---------- 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: North_Dakota
Specified wind farm density: 65 acres/MW
Specified wind power capacity: 80th percentile (202 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
Avg_25	0.001877
Unem_15_19	0.004832
Fem_15_19	0.033992
Whit_15_19	0.059775
Near_Hosp	0.069884
Prop_Rugg	0.108065
Near_Trans	0.118142
Type_15_19	0.124240
Near_Roads	0.127029
Hisp_15_19	0.133392
Near_Plant	0.151022
Avg_Wind	0.293369
Avg_Temp	0.321353
Dens_15_19	0.423367
Near_Air	0.581573
Avg_Elevat	0.679404
Near_Sch	0.795061
Undev_Land	0.864468

Predictors to be removed based on a non-linear relationship with the logit of likelihood of wind farm occurrence: None

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

Grouped Multicollinearity Test Results:

Predictor	VIF
Military	1.034294
Nat_Parks	1.110143
Historical	1.158101
Bat_Count	1.331720
Near_Trans	1.342415
Farm_Year	1.347277
Near_Hosp	1.387304
Undev_Land	1.408103
Near_Roads	1.429112
Critical	1.487604
Trib_Land	1.635895
Near_Sch	1.640873
Near_Air	1.663229
Plant_Year	1.675105
Fem_15_19	1.714971
Prop_Rugg	1.814857
Avg_Wind	1.835969
Mining	2.122677
Unem_15_19	2.208655
Dens_15_19	2.343685
Hisp_15_19	2.375574
Wild_Refug	2.594833
Near_Plant	2.716043
Type_15_19	2.978334
Avg_Temp	4.102340
Bird_Count	4.676620
supp_2018	5.701869
Avg_25	5.890628
Avg_Elevat	7.131514
Dem_Wins	8.396979
Whit_15_19	12.118046

Pairwise Multicollinearity Test Results:

Predictor1	Predictor2	VIF
Near_Sch	Wild_Refug	1.000000
Near_Roads	Trib_Land	1.000001
Avg_Wind	Near_Roads	1.000002
Unem_15_19	Historical	1.000002
Farm_Year	Mining	1.000004
...	...	...
Dens_15_19	Type_15_19	2.089288
Avg_Temp	Bird_Count	2.237688
Dem_Wins	supp_2018	2.265315
Avg_25	Whit_15_19	2.296532
Dem_Wins	Whit_15_19	3.713336

Predictors to be removed from the model based on multicollinearity:
['Whit_15_19']

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', 'Whit_15_19']

################## 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: 302.0187499099202
Median Score: 292.3315457221197
Minimum Score: 279.34424729691455

Range of log-likelihood scores of the Null model:
Maximum Score: 192.06115078192624
Median Score: 192.06115078192624
Minimum Score: 192.0611507819258

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: 200.5407898803869
p-value of the Median Log-Likelihood Ratio: 3.923253855863396e-27

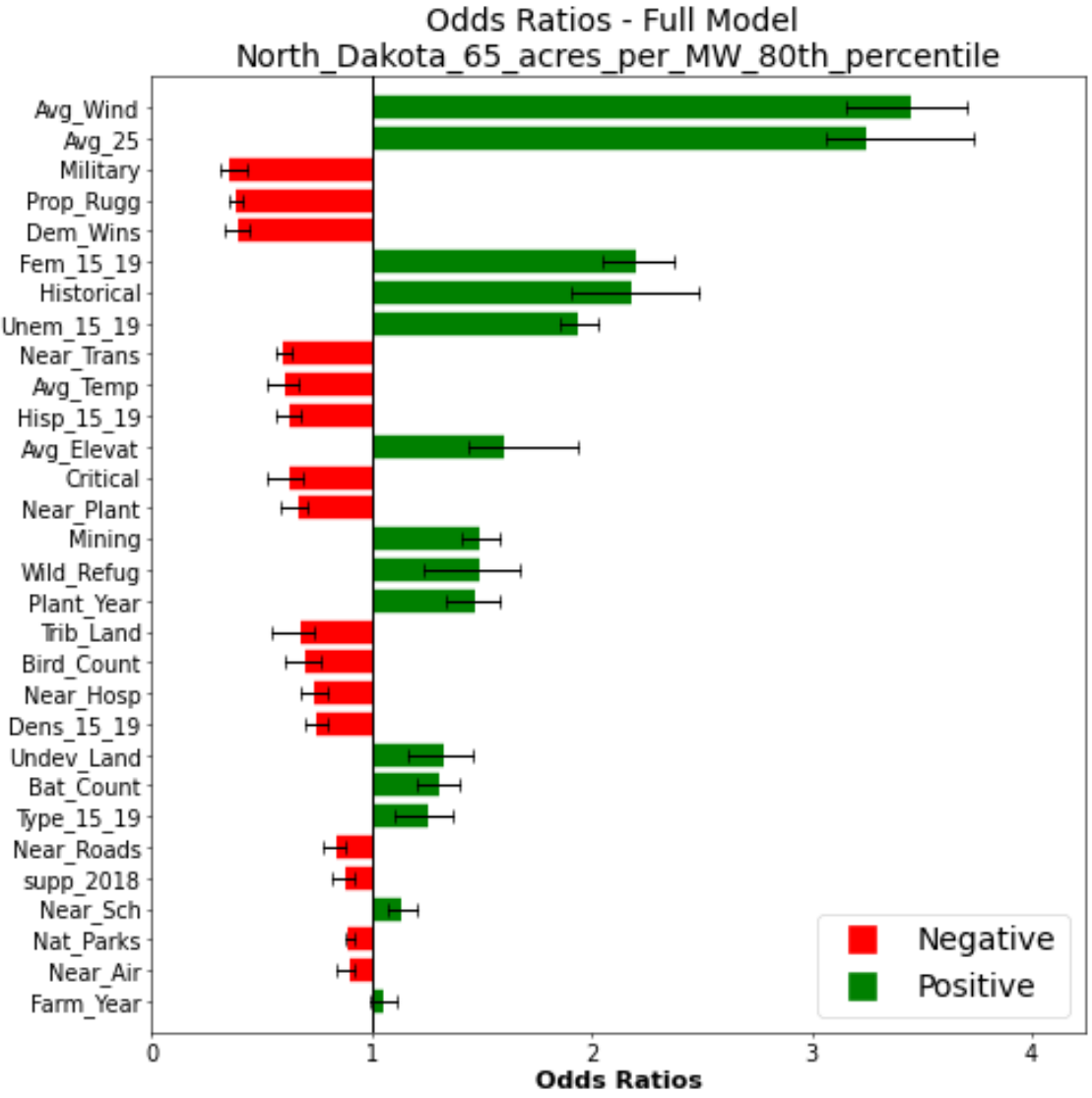
Range of McFadden Adjusted Psuedo R-Squared statistics for the Full model:
Minimum Pseudo R-Squared: -0.41110656062685735
Median Pseudo R-Squared: -0.36066843637131896
Maximum Pseudo R-Squared: -0.2930477938190339

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
Avg_Wind	3.156202	3.453752	3.704190	1.239461	1
Avg_25	3.065122	3.252726	3.739624	1.179493	2
Military	0.308079	0.354573	0.432531	-1.036840	3
Prop_Rugg	0.351293	0.380552	0.417570	-0.966132	4
Dem_Wins	0.328983	0.391059	0.441491	-0.938898	5
Fem_15_19	2.052538	2.203267	2.372954	0.789941	6
Historical	1.905713	2.179593	2.482006	0.779138	7
Unem_15_19	1.857671	1.939081	2.029338	0.662214	8
Near_Trans	0.565370	0.600792	0.636259	-0.509507	9
Avg_Temp	0.525480	0.601239	0.666277	-0.508763	10
Hisp_15_19	0.567423	0.621758	0.676950	-0.475204	11
Avg_Elevat	1.443469	1.606744	1.939496	0.474210	12
Critical	0.529578	0.623393	0.682443	-0.472579	13

