Developing a cross-site system to improve access to vegetation synthetic databases:

VegDB Workshop I

Harvard Forest
19-20 June 2012

Introduction

A workshop was held at Harvard Forest to consider whether the development of a system to provide access to vegetation-related synthetic databases would be a desirable and feasible step in fostering cross-site research in the LTER network. Participants from a range of sites representing very different biomes and measurement methods met for two days to discuss this topic. Additionally participants considered how this system might be structured, what it might provide, and how it might be implemented. The following report summarizes the findings of this working group.

Overview of Problem

In the last decade there has been significant progress within the LTER network in making climate, hydrologic, and stream chemistry data more available for cross-site synthesis via a series of data harvester/processors. These early "cen-tributed" systems harvest data that is maintained at individual sites and then process and add it into a central database that can be readily accessed through a single web-based portal. This eliminates the need to go to each LTER sites' webpage, transform the many formats, interpret the various metadata, and perform the underlying calculations. Significant time and confusion is eliminated, speeding synthesis.

As these database systems have moved from climate to hydrology to steam chemistry, the complexity of the problem addressed has increased. For climate the challenge was to transform variously formatted but largely similar data into one form that can be aggregated over various time periods (day, month, and year). For hydrological data the additional step of transforming stage height to discharge using rating curve regressions was added. Stream chemistry, particularly fluxes, required the combination of discharge and analyte concentration as well as considering chemistry analysis comparability and QA/QC. Vegetation data represent the next step in complexity both in terms of data structure and in metadata. Unlike climate, hydrological, and stream chemistry data, vegetation measurements and attributes are highly varied. For many vegetation-related data products such as biomass, the raw data needs to be significantly transformed into a secondary product that requires many types of data not needed for climate, hydrology, and stream chemistry databases. Vegetation data also adds in the complexity associated with biological taxa. Vegetation-related data therefore represents ecological data in all its complexity; and a system that can deal with this level of complexity will likely represent a major transition from providing largely abiotic simple data to biotic complex data within the LTER Network.

While technologies analogous to these earlier database harvester systems may be employed to create Veg-DB, the development of this system comes as PASTA (Provenance Aware Synthesis Tracking Architecture) becomes operational. At this point it is not entirely clear how a data portal such as Veg-DB would relate to a centralized system such as PASTA, although several potential models are possible and discussed below. Regardless of the method of integration with PASTA used to achieve Veg-DB, it could serve as an example for future efforts.

Objective of the Veg-DB System

The objective of Veg-DB would be to deliver reliable and consistent vegetation-related data to users via a single web-based portal. The focus would not be on primary/raw data which can be either currently gathered from individual sites or in the future from a network system such as PASTA. Rather the system would provide access to a value-added, secondary data product with standardized units as well as the ancillary information needed to interpret these data.

There are many potential types of vegetation-related data ranging from phenology to biomass. It would be difficult if not impossible for any system to address all of them. The emphasis of Veg-DB would be on ecosystem and plant community attributes, with some important population levels ones as well (see output variables below). Aside from limiting the scope of the system to a practical size, the raw data required for population, community, and ecosystem level data is very similar; and often involves the tracking of changes in size and fate of individuals. Important aspects of vegetation not addressed would include seed production, germination, flowering, phenology, and spatial arrangements such as those represented on stem maps. The system would also not consider biomass, cover, and other vegetation variables gathered through remote sensing methods such as LIDAR.

Benefits of Veg-DB

There are several benefits that Veg-DB could provide:

This system may help sites process their data particularly when the value-added data being considered in Veg-DB is not currently a site's primary focus.

The system would help inform investigators what data is being collected at which sites.

It would help the LTER share the long-term data it collects in a meaningful and useful manner to all investigators and students.

It would enhance the LTER network capacity to lead ecological synthesis efforts, an achievement that NSF and broader scientific community is expecting.

It could help address research problems that currently viewed as data limited (see below).

Uses of Veg-DB

LTER sites have become a significant reservoir of long-term data on changes in vegetation structure, composition, and productivity. This information is essential to detect long-term trends, test hypotheses, and to evaluate simulation models; yet it is extremely difficult to process on a site by site basis. The kinds of topics and related hypotheses that Veg-DB could enable addressing include:

Individual plant growth rates versus size/age of plant. There has been a reexamination of the long-held assumption that tree growth rate decreases as the size of trees increases. This has major implications for understanding long-term temporal changes in plant biomass.

