Representativeness-Based Sampling Network Design for the Arctic

Forrest M. Hoffman[†‡], Jitendra Kumar[†], Richard T. Mills[†], and William W. Hargrove*

[†]Oak Ridge National Laboratory,
[‡]University of California-Irvine, and
*USDA Forest Service, Eastern Forest Environmental Threat Assessment Center (EFETAC)

98th Annual Ecological Society of America Meeting, Minneapolis, MN, USA

August 5, 2013

U.S. DEPARTMENT OF ENERGY

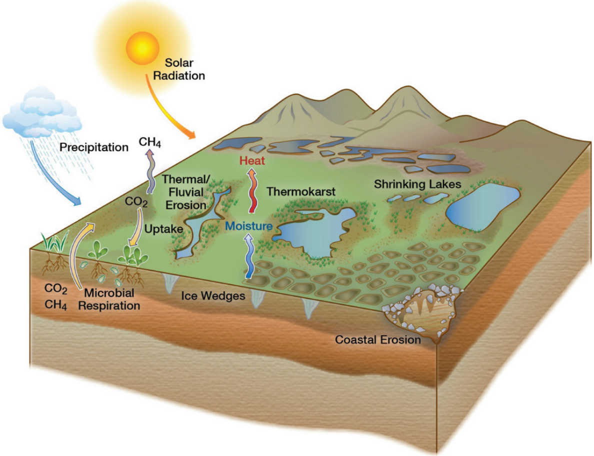
FOREST SERVICE U.S. DEPARTMENT OF AGRICULTURE

Climate Change Science Institute AT OAK RIDGE NATIONAL LABORATORY

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE U.S. DEPARTMENT OF ENERGY

Next-Generation Ecosystem Experiments (NGEE Arctic)

http://ngee.ornl.gov/

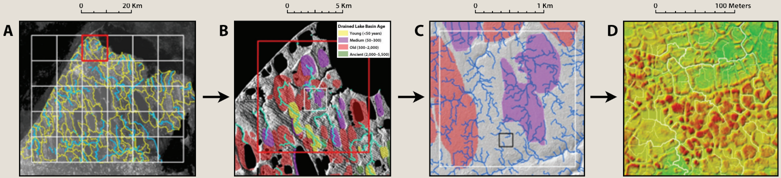


The Next-Generation Ecosystem Experiments (NGEE Arctic) project is supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Integrating Across Scales

- NGEE Arctic process studies and observations are strongly linked to model development and application for improving process representation, initialization, calibration, and evaluation.
- A hierarchy of models will be deployed at fine, intermediate, and climate scales to connect observations to models and models to each other in a quantitative up-scaling and down-scaling framework.

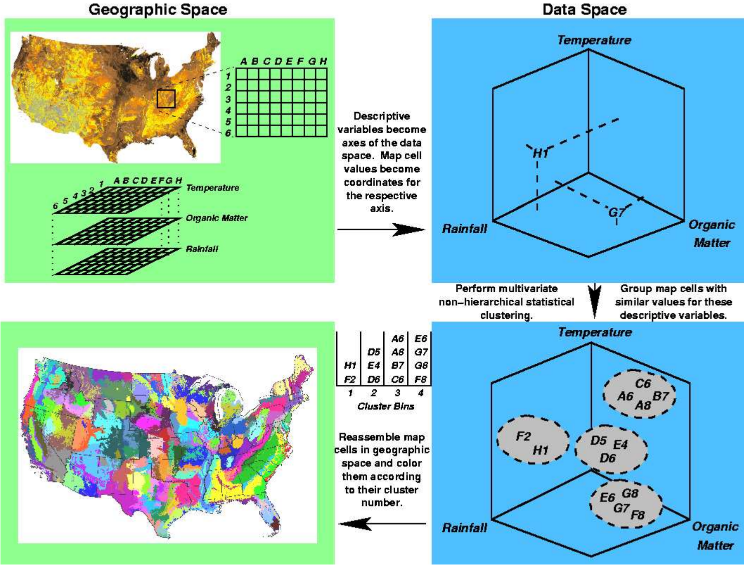
Hydrologic and Geomorphic Features at Multiple Scales. *At the scale of (A) a high-resolution ESM, (B) a single ESM grid cell, (C) a 2 × 2 km domain of high-resolution Light Detection and Ranging (LiDAR) topographic data, and (D) polygonal ground. Yellow outlines in panel A show geomorphologically stable hydrologic basins, connected by stream channels (blue). Colored regions in panels B and C show multiple drained thaw lake basins within a single 10 × 10 km grid cell (B) or a 2 × 2 km domain (C), with progressively more detailed representation of stream channels (blue). Colors in panel D represent higher (red) to lower (green) surface elevations for a fine-scale subregion, with very fine drainage features (white). [Los Alamos National Laboratory, University of Alaska Fairbanks, and University of Texas at El Paso]*



