TEAM Network: Building Web-based Data Access and Analysis Environments for Ecosystem Services

Choonhan Youn[a,*], Sandeep Chandra[a], Eric H. Fegraus[b], Kai Lin[a], Chaitan Baru[a]

[a]San Diego Supercomputer Center, University of California at San Diego, 9500 Gilman Drive La Jolla, CA 92093-0332, USA
[b]Conservation International, 2011 Crystal Drive, Suite 500, Arlington, VA 22202, USA

Abstract

A team of ecologists, computer scientists, and engineers from *Conservation International (CI)* and *San Diego Supercomputer Center at the University of California San Diego (SDSC, UCSD)* has been collaborating over the past 3 years to develop cyberinfrastructure for the *TEAM (Tropical Ecology, Assessment and Monitoring)* Network of tropical forest field sites. The TEAM project provides real-time data to understand how tropical forest ecosystems are being impacted by global climate change and land cover change and to improve conservation decisions. A major objective of this project is to provide information and services on tropical forest data disseminated by TEAM sites within countries participating in the TEAM network. The cyberinfrastructure provides a pervasive computational ecosystem, integrating grid computing infrastructure with high-performance backend resources, data warehouses, sophisticated client applications, new instruments, and embedded sensors, thus enabling a new paradigm for monitoring, understanding, and managing ecological and environmental systems.

This paper presents the TEAM data management, access, and analysis system that provides end-to-end solutions for sensor-based data and field observations collected at TEAM sites. Specifically, two major applications are presented, viz., the "*Data Query and Download Application (DQA)*" application, which allows users to navigate and download diverse TEAM datasets such as *Tree and Liana Biodiversity, Terrestrial Vertebrate, Climate and Forest Carbon data* using a Google Maps based interface; and, the "*Forest Carbon Calculator (FCC)*" application, which calculates tree biomass using equations for forests with different precipitation regimes, thereby predicting relationships between tree biomass, tree diameter and wood density to estimate the amount of above ground carbon in the forests.

Keywords: Tropical ecology; Forest carbon; Cyberinfrastructure; Ecosystem services; Ecoinformatics

1. Introduction

Science communities are increasingly becoming dependent upon cyberinfrastructure for their research and education. Some examples include the GEON grid [1], Polar grid [2], and SIDGrid [3]. The cyberinfrastructure provides advanced scientific computing and information processing services tailored to the needs, requirements, and

---

* Corresponding author. Tel.: +1-858-534-5071; fax: +1-858-534-5152.
E-mail address: cyoun@sdsc.edu.

challenges of applications in each science domain. These applications require increasingly complex data and information management strategies, which are being facilitated through these cyberinfrastructures. For observation networks such as TEAM, the data are gathered via a multitude of sensors and instruments, for which the data grid infrastructure is required to support appropriate services and capabilities.

A pervasive computational ecosystem integrates the grid infrastructure with high-performance backend resources, data warehouses, sophisticated client applications, new instruments, and embedded sensors [4][5]. This enables new computational paradigms for monitoring, understanding, and managing ecological and environmental systems. A team of ecologists, computer scientists, and engineers from Conservation International (CI) and the San Diego Supercomputer Center (SDSC), University of California San Diego are collaborating to develop the cyberinfrastructure to support the TEAM (Tropical Ecology, Assessment and Monitoring) Network. The TEAM Network [6] is an early warning system for climate change and biodiversity, akin to the USArray [7] system of broadband sensors for tracking seismic activity at a continental scale to warn of impending earthquakes. Data collected within the TEAM Network is required to be made publically accessible in as near real-time as possible. This open data policy, along with a comprehensive set of business rules and requirements, informs the system specifications.

Based on our experiences from other data management and access environments, we have developed a domain-specific data management and computing infrastructure that provides a unified, single-point of access and analysis environment for TEAM data. The system features basic search and data access capabilities to allow users to discover and access TEAM data using metadata-based search conditions (e.g., using time range, data type, and site information). It also provides an analysis environment for estimating aboveground tree biomass stocks and changes across a broad range of tropical forests, using TEAM vegetation data.

