

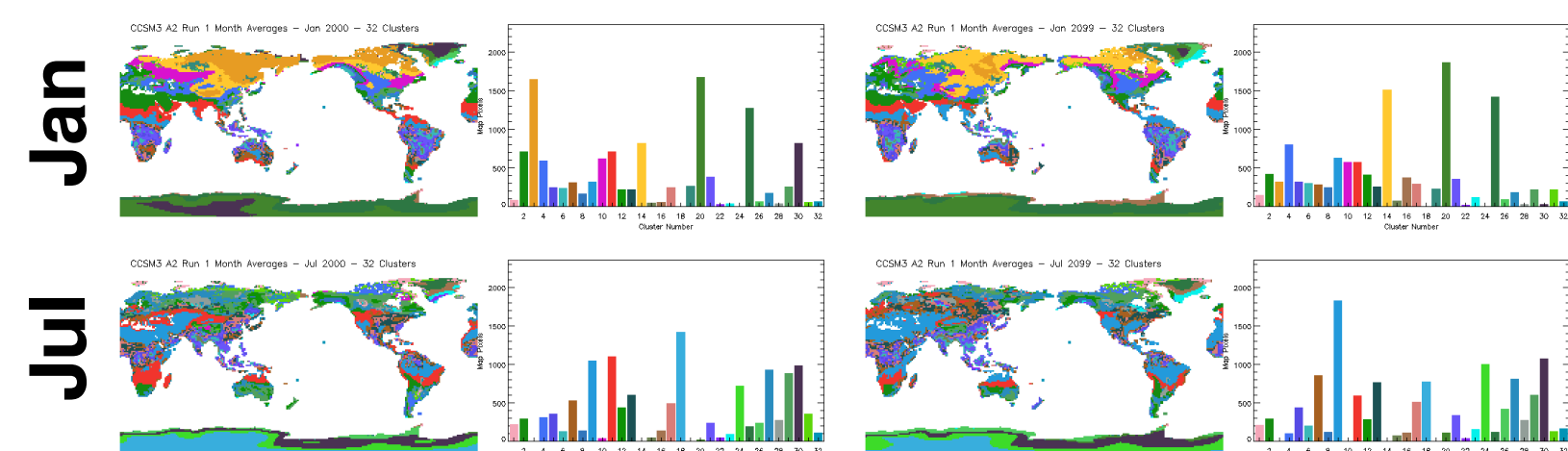
# Large Scale Climate Data Analytics

Forrest M. Hoffman, Jitendra Kumar, William W. Hargrove, and Richard T. Mills

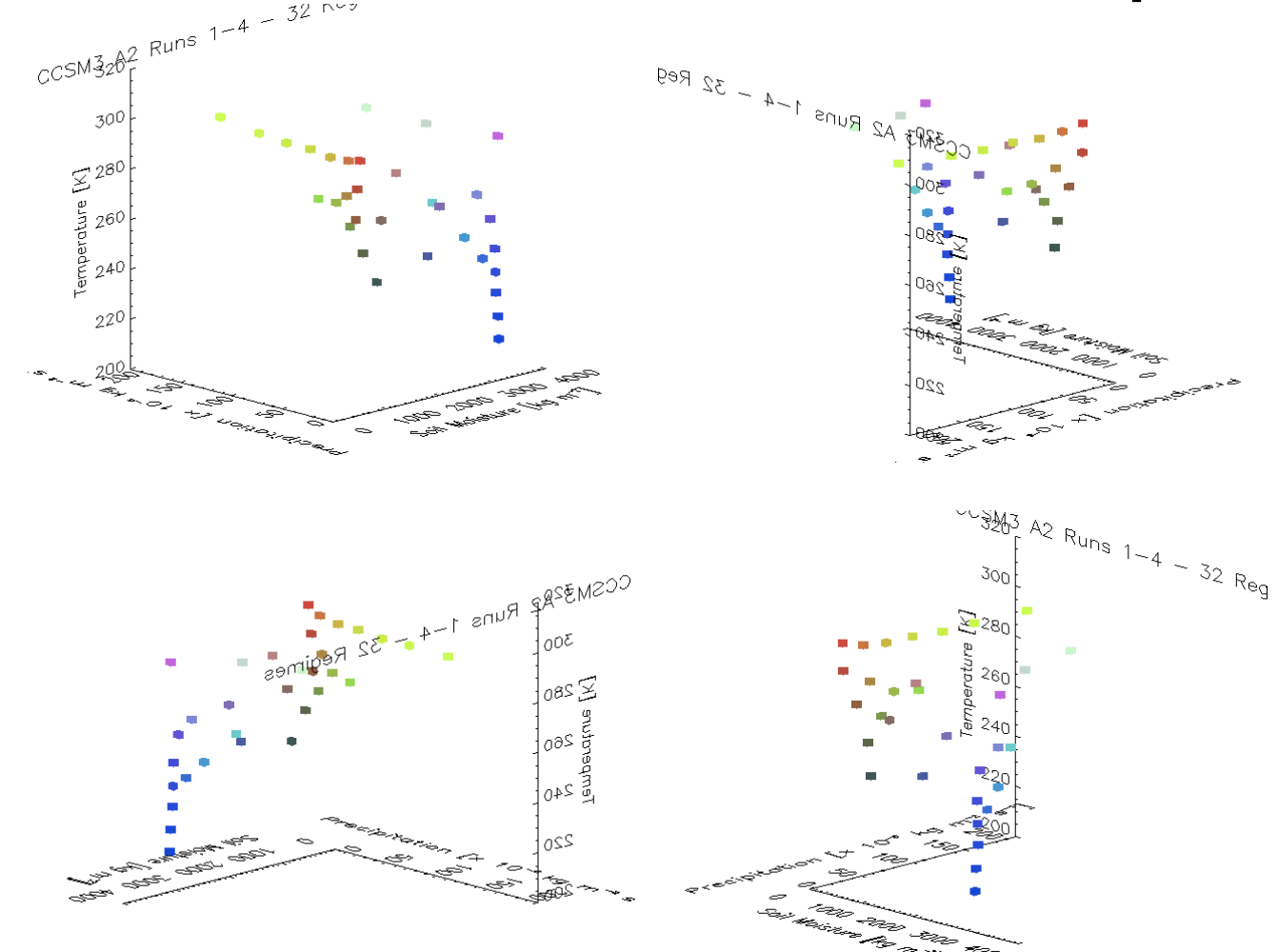
## Coupled Climate Model Future Projection Analysis

