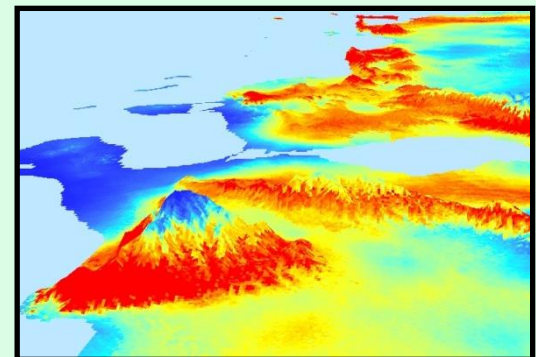
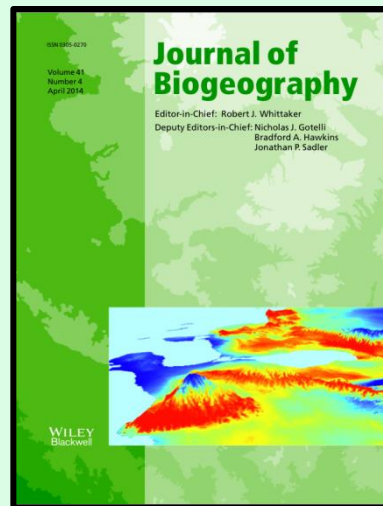


Spatially independent evaluations and estimates of optimal model complexity

Robert P. Anderson
Department of Biology
City College of CUNY



Model complexity

Occurrence data suffer from (among other things):

biased sampling across geography ...

which may also be biased in environmental space

Model complexity

Overly complex models overfit to bias or noise, and transfer poorly

In Maxent, complexity depends on (at least):

- 1) number of variables
- 2) feature classes considered
- 3) level of regularization



SPECIAL
PAPER

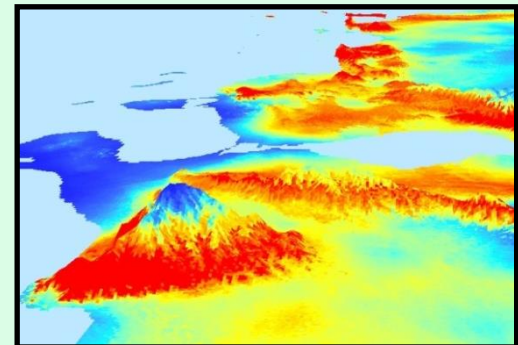
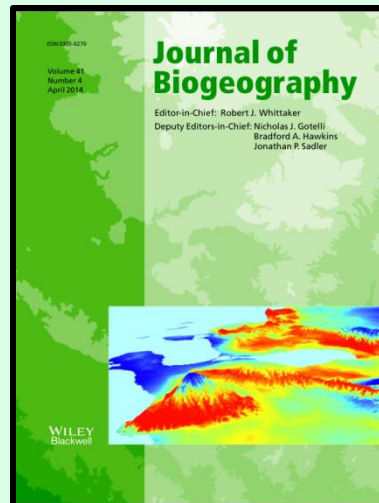


Making better MAXENT models of species distributions: complexity, overfitting and evaluation