Temporal trends in mortality related to climate variability and change. There have been recent indications that mortality rates of trees have increased in North America possibly in response to a changing climate. How consistent and widespread is this pattern?

Temporal trends in NPP related to climate variability and change. Understanding the way NPP is controlled by climate is essential to predicting ecosystem responses to future climate. Exploring this topic would be enhanced with access current and widespread data.

Successional patterns of biomass accumulation and NPP. There has been a recent reexamination of whether older forests reach a balance with respect to biomass and carbon stores. NPP is often assumed to asymptote, yet there is some evidence it declines as forest reach older stages. Surprisingly little time series data is available to examine either of these important questions.

The relationship between diversity (richness, evenness, etc) and NPP. This is a long standing topic that requires significantly more data from a wide range of ecosystems to fully examine.

Correlation of temperature, precipitation, and other abiotic factors with broad-scale patterns of NPP and biomass. Many of the data currently used to address these relationships was collected decades ago. More current, high quality data would strengthen these efforts.

Veg-DB could also be a helpful resource in educational activities. By lowering the barriers to data access, students could be incorporating LTER Network data on vegetation into laboratory and course projects on a routine basis.

Finally, simulation models are key tools to enhance understanding of how ecosystems function and may respond to changing environment and management. However, before being used most simulation models require data for calibration, verification, and validation/corroboration. As with the efforts described above, having a readily available set of data from the LTER Network would greatly enable this activity.

Types of Data

Veg-DB will need to do more than deliver raw data. To function it will need to integrate the raw measurement data and supporting data to generate value-added output data. It will also need to connect to other database systems to provide the ancillary data needed to interpret these output data.

Raw data. The raw data to build the secondary data products such as biomass stores are highly heterogeneous among LTER sites (see Range of Problem section and Site Reports below). In some cases variables such as cover and biomass are collected directly at an aggregated level (e.g., all individuals within a plot) and aside from scaling to a common area (e.g., m^2 or ha) very little transformation is required. This would be typical of sites dominated by herbaceous vegetation. In other cases the raw data consist of size measurements of individuals (height or diameter) and calculation of a variable such as biomass would require additional supporting data describing size-biomass relationships.

Supporting data. Without certain kinds of supporting data, the raw data cannot be used to estimate other variables of interest. An exception might be plant cover, which is often recorded as a percent and therefore is relatively scale independent. For most other types of raw data the following must be supplied: 1) Plot size to scale results to a common area and 2) Slope steepness to put results on a horizontal distance basis. In addition when the sizes of individuals comprise the raw data, data describing the size to biomass relationship are required to calculate biomass and NPP as well as other fluxes such as mortality. Given that many species may not have these relationships, a database describing species equation substitutions is also necessary. Finally, for some output variables, it is necessary to have data describing conversion factors such as carbon concentration (to estimate carbon stores from biomass), and life-span of plant parts (to estimate litterfall from biomass).

Ancillary Data. To interpret the output data from the Veg-DB system it would be extremely helpful to have the following ancillary data available:

Location name, latitude, and longitude
Biome type using an internationally accepted system
Vegetation type using a locally based system
Stand age or age-class
Treatment history in terms of harvest and other land-uses
Disturbance history in terms of type, severity, and time

Additional variables that would be helpful to have would include:

Elevation
Slope
Aspect
Landform and topographic position
Geology (surface and subsurface)
Drainage class (e.g., poor, good, well-drained)
Moisture class (e.g., xeric, mesic, hydric)
Soil characteristics including depth, rockiness, water holding capacity, texture, pH, carbon content, and organic layer depth

Ideally the latter set of variables would not be provided by Veg-DB *per se*, but would be available through Site-DB linked from the Veg-DB portal. To function in this way SiteDB would have to be developed so as to contain this information down to the plot level.

Output Data. Compared to raw data, value-added output data will be very similar across sites. The variables to be made available via Veg-DB will include ecosystem, community, and some key population variables that can be derived from a common set of raw data:

At the ecosystem level: 1) Live biomass and carbon stores; 2) NPP; 3) Net change in live biomass; 4) mortality and litterfall; 5) ingrowth/birth of new biomass; and 6) herbivory.

