

----- DATASET SELECTION AND SETUP -----

NOTE: The desired study region must be specified as 'CONUS' if one wishes to execute the logistic regression model over states that contain zero commercial wind farms (Louisiana, Mississippi, Alabama, Georgia, South Carolina, Kentucky), states that possess wind farms in only one grid cell at all but the highest spatial resolutions (Arkansas, Florida, Virginia, Delaware, Connecticut, New Jersey, Tennessee), or states at low spatial resolutions at which too many predictors were removed due to collinearity (Rhode Island at the 100th or 80th percentile).

Specified study region: Indiana

Specified wind farm density: 85 acres/MW

Specified wind power capacity: 100th percentile (525 MW)

Predictor configurations specified by the user: ['Full', 'No_Wind', 'Wind_Only', 'Reduced']

Predictors removed from the model based on having a constant value in all grid cells: ['ISO_YN']

----- TESTING ASSUMPTIONS -----

Assumption #1: All continuous predictors have a linear relationship with the logit of the dependent variable, based on a Box-Tidwell test.

Bonferroni-corrected p-value: 0.001388888888888889

Results of the Box-Tidwell test:

Predictor	pval
Undev_Land	0.001320
Avg_Temp	0.008362
Near_Plant	0.013838
Hisp_15_19	0.021807
Whit_15_19	0.037904
Avg_25	0.062905
Type_15_19	0.071564
Prop_Rugg	0.097851
Fem_15_19	0.215330
Avg_Elevat	0.239317
Near_Sch	0.277603
Near_Trans	0.283929
Near_Air	0.286025
Avg_Wind	0.492689
Dens_15_19	0.540025
Near_Roads	0.836351
Near_Hosp	0.871329
Unem_15_19	0.929700

Predictors removed from the model based on the results of the Box-Tidwell test: ['Undev_Land']

Assumption #2: There is no multicollinearity, or pairwise collinearity, between the model's predictors, based on Variance Inflation Factors (VIF).

Grouped Multicollinearity Test Results:

Predictor	VIF
Trib_Land	1.081770
Farm_Year	1.106703
Plant_Year	1.176663
Military	1.189472
Wild_Refug	1.192464
Near_Roads	1.219197
Near_Trans	1.284194
Near_Plant	1.336924
Near_Air	1.353504
Nat_Parks	1.447600
Fem_15_19	1.456247
Bat_Count	1.471641
Near_Hosp	1.490045
Near_Sch	1.528293
Critical	1.575129
Dens_15_19	1.686932
Avg_25	1.895826
Unem_15_19	1.940115
Hisp_15_19	2.306988
Bird_Count	2.337182
Prop_Rugg	2.732418
Dem_Wins	2.812159
Historical	3.619759
Mining	3.801296
Avg_Elevat	3.899186
Avg_Wind	3.919154
supp_2018	4.943697
Avg_Temp	5.186219
Type_15_19	5.419502
Whit_15_19	9.579433

Pairwise Multicollinearity Test Results:

Predictor1	Predictor2	VIF	
Plant_Year	Trib_Land	1.000000	
Dem_Wins	Historical	1.000000	
Avg_Temp	Dem_Wins	1.000000	
Mining	Wild_Refug	1.000000	
Avg_Wind	Bat_Count	1.000002	

Type_15_19	supp_2018	1.738395	
Avg_Elevat	Avg_Temp	1.844099	
Avg_Temp	Avg_Wind	2.069999	
Whit_15_19	supp_2018	2.732274	
Type_15_19	Whit_15_19	4.067899	

Predictors to be removed based on multicollinearity: None

Assumption #3: None of the grid cells contain data that represent extreme outliers, based on a Cook's distance test.

Number of grid cells removed due to outlying observations according to a Cook's distance test: 0

Final list of predictors that did not pass the model's three assumptions:
['ISO_YN', 'Undev_Land']

Full Configuration Output Begins

----- MODEL CALIBRATION (Training Data): Full Configuration -----

Range of log-likelihood scores from 30 training runs of the Full model:

Maximum Score: 8.681495295400396

Median Score: -4.668437921798613

Minimum Score: -14.8485132295433

Range of log-likelihood scores of the Null model:

Maximum Score: -32.77893094432278

Median Score: -32.77893094432278

Minimum Score: -38.53070767895383

Number of times (out of 30) the Full model possesses a greater goodness-of-fit: 30

Number of times (out of 30) the Full model's outperformance of the Null model is statistically significant: 30

Median Log-Likelihood Ratio, Full model vs. Null model: 56.220986045048335

p-value of the Median Log-Likelihood Ratio: 0.0025751160528970567

Range of McFadden Adjusted Psuedo R-Squared statistics for the Full model:

Minimum Pseudo R-Squared: -0.25700564635266665

Median Pseudo R-Squared: -0.0594562890030631

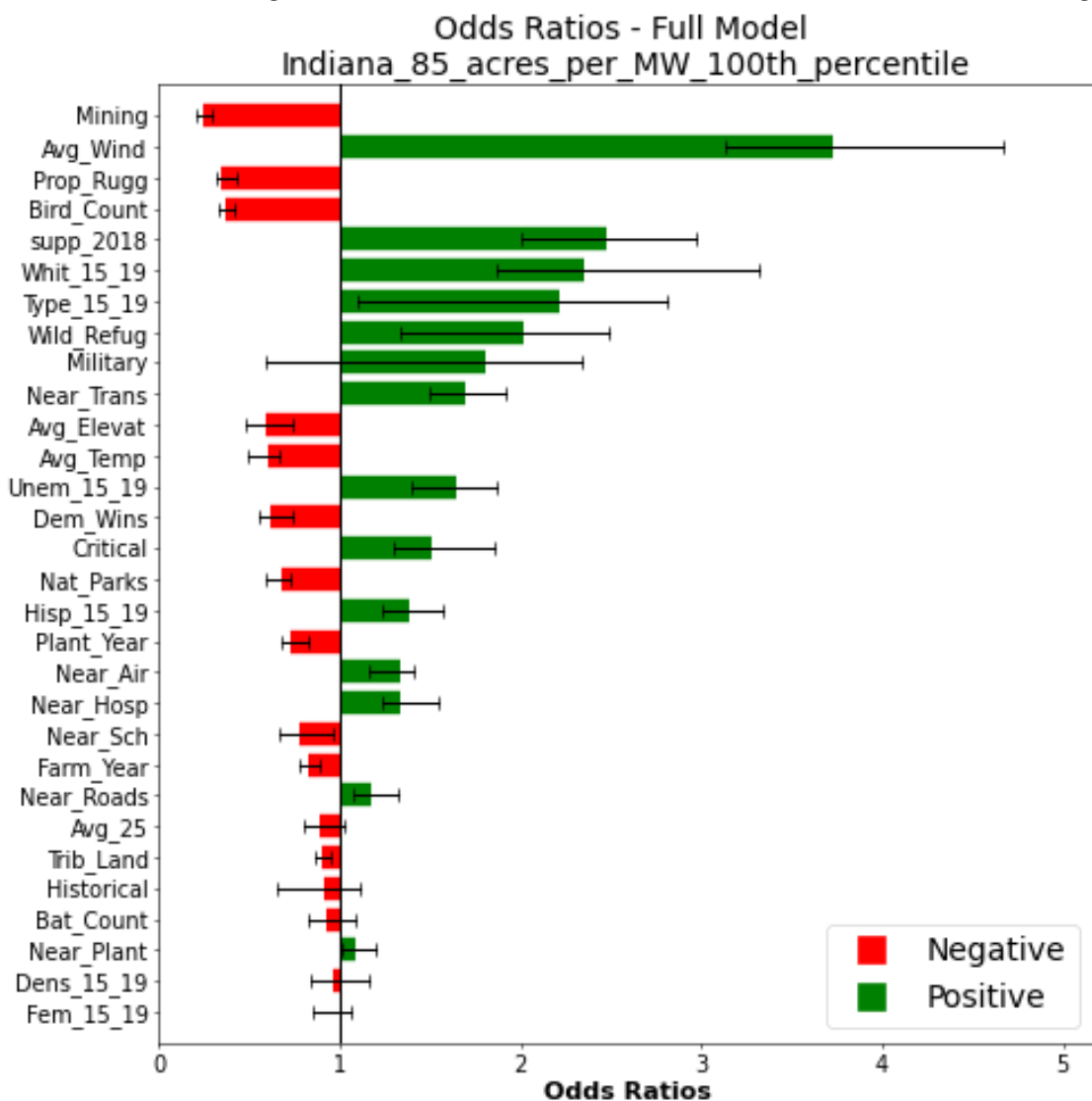
Maximum Pseudo R-Squared: 0.319120420903626

The following dataframe summarizes the coefficients and odds ratios obtained from fitting the Full model to the aggregated dataset. Predictors are ranked by the magnitude of their coefficients to convey strength of association:

Predictor	Odds_Low	Odds_Med	Odds_Upp	Coef_Med	Rank
Mining	0.201828	0.238904	0.287230	-1.431692	1
Avg_Wind	3.128274	3.725337	4.670841	1.315157	2
Prop_Rugg	0.315848	0.346025	0.425316	-1.061243	3
Bird_Count	0.332443	0.367995	0.422256	-0.999687	4
supp_2018	2.003476	2.472616	2.972142	0.905277	5
Whit_15_19	1.861646	2.356795	3.315121	0.857303	6
Type_15_19	1.099253	2.211397	2.810572	0.793625	7
Wild_Refug	1.335401	2.014301	2.488554	0.700272	8
Military	0.596042	1.803217	2.342845	0.589572	9
Near_Trans	1.498417	1.688049	1.914127	0.523574	10
Avg_Elevat	0.478585	0.592454	0.741872	-0.523482	11
Avg_Temp	0.497003	0.606967	0.662122	-0.499281	12
Unem_15_19	1.397998	1.647053	1.866034	0.498987	13
Dem_Wins	0.554064	0.616700	0.736966	-0.483373	14
Critical	1.292148	1.504283	1.857559	0.408316	15

