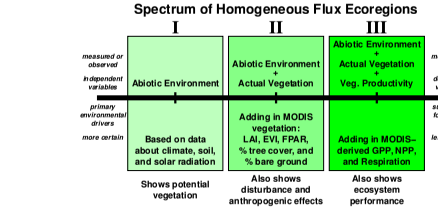
Quantifying Importance Values for AmeriFlux Sites

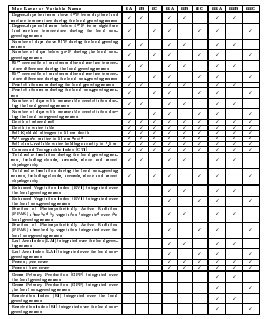
Introduction

Multivariate clustering based on fine spatial resolution maps of flux-relevant ecosystem characteristics has been used at several specified levels of division to produce nine sets of **quantitative flux-relevant ecoregion maps** for the conterminous United States. Within any single flux ecoregion, the carbon flux from terrestrial ecosystems is expected to be relatively **homogeneous**.

Choosing any statistical ecoregion establishes an origin in data space. Using the Euclidean distance from this origin to each other ecoregion, a pairwise similarity measure can be calculated. Within the context of an **existing network of sites or samples**, quantitative difference in data space can form the basis for a tool which quantifies the degree of representativeness coverage for a particular established network. A network in this sense consists of a geographic constellation of existing installations, facilities, or locations where samples have been taken. Quantitative similarity is based on comparison with the most-similar of all of the sites within an established network. The best place to locate an **additional site** is the place that is the **least well-represented** by the network of existing sites.

	Abiotic Environment	*Abiotic Environment + Actual Vegetation*	*Abiotic Environment + Actual Vegetation + Vegetation Productivity*
Annual (growing and nongrowing)	**IA** 18 input maps	**IIA** 26 input maps	**IIIA** 30 input maps
Local Growing Season	**IB** 12 input maps	**IIB** 17 input maps	**IIIB** 19 input maps
Local Non-growing Season	**IC** 12 input maps	**IIC** 17 input maps	**IIIC** 19 input maps





List of Input Maps

Network Analysis

Produces a map of "representativeness" of your current network of installations or samples, **showing areas of good and poor representation**,

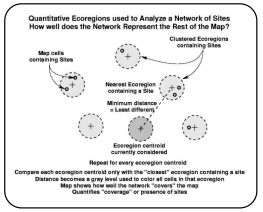
Areas which are poorly represented by the current network configuration are the best candidate locations for additional stations,

Marginal representation of each station is determined by **first adding, then removing it** to calculate the representation that it contributes to the network – Provides a quantitative basis for minimizing the impact if a station must be removed from the network.

A **theoretically optimum network** with a given number of stations can be designed that has the highest possible representation for the combinations of conditions in the map that comprise it,

Network Analysis is **ecoregion-based**, and operates at the scale of the entire sampling network,

Results are **with respect to the selected input variables only**, and depend on the intelligent selection of the environmental variables being considered, and on the quality of the input data,



Importance Values

A single number inversely reflecting the overall **representativeness of a network** can be **calculated** as the sum of the shortest Euclidean distances from every ecoregion to the closest ecoregion containing a site,

By calculating the network representativeness **with and without a particular site**, the **marginal contribution of that site to the representativeness of that network** can be calculated,

The difference between the representativeness of the network with and without this site is called the **Site Representativeness Importance Value**, and reflects **the marginal contribution of that site to the representativeness of the network**,

Site Representativeness Importance Value has a very specialized meaning with regard to network representativeness,

Site Representativeness Importance Values **allow direct comparison of sites** in terms of their contribution toward network representativeness,

Site Representativeness Importance Values are **dynamic if the network changes – If any sites are added or removed**, the Site Representativeness Importance Value will change for all sites, and **must be recalculated**,

Equal Weighting

All environmental characteristics are weighted **equally** – Assumes that if a particular combination of flux-relevant conditions is present in the conterminous United States, **then it should be sampled** – i.e., presence is enough to require sampling,

It is more likely that common combinations of flux conditions will be sampled with towers, just because they are common – Unusual combinations are less likely to be sampled with towers,

The method does **NOT focus on uniqueness**, *per se*, but instead focuses on sample **coverage**,

