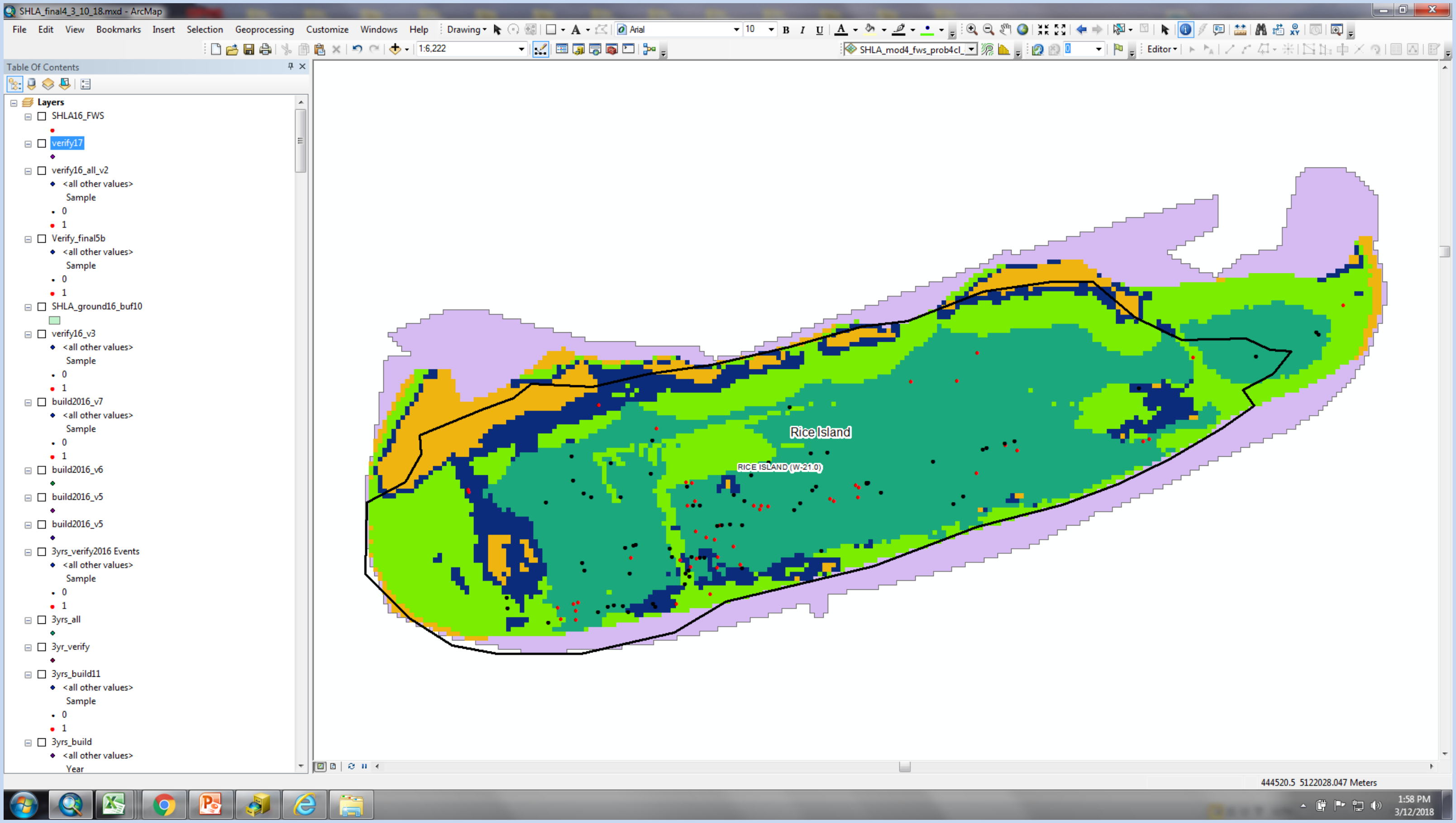
USGS
science for a changing world

Overview of SHLA Modeling Project



April 10, 2018

Jim Hatten - USGS
Gary Slater - CNLM
Jerrmaine Treadwell - CNLM
Matt Stevenson - CoreGIS

Project objectives

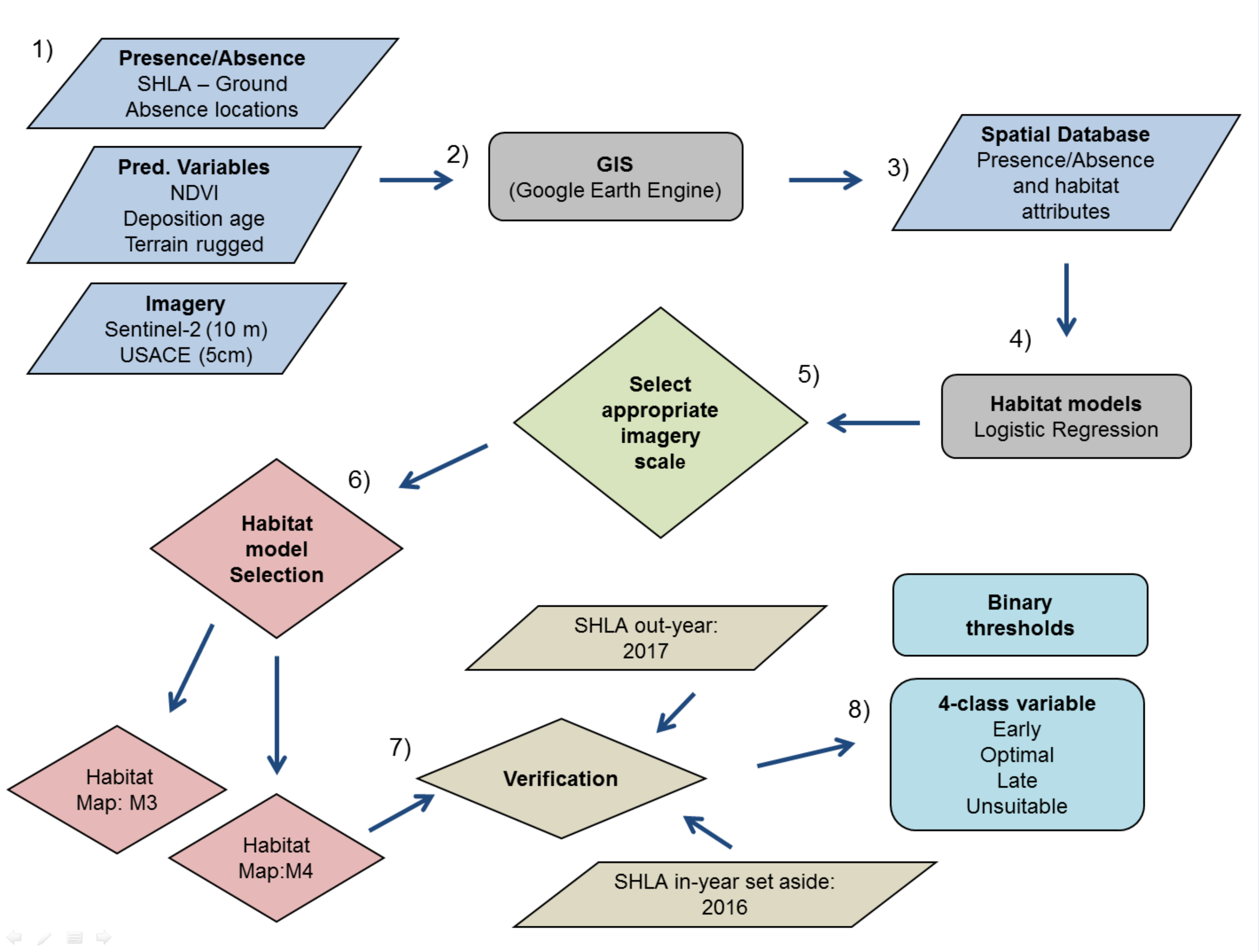
Goal: Inform agencies responsible for lark habitat management & conservation

- create a GIS tool to quantify & track SHLA habitat across lower 169 km of Columbia River

- develop a probability model to remotely identify & monitor lark habitat
- make maps of lark habitat at all 25 USACE dredge disposal sites
- develop a variable that can identify lark habitat by seral stage (young, mature, old)
- characterize the distribution and stability of lark habitat within and between years

USGS

Modeling Flow Diagram



USGS

Logistic Regression Modeling process

- Build main-effects model from conceptual diagram
  - Contains all significant variables from conceptual diagram (e.g., seasonal climate variables, topography, vegetation - when necessary)
  - Examine scale of covariates
  - Test for interactions
  - Look for autocorrelation (i.e., autologistic regression)
  - Calculate model-fit stats (e.g., ROC, AIC, classification table)
- Develop competing sub-models
  - Systematically remove one or more variables
  - Recompute model-fit stats
- Select most parsimonious model
  - "an accurate approximation to the interpretable information in the data at hand."
  - Keep it as simple as possible, don't overfit

References

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information theoretic approach, 2nd edition. Springer-Verlag, New York.

Hatten, J.R., Tiffan, K., Anglin,D., Haeseker, S., Skalicky, J., and H. Schaller. 2009. A Spatial Model to Assess the Effects of Hydropower Operations on Columbia River Fall Chinook Salmon Spawning Habitat. North American Journal of Fisheries Management 29:1379–1405, 2009.

Hosmer, D. W., and S. Lemeshow. 2000. Applied logistic regression. Wiley, New York.

