System for Mapping and Predicting Species of Concern

John Olson, Emily Campbell, Jessie Doyle, Arev Markarian

California State University, Monterey Bay

Chuck Hawkins, Torrey Rodgers, Karen Mock, James Eddings

Utah State University

Matthew Varner, Matthew Whitman

Alaska BLM

California State University MONTEREY BAY

S.J. & JESSIE E. QUINNEY COLLEGE of NATURAL RESOURCES UtahStateUniversity

NATIONAL SYSTEM OF PUBLIC LANDS U.S. DEPARTMENT OF THE INTERIOR BUREAU OF LAND MANAGEMENT

National Petroleum Reserve - Alaska

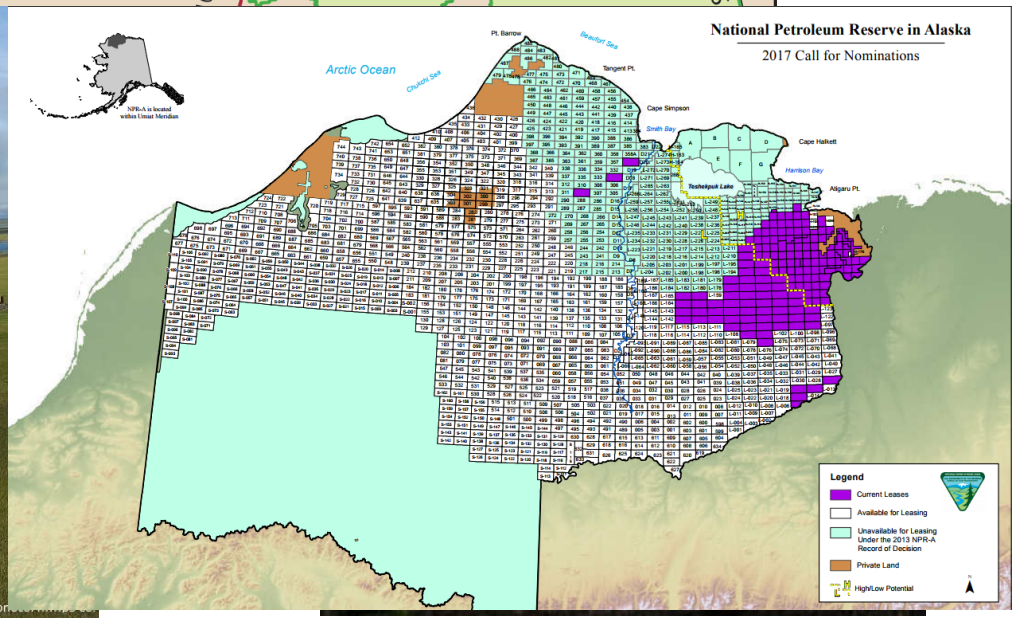
23 million acres

1000s of miles of streams