Aleksandar Radosavljevic^{1*} and Robert P. Anderson^{1,2,3}



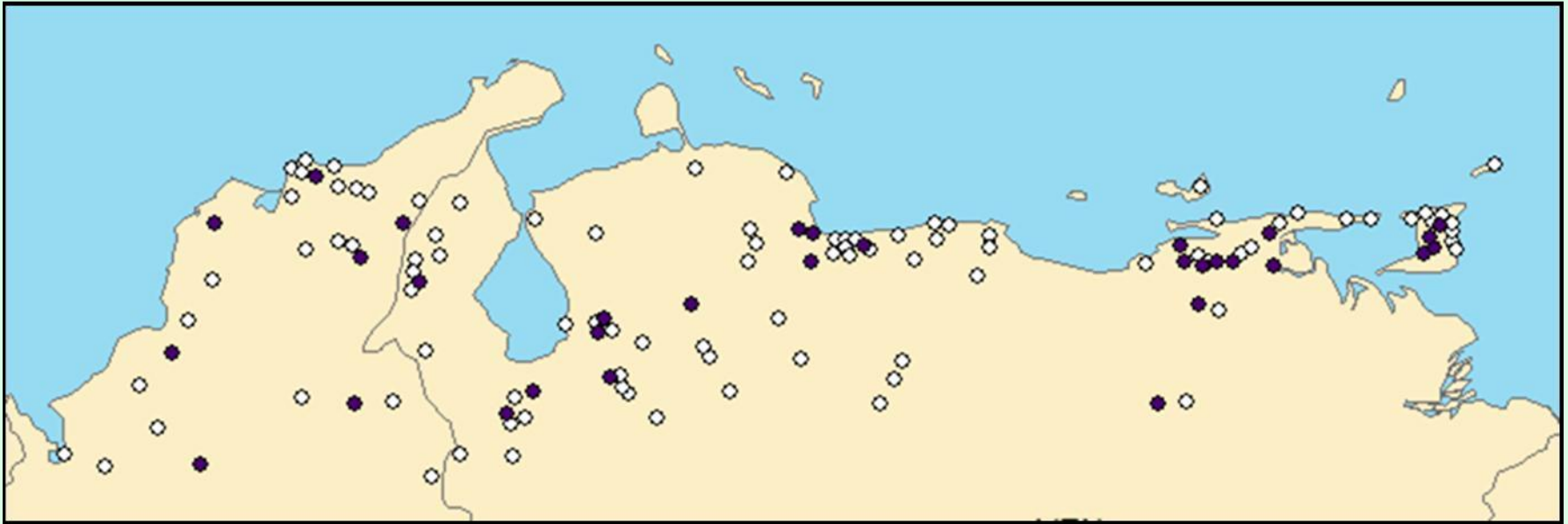
Aleks
Radosavljevic



Heteromys anomalus

Random subsets

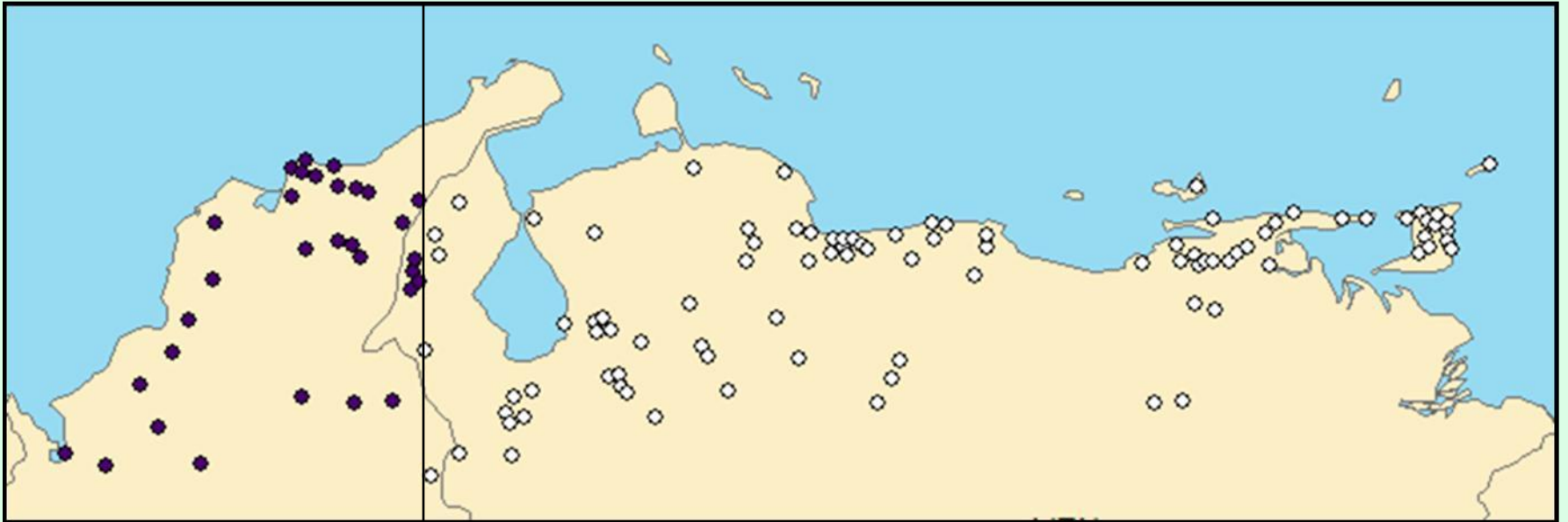
random split-sample approach: easy test, cannot detect overfitting to bias