At the community level: 1) Presence/absence of species; 2) dominance expressed as cover, basal area, density, volume, biomass, and carbon; and 3) diversity expressed as richness and evenness.

At the population level: 1) density of individuals; 2) recruitment into minimum size class measured; and 3) mortality of individuals.

It is unlikely all output variables will be available for all sites. Therefore sites will initially provide the data they have and will be encouraged to supply the missing data as resources allow.

Data Resolution

Veg-DB will not be able to provide data at every level of spatial, temporal, and taxonomic resolution. Therefore there will be cases in which users will have to go to the original raw data. However, data resolutions that can address many commonly posed questions might be as follows:

Temporal. While one-time measurements would be accepted into the system, multi-date data would be preferred to allow examination of temporal trends. The minimum time step of the data would be one year, either as a cumulative value (NPP and other fluxes), an annual average (biomass), or peak value (cover). For some ecosystems this would require averaging over a year (e.g., grassland), whereas for others it would require averaging over multiple years (e.g., forest).

Spatial. The spatial resolution of the data might range from individuals to subdivisions of plots, to plots, to a level of “logical” plot aggregation such as a watershed, marsh, stand or tract. The latter would vary by site and perhaps study, but would allow users to receive data aggregated above the plot level. While including data at the level of individuals would be ideal, the costs of providing data at this level need to be explored. Landscape level aggregation will not be provided by Veg-DB, but could be derived from the levels data that it does supply.

Taxonomic. Data will be available at the species, life-form (e.g., herbs, shrubs, trees), and fully aggregated levels (i.e., all plants). Additional levels such as genera should be easily derived from Veg-DB resolutions.

Plant Size. The minimum size of plants in terms of height or diameter will be determined by the sites supplying the data. Veg-DB could be constructed so as to provide either the minimum size supplied by a site or to use a fixed cut-off to create a uniform database for the user. Data will not be delivered by size classes as there is no uniform way to do this even for a given life form. If aggregation by size classes is desired by the user, then an individual plant resolution will need to be made available.

Plant age. Veg-DB would not provide data by plant age or age class other than that associated with a given plot via ancillary data.

Units. Data would be provided in standard units, although there may be a choice in the area reported as m^2 make sense for some systems (grasslands) and hectares make more sense for others (forests).

Range of Complexity of the Problem

To understand the range of complexity of developing Veg-DB two sites were compared: Plum Island Estuary’s (PIE) marsh dominated ecosystem versus the Andrews’ (AND) forest ecosystem. PIE represents the simplest case in that there is one life-form (grasses) and one species present that has biomass sampled directly. AND is more complex in that multiple life-forms of vastly different size co-occur on sloping terrain and variables such as biomass and NPP are indirectly estimated via regressions.

At PIE biomass is directly harvested in small plots for each month of the growing season and total dry weight of plants (Figure 1). These 5-6 measurements per year are used to calculate peak biomass, which is assumed to equal aboveground NPP. Given the fact that harvested plots cannot be remeasured in a year there are approximately 5 replications for each measurement to reduce the effect of spatial variation on biomass and NPP estimates. Estimates are currently made using a FORTRAN program and routinely added to the site's online database.

At AND attributes such as biomass and NPP cannot be measured directly (Figure 2). The raw data for biomass, NPP, ingrowth, and mortality estimates include tree diameter, appearance of trees reaching the minimum diameter (i.e., in-growth), identifying the trees that have died over the measurement period (i.e., mortality), shrub cover and diameters, as well as herb cover. Tree data are gathered periodically with a 3 to 6 year remeasurement interval, whereas shrub and herb data are typically one time measurements in stand ages greater than 100 years. If stands are younger than 100 years in age, then shrub and herb variables are measured on the same interval as trees. There are currently no measures of epiphytes despite the fact this life-form probably has about the same biomass as vascular plant foliage. Allometric equations based on the most likely fit for the plot are used to estimate biomass, changes in biomass, volume, ingrowth, mortality. These allometric equations are likely the largest source of uncertainty in estimates of biomass and variables derived from biomass. In addition information on plot area, slope corrections, and a total species list (used to generate the list of equations and other conversion factors) is required. NPP of woody parts is calculated by adding ingrowth and mortality to the net change in biomass from one measurement time to the next. NPP of non-woody parts is calculated from the biomass and estimates of part or plant life-span. Currently there are no measurements of fine root biomass or NPP. These analyses are currently performed on demand for investigators using FoxPro- and SAS-based programs, and are reported by species, life form, plot, and year of measurement.