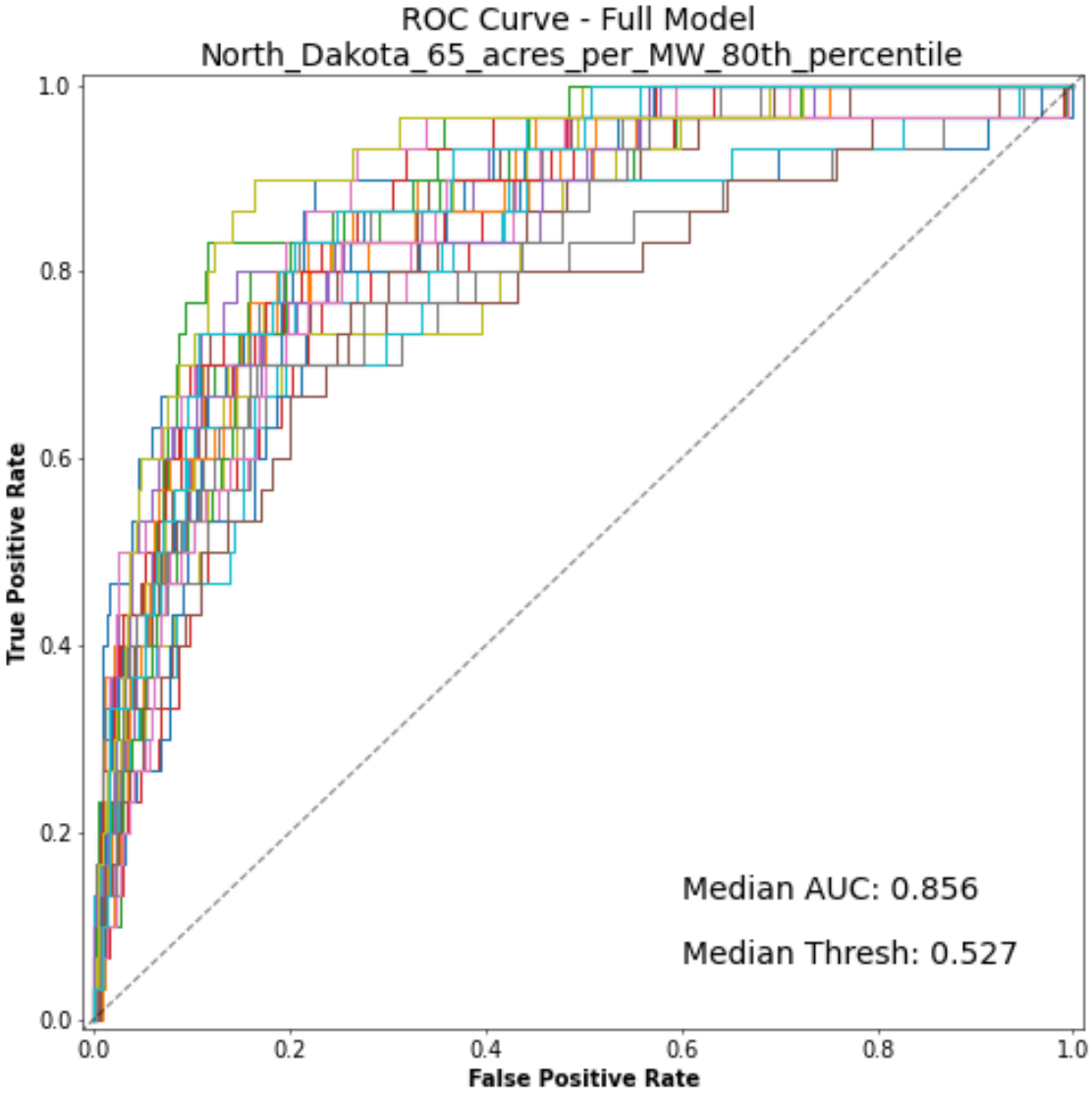
Near_Plant	0.590149	0.668124	0.709614	-0.403282	14
Mining	1.410468	1.490720	1.584522	0.399259	15
Wild_Refug	1.240616	1.487540	1.671038	0.397124	16
Plant_Year	1.337017	1.472850	1.583418	0.387199	17
Trib_Land	0.540087	0.680276	0.736901	-0.385257	18
Bird_Count	0.602955	0.700727	0.769958	-0.355637	19
Near_Hosp	0.678492	0.737678	0.798509	-0.304248	20
Dens_15_19	0.693030	0.748509	0.801144	-0.289673	21
Undev_Land	1.165081	1.325806	1.459472	0.282020	22
Bat_Count	1.204215	1.311533	1.397869	0.271197	23
Type_15_19	1.107250	1.257239	1.368659	0.228918	24
Near_Roads	0.780826	0.840242	0.877964	-0.174066	25
supp_2018	0.817587	0.881983	0.922779	-0.125583	26
Near_Sch	1.073182	1.131168	1.205433	0.123251	27
Nat_Parks	0.883055	0.893085	0.924381	-0.113074	28
Near_Air	0.841282	0.898083	0.920183	-0.107493	29
Farm_Year	0.987988	1.049037	1.110197	0.047873	30

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



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

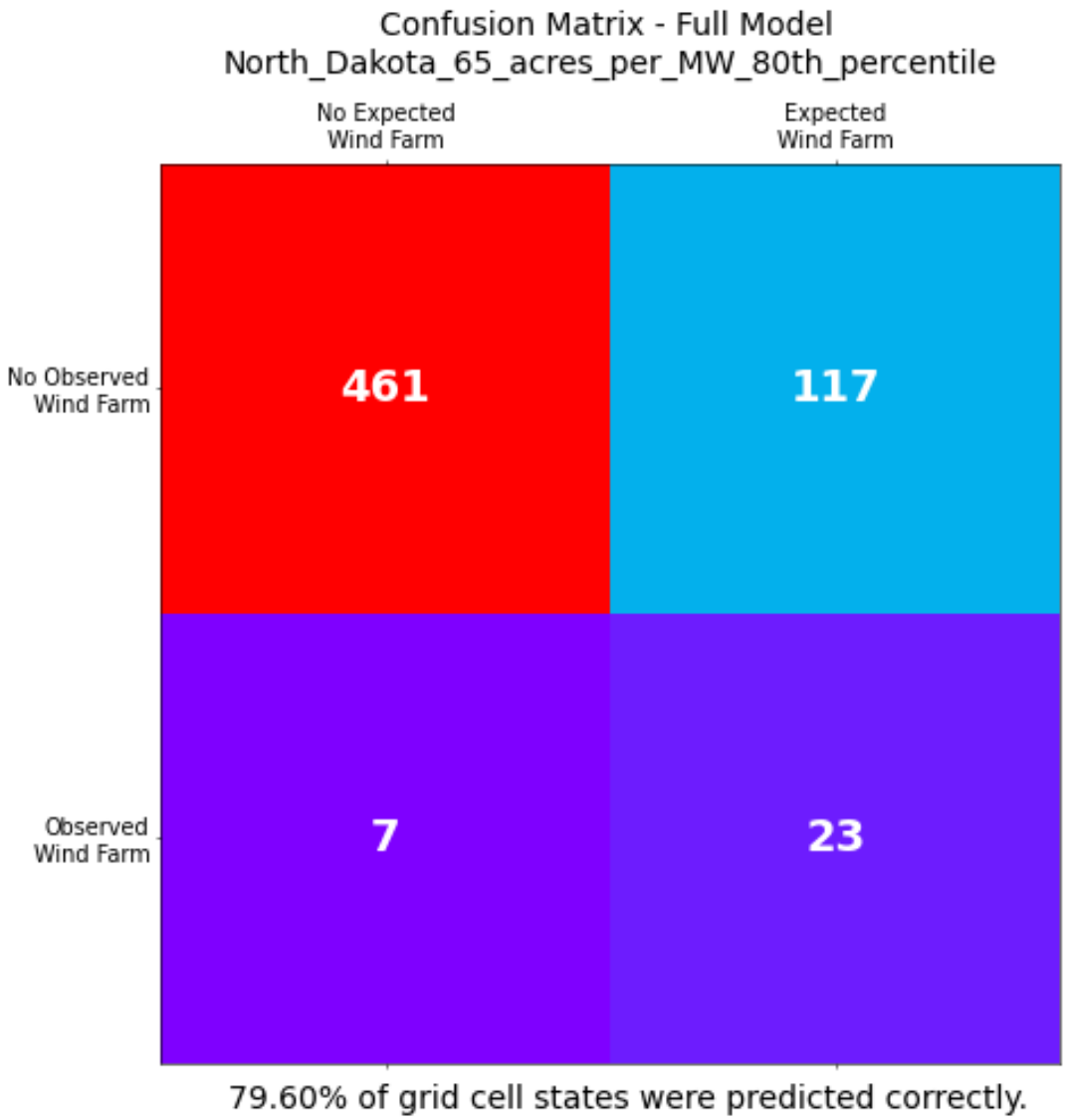
ROC curves generated from the 30 Full model runs with the testing data:



Range of Area Under Curve (AUC) statistics for the Full model:
Minimum AUC: 0.7643021914648211
Median AUC: 0.8561707035755479
Maximum AUC: 0.9097462514417531

Range of optimal threshold classifications for the Full model:
Minimum Threshold: 0.3514546008308326
Median Threshold: 0.5277309314921165
Maximum Threshold: 0.7236370259284025

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



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

Lower Quartile confusion matrix:
[[437 141]
 [5 25]]
Lower Quartile proportion of correctly predicted grid cell states by the Full model:
0.7598684210526315

Median confusion matrix:
[[461 117]
 [7 23]]
Median proportion of correctly predicted grid cell states by the Full model:
0.7960526315789473

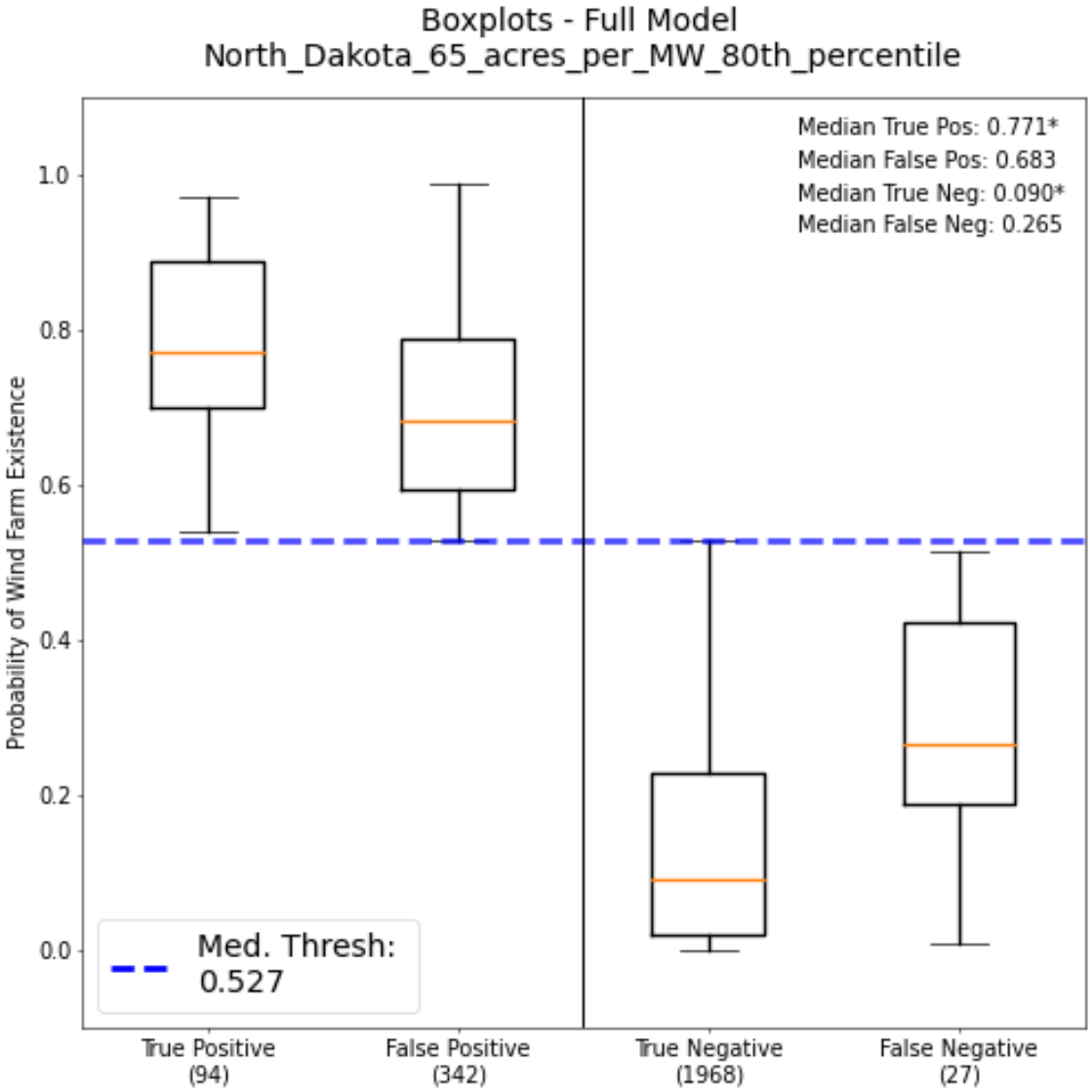
Upper Quartile confusion matrix:
[[486 92]
 [7 23]]
Upper Quartile proportion of correctly predicted grid cell states by the Full model:
0.837171052631579