White: calibrate the model

Black: evaluate the model

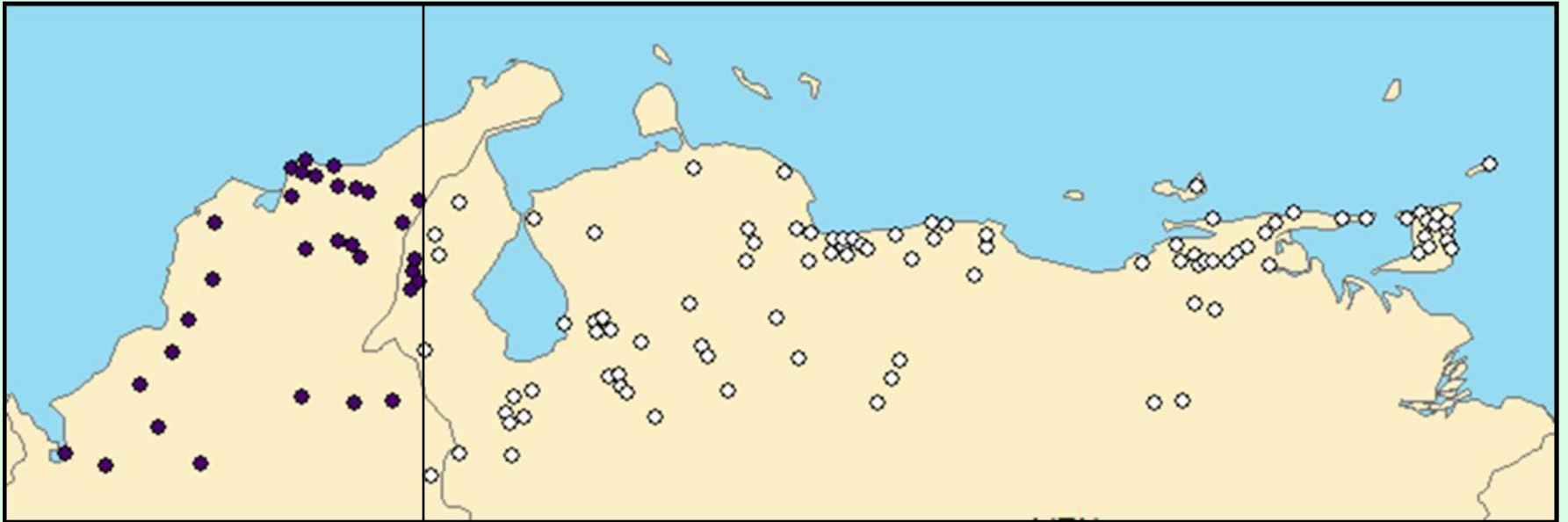
Spatial subsets



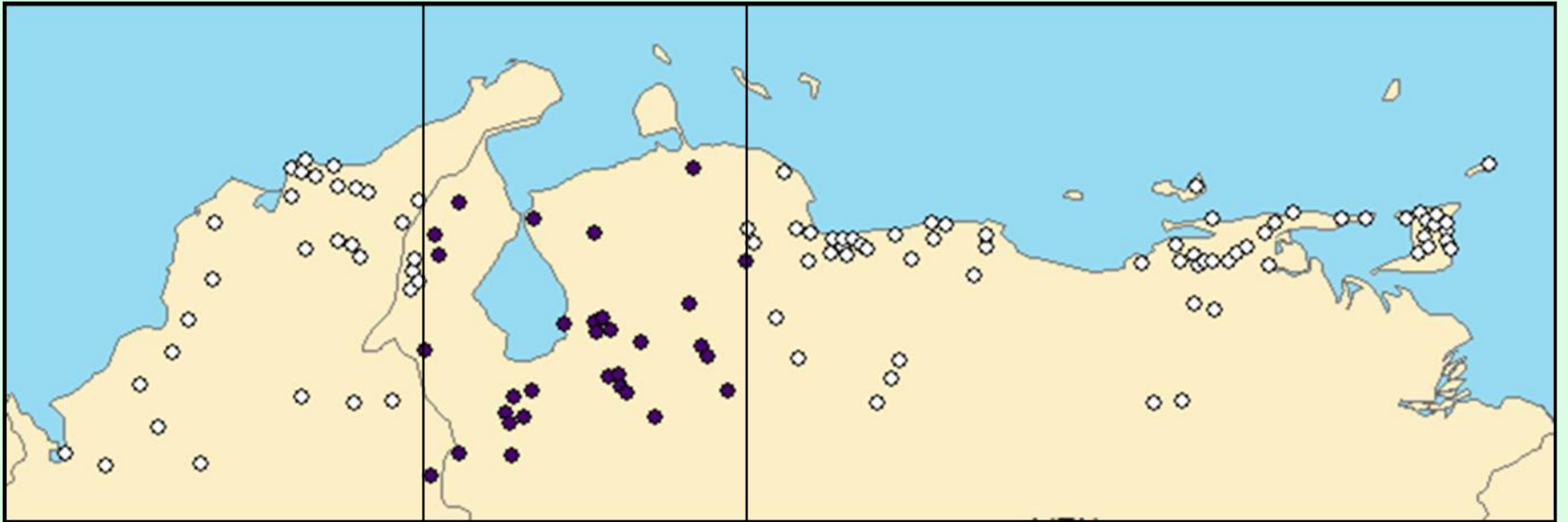
White: calibrate the model

Black: evaluate the model

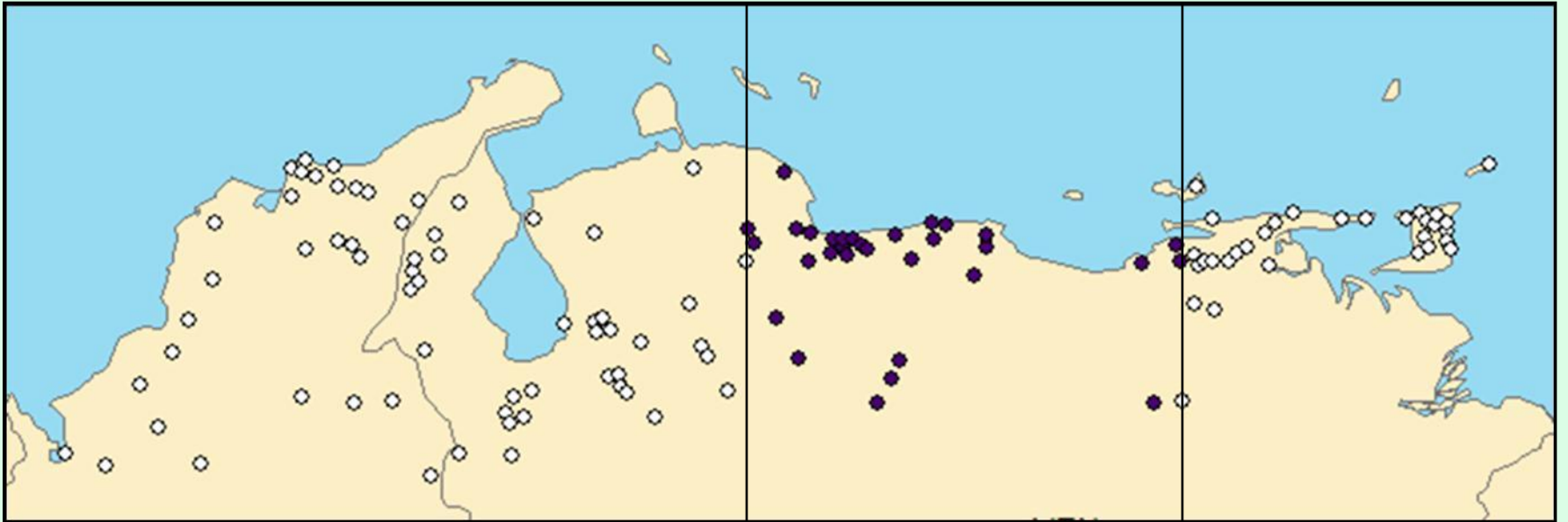
Bin A for evaluation



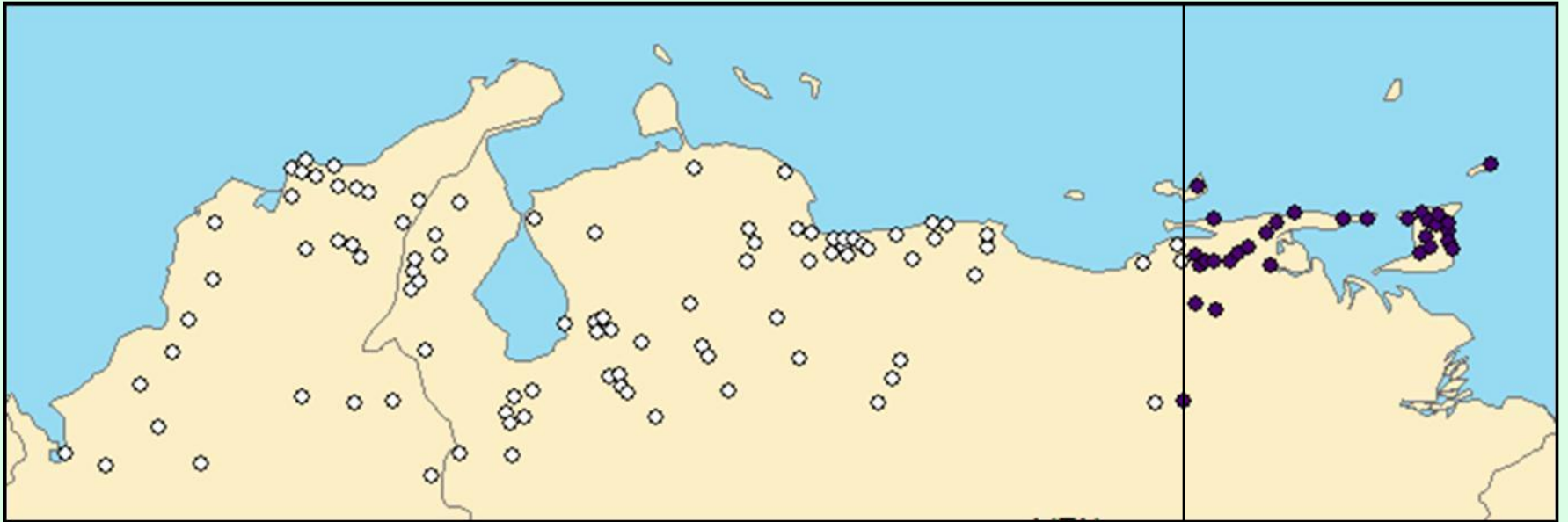
Bin B for evaluation



Bin C for evaluation



Bin D for evaluation




Measures of performance

AUC (Area under the ROC curve):


performance of the model in predicting a presence locality over a random pixel

(RELATIVE measure of performance, in a single study region, for a single species)