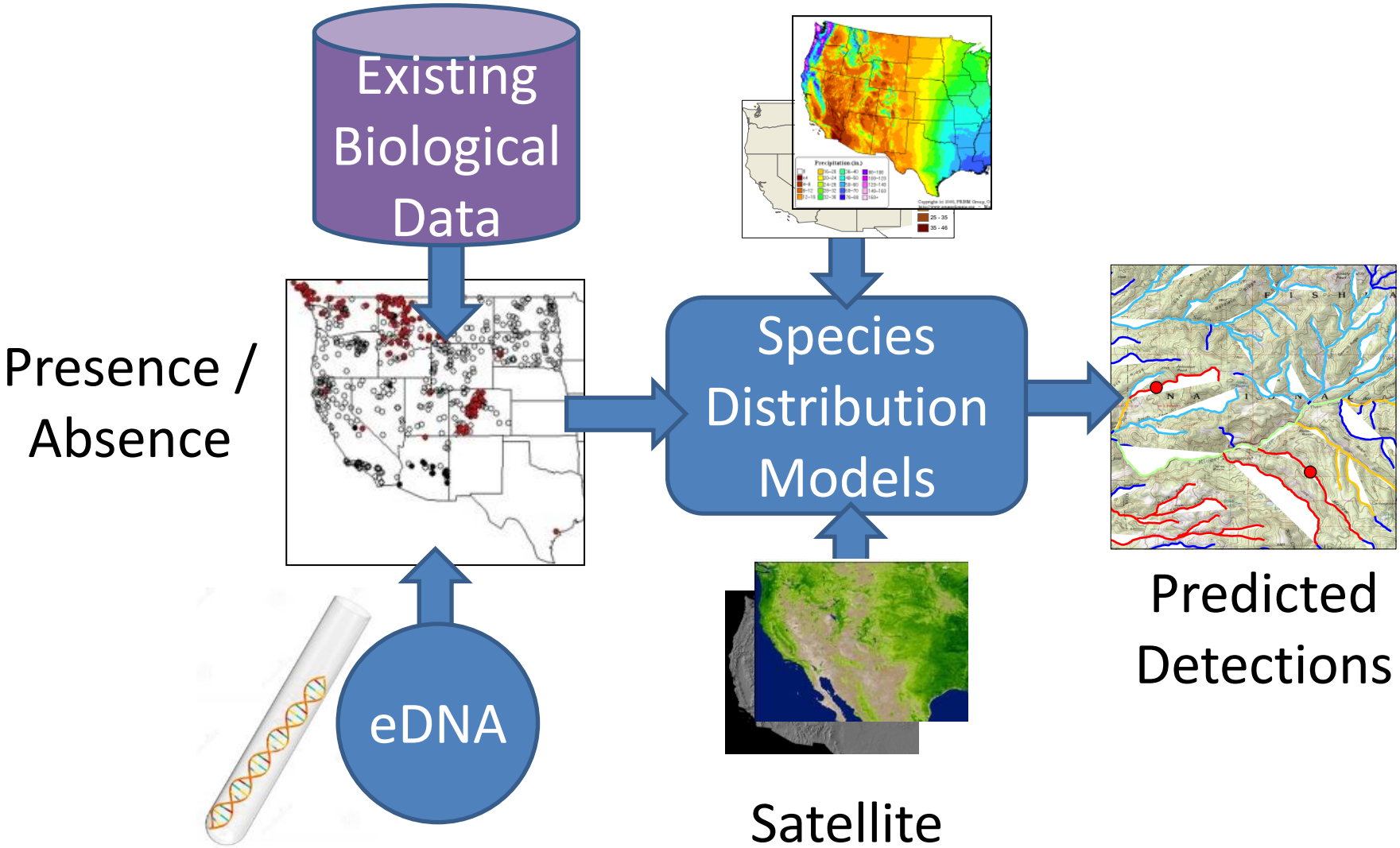
Being rapidly developed

Oil: 15^9 barrel

Gas: 45^{12} ft^3



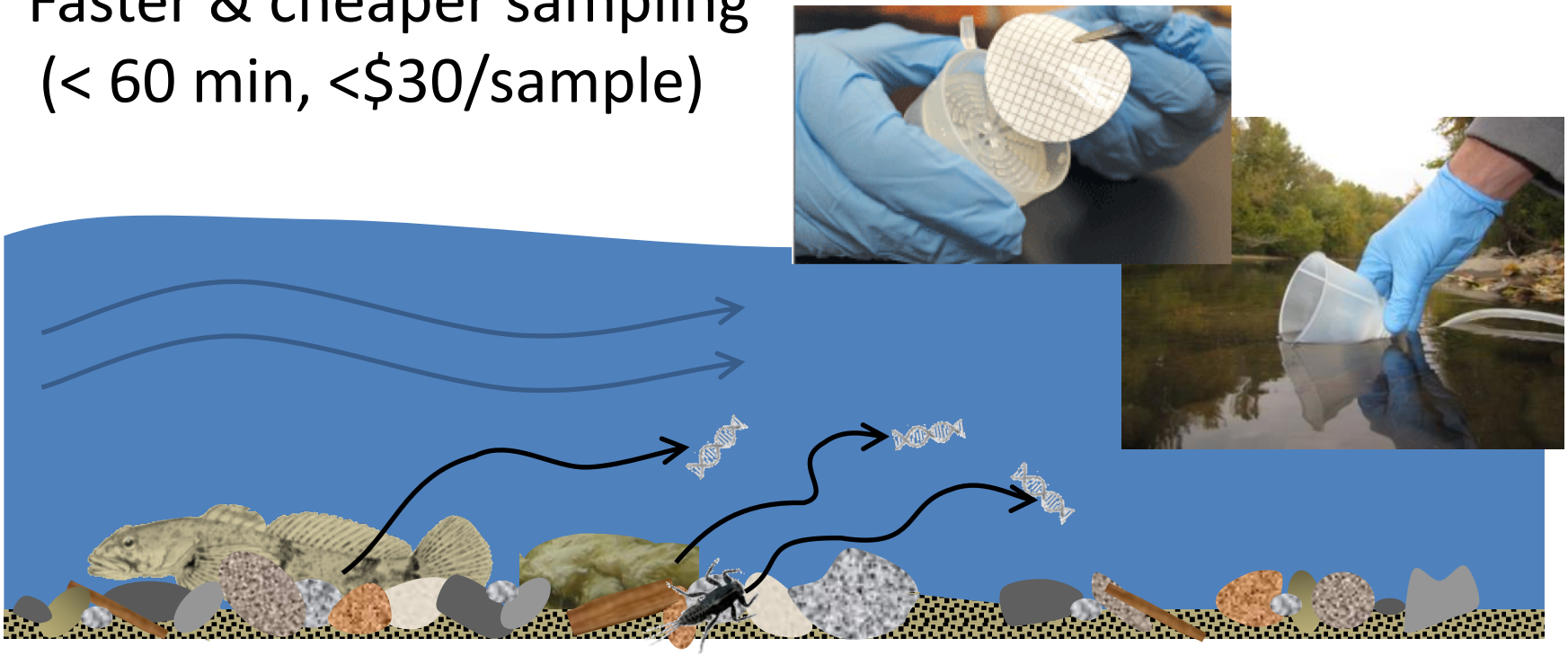
System for Mapping and Predicting Species of Concern



System for Mapping and Predicting Species of Concern

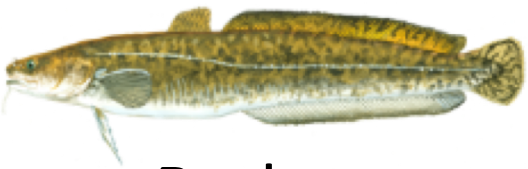
Environmental DNA :

- COI assessed using qPCR
- Detection rates 80-96%
- Faster & cheaper sampling (< 60 min, <$30/sample)

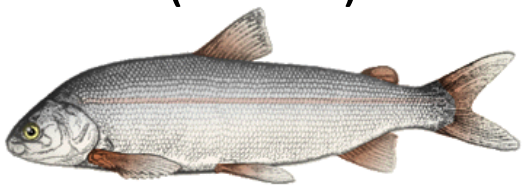


Develop & apply eDNA assays

RNase H–dependent PCR
(rhPCR)



Burbot
(*Lota lota*)



Broad Whitefish
(*Coregonus nasus*)

Chum Salmon
(*Oncorhynchus keta*)

