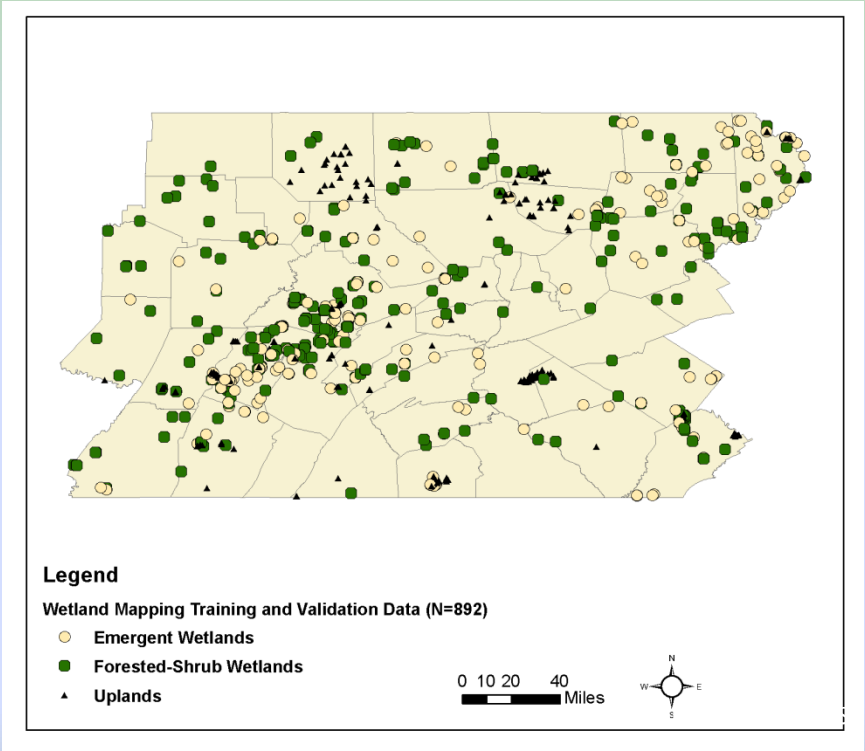
Upper Susquehanna Wetland Mapping Update: A hydrogeologic approach for identifying wetland plant communities in Pennsylvania

September 7th, 2016



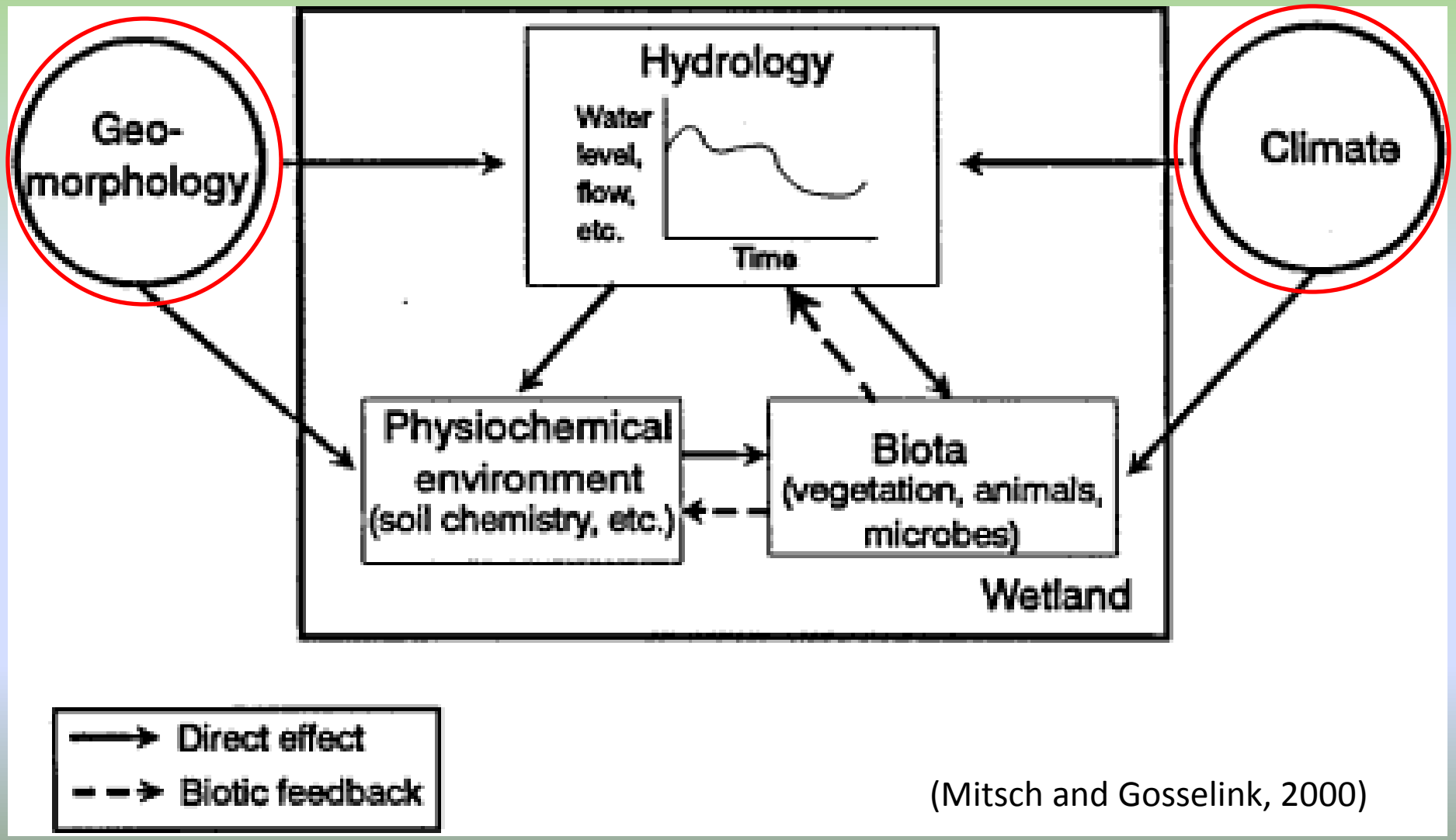
Patrick A. Raney, Ph.D.