- Cluster analysis makes large, multivariate time-series projections from Earth System Models understandable.
- Community Climate System Model (CCSM3) results for years 2000–2099 were analyzed.
- Temperature, precipitation, and soil moisture were used in unsupervised classification.

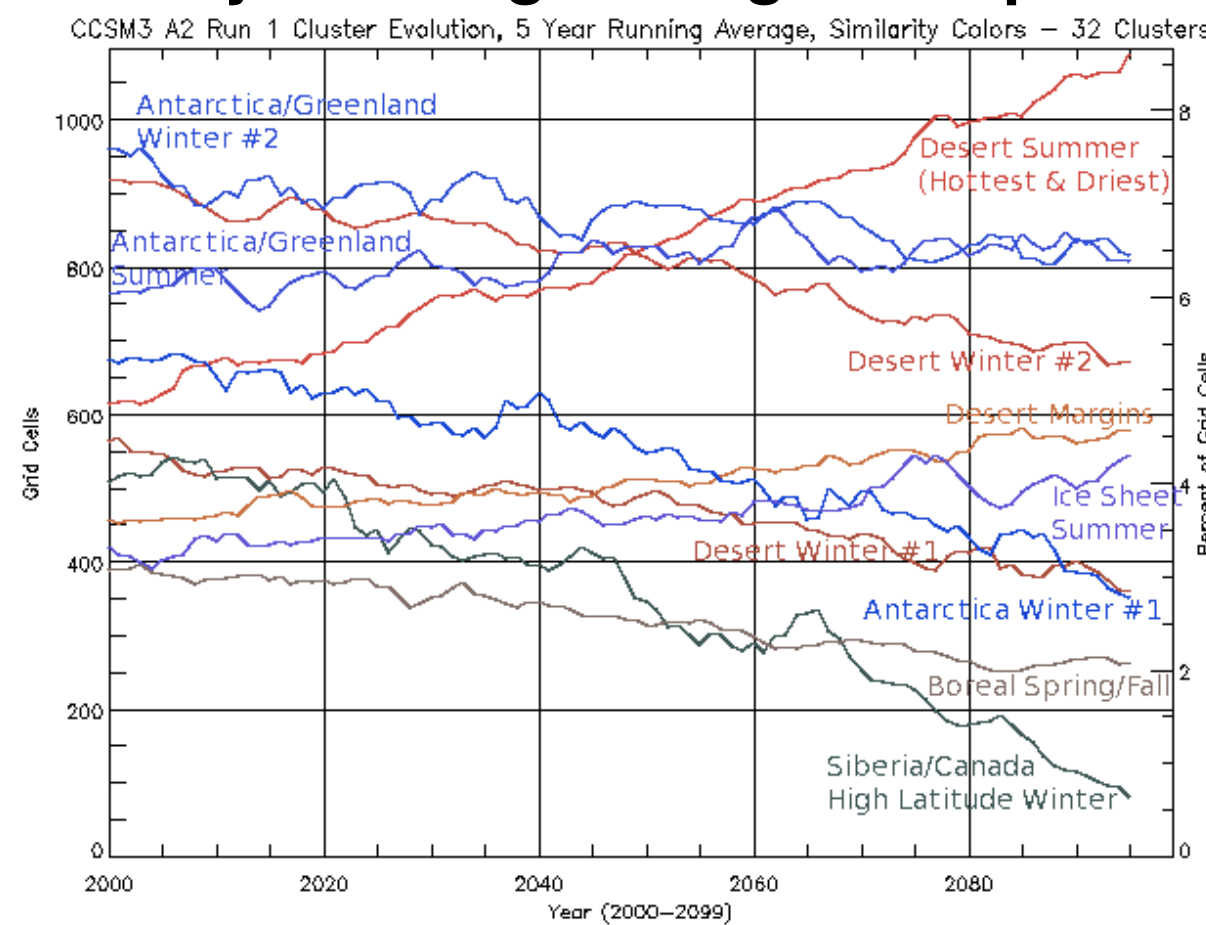
### Shifting Climate Regimes Defined Using Clustering