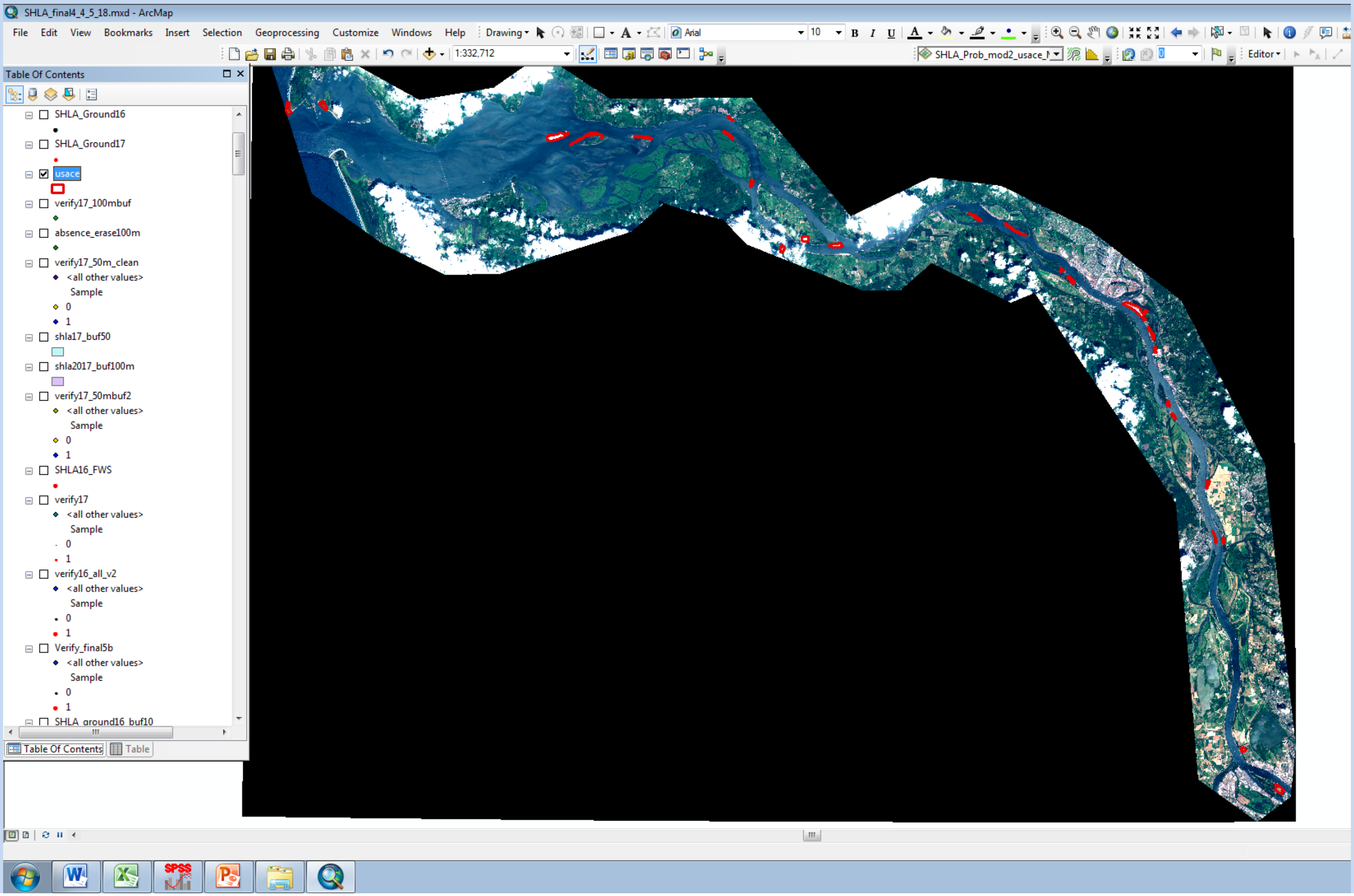
Menard, S. 2001. Applied logistic regression analysis. Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-106. Thousand Oaks, CA: Sage.

Use Sentinel-2 Satellite Imagery

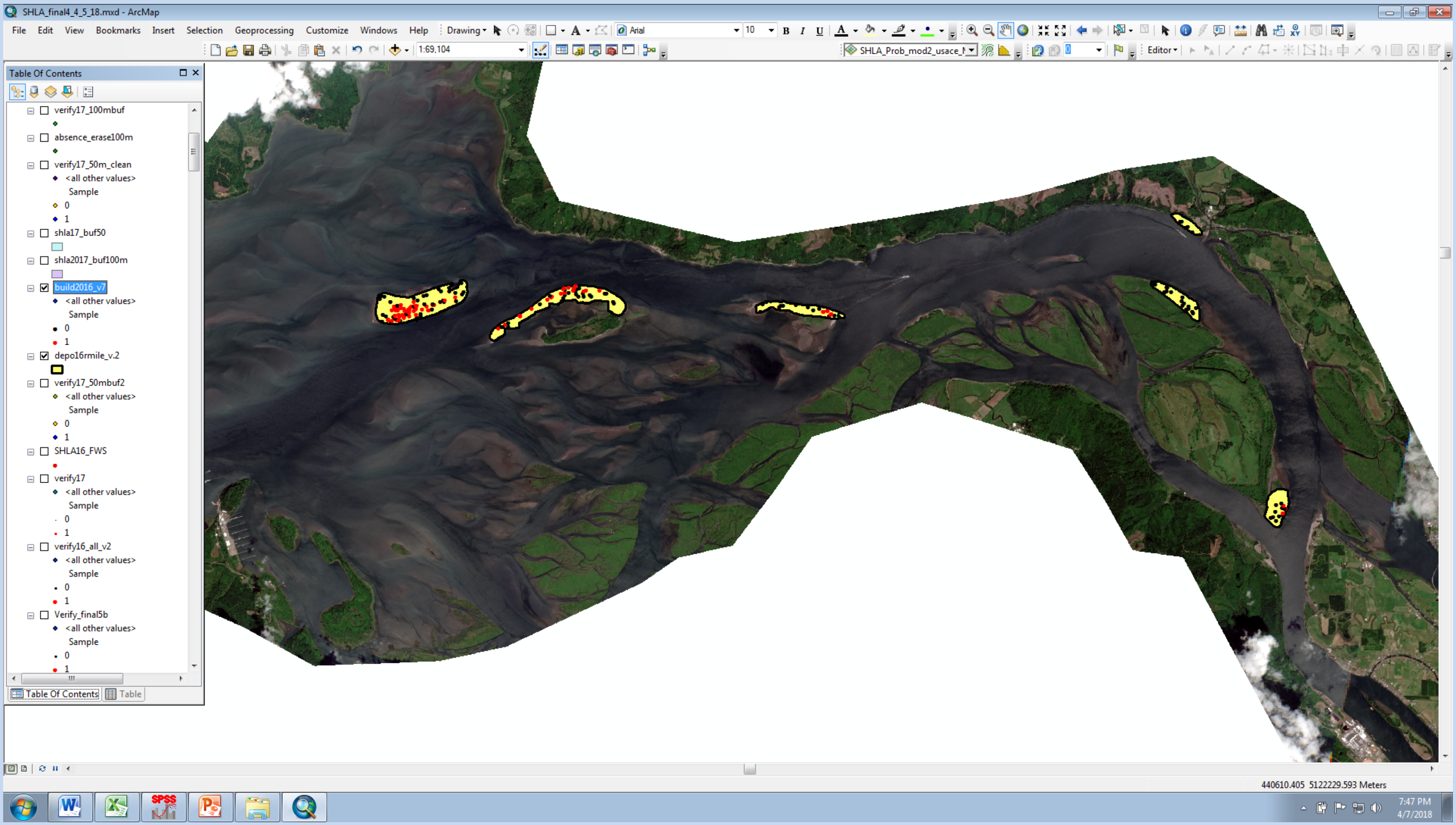
- European Space Agency satellite (launched in summer 2015)
- 10-m spatial resolution
- Updates every 4-8 days
- Freely available to everyone
- Google Earth Engine seamlessly mosaics and atmospherically corrects it

Project area summer 2016

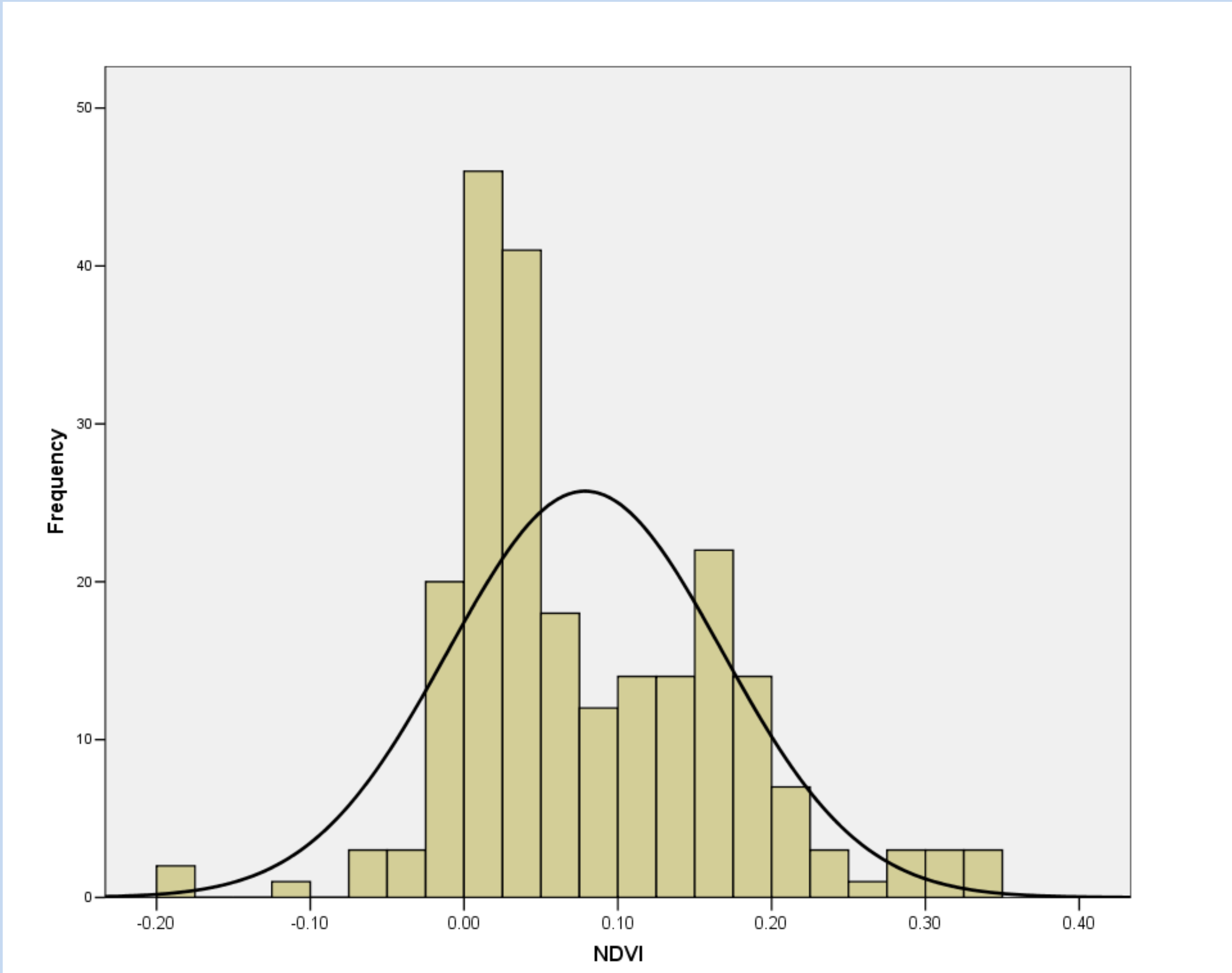
Sentinel-2 natural color image & Corps deposition sites