Wetland Scientist

Upper Susquehanna Coalition

paraney@u-s-c.org

Climate, geomorphology as predictors



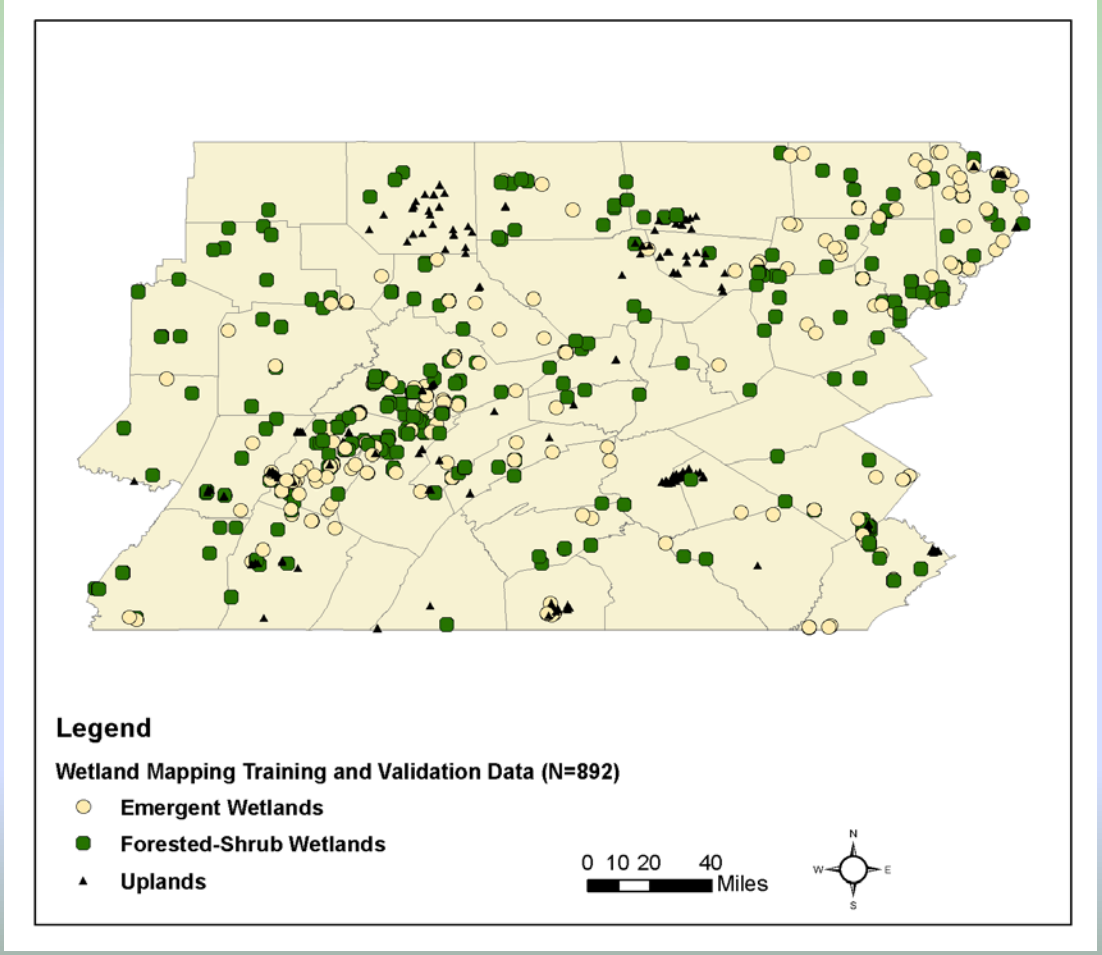
(Mitsch and Gosselink, 2000)

Modeling steps

- Lit review → informs predictors & scale
  - compile predictors in GIS
- Multi-collinearity screening
  - (i.e., identify correlations among predictors)
  - Reduce as necessary (e.g., PCA)
- Model training and validation Identify variables significantly related to function of interest
  - Overall model fit (R^2; model AIC; p-values)
- Review model statistics / visualize results and refine approach / predictors / models
  - Revisit above as needed…

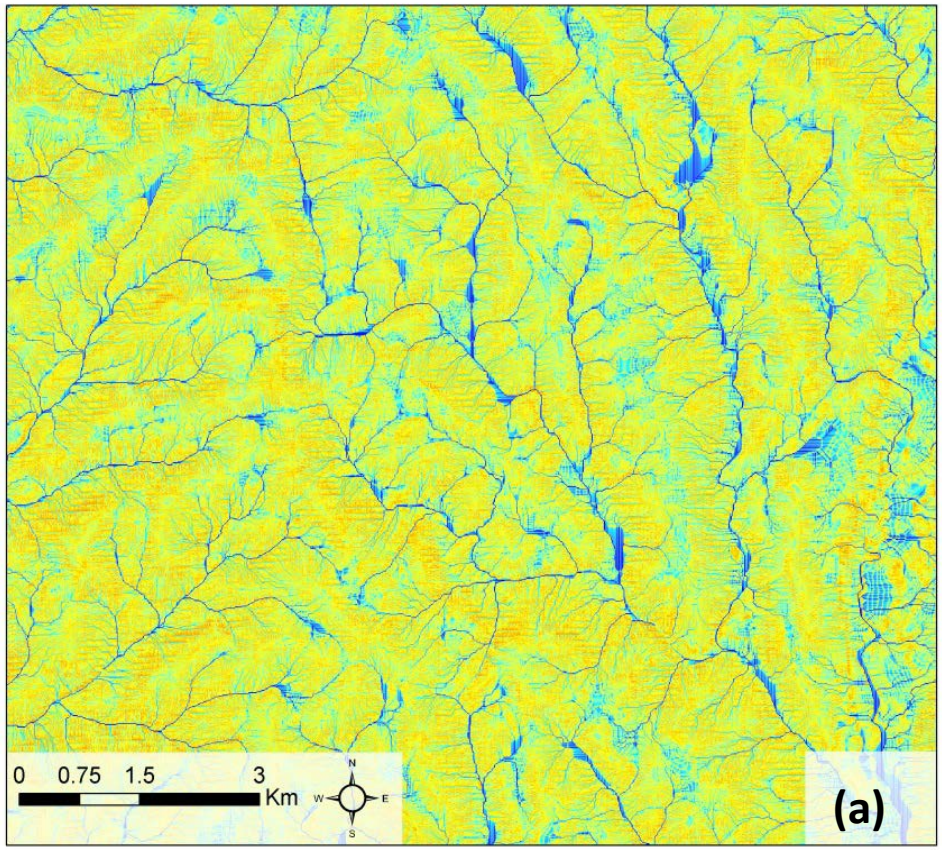
Compiled training/validation observations from Riparia and PA Natural Heritage Conservancy