### Centroids Form a Skeleton in State Space



### Global 5-y Running Averages of Spatial Area



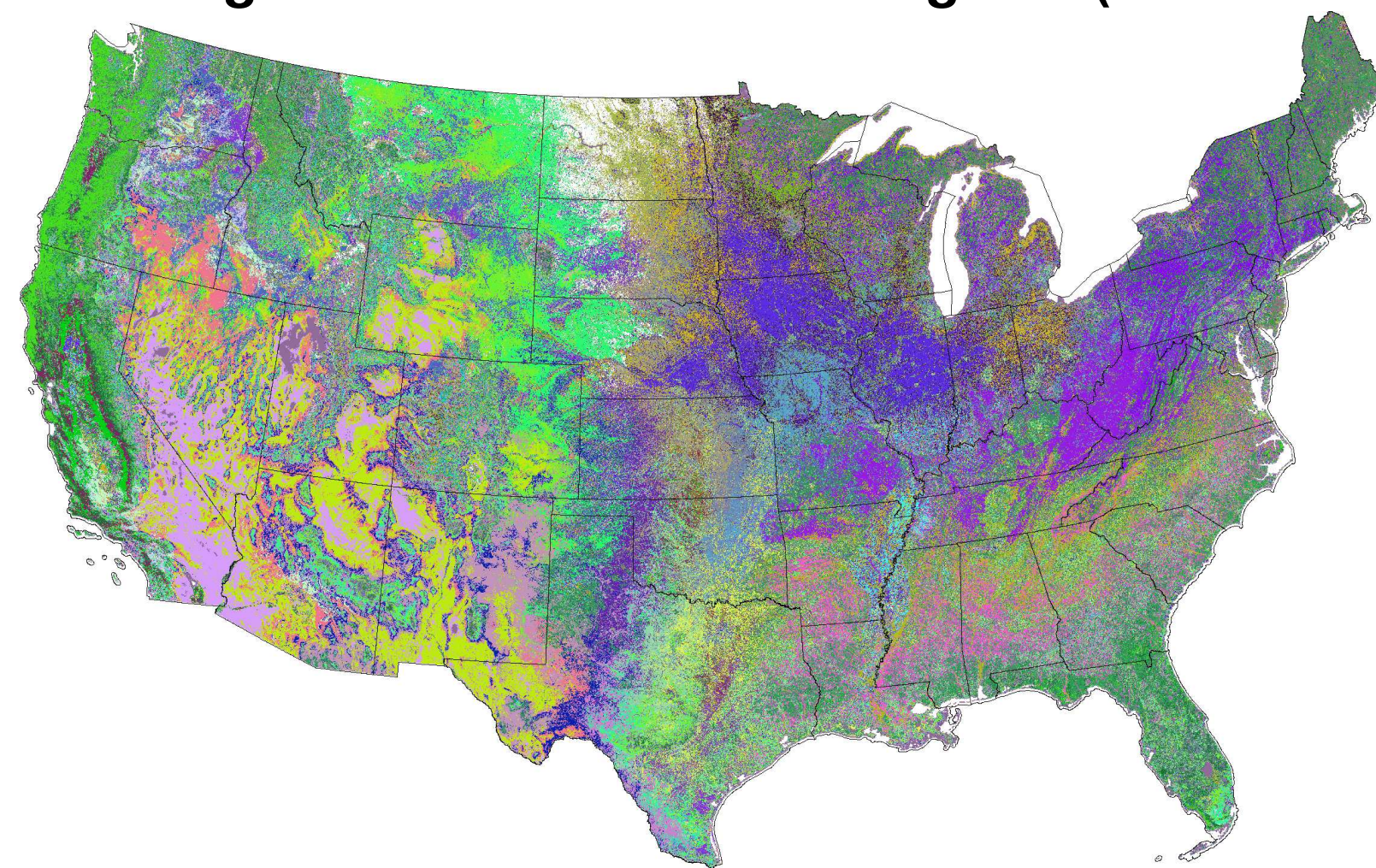
### CCSM3 Climate Regime Definitions

Cluster Number	Temperature [K]	Precipitation [ $\times 10^6 \text{ kg m}^{-2} \text{ s}^{-1}$ ]	Soil Moisture [ $\text{kg m}^{-3}$ ]	Name
-18	218.03	1.00	3148.15	Antarctica Winter #1 (Coldest and Driest)
24	229.01	3.00	3148.08	
-30	239.94	5.00	3142.49	Antarctica/Greenland Winter #2
-3	249.57	8.00	1000.00	Siberia/Canada High Latitude Winter
+20	250.13	5.00	3131.96	Antarctica/Greenland Summer
26	250.16	23.00	3074.85	
19	254.71	10.00	2000.09	
14	260.99	13.00	852.52	
+25	262.97	7.00	3118.23	Ice Sheet Summer Coastal Margins
18	266.14	24.00	3116.24	
10	268.99	27.00	849.38	
-1	272.12	16.00	2348.85	
-31	272.51	16.00	1281.38	Boreal Spring/Fall