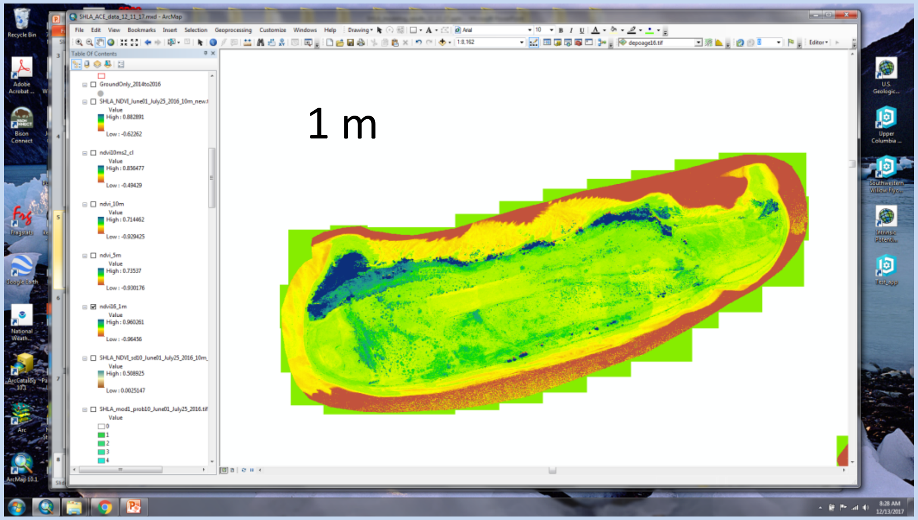
Built predictive models with Lark presence/absence data & logistic regression

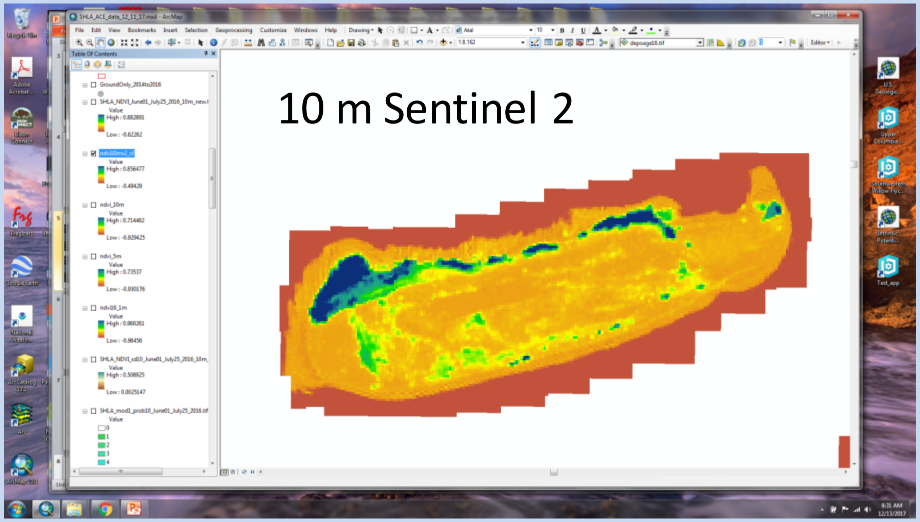


Importance of Normalized Difference Vegetation Index NDVI: values at lark locations: 2016 (n = 230)

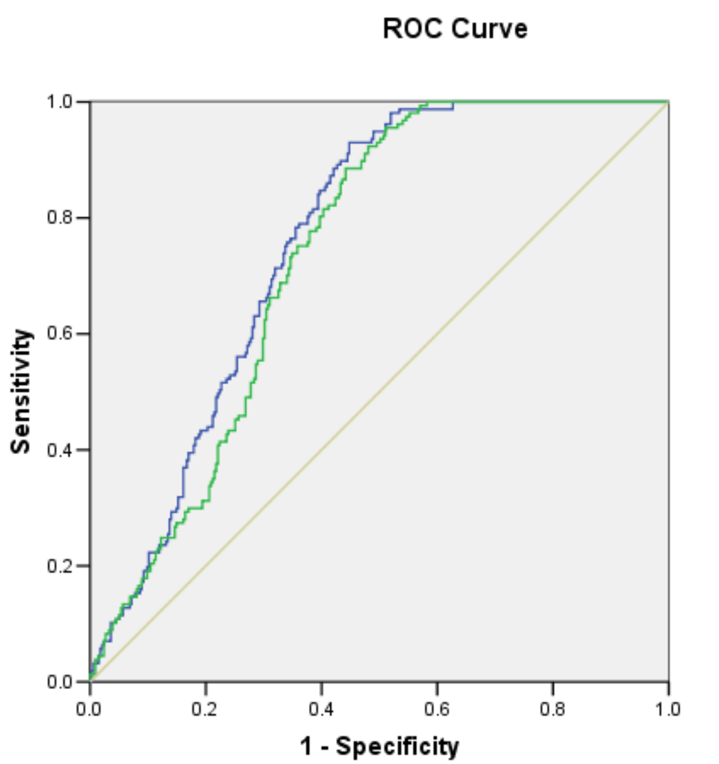


What spatial resolution to model lark habitat (1m – 10m)?





Compare 1m & 10m NDVI: Corps versus Sentinel 2



10m spatial resolution imagery was a Better classifier of lark habitat than 1m resolution

ΔAIC = 5.66

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B) Lower	Upper
Step 1[a]	ndvi16_s2	7.302	1.914	14.550	1	.000	1483.511	34.819	63207.934
	nd16s2_2	-33.388	6.610	25.517	1	.000	.000	.000	.000
	Constant	-.236	.137	2.945	1	.086	.790		

a. Variable(s) entered on step 1: ndvi16_s2, nd16s2_2.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B) Lower	Upper
Step 1[a]	ndvi16_1m	5.303	2.553	4.316	1	.038	201.030	1.350	29929.455
	nd161m_2	-45.614	12.914	12.477	1	.000	.000	.000	.000
	Constant	-.063	.124	.257	1	.612	.939		

a. Variable(s) entered on step 1: ndvi16_1m, nd161m_2.

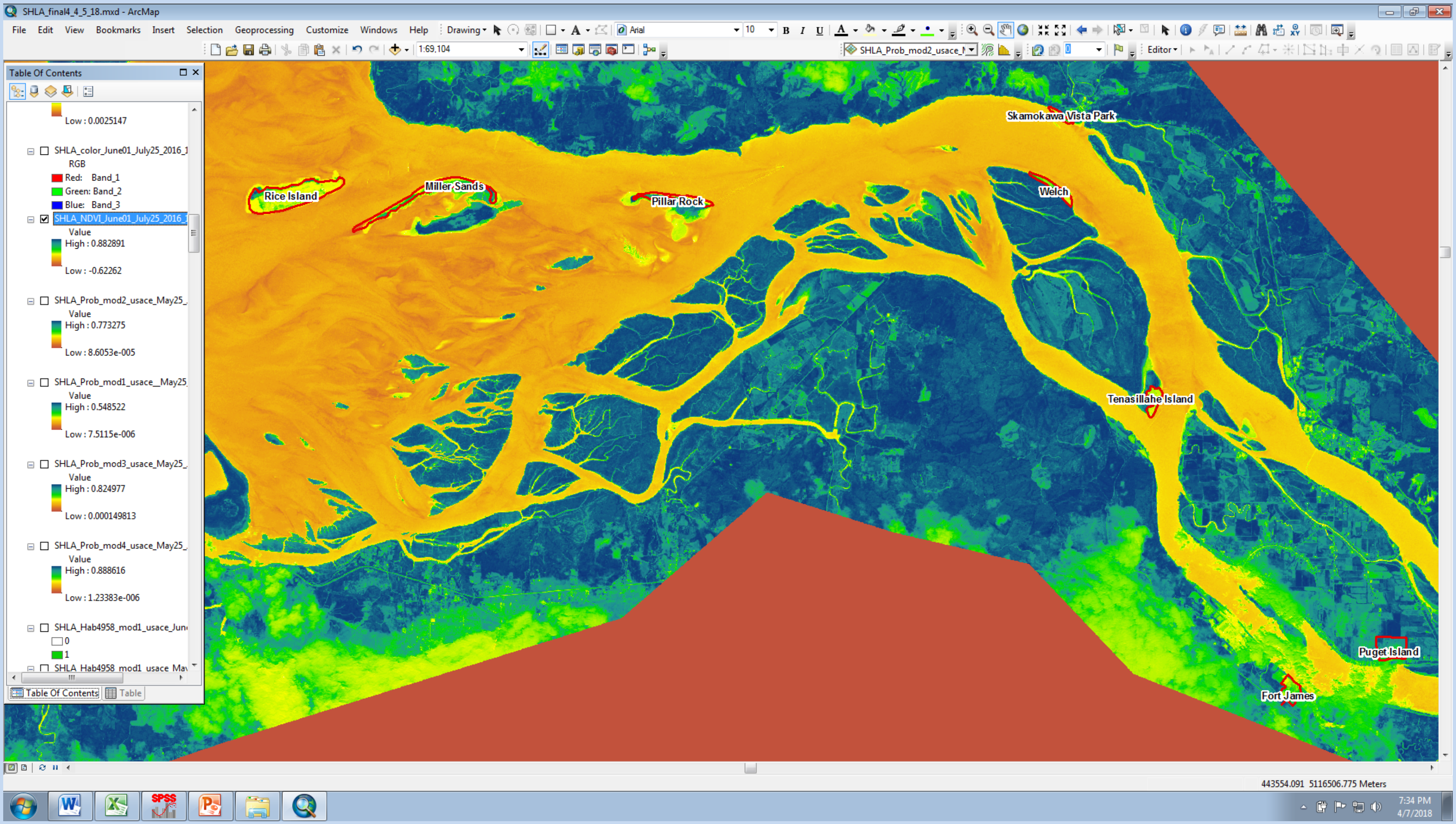
Area Under the Curve

Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval Lower Bound	Upper Bound
Pre_nd10m	.760	.021	.000	.720	.801
Pre_nd1m	.738	.022	.000	.696	.780