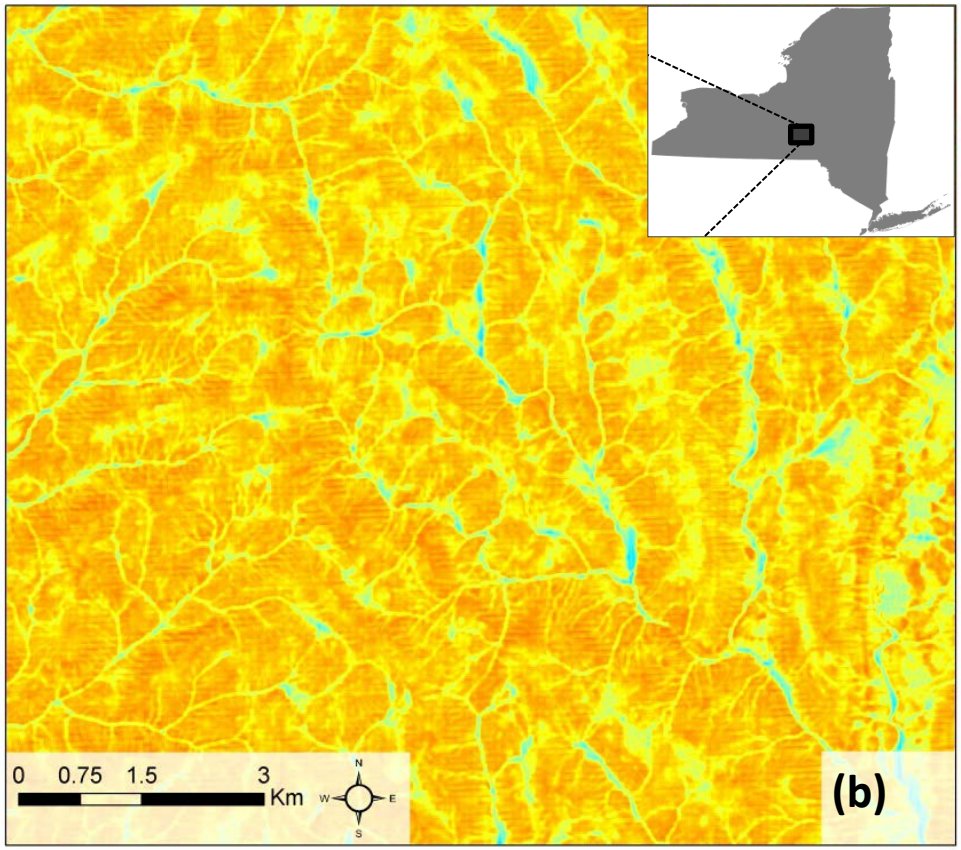
Absence locations – are upland sampling locations from the Heritage Database



Improving accuracy of wetland basin descriptors

(Raney et al. in prep)