The TEAM Network software exploits recent trends in Web technologies, including the use of Web 2.0 technology to provide highly interactive interfaces to complex services invoked by end users or applications [8]. These Web 2.0 applications take advantage of Ajax-based tools to interact with distributed services [9], and with other implementation that use JavaScript with callback functions to remote servers over HTTP to provide rich interactivity. The Web 2.0-based application framework enables users to easily access and analyze the TEAM Network data by effectively integrating the TEAM scientific database with external data sources, and providing improved support for the analysis environment. The system is built using a *portlet*-based architecture for reusability and interoperability of each constituent portlet component. The ability to customize the user interface; aggregate content from different sources; and reuse pluggable user interface components, allows other researchers and portal developers to build larger composite applications from these portlets by reusing each of the portlet components as a service in different scenarios [10].

1.1. The TEAM network

The Tropical Ecology Assessment and Monitoring (TEAM) Network (www.teamnetwork.org) is a partnership among Conservation International, The Missouri Botanical Garden, The Wildlife Conservation Society and the Smithsonian Institution—with the cyberinfrastructure being developed by the San Diego Supercomputer Center—that seeks to monitor long term trends in the status of biodiversity and ecosystem services through a tropical forest monitoring networks. Data is collected at *TEAM Sites* using scientifically accepted standardized *monitoring protocols* for *Climate* [11], *Vegetation* [12], and *Terrestrial Vertebrate* [13], which facilitates comparisons among TEAM sites, regions and continents using metrics that quantify changes in climate, land use/land cover, biodiversity and carbon sequestration.

The monitoring protocols specify to some level of detail the actions, settings, constraints and procedures needed to ensure that the protocol is implemented in a consistent fashion both in space and time. To ensure consistency and high data quality, it is paramount that the protocols be well-specified and properly implemented, with the implementation details being tracked and managed over time. A well-designed protocol should be able to respond to the following (i) *why*: i.e., the scientific background, questions and rationale as to why it is being implemented; (ii) *how:* the field methodology of how it should be implemented, including the *sampling design*, which defines the appropriate spatial and temporal scales; (iii) *when:* the workflows describing the timing (when) and ordering of the activities that need to occur in order to successfully carry out a field campaign; (iv) *who:* the clearly defined roles and responsibilities of individuals, defining who will do what, and (v) the quality control and calibration measures

required at each step to ensure that the data are of high scientific quality.

To ensure that the protocols are properly implemented; that the data are properly collected according to these protocols; and, that the QA/QC processes are properly applied to the data, the TEAM cyberinfrastructure is a set of tools and services for overall network administration and operation; data collection based on enforcement of constraints specified in each of the TEAM protocol specifications; data curation (including QA/QC) and dissemination. Fig. 1 depicts all of these major functions of the TEAM cyberinfrastructure, including tools to support network operations, data ingestion and data management functions, and applications.
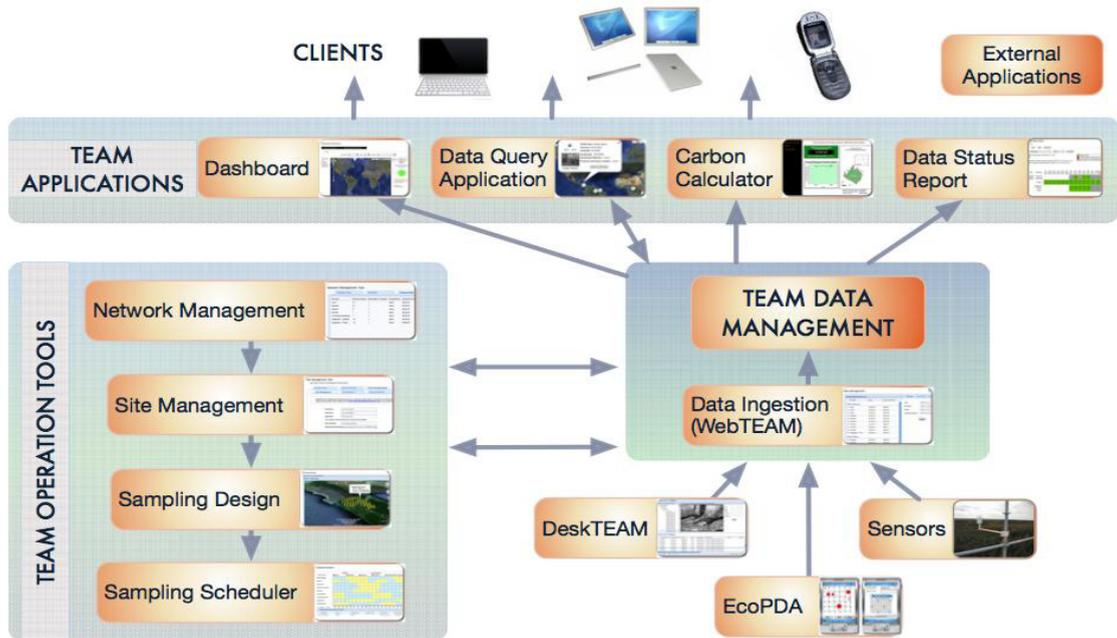