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

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

Number of True Positive Grid Cells: 94
Number of False Positive Grid Cells: 342
Number of True Negative Grid Cells: 1968
Number of False Negative Grid Cells: 27

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: 0.683
Median True Pos: 0.771*
Median False Neg: 0.265
Median True Neg: 0.090*

Mann-Whitney U-test results:

Mann-Whitney Statistic - True Positive vs False Positive:
U-statistic = 10600.0
p-value = 4.221469902237081e-07

Mann-Whitney Statistic - True Negative vs False Negative:
U-statistic = 42112.0
p-value = 1.710285057888519e-07

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

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Wind_Farm_Predictor_Maps\Hexagon_Grid_65_acres_per_MW_80th_percentile_North_Dakota_Full.gdb\Hexagon_Grid_65_acres_per_MW_80th_percentile_North_Dakota_Full_Map

Total (Percentage) of all grid cells over North_Dakota that exist in hotspots:
154 (6.33%)
Total (Percentage) True Positive grid cells over North_Dakota that exist in hotspots:
49 (52.13%)
Total (Percentage) False Positive grid cells over North_Dakota that exist in hotspots:
105 (30.7%)

########## 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: 286.97479536429137
Median Score: 270.00621015141473
Minimum Score: 262.02694750640603

Range of log-likelihood scores of the Null model:
Maximum Score: 192.06115078192624
Median Score: 192.06115078192624
Minimum Score: 192.0611507819258

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: 155.89011873897698
p-value of the Median Log-Likelihood Ratio: 2.545708201907264e-19

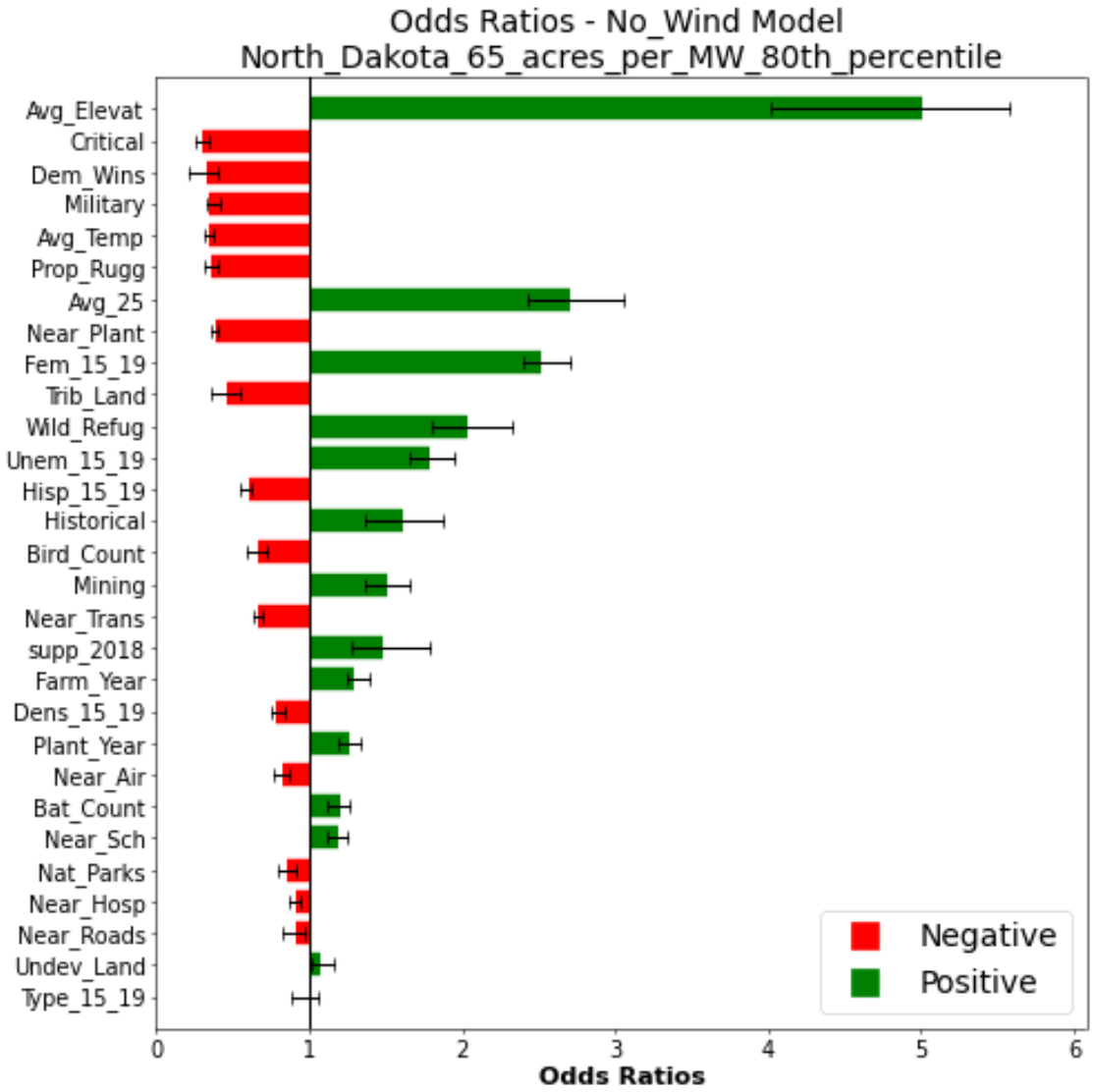
Range of McFadden Adjusted Psuedo R-Squared statistics for the No_Wind model:
Minimum Pseudo R-Squared: -0.3379842530261137
Median Pseudo R-Squared: -0.24963434392792527
Maximum Pseudo R-Squared: -0.208088916273643

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:

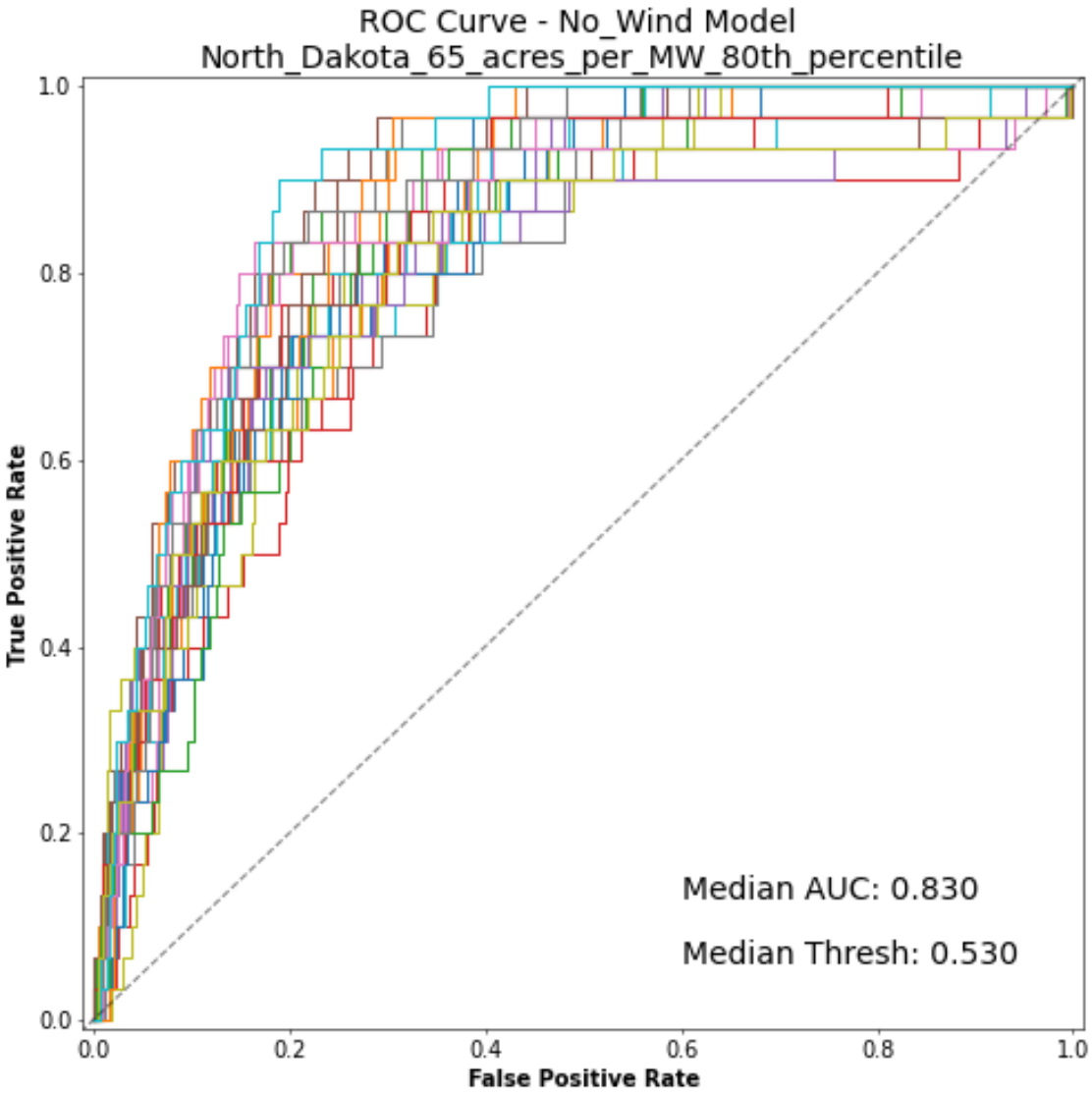
Predictor	Odds_Low	Odds_Med	Odds_Upp	Coef_Med	Rank
Avg_Elevat	4.012807	5.005547	5.579372	1.610547	1
Critical	0.262530	0.296811	0.338081	-1.214658	2
Dem_Wins	0.219365	0.334952	0.402102	-1.093768	3
Military	0.323754	0.341914	0.414499	-1.073195	4
Avg_Temp	0.313477	0.350042	0.379158	-1.049702	5
Prop_Rugg	0.312233	0.360579	0.400939	-1.020044	6
Avg_25	2.431311	2.704349	3.047880	0.994861	7
Near_Plant	0.357909	0.382965	0.397019	-0.959811	8
Fem_15_19	2.402663	2.518438	2.704495	0.923639	9
Trib_Land	0.364038	0.463483	0.546539	-0.768985	10
Wild_Refug	1.804858	2.038601	2.333639	0.712264	11
Unem_15_19	1.653882	1.782681	1.948565	0.578118	12
Hisp_15_19	0.543417	0.603674	0.626642	-0.504721	13
Historical	1.368876	1.608571	1.869442	0.475346	14
Bird_Count	0.590637	0.661795	0.721620	-0.412799	15
Mining	1.364922	1.504274	1.655726	0.408310	16
Near_Trans	0.633157	0.670805	0.698663	-0.399277	17
supp_2018	1.283654	1.486629	1.780895	0.396511	18
Farm_Year	1.243954	1.293195	1.388612	0.257116	19
Dens_15_19	0.751874	0.781696	0.838183	-0.246289	20
Plant_Year	1.193400	1.261435	1.334063	0.232250	21
Near_Air	0.767748	0.829386	0.871916	-0.187070	22
Bat_Count	1.116755	1.204547	1.255495	0.186104	23
Near_Sch	1.120345	1.191447	1.250764	0.175169	24
Nat_Parks	0.797006	0.848761	0.909746	-0.163977	25
Near_Hosp	0.861848	0.905860	0.946432	-0.098871	26
Near_Roads	0.831402	0.917508	0.964095	-0.086094	27
Undev_Land	1.013849	1.079854	1.155823	0.076825	28
Type_15_19	0.890451	0.997126	1.051825	-0.002878	29