Quantitative Sampling Network Design

- Resource and logistical constraints limit the frequency and extent of observations, necessitating the development of a systematic sampling strategy that objectively represents environmental variability at the desired spatial scale.
- Required is a methodology that provides a quantitative framework for informing site selection and determining the representativeness of measurements.
- Multivariate spatiotemporal clustering (MSTC) was applied at the landscape scale (4 km^2) for the State of Alaska to demonstrate its utility for representativeness and scaling.
- An extension of the method applied by Hargrove and Hoffman for design of National Science Foundation's (NSF's) National Ecological Observatory Network (NEON) domains.

Multivariate Spatiotemporal Clustering (MSTC)

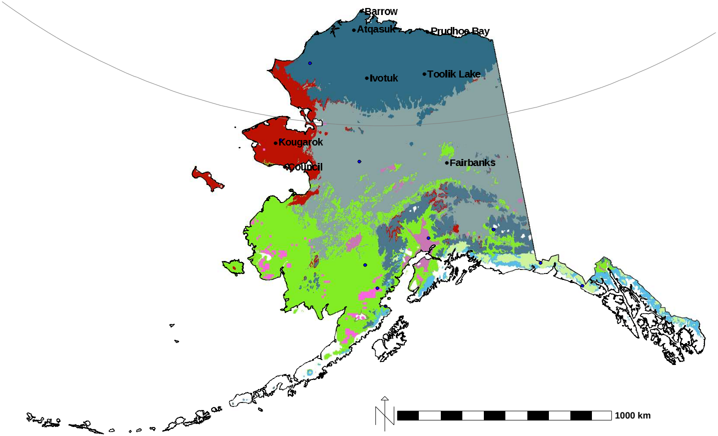


Data Layers

Table: 37 characteristics averaged for the present (2000–2009) and the future (2090–2099).

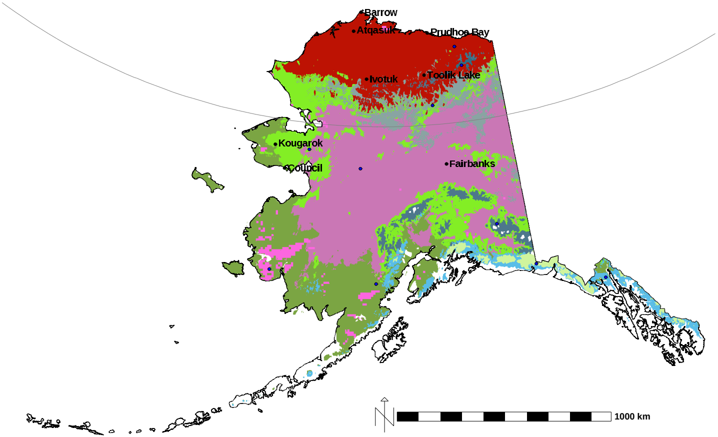
Description	Number/Name	Units	Source
Monthly mean air temperature	12	°C	GCM
Monthly mean precipitation	12	mm	GCM
Day of freeze	mean	day of year	GCM
	standard deviation	days	
Day of thaw	mean	day of year	GCM
	standard deviation	days	
Length of growing season	mean	days	GCM
	standard deviation	days	
Maximum active layer thickness	1	m	GIPL
Warming effect of snow	1	°C	GIPL
Mean annual ground temperature at bottom of active layer	1	°C	GIPL
Mean annual ground surface temperature	1	°C	GIPL
Thermal offset	1	°C	GIPL
Limnicity	1	%	NHD
Elevation	1	m	SRTM