Fig. 1. A schematic overview of TEAM Network cyberinfrastructure

2. Related work

There is, indeed, a significant body of research focused on cyberinfrastructure support for ecosystems [14]. Two example projects that have designed a cyberinfrastructure to manage data, enable analysis, experimentation and collaboration for ecological and environmental data are the Digital Moorea and Knowledge Network for Biocomplexity. The Digital Moorea project has developed an integrated cyberinfrastructure to support the complete lifecycle of data products (from acquisition to publication) and the full range of scientific activities (including monitoring and analysis) for coral reefs on Moorea [15]. A coral reef ecosystem is instrumented with sensors that transmit data in real-time to a data server and viewer. The data server is a real-time streaming data engine that is linked to an archival database. This system supports a client application. For example, the data viewer operates on real-time data streams and the archived data. However, this system has been emphasized to build a sensor-based system for the data acquisition.

The Knowledge Network for Biocomplexity (KNB) project developed tools for data documentation, discovery, access, interpretation, integration, and analysis of diverse ecological data [16]. Specifically, KNB uses the standard format, EML (Ecological Metadata Language) for documentation of the ecological data. KNB tools include a data management system for ecologists, based on the Morpho client and Metacat server software. The Morpho client application is a desktop application and provides the capabilities to generate compliant metadata for ecological datasets. Metacat, is an XML database and repository for ecological data and generates EML compliant metadata [17][18]. However, the data collected in TEAM is much more diverse, and this approach is not easily adapted to

manage data in such environments. TEAM Network uses a relational database schema to describe its metadata.

3. Web-based architecture

The TEAM project has implemented an integrated platform for hierarchical data access using a "*data cart*" model for packaging various data types, forest carbon estimations, and visual graph presentations. The system has been implemented using a traditional three-tier architecture consisting of user interfaces in the front-end, services in the middleware, and resources in the back end. These tiers separate functionality and responsibility, allowing us to develop services independently as illustrated in Fig. 2 (a).

The front end supports two key functions; (1) provision of a single point of entry, which simplifies access to the applications and tools, and (2) enabling the development of applications and services in different backend environments. The underlying middleware provides an in-network data processing engine, which supports data dissemination, aggregation, collaboration and analysis in a dynamic, heterogeneous sensor-based network. In the back end, data from sensors, with measurements focused on a specific area are collected from the distributed sites deployed by TEAM Network. At the lab or field station, they are transferred to the TEAM scientific database and the central TEAM data repository, using the data management application [19].
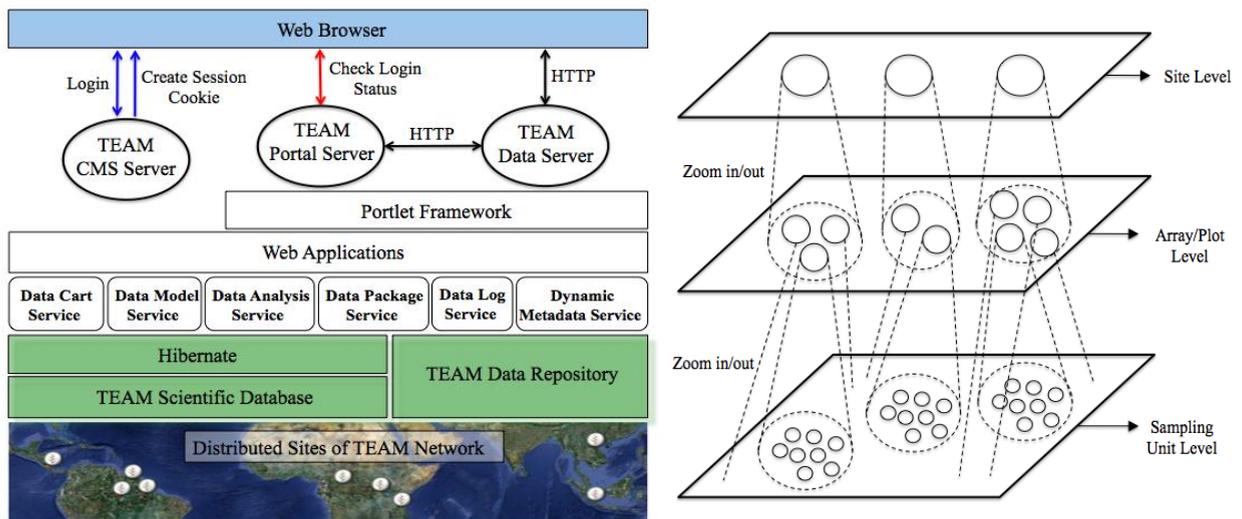