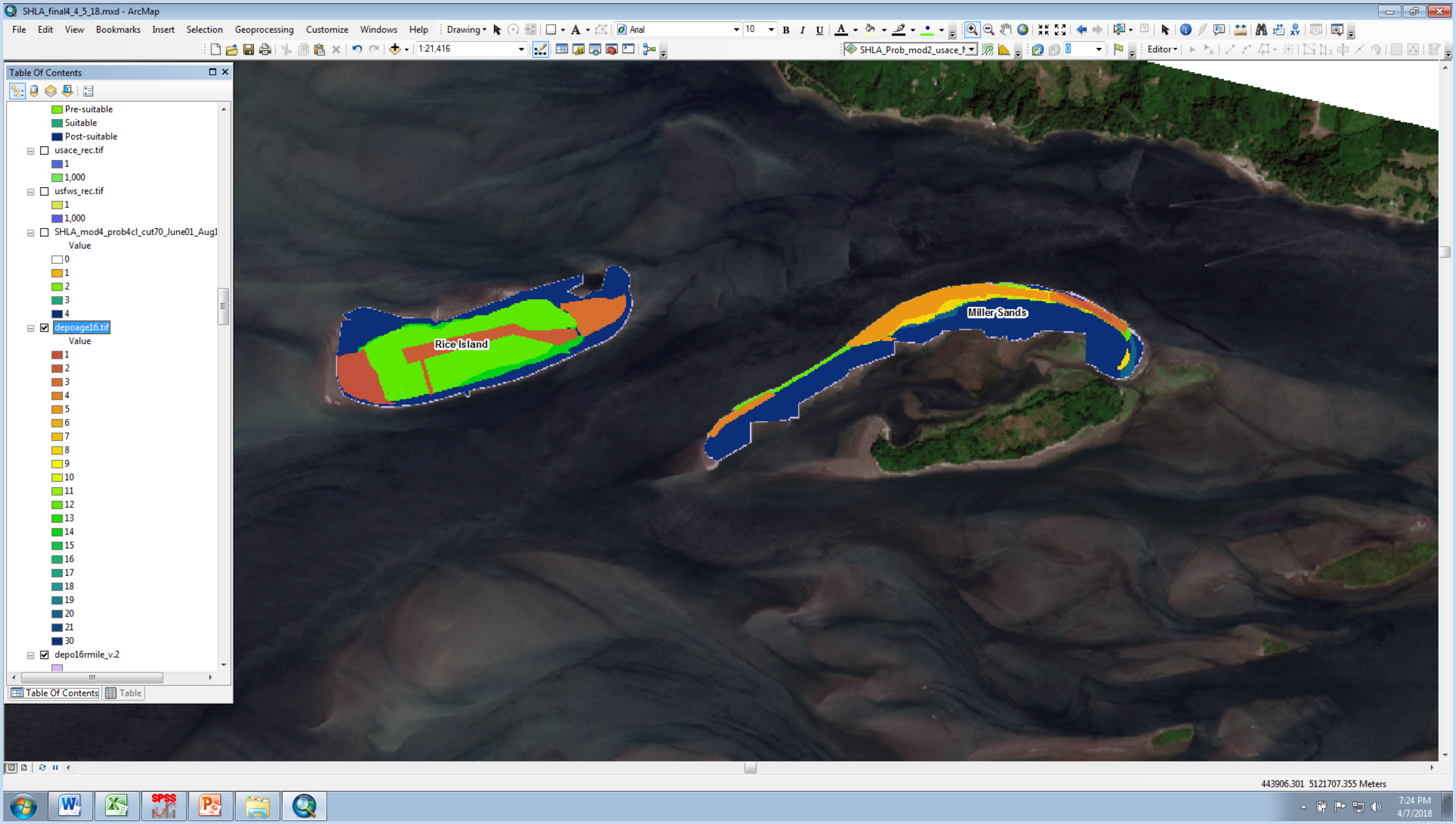
a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

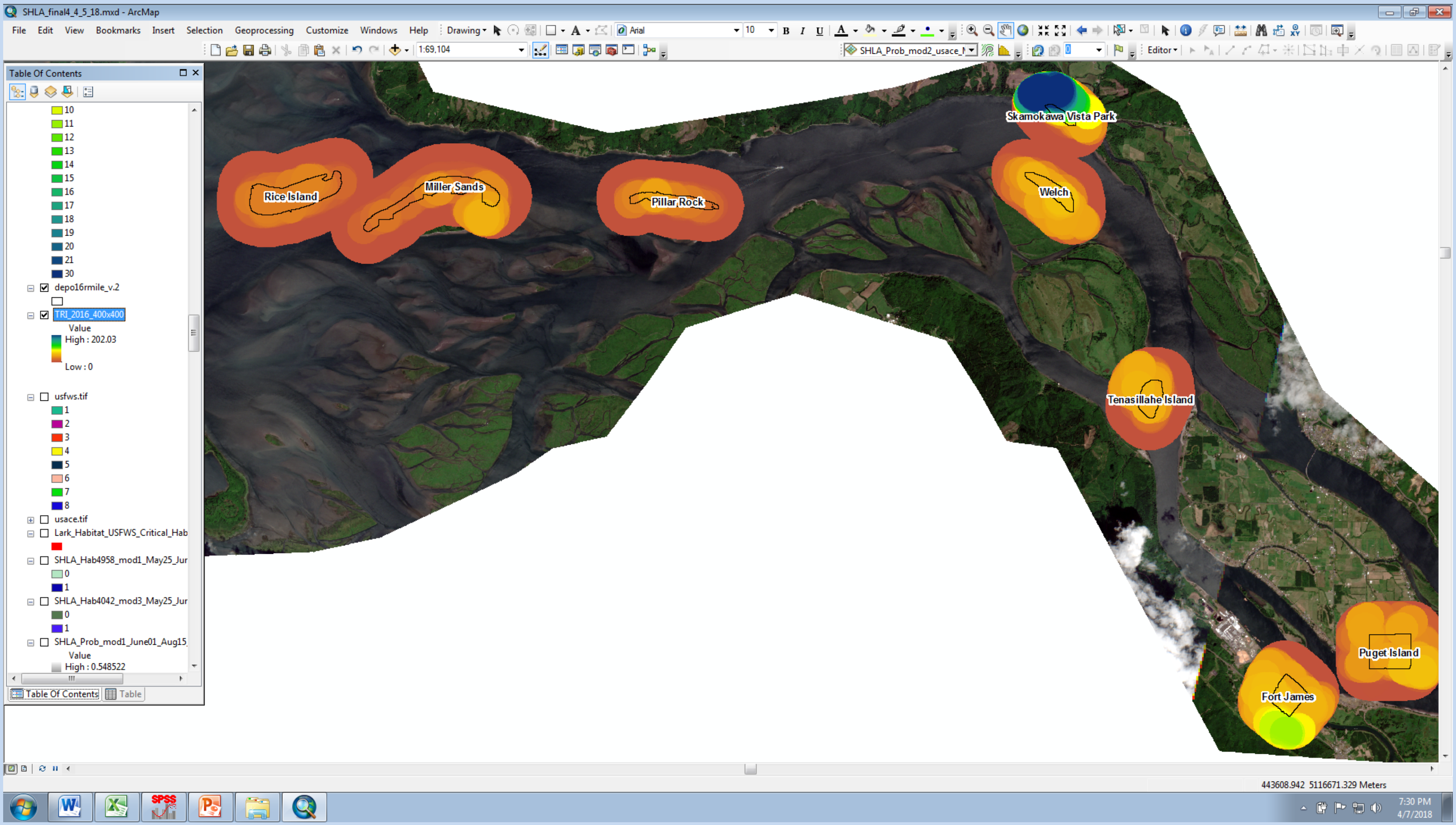
Predictor variable: Sentinel-2
Normalized Difference Vegetation Index (NDVI)