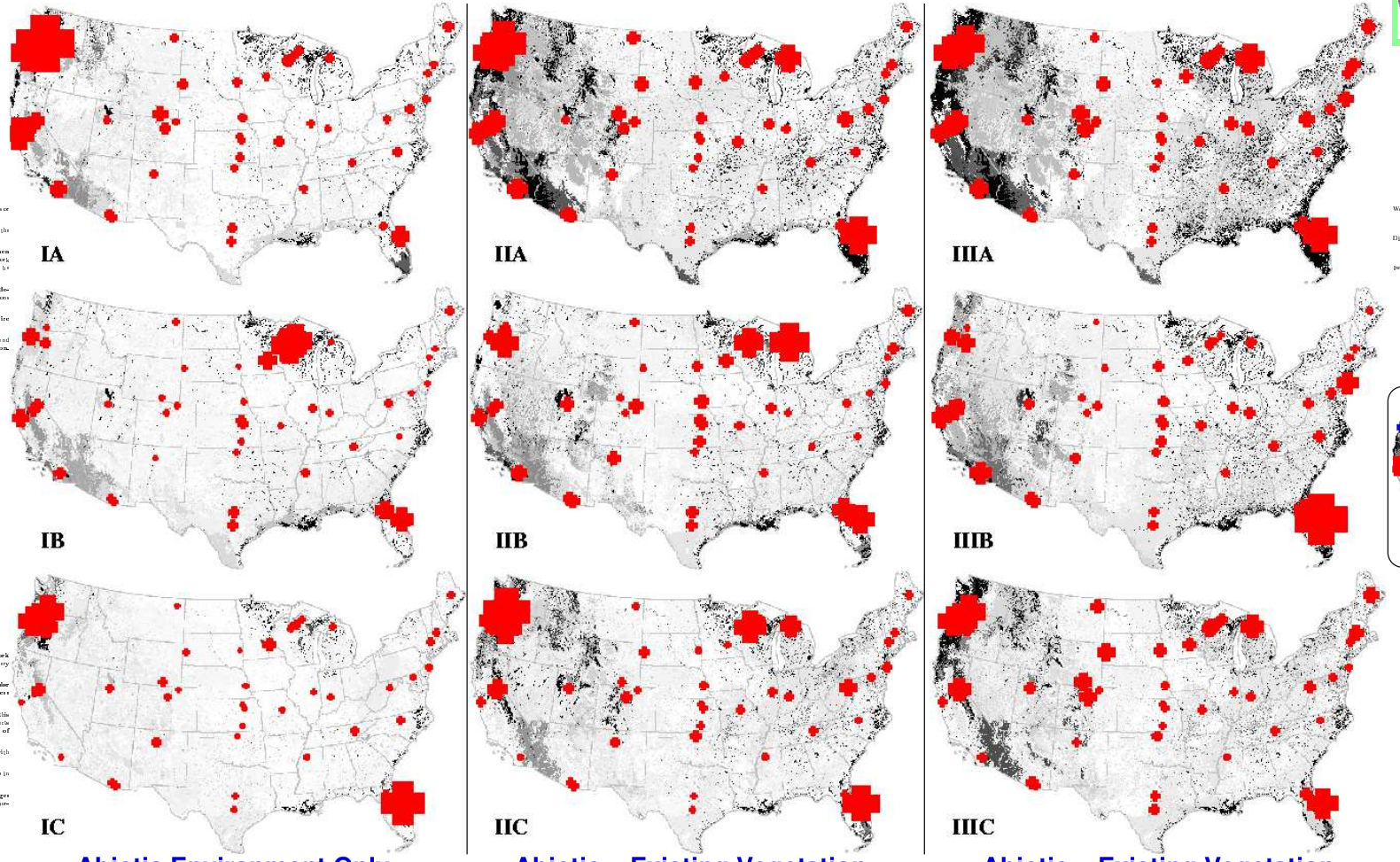
The need to sample is **irrespective of the size of the flux ecoregion**,

Weighting by Area would imply that smaller unique flux ecoregions are less important to sample than larger ones,

Weighting by Magnitude of Carbon Flux would imply that unique flux ecoregions with less carbon exchange are less important to sample than ones with greater carbon exchange,