Fig. 2. (a) A schematic overview of data computing architecture and its services; (b) A hierarchical level structure

3.1. System architecture

The system infrastructure shown in Fig. 2 (a) is organized around three separate physical servers. First, a *TEAM CMS server* designed to tailor the information according to the user's role. It also provides content management services and the main point of entry for access to all of the applications and tools deployed by these servers. We use the *Drupal* content management framework to manage user interactions [20]. In the middle tier, the portal framework, the *TEAM Portal Server* provides the required resources for managing the various components. It also provides a flexible, comprehensive, collaborative user environment for scientists to share data, tools and research. In our implementation, we use the *Gridsphere* web application framework, which supports the portlet framework [21]. Lastly, the *TEAM Data Server*, which manages access to the *TEAM data repository* and uses the same portal framework as the TEAM Portal server. The TEAM Portal server uses the *Apache Http Components Client* library [22] in order to interact with the TEAM Data server, to run data services.

The front end provides virtually identical interfaces as the backend servers and consists of two key integration aspects. The first establishes *Drupal*-based layouts that are consistent and standardized on all pages. This is done using Drupal modules and themes. We have edited and modified the layout template of the default page on the

TEAM Portal server to correspond to the *Drupal*-based layout used by the TEAM CMS server. The second part configures multiple servers on one public IP address logically, and enables the single sign-on authentication allowing users to transparently login to different systems or applications. We have configured Apache HTTP server [23] to run multiple publicly accessible web servers on a single public IP address, using the Tomcat connector that handles the communication between Tomcat and Apache. Since the TEAM CMS server and TEAM Portal server each have different underlying security infrastructures, the ideal and most efficient authentication solution is to use a single sign-on authentication mechanism, in which a user has to authenticate only once and is authenticated to all of computing and data resources. The TEAM CMS server has been configured for the single sign-on authentication and provides the application's login page, performs the authentication using the user's login information, and creates the session cookie in the user's browser and saves the user's session ID in the database after login. When the user attempts to access an application deployed by the TEAM Portal server, the TEAM Portal server checks the session cookie sent by the web browser for user's authentication and then validates that this is the correct user's session ID and confirms a valid user by querying the user information on the database. If this user is presented with the login information, it indicates that this user has already successfully logged in. If authentication fails, a redirect to the application's login page is sent back to the browser. When a user or a proxy accesses the TEAM Data server, some services are restricted to the source IP addresses only for the authentication. Thus, the front end can make it more efficient to construct the dynamically changing physical environments.

3.2. Services

From Fig. 2 (a), based on consideration of the design principles, and the real world requirements and challenges in observing systems infrastructure, we have defined a set of data services to build web applications and tools. These include:

- *Data Model* service: TEAM scientific data along with corresponding metadata are stored in a relational database (PostgreSQL). This service represents the primitive data types, which map data from an object model to a relational data model (and vice versa) using Hibernate, which is an object/relational mapping solution for Java environments [24]. Hibernate provides data query and retrieval functions as well.
- *Data Package* service: When data products are selected and placed in the data cart the data package service is invoked. The general process is as follows:
  (1) The data entities for a selected data product are retrieved from the database and then saved into a data file in a user-selected format (e.g., CSV). Other data types, such as image data files associated with this data product are also included. In the case of data products with many images (e.g., Camera trap data) the image data is retrieved from the data repository and copied into the temporary directory for packaging. Camera trap images are organized hierarchically to facilitate analysis by the user.
  (2) A PDF document file that describes the selected data product is generated using the *Dynamic Metadata* service.
  (3) The data file, accompanying data objects (e.g., images), the metadata file and README files are created by zipping these together and storing in a temporary directory. The final data package filename serves as a unique identifier and is stored in the data log. Prior to distribution, the data package service evaluates the file size so that we determine the proper data package distribution method. The default consists of an immediate download. If the size is large (i.e., greater than 30 MB) the user is notified that they will receive an email when the data package is created. The data package is created on the backend server using Java threads and is stored on the TEAM Data server. Once complete and ready for download and an email that contains a download link to that data package is sent out to the user. Large data packages are stored on the TEAM Data server that is available for download for a restricted amount of time and eventually deleted from the server. Currently, access to this service running on the TEAM Data server is restricted to the TEAM Portal server to prevent anonymous access of this service.
- *Data Cart* service: The TEAM Network distributes many data products in a variety of data formats. Filters can be applied to data products to subset them spatially (e.g., by TEAM Sites or smaller logical blocks of data), time range and taxonomic information. The Data Cart service operates in a similar manner as a shopping cart in Amazon.com where data products are added, removed and eventually downloaded [25].
- *Dynamic Metadata* service: Data documentation by XML is an important information technology resource as it

can be used to build and organize metadata to describe resources and the raw data generated either by codes or by scientific instruments. It promotes the maximum reuse of software, services, information and knowledge. We have designed and implemented the protocol XSL formatting objects (XSL-FO [26]) metadata, which defines the information necessary to generate the file for protocol metadata using Apache FOP (Formatting Objects Processor). Apache FOP [27] is a print formatter driven by XSL-FO and renders the resulting pages to specified outputs such as PDF, PS and XML. We process XSL-FO generated from the XML protocol metadata driven by the database as the input source, using XSLT transformation. FOP transforms the XSL-FO from the XSL transformation to PDF file as the output form. For instance, the PDF document for each selected data type is created from the protocol XSL-FO metadata file generated from protocol XML metadata driven by the database dynamically using Apache FOP.

- *Data Log* service: This service records the downloaded data information captured on the *Data Query and Download Application (DQA)* application for the TEAM data analysis. It provides the TEAM data's usages statistics. For example, when a user downloads the Terrestrial Vertebrate data, this service registers the user's name and organization, intended use, a TEAM site name, time range, species, genus, and the status of photo data included.

- *Data Analysis* service: This service uses descriptive statistics that describe and summarize the collected raw data. Simple graphs and analyses make the data much easier to understand and interpret than raw data. For example, to easily visualize climate data, the raw data are aggregated and viewed in five categories such as dry temperature, precipitation, relative humidity, wind direction, and wind speed.

4. Implementation of user interfaces

We have developed two major applications that serve as the user interfaces. The "*Data Query and Download Application (DQA)*" application navigates and downloads diverse TEAM datasets such as *Tree and Liana Biodiversity, Terrestrial Vertebrate, Climate and Forest Carbon data* using a Google Maps based interface. The "*Forest Carbon Calculator (FCC)*" application calculates the tree biomass using equations for forests with different precipitation regimes, thereby predicting relationships between tree biomass, tree diameter and wood density to estimate the amount of above ground carbon in the forests.

4.1. Data Query and Download Application (DQA)

The design of the DQA user interface takes into account the many TEAM sites and the associated hierarchy of blocks, plots, and sampling units at each site. The DQA disseminates data using this same hierarchy. Thus, it is possible to view the data at each TEAM site by traversing down this hierarchy. The views consist of three levels namely, a site, array/plot, and sampling unit, as shown in Fig. 2 (b). Each site in the site level is decomposed into groups of array/plot, and each array/plot is broken down into groups of sampling units. Each level is responsible for only a finite amount of functions. Any function that cannot be done by a particular level gets delegated to a level more appropriate for handling the task. The site level displays the TEAM site information including the selectable data types for the data download, and viewing images and graphing data for analysis. Each TEAM site is broken into smaller more manageable pieces on the array/plot level, which provides focused region information including the same functions on the site level. For example, the Terrestrial Vertebrate protocol in a particular region of the TEAM site has a set of camera trap arrays, each of which has a set of camera traps. The sampling unit level displays the information for a basic unit used in collecting TEAM data and also provides the data graph function and image viewer for analysis. Each camera trap in the array has a unique ID and geographic coordinates associated with each camera trap and the images generated from them.