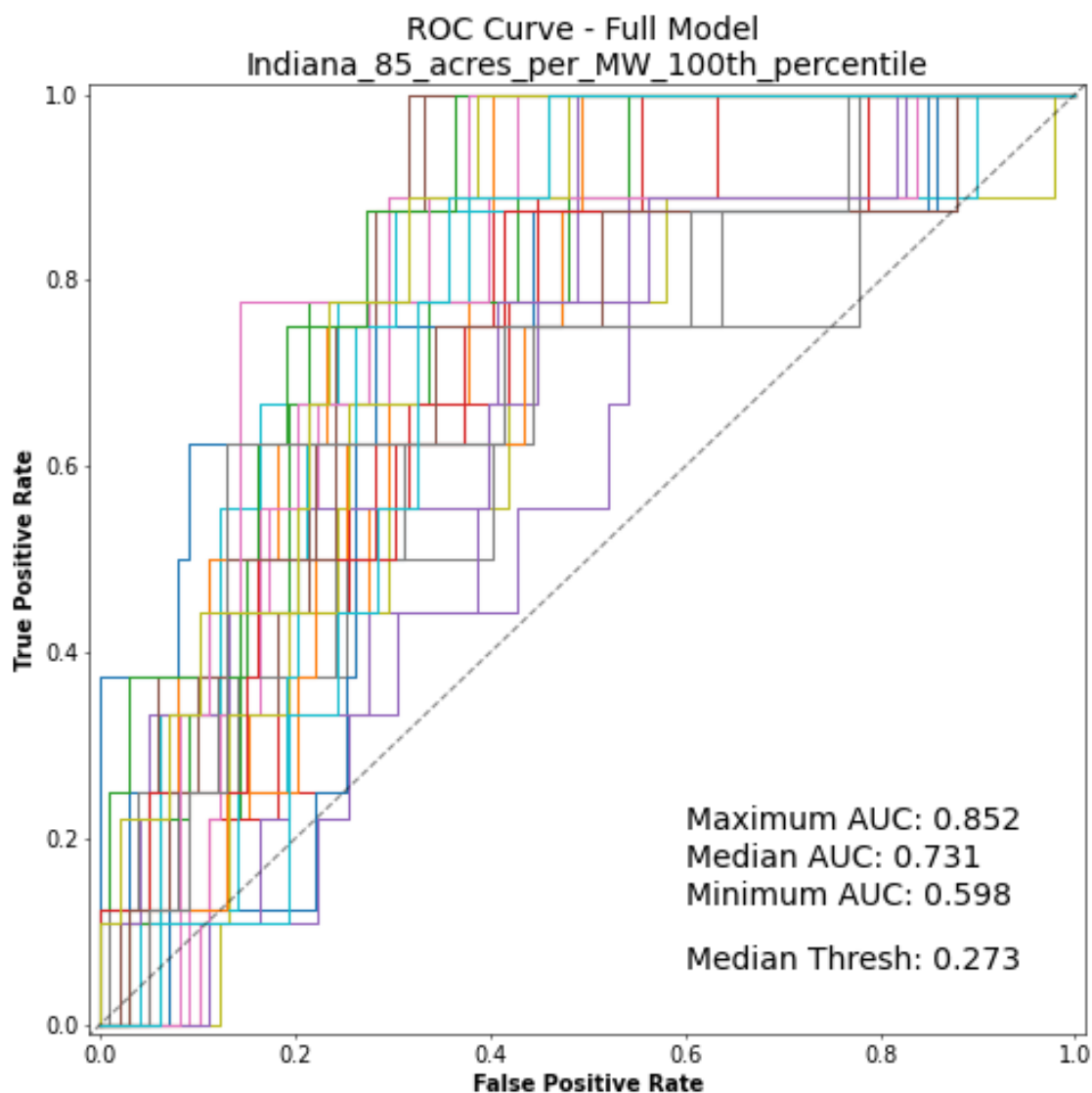
Nat_Parks	0.593630	0.680947	0.731568	-0.384270	16
Hisp_15_19	1.238561	1.389101	1.568186	0.328657	17
Plant_Year	0.673787	0.722775	0.830313	-0.324657	18
Near_Air	1.161931	1.330544	1.405315	0.285588	19
Near_Hosp	1.235778	1.328770	1.539975	0.284254	20
Near_Sch	0.662444	0.771993	0.965802	-0.258780	21
Farm_Year	0.771318	0.821359	0.887232	-0.196794	22
Near_Roads	1.078422	1.169210	1.327782	0.156328	23
Avg_25	0.805037	0.888935	1.023628	-0.117731	24
Trib_Land	0.858642	0.896932	0.953465	-0.108775	25
Historical	0.656767	0.907008	1.110499	-0.097604	26
Bat_Count	0.830935	0.922302	1.091640	-0.080883	27
Near_Plant	1.006093	1.084190	1.198305	0.080833	28
Dens_15_19	0.835709	0.966289	1.159347	-0.034292	29
Fem_15_19	0.851243	1.004098	1.061889	0.004090	30

Odds Ratio chart generated from the 30 Full model runs with the training data:



----- MODEL Validation (Testing Data): Full Configuration -----

ROC curves generated from the 30 Full model runs with the testing data:



Range of Area Under Curve (AUC) statistics for the Full model:

Minimum AUC: 0.5986394557823129

Median AUC: 0.7312925170068028

Maximum AUC: 0.8522727272727273

Range of optimal threshold classifications for the Full model:

Minimum Threshold: 0.03188237194776313

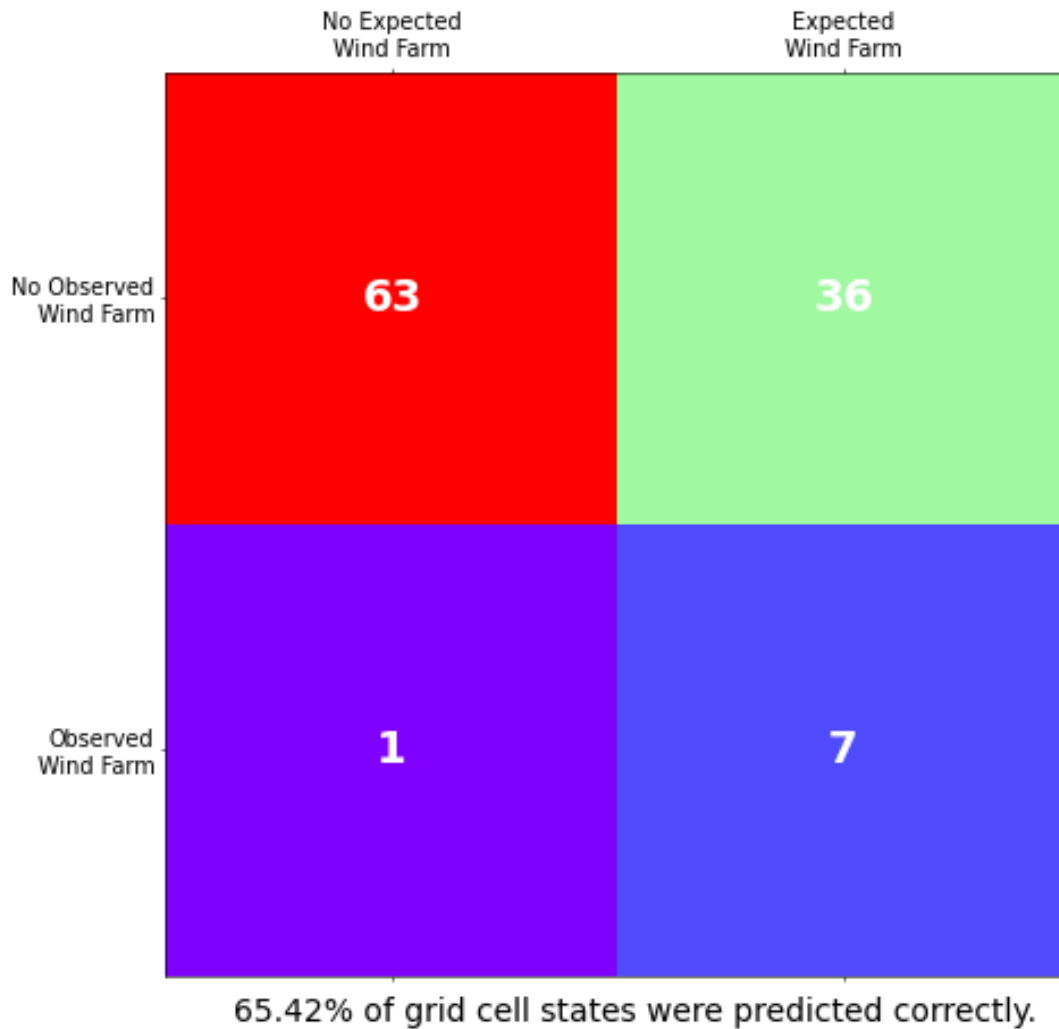
Median Threshold: 0.2735953276986943

Maximum Threshold: 0.7258284359531145

Median Confusion Matrix of the Full model's predictive accuracy:

Confusion Matrix - Full Model

Indiana_85_acres_per_MW_100th_percentile



Below are the range of confusion matrix results from the 30 Full model runs with the testing data:

Lower Quartile confusion matrix:

```
[[53 45]
 [ 1  8]]
```

Lower Quartile proportion of correctly predicted grid cell states by the Full model:
0.5700934579439252

Median confusion matrix:

```
[[63 36]
 [ 1  7]]
```

Median proportion of correctly predicted grid cell states by the Full model:
0.6542056074766355

Upper Quartile confusion matrix:

```
[[69 29]
 [ 2  7]]
```

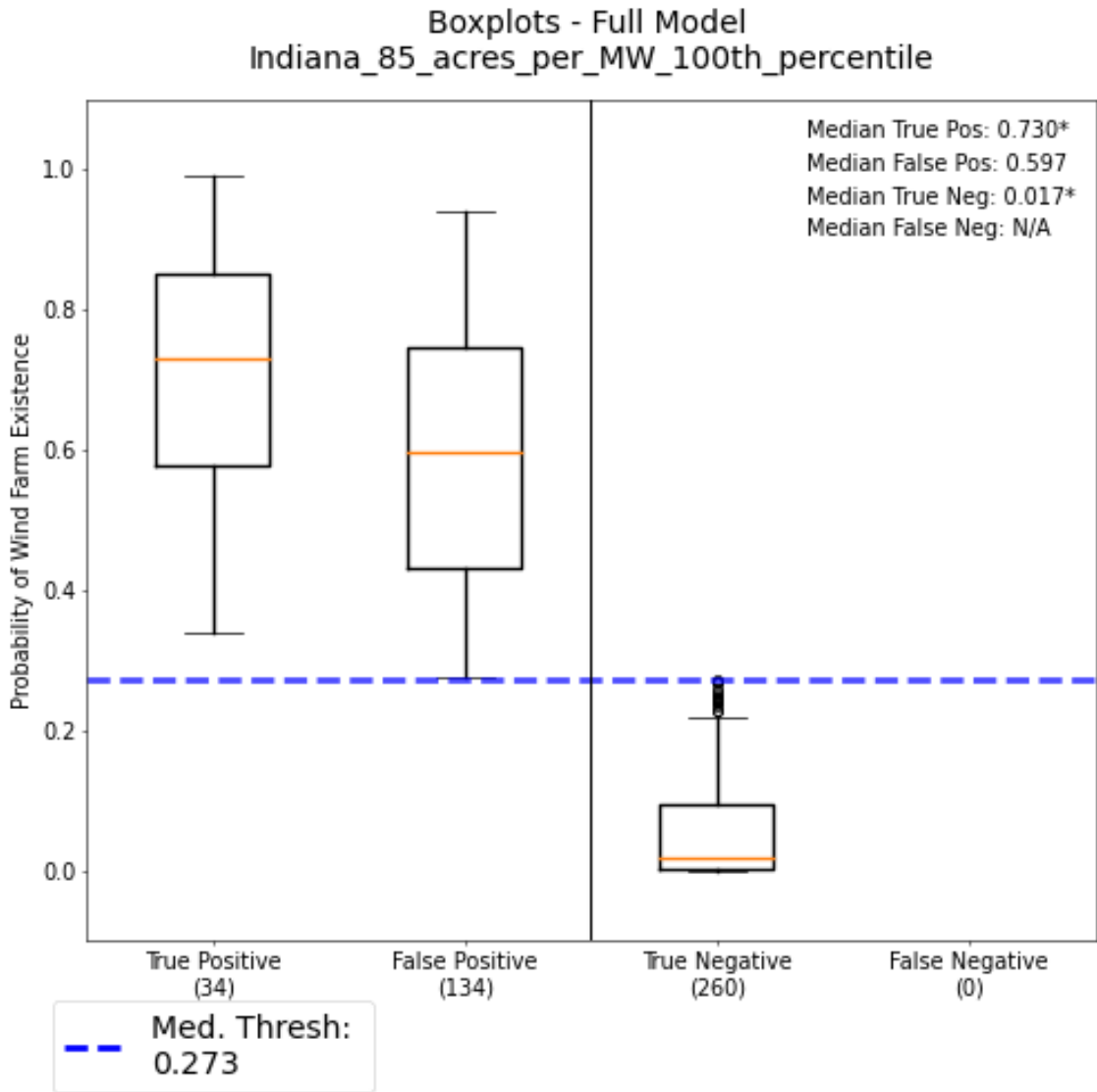
Upper Quartile proportion of correctly predicted grid cell states by the Full model:
0.7102803738317757