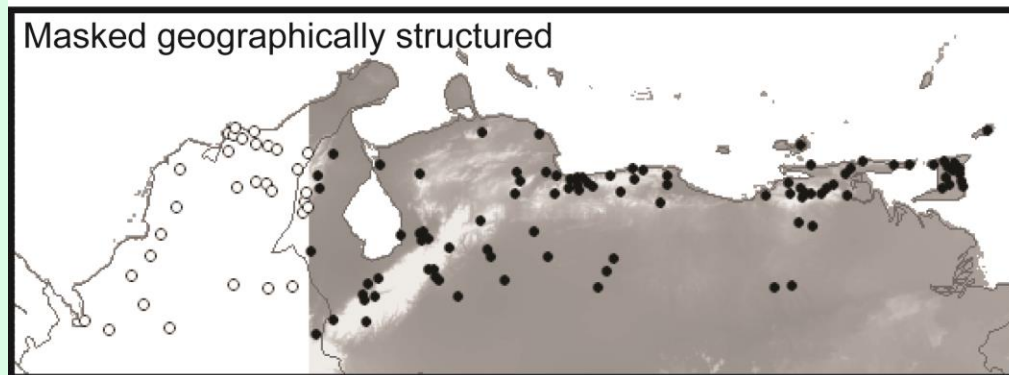
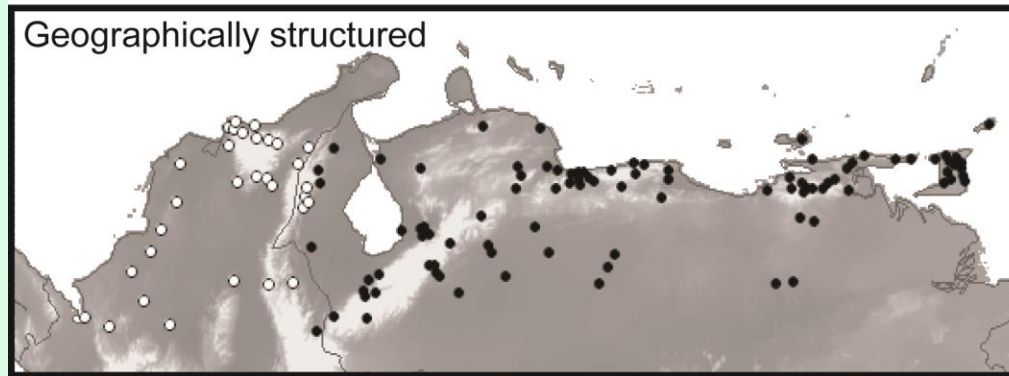
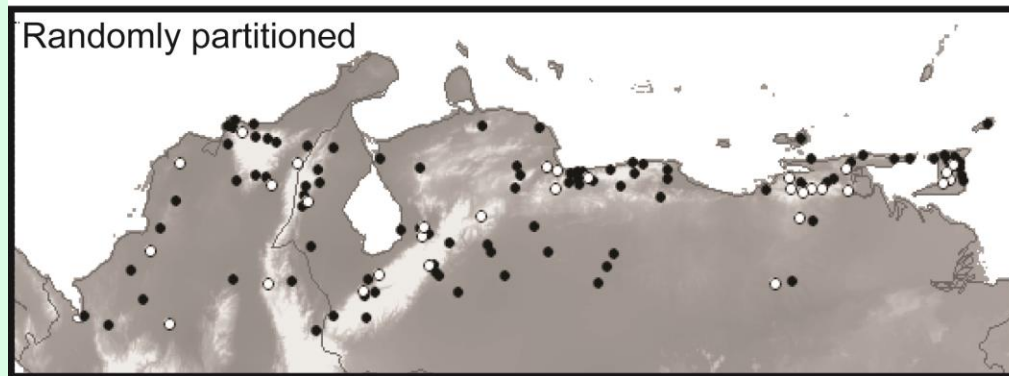
high = 

Omission:

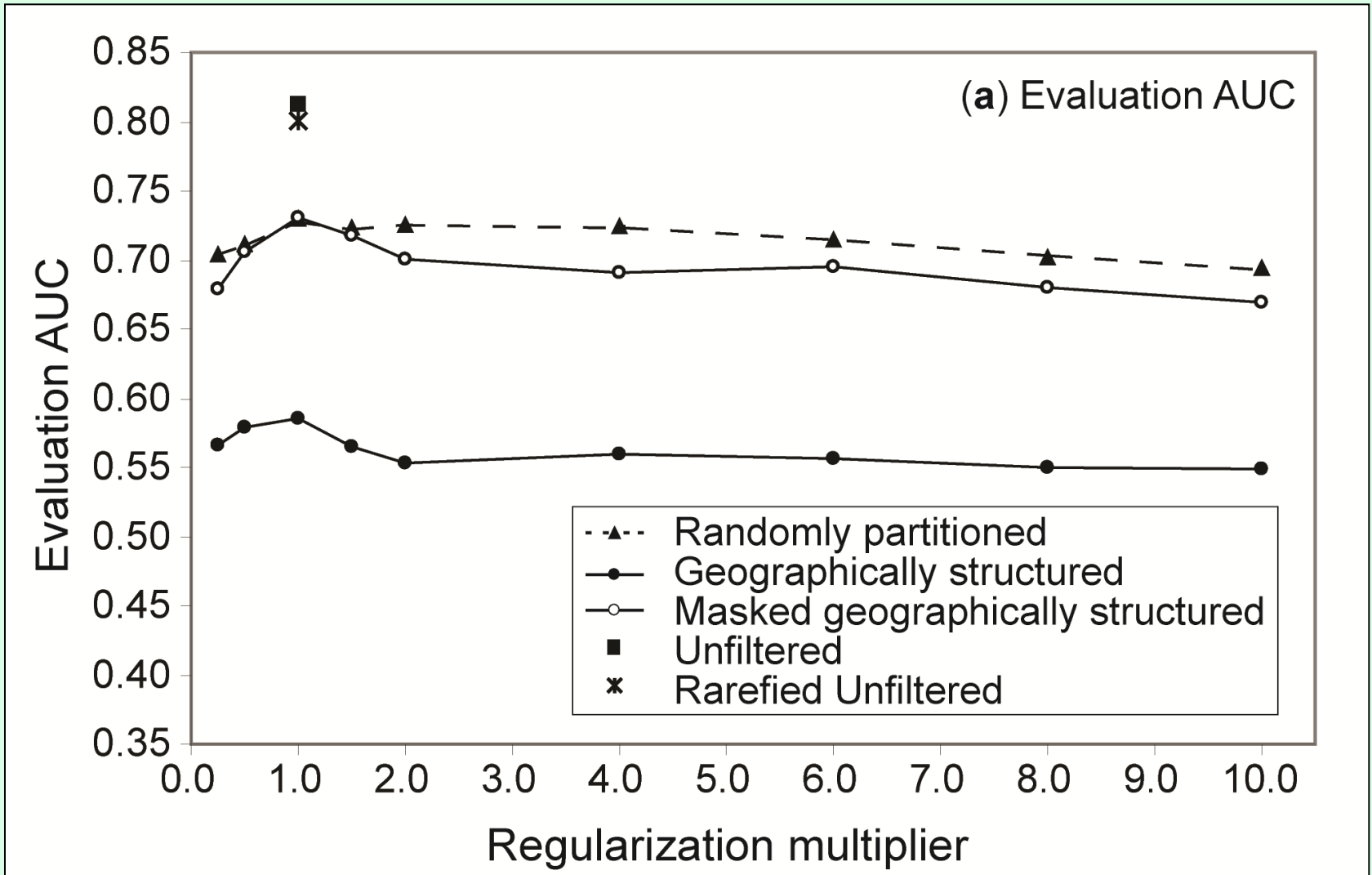
proportion of localities falling outside the prediction