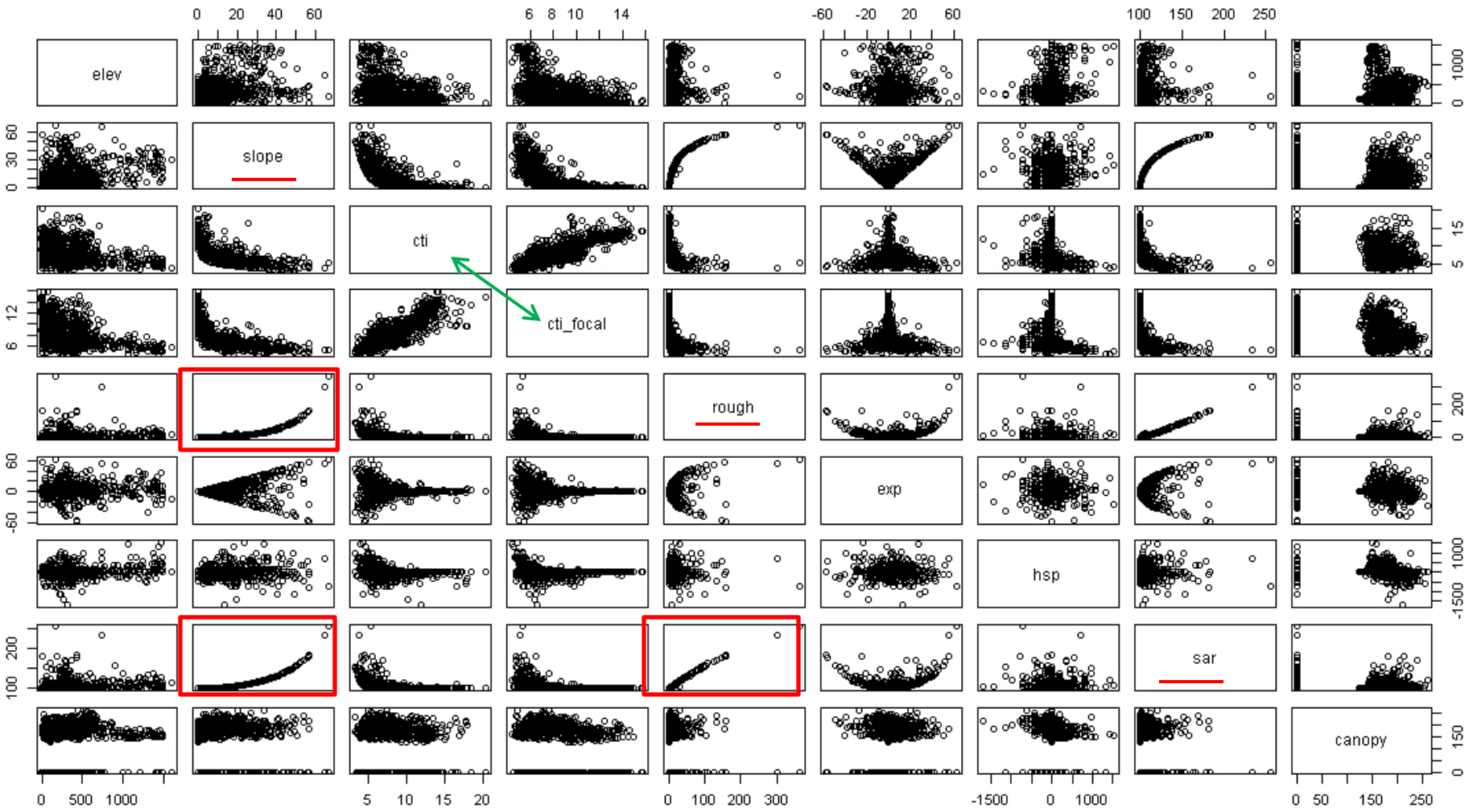


a) Standard Compound Topographic Index – **CTI (Beven and Kirkby 1979)**

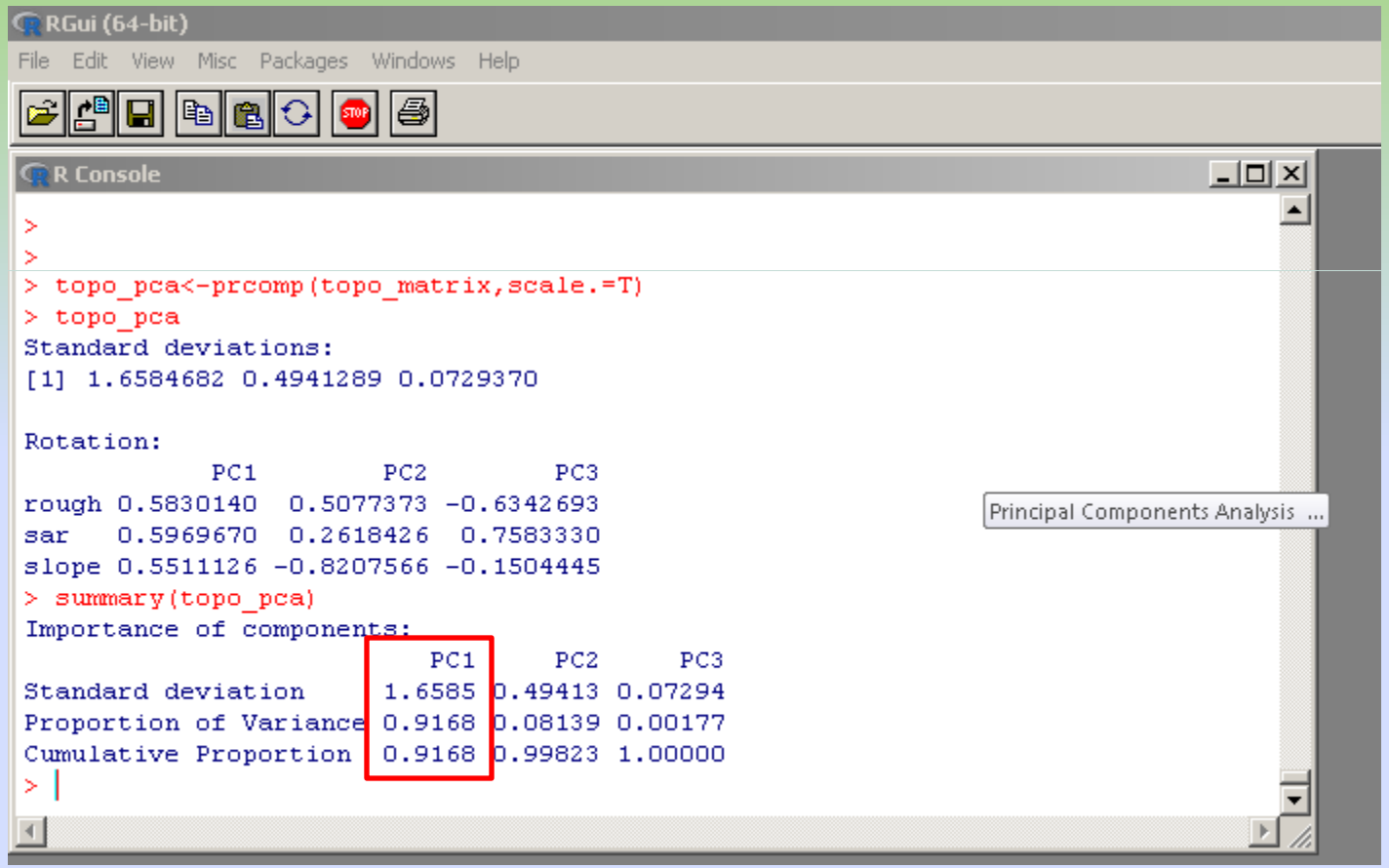
b) Five – cell neighborhood average calculated within window surrounding central cell – **CTI Focal Mean**

Screening continuous variables (predictors) for multicollinearity

```
> pairs(topographic)
> cor(sar,slope)
[1] 0.8518293
> cor(rough,sar)
[1] 0.9871922
```