For efficient and easy navigation based on this structure, we employ a Google Maps based interface to allow users to browse and download various data products. We built the *Data Query and Download Application (DQA)* portlet in the TEAM Portal to provide a highly interactive user interface to access TEAM data, as shown in Fig. 3. This portlet consists of three main components: *navigation*, *search*, and *quick download*. In the *search* panel, users are able to search and discover TEAM data using specified metadata-based search parameters that include data product(s), TEAM site(s), temporal coverage, and other options such as *genus* and *species*, which are taxonomic terms. For search results, the *navigation* panel is organized in a hierarchical structure as described in the level

structure. The interface allows users to explore the various TEAM data hierarchically, add the selected data to a data cart, and request the selected data to be packaged for the download. In addition, pre-defined graphs and image viewers at all of the levels are available. The *quick download* panel supports only the capability to find the data sets that are available and save them into the data cart for the download quickly using the pre-specified search conditions, without navigating the data.

Fig. 3. A screenshot of the *DQA* portlet page

4.2. Forest Carbon Calculator (FCC) Application

Tropical forests hold large stores of carbon, yet uncertainty remains regarding their quantitative contribution to the global carbon cycle. One approach to quantifying carbon biomass stores consist in inferring changes from long-term forest inventory plots. A regression model involving wood density and stem diameter only is used to convert inventory data into an estimate of aboveground biomass (AGB) [28]. The purpose of this application is to produce Biomass and Carbon estimates for tropical forest vegetation plot data. We have designed this application using the TEAM Vegetation protocol and the vegetation data in the TEAM scientific database.

4.2.1. Technical details for carbon estimation from field measurements

Above ground carbon was estimated following the GOLD-GOFC guidelines for monitoring and measuring carbon [29] by using statistical relationships between tree biomass, tree diameter and wood density. Above ground carbon has been shown to vary between tropical forests as a function of climatic variables, in particular precipitation. The power of allometric relationships for predicting forest carbon increases when these relationships are fit separately for forests with different precipitation regimes (wet forests, moist forests, dry forests; see Ref. [28] for definitions).

The recommended equations for wet and moist forests are:

Wet Forests: $AGB = \rho \times \exp(-1.239 + 1.980 \ln(D) + 0.207(\ln(D))^2 - 0.0281(\ln(D))^3)$

Moist Forests: $AGB = \rho \times \exp(-1.499 + 2.148 \ln(D) + 0.207(\ln(D))^2 - 0.0281(\ln(D))^3)$

where: ρ = wood density (g/cm^3), D = tree diameter (cm), AGB = above ground biomass (kg)

Summing AGB (Above Ground Biomass) for all trees over the plot results in plot AGB (*plot_AGB*). Assuming that 50% of tree biomass is Carbon, convert to Tons of Carbon (*C*):

$$C = \frac{plot\_AGB}{2000}$$

Ideally, wood density values should be obtained for each species or genus. In the absence of genus or species-specific wood densities, average stand values can be used. The following average stand values of wood density (L) were used for the TEAM sites [29]: *Caxiuana and Manaus:* 0.69 g/cm^3, *Volcan Barva:* 0.60 g/cm^3