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



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

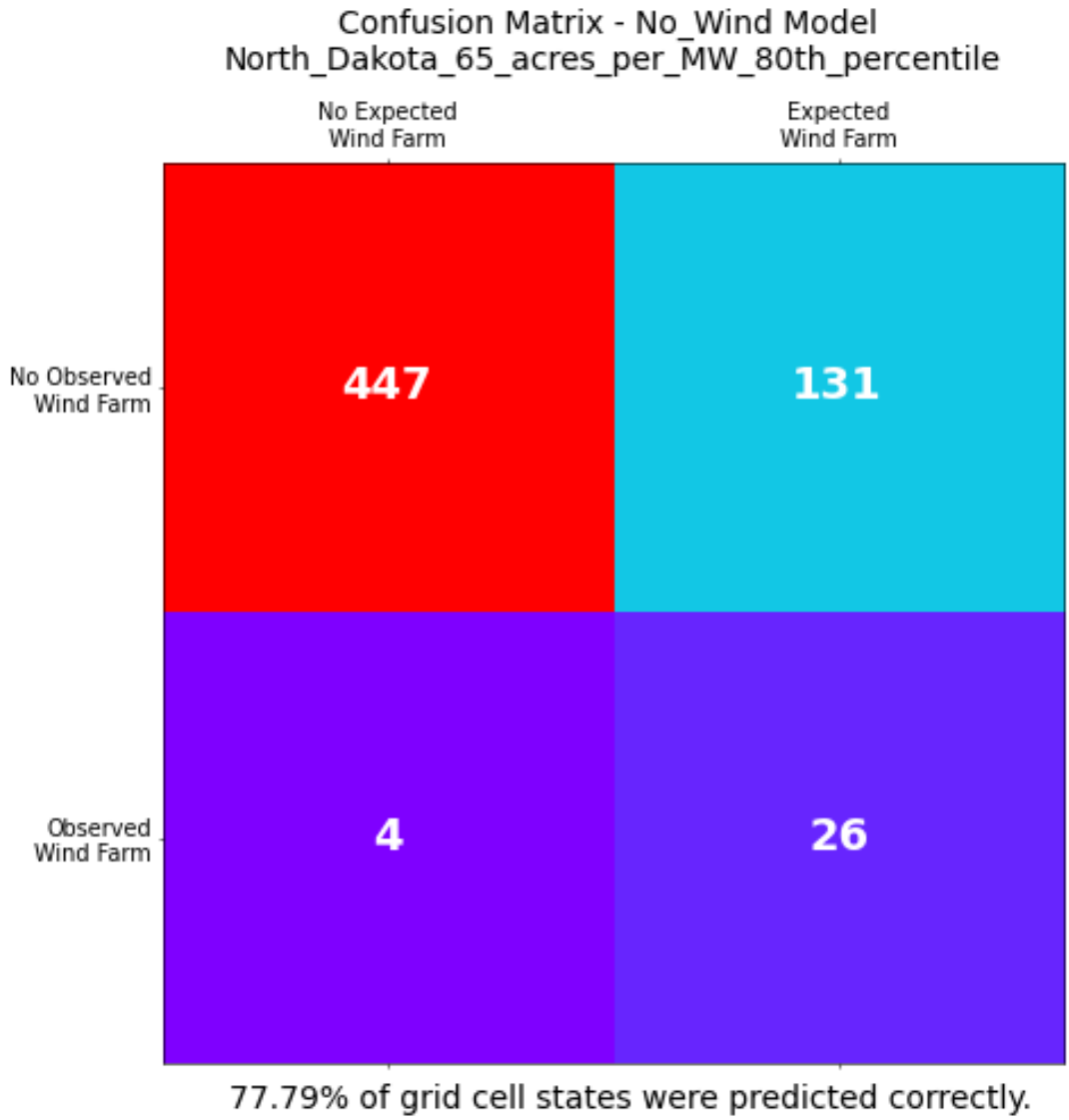
ROC curves generated from the 30 No_Wind model runs with the testing data:



Range of Area Under Curve (AUC) statistics for the No_Wind model:
Minimum AUC: 0.7544405997693194
Median AUC: 0.830565167243368
Maximum AUC: 0.8972895040369089

Range of optimal threshold classifications for the No_Wind model:
Minimum Threshold: 0.23254589420840086
Median Threshold: 0.5305602366455018
Maximum Threshold: 0.6876002912676701

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



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

Lower Quartile confusion matrix:
[[400 178]
 [3 27]]
Lower Quartile proportion of correctly predicted grid cell states by the No_Wind model:
0.7023026315789473

Median confusion matrix:
[[447 131]
 [4 26]]
Median proportion of correctly predicted grid cell states by the No_Wind model:
0.7779605263157895

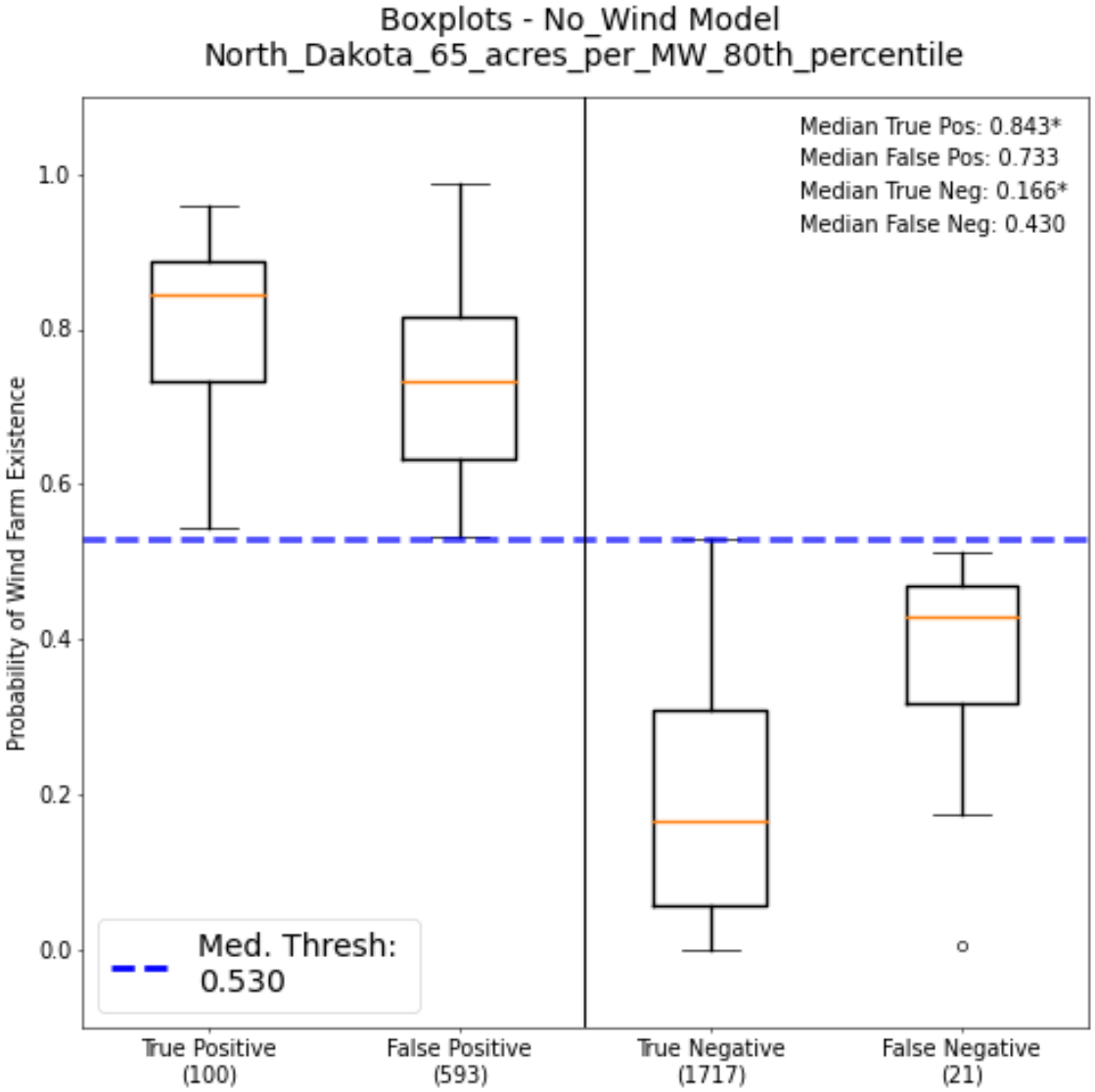
Upper Quartile confusion matrix:
[[463 115]
 [8 22]]
Upper Quartile proportion of correctly predicted grid cell states by the No_Wind model:
0.7976973684210527