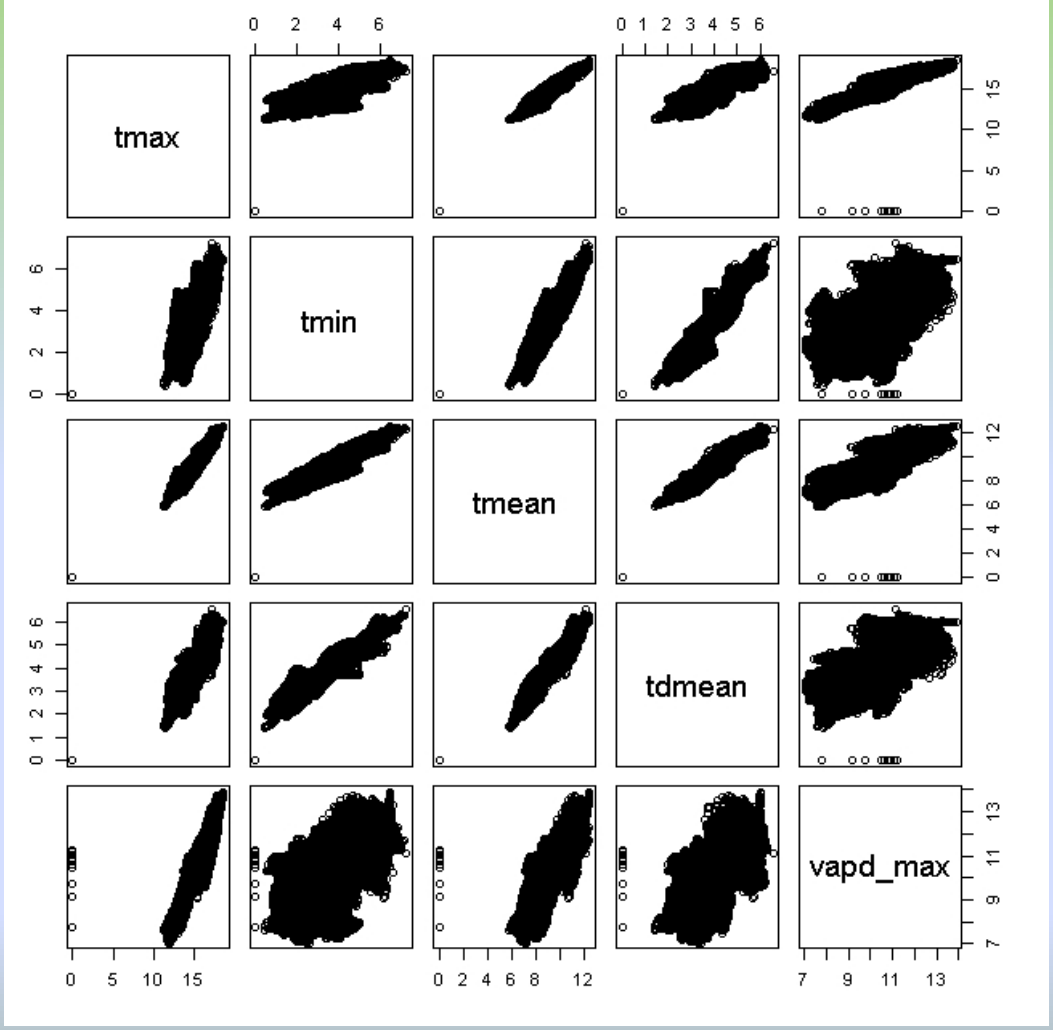
Principal components analysis:
correlated topographic variables



So – new variable called “topo_pca” describes slope, surface area ratio and surface roughness, and can be used as a possible predictor.

I’ll apply the same technique to 5 correlated PRISM climate variables.

Principal Components Analysis
and
Statistical Downscaling of PRISM Climate Data (800m to 10m)