----- BOXPLOT CONSTRUCTION (All Data): Full Configuration -----

Grid cell classifications from executing the trained and tested Full model over all grid cells in Indiana:

Number of True Positive Grid Cells: 34
Number of False Positive Grid Cells: 134
Number of True Negative Grid Cells: 260
Number of False Negative Grid Cells: 0

Boxplot of grid cell probabilities in each classification:



Median probabilities of wind farm existence for each grid cell classification.
An asterisk indicates a Mann-Whitney U-test result that is statistically significant ($p < 0.05$):

Median False Pos: 0.597
Median True Pos: 0.730*
Median False Neg: N/A
Median True Neg: 0.017*

Mann-Whitney U-test results:

Mann-Whitney Statistic - True Positive vs False Positive:

U-statistic = 1524.0

p-value = 0.0029330469861638345

Mann-Whitney Statistic - True Negative vs False Negative:

U-statistic = 0.0

p-value = 0.0

----- MAP CONSTRUCTION: Full Configuration -----

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Model_Testing\WiFSS_Surfaces\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana_Full.gdb\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana_Full_Map

Total (Percentage) of all grid cells over Indiana that exist in hotspots:

20 (4.67%)

Total (Percentage) True Positive grid cells over Indiana that exist in hotspots:

6 (17.65%)

Total (Percentage) False Positive grid cells over Indiana that exist in hotspots:

14 (10.45%)

No_Wind Configuration Output Begins

----- MODEL CALIBRATION (Training Data): No_Wind Configuration -----

Range of log-likelihood scores from 30 training runs of the No_Wind model:

Maximum Score: -0.6830272882328927

Median Score: -9.098394303829934

Minimum Score: -14.118426872707118

Range of log-likelihood scores of the Null model:

Maximum Score: -32.77893094432278

Median Score: -38.530707678953775

Minimum Score: -38.53070767895383

Number of times (out of 30) the No_Wind model possesses a greater

goodness-of-fit: 30

Number of times (out of 30) the No_Wind model's outperformance of the Null model is statistically significant: 30

Median Log-Likelihood Ratio, No_Wind model vs. Null model: 58.86462675024768

p-value of the Median Log-Likelihood Ratio: 0.000853228493643272

Range of McFadden Adjusted Psuedo R-Squared statistics for the No_Wind model:

Minimum Pseudo R-Squared: -0.18024654722701539

Median Pseudo R-Squared: -0.07512674420207333

Maximum Pseudo R-Squared: 0.09453983797361898

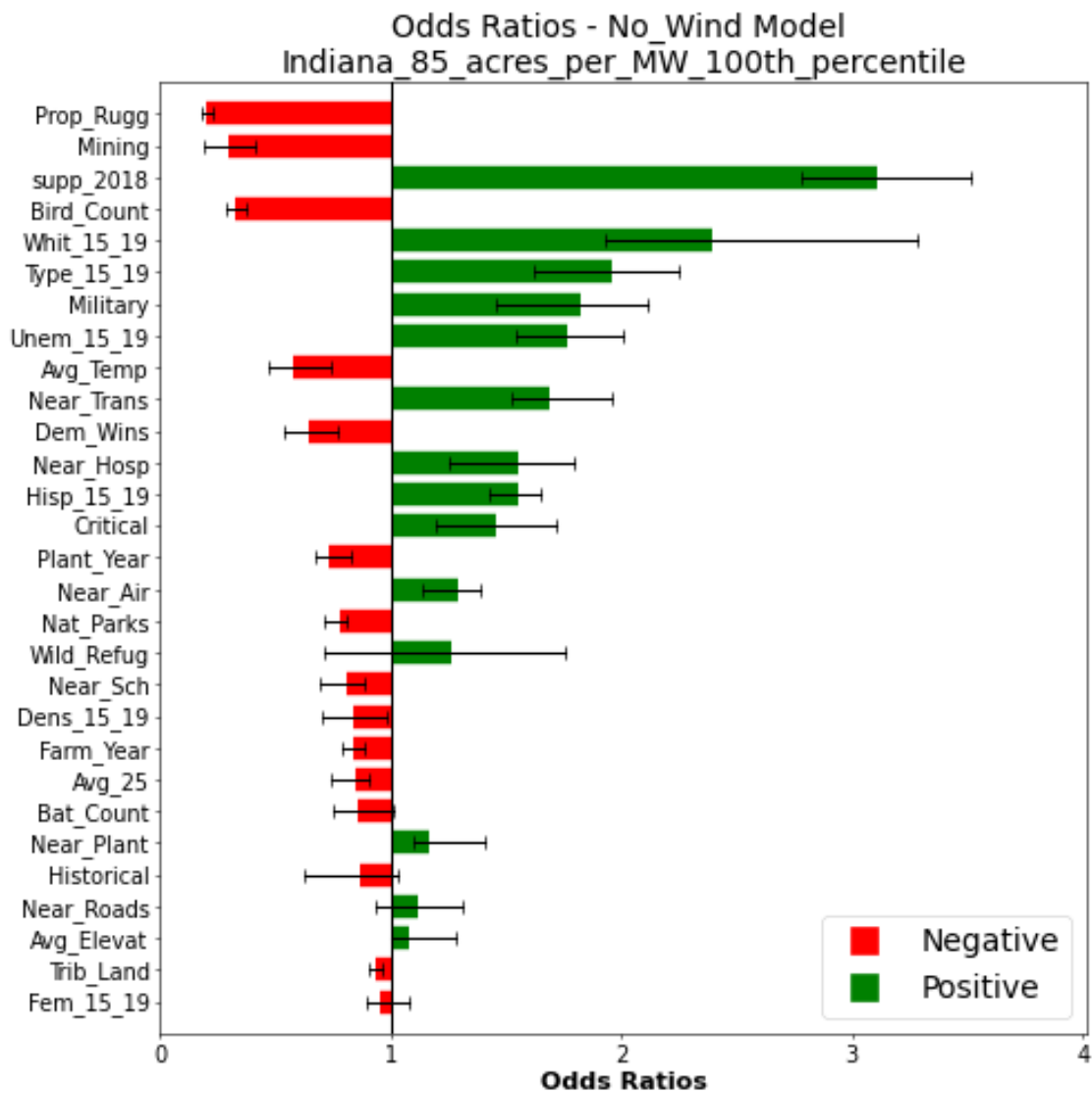
The following dataframe summarizes the coefficients and odds ratios

obtained from fitting the No_Wind model to the aggregated dataset. Predictors are

ranked by the magnitude of their coefficients to convey strength of association:

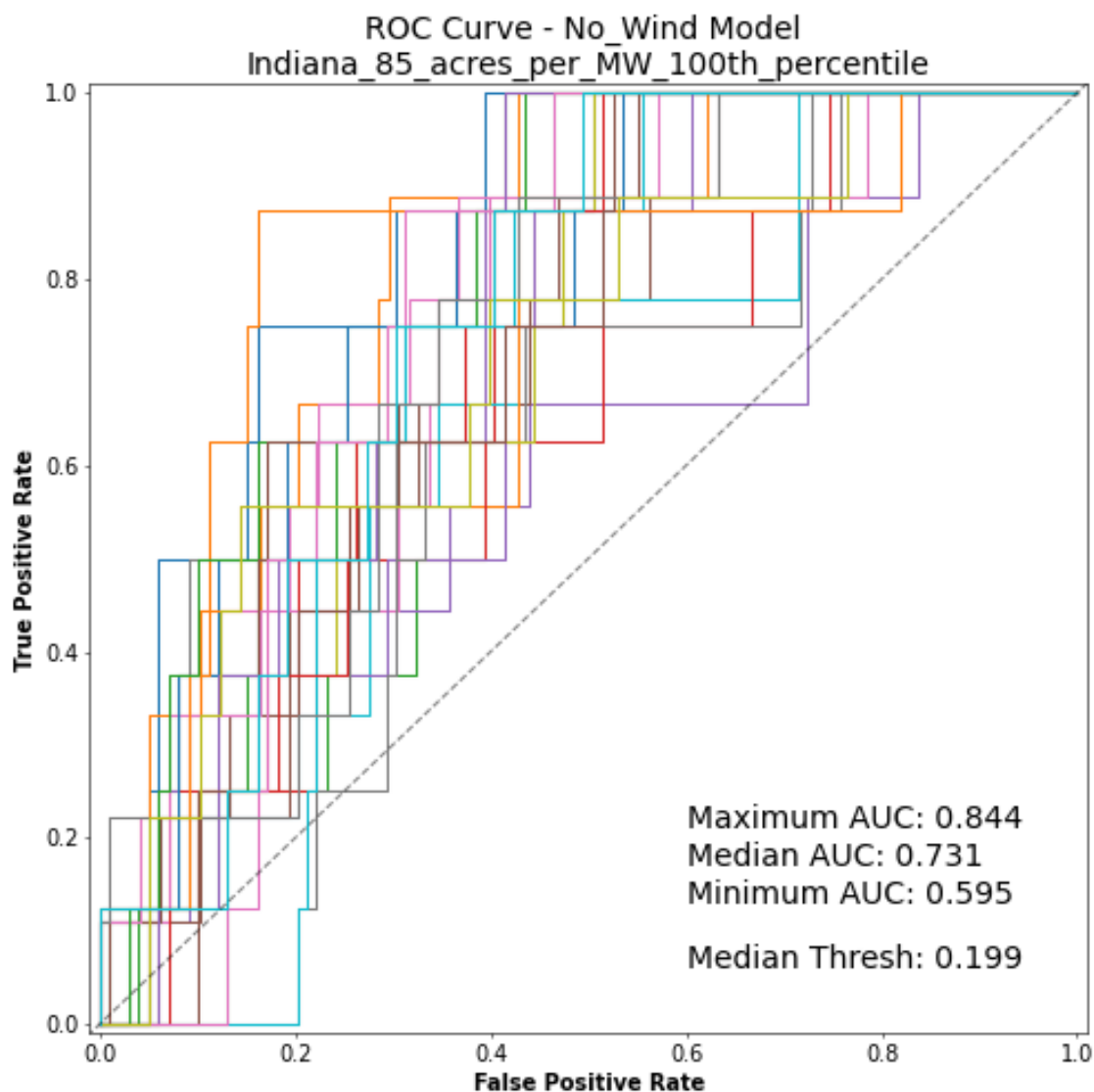
Predictor	Odds_Low	Odds_Med	Odds_Upp	Coef_Med	Rank
Prop_Rugg	0.184133	0.200837	0.224275	-1.605261	1
Mining	0.192483	0.294128	0.410549	-1.223741	2
supp_2018	2.782048	3.108851	3.518253	1.134253	3
Bird_Count	0.286963	0.326498	0.372611	-1.119332	4
Whit_15_19	1.932125	2.397382	3.281382	0.874377	5
Type_15_19	1.616429	1.962822	2.249723	0.674383	6
Military	1.458856	1.824005	2.115757	0.601035	7
Unem_15_19	1.542104	1.762016	2.007594	0.566459	8
Avg_Temp	0.470386	0.580700	0.737472	-0.543521	9
Near_Trans	1.525580	1.688864	1.958877	0.524056	10
Dem_Wins	0.535551	0.639286	0.771035	-0.447403	11
Near_Hosp	1.250936	1.556385	1.791913	0.442366	12
Hisp_15_19	1.431086	1.549127	1.650367	0.437691	13
Critical	1.189955	1.456794	1.718048	0.376238	14
Plant_Year	0.676807	0.726016	0.823754	-0.320183	15
Near_Air	1.134483	1.292320	1.389425	0.256439	16
Nat_Parks	0.707345	0.776584	0.809690	-0.252850	17
Wild_Refug	0.709796	1.263512	1.753437	0.233895	18
Near_Sch	0.690559	0.812774	0.881439	-0.207303	19
Dens_15_19	0.700238	0.835471	0.984779	-0.179760	20
Farm_Year	0.785268	0.838814	0.881736	-0.175766	21
Avg_25	0.736753	0.841999	0.905555	-0.171976	22
Bat_Count	0.752216	0.853482	1.015696	-0.158431	23
Near_Plant	1.096997	1.163464	1.409584	0.151401	24
Historical	0.627397	0.864131	1.028498	-0.146030	25
Near_Roads	0.936308	1.119120	1.312828	0.112543	26
Avg_Elevat	0.996716	1.077866	1.283522	0.074983	27
Trib_Land	0.904944	0.931670	0.958843	-0.070777	28
Fem_15_19	0.895388	0.955817	1.076581	-0.045189	29

Odds Ratio chart generated from the 30 No_Wind model runs with the training data:



----- MODEL Validation (Testing Data): No_Wind Configuration -----

ROC curves generated from the 30 No_Wind model runs with the testing data:



Range of Area Under Curve (AUC) statistics for the No_Wind model:

Minimum AUC: 0.595959595959596

Median AUC: 0.7316919191919191

Maximum AUC: 0.8446969696969697

Range of optimal threshold classifications for the No_Wind model:

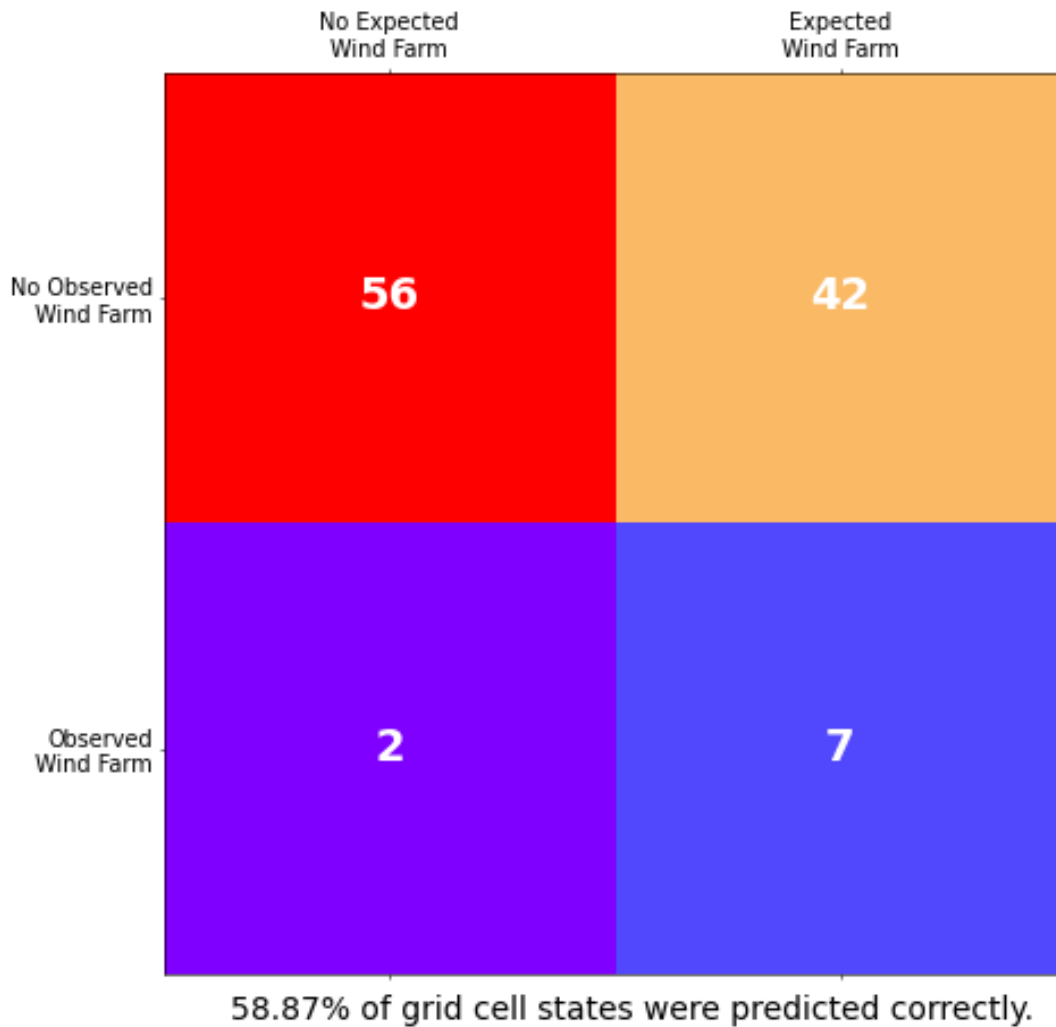
Minimum Threshold: 0.04047361463769025

Median Threshold: 0.19926365146575617

Maximum Threshold: 0.6899307250475301

Median Confusion Matrix of the No_Wind model's predictive accuracy:

Confusion Matrix - No_Wind Model Indiana_85_acres_per_MW_100th_percentile



Below are the range of confusion matrix results from the 30 No_Wind model runs with the testing data:

Lower Quartile confusion matrix:

[[50 49]
[1 7]]

Lower Quartile proportion of correctly predicted grid cell states by the No_Wind model:
0.5327102803738317

Median confusion matrix:

[[56 42]
[2 7]]

Median proportion of correctly predicted grid cell states by the No_Wind model:
0.5887850467289719

Upper Quartile confusion matrix:

[[62 36]
[2 7]]

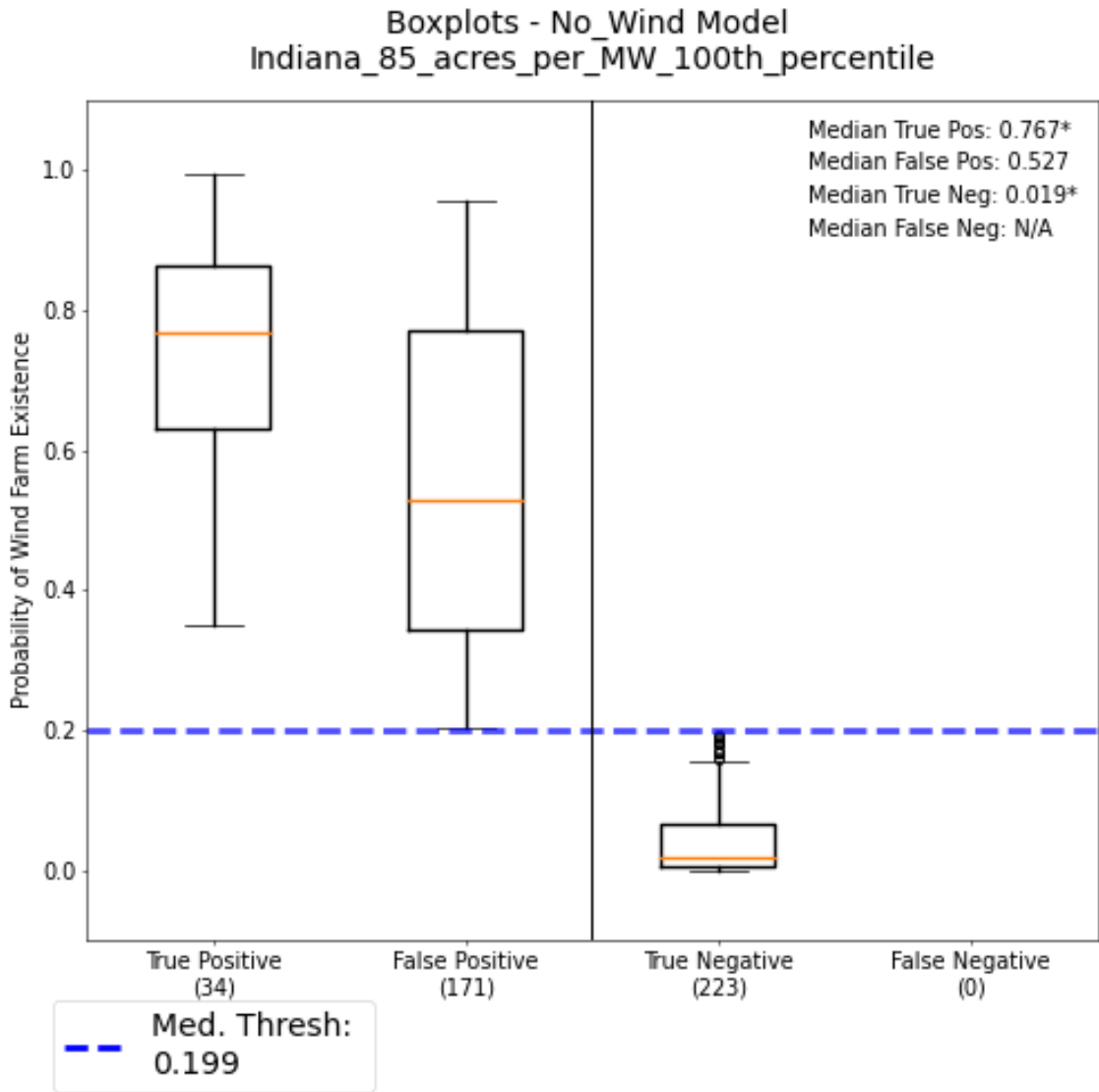
Upper Quartile proportion of correctly predicted grid cell states by the No_Wind model:
0.6448598130841121