Limitations to Developing Veg-DB

There are several factors that might limit the development of Veg-DB. As in any form of network science there is a social dimension: e.g. willingness of investigators to share data. Workshop participants were more than willing to share vegetation data, but sharing data brings additional costs that can make data sharing a burden and hence unattractive in practice. To be successful rewards need to be at least equivalent to costs. Rewards could take many forms including additional resources to transform the raw data to forms users desire, participation in publications, and support for analysis. Costs would include additional unfunded work, uncertainty about data use (e.g., inappropriate analyzes and secondary sharing outside the original agreement), and unacknowledged efforts when data users do publish.

Another limitation relates to resources and their timing relative to new demands. It would be less burdensome for sites to provide the value-added data that they have readily available rather than all the data that is potentially available. Decisions concerning which data to include in Veg-DB are best left at the site level. The eventual goal would be to have all the relevant data available, but a longer time line would be more realistic than a short one when resources are limited.

As with any new activity new resources to conduct the work are necessary. Several efforts need to be considered when seeking funding:

Creating the value added databases used by Veg-DB will require additional work by LTER sites that goes beyond providing and documenting raw data. This includes the supporting data and programs needed to transform the raw data into the value added product as well as ancillary data that will enhance its interpretation. Some of these costs could be spread among sites with common raw data types and life-forms. This would have the added benefit of providing sites with data relevant to Veg-DB resources to calculate variables that are not part of their current emphasis.

Creating a user friendly and transparent user interface for Veg-DB will be essential. This will not only cost programing time, but the system will need to be tested and improved so that it can intuitively be used by “non-experts” such as graduate students.

Creating scripts to harvest the value-added data for PASTA will be a necessary step. As with the other database systems, a script to routinely harvest and update the data relevant to Veg-DB would be highly desirable. Sites would have the responsibility to update raw data and the value-added data sets depending on the structure adopted. Ideally a network system would harvest data and metadata, but this system needs to be developed and maintained.

It would be extremely helpful if there was simultaneous funding for a science-based project that would use the system in a real application. This would push the development of the system and provide insights into how it could be improved.

Possible Structures of Veg-DB

The current structure for retrieving raw and value-added vegetation data is to go to individual LTER site websites. If the value-added data such as biomass does not exist, then users have to derive it on their own. With the implementation of PASTA users will have one location at which to find the raw or value-added data. However, users will have to create their own tool to aggregate or filter the data.

Veg-DB could have two possible roles as part of a new system. In the simplest configuration, Veg-DB would primarily be an aggregation and filtering tool (Figure 3). Raw data (data level 1) would be processed at each site using their existing or newly created data workflows to create the value-added databases (data level 2). This newly created data and relevant metadata would in turn be uploaded into PASTA. Veg-DB would draw upon the data within PASTA, filtering and aggregating it as the user needs. This configuration builds upon existing infrastructure at sites and will be faster to implement. However, it will exclude sites that do not have a suitable workflow infrastructure.

A more complicated configuration would have Veg-DB serve several roles (Figure 3). It would continue to serve as a data filter and aggregator for users, but it would also create the value-added databases (data level 2), removing that responsibility from individual sites. Instead sites would be responsible for providing all the raw and supplemental data as well as work scripts and programs required to create the higher level database (data level 2). It would be the network responsibility to periodically update the value-added databases. This system would take longer to create, but would have the advantage of reducing each sites workload and likely allow the system to be more inclusive and current.

Evolution Over Time

As with any complex activity, the development of Veg-DB will take time. The workshop participants recommended a two stage development process. The first phase would have the sites create the value-added databases and upload them into PASTA (Figure 3). Veg-DB would then be used to extract and aggregate the data in the form required by the user. The downside of this approach is that sites would have to run the programs to create the secondary value-added data. If sites are already doing this activity, then the additional burdens will be minimal. However, if sites are not then it means additional, and from a site's perspective, unnecessary work. The likely outcome is that the data Veg-DB accesses would be incomplete in terms of sites and potentially not current. The upside is that phase I could be implemented fairly rapidly at low cost which would have the benefit of demonstrating the system's potential utility.