low = 

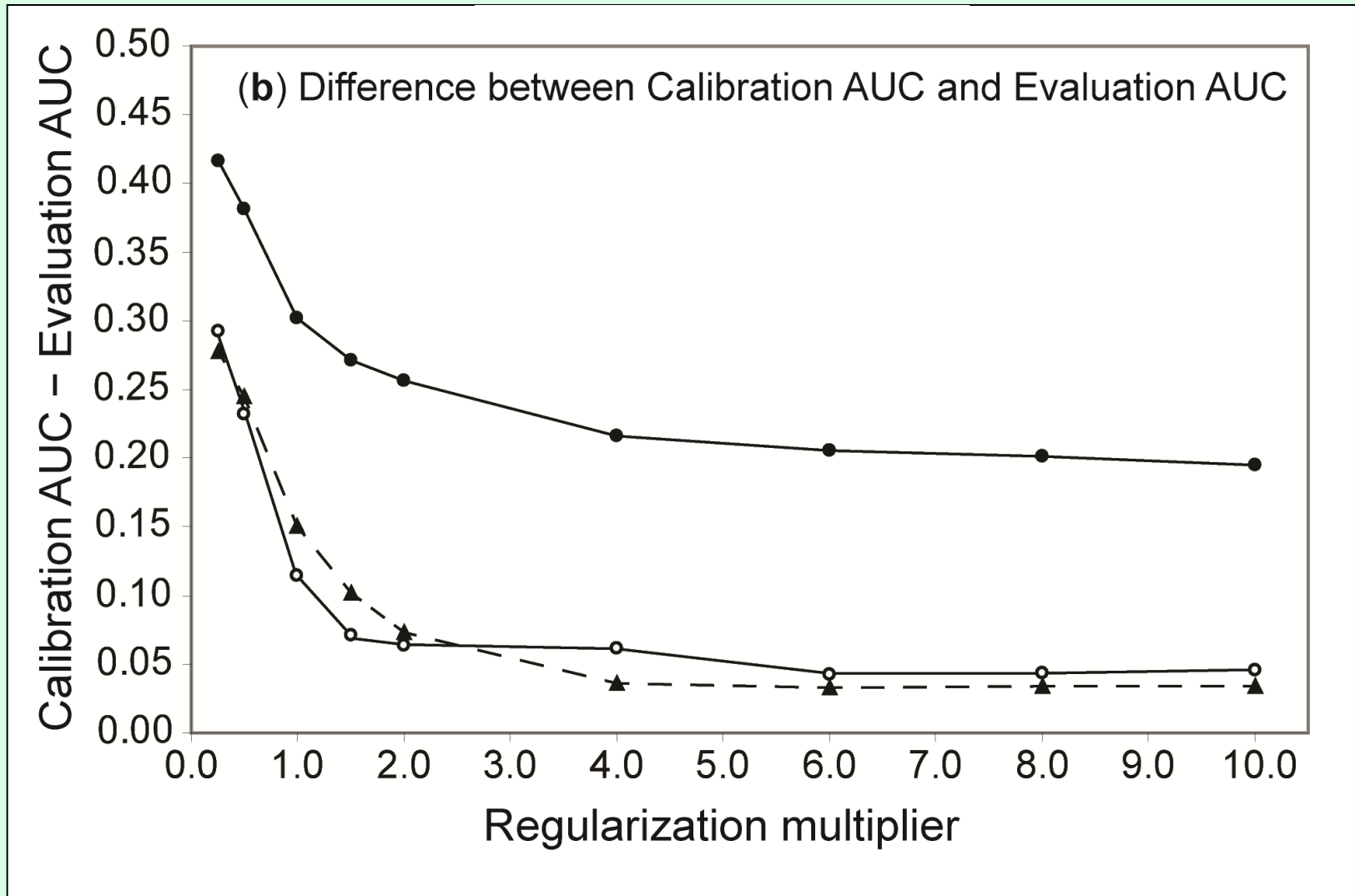
Presence/background sampling



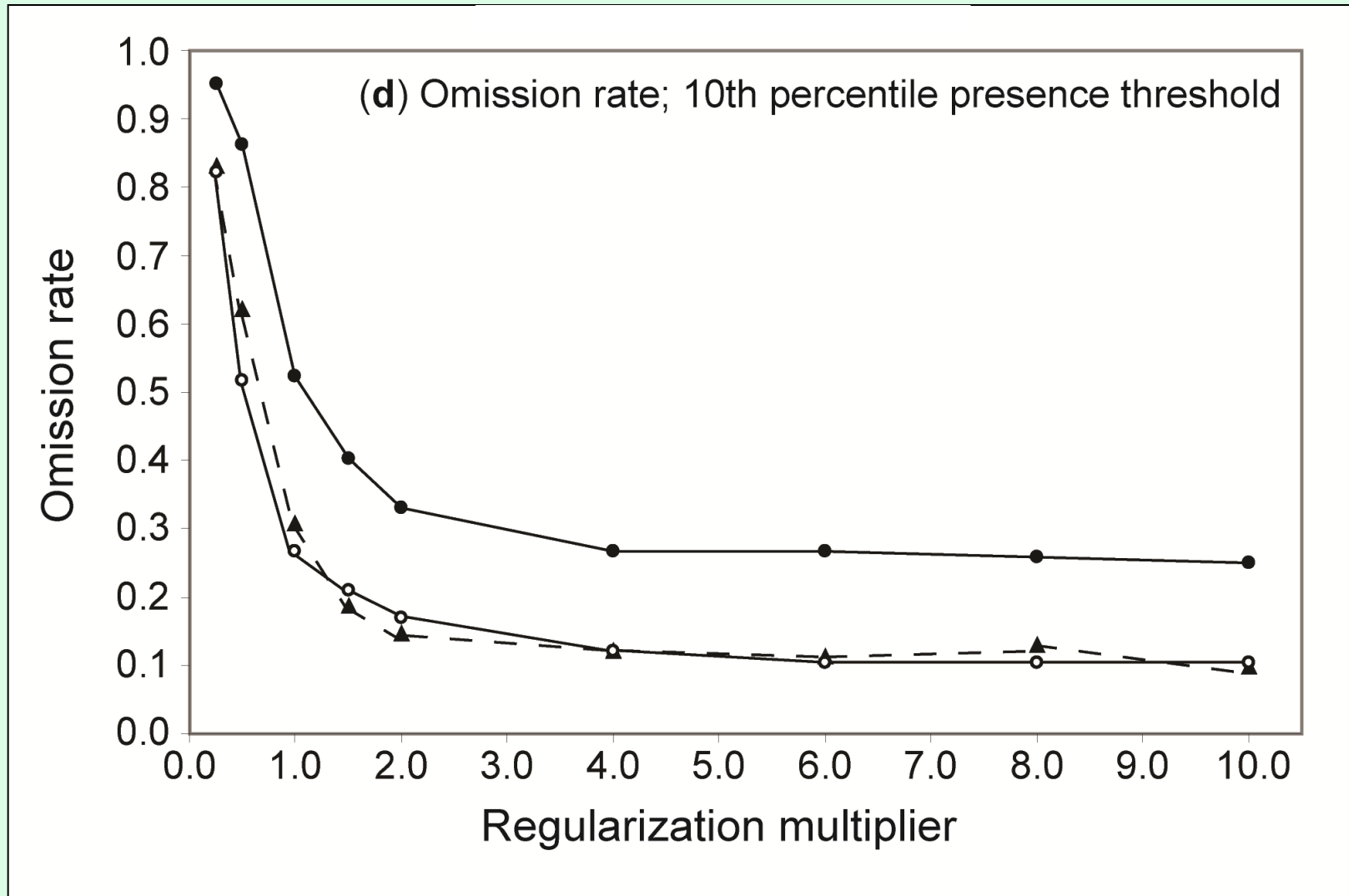
Results for AUC: random partitions (especially with unfiltered localities) overestimate performance