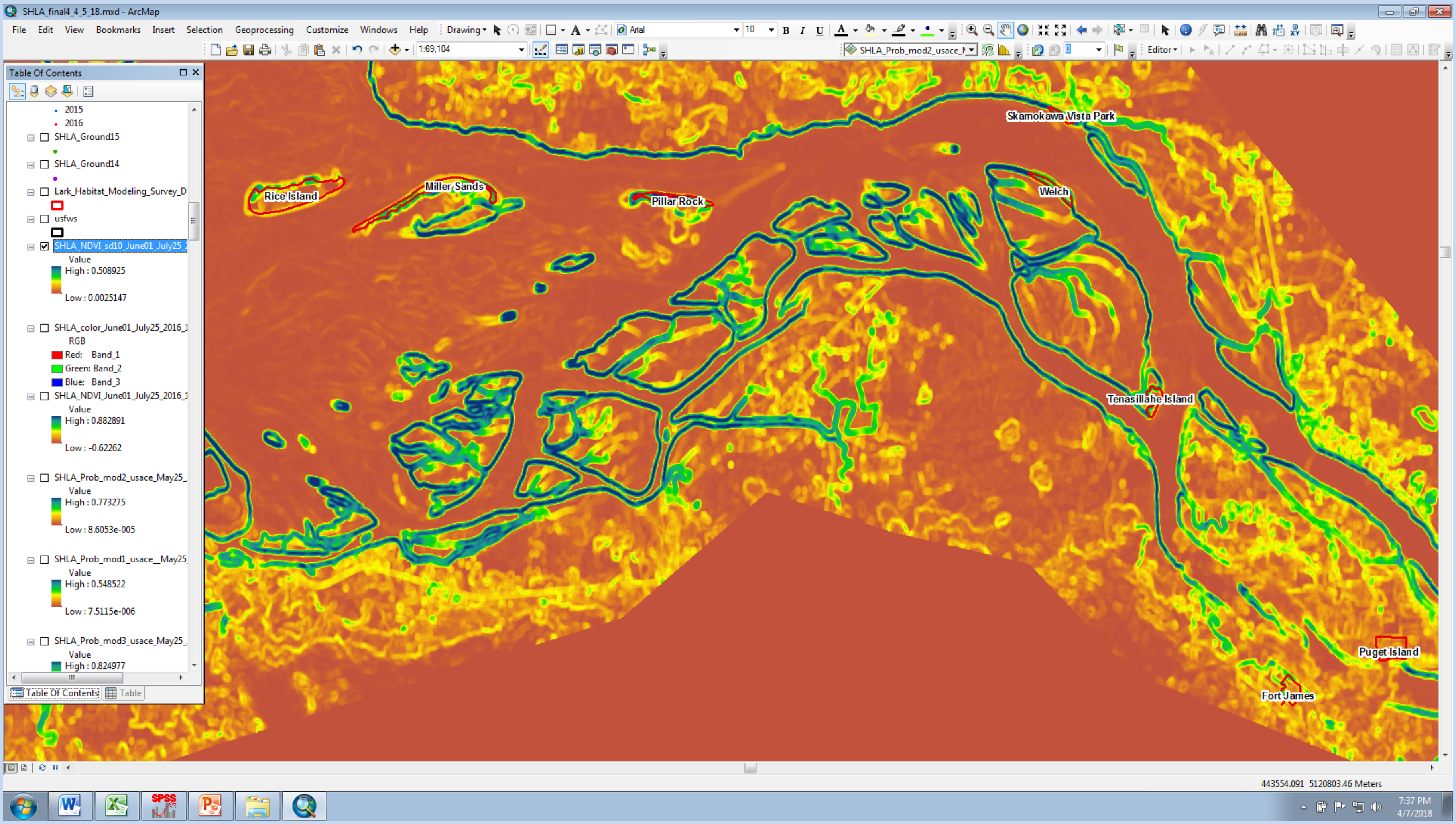
Predictor variable: deposition age



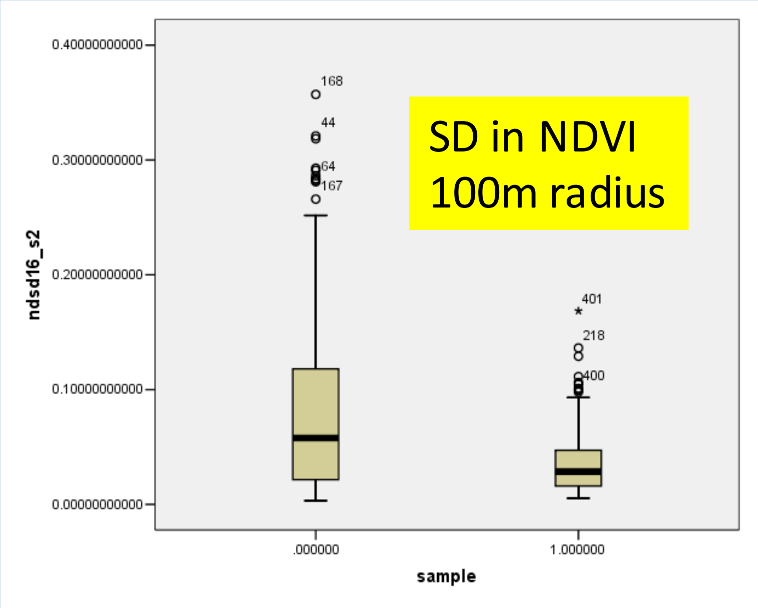
Predictor variable: terrain ruggedness index (TRI) 400m radius

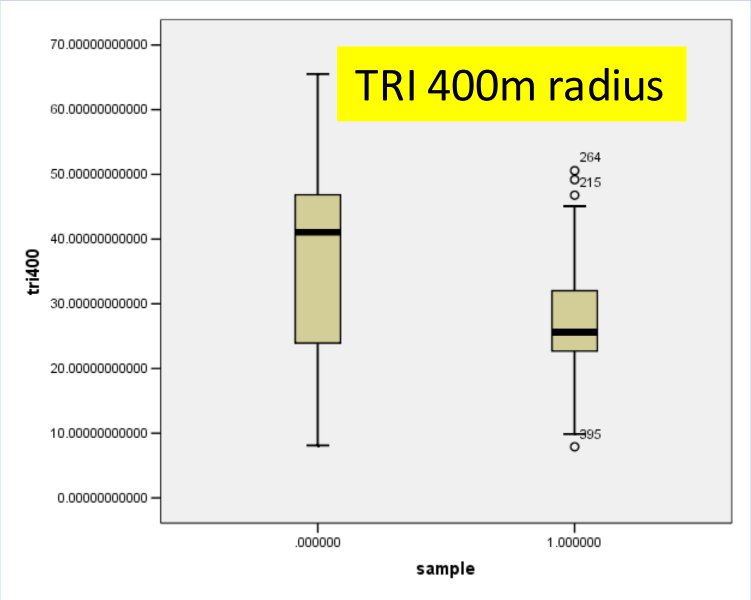


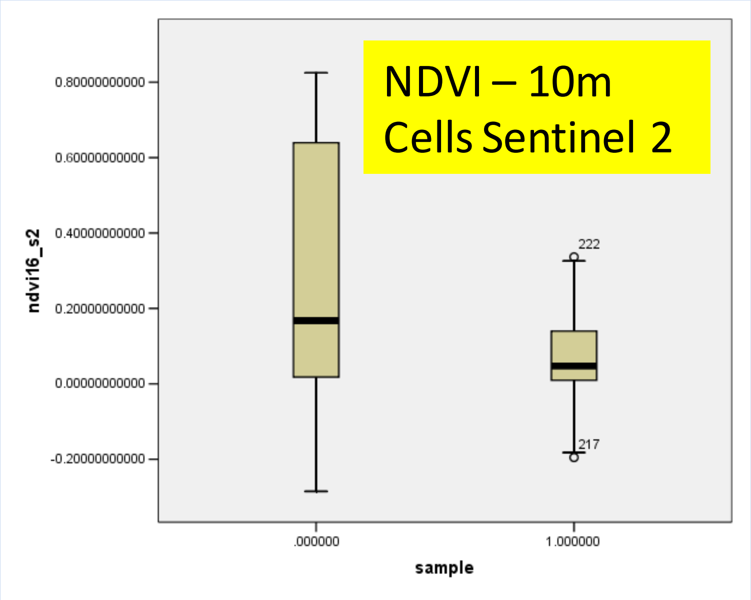
Predictor variable: Sentinel-2 SD of NDVI – 100m radius



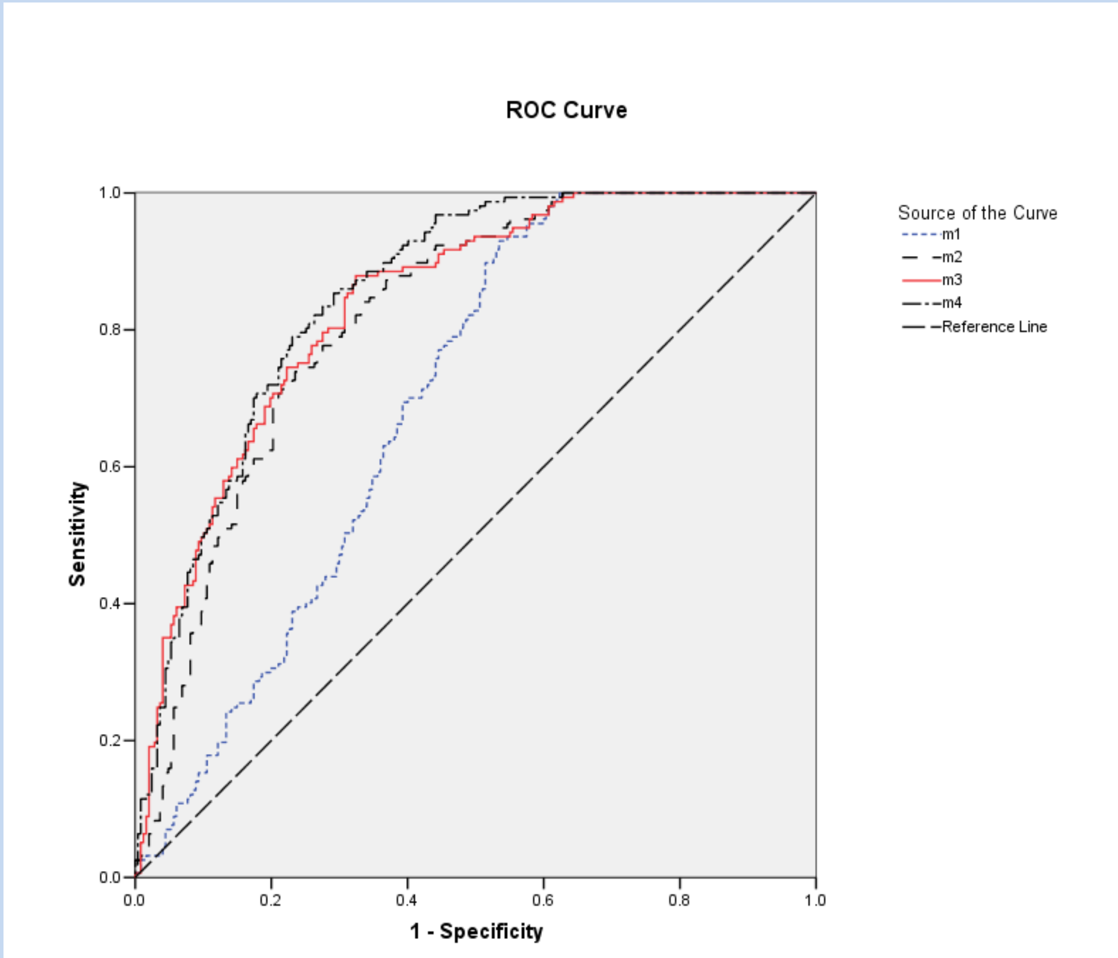
Boxplots – Environmental values at lark presence/absence 2016







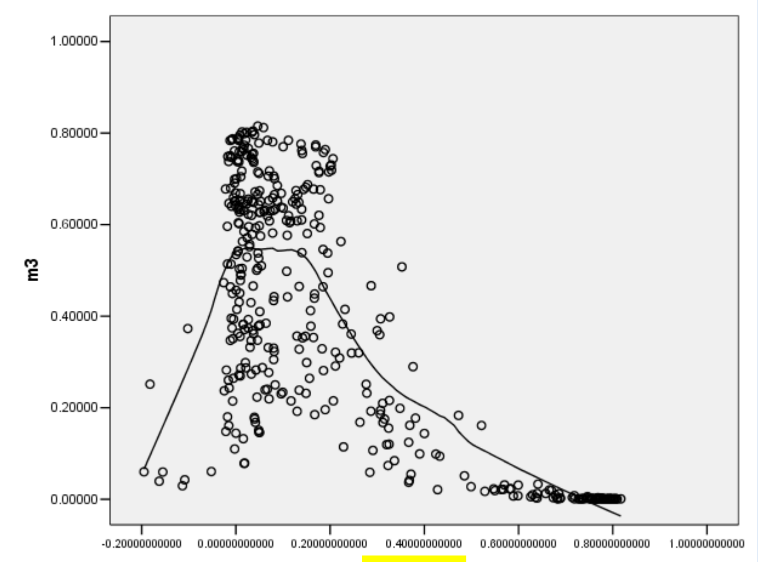
Model selection & accuracy assessment



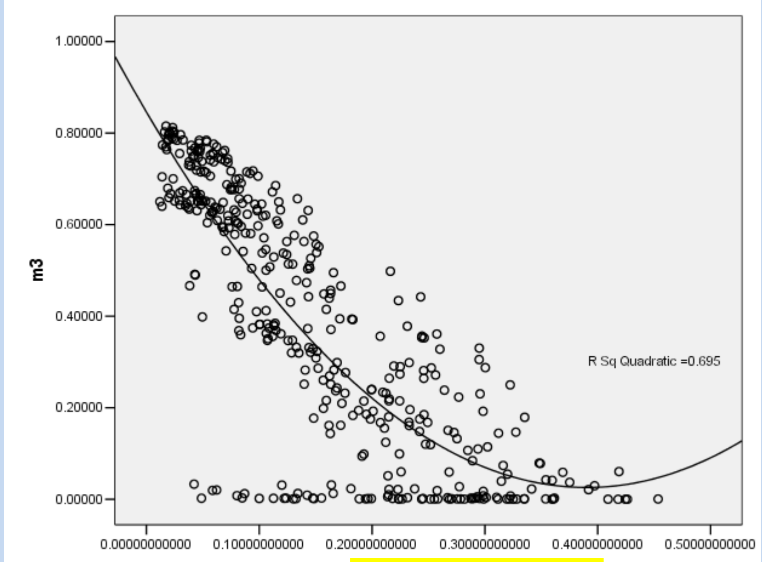
Area Under the Curve

Test Result Variable(s)	Area	Std. Error(a)	Asymptotic Sig.(b)	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
m1	.695	.025	.000	.646	.745
m2	.817	.021	.000	.776	.857
m3	.838	.019	.000	.800	.876
m4	.852	.018	.000	.816	.888