N. Pike
(*Esox lucius*)

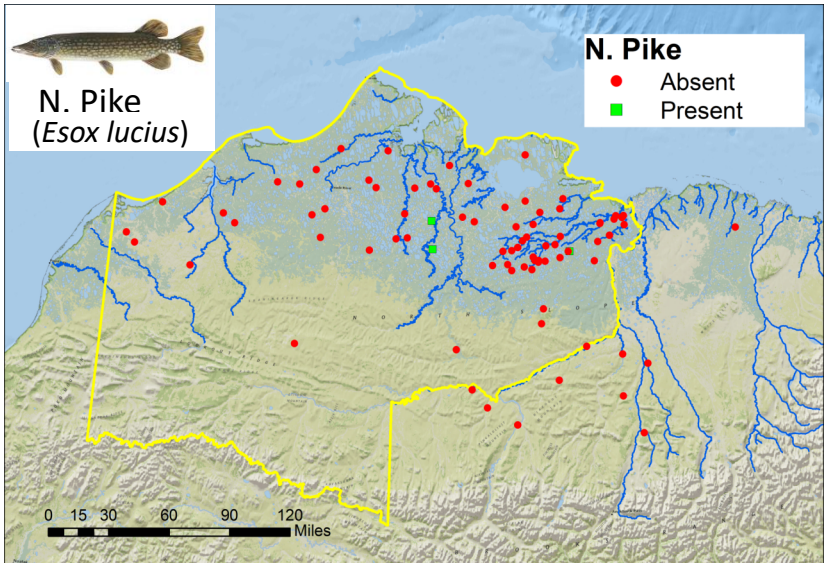
Least Cisco
(*Coregonus sardinella*)

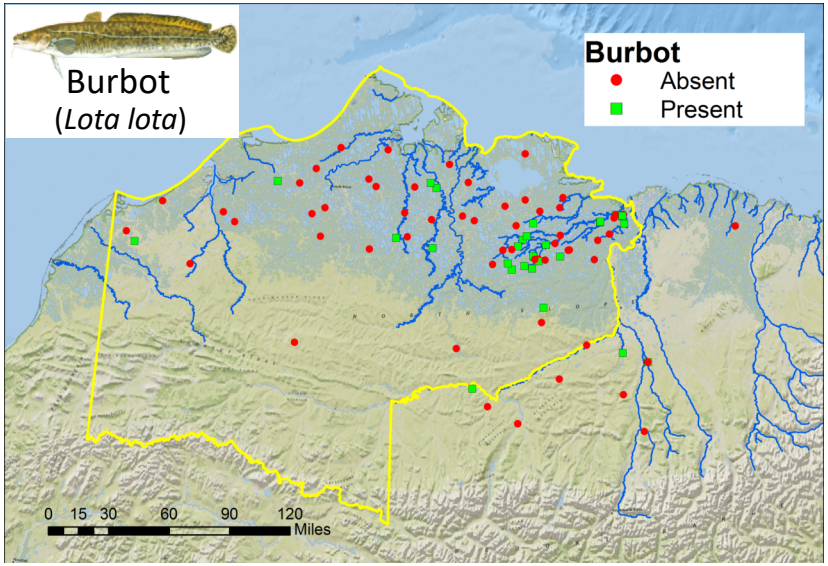
Chinook Salmon
(*Oncorhynchus tshawytscha*)

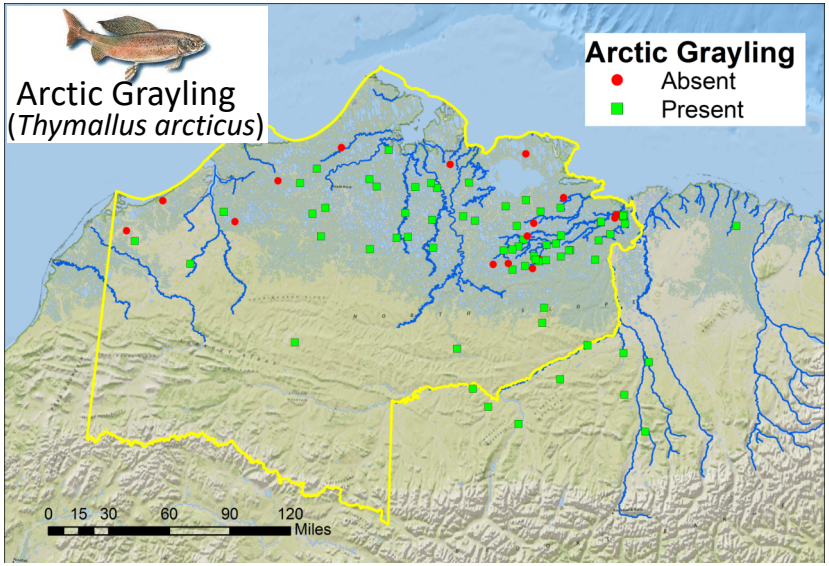
Arctic Grayling
(*Thymallus arcticus*)

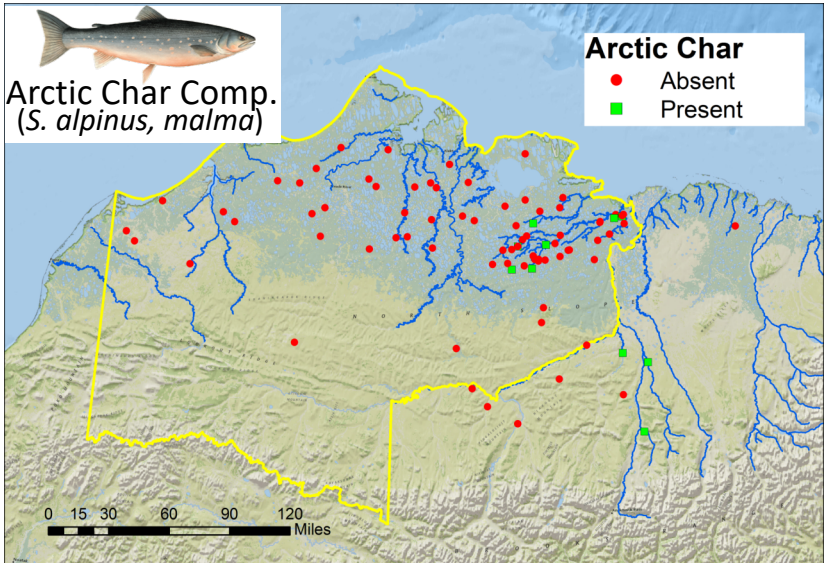
Arctic Char Complex
(*Salvelinus alpinus, malma*)