----- BOXPLOT CONSTRUCTION (All Data): No_Wind Configuration -----

Grid cell classifications from executing the trained and tested No_Wind model over all grid cells in Indiana:

Number of True Positive Grid Cells: 34
Number of False Positive Grid Cells: 171
Number of True Negative Grid Cells: 223
Number of False Negative Grid Cells: 0

Boxplot of grid cell probabilities in each classification:



Median probabilities of wind farm existence for each grid cell classification.
An asterisk indicates a Mann-Whitney U-test result that is statistically significant ($p < 0.05$):

Median False Pos: 0.527
Median True Pos: 0.767*
Median False Neg: N/A
Median True Neg: 0.019*

Mann-Whitney U-test results:

Mann-Whitney Statistic - True Positive vs False Positive:

U-statistic = 1677.0

p-value = 9.951237717639365e-05

Mann-Whitney Statistic - True Negative vs False Negative:

U-statistic = 0.0

p-value = 0.0

----- MAP CONSTRUCTION: No_Wind Configuration -----

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Model_Testing\WiFSS_Surfaces\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana_No_Wind.gdb\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana_No_Wind_Map

Total (Percentage) of all grid cells over Indiana that exist in hotspots:

14 (3.27%)

Total (Percentage) True Positive grid cells over Indiana that exist in hotspots:

3 (8.82%)

Total (Percentage) False Positive grid cells over Indiana that exist in hotspots:

11 (6.43%)

Wind_Only Configuration Output Begins

----- MODEL CALIBRATION (Training Data): Wind_Only Configuration -----

Range of log-likelihood scores from 30 training runs of the Wind_Only model:

Maximum Score: -21.864600924778813

Median Score: -26.47186680319615

Minimum Score: -30.083934281517656

Range of log-likelihood scores of the Null model:

Maximum Score: -32.77893094432278

Median Score: -35.65481931163828

Minimum Score: -38.53070767895383

Number of times (out of 30) the Wind_Only model possesses a greater

goodness-of-fit: 30

Number of times (out of 30) the Wind_Only model's outperformance of the Null model is statistically significant: 30

Median Log-Likelihood Ratio, Wind_Only model vs. Null model: 18.365905016884255

p-value of the Median Log-Likelihood Ratio: 1.8229087488656115e-05

Range of McFadden Adjusted Psuedo R-Squared statistics for the Wind_Only model:

Minimum Pseudo R-Squared: 0.11251445059706733

Median Pseudo R-Squared: 0.21322320400377492

Maximum Pseudo R-Squared: 0.27195304308995183

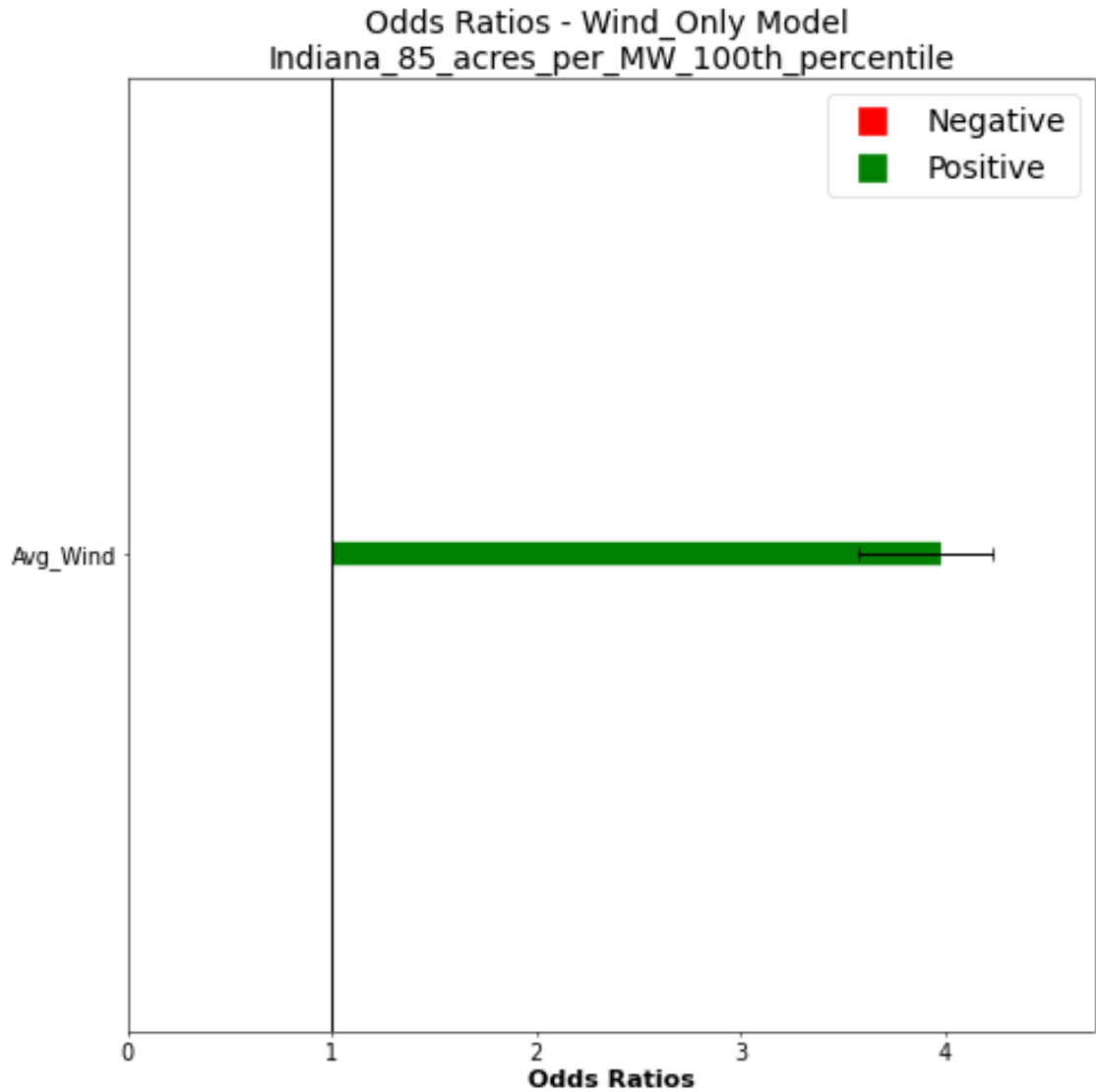
The following dataframe summarizes the coefficients and odds ratios

obtained from fitting the Wind_Only model to the aggregated dataset. Predictors are

ranked by the magnitude of their coefficients to convey strength of association:

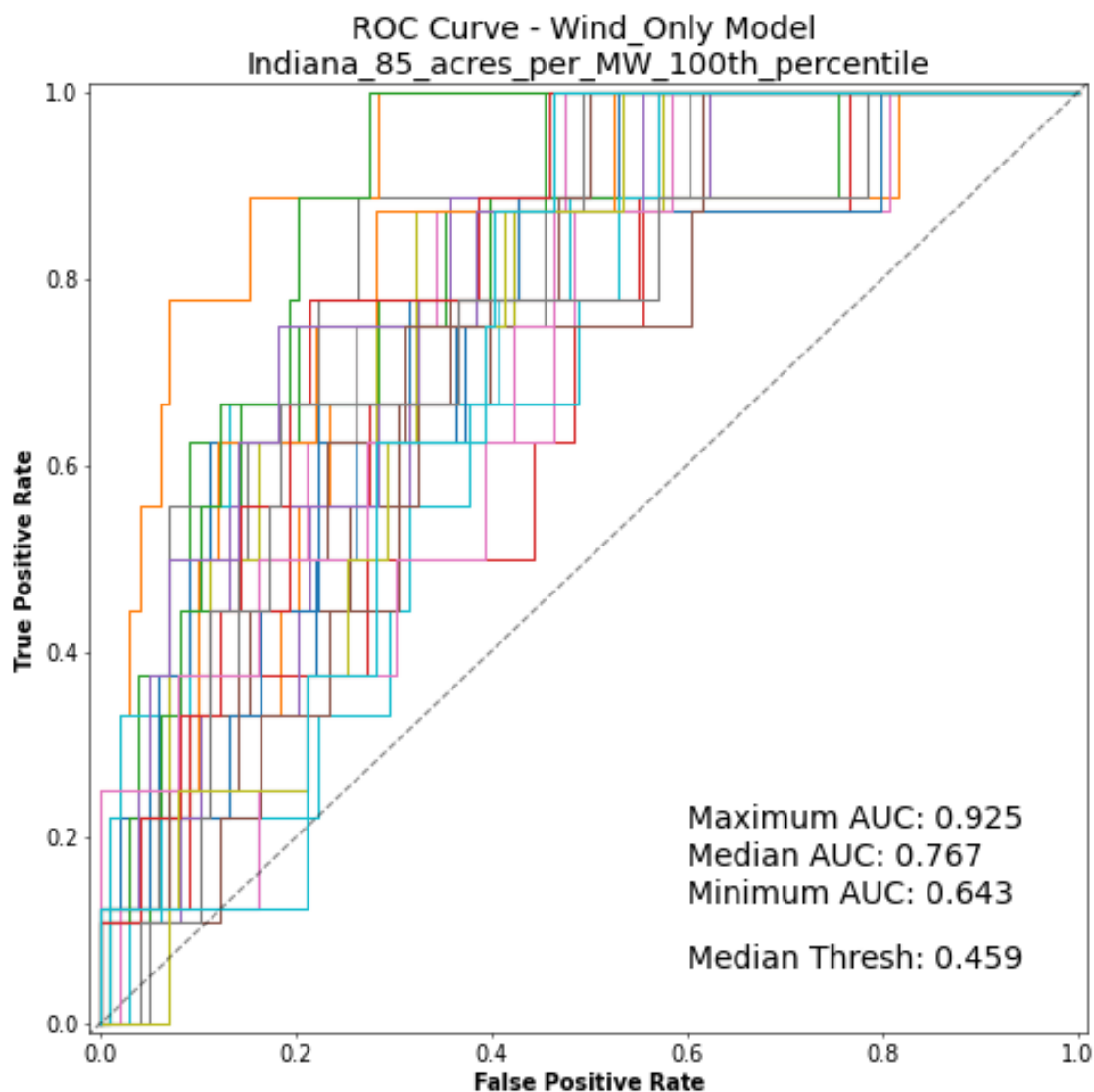
Predictor	Odds_Low	Odds_Med	Odds_Upp	Coef_Med	Rank
Avg_Wind	3.576378	3.978295	4.23095	1.380853	1