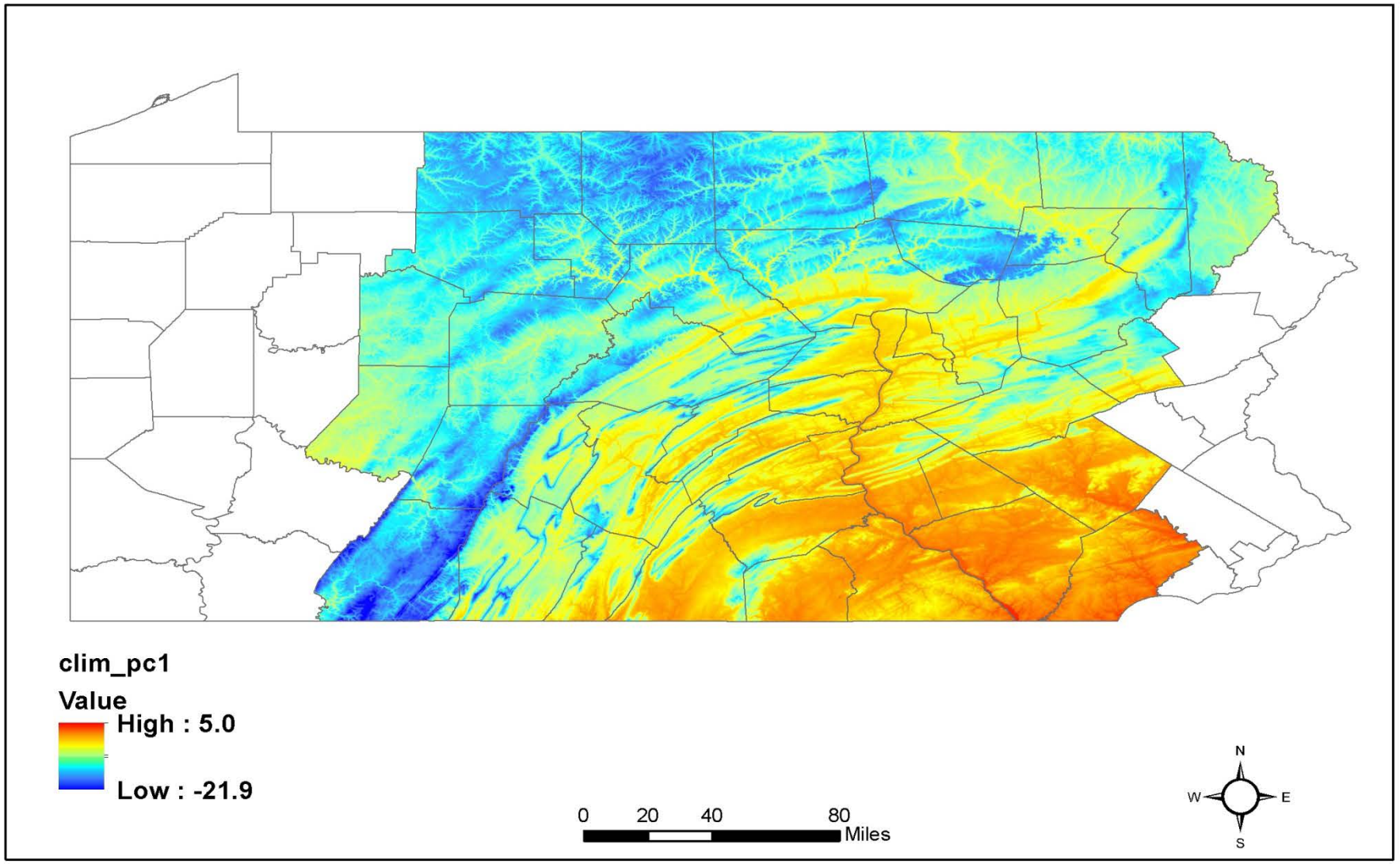


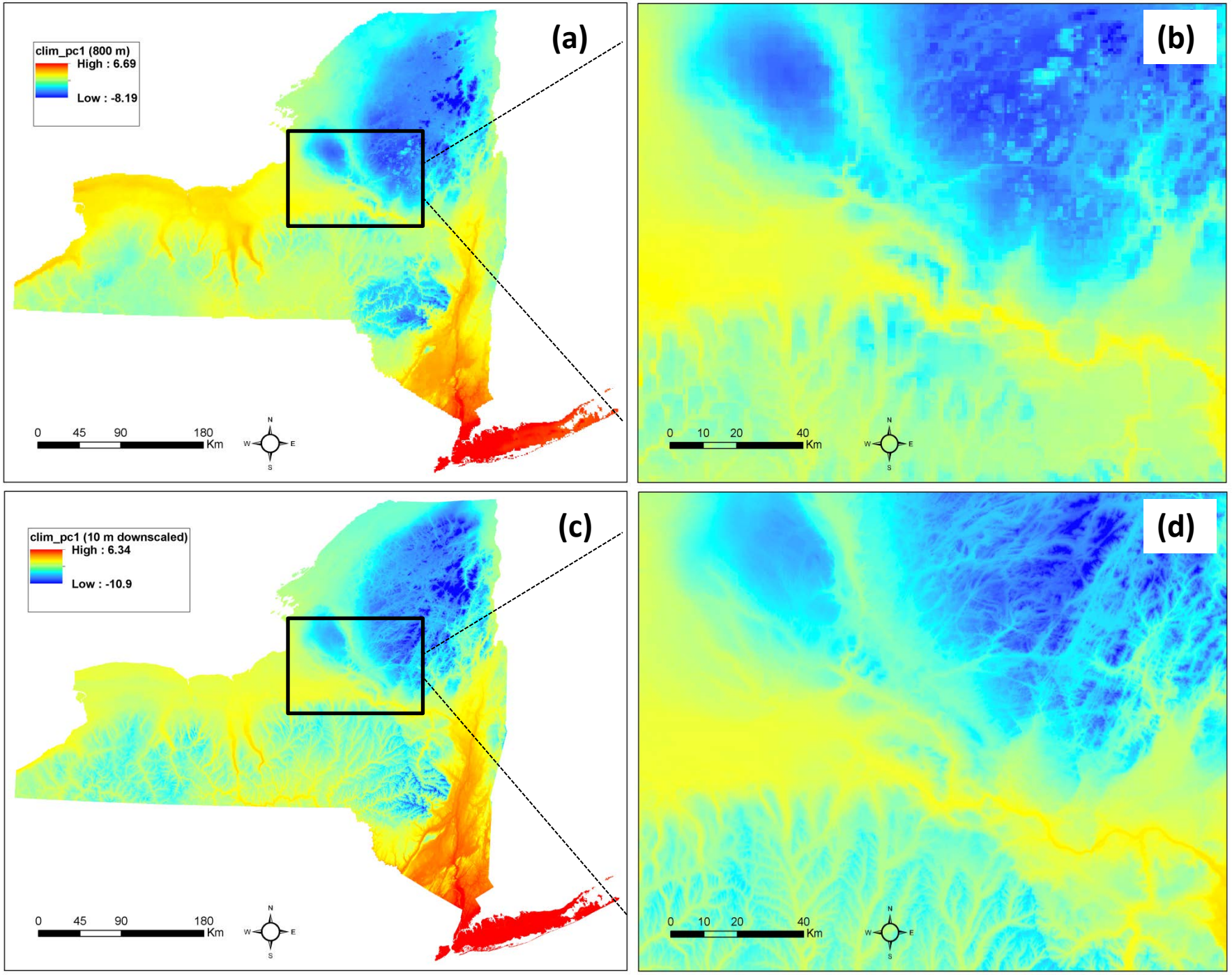
- Tmax, Tmin, Tmean, Dew Point, Vapor Pressure Deficit
- Data dimensionality reduced through a Principal Components Analysis in R version 3.3.1, retaining > 90% of variation in the dataset.
- Linear regression model:
- clim_pc1 ~ elev*northing*aspect
- (R^2 =0.97; on 3 and 60,814 DF on withheld data)

(Example from Central PA – similar results for NY)

PCA axis 1 from five correlated PRISM Climate Variables, downscaled to 10 meter resolution.