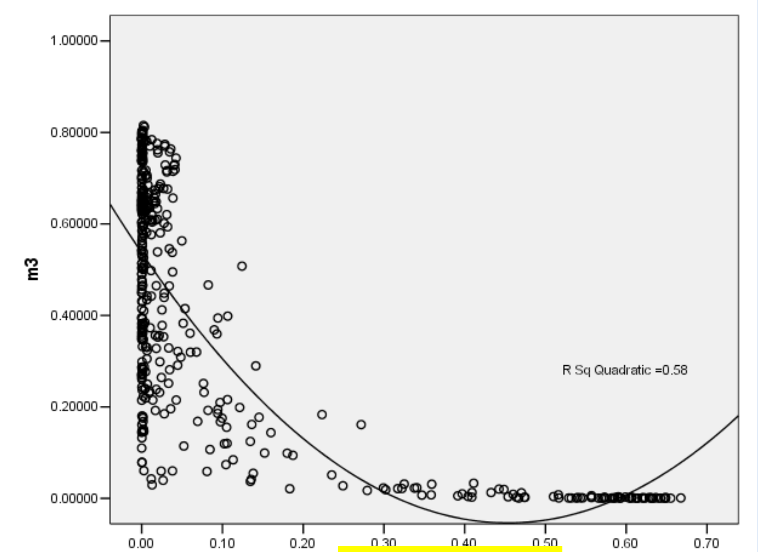
Predictor/model relationships – model 3



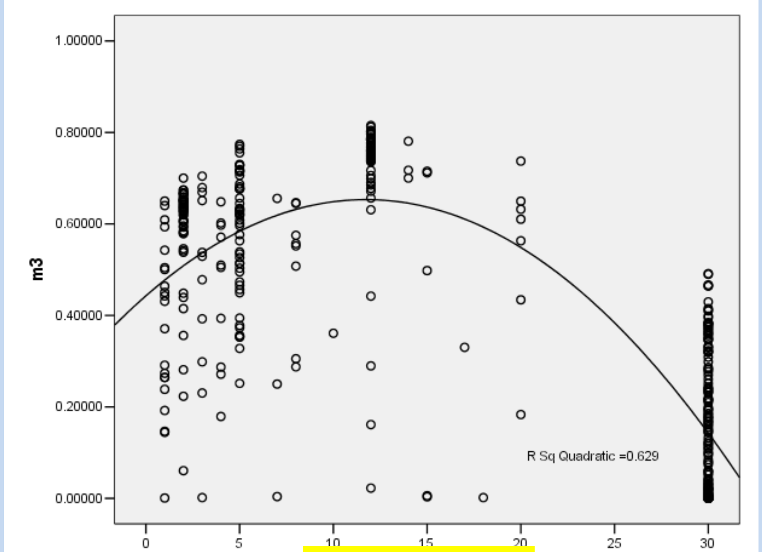
NDVI



NDVI_SD_100mr

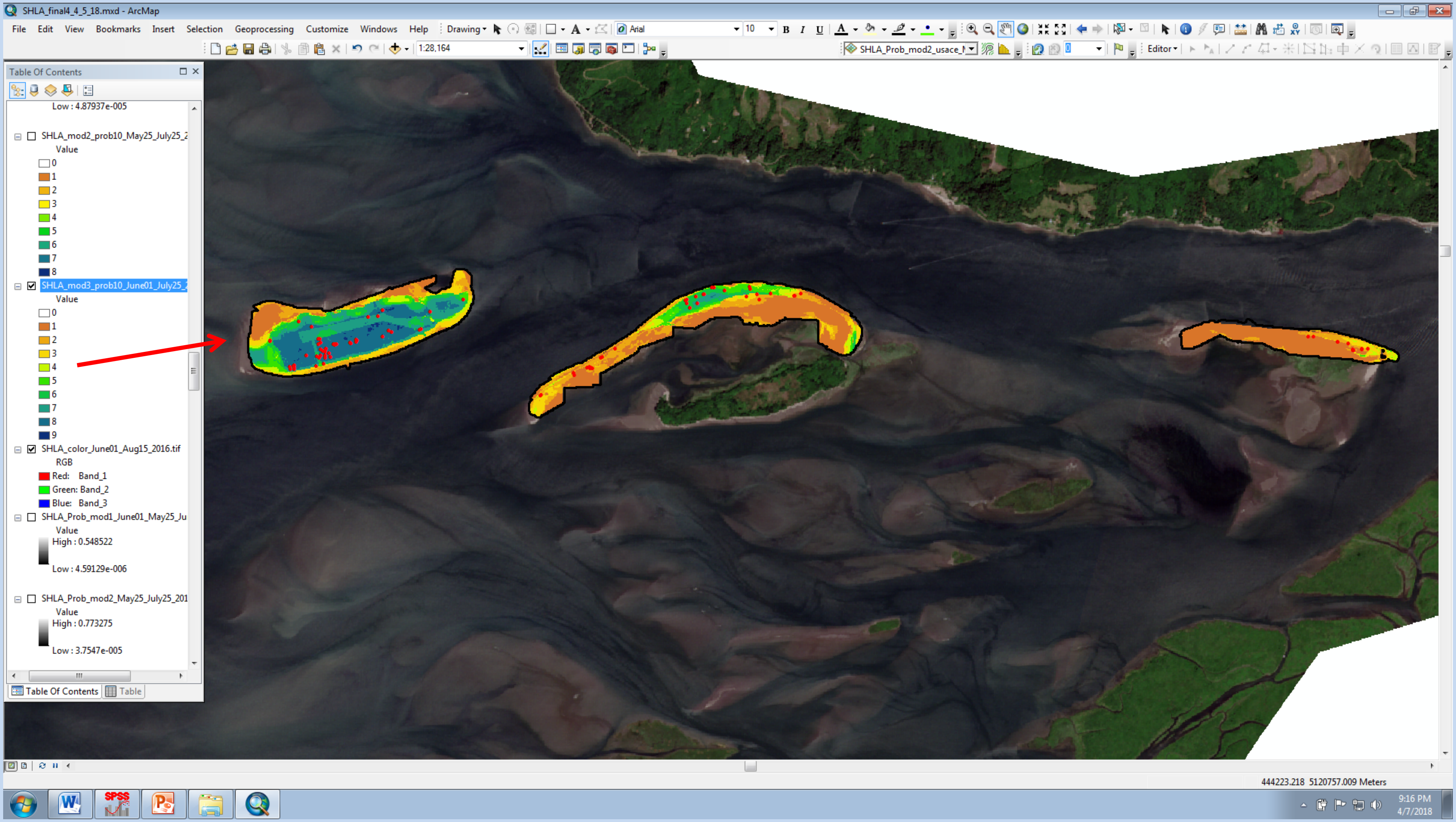


NDVI_squared

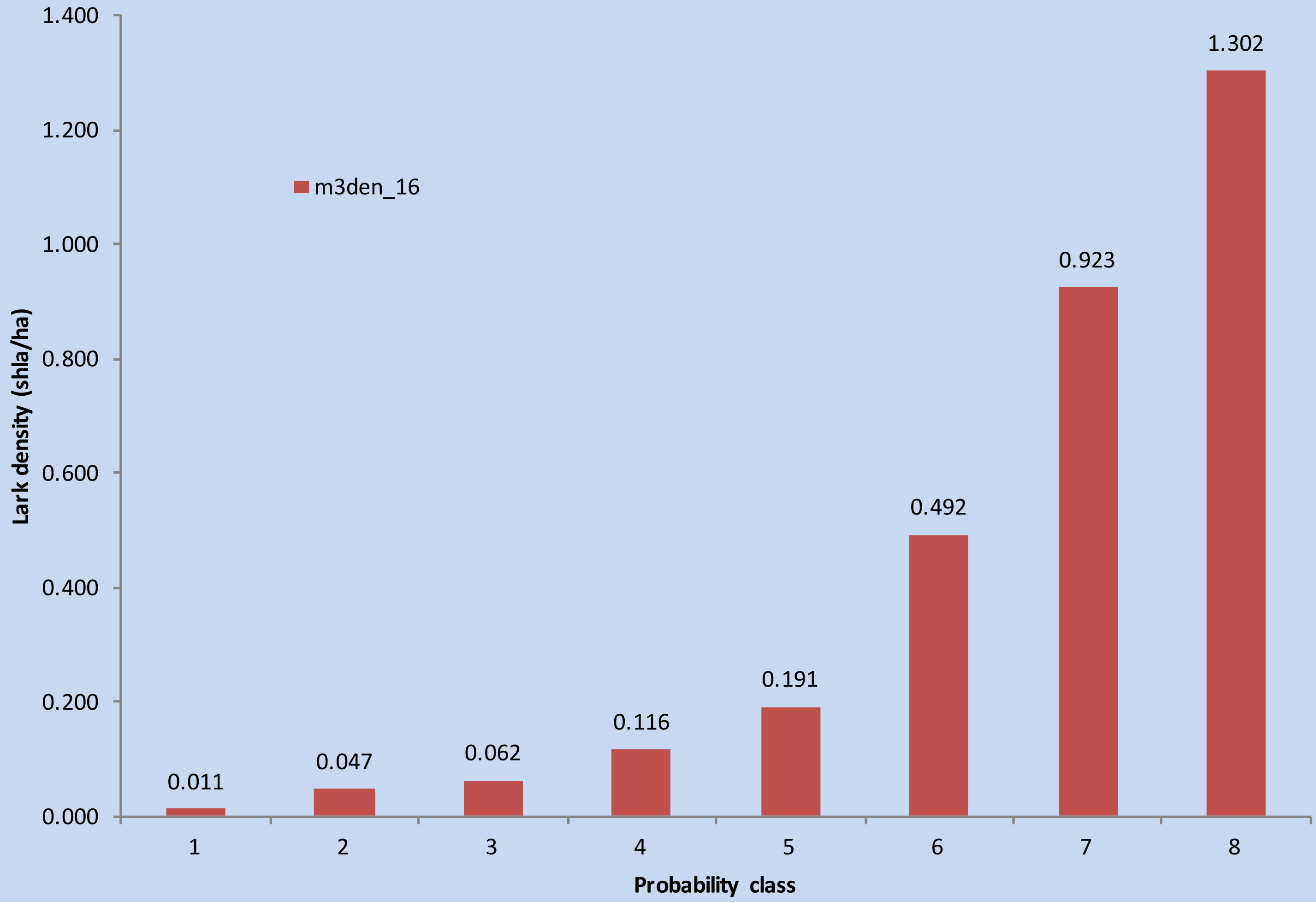


Deposition age

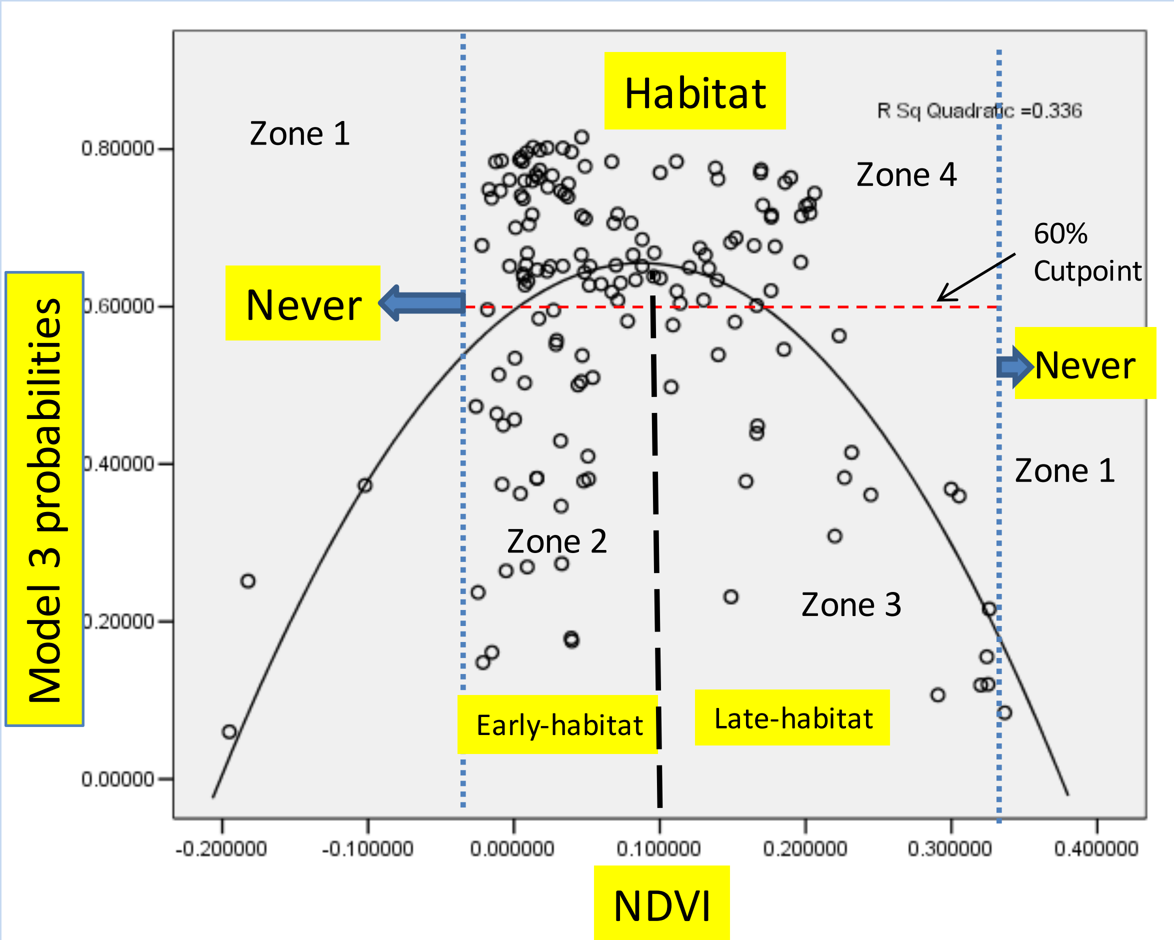