4	273.47	13.00	704.89	
23	274.05	52.00	2915.14	
8	278.24	64.00	886.54	
12	278.51	42.00	845.17	
29	281.45	28.00	747.28	
-2	285.59	5.00	542.59	Desert Winter #1
27	286.52	17.00	678.39	
28	289.53	22.00	1640.25	
32	294.04	15.00	3373.13	
-11	295.05	3.00	547.47	Desert Winter #2
21	298.00	97.00	814.77	
13	298.05	35.00	677.27	
17	298.31	54.00	741.58	
22	298.66	99.00	2861.80	
5	298.68	74.00	784.84	
6	298.82	124.00	866.06	
+7	299.73	18.00	650.07	Desert Margins
15	299.76	168.00	1029.57	
+9	303.06	2.00	583.05	Desert Summer (Hottest and Driest)

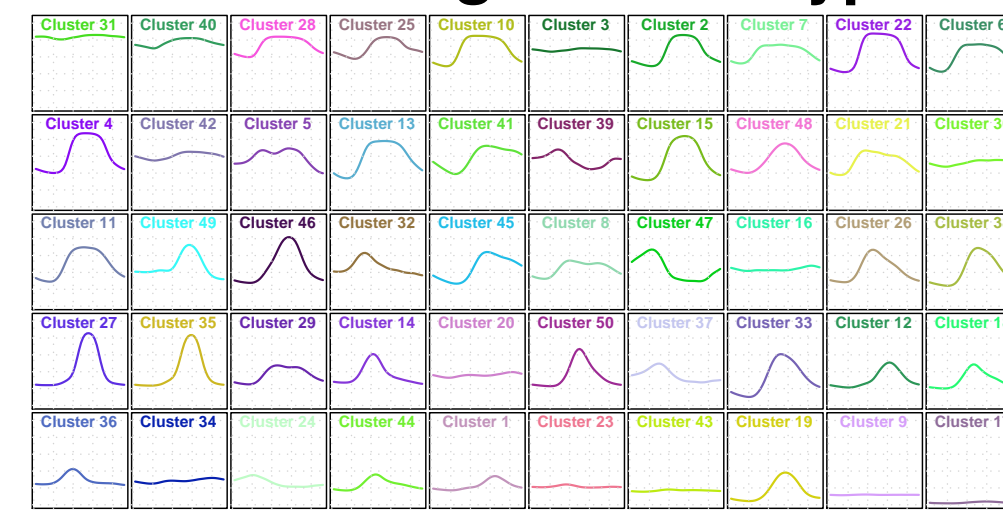
## Data Mining for Detecting Threats to U.S. Forest Health

- USDA Forest Service, NASA, DOE ORNL, and USGS developed an early warning system for forest threats.
- ForWarn system uses phenology derived from NDVI observations from MODIS every 8 days.