Warmer colors = greater evaporative demand





(Raney et al. *in prep*)

Wetland model fit statistics

Chesapeake Bay Wetland Mapping For Pennsylvania 2016 - Upper Susquehanna Coalition

Wetland predictive models of forested/scrub-shrub wetlands (PFO-PSS) and emergent (PSS) wetlands fitted by GAM method (MGCV library, Wood 2006) in *R* version 3.3.0 (R core Team, 2016).

- **PFO-PSS Validation**: R^2=0.86; AUC=0.99; n=350.
- **PEM Validation**: R^2=0.85; AUC=0.99; n=295.

*Model training and validation observations provided by Riparia and The Pennsylvania Natural Heritage Program.

Contact: Patrick A. Raney, Ph.D. USC Wetland Scientist, paraney@u-s-c.org

emergent wetland model validation statistics

```
Family: binomial
Link function: logit

Formula:
pem > 0 ~ s(cti_focal) + s(clim_pc1) + s(slope) + s(elev) + ssurgo +
    s(dissect) + s(expos) + s(hsp) + s(dissect) + s(cti) + s(moist) +
    s(rough) + forest

Parametric coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.072e+03  1.869e+03  -0.574   0.5660
ssurgo3      5.615e-01  1.741e+00   0.322   0.7471
ssurgo5      6.549e+02  3.875e+07   0.000   1.0000
ssurgo11     2.260e+00  2.887e+00   0.783   0.4338
ssurgo13    -2.145e+00  3.648e+00  -0.588   0.5565
forest1     -2.555e+00  1.094e+00  -2.336   0.0195 *  ←forest cover
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
               edf Ref.df Chi.sq p-value
s(cti_focal) 1.000  1.000  4.040  0.0445 * ←moisture index
s(clim_pc1)  1.000  1.000  2.063  0.1509
s(slope)     6.492  6.773 11.933  0.0909 .
s(elev)      8.308  8.594 14.913  0.0903 .
s(dissect)   6.578  7.645 14.984  0.0465 * ←topographic dissection
s(expos)     1.000  1.000  1.987  0.1587
s(hsp)       1.000  1.000  3.208  0.0733 .
s(cti)       6.042  7.084  8.765  0.2778
s(moist)     2.140  2.430  3.758  0.1927
s(rough)     1.000  1.000  4.585  0.0323 * ←surface roughness
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.854   Deviance explained = 85.6%
```

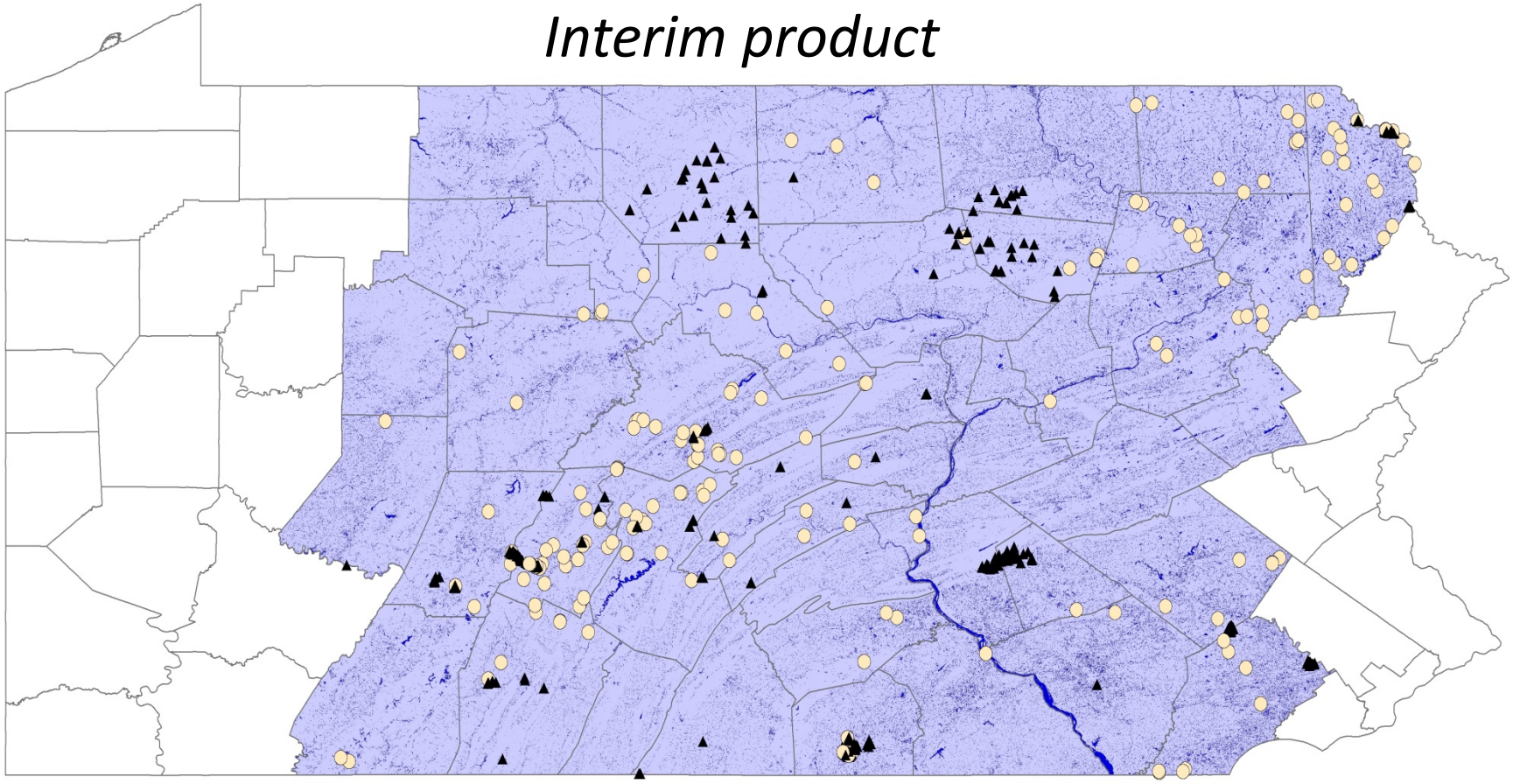
- Variables inclusion selected based on model AIC.