Results for AUC: overfitting decreases with higher regularization



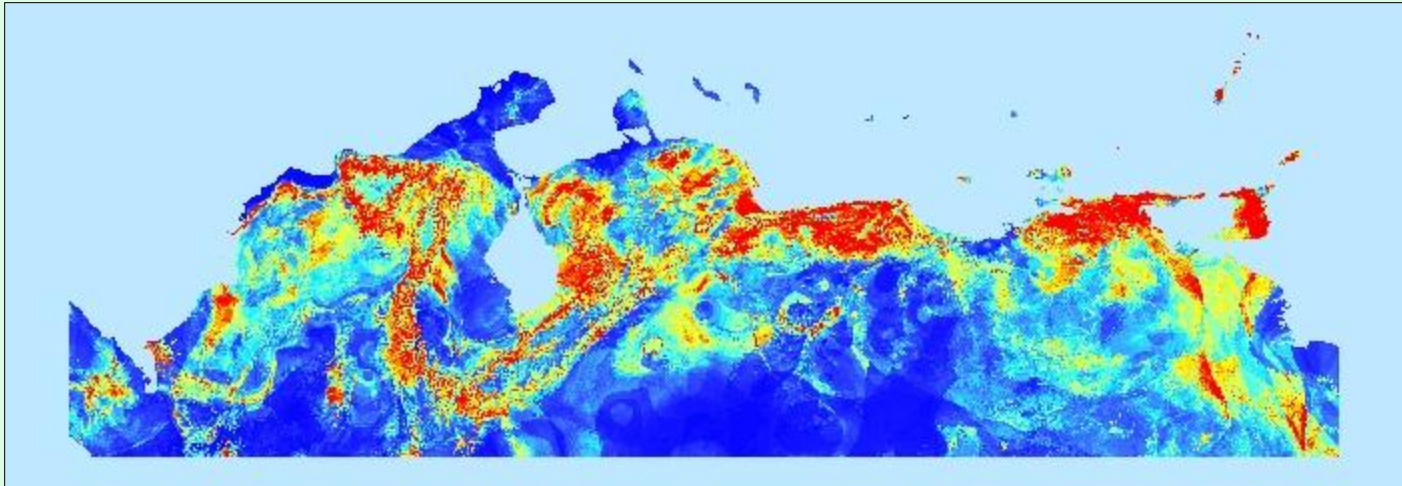
Results for omission rate: overfitting decreases with higher regularization