Develop & apply eDNA assays







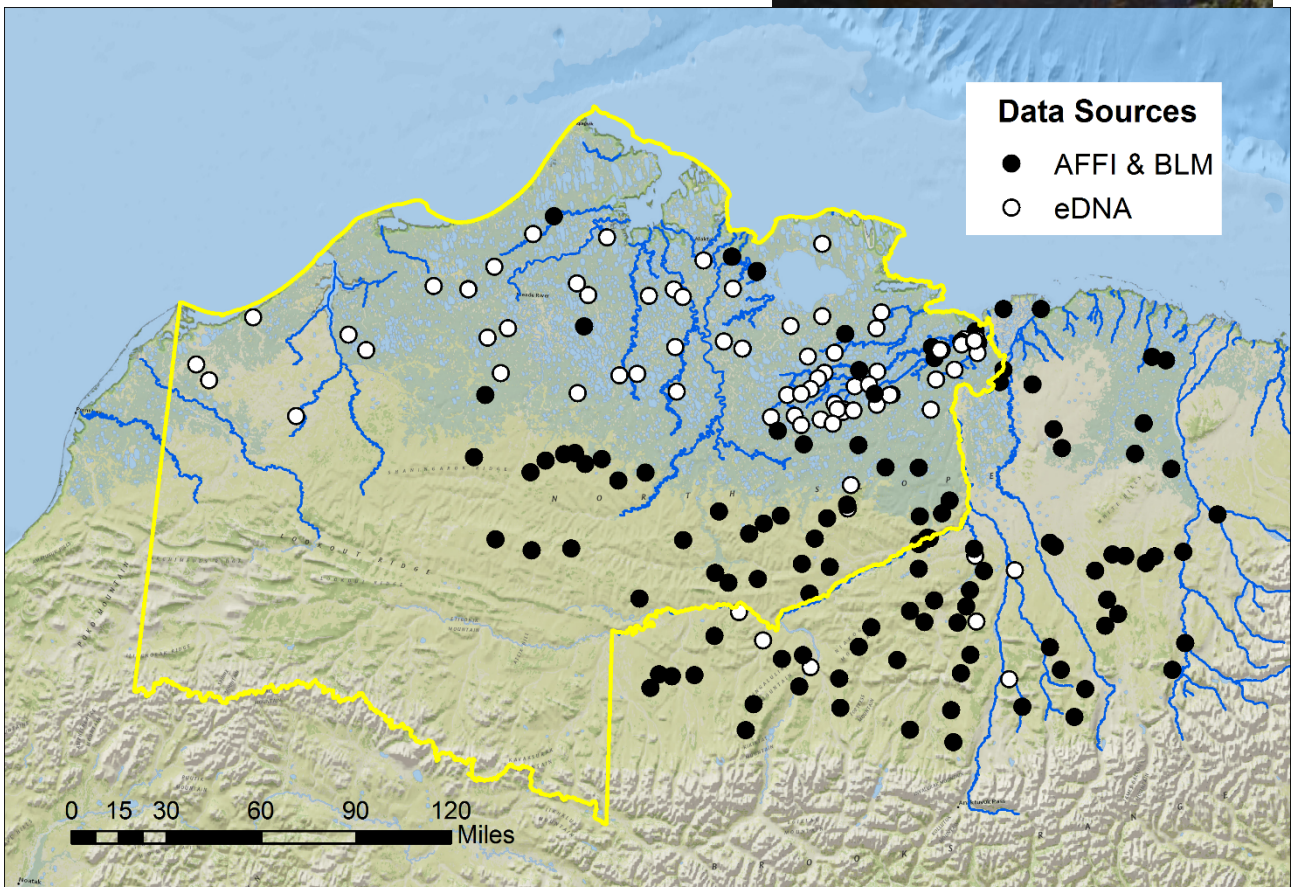


Combine with traditional data

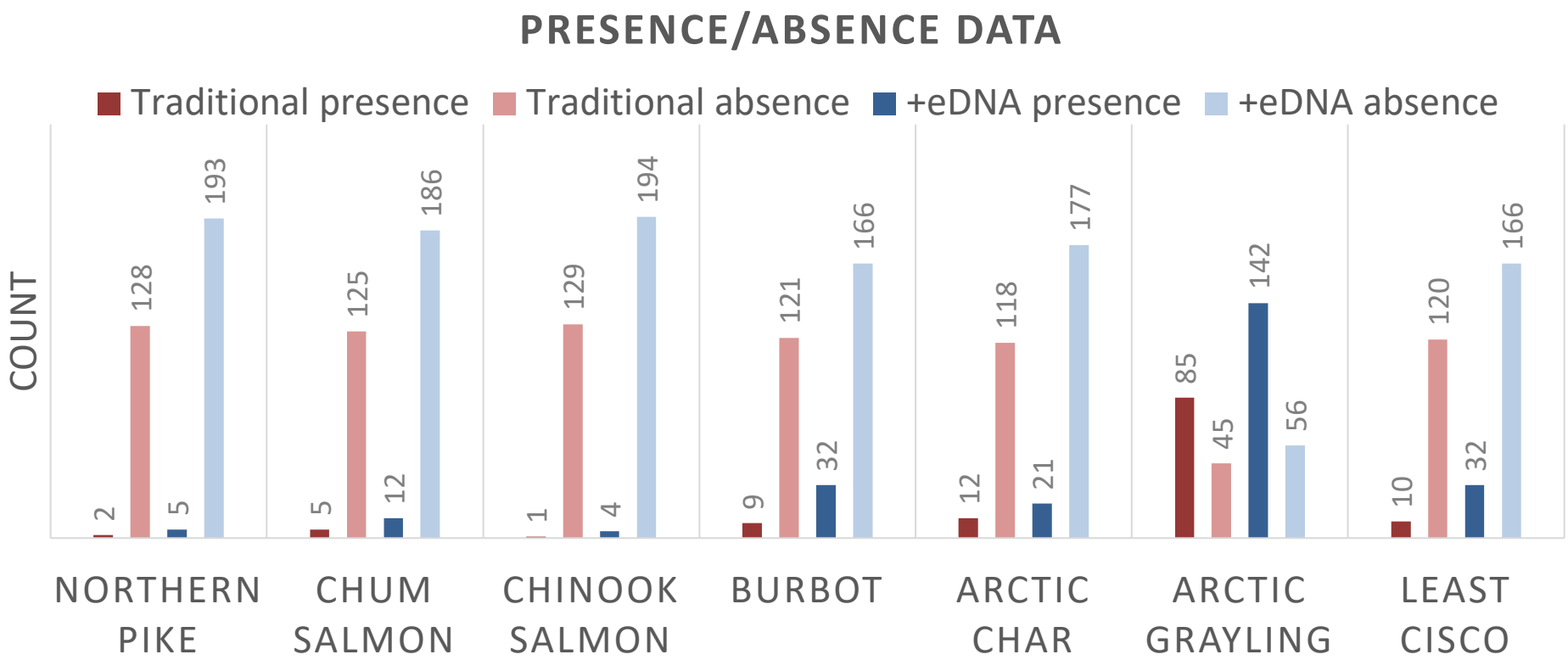
68 eDNA sites + 130 sites sampled 2001 – 2016 using traditional methods

Alaska Freshwater Fish Inventory Database - 100 sites

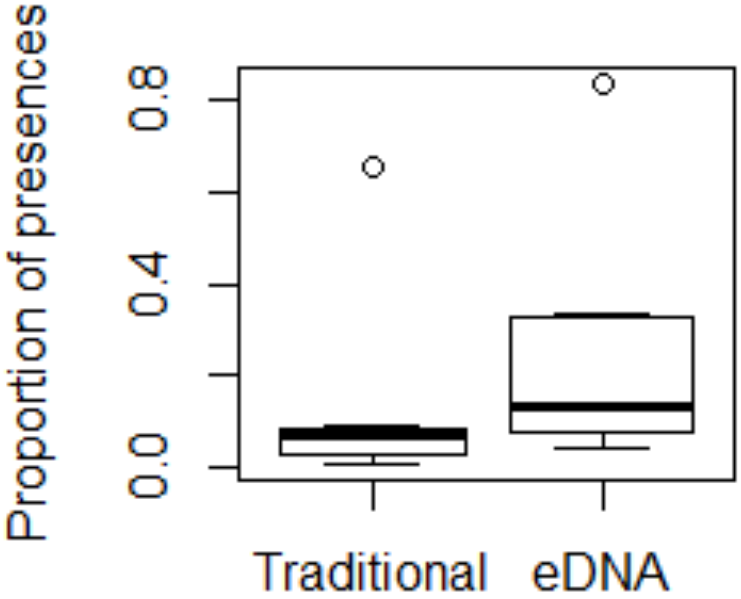