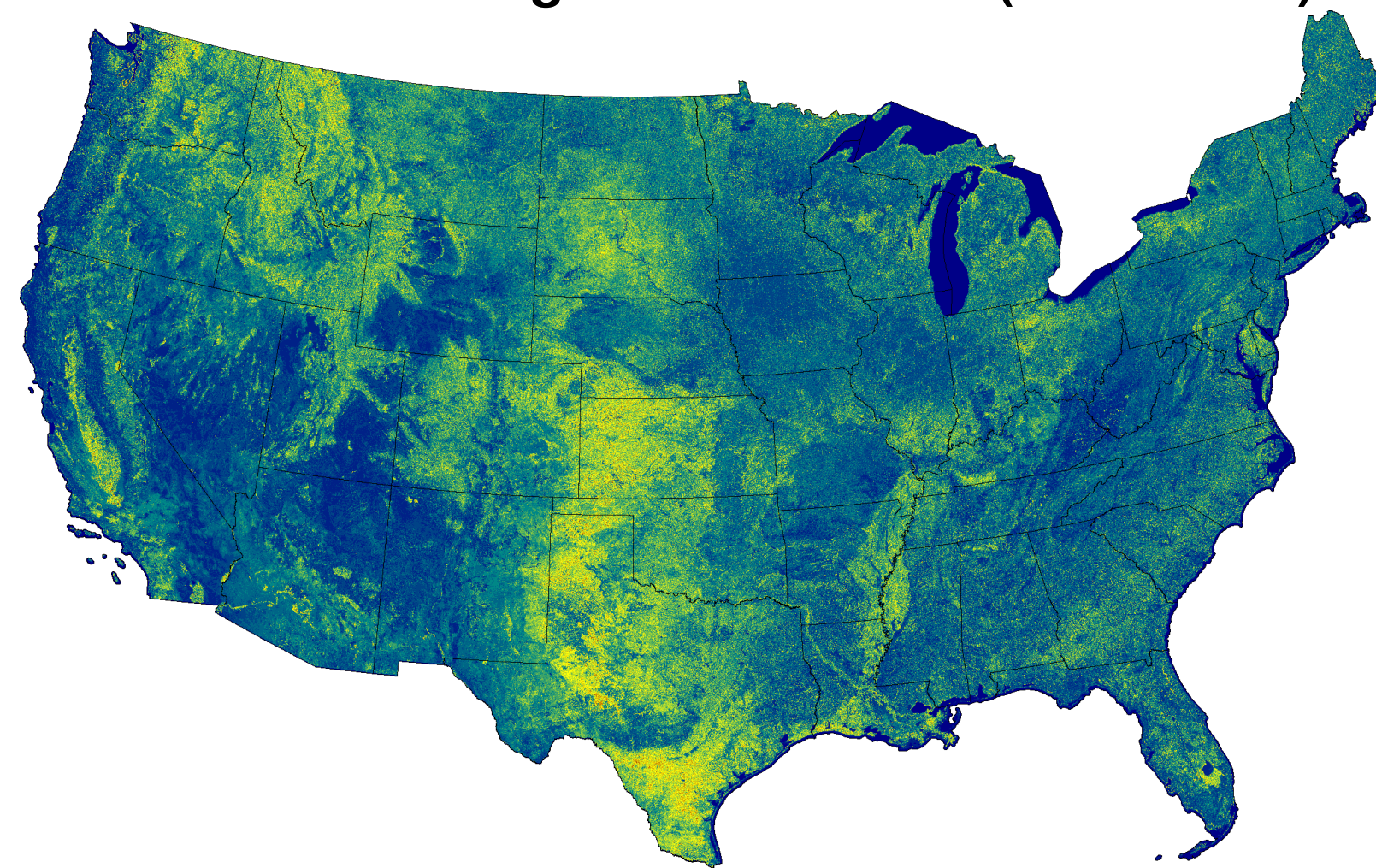
### Phenological Mode for 50 Phenoregions (2000–2011)



