alike** license provided by Creative Commons (summary: http://creativecommons.org/licenses/by-nc-sa/3.0; legal code: http://creativecommons.org/licenses/by-nc-sa/3.0/legalcode), which observes the following rules:

Attribution: The data user must give credit to the project leader (or project) in the manner specified by him/her (but not in any way that suggests that the project leader endorses the data user or his/her use of the data);

Noncommercial: The data user may not use TerraC data for commercial purposes; data should be used for reserach and non-profit applications;

Share Alike: If the data is modified in any manner or used to derive other products, the data user may distribute the resulting work only under the same or similar license to this one;

Credits and publications derived from TerraC usage:

- The data user must inform or consult the project leader about his/her intentions to use the data for publication well in advance of submission of the publication; the project leader should be given the opportunity to read the manuscript and, if appropriate, be offered co-authorship;

- The data user must give credit to the project leader (or project), which can be in the form of co-authorship, citation, or acknowledgement, according to the requirements imposed by the project leader; any deviation from this rule must be formally agreed between the data user and project leader;

- The data user must cite or acknowledge TerraC as the data host used to obtain the data;

- Any modification to the data originally obtained from TerraC by the data user must be fully documented.

Carbon Data and Associated Environmental Data

(1) Core Data Fields:

- Identification number for each observation (ID)
- Replication number (REPN)
- X coordinate (X) {Geographic Coordinate - longitude in decimal degrees, World Geographic System 1984, WGS 1984}
- Y coordinate (Y) {Geographic Coordinate - latitude in decimal degrees, World Geographic System 1984, WGS 1984}
- Sample date (DATE) {MM/DD/YYYY}
- Time (TIME) {HH:MM:SS}
- Height or depth of measurement (Z) {in cm; below the soil surface negative numbers; above the soil surface positive numbers}
- Carbon measurements (variable names, data values, and meta data: analytical methods & units of measurement in Standard International Units)

- Biogeochemical or other environmental data (variable names, data values, and meta data: analytical methods & units of measurement in Standard International Units)

(2) Project Elements (meta data):

- Project title
- Project description (description of sampling design, sampling protocol, quality assessment, data constraints such as below detection limit treatment, missing values, etc.)
- Project owner (typically Principal Investigator of a research project; or Project Leader for agency lead project)
- Project contributor (optional)
- Project user (optional)
- Contact information (Project Owner)
- Funding source
- Project location (description of geographic location of project; size of project area)
- Project period (YYYY to YYYY)
- Link to project homepage
- Publications from project
- Acknowledgements

Data Quality and Standards

Data format: TerraC focuses on terrestrial carbon and related environmental data. Data submitted to TerraC must contain carbon data and have the following format:

- Data organized n rows and columns, with cases (observations) listed in the rows and properties (attributes) listed in the columns;

- Carbon and other measured properties must be presented as variables in specific columns:
    - Each column must only contain properties measured using the same method; if the same property was measured using more than one method (e.g. total carbon vs. carbon fractions), each method must be presented as a separate column;

- Spatial coordinates (horizontal and vertical) and time stamps must be presented, whenever available, as variables in specific columns;

- Repeated measures (e.g., the same property collected at different times or replicated) must be treated as separate cases (i.e. listed in separate rows):
    - A column indicating that the cases are repeated measures of the same property must be included (e.g. using the same sample identifier for the repetitions);
    - A column indicating the number of the repetition (i.e. 1, 2, 3…) must be included;

- Quality assurance/quality control (QA/QC) data should not be included in the dataset, but instead in the metadata of the property it pertains to.

Data preparation: You can use any relational data base software of spreadsheet program to prepare your data for upload into Terra C. For example, Excel, MS Access, SQL or similar to organize your data listing cases (observations) in rows and properties in columns. A data template provided by TerraC helps to setup datasets ready for upload into the system.