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

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

Number of True Positive Grid Cells: 100
Number of False Positive Grid Cells: 593
Number of True Negative Grid Cells: 1717
Number of False Negative Grid Cells: 21

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: 0.733
Median True Pos: 0.843*
Median False Neg: 0.430
Median True Neg: 0.166*

Mann-Whitney U-test results:

Mann-Whitney Statistic - True Positive vs False Positive:
U-statistic = 18353.0
p-value = 1.0609071761299646e-09

Mann-Whitney Statistic - True Negative vs False Negative:
U-statistic = 29428.0
p-value = 6.142096843402537e-07

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

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Wind_Farm_Predictor_Maps\Hexagon_Grid_65_acres_per_MW_80th_percentile_North_Dakota_No_Wind.gdb\Hexagon_Grid_65_acres_per_MW_80th_percentile_North_Dakota_No_Wind_Map

Total (Percentage) of all grid cells over North_Dakota that exist in hotspots:
62 (2.55%)
Total (Percentage) True Positive grid cells over North_Dakota that exist in hotspots:
21 (21.0%)
Total (Percentage) False Positive grid cells over North_Dakota that exist in hotspots:
41 (6.91%)

########## 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: 247.68000458245479
Median Score: 241.7327942321624
Minimum Score: 233.1615568807847

Range of log-likelihood scores of the Null model:
Maximum Score: 192.06115078192624
Median Score: 192.06115078192624
Minimum Score: 192.0611507819258

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: 99.34328690047232
p-value of the Median Log-Likelihood Ratio: 2.123163299796115e-23

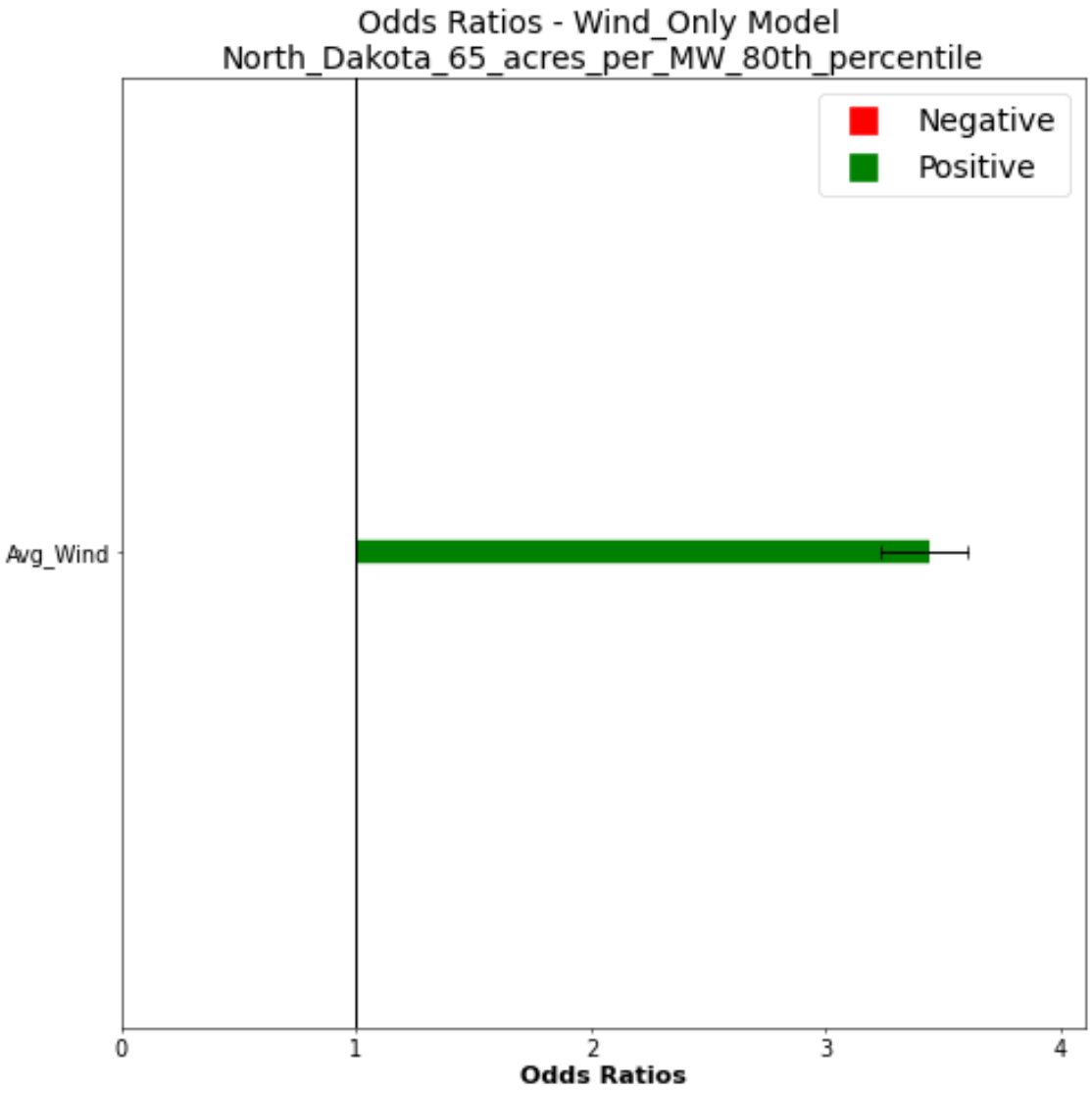
Range of McFadden Adjusted Psuedo R-Squared statistics for the Wind_Only model:
Minimum Pseudo R-Squared: -0.2791759477761855
Median Pseudo R-Squared: -0.2482107560855158
Maximum Pseudo R-Squared: -0.20358310850305483

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:

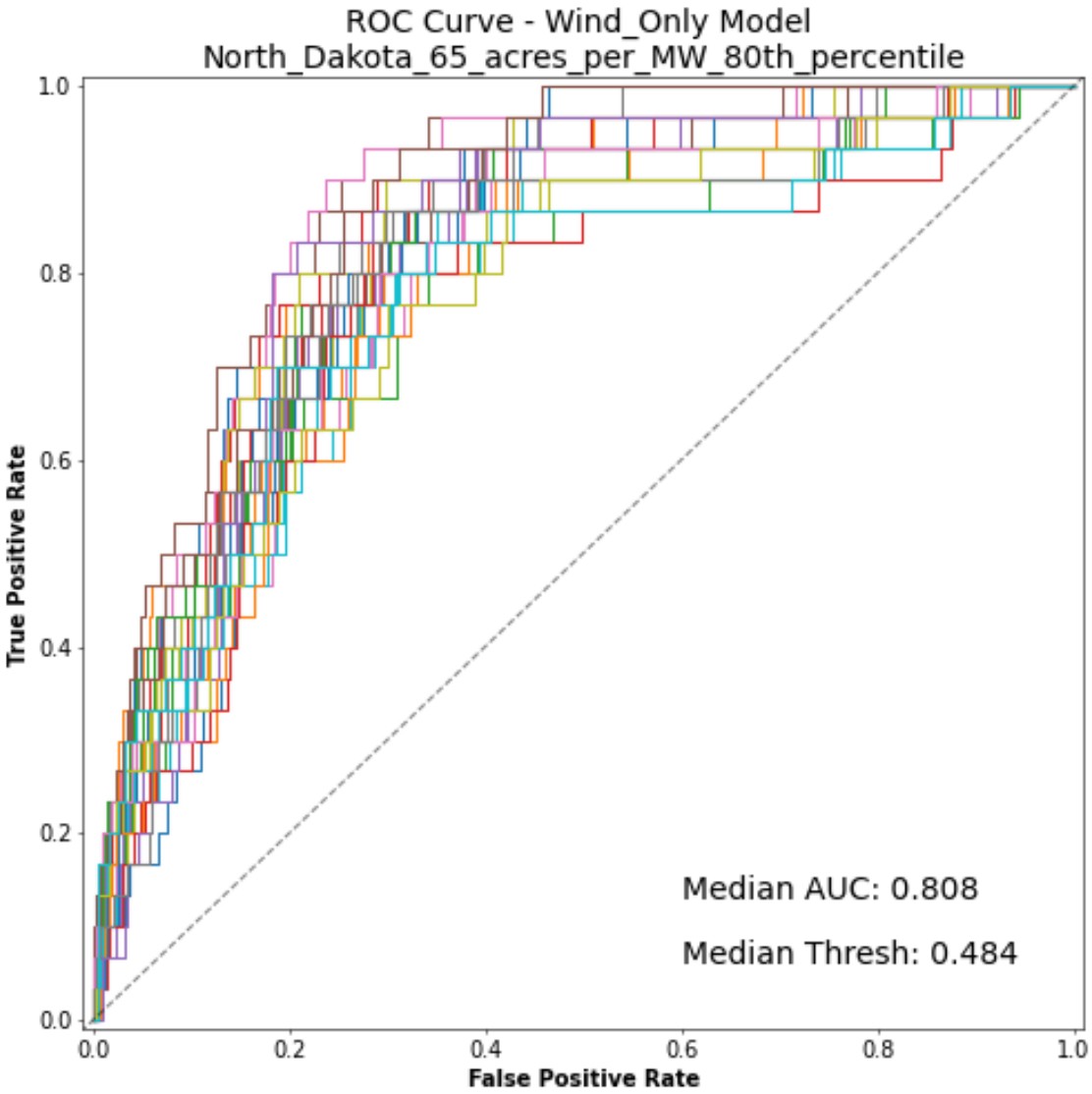
Predictor	Odds_Low	Odds_Med	Odds_Upp	Coef_Med	Rank
Avg_Wind	3.236745	3.437126	3.599995	1.234636	1