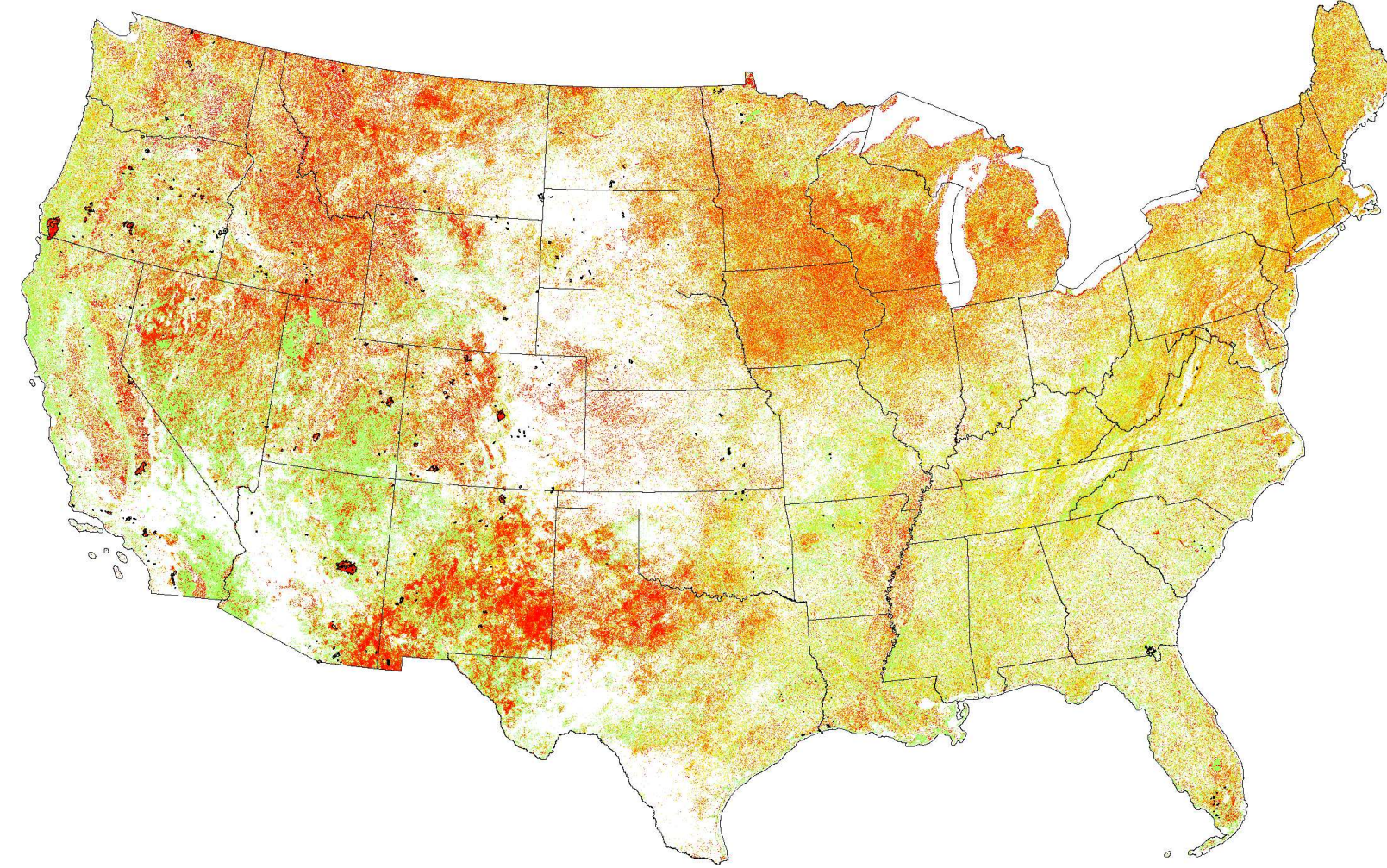
### 50 Phenoregion Prototypes



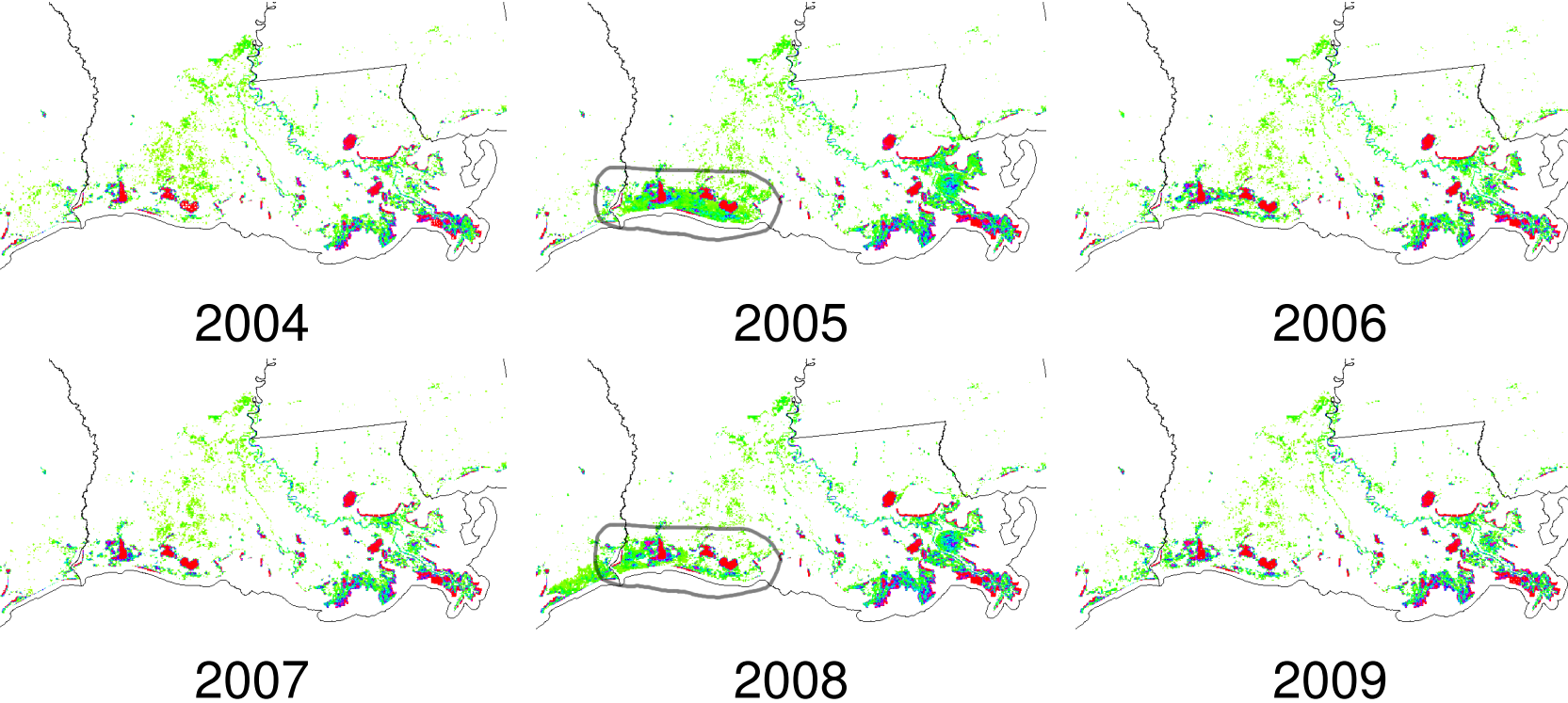
### Annual Phenoregion Persistence (2000–2011)