Notes:

- Data values can be represented in form of strings, boolean (yes or no; 1 or 0), continuous (float), or discrete (integer) data
- Null values should be represented in the data as "0" or "0.0"
- To represent missing data leave fields empty (blank); do not use "N/A"
- To mark any special data values you may use codes "-999", "-9999" or similar (outside the data range for a specific measurement). The meaning of such special values should be included in the meta data
- Below detection limits (BDL) should be marked in data fields with a value (numeric data value) instead of listing "BDL" (string). The numeric data value can either represent the "true" BDL for a specific analytical method (e.g. 0.0005) or a designated arbitrary value (e.g. -99999). Mixing of different data types in one variable should be avoided (e.g. mixing of numeric and string values; or decimal values and text)

4. Datasets incorporated in TerraC Information System

Several large carbon and environmental datasets were integrated into TerraC covering Florida and the southeastern U.S. Environmental data are focused on climate and other relevant information, such as land use and vegetation, pertaining to the carbon cycle.

Soil Carbon Data – Florida

Soil carbon data from various projects, covering the Santa Fe River Watershed and the State of Florida, from different time periods (1965 to current) are included in TerraC.

PRISM (Parameter-elevation Regressions on Independent Slopes Model)

These data sets were created using the PRISM climate mapping system, developed by Dr. Christopher Daly, PRISM Climate Group director. PRISM is a unique knowledge-based system that uses point measurements of precipitation, temperature, and other climatic factors to produce continuous, digital grid estimates of monthly, yearly, and event-based climatic parameters. Continuously updated, this unique analytical tool incorporates point data, a digital elevation model, and expert knowledge of complex climatic extremes, including rain shadows, coastal effects, and temperature inversions. PRISM data sets are recognized world-wide as the highest- quality spatial climate data sets current l y available. PRISM is the USDA’s official clima-tological data. http://prism.oregonstate.edu
Variables representing 1970-2010 included in the PRISM dataset are: Precipitation (average

monthly), Minimum Temperature (average monthly), Maximum Temperature (average monthly), Dew Point Temperature (average monthly), Mean Temperature (average of Minimum and Maximum Temperature)(average monthly) , Vapor Pressure (average monthly) .

Idaho Geospatial

This gridded data set was developed by Dr. John Abatzoglou from the University of Idaho. These climate data combines spatial attributes of gridded climate data from PRISM with temporal attributes of a regional-scale and daily gauge -based precipitation. The gridded was based on observations from various weather stations including RAWS, AgriMet, AgWeatherNet and USHCN-2. The dataset is intended for users who require daily climate data to drive ecological or hydrological models as well as other applications. The original files came in a netCDF format. Then these netCDF were reformat ted and imported in ArcGIS.
http://inside.uidaho.edu/webapps/search/epscor_browse.aspx
Variables representing 1979-2011 included in the Idaho Geospatial dataset are: Precipitation (average monthly) , Precipitation(total monthly accumulations), Maximum Relative Humidity (average monthly) , Minimum Relative Humidity (average monthly), Specific Humidity (average monthly), Downwelling Short wave Radiation at Sur face (average monthly) , Wind Direction (average monthly), Minimum Temperature (average monthly), Maximum Temperature (average monthly) , Wind Speed (average monthly) .

NARCCAP (The North American Regional Climate Change Assessment Program)

NARCCAP is a joint international program that aims to produce climate change simulations in order to investigate uncertainties in regional scale projections of future climate and generate climate change scenarios. NARCCAP is a database that hosts climate change projections for North America. NACCAP dataset are generated by various GCMs (Global Climate Models) , from which, various climate change projections (scenarios) are derived. GCMs use grids of spatial resolution e.g., 300 km * 150 km grids. These GCMs are downscaled by various RCMs (Regional Climate Models) to spatial resolution of 50 * 50 km grids.
NARCCAP projections have been made for two 30-year time period using each GCM-RCM combination. These are: 1) Current time period of 1971-2000, usually known as the “baseline” projection, and is received by forcing the GCMs with historic CO2 emissions, till the year 2000. A perfect GCM-RCM combination should simulate a climate almost identical to the climate that was actually observed during the period of 1970-2000. 2) The future time period of 2041-2070 representing projections into the future under various assumptions for scenarios.
The NARCCAP climate variables are projected in a very high temporal resolution. For example, temperature, precipitation and surface pressure are all represented on a 3-hourly time scale. However, other variables such as minimum/maximum surface air temperature are represented on a daily scale (http://www. narccap.ucar.edu/data/data-tables.html).
The original data set comes in netCDF format. Ferret (from NOAA) and Cdat (from NCAR) were used to aggregate the data into the monthly scale. Some variables such as the monthly minimum temperature were derived from the NARCCAP daily data. Then these aggregate monthly scales were imported to ArcGIS.
Variables available in the NARCCAP dataset are: Sur face Air Temperature, Precipitation, Downwelling Shortwave Radiation at Surface, Surface Pressure, Specific Humidity, Minimum Surface Air Temperature, Maximum Surface Air Temperature. PINEMAP has derived the following: Number of Frost Days, Average Minimum Surface Air Temperature, Average Maximum Daily Sur face Air Temperature.