BLM Assessment, Inventory, & Monitoring – 30 sites



Combine with traditional data



Combine with traditional data



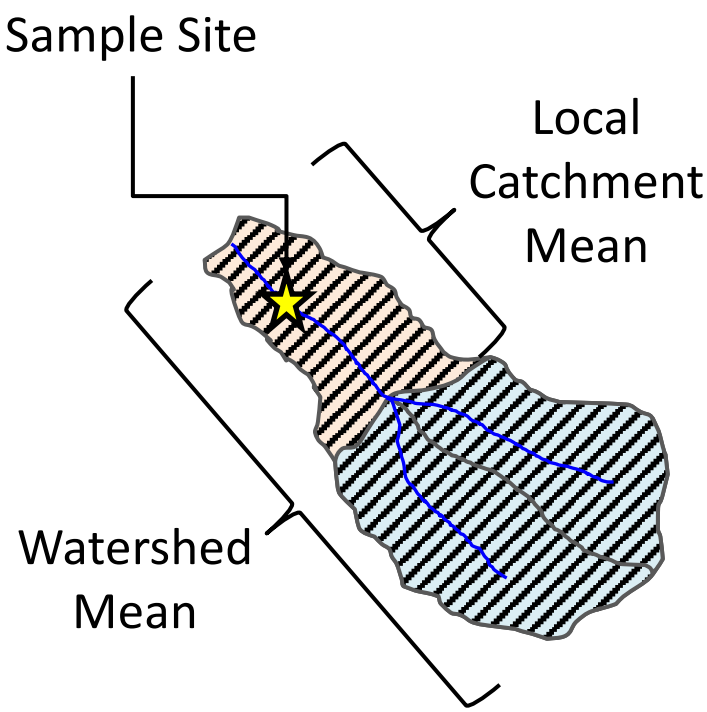
Watershed Earth Observations

Static Predictors
Unfrozen Lakes (SAR)
Drainage Area
Stream Slope
Soil Characteristics
Vegetation Type (Landsat)
Thermokarst Activity
Long-term Averages of Dynamic Predictors