Larks increase within higher model probability classes



Lark density increases in higher probability classes: 2016 (n = 230)

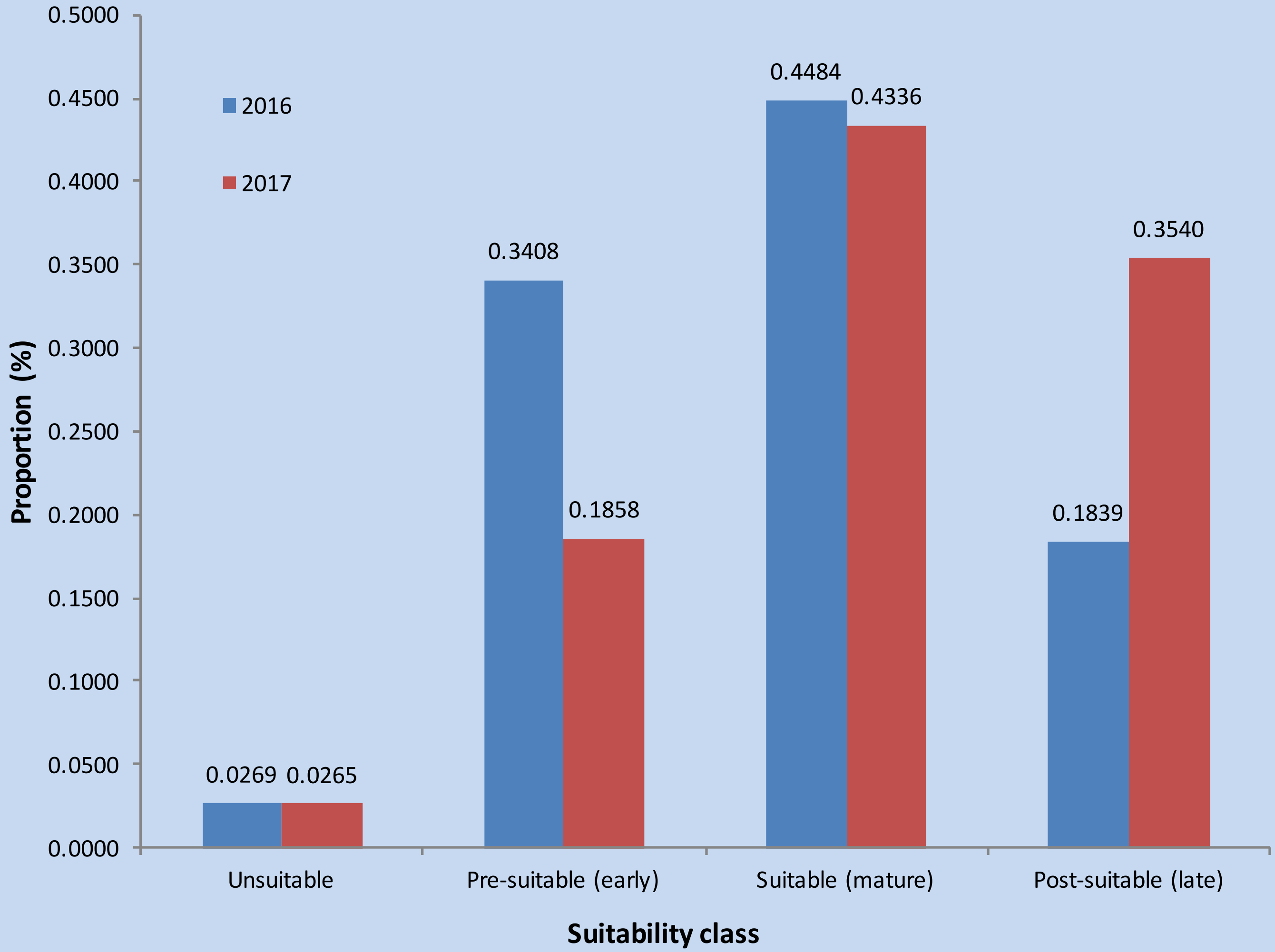


Create a designer variable to predict habitat suitability based upon seral stage (young, mature, old)