Tier 1 and 2 site-specific data from PINEMAP project have been included in TerraC. Those will be complemented by tier 3 site-specific data as soon as measurements become available.

5. **Building of a TerraC User Community and Visualization of Carbon Data**

TerraC has an intricate user structure (data/project owners, data contributors, and data users) [levels 1 to 3] that have different permissions to control/access data in the system. Users can come together in TerraC not only to share carbon and environmental data but also to collaborate. The system provides citation lists, where project owners can add citations to articles that have been published using the data in their project. The system also provides the ability to upload meta data and additional descriptions in form of PDF documents attached to each data table so that other users can fully understand project data. Graphing functions allow viewing and exploring data, which not only work in PC-based web browsers, but also on tablets and handheld portable devices.

Terra C is enhanced with a Google Earth apps that allows creating maps using global positioning system (GPS) points from client data to give a brief view of where the data is being collected. The idea is that users can use this tool to visualize their data and superimpose them onto other readily available aerial photographs and spatial maps.

1. **Funds leveraged/new partnerships created**

TerraC is now providing the data infrastructure for a $20 million integrated research, education, and extension project. This large-scale project funded by the United State Department of Agriculture (USDA) – National Institute of Food and Agriculture (NIFA) – Agriculture and Food Research Initiative (AFRI) Regional Project "PINEMAP: Integrating Research, Education and Extension for Enhancing Southern Pine Climate Change" (2011-2016) allows to populate TerraC and cross-fertilizes several research idea centered around carbon budgets and assessments, carbon change in dependence of global climate change and other stressors, and carbon sequestration and regulation as an ecosystem service. Many other similar synthesis projects are facilitated through TerraC-PINEMAP. Dr. Martin (co-PI of this FESC project) is the PI of the PINEMAP project and Dr. Grunwald (PI of this FESC project) is one of the 50+ co-PIs of the PINEMAP project. More information about PINEMAP can be found at: http://www.pinemap.org.

The PINEMAP project goals are to create, synthesize, and disseminate the necessary knowledge to enable southern forest landowners to:

- harness pine forest productivity to mitigate atmospheric carbon dioxide
- more efficiently utilize nitrogen and other fertilizer inputs
- adapt their forest management approaches to increase resilience in the face of changing climate.

PINEMAP has a multi-tier data structure representing different scales including:

- Tier 1 (historic measurements of tree response in dependence of treatments at about 700 locations across the southeastern U.S.)
- Tier 2 (new base measurements at hundreds of sites across the southeastern U.S.)

- Tier 3 (high-intensity measurements to capture water and carbon cycle at 4 sites)

The PINEMAP project has added almost 4GB of data which accounts for millions of rows for vegetation, climate, and atmospheric data. All of this data is being consumed by faculty, research staff, post-docs and students spread throughout the southeastern United States.

Concluding Remarks
All project goals and objectives of this project were met. We like to remark that this FESC-funded project was able to spawn a $20 million research, extension and education project, which is itself a big success. Carbon and climate related research is profoundly important and TerraC plays a major role in the southeastern U.S. to facilitate to make progress towards better understanding of the terrestrial carbon balance threatened by numerous human-made disturbances, including global climate and land use changes.

Patents
None

Publications

Peer-reviewed Publications:
Ross C.W., S. Grunwald, and D.B. Myers. 2013. Spatiotemporal modeling of soil carbon stocks across a subtropical region. Science of Total Environ. J. 461-462: 149-157.

Cao B., S. Grunwald and X. Xiong. 2012. Cross-regional digital soil carbon modeling in two contrasting soil-ecological regions in the U.S. *In* Minasny B**.,** B.P. Malone, and A.B. McBratney (eds.). CRC Press, Taylor and Francis, 2012. ISBN: 978-0-415-62155-7.