The second phase of development would transfer the responsibility of creating the value-added data from sites and make that a network level activity (Figure 3). Sites would provide all the raw and supplemental data to create the value added database to PASTA. The Veg-DB system would periodically rerun the calculations to produce the value-added database and enter it into PASTA. Veg-DB would also be used to retrieve, aggregate, and report the value-added data. The upside of this configuration is that it would only require sites to update their raw and supplementary data. This might encourage more sites to participate and for data to be more current. The downside is that this would require significant investments in creating a centralized processing system. If the value-added database is created "on the fly" it would potentially be more current, but might also slow down processing time to an undesirable degree.

Another aspect of evolution to consider is the number of sites involved. While including all sites with vegetation data is a desirable final goal, it would be better to start with a smaller set of prototype sites. These should represent as wide a range of ecosystem types as possible, but regardless of the ecosystems selected they should represent the full range of plant life-forms and types of raw data measurements that are taken within the LTER network. This will not only assure that the system will be general enough, but will share expertise among sites. For example, NPP estimated in forested ecosystems focus on tree-related NPP, but herbaceous and shrubby life-forms may contribute a significant share of total forest NPP. Expertise to improve estimates for herbaceous and shrub life-forms is far more likely in grassland or desert regions than in forest ones; hence non-forest site expertise could improve forest ecosystem estimates.

While initially designed for the LTER network, Veg-DB could be used by other entities with similar data and missions such as the US Forest Services experimental forests. This has occurred for other databases systems and would be a natural step for Veg-DB.

Relationship to Other Efforts

There are a number of other efforts that Veg-DB should take advantage of including:

Existing site level scripts and programs to create value-added data. More information needs to be gathered about what already exists at sites and what can be used to build upon.

The Veg-X exchange format for vegetation data has developed a common XML-based language which describes entities (projects, plots, etc) and resolves issues related to taxonomy of organisms and communities. Veg-DB should employ Veg-X standards whenever possible.

Veg-DB designer could learn from other previous efforts such as SiteDB, ClimDB, HydroDB, and CTFS-SIGEO. This will help designers anticipate problems and create solutions.

Use other database systems in a complementary manner. For example, much of the ancillary site data required by users could be supplied by Site-DB. Likewise data on climate could be provided by Climb-DB. The key will be to develop user friendly linkages that are obvious and useful.

Next Steps

Inform the LTER network of plans for Veg-DB. This will be achieved at the upcoming 2012 All-Scientists meeting and via the network newsletter.

Conduct a LTER community survey to determine the kinds of vegetation data and processing infrastructure at sites, the degree of interest in participating in Veg-DB, and information on how scientists and educators might use VegDB.

Design a user interface mock-up that can be used to demonstrate how the system would work by showing the kinds of possible queries.

Determine the optimal integration method(s) with PASTA.

Select the prototype sites based level of interest, ecosystem and life-forms present, and existing infrastructure (scripts and programs) that can be used.

Secure funding and resources for development of Phase I of the system.

Summary

The working group participants agreed that the creation of Veg-DB would be both feasible and desirable. While raw data describing vegetation attributes is quite heterogeneous, derived products such as biomass, NPP, relative cover, etc are quite similar for a wide range of ecosystem types. Several configurations of Veg-DB were considered, and a two phase development was recommended. Both phases would provide a single web-based portal to value-added vegetation databases, but dividing the development would allow users to explore the utility of the proposed system quicker and at a lower cost. The working group recommended that Veg-DB should begin the next steps of design and prototype implementation.