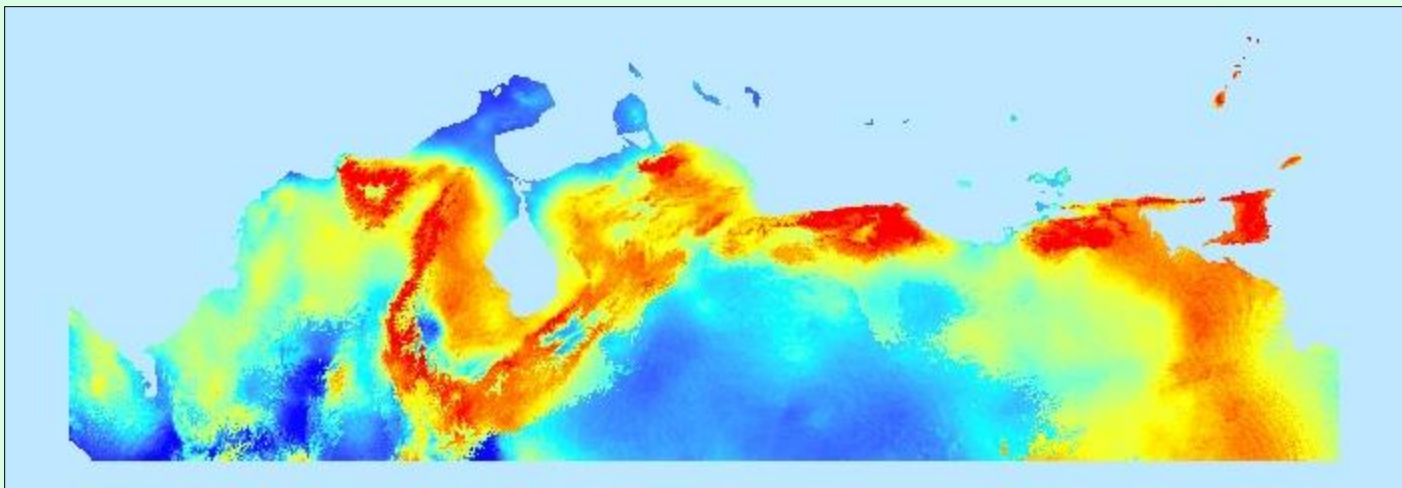
Geographic predictions

variation among regularization values

0.5



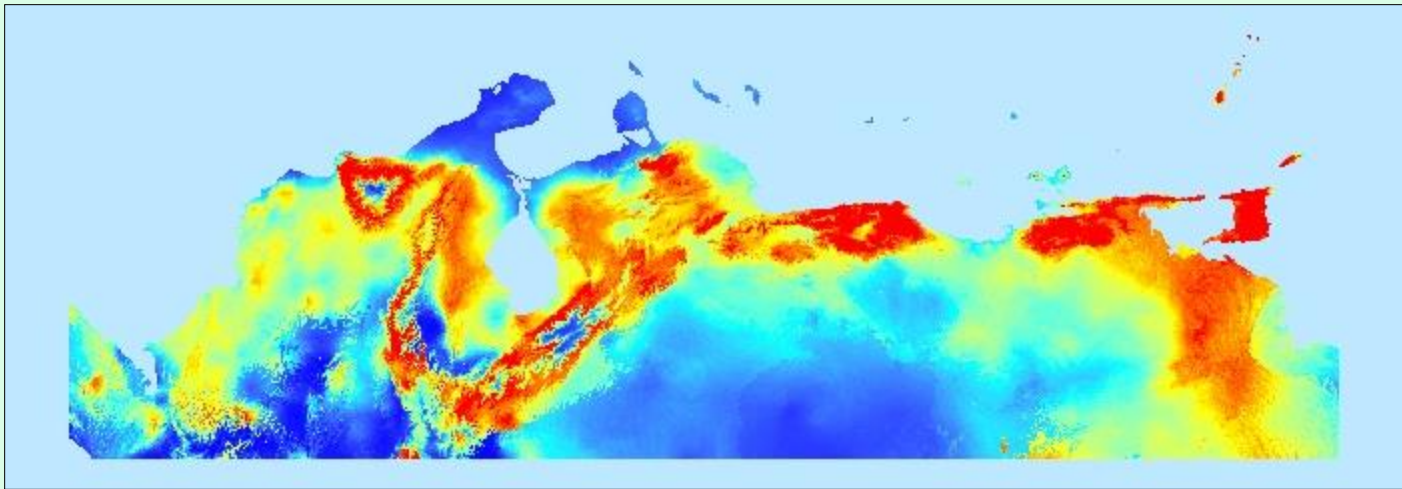
8.0



Optimal settings

Regularization 2.0

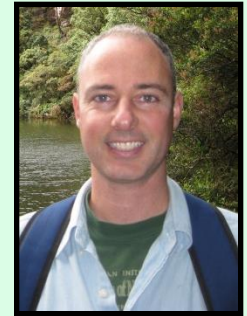
2.0



Thank you

NSF DEB-0717357 and DEB-1119915

PSC-CUNY 64215-00-42



<http://web.sci.ccny.cuny.edu/~anderson>