### Δ Integrated NDVI with MTBS Fires Identified



### Detecting Hurricane-Driven Salt-Water Infiltration Using PCA



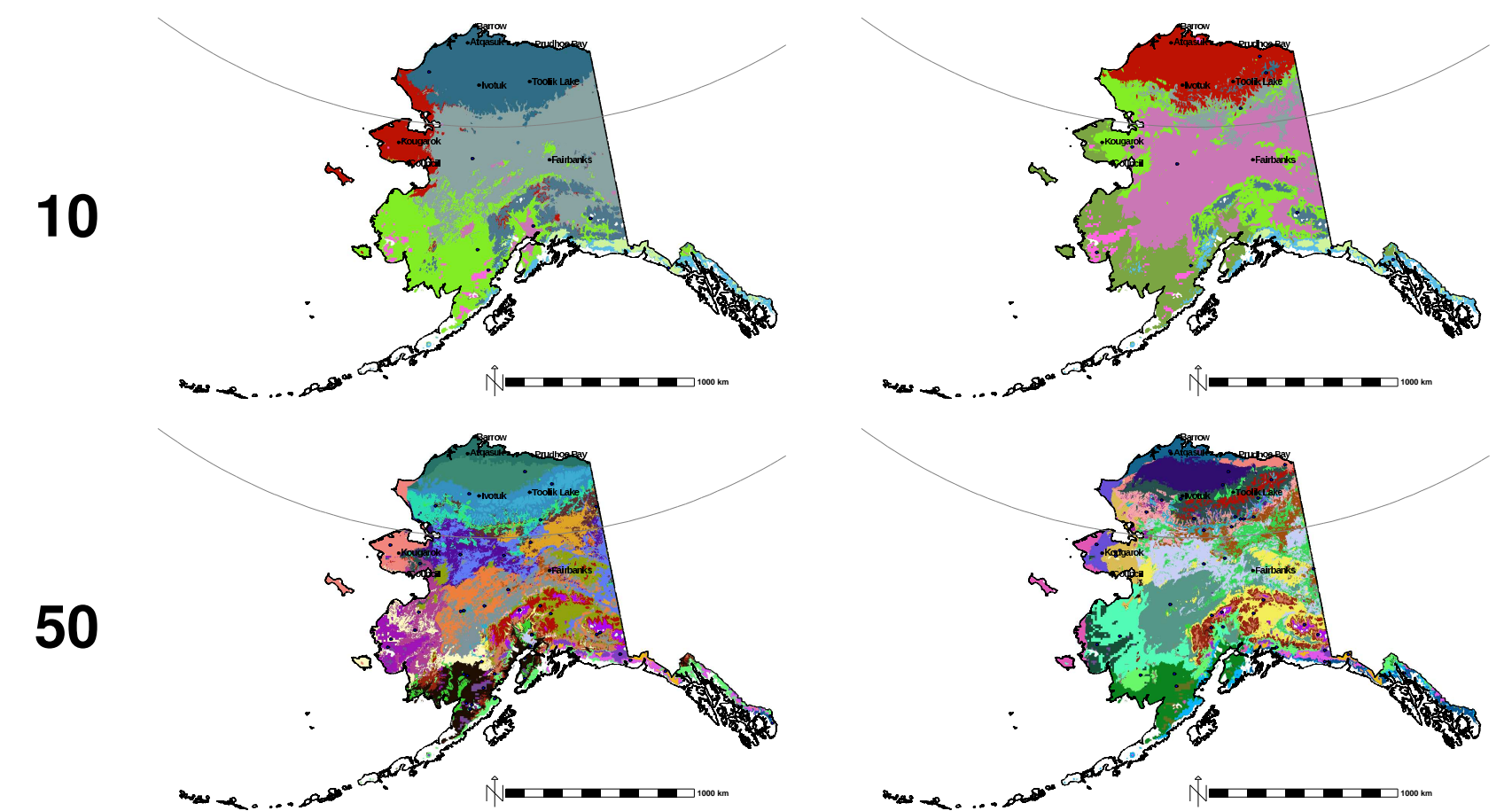
## Next Generation Ecosystem Experiments (NGEE)

- NGEE is a model-inspired field measurement program focused on the Arctic and other critical regions.
- Quantitative methodology developed for stratifying domains and determining representativeness of sites.