Xiong X., S. Grunwald, D.B. Myers, J. Kim, W.G. Harris and N.B. Comerford. 2012. Which soil, environmental and anthropogenic covariates for soil carbon models in Florida are needed? *In* Minasny B**.,** B.P. Malone, and A.B. McBratney (eds.). CRC Press, Taylor and Francis, 2012. ISBN: 978-0-415-62155-7.

Xiong X., S. Grunwald, D.B. Myers, J. Kim*, W.G. Harris and N.B. Comerford. 2012. Which soil, environmental and anthropogenic covariates for soil carbon models in Florida are needed? The 5th Global Workshop on Digital Soil Mapping 2012, Sydney, Australia, April 10-13, 2012.

Xiong X., S. Grunwald, D.B. Myers, J. Kim, W.G. Harris and N.B. Comerford. 20__. Optimal selection of predicting variables for soil carbon modeling in Florida, USA. Environ. Modeling and Software J. (in review).

Xiong X., S. Grunwald, D.B. Myers, J. Kim, W.G. Harris, N.B. Comerford, and N. Bliznyuk. 20__. Bayesian geostatistical modeling of soil organic carbon with uncertainty analysis across a highly heterogeneous landscape. Biogeoscience J. (in review)

Xiong X., S. Grunwald, D.B. Myers, W.G. Harris, and N.B. Comerford. 20__. Soil organic carbon stock change and its link to land use and land cover conversion and climate gradient. Science of Total Environ. J. (in review)

Patarasuk R., S. Grunwald, T.A. Martin and B. Hoover. 20__. Integrative modeling of tree response along geographic and ecological trajectories in the southeastern U.S. Ecological Modeling J. (in preparation).

Theses and Dissertations

Ross C.W. 2011. Spatiotemporal modeling of soil organic carbon across a subtropical region. M.S. thesis. University of Florida, Gainesville, FL.

Xiong X. 2013. Geospatial modeling of soil organic carbon and its uncertainty. Ph.D. dissertation. University of Florida, Gainesville, FL.

Presentations

Grunwald S. 2012. Soil carbon variability across large landscapes. Soil and Water Science Research Forum, Gainesville, FL, Sept. 7, 2012.

Grunwald S., B. Hoover, and R. Patarasuk. 2012. Terra C and Pinemap data resources. Webinar series Pinemap project. Gainesville, FL, July 13, 2012.

Grunwald S. 2011. Geospatial and spectral soil carbon modeling across large regions. NRCS, National Soil Survey Center (NSSC), Lincoln, NE, May 13, 2011.

Grunwald S., T. A. Martin, B. Hoover, G.M. Vasques, B. Zhong, and D.L. DePatie Jr. 2010. Terrestrial carbon (TerraC) information system. 2010 Florida Energy Systems Consortium (FESC) Summit, Orlando, FL, Sep. 27-29, 2010.

Grunwald S., T.A. Martin, G.M. Vasques and B. Hoover. 2009. Database infrastructure for integrative carbon science research. Florida Energy Systems Consortium Summit, Tampa, FL, Sept. 29-30, 2009.

Hoover B., S. Grunwald, T.A. Martin, G.M. Vasques, N.M. Knox, J. Kim, X. Xiong, P. Chaikaew, J. Adewopo, B. Cao and C.W. Ross. 2011. The Terrestrial Carbon (Terra C) Information System to facilitate carbon synthesis across heterogeneous landscapes No. 264-10. Symposia Spatial Predictions in Soils, Crops and Agro/Forest/Urban/Wetland Ecosystems, ASA-CSSA-SSSA Int. Meeting, San Antonio, TX, Oct. 16-19, 2011.

Hoover B., N.M. Knox, S. Grunwald, T.A. Martin, X. Xiong, P. Chaikaew, J. Kim, and B. Cao. 2011. Synthesis tools for carbon assessment in ecosystems. FESC Summit, University of Florida, Gainesville, FL, Sept 28-29, 2011.

Hoover B., G.M. Vasques, B. Zhong, S. Grunwald, T. A. Martin, and D.L. DePatieJr. 2010. The terrestrial carbon (TerraC) information system Vers. 1.0. 11th Annual Soil and Water Science Research Forum, Gainesville, FL, Sep. 10, 2010.