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



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

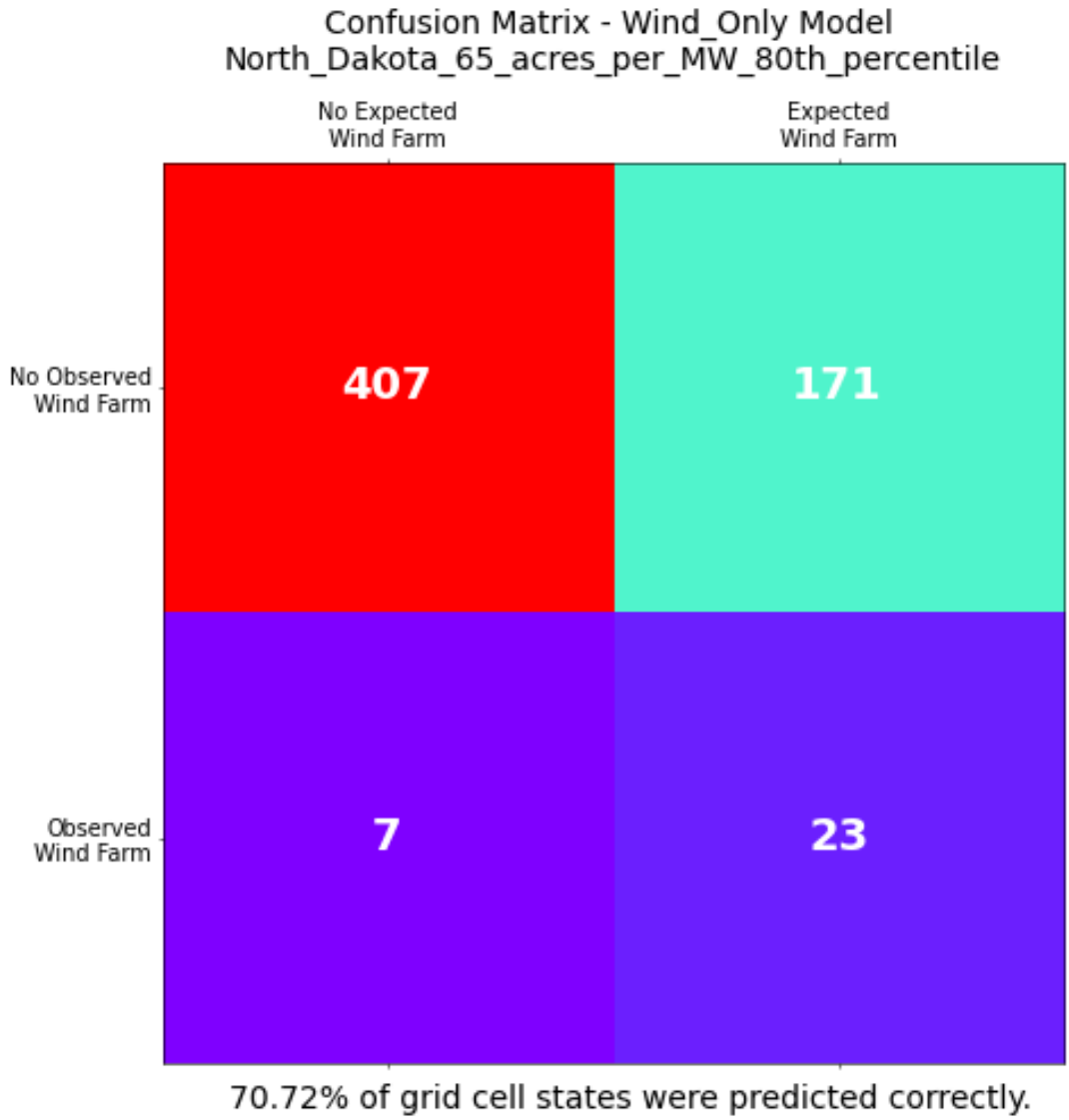
ROC curves generated from the 30 Wind_Only model runs with the testing data:



Range of Area Under Curve (AUC) statistics for the Wind_Only model:
Minimum AUC: 0.7504613610149943
Median AUC: 0.8086505190311419
Maximum AUC: 0.8754901960784314

Range of optimal threshold classifications for the Wind_Only model:
Minimum Threshold: 0.3798597129522466
Median Threshold: 0.48467038391871253
Maximum Threshold: 0.647394846763038

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



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

Lower Quartile confusion matrix:
[[385 193]
 [4 26]]
Lower Quartile proportion of correctly predicted grid cell states by the Wind_Only model:
0.6759868421052632

Median confusion matrix:
[[407 171]
 [7 23]]
Median proportion of correctly predicted grid cell states by the Wind_Only model:
0.7072368421052632

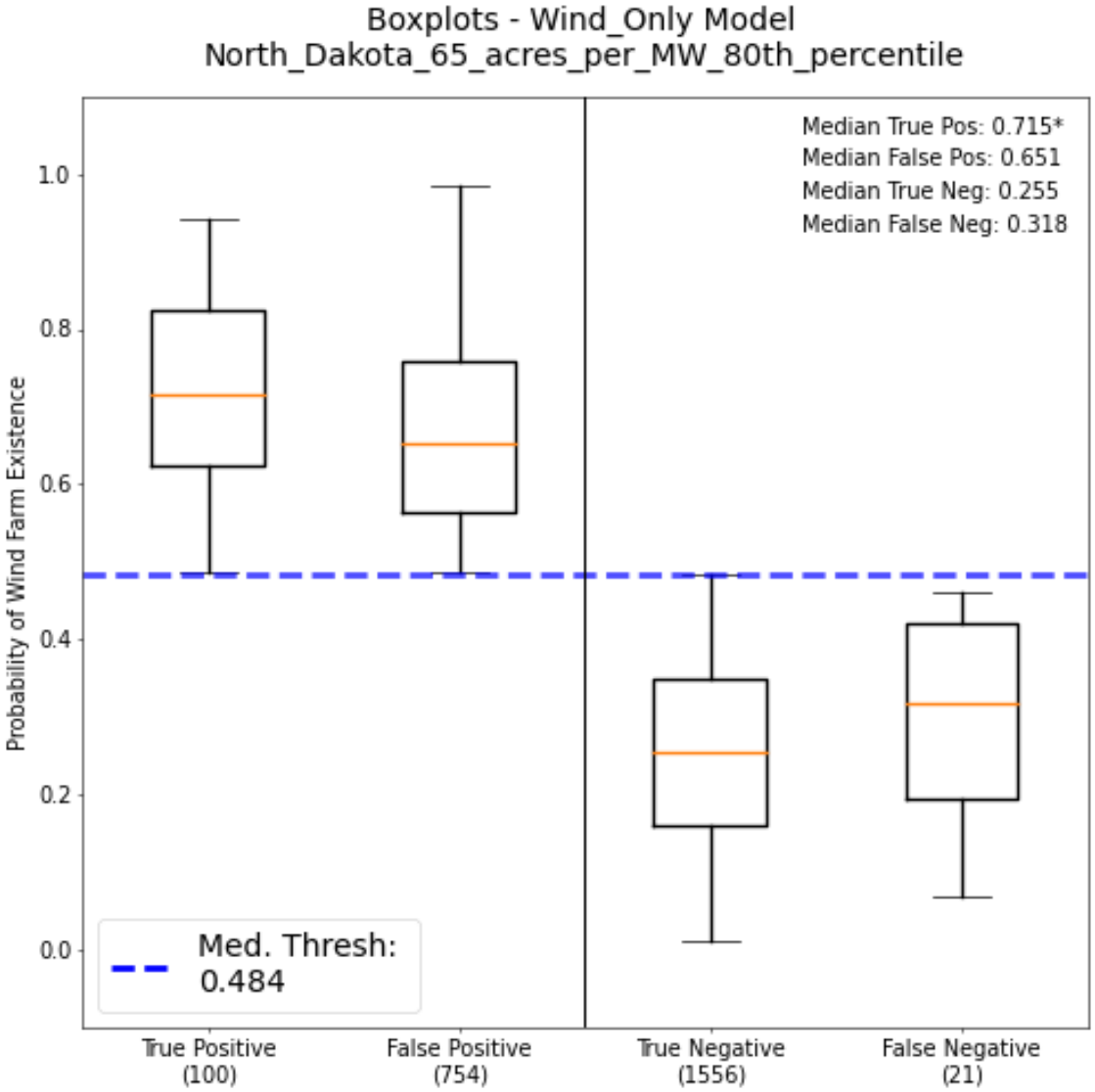
Upper Quartile confusion matrix:
[[433 145]
 [6 24]]
Upper Quartile proportion of correctly predicted grid cell states by the Wind_Only model:
0.7516447368421053

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

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

Number of True Positive Grid Cells: 100
Number of False Positive Grid Cells: 754
Number of True Negative Grid Cells: 1556
Number of False Negative Grid Cells: 21

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: 0.651
Median True Pos: 0.715*
Median False Neg: 0.318
Median True Neg: 0.255

Mann-Whitney U-test results:

Mann-Whitney Statistic - True Positive vs False Positive:
U-statistic = 27462.0
p-value = 1.0013480870324004e-05

Mann-Whitney Statistic - True Negative vs False Negative:
U-statistic = 19696.0
p-value = 0.10529343470081766

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

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Wind_Farm_Predictor_Maps\Hexagon_Grid_65_acres_per_MW_80th_percentile_North_Dakota_Wind_Only.gdb\Hexagon_Grid_65_acres_per_MW_80th_percentile_North_Dakota_Wind_Only_Map

Total (Percentage) of all grid cells over North_Dakota that exist in hotspots:
63 (2.59%)
Total (Percentage) True Positive grid cells over North_Dakota that exist in hotspots:
19 (19.0%)
Total (Percentage) False Positive grid cells over North_Dakota that exist in hotspots:
44 (5.84%)

########## 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
Avg_Wind	30	27
Avg_25	30	17
Unem_15_19	30	14
Avg_Temp	30	13
Near_Plant	30	10
Fem_15_19	29	11
Wild_Refug	28	13
Bird_Count	28	10
Mining	27	9
Dem_Wins	27	4
Near_Hosp	27	1
Near_Roads	27	1
supp_2018	27	0
Trib_Land	27	0
Type_15_19	26	8
Historical	26	3
Avg_Elevat	25	6
Near_Air	25	4

Hisp_15_19	25	4
Prop_Rugg	25	3
Near_Trans	25	3
Military	25	3
Bat_Count	25	2
Critical	25	0
Undev_Land	24	2
Near_Sch	24	2
Nat_Parks	23	1
Dens_15_19	23	0
Farm_Year	22	1
Plant_Year	20	3