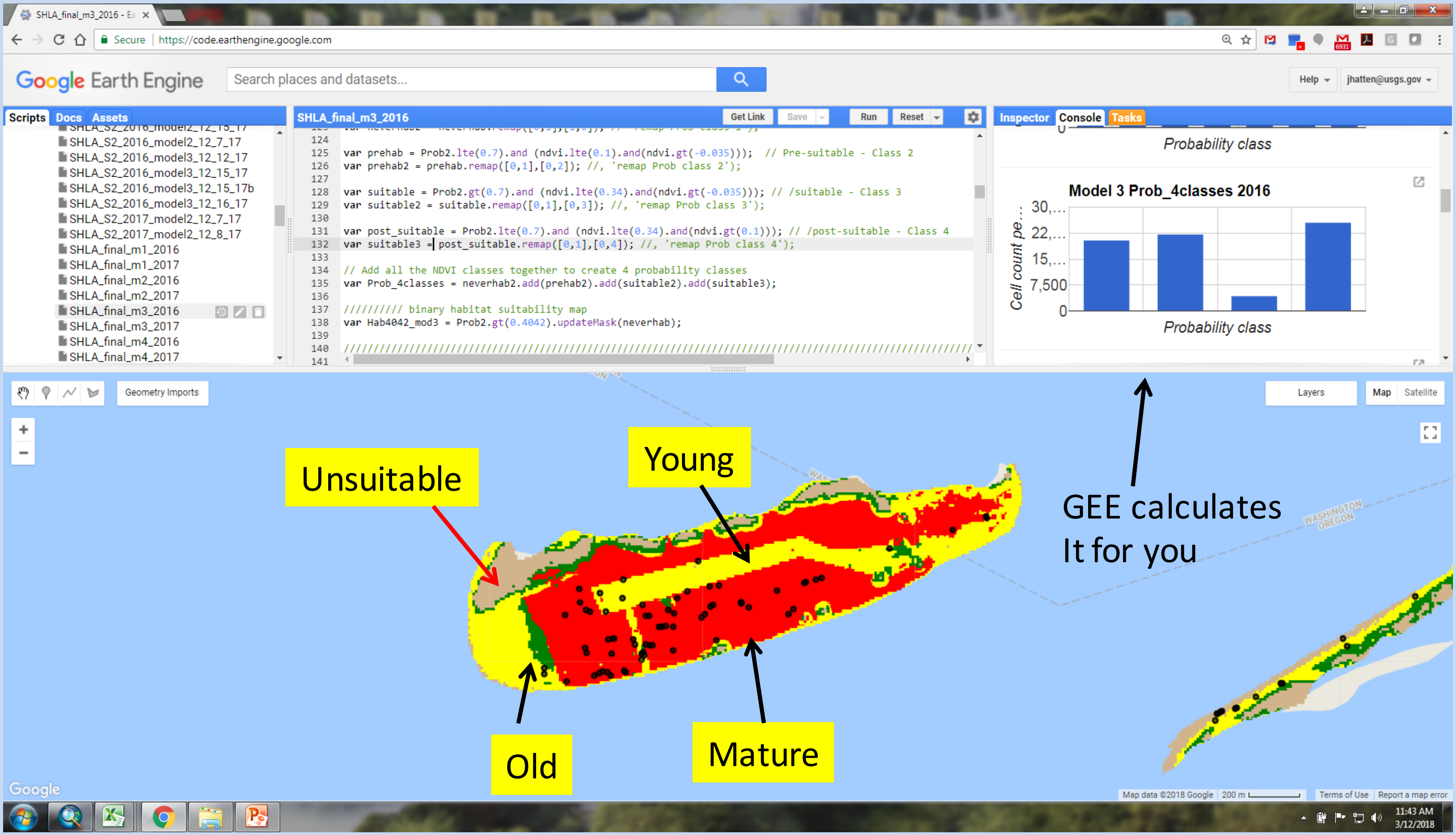


Creating a 4-class predictor variable: unsuitable,
Pre-suitable (young), suitable (mature), post-suitable(old)

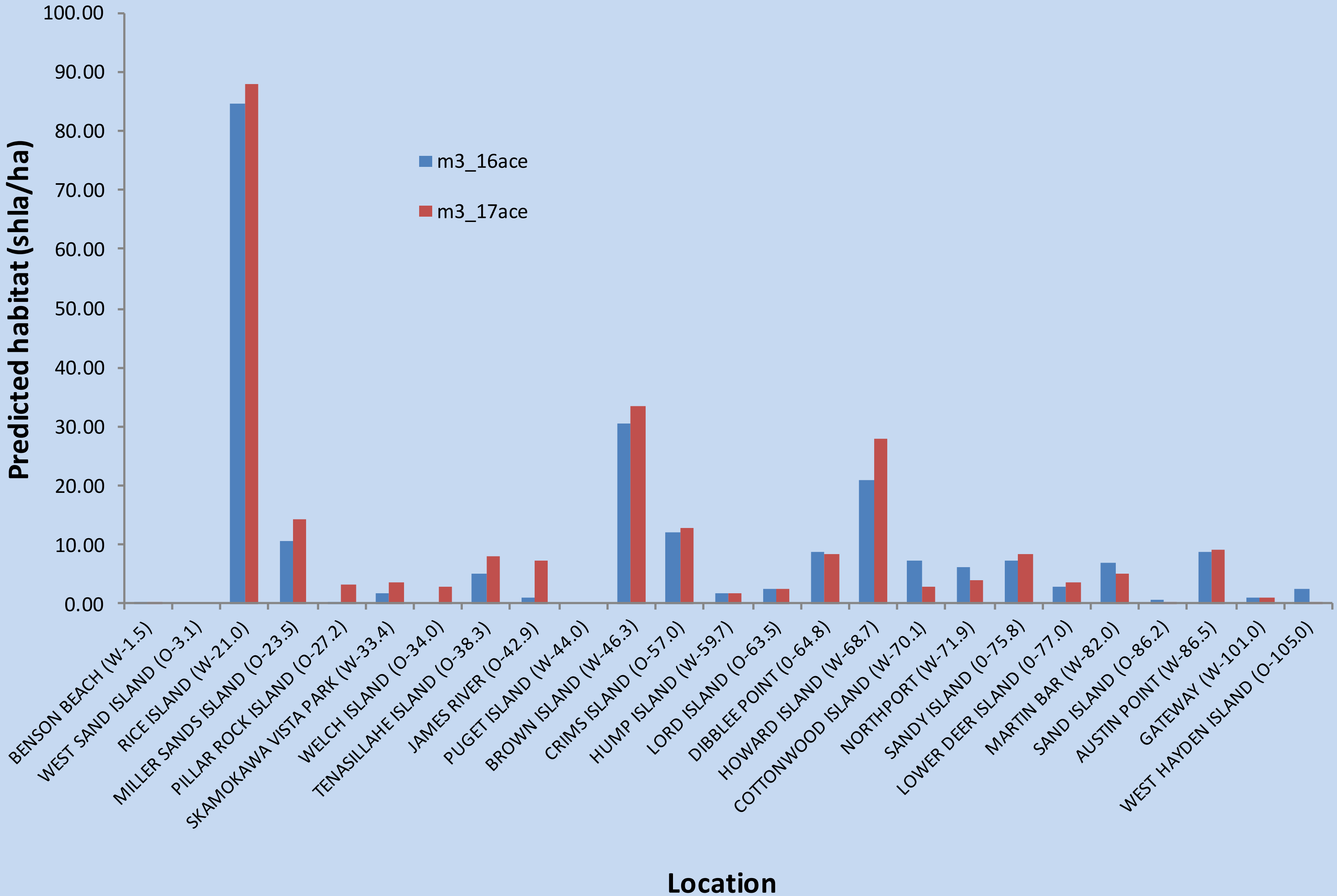
Proportion of larks within each suitability class



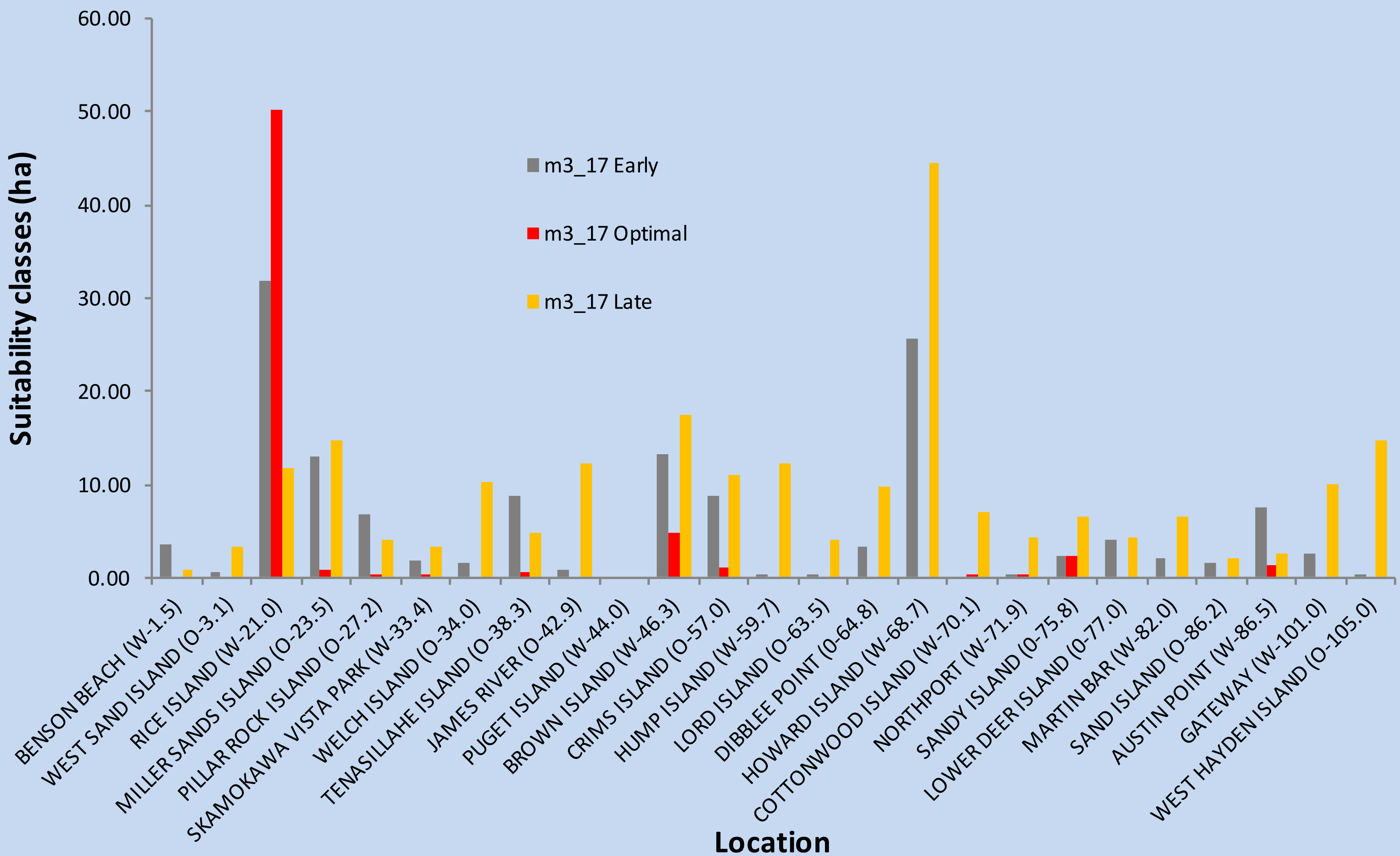
4-class designer variable in Google Earth Engine (2016)



Habitat predictions for years 2016 – 2017)



Designer variable: habitat suitability classes: 2017



Summary

- Successfully created a new predictive model of lark habitat
- Model runs seamlessly in GEE or ArcGIS
- USACE will begin using model this summer
- Model will inform FWS' Biop and Corp disposal schedule
- Model can be adapted for many other species and locations
- Data and GEE are free!