10 Alaska Ecoregions (2000–2009)



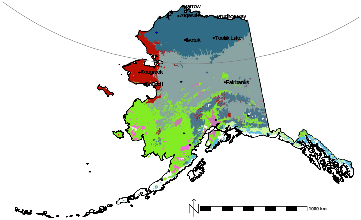
Each ecoregion is a different random color. Blue filled circles mark locations most representative of mean conditions of each region.

10 Alaska Ecoregions (2090–2099)

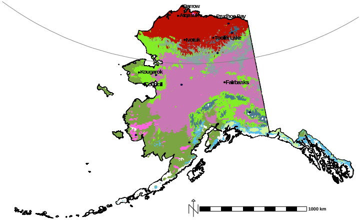


Each ecoregion is a different random color. Blue filled circles mark locations most representative of mean conditions of each region.

10 Alaska Ecoregions, Present and Future



2000–2009

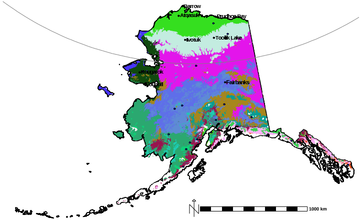


2090–2099

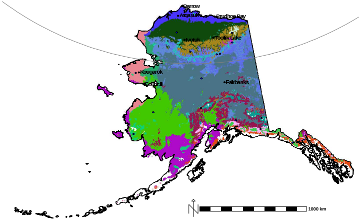
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

At this level of division, the conditions in the large boreal forest become compressed onto the Brooks Range and the conditions on the Seward Peninsula "migrate" to the North Slope.

20 Alaska Ecoregions, Present and Future



2000–2009

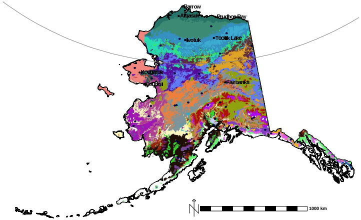


2090–2099

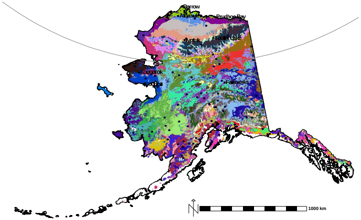
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

At this level of division, the two primary regions of the Seward Peninsula and that of the northern boreal forest replace the two regions on the North Slope almost entirely.

50 and 100 Alaska Ecoregions, Present



$k = 50$, 2000–2009

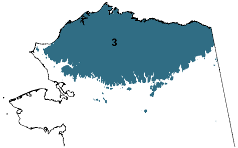


$k = 100$, 2000–2009

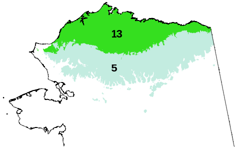
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

At high levels of division, some regions vanish between the present and future while other region representing new combinations of environmental conditions come into existence.

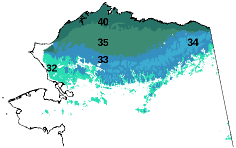
A Hierarchy of Ecoregions



(a) At $k = 10$, the North Slope is occupied by Ecoregion #3, which corresponds to the Arctic Tundra Level 2 ecological group.



(b) At $k = 20$, the North Slope is occupied by Ecoregion #5, corresponding to the Brooks Range ecoregion; and Ecoregion #13, corresponding to the Beaufort Coastal Plains ecoregion.