Weighting by both Area and Magnitude of Carbon Flux would imply that unique flux ecoregions with less carbon exchange are less important to sample than ones with greater carbon exchange, **UNLESS they are very large**,

We have used **equal weighting** rather than any of these other schemes,



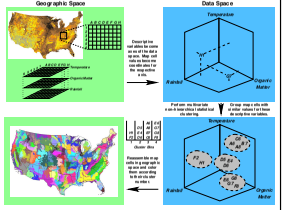
Abiotic Environment Only
Shows Potential Vegetation

Abiotic + Existing Vegetation
Shows Disturbance /Human Effects

Abiotic + Existing Vegetation
+ Ecosystem Performance

Size of red cross indicates importance value of this AmeriFlux station

Darker areas are more poorly represented by existing AmeriFlux stations



Multivariate Geographic Clustering

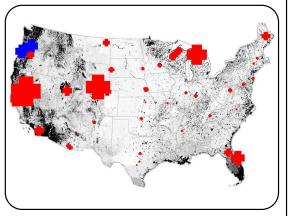
High–Res Analysis

We resolved all the IIIA set of flux ecoregions to increase the number of flux ecoregions from 90 to 1500 for increased resolution of both network representativeness and site importance values,

Dividing into many more flux ecoregions **dramatically increases Site Importance values**, since fewer flux ecoregions actually comprise AmeriFlux towers,

Importance values of all sites were **scaled down by 80 times**; the two blue sites were **scaled down by an additional factor of ten**,

High–Res Map



Site Importance List

Tower Site	Importance Value
Florida	377.6148
Wind River Crane Site	376.9994
Blodgett Forest	91.7266
Glacier Lake	67.9970
Univ. of Mich. Biological Station	58.1972
Jasper Ridge	43.3156
Florida-Kennedy Space Flight Center (scrub oak)	38.0088
Vaira Ranch	37.7666
Tonzi Ranch	37.7666
Utah Sagebrush	26.8569
Utah Juniper woodland	26.8569
Utah Coastal sheatgrass	26.8569
Sky Oaks Biological Station	25.9483
Howland Forest (east tower)	24.1695
Gainesville (Deerlabless) - slash pine (mid-rotation)	23.0645
CRP unmowed	23.0496
CRP minimum-till	23.0496
CRP grazed	23.0496
Metolius - Intermediate (69 yrs) ponderosa pine	21.1709
Willow Creek	19.6797
Sylvania Wilderness Area (Watersmeet)	19.6797
Fort Peck	19.4238
Northern Michigan Jack Pine Stand	19.2344
Gainesville (Austin Cary) - 46yr naturally regenerated slash/loblolly pine	18.1486
Niwot Ridge Forest (2)	17.8799
Niwot Ridge Forest	17.8799
Florida-Kennedy Space Flight Center (slash pine)	17.7496
Metolius - new young pine	15.7008
Black Hills	15.4785
Park Falls WLEF	14.9558
Metolius - Eyerly burn	13.0127
Santa Rita Mesquite	12.9434
Audubon Research Ranch	12.9434
Rosemount	12.2949

Conclusions

The large-scale patterns of representation indicated for the AmeriFlux network are **robust to prediction**, regardless of which statistical ecoregions are selected for the basis of the analysis,

We prefer using the **IIIA flux ecoregions**, which are based on the greatest number of flux-relevant environmental characteristics (30 map layers),

The **Pacific Northwest**, the **Sierra Nevada mountains**, and the **Sonoran desert region** are poorly represented by existing AmeriFlux towers,

Flux environments within the **Florida peninsula** are also shown to be poorly represented by AmeriFlux towers,

The poorly represented ecoregions are probably **unusual and unique combinations** of flux-relevant environmental characteristics, making them difficult to sample,

All of these poorly sampled regions are **located around the geographic periphery** of the conterminous United States,

The **most important sites in the network also tend to be located around the periphery** of the lower 48 states – **consistent set of results**,

It may be relatively easy to capture representation of the more average **continental interior**, but much **more difficult** to achieve representation of unusual regions on the **margins**,

Complete Presentation available at
http://research.esd.ornl.gov/~hnw/ameriflux/oct2005

William W. Hargrove
hnw@fire.esd.ornl.gov
Forrest M. Hoffman
forrest@climate.ornl.gov