4.2.2. Results

The general process for the TEAM vegetation data consists of three components: (1) use all vegetation data stored in the TEAM scientific database as the vegetation data resource; (2) analyze them using allometric equations; and (3) visualize the results. TEAM sites and 1ha plots are initially selected to start the process. The finest level of the selection at a TEAM site is the 1ha plot. The highest level of the selection is all the 1ha plots at a TEAM site. In the user interface, after selecting a TEAM site, initial census and final census date are loaded and then selected. These selected parameters are used for getting the vegetation data from the TEAM scientific database. In order to predict aboveground tree biomass across a broad range of tropical forests, we utilize allometric equations based on tropical forest types described above. Fig. 4 shows results for the applied allometric equations: (1) calculate the total forest carbon in the protected area for the final census date, (2) plot out the biomass over time, for example, multiple years, (3) load the image of the protected area in the regional map, and (4) download results which specifies TEAM site name, plot number, the number of trees, the forest type, the initial date, AGB, and the forest
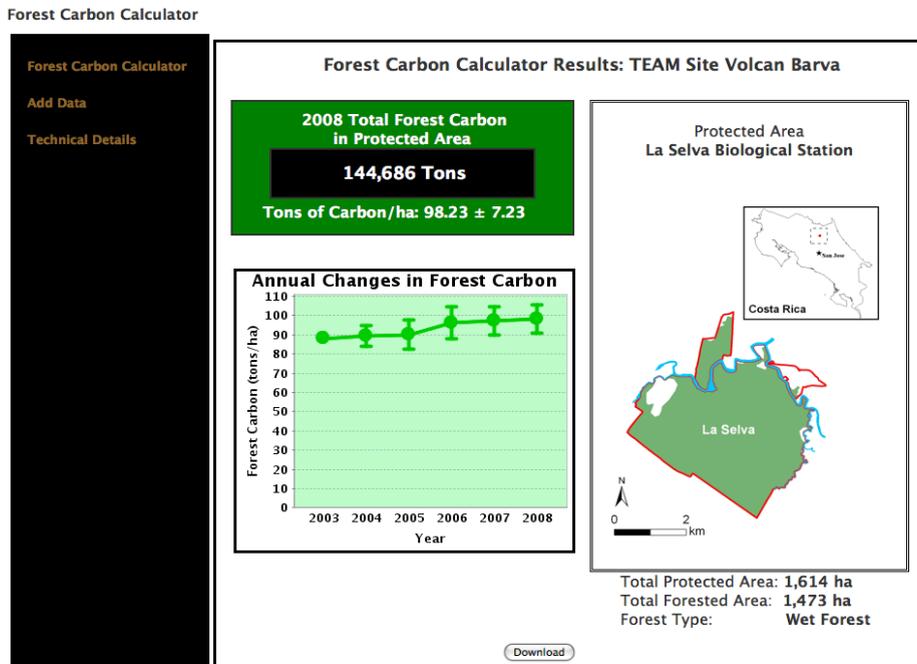
carbon.

Fig. 4. A screenshot of the result page in *FCC* application

5. Conclusion

This paper has presented a web based infrastructure for the TEAM network, which includes the development of two important applications, viz., the DQA application for integrated and simplified TEAM data distribution, and the Forest Carbon Calculator application. The TEAM Network is a global network with rapidly expanding data volumes, especially for camera trap image data. Future developments will include scalable support for managing such large datasets using, for example, Cloud Computing systems that promise infrastructure resources to support application scalability [30].

Acknowledgements

References

1. C. Youn, C. Baru, K. Bhatia, S. Chandra, K. Lin, A. Memon, G. Memon, D. Seber, GEONGrid portal: design and implementations. Concurrency and Computation: Practice and Experience. 19, 12, pp. 1597-1607, (2007).

2. Z. Guo, R. Singh, M. Pierce, Building the PolarGrid portal using web 2.0 and OpenSocial. In Proceedings of the 5th Grid Computing Environments (GCE) Workshop, Portland, Oregon, USA, 2009).

3. W. Wu, T. Uram, M. E. Papka, Web 2.0-based social informatics data grid. In Proceedings of the 5th Grid Computing Environments (GCE) Workshop, Portland, Oregon, 2009.

4. C. L. Borgman, J. C. Wallis, N. Enyedy, Building digital libraries for scientific data: An exploratory study of data practices in habitat ecology. In 10th European Conference on Digital Libraries, Alicante, Spain, Springer , 2006.

5. G. Briscoe, A. Marinos, Digital Ecosystems in the Clouds: Towards Community Cloud Computing. In Proceedings of IEEE International Conference on Digital Ecosystems and Technologies (IEEE-DEST'09), pp. 103-108, 2009.