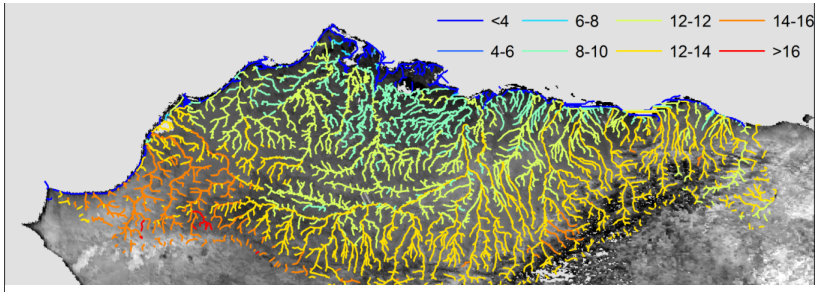
Dynamic Predictors
Active Layer Thickness
Evapotranspiration (MODIS)
Land Surface Temperature (MODIS)
EVI/GPP (MODIS)
Fire Activity (MODIS)
Oil & Gas Development
Snow cover (MODIS)

Watershed Earth Observations

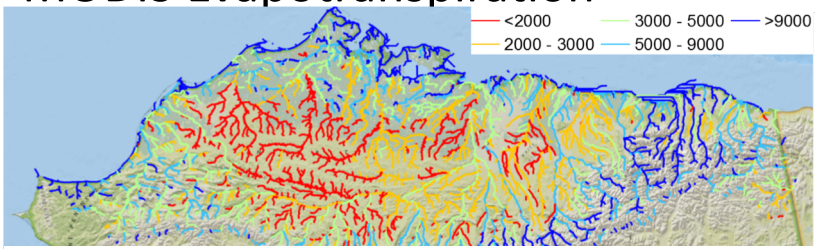
Summarize Earth observation data using StreamCat



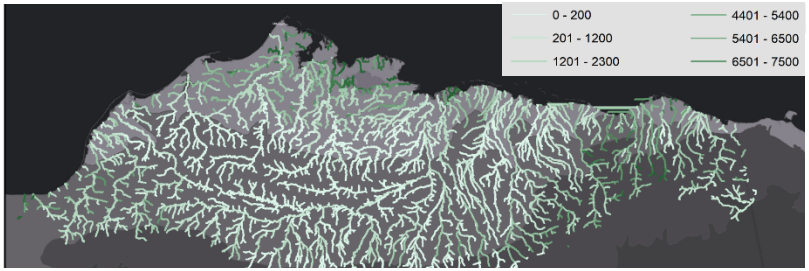
MODIS Land Surface Temperature



MODIS Evapotranspiration



MODIS Gross Primary Production



Species Distribution Models

Developed two types of models to predict probability of detection from local and watershed conditions:

1. Presence/Absence using Random Forests (RF) – weighting minority classes
2. Presence only using Maximum Entropy (MaxEnt)

Species Distribution Models

Cross validation model performance using both traditional and eDNA data

	n		AUC		CCR	
Species Name	Presence	Absence	MaxEnt	RF	MaxEnt	RF
Arctic Grayling	142	56	0.53	0.41	0.59	0.68
Arctic Char Complex	21	177	0.77	0.83	0.81	0.94
Chinook Salmon	4	194	0.84	0.79	0.59	0.96
Burbot	32	166	0.87	0.71	0.91	0.66
Least Cisco	32	166	0.91	0.84	0.92	0.69
Northern Pike	5	193	0.93	0.78	0.92	0.94
Chum Salmon	12	186	0.96	0.80	1.00	0.85

AUC: Area Under the ROC Curve, CCR: Correct Classification Rate

Species Distribution Models

Cross validation model performance using traditional data only

	n		AUC		CCR	
Species Name	Presence	Absence	MaxEnt	RF	MaxEnt	RF
Slimy Sculpin	32	98	0.56	0.63	0.67	0.71
Longnose Sucker	3	127	0.98	0.65	0.88	0.96
Ninespine Stickleback	39	91	0.87	0.78	0.88	0.71
Round Whitefish	10	120	0.95	0.73	0.96	0.90
Broad Whitefish*	13	117	0.94	0.92	0.94	0.88
Alaska Blackfish	8	122	0.96	0.90	0.92	0.94
Bering Cisco	7	123	0.97	0.83	0.98	0.94
Humpback Whitefish	10	120	0.98	0.91	0.98	0.93