Figure 1: VegDB – simple example for Plum Island (PIE)

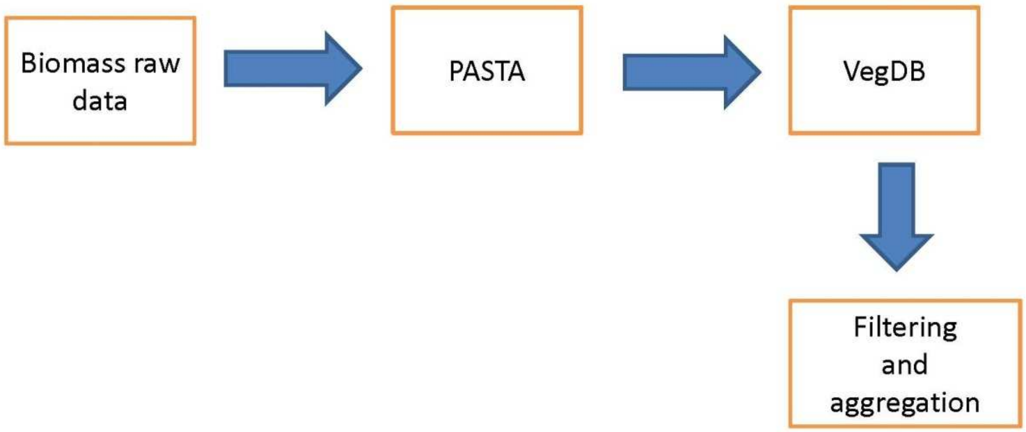


Figure 2. VegDB – complex example for Andrews Forest

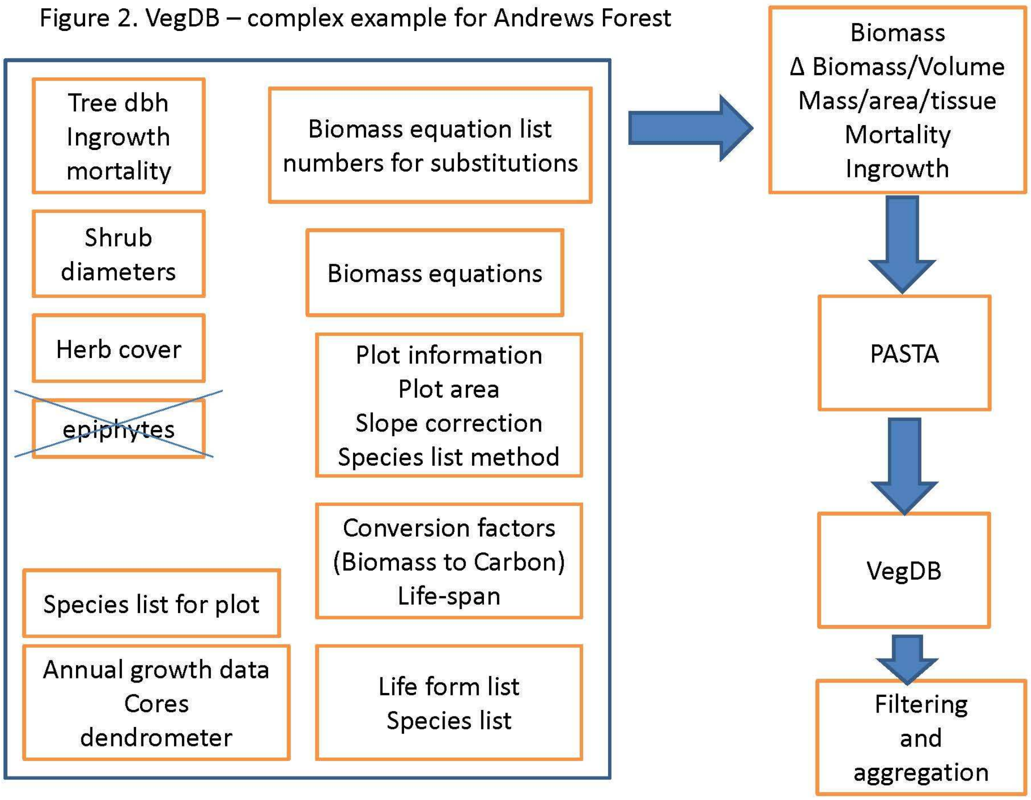
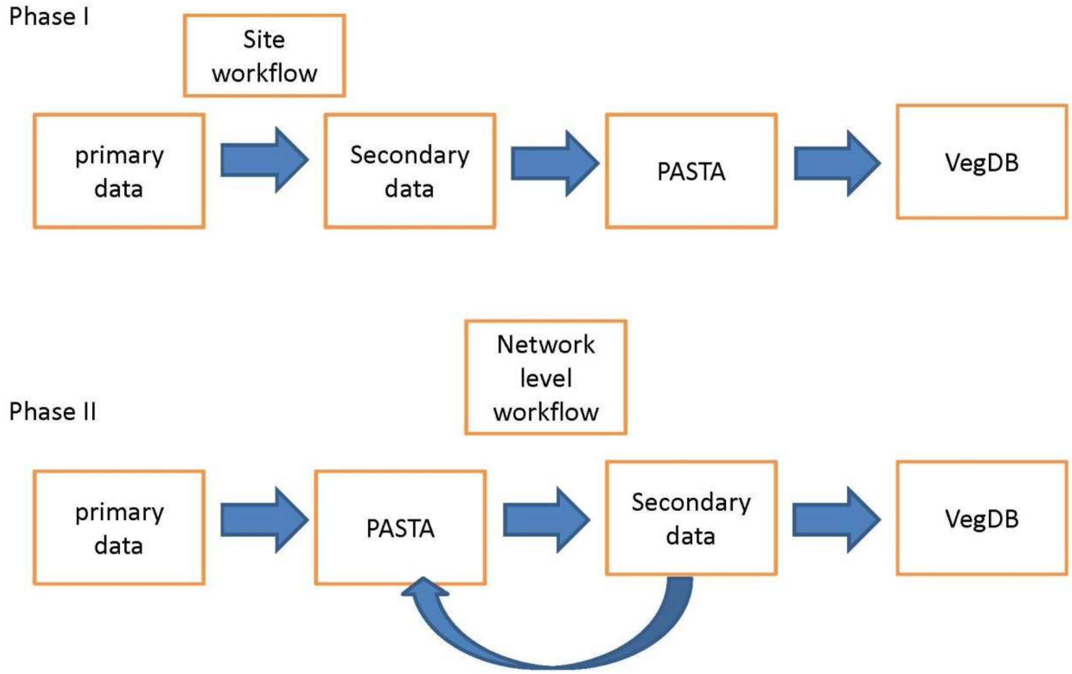


