Predicting West Nile Virus in Chicago

West Nile Virus Data

Data sourced by Chicago Department of Public Health

Datasets

- Mosquito West Nile Virus Training (2007, 2009, 2011, 2013)
- Spray Information (2011 - 2013)
- Weather Information (2007 - 2014, from two Stations)
- Mosquito West Nile Virus Test (2008, 2010, 2012, 2014)

Preparing the Data

Combined train and test data with weather data.

Hot and dry conditions are favorable to the virus, so features of interest:

- Heat (Days above 65)
- Cool (Days below 65)
- Wet Bulb
- Dew Point
- Total Precipitation

Preparing