AUC: Area Under the ROC Curve, CCR: Correct Classification Rate

* - eDNA data to be included later

Species Distribution Models

Comparison of RF models using just eDNA vs. eDNA + Traditional

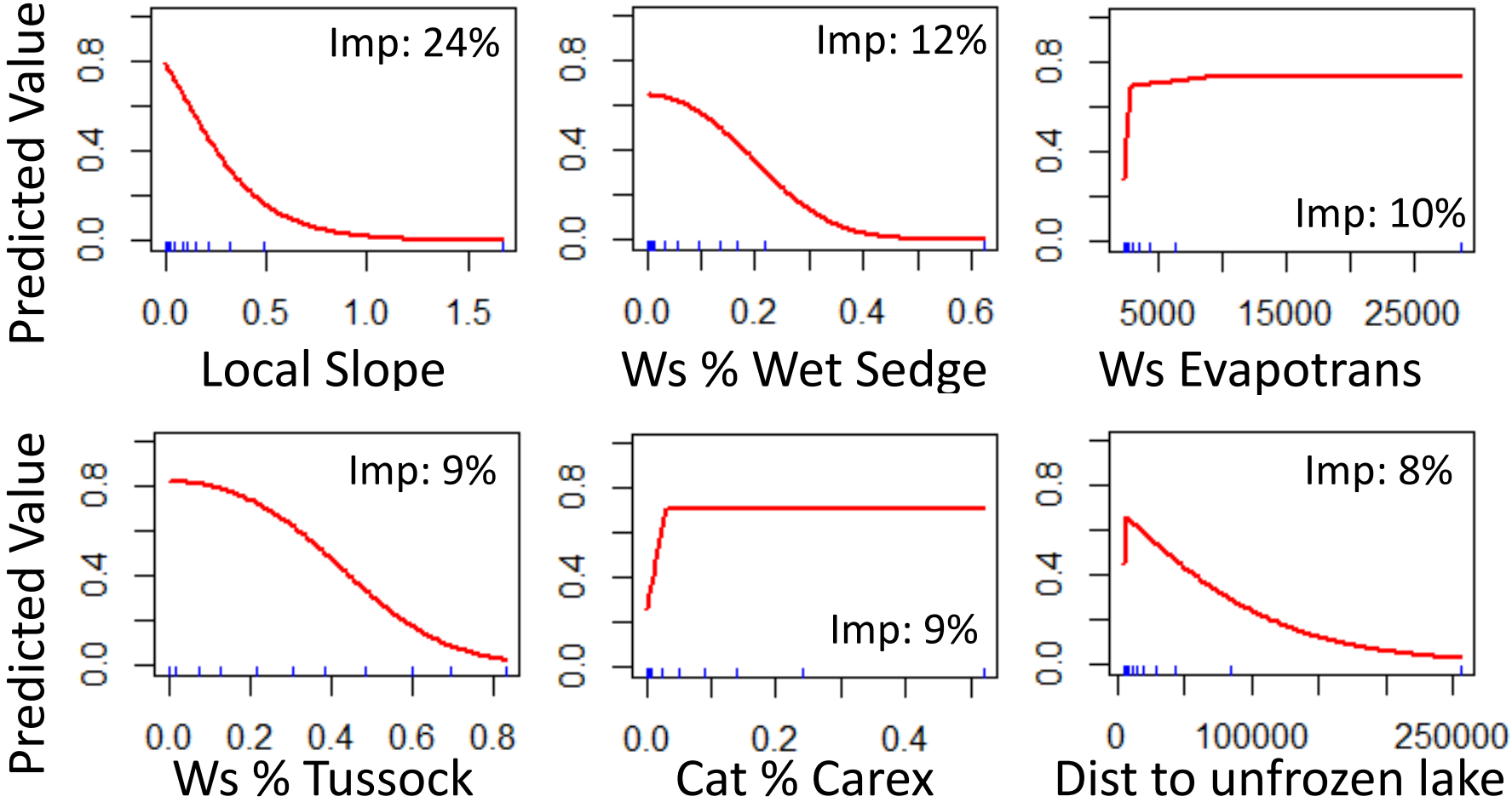
	eDNA					eDNA+Traditional				
Species Name	P	A	Sensitivity	Specificity	AUC	P	A	Sensitivity	Specificity	AUC
eDNA < eDNA+Traditional										
Burbot	23	45	88%	0%	0.36	32	166	67%	67%	0.71
Northern Pike	3	65	0%	98%	0.40	5	193	0%	97%	0.78
Chum Salmon	7	61	60%	70%	0.58	12	186	50%	88%	0.80
Arctic Char Comp.	9	59	43%	73%	0.63	21	177	70%	82%	0.83
Least Cisco	22	46	94%	21%	0.61	32	166	90%	65%	0.84
eDNA > eDNA+Traditional										
Arctic Grayling	57	11	96%	0%	0.78	142	56	89%	10%	0.43
Chinook Salmon	3	65	33%	92%	0.87	4	194	0%	100%	0.79

P: Presence, A: Absence, AUC: Area Under the ROC Curve

Species Distribution Models

Burbot
(*Lota lota*)