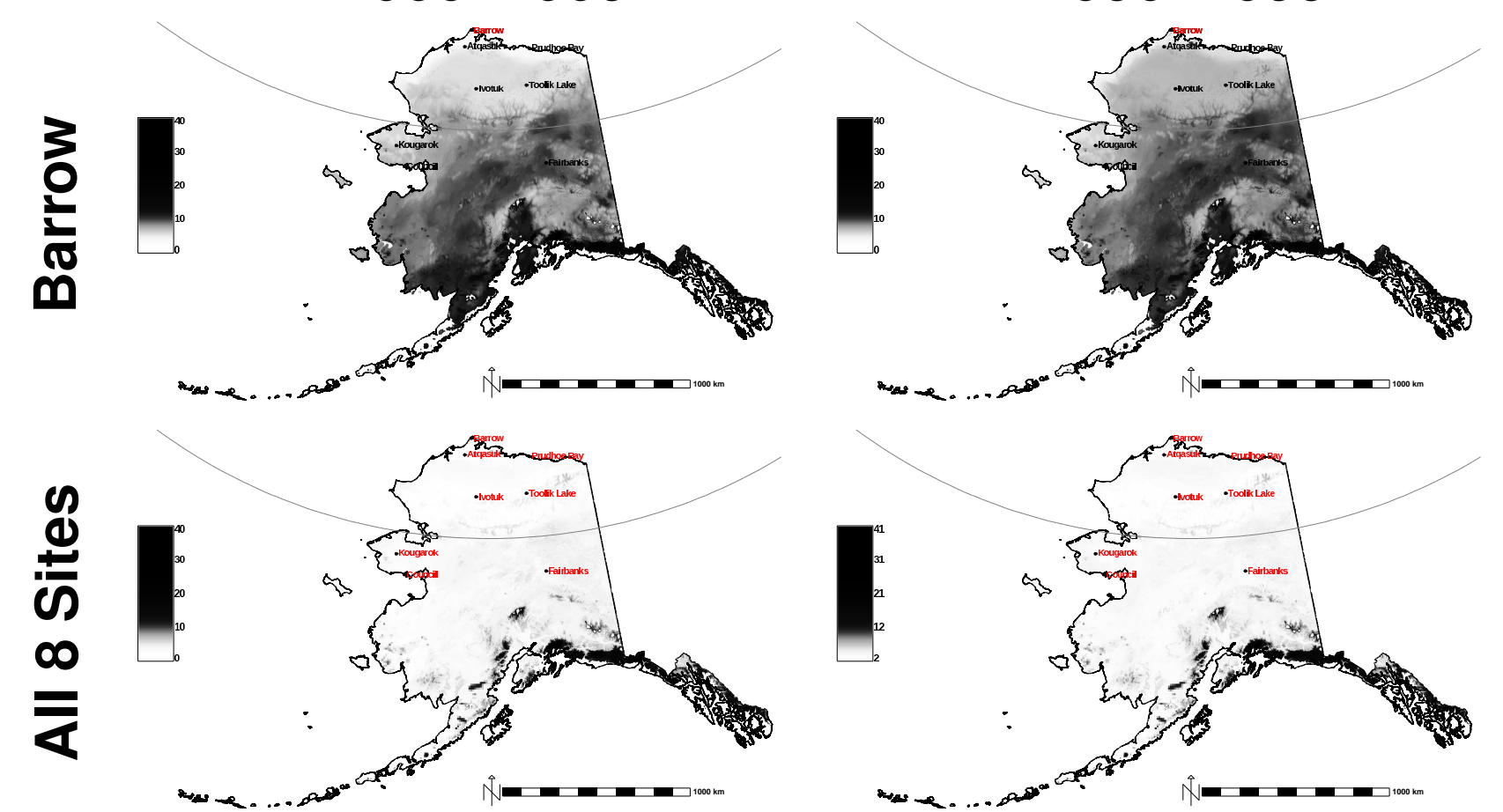
### 37 Characteristics for the State of Alaska

Description	Number or 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	SRTM30

### Ecoregions and Representative Sites 2000–2009



### Site and Network Representativeness 2000–2009



### Site Dissimilarity Between Present and Future

Sites	Future (2090–2099)								
	Barrow	Council	Atqasuk	Ivotuk	Lake Kougarak	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
	Kougarak	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

Dissimilarity values are calculated as the Euclidean distance between sites in the 37-dimensional state space.

Hoffman, F. M., J. Kumar, R. T. Mills, and W. W. Hargrove (2012) "Representativeness-Based Sampling Network Design for the Arctic." *Landscape Ecol.*, in review.