Xiong X., S. Grunwald, D.B. Myers, W.G. Harris, A. Stoppe and N.B. Comerford. 2011. Are soil carbon models transferable across distinct regions or scales in Florida? No. 262-8. Symposia Spatial Predictions in Soils, Crops and Agro/Forest/Urban/Wetland Ecosystems, ASA-CSSA-SSSA Int. Meeting, San Antonio, TX, Oct. 16-19, 2011.

Attachments

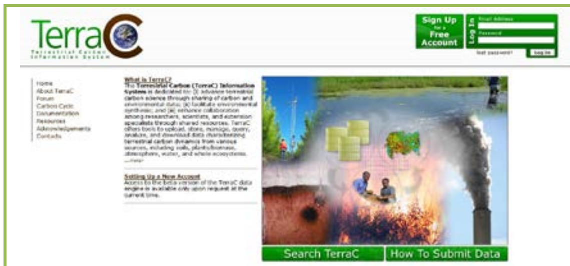

Fig. 1. Snapshot of the Terrestrial Carbon (TerraC) Information System website.

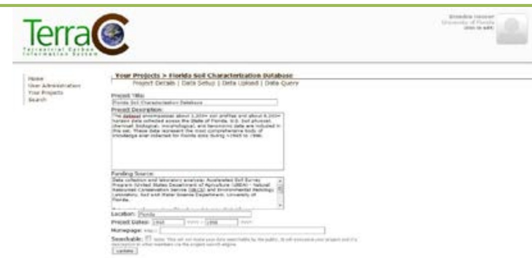

Fig. 2. Project setup in TerraC.

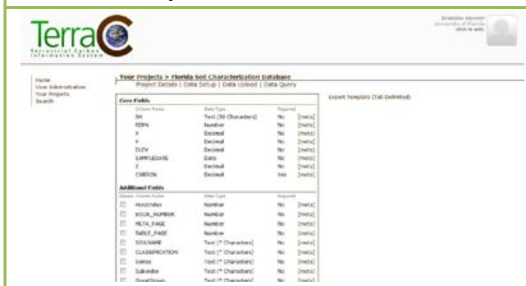

Fig. 3. Data setup in TerraC.

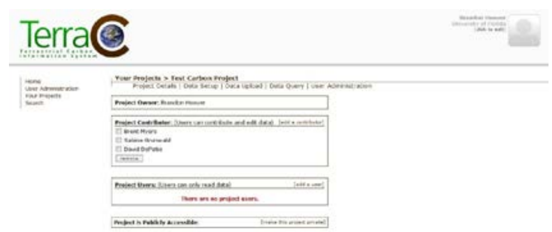

Fig. 4. User administration tools.

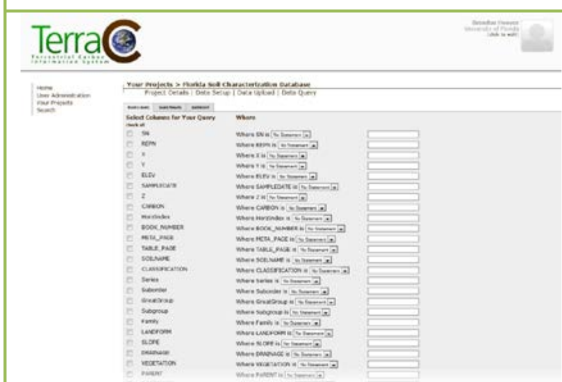

Fig. 5. Data query.

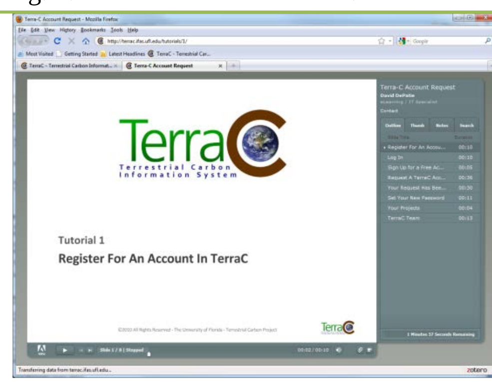

Fig. 6. TerraC tutorials.