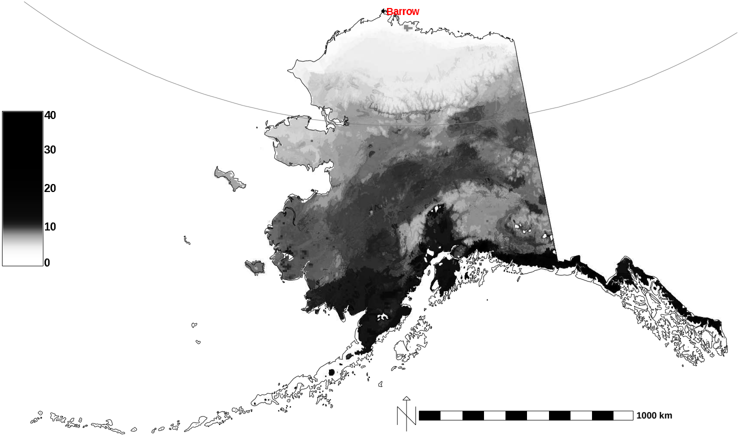


(c) At $k = 50$, the North Slope is occupied by Ecoregion #32, corresponding to the Intermontane Boreal ecological group; Ecoregions #33 and #34, corresponding to mid- and high-elevation of the Brooks Range ecoregion; Ecoregion #35, corresponding to the Brooks Foothills ecoregion; and Ecoregion #40, corresponding to the Beaufort Coastal Plains ecoregion.

NGEE Arctic Site Representativeness

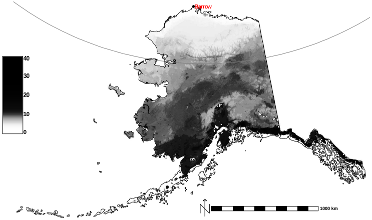
- This representativeness analysis uses the standardized n-dimensional data space formed from all input data layers.
- In this data space, the Euclidean distance between a sampling location (like Barrow) and every other point is calculated.
- These data space distances are then used to generate grayscale maps showing the similarity, or lack thereof, of every location to the sampling location.
- In the subsequent maps, white areas are well represented by the sampling location or network, while dark and black areas as poorly represented by the sampling location or network.
- This analysis assumes that the climate surrogates maintain their predictive power and that no significant biological adaptation occurs in the future.

Present Representativeness of Barrow or “Barrow-ness”

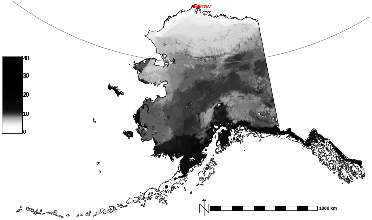


Light-colored regions are well represented and dark-colored regions are poorly represented by the sampling location listed in **red**.

Present vs. Future Barrow-ness



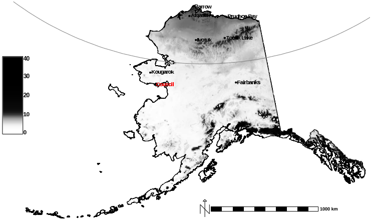
2000–2009



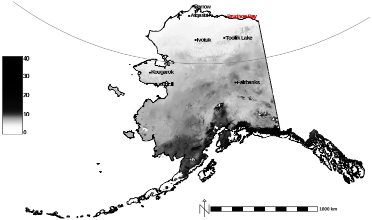
2090–2099

As environmental conditions change, due primarily to increasing temperatures, climate gradients increase and the representativeness of Barrow will be diminished in the future.