Odds Ratio chart generated from the 30 Wind_Only model runs with the training data:



----- MODEL Validation (Testing Data): Wind_Only Configuration -----

ROC curves generated from the 30 Wind_Only model runs with the testing data:



Range of Area Under Curve (AUC) statistics for the Wind_Only model:

Minimum AUC: 0.643939393939394

Median AUC: 0.7670454545454546

Maximum AUC: 0.9251700680272108

Range of optimal threshold classifications for the Wind_Only model:

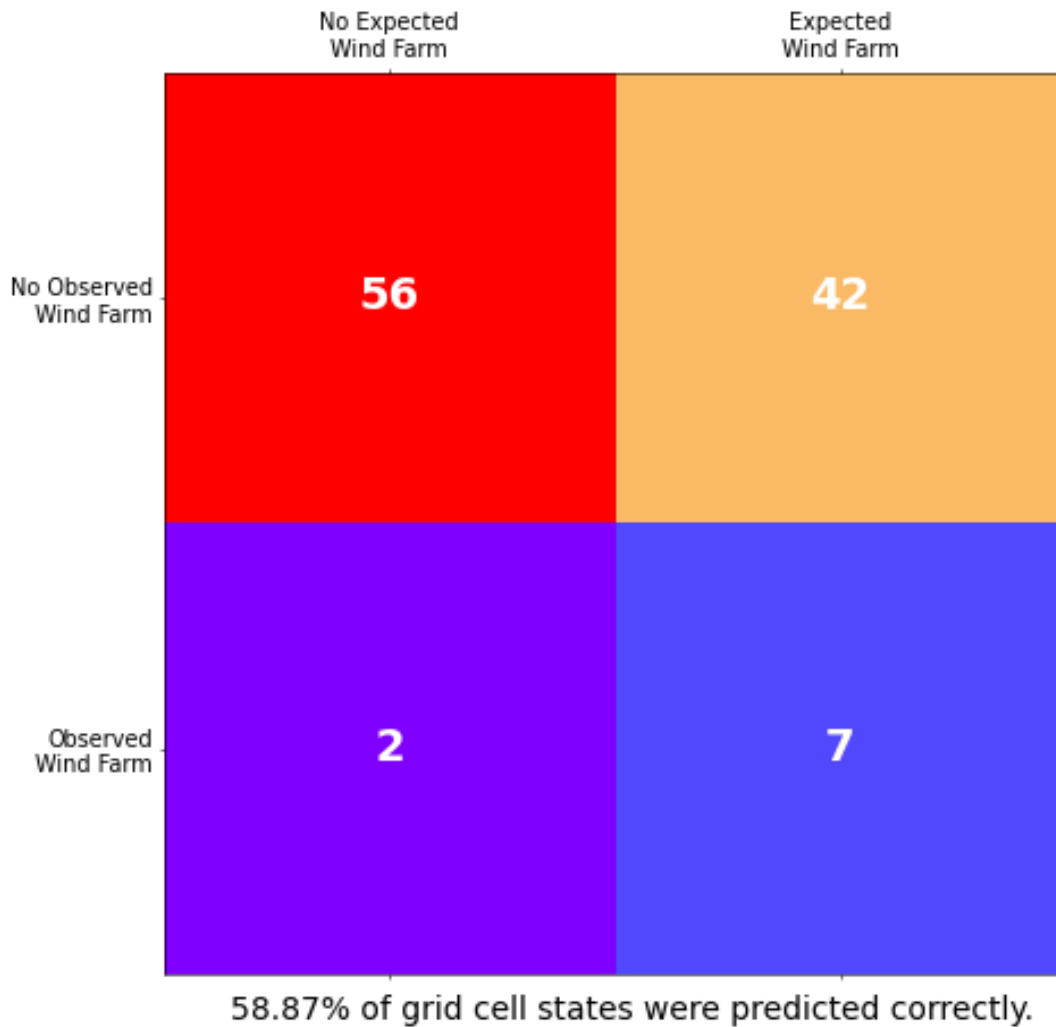
Minimum Threshold: 0.27484629518449755

Median Threshold: 0.45909068027215766

Maximum Threshold: 0.7226069418712211

Median Confusion Matrix of the Wind_Only model's predictive accuracy:

Confusion Matrix - Wind_Only Model Indiana_85_acres_per_MW_100th_percentile



Below are the range of confusion matrix results from the 30 Wind_Only model runs with the testing data:

Lower Quartile confusion matrix:

```
[[49 49]
 [ 1  8]]
```

Lower Quartile proportion of correctly predicted grid cell states by the Wind_Only model:
0.5327102803738317

Median confusion matrix:

```
[[56 42]
 [ 2  7]]
```

Median proportion of correctly predicted grid cell states by the Wind_Only model:
0.5887850467289719

Upper Quartile confusion matrix:

```
[[77 21]
 [ 3  6]]
```

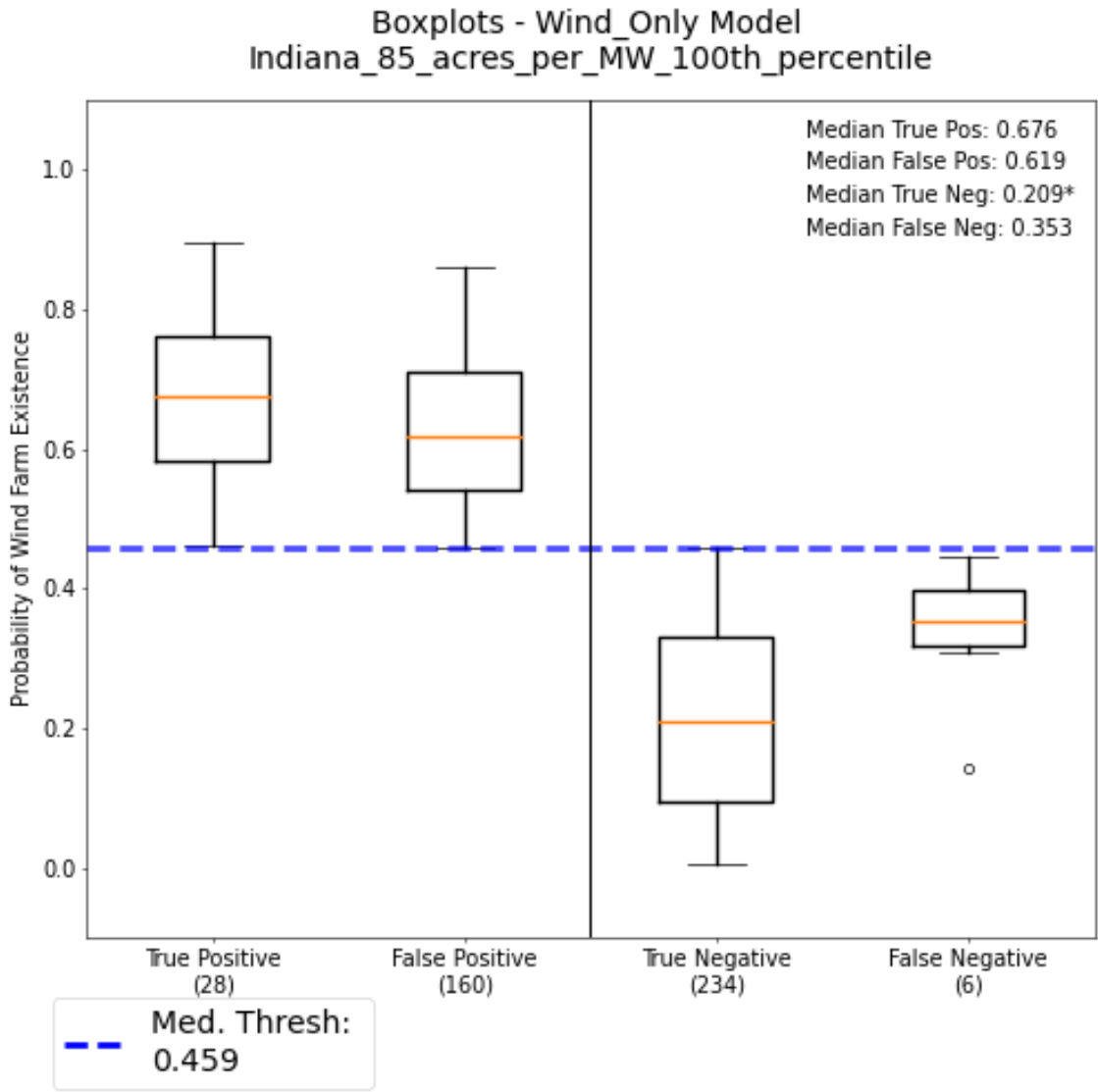
Upper Quartile proportion of correctly predicted grid cell states by the Wind_Only model:
0.7757009345794392

----- BOXPLOT CONSTRUCTION (All Data): Wind_Only Configuration -----

Grid cell classifications from executing the trained and tested Wind_Only model over all grid cells in Indiana:

Number of True Positive Grid Cells: 28
Number of False Positive Grid Cells: 160
Number of True Negative Grid Cells: 234
Number of False Negative Grid Cells: 6

Boxplot of grid cell probabilities in each classification:



Median probabilities of wind farm existence for each grid cell classification.
An asterisk indicates a Mann-Whitney U-test result that is statistically significant ($p < 0.05$):

Median False Pos: 0.619
Median True Pos: 0.676
Median False Neg: 0.353
Median True Neg: 0.209*

Mann-Whitney U-test results:

Mann-Whitney Statistic - True Positive vs False Positive:

U-statistic = 1771.0

p-value = 0.07777867420663816

Mann-Whitney Statistic - True Negative vs False Negative:

U-statistic = 1052.0

p-value = 0.037401614049936104

----- MAP CONSTRUCTION: Wind_Only Configuration -----

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Model_Testing\WiFSS_Surfaces\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana_Wind_Only.gdb\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana_Wind_Only_Map

Total (Percentage) of all grid cells over Indiana that exist in hotspots:

1 (0.23%)

Total (Percentage) True Positive grid cells over Indiana that exist in hotspots:

1 (3.57%)

Total (Percentage) False Positive grid cells over Indiana that exist in hotspots:

0 (0.0%)

Reduced Configuration Output Begins

----- MODEL CALIBRATION (Training Data): Reduced Configuration -----

Dataframe showing the lowered goodness-of-fit caused by removing each predictor with replacement over 30 model runs. The columns show the number of times removal of each predictor reduced the model's goodness-of-fit, and the number of times this reduction exceeded a $p < 0.5$ stopping criterion:

Predictors	Reduced_Fit	Stop_Criterion
Bird_Count	27	0
Near_Trans	27	0
Bat_Count	25	0
Near_Air	25	0
Dens_15_19	24	0
Near_Sch	24	0
supp_2018	24	0
Mining	24	0
Avg_Temp	23	0
Wild_Refug	23	0
Prop_Rugg	22	0
Near_Roads	22	0
Farm_Year	22	0
Near_Plant	22	0
Dem_Wins	22	0
Avg_Elevat	21	0
Avg_Wind	21	0
Fem_15_19	21	0

Avg_25	21	0
Whit_15_19	21	0
Trib_Land	21	0
Type_15_19	20	0
Historical	20	0
Military	20	0
Nat_Parks	20	0
Near_Hosp	19	0
Plant_Year	19	0
Unem_15_19	18	0
Hisp_15_19	18	0
Critical	15	0

Dataframe of model performance for each set of predictors, showing the Number of predictors in each combination, the median number of accurately predicted grid cell states, and the ratio of true-to-false positive predictions:

index	Num_Pred	Accuracy	True_False
0	2	0.728972	0.222222
1	5	0.719626	0.250000
2	3	0.710280	0.241379
3	4	0.705607	0.233333
4	7	0.705607	0.233333
5	11	0.696262	0.218750
6	9	0.686916	0.218750
7	28	0.682243	0.187500
8	17	0.677570	0.205882
9	23	0.672897	0.212121
10	25	0.672897	0.212121
11	6	0.672897	0.205882
12	12	0.672897	0.205882
13	14	0.672897	0.205882
14	22	0.672897	0.205882
15	13	0.663551	0.205882
16	19	0.658879	0.205882
17	8	0.658879	0.200000
18	21	0.658879	0.200000
19	10	0.658879	0.194444
20	20	0.654206	0.194444
21	26	0.644860	0.194444
22	27	0.644860	0.194444
23	18	0.640187	0.184211
24	24	0.635514	0.189189
25	29	0.626168	0.184211
26	16	0.621495	0.179487
27	1	0.621495	0.175000
28	30	0.621495	0.175000
29	15	0.607477	0.170732

Set of predictors (2 total) to be used in the Reduced Model:

['Bird_Count', 'Near_Trans']

Range of log-likelihood scores from 30 training runs of the Reduced model:

Maximum Score: -27.07230310822922

Median Score: -37.104689350660635

Minimum Score: -43.10294723634837

Range of log-likelihood scores of the Null model:

Maximum Score: -32.77893094432278

Median Score: -38.530707678953775

Minimum Score: -38.53070767895383

Number of times (out of 30) the Reduced model possesses a greater goodness-of-fit: 12

Number of times (out of 30) the Reduced model's outperformance of the Null model is statistically significant: 2

Median Log-Likelihood Ratio, Reduced model vs. Null model: 2.8520366565862787

p-value of the Median Log-Likelihood Ratio: 0.09125826236663705

Range of McFadden Adjusted Psuedo R-Squared statistics for the Reduced model:

Minimum Pseudo R-Squared: -0.17057147281477225

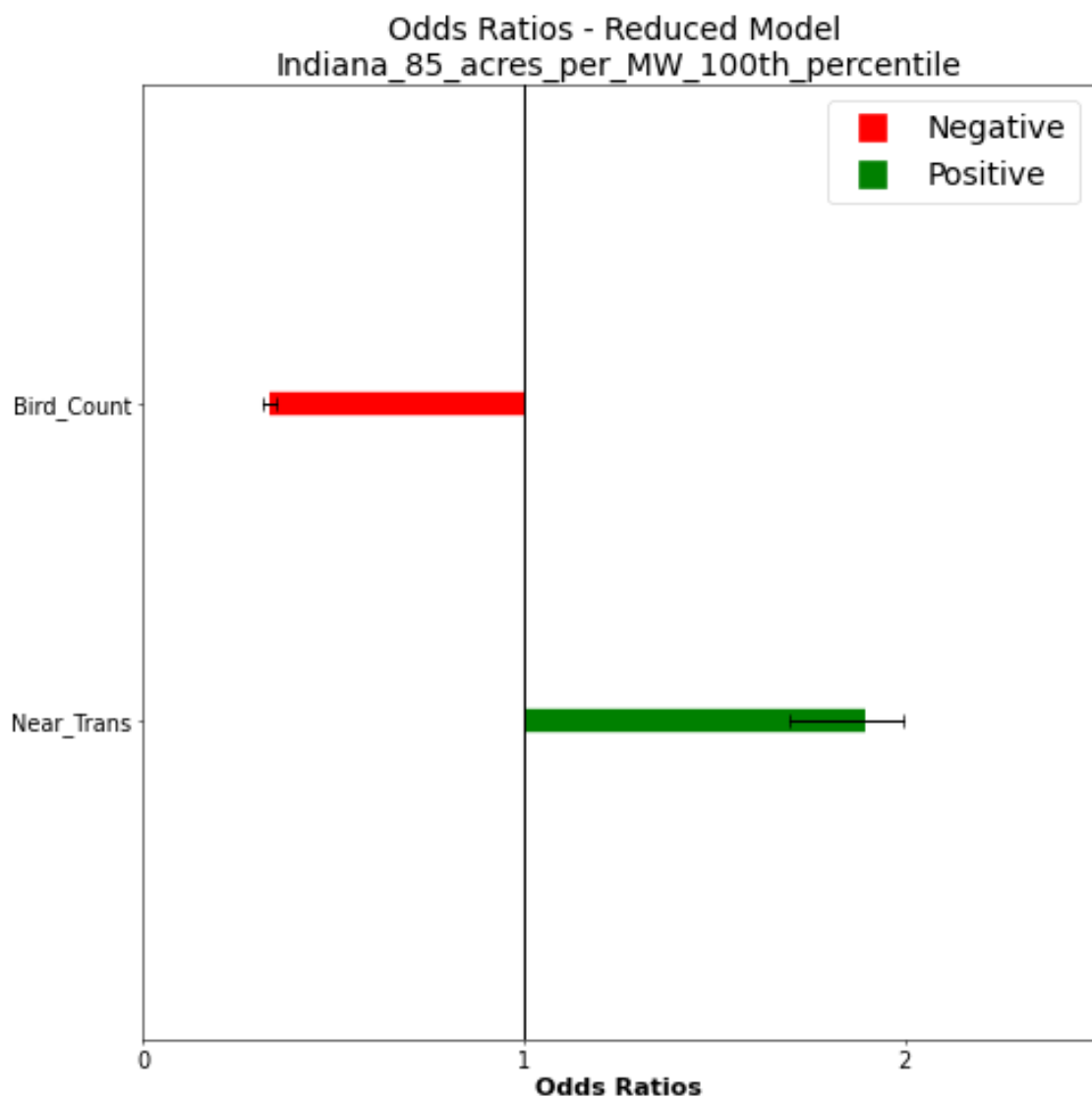
Median Pseudo R-Squared: -0.059736790185587085

Maximum Pseudo R-Squared: 0.1130795828085277

The following dataframe summarizes the coefficients and odds ratios obtained from fitting the Reduced model to the aggregated dataset. Predictors are ranked by the magnitude of their coefficients to convey strength of association:

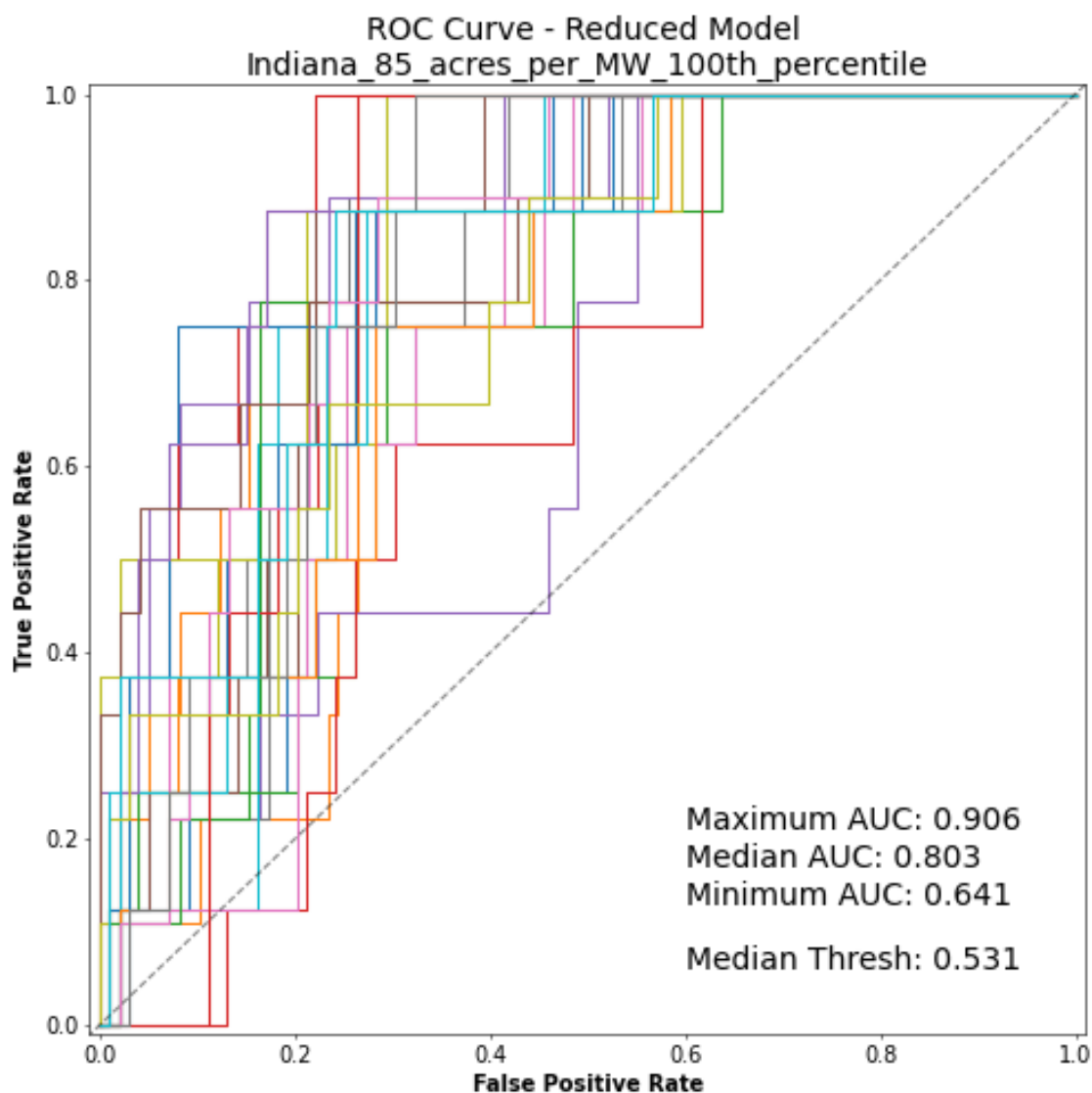
Predictor	Odds_Low	Odds_Med	Odds_Upp	Coef_Med	Rank
Bird_Count	0.314853	0.332189	0.34763	-1.102052	1
Near_Trans	1.696994	1.898476	1.99662	0.641051	2