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	10	0.831414	0.229167
1	11	0.818257	0.227723
2	19	0.814967	0.216981
3	13	0.794408	0.198276
4	22	0.793586	0.205128
5	30	0.790296	0.196721
6	18	0.790296	0.188525
7	17	0.788651	0.196721
8	15	0.788651	0.188525
9	12	0.787007	0.196721
10	6	0.787007	0.175000
11	14	0.785362	0.192000
12	9	0.784539	0.185484
13	8	0.783717	0.188525
14	20	0.782895	0.190476
15	29	0.778783	0.178295
16	16	0.774671	0.184615
17	28	0.767270	0.177778
18	27	0.763158	0.175182
19	24	0.763158	0.172662
20	21	0.763158	0.171429
21	7	0.762336	0.159420
22	23	0.756579	0.176056
23	4	0.751645	0.159722
24	25	0.748355	0.162162
25	5	0.739309	0.157895
26	26	0.739309	0.169935
27	2	0.722862	0.148148
28	1	0.717105	0.144578
29	3	0.696546	0.134831

Set of predictors (10 total) to be used in the Reduced Model:

['Avg_Wind', 'Avg_25', 'Unem_15_19', 'Avg_Temp', 'Near_Plant', 'Fem_15_19', 'Wild_Refug', 'Bird_Count', 'Mining', 'Dem_Wins']

Range of log-likelihood scores from 30 training runs of the Reduced model:
Maximum Score: 282.29590687118616
Median Score: 268.9696065606254
Minimum Score: 258.656328739914

Range of log-likelihood scores of the Null model:
Maximum Score: 192.06115078192624
Median Score: 192.06115078192624
Minimum Score: 192.06115078192624

Number of times (out of 30) the Reduced model possesses a greater goodness-of-fit: 30
Number of times (out of 30) the Reduced model's outperformance of the Null model is statistically significant: 30

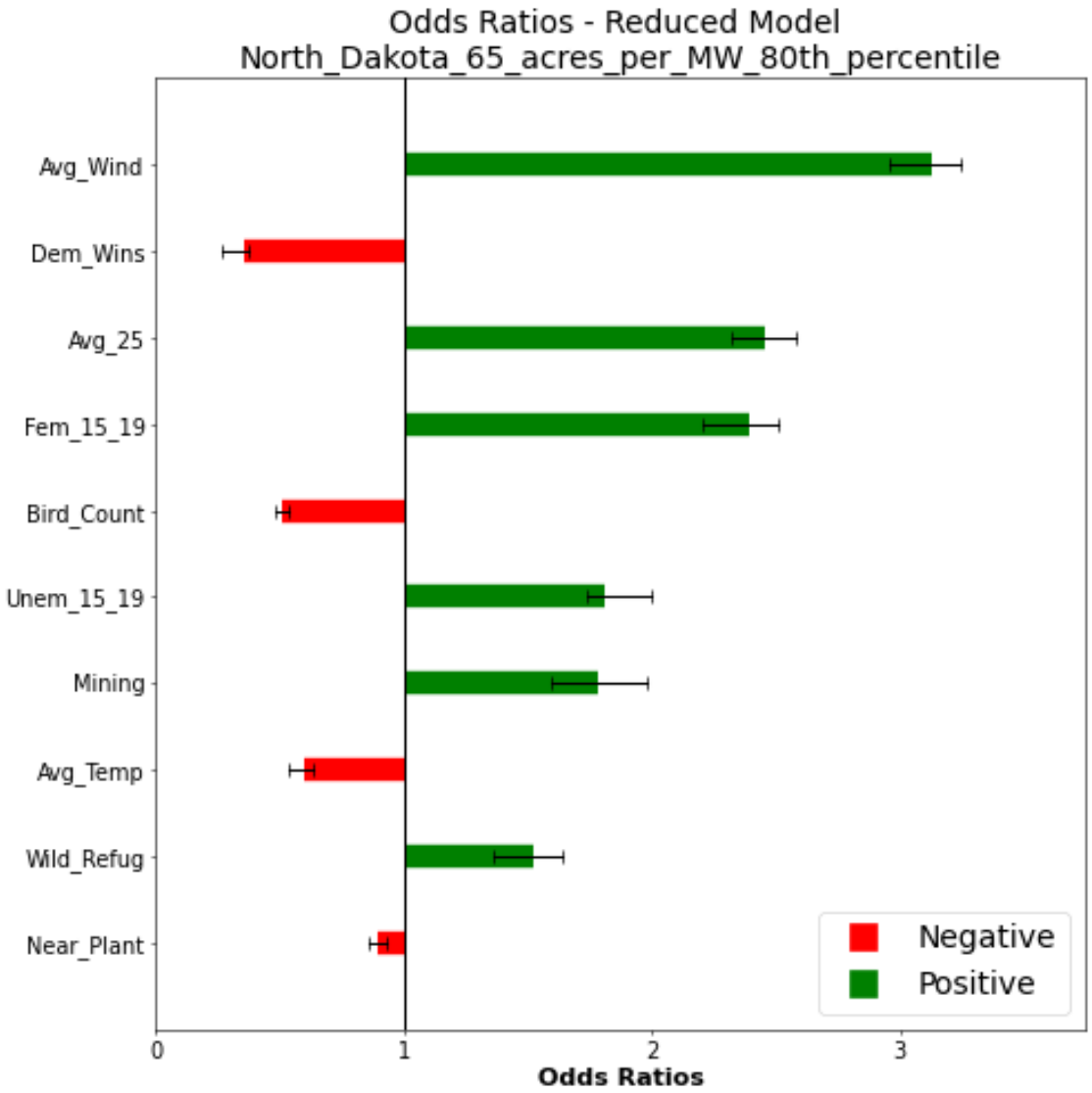
Median Log-Likelihood Ratio, Reduced model vs. Null model: 153.81691155739827
p-value of the Median Log-Likelihood Ratio: 1.42659548174673e-28

Range of McFadden Adjusted Psuedo R-Squared statistics for the Reduced model:
Minimum Pseudo R-Squared: -0.417756301899709
Median Pseudo R-Squared: -0.34837058669230625
Maximum Pseudo R-Squared: -0.2946727004788601

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:

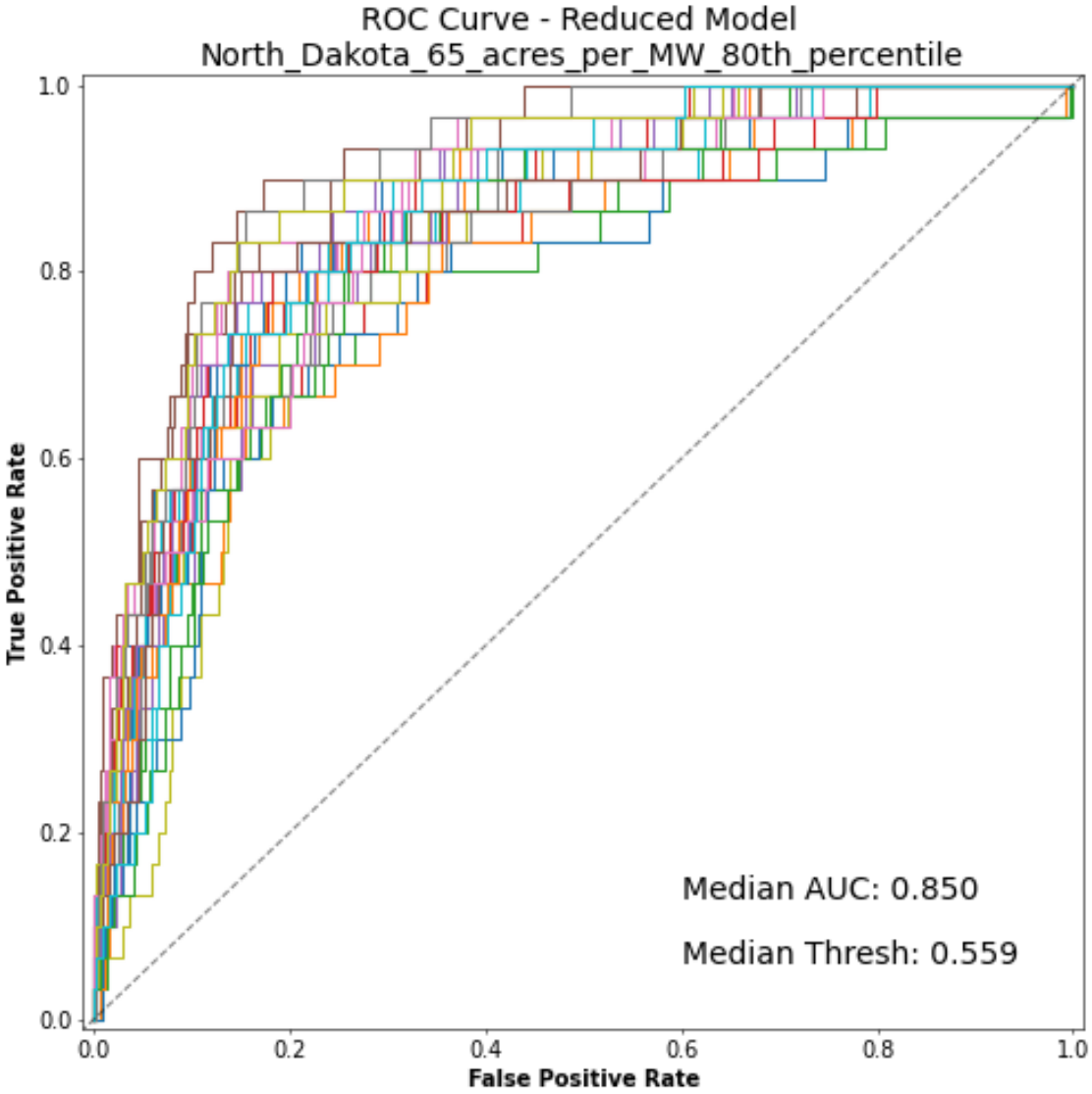
Predictor	Odds_Low	Odds_Med	Odds_Upp	Coef_Med	Rank
Avg_Wind	2.956537	3.132174	3.245010	1.141727	1
Dem_Wins	0.270046	0.351637	0.376155	-1.045156	2
Avg_25	2.324771	2.460028	2.578563	0.900173	3
Fem_15_19	2.201468	2.397556	2.510154	0.874450	4
Bird_Count	0.480328	0.507257	0.538667	-0.678737	5
Unem_15_19	1.738341	1.808947	1.999943	0.592745	6
Mining	1.591236	1.786420	1.983953	0.580214	7
Avg_Temp	0.531693	0.595573	0.637661	-0.518231	8
Wild_Refug	1.363430	1.522149	1.643509	0.420123	9
Near_Plant	0.854312	0.890513	0.927312	-0.115958	10