Figure 3. Possible VegDB configurations.



Workshop Participants

Attending:

Audrey Barker Plotkin (HFR)
Emery Boose (HFR)
David Foster (HFR)
Mark Harmon (AND)
Jim Morris (PIE)
Dave Orwig (HFR)
Suzanne Remillard (AND)
Roger Ruess (BNZ)

Conference call:

John Blair (KNZ)
Laura Gough (ARC)

Site Reports

ARC (Laura)

- Site interests include: community composition, productivity, and biomass for long-term monitoring and determining responses to experimental treatments (i.e., warming, fertilization).
- Annual non-destructive sampling includes cover for describing plant communities. Sampling is driven by funded projects.
- There are also periodic biomass harvests of all plants: herbs, shrubs, mosses, in which above ground biomass, root biomass, nitrogen, and carbon are determined every 5-10 years. Harvests are used to estimate productivity. This sampling regime also varies with funding.
- Weekly NDVI measurements with hand-held instrument (greenness) are now being taken. Regressions between NDVI and biomass are used to predict biomass, but the relationship is not great.
- The point-frame method is also used for estimating biomass.
- Traditional phenology measurements and flowering data have been taken, but the sampling intensity varies over time.
- No concerns about sharing the data, although retaining a lag for graduate students is desirable.
- Time required to reformat data is a concern. Maintaining ARC data is enough.
- Synthetic efforts usually involve emailing the PIs responsible for collecting the data. It would help to have a better sense of what is available at each site to facilitate cross-site analyses.

AND (Mark)

- Andrews manages data for a larger region, here the focus is on AND-LTER data.
- There are two kinds of plots: (1) Watershed transects of plots, circular plots (0.1 ha), not slope corrected. (2) Reference stands (0.25 to 2 ha), slope corrected and designed to represent stages of forest succession in which trees mapped.
- Both systems of plots are re-measured every 3-6 years. Trees are tagged in field.
  Past data is used to check new measurements.
- On some plots and time periods annual mortality checks were made that identified trees that have died, their position (standing or fallen), and cause (where possible).
- Shrubs, small trees, herbaceous layers have been sampled in most plots, but most stands do not have regular estimates of understory biomass.
- Data used to estimate carbon stores in forests, forest dynamics, etc.
- DEMS (from LIDAR) used for slope corrections.
- Biomass equations are used to estimate biomass, NPP, mortality but there are different versions that can lead to uncertainty as high as 50%.
- Litter traps were used in selected stands in the past, but currently this is not being done. Current estimate leaf production and that of understory plants comes from estimates of biomass and part life-span.
- There are estimates of macronutrient stores in the vegetation within the watersheds sampled based a recent determination of nutrient concentrations.
- Not much below-ground work as root production is difficult to measure in these rocky soils.