Effect of top six predictors in MaxEnt model

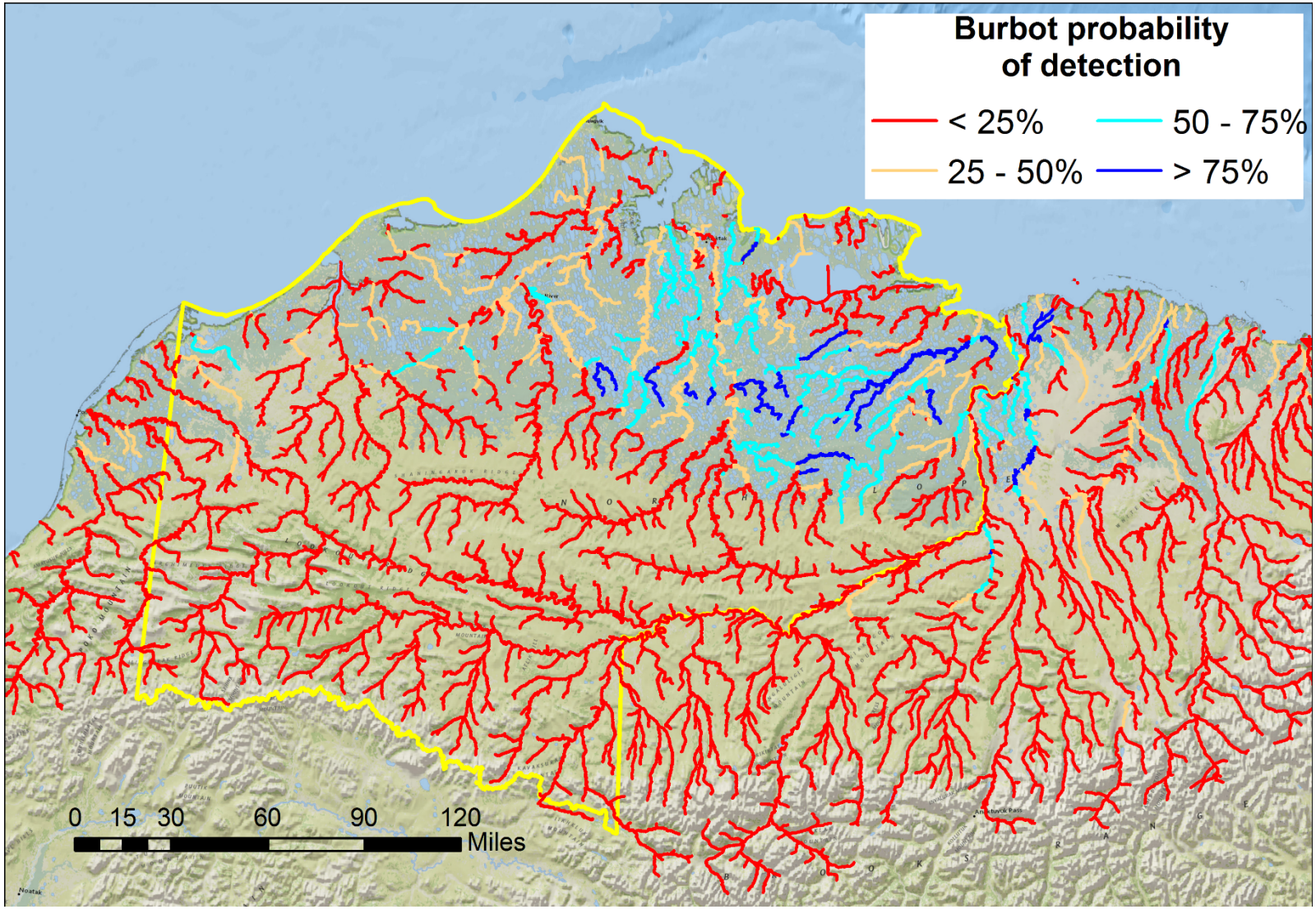


Imp: Variable importance as increase in error when permuted
Ws: Watershed average, Cat: Local catchment average

Species Distribution Models

Burbot
(*Lota lota*)

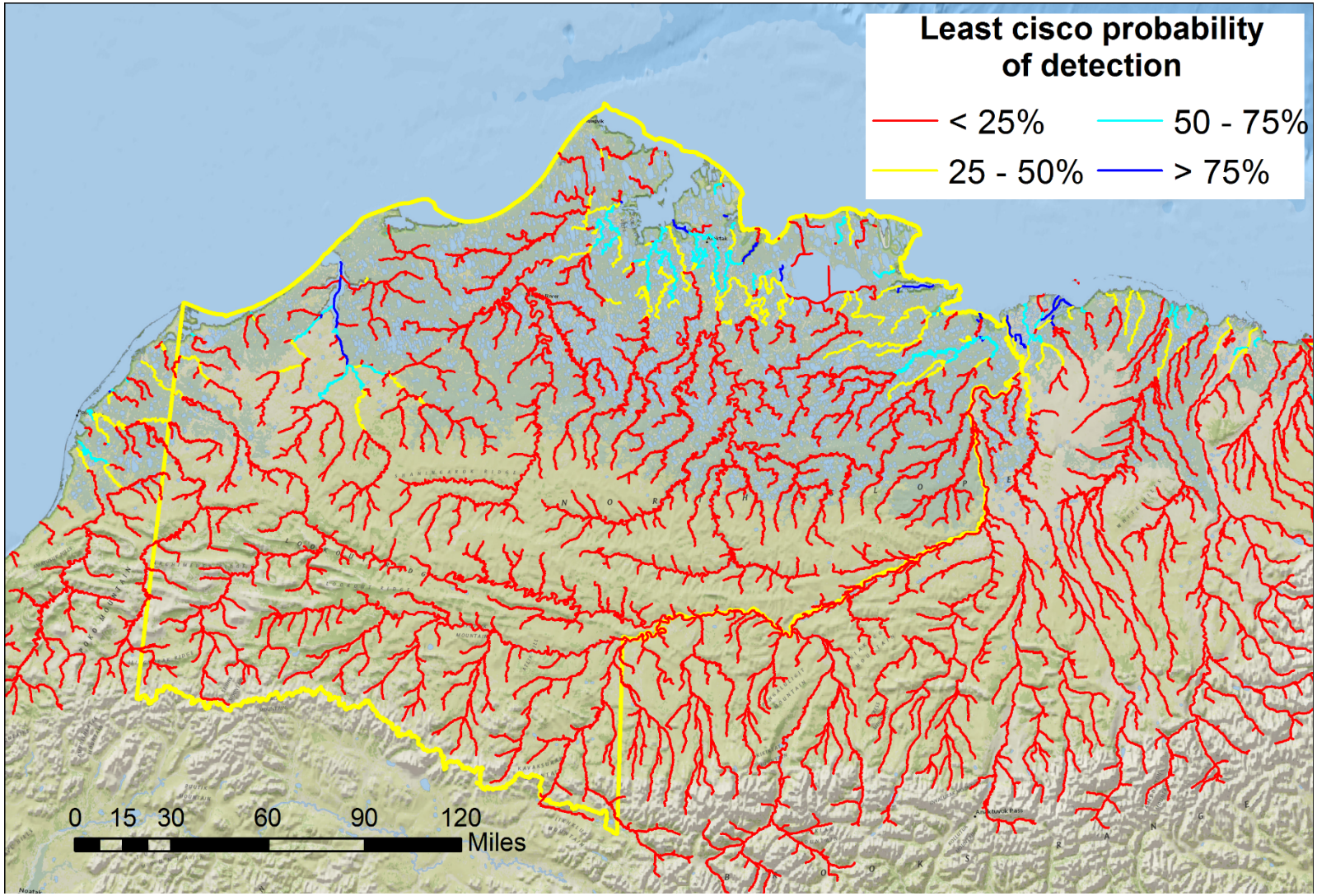
Apply MaxEnt model to all stream data



Species Distribution Models

Least Cisco
(*Coregonus sardinella*)

Apply MaxEnt model to all stream data



Next Steps

- Use dynamic (temporally specific) data in models
- Validate models using 2017-18 eDNA data
- Examine using ensemble model predictions
- Integrate map outputs into existing BLM decision support framework

Questions