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



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

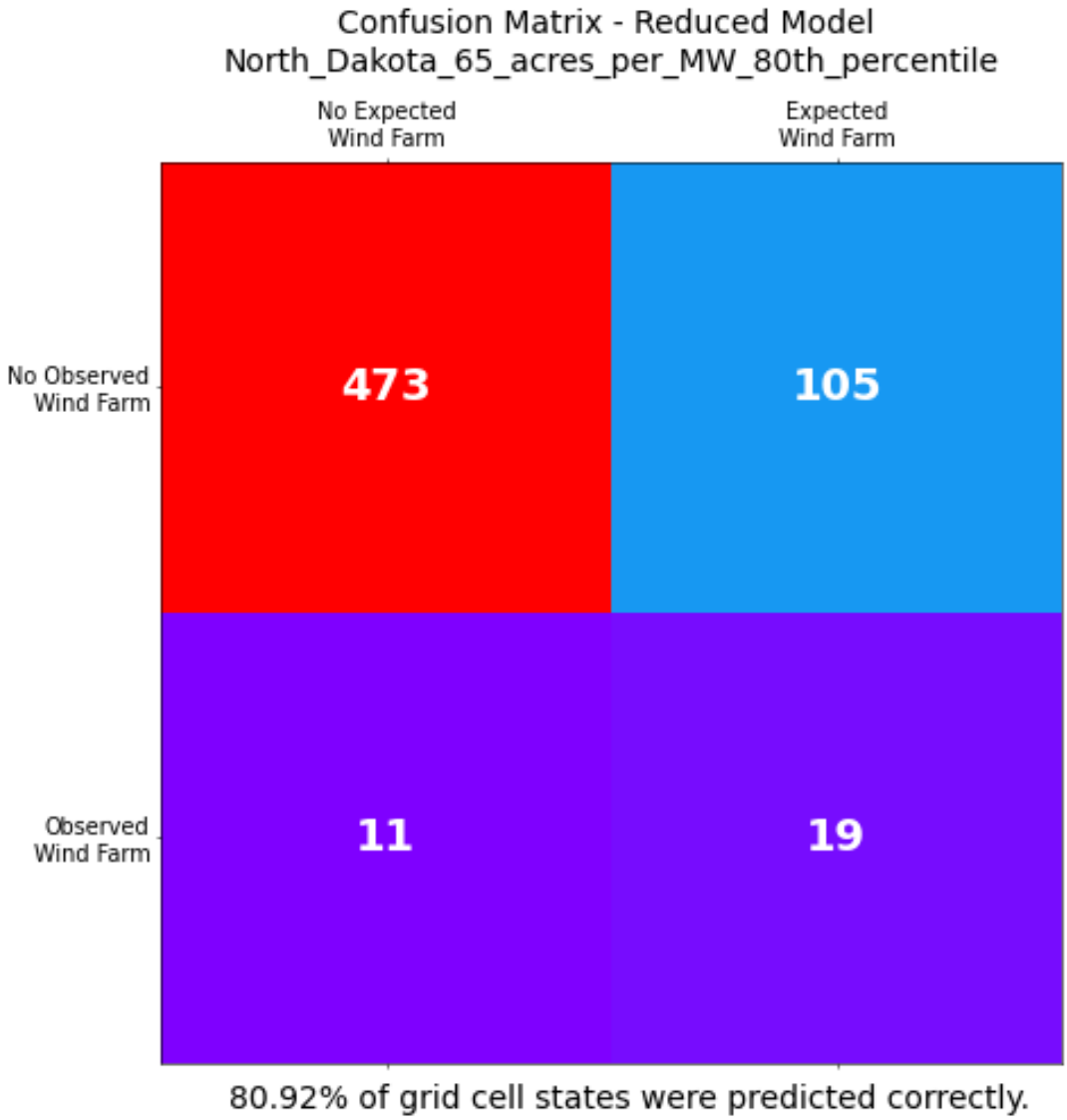
ROC curves generated from the 30 Reduced model runs with the testing data:



Range of Area Under Curve (AUC) statistics for the Reduced model:
Minimum AUC: 0.7751441753171857
Median AUC: 0.8503748558246829
Maximum AUC: 0.9141868512110726

Range of optimal threshold classifications for the Reduced model:
Minimum Threshold: 0.345964586905432
Median Threshold: 0.5591590325001985
Maximum Threshold: 0.6671842596003343

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



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

Lower Quartile confusion matrix:
[[428 150]
 [6 24]]
Lower Quartile proportion of correctly predicted grid cell states by the Reduced model:
0.743421052631579

Median confusion matrix:
[[473 105]
 [11 19]]
Median proportion of correctly predicted grid cell states by the Reduced model:
0.8092105263157895

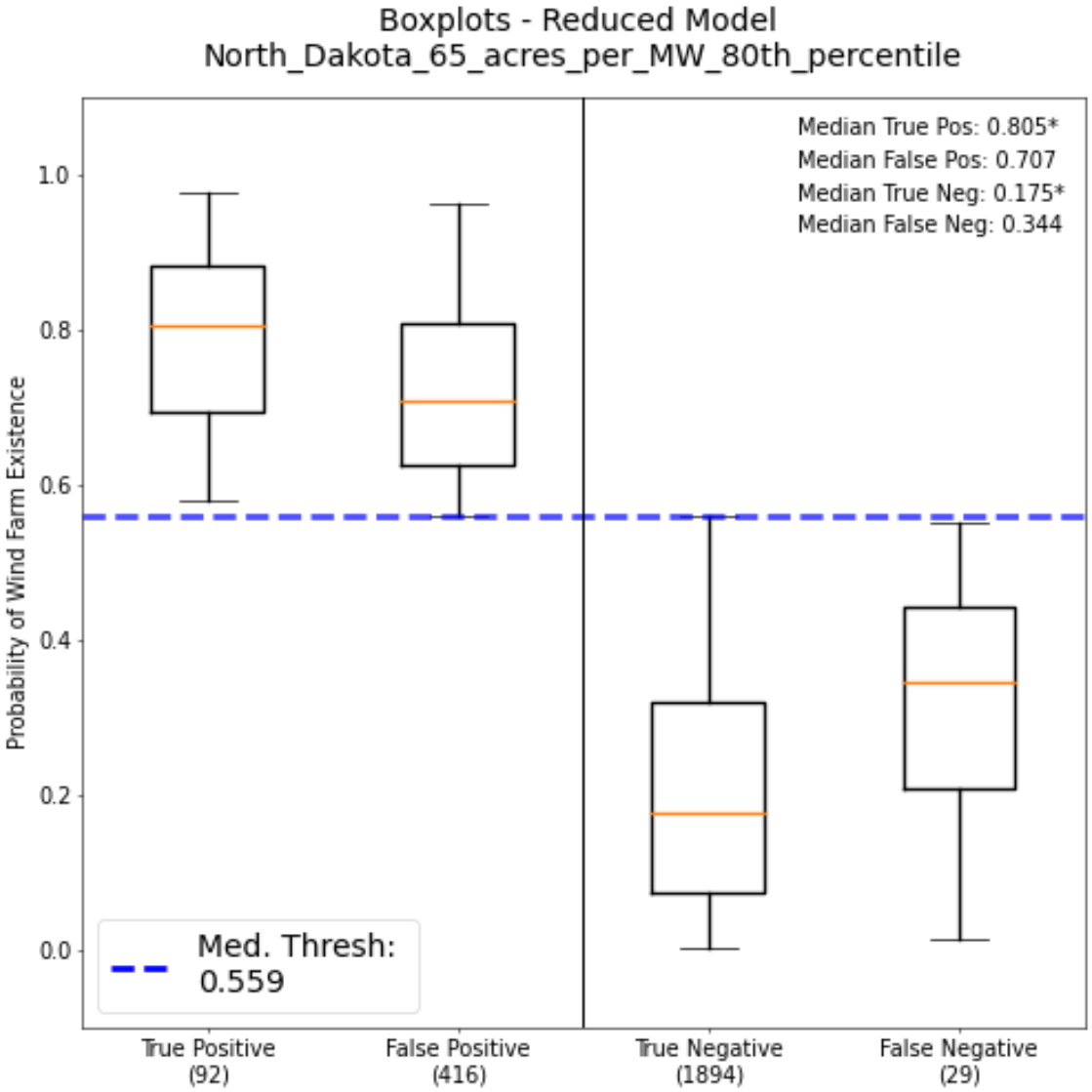
Upper Quartile confusion matrix:
[[487 91]
 [8 22]]
Upper Quartile proportion of correctly predicted grid cell states by the Reduced model:
0.837171052631579

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

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

Number of True Positive Grid Cells: 92
Number of False Positive Grid Cells: 416
Number of True Negative Grid Cells: 1894
Number of False Negative Grid Cells: 29

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: 0.707
Median True Pos: 0.805*
Median False Neg: 0.344
Median True Neg: 0.175*

Mann-Whitney U-test results:

Mann-Whitney Statistic - True Positive vs False Positive:
U-statistic = 12146.0
p-value = 4.1168204076561793e-08

Mann-Whitney Statistic - True Negative vs False Negative:
U-statistic = 39484.0
p-value = 5.1081834228640846e-05

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

Filepath to the constructed hexagonal grid map:

D:\Dissertation_Resources\Wind_Farm_Predictor_Maps\Hexagon_Grid_65_acres_per_MW_80th_percentile_North_Dakota_Reduced.gdb\Hexagon_Grid_65_acres_per_MW_80th_percentile_North_Dakota_Reduced_Map

Total (Percentage) of all grid cells over North_Dakota that exist in hotspots:
119 (4.9%)
Total (Percentage) True Positive grid cells over North_Dakota that exist in hotspots:
37 (40.22%)
Total (Percentage) False Positive grid cells over North_Dakota that exist in hotspots:
82 (19.71%)