- Sharing data with people who have time and interest to conduct analyses has been productive, so there are few concerns on data sharing.
- Concerns about simply supplying raw data and then having to transform it for users. Some users have not understood the complexity of the datasets and not used it appropriately.
- There are concerns about time, effort, and resources. Not looking to increase workload.

BNZ (Roger)

- BNZ set up to study successional dynamics at different sites: floodplains (fluvial dynamics), uplands (fire).
- Monitoring program focused on 35 sites. Most sites have been sampled annually since 1988.
- Control plots 50 x 60m in which plant cover, seedlings, trees, litterfall, and seedfall, etc. are measured at different time steps depending on successional stage.
- To determine the portion of NPP consumed and not measured hare pellets and browsing locations on plants are noted. Moose pellets not measured.
- Trees are tagged, mapped, and re-measured every 5 years. Dendrometers have been placed on a subset of trees to measure changes over a shorter period.
- Data are used for carbon & nitrogen budgets, synthesis efforts, modeling.
- Other data relevant to vegetation include biomass equations, moss production, and shrub NPP.
- BNZ is involved in a number of network analyses including: biomass, AGNPP, tissue turnover by functional type, stand demographics and successional dynamics, sensitivity of AGNPP to direct and interactive press / pulse drivers, diversity / species change (invasives) impacts on AGNPP.
- No concerns about sharing vegetation databases.

HFR (Audrey)

- Focus of HFR is on woody plants and describing these populations communities in space and time (including a historical perspective).
- There are four major plot types: 20x20m, large stem-mapped, experimental controls, and tower plots.
- In 20x20m plots there are measurements of various strata: tree diameter are taken, saplings / seedlings, understory vegetation cover class is determined. Nearly all these plots were established for a particular objective. Related foci on decomposition, CWD, hemlock mortality, N cycling, tree cores, etc.
- Large stem-mapped plots such as SIGEO. Individual trees tagged and mapped with dendrometer bands in SIGEO plots (some measured every 2 weeks). SIGEO plot was surveyed to 20x20m plots.
- Experimental & tower plots. E.g. experimental hurricane, hemlock removal experiment, EMS tower biometry plots. Wider range of measurement that include tree diameters.
- Ground-based LIDAR has also been used in some of the plots.
- A list of datasets can be taken from HF Data Archive.

KNZ (John)

- Focus on herbaceous grasslands, although there is some sampling in gallery forests.

- Site interests include long-term trends in productivity associated with fire return intervals, response to nutrient and rainfall manipulations, identifying the primary drivers of plant distribution (fire, climate, topographic position).
- Productivity estimated by clipping once per year at end of growing season, although earlier biomass measurements made three times per year and in a few cases bi-weekly clippings were taken.
- Most data collected annually (woody plants every 5 years). Infrequent measurements of root biomass, and nitrogen and phosphorus content at various times.
- Species cover composition is determined on permanent transects.
- Some plant phenology data is collected, mostly by docents at the watershed scale.
- Need to improve sampling of shrub communities (e.g. biomass).
- No site-level concerns about sharing data.
- Only concern is additional investment of time and effort.

PIE (Jim)

- The focus at PIE is on marsh plant productivity on various plots.
- On plots with one species present (Spartina) bBiomass and NPP is determined via destructive sampling. Access database includes plot, subplot, plant and sib IDs. System includes some QA/QC for data entry.
- On mixed plots there is non-destructive sampling in permanent plots of 100 cm squared in which stems are tagged with bird bands. Annual measurements of height with ruler are recorded on tape recorder and transcribed in office. These data are converted to biomass using regression equations. The output is ASCII file. A FORTRAN program calculates biomass, births, deaths, density, and age. Final FOTRAN program calculates aggregate data for online publication as received.
- Also collect data on snails, water chemistry of pore water (daily), site elevation, etc.