Odds Ratio chart generated from the 30 Reduced model runs with the training data:



----- MODEL Validation (Testing Data): Reduced Configuration -----

ROC curves generated from the 30 Reduced model runs with the testing data:



Range of Area Under Curve (AUC) statistics for the Reduced model:

Minimum AUC: 0.6414141414141414

Median AUC: 0.803030303030303

Maximum AUC: 0.9065656565656566

Range of optimal threshold classifications for the Reduced model:

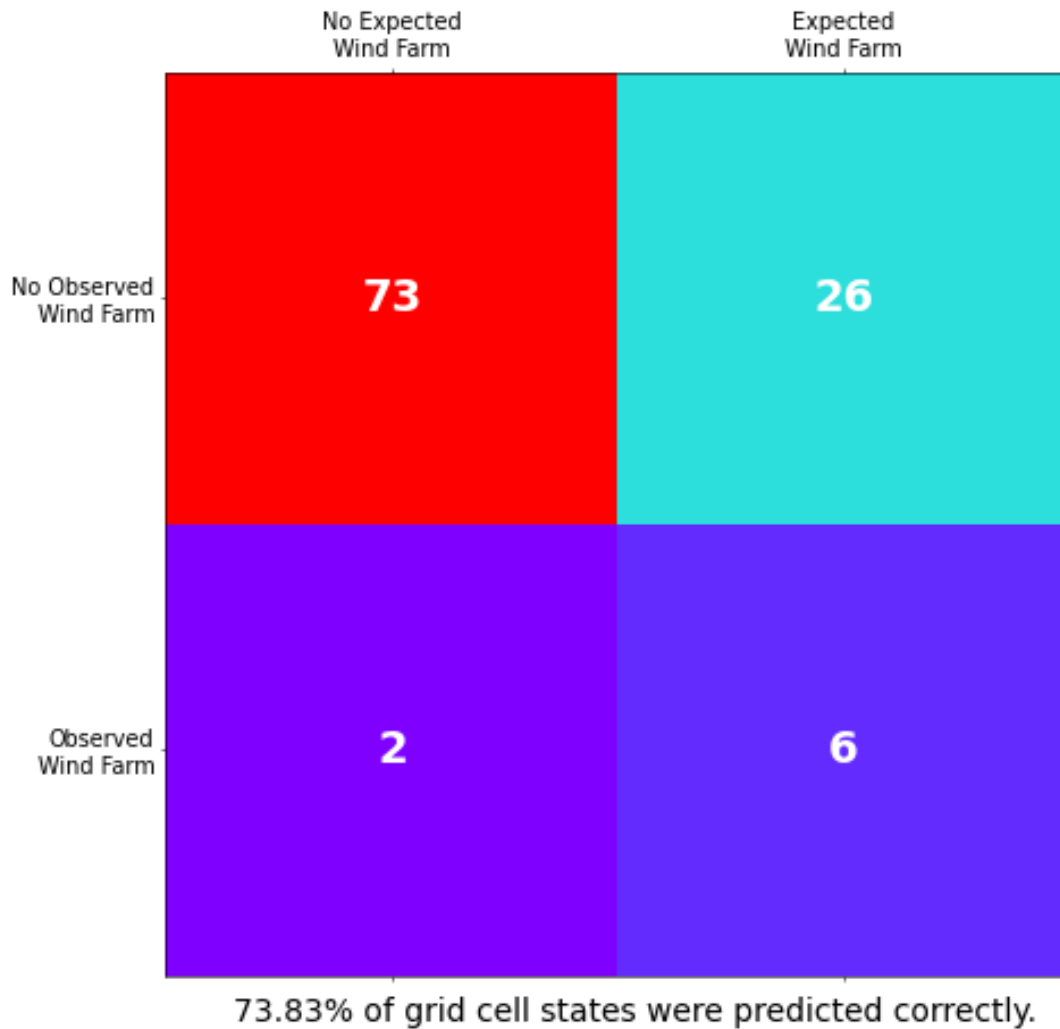
Minimum Threshold: 0.13385218617491382

Median Threshold: 0.5319625951516664

Maximum Threshold: 0.7569605823840615

Median Confusion Matrix of the Reduced model's predictive accuracy:

Confusion Matrix - Reduced Model Indiana_85_acres_per_MW_100th_percentile



Below are the range of confusion matrix results from the 30 Reduced model runs with the testing data:

Lower Quartile confusion matrix:

```
[[70 29]
 [ 2  6]]
```

Lower Quartile proportion of correctly predicted grid cell states by the Reduced model:
0.7102803738317757

Median confusion matrix:

```
[[73 26]
 [ 2  6]]
```

Median proportion of correctly predicted grid cell states by the Reduced model:
0.7383177570093458

Upper Quartile confusion matrix:

```
[[77 22]
 [ 2  6]]
```

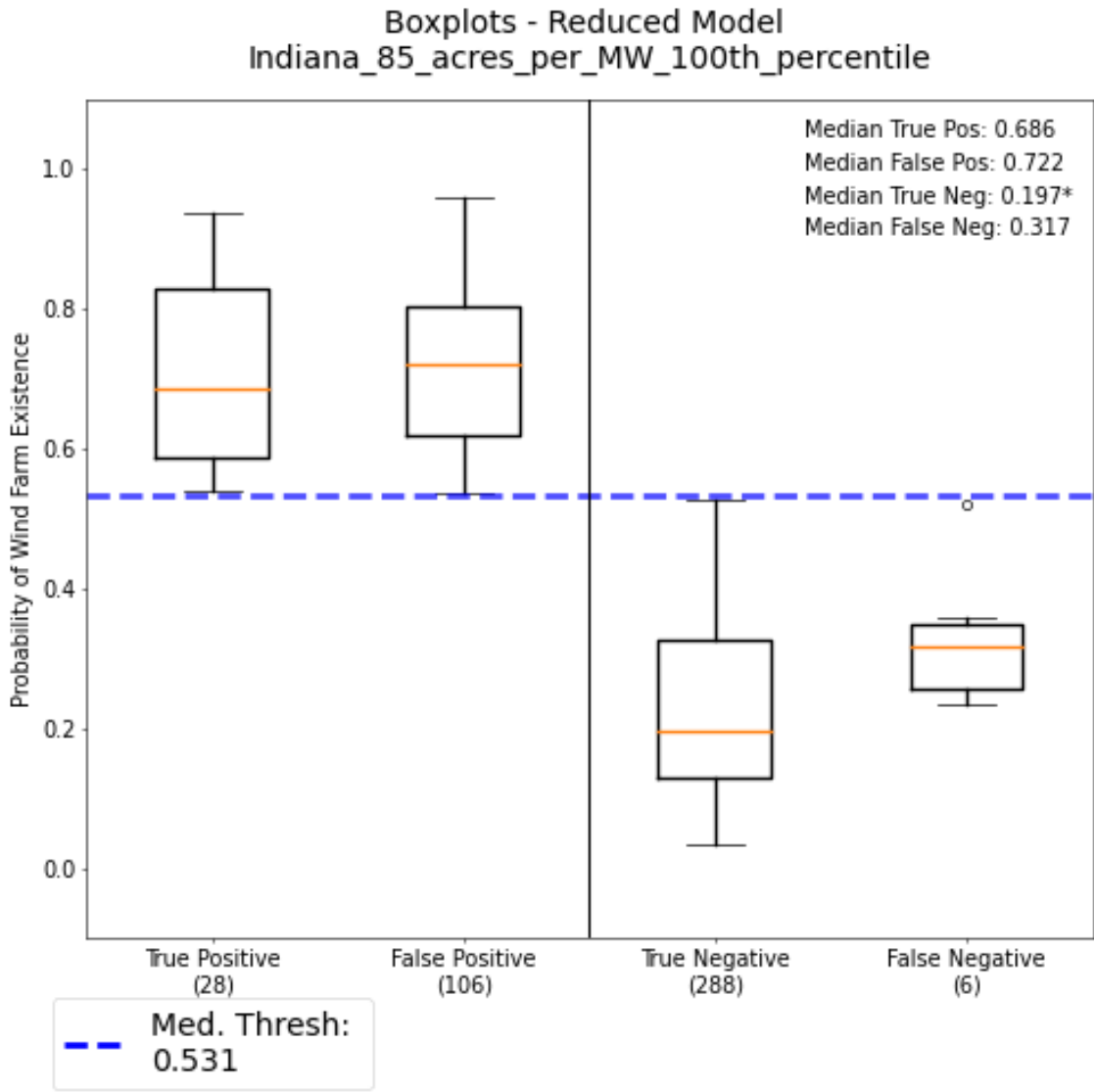
Upper Quartile proportion of correctly predicted grid cell states by the Reduced model:
0.7757009345794392

----- BOXPLOT CONSTRUCTION (All Data): Reduced Configuration -----

Grid cell classifications from executing the trained and tested Reduced model over all grid cells in Indiana:

Number of True Positive Grid Cells: 28
Number of False Positive Grid Cells: 106
Number of True Negative Grid Cells: 288
Number of False Negative Grid Cells: 6

Boxplot of grid cell probabilities in each classification:



Median probabilities of wind farm existence for each grid cell classification.
An asterisk indicates a Mann-Whitney U-test result that is statistically significant ($p < 0.05$):

Median False Pos: 0.722
Median True Pos: 0.686
Median False Neg: 0.317
Median True Neg: 0.197*

Mann-Whitney U-test results:

Mann-Whitney Statistic - True Positive vs False Positive:

U-statistic = 1569.0

p-value = 0.6437710825449865

Mann-Whitney Statistic - True Negative vs False Negative:

U-statistic = 1291.0

p-value = 0.038516458496435216

----- MAP CONSTRUCTION: Reduced Configuration -----

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Model_Testing\WiFSS_Surfaces/Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana_Reduced.gdb\Hexagon_Grid_85_acres_per_MW_100th_percentile_Indiana_Reduced_Map

Total (Percentage) of all grid cells over Indiana that exist in hotspots:

13 (3.04%)

Total (Percentage) True Positive grid cells over Indiana that exist in hotspots:

4 (14.29%)

Total (Percentage) False Positive grid cells over Indiana that exist in hotspots:

9 (8.49%)