6. TEAM Network - Early warning system for nature, Available: http://www.teamnetwork.org/en/, 2011.

7. USArray: A continental-scale seiamic observatory, Available: http://usarray.org/, 2011.

8. M. E. Pierce, G. C. Fox, J. Y. Choi, Z. Guo, X. Gao, Y. Ma, Using Web 2.0 for scientific applications and scientific communities. Concurrency and Computation: Practice and Experience. 21, 5, pp. 583-603, (2009).

9. I. Foster, Y. Zhao, I. Raicu, and S. Lu, Cloud Computing and Grid Computing 360-degree compared. in IEEE Proceedings of Grid Computing Environments (GCE) Workshop, Austin, TX, USA, 2008.

10. A. Akram, D. Chohan, X. D. Wang, X. Yang, R. Allan, A Service Oriented Architecture for Portals Using Portlets. In UK e-Science All Hands Conference, 2005.

11. Climate Monitoring Protocol. V 3.0. Conservation International, Arlington, VA, 2010.

12. Vegetation Monitoring Protocol. V 1.5. Conservation International, Arlington, VA, 2010.

13. Terrestrial Vertebrate Monitoring Protocol. V 3.0. Conservation International, Arlington, VA, 2008.

14. J. K. Hart, K. Martinez, Environmental Sensor Networks: A revolution in the earth system science? Earth-Science Reviews 78, pp. 177-191, 2006.

15. T. Fountain, S. Tilak, P. Shin, S. Holbrook, R. J. Schmitt, A. Brooks, L. Washburn, D. Salazar. Digital Moorea Cyberinfrastructure for Coral Reef Monitoring. In 5th International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2009.

16. D. Higgins, C. Berkley, M. B. Jones, Managing heterogeneous ecological data using morpho. Proc. of the 14th Intl. Conf. on Scientific and Statistical Database Management. 2002.

17. C. Berkley, M. Jones, J. Bojilova, D. Higgins, Metacat: A schema-independent XML database system. In Proceedings of the 13th International Conference on Scientific and Statistical Database Management (SSDBM '01) (July), Fairfax, VA, L. Kerschberg and M. Kafatos, Eds. IEEE Computer Society. 171–179, 2001.

18. E. H. Fegraus, S. Andelman, M. B. Jones, M. Schildhauer, Maximizing the value of ecological data with structured metadata: an

introduction to ecological metadata language (EML) and Principles for metadata creation. Bull. Ecol. Soc. Am. 86, pp. 158–168, 2005.

19. E. H. Fegraus, K. Lin, J. Ahumada, C. Baru, S. Chandra, C. Youn, Data acquisition and management software for camera trap data: A case study from the TEAM Network, this paper has been submitted to Ecological Informatics Journal, 2011.

20. Drupal - Open Source CMS, Available: http://drupal.org/, 2011.

21. GridSphere portal framework, Available: http://www.gridsphere.org, 2011.

22. Apache HttpComponents Clinet, Available: http://hc.apache.org/, 2011.

23. Apache HTTP server, Available: http://httpd.apache.org/, 2011.

24. Hibernate - Relational persistence for Java and .NET, Available: http://www.hibernate.org/, 2011.

25. C. Baru, S. Chandra, K. Lin, A. Memon, C. Youn, The GEON service-oriented architecture for Earth Science applications, International Journal of Digital Earth, 2,1, pp 62 – 78, 2009.

26. XSL-FO Version 1.1 W3C, Available: http://www.w3.org/TR/xsl11/, 2011.

27. Apache FOP (Formatting Objects Processor), Available: http://xmlgraphics.apache.org/fop/, 2011.

28. J. Chave, C. Andalo, S. Brown, M. A. Cairns, J. Q. Chambers, D. Eamus, H. Fölster, F. Fromard, N. Higuchi, T. Kira, J. P. Lescure, B. W. Nelson, H. Ogawa, H. Puig, B. Riéra, and T. Ymakura, Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia, 145:97-99, 2005.

29. Reducing greenhouse gas emissions from deforestation and degradation in developing countries: a sourcebook of methods and procedures for monitoring, measuring and reporting, GOFC-GOLD Report version COP13-2, (GOFC-GOLD Project Office, Natural Resources Canada, Alberta, Canada), 2008.

30. M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, G.A. Patterson, A. Rabkin, I. Stoica, M. Zaharia, Above the clouds: A Berkeley view of cloud computing. University of California, Berkeley, Tech. Rep., 2009, http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf.