

Representativeness- Based Sampling Network Design for the Arctic

Forrest M. Hoffman^{1,2},
Jitendra Kumar²,
Damian Maddalena²,
Richard T. Mills³, and
William W. Hargrove⁴

¹University of California-Irvine,
²Oak Ridge National Laboratory,
³Intel Corp., and ⁴USDA Forest
Service, Eastern Forest
Environmental Threat Assessment
Center (EFETAC)

2014 Annual US-IALE Symposium
Anchorage, Alaska, USA

May 19, 2014



Climate Change
Science Institute

AT OAK RIDGE NATIONAL LABORATORY

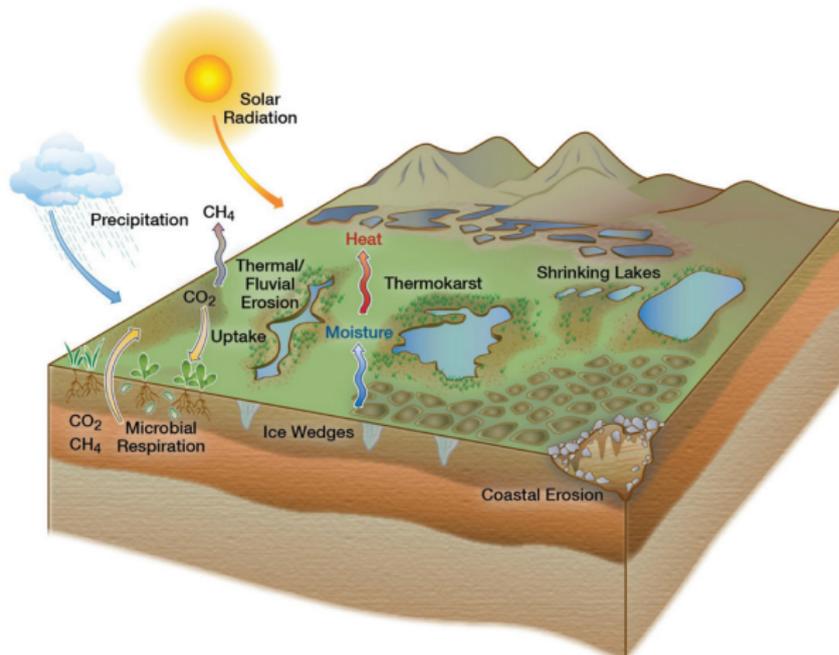


OAK RIDGE NATIONAL LABORATORY

MANAGED BY UT-BATTELLE FOR THE U.S. DEPARTMENT OF ENERGY

Next-Generation Ecosystem Experiments (NGEE Arctic)

<http://ngee.ornl.gov/>



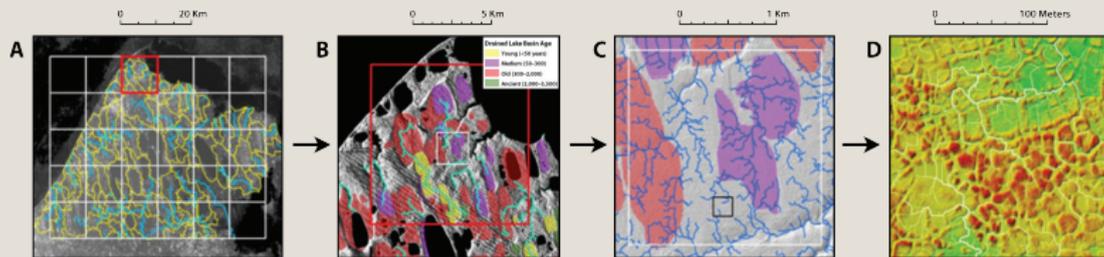
The Next-Generation Ecosystem Experiments (NGEE Arctic) project is supported by the Office of Biological and Environmental Research in the DOE Office of Science.



Integrating Across Scales

- ▶ NGE 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 Earth and Space Observing (ESMO) topographic data, (B) a single ESMO 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]

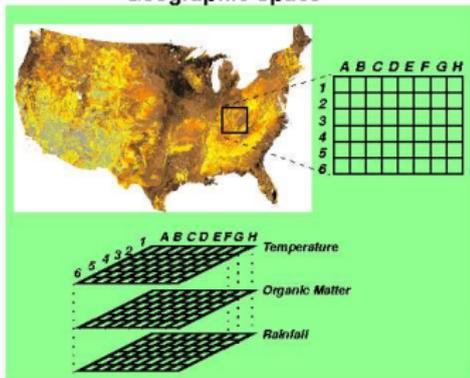


Quantitative Sampling Network Design

- ▶ 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 km²) 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 (Schimel et al., 2007; Keller et al., 2008).

Multivariate Spatiotemporal Clustering (MSTC)

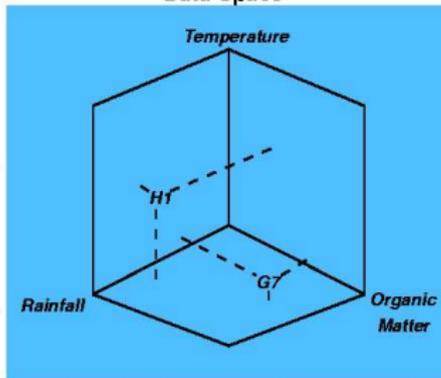
Geographic Space



Descriptive variables become axes of the data space. Map cell values become coordinates for the respective axis.

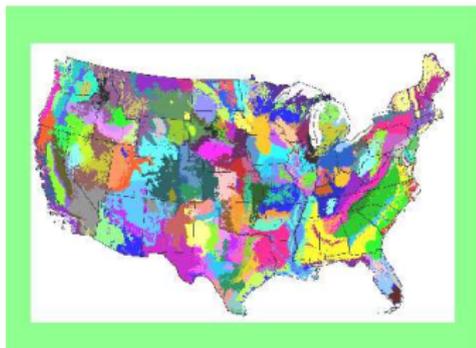


Data Space



Perform multivariate non-hierarchical statistical clustering.

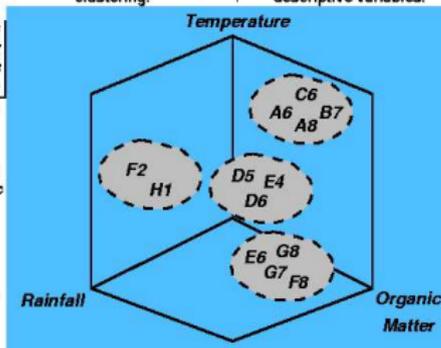
Group map cells with similar values for these descriptive variables.



	A6	E6		
	D5	A8	G7	
H1	E4	B7	G8	
F2	D6	C6	F8	
1	2	3	4	

Cluster Bins

Reassemble map cells in geographic space and color them according to their cluster number.

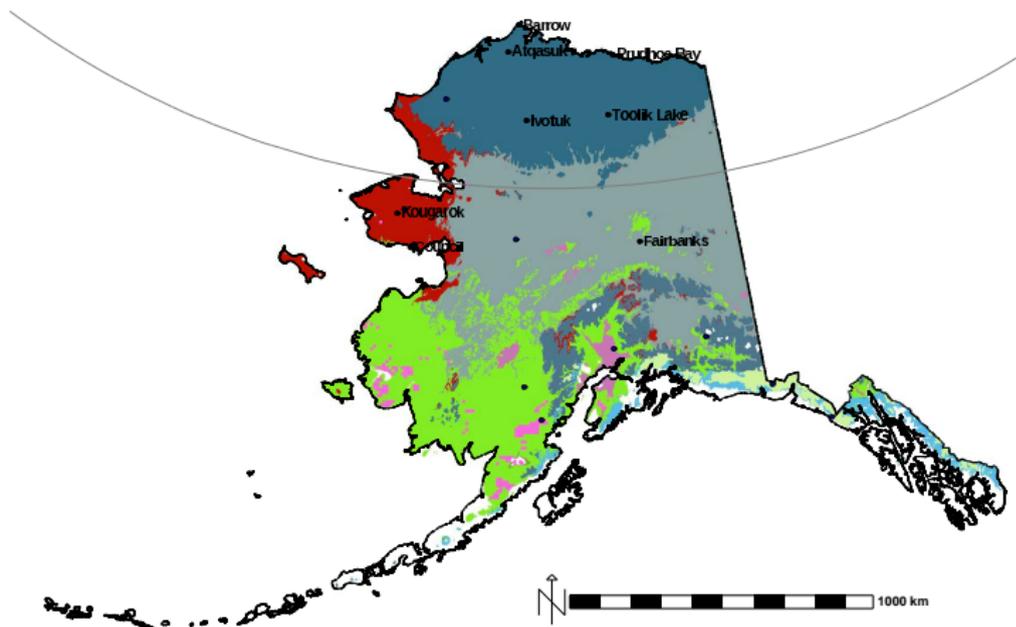


Data Layers

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

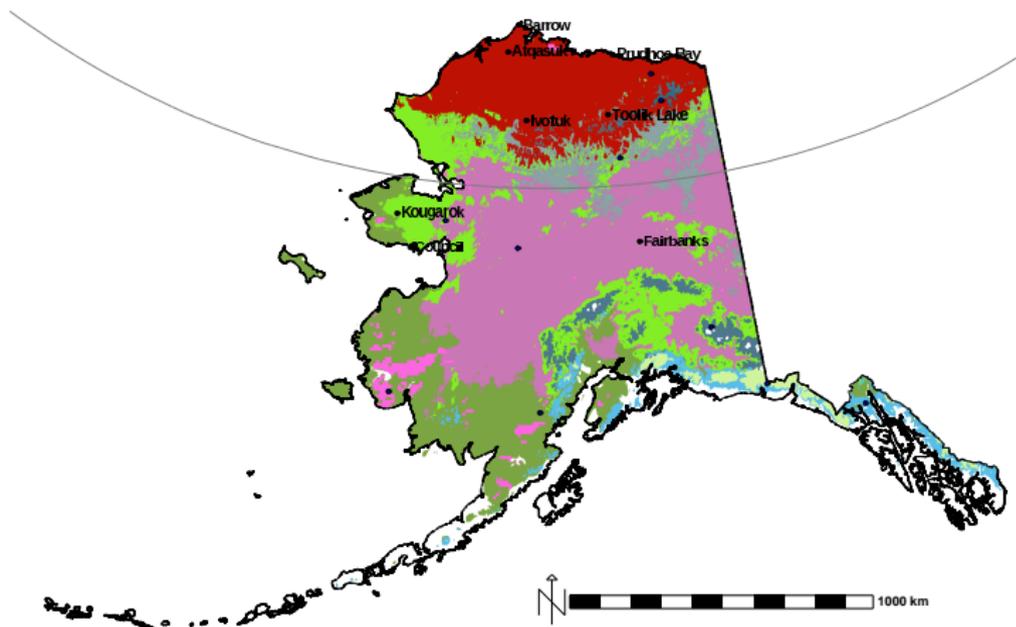
10 Alaska Ecoregions (2000–2009)



(Hoffman et al., 2013)

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

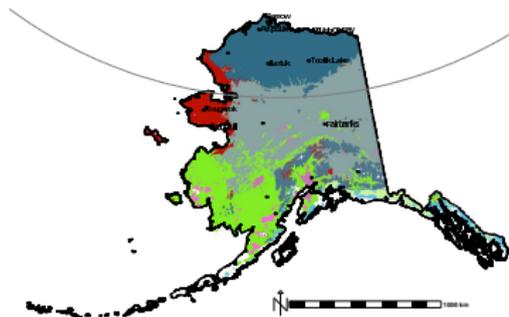
10 Alaska Ecoregions (2090–2099)



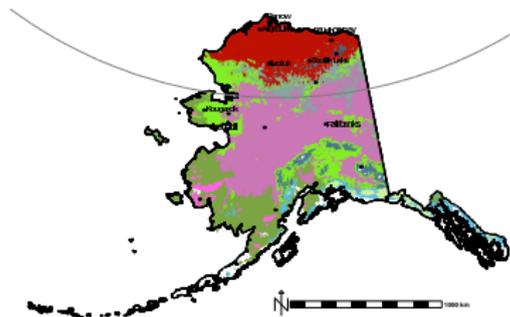
(Hoffman et al., 2013)

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



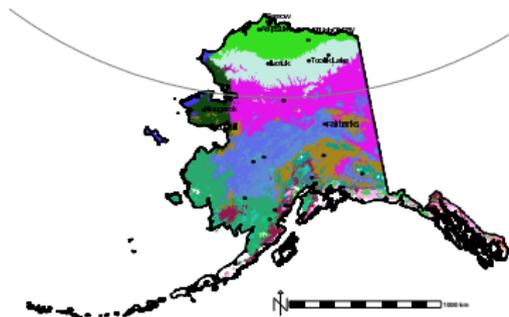
2090–2099

(Hoffman et al., 2013)

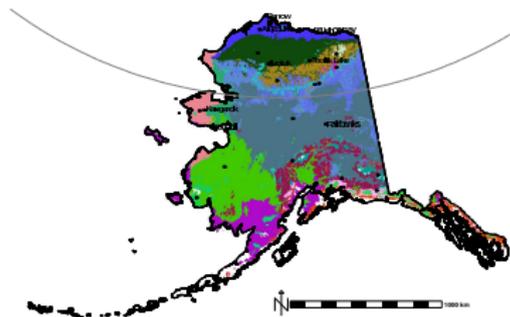
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

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



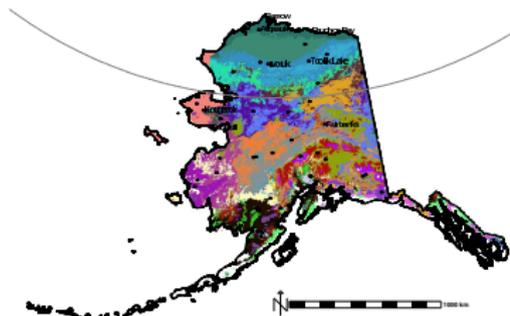
2090–2099

(Hoffman et al., 2013)

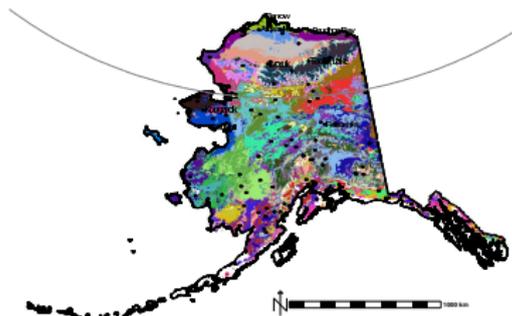
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

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

(Hoffman et al., 2013)

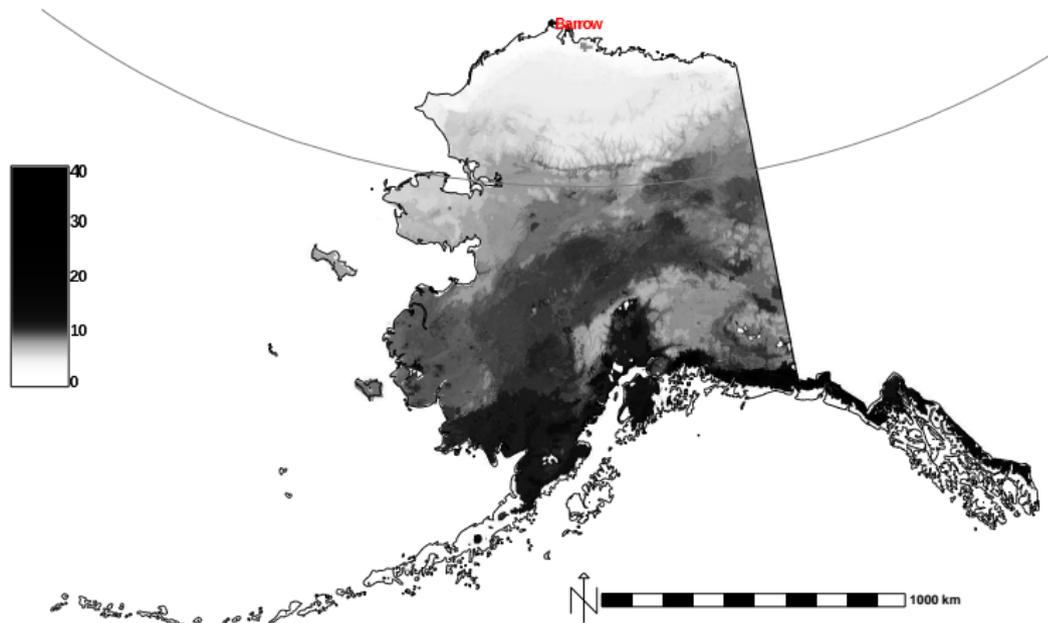
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

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.

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.

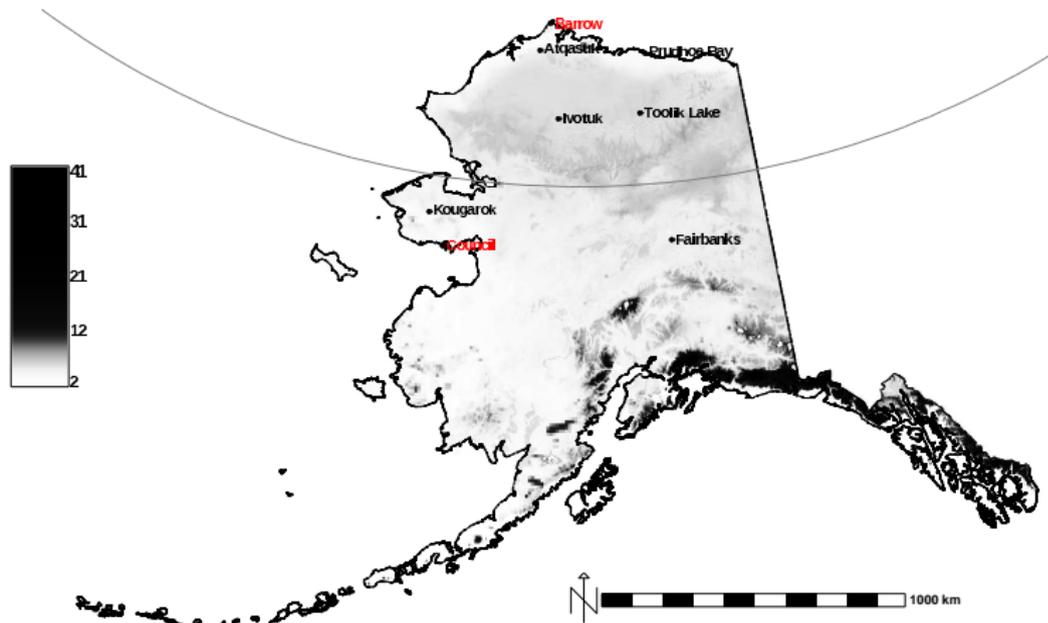
Present Representativeness of Barrow or “Barrow-ness”



(Hoffman et al., 2013)

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

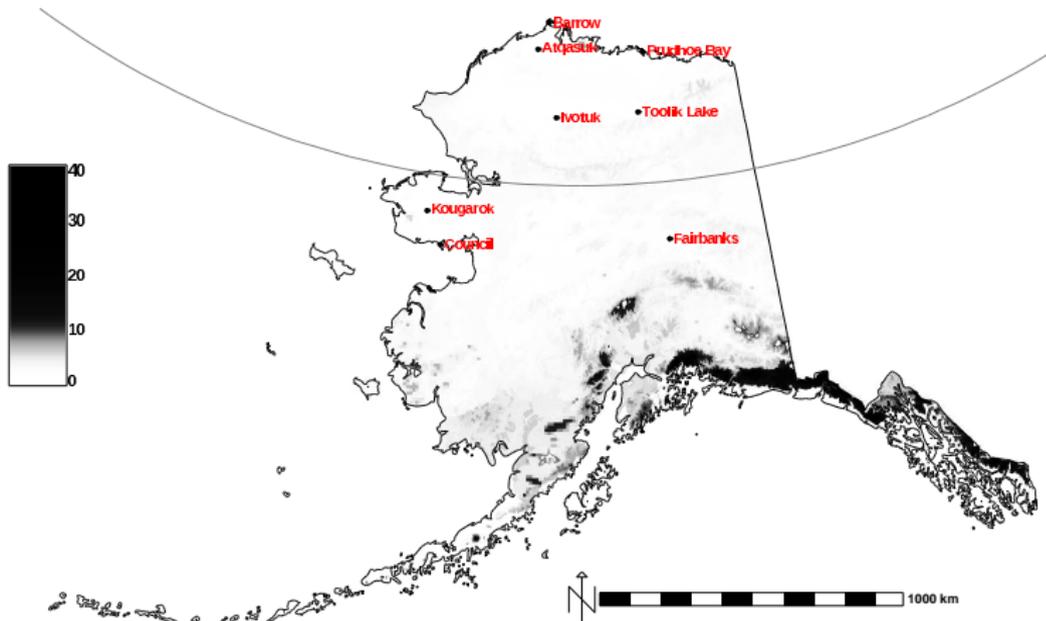
Network Representativeness: Barrow + Council



(Hoffman et al., 2013)

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

Network Representativeness: All 8 Sites



(Hoffman et al., 2013)

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

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

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

Sites				Toolik		Prudhoe	
	Council	Atqasuk	Ivotuk	Lake	Kougarok	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				Toolik		Prudhoe	
	Council	Atqasuk	Ivotuk	Lake	Kougarok	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

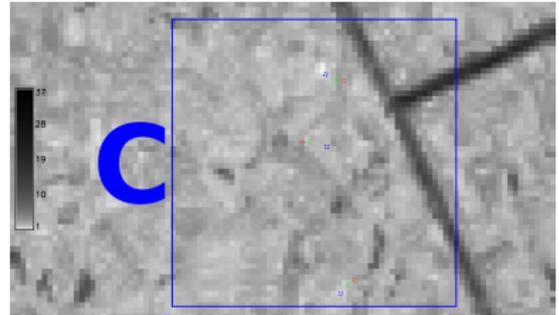
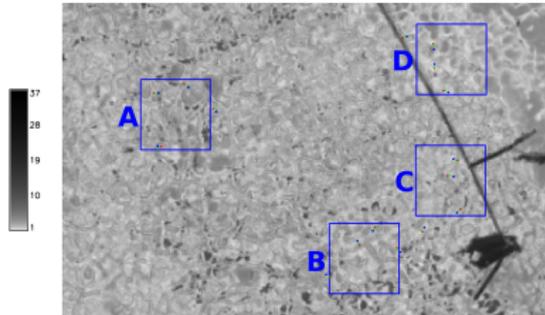
Table: Site state space dissimilarities between the present (2000–2009) and the future (2090–2099).

		<i>Future (2090–2099)</i>							
		Sites	Barrow	Council	Atqasuk	Ivotuk	Toolik		Prudhoe
Lake	Kougarok						Bay	Fairbanks	
<i>Present (2000–2009)</i>	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.
- ▶ Methodology is described in an Open Access paper: 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.*, 28(8):1567–1586. doi:10.1007/s10980-013-9902-0.
- ▶ Resulting maps and data are available from (the first Ngee Arctic Data DOI): doi:10.5440/1108686.

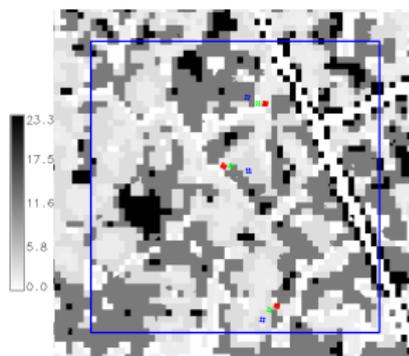
Barrow Environmental Observatory (BEO)



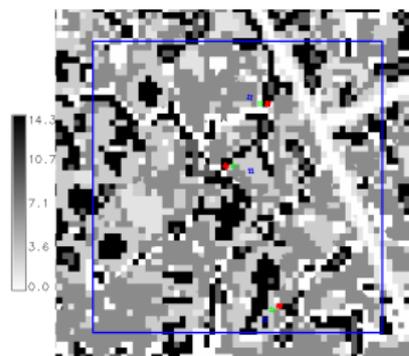
(Kumar et al., in prep)

Representativeness map for vegetation sampling points for A, B, C, and D sampling area (left) and zoomed in on the C sampling area (right) developed from WorldView2 satellite images for the year 2010 and LiDAR data.

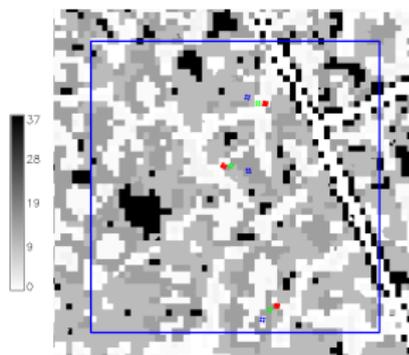
Vegetation sampling locations represent polygon troughs (red), edges (green), and centers (blue).



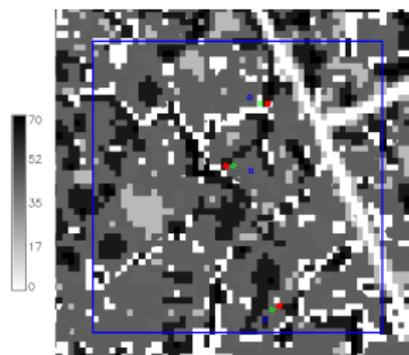
(a) dry tundra gramanoid



(b) forb



(c) lichen



(d) moss
(Kumar et al., in prep)

Example plant functional type (PFT) distributions scaled up from vegetation sampling locations.

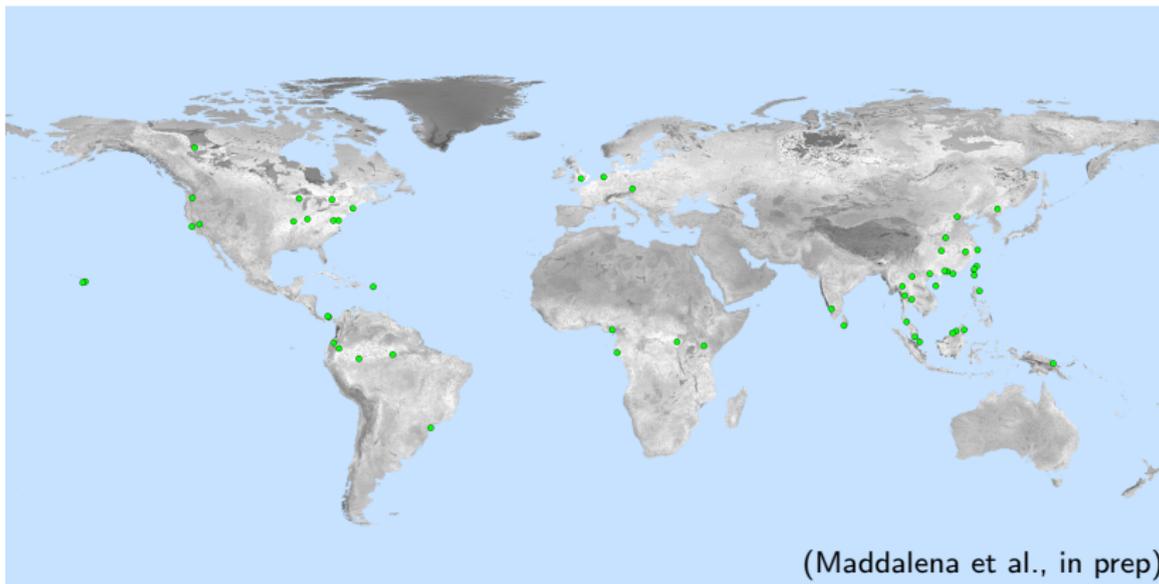
Stay Tuned for More!

See the talk at 2:20 p.m. in this session/room:

**Mapping plant functional type distributions in
Arctic ecosystems using multi-spectral remote
sensing and vegetation survey datasets**

by Jitendra Kumar

ForestGEO Network Global Representativeness

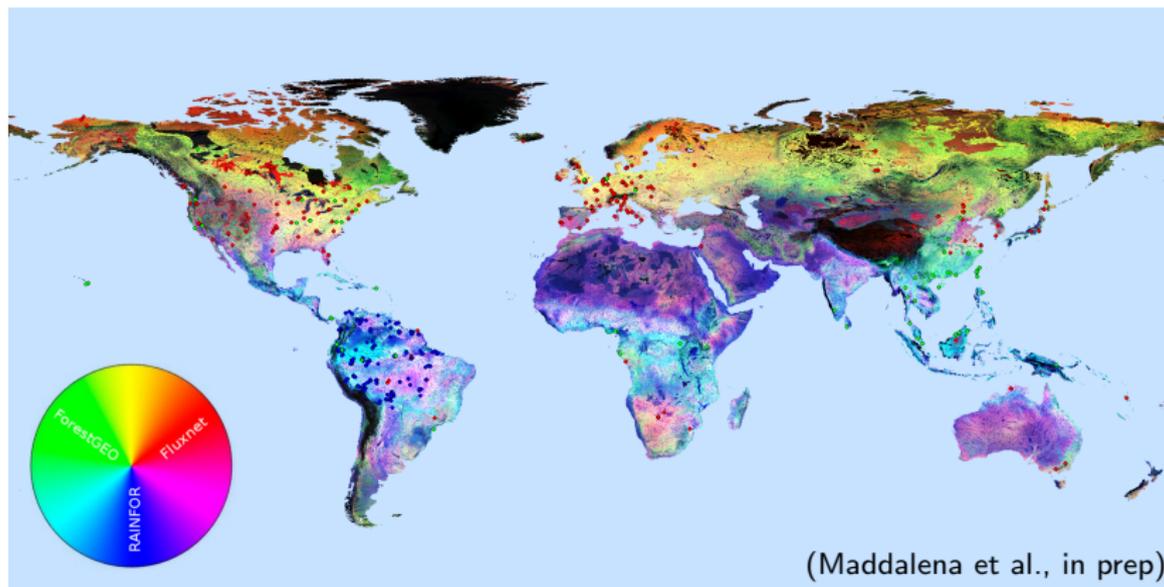


Light-colored regions are well represented and dark-colored regions are poorly represented by the ForestGEO sampling network.

Animation of the time evolution of the ForestGEO network:

https://climate.ornl.gov/~jkumar/share/forestGEOall_years.gif

Triple-Network Global Representativeness



Map indicates which sampling network offers the most representative coverage at any location. Every location is made up of a combination of three primary colors from Fluxnet (red), ForestGEO (green), and RAINFOR (blue). Light-colored regions are well represented and dark-colored regions are poorly represented by the sampling networks.

Acknowledgments



U.S. DEPARTMENT OF
ENERGY

Office of Science



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 Oak Ridge Leadership Computing Facility (OLCF) 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.*, 28(8):1567–1586, Oct. 2013. doi: 10.1007/s10980-013-9902-0.
- M. Keller, D. Schimel, W. Hargrove, and F. Hoffman. A continental strategy for the National Ecological Observatory Network. *Front. Ecol. Environ.*, 6(5):282–284, June 2008. doi: 10.1890/1540-9295(2008)6[282:ACSFNT]2.0.CO;2. Special Issue on Continental-Scale Ecology.
- D. Schimel, W. Hargrove, F. Hoffman, and J. McMahon. NEON: A hierarchically designed national ecological network. *Front. Ecol. Environ.*, 5(2):59, Mar. 2007. doi: 10.1890/1540-9295(2007)5[59:NAHDNE]2.0.CO;2.