Council and Prudhoe Bay Representativeness



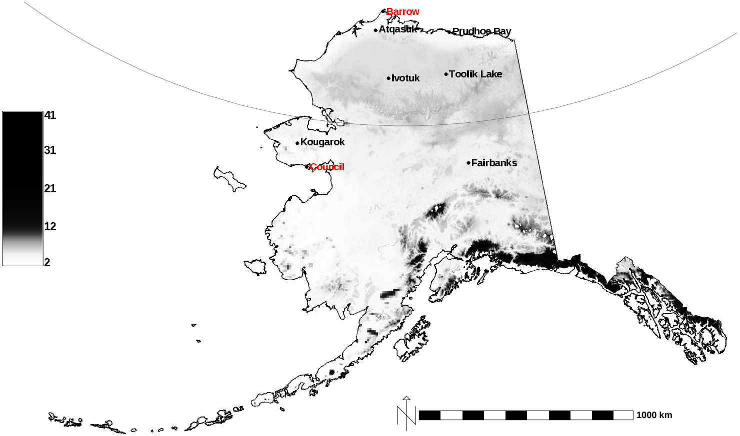
Council



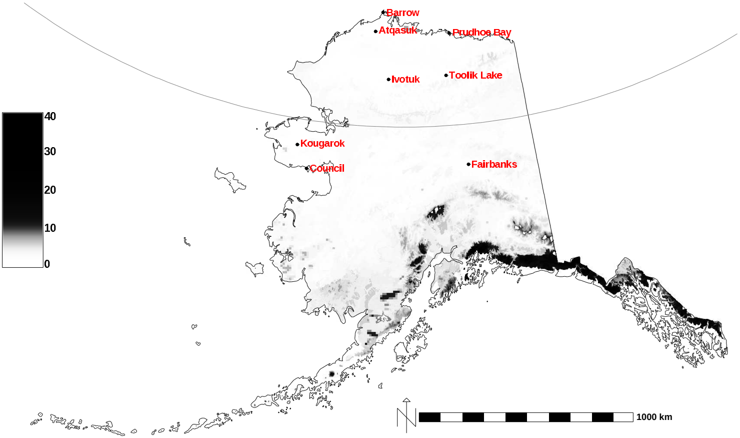
Prudhoe Bay

Representativeness analysis was performed for sites at Barrow, Council, Atqasuk, Ivotuk, Kougarok, Prudhoe Bay, Toolik Lake, and Fairbanks.

Network Representativeness: Barrow + Council



Network Representativeness: All 8 Sites



State Space Dissimilarities: 8 Sites, Present (2000–2009)

Table: Site state space dissimilarities for the present (2000–2009).

Sites	Council	Atqasuk	Ivotuk	Toolik Lake	Kougarok	Prudhoe Bay	Fairbanks
Barrow	9.13	4.53	5.90	5.87	7.98	3.57	12.16
Council		8.69	6.37	7.00	2.28	8.15	5.05
Atqasuk			5.18	5.23	7.79	1.74	10.66
Ivotuk				1.81	5.83	4.48	7.90
Toolik Lake					6.47	4.65	8.70
Kougarok						7.25	5.57
Prudhoe Bay							10.38

State Space Dissimilarities: 8 Sites, Future (2090–2099)

Table: Site state space dissimilarities for the future (2090–2099).

Sites	Council	Atqasuk	Ivotuk	Toolik Lake	Kougarok	Prudhoe Bay	Fairbanks
Barrow	8.87	4.89	6.88	6.94	8.04	4.18	11.95
Council		8.82	6.93	7.74	2.43	8.24	5.66
Atqasuk			5.86	5.84	8.15	2.30	10.16
Ivotuk				2.01	7.27	4.75	7.51
Toolik Lake					7.81	5.00	8.33
Kougarok						7.89	6.42
Prudhoe Bay							9.81

State Space Dissimilarities: 8 Sites, Present and Future

Table: Site state space dissimilarities between the present (2000–2009) and the future (2090–2099).

		Future (2090–2099)							
	Sites	Barrow	Council	Atqasuk	Ivotuk	Toolik Lake	Kougarok	Prudhoe Bay	Fairbanks
Present (2000–2009)	Barrow	3.31	9.67	4.63	6.05	5.75	9.02	3.69	11.67
	Council	8.38	1.65	8.10	5.91	6.87	3.10	7.45	5.38
	Atqasuk	6.01	9.33	2.42	5.46	5.26	8.97	2.63	10.13
	Ivotuk	7.06	7.17	5.83	1.53	2.05	7.25	4.87	7.40
	Toolik Lake	7.19	7.67	6.07	2.48	1.25	7.70	5.23	8.16
	Kougarok	7.29	3.05	6.92	5.57	6.31	2.51	6.54	5.75
	Prudhoe Bay	5.29	8.80	3.07	4.75	4.69	8.48	1.94	9.81
	Fairbanks	12.02	5.49	10.36	7.83	8.74	6.24	10.10	1.96