Forested wetland model validation

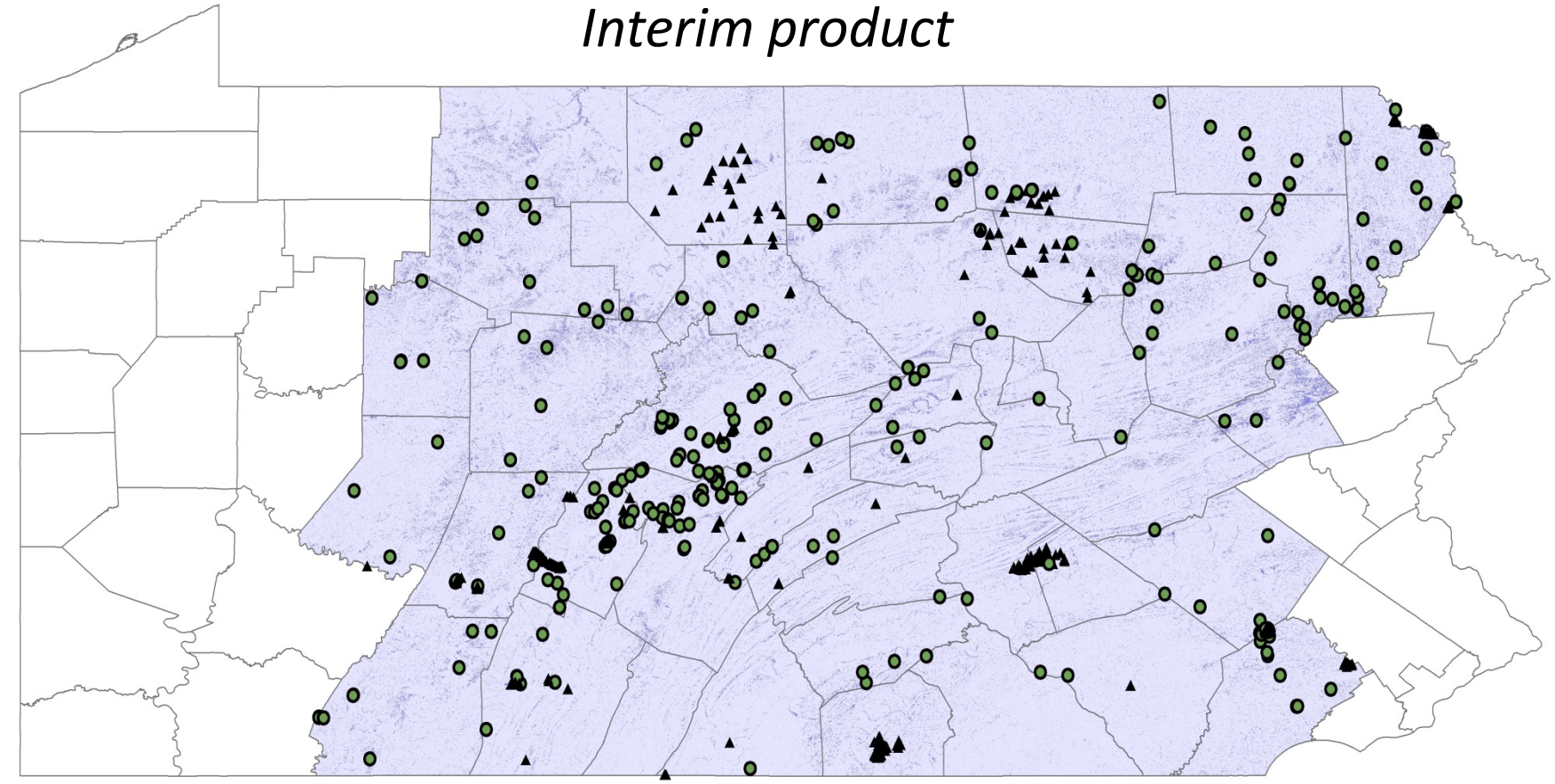
```
Formula:
pfo.pss > 0 ~ s(cti_focal) + s(clim_pc1) + (forest) + ssurgo +
    s(slope) + s(elev) + s(dissect) + s(hsp) + s(rough) + s(expos) +
    s(moist)

Parametric coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.082e+02  1.345e+02   1.547   0.1218
forest1      2.457e+00  1.437e+00   1.709   0.0874 .
ssurgo3     -3.254e+00  1.792e+00  -1.816   0.0693 .
ssurgo5      2.375e+02  1.861e+07   0.000   1.0000
ssurgo11     1.226e+00  3.952e+00   0.310   0.7563
ssurgo13    -2.394e+02  3.001e+07   0.000   1.0000
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
               edf Ref.df Chi.sq  p-value
s(cti_focal) 1.000  1.000 13.511 0.000237 ***  ←moisture index
s(clim_pc1)  4.741  5.874 18.824 0.004425 **   ←climate index
s(slope)     5.845  6.748 12.200 0.115845
s(elev)      9.000  9.000 21.005 0.012623 *  ←elevation
s(dissect)   6.703  7.532  8.842 0.338518
s(hsp)       2.358  2.984  5.682 0.122486
s(rough)     1.000  1.000  1.717 0.190044
s(expos)     9.000  9.000 16.555 0.056158 .
s(moist)     8.912  8.986 13.580 0.137372
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =   0.86   Deviance explained = 86.5%
UBRE = -0.50534  Scale est. = 1         n = 350
```

- Variables inclusion selected based on model AIC.

Interim product



Interim product



PFO-PSS Wetland Model - GAM Method

High : 1

Low : 0

PFO-PSS

Absences

0 20 40 80 Miles

N

W E

S

Next steps

- Filter model output with high-resolution CBP landcover product from UVM
- Provide UVM with updated predictive surfaces
- Evaluate accuracy of UVM's intermediate product using the withheld model testing observations