Geophysical Research Abstracts, Vol. 9, 11174, 2007
SRef-ID: 1607-7962/gra/EGU2007-A-11174

A Next Generation Flux Network Data Server

D. Agarwal (1,2), D. Baldocchi (1), and C. van Ingen (3)

(1) University of California, Berkeley, California, USA, (2) Lawrence Berkeley Laboratory, Berkeley, California, USA, (3) Microsoft Corporation, San Francisco, California, USA (daagarwal@lbl.gov, baldocchi@nature.berkeley.edu, vaningen@windows.microsoft.com)

A continental and global network of measurement towers provide biogeoscience researchers with a wealth of long-term carbon, water and energy flux data and metadata. The data from these towers has already led to many impressive analysis results. One of the next big challenges for the networks that make up FLUXNET is to enable easy cross-site, regional, ecosystem, and global-scale analyses. Tackling this challenge requires tools beyond those typically in use today to perform analysis at a single site. We are prototyping a new *scientific data server* for use by the researchers in a collaborating group to *jointly analyze* data *across sites.* A database server provides the core of the new data server.

Working with the data and metadata from the Oak Ridge Ameriflux data repository, we at the Berkeley Water Center working with Microsoft's E-Science initiative have developed the schema and architecture of the server. This prototype server provides a framework to allow easy data download, quality checking, cleaning, and storage. The server includes all of the half-hourly Ameriflux data and also includes scientifically important metadata such as site biome or climate as well as the actual data. Data from other related data sets can also be included as needed. From the database we derive a data cube that provides a multidimensional view of the data. Use of a cube allows users to browse the data using a number of tools including Excel Pivot Tables and the web. The cube pre-calculates and stores simple statistics at yearly, monthly, and daily intervals. Through the use of the combination of the data cube and PivotTables, browsing and plotting the data is as simple as designating a set of sites, a time period, and the measurement types; the results can be plotted in a few seconds to a few minutes in the case of a very large data request. This scientific data server work will be used to support analysis efforts at the upcoming Fluxnet workshop.