Representativeness: A Quantitative Approach for Scaling

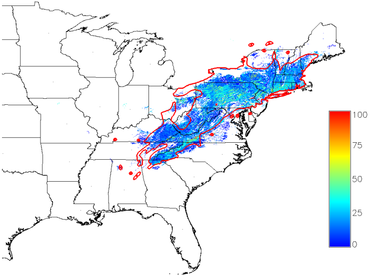
- MSTC provides a quantitative framework for stratifying sampling domains, informing site selection, and determining representativeness of measurements.
- Representativeness analysis provides a systematic approach for up-scaling point measurements to larger domains.
- Methodology is independent of resolution, thus can be applied from site/plot scale to landscape/climate scale.
- It can be extended to include finer spatiotemporal scales, more geophysical characteristics, and remote sensing data.
- Paper describing the methodology is in press:

Hoffman, F. M., J. Kumar, R. T. Mills, and W. W. Hargrove (2013) "Representativeness-Based Sampling Network Design for the State of Alaska." *Landscape Ecol.*, in press. doi:10.1007/s10980-013-9902-0.

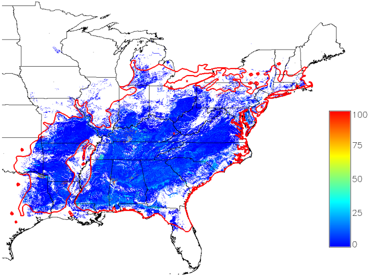
A Scaling Example: Estimating Tree Species Ranges

We used cluster analysis in the scaling of tree productivity measures in the continental United States using sparse Forest Inventory and Analysis (FIA) measurements.
(Kumar et al., in prep.)

For more information, attend:
COS 26-4
Imputation of continuous tree suitability over the Continental United States from sparse measurements
Jitendra Kumar et al.
Tuesday, August 6 at 8:40 a.m.
Room L100F, Minneapolis Convention Center



Sweet Birch (*Betula Lenta*)



Flowering Dogwood (*Cornus florida*)

AGU Fall Meeting Session

IN006. Big Data in the Geosciences: New Analytics Methods and Parallel Algorithms

Co-conveners: Jitendra Kumar (ORNL), Robert Jacob (ANL), Don Middleton (NCAR), and Forrest Hoffman (ORNL)

Confirmed Invited Presenters:

- Gary Geernaert (U.S. Dept. of Energy)
- Matt Hancher (Google Earth Engine)
- Jeff Daily (Pacific Northwest National Laboratory)
- William Hargrove (USDA Forest Service)

Earth and space science data are increasingly large and complex, often representing long time series or high resolution remote sensing, making such data difficult to analyze, visualize, interpret, and understand. The proliferation of heterogeneous, multi-disciplinary observational and model data have rendered traditional means of analysis and integration ineffective. This session focuses on development and applications of data analytics (statistical, data mining, machine learning, etc.) approaches and software for the analysis, assimilation, and synthesis of large or long time series Earth science data that support integration and discovery in climatology, hydrology, geology, ecology, seismology, and related disciplines.

Abstract submissions are due August 6, 2013.

Acknowledgments

The Next-Generation Ecosystem Experiments (NGEE Arctic) project is supported by the Office of Biological and Environmental Research in the U.S. Department of Energy (DOE) Office of Science. This research was sponsored by the Office of Biological and Environmental Research in the U.S. Department of Energy Office of Science and the U.S. Department of Agriculture (USDA) Forest Service, Eastern Forest Environmental Threat Assessment Center (EFETAC). This research used resources of the National Center for Computational Science at Oak Ridge National Laboratory, which is managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

References

F. M. Hoffman, J. Kumar, R. T. Mills, and W. W. Hargrove. Representativeness-based sampling network design for the State of Alaska. *Landscape Ecol.*, 2013. doi:10.1007/s10980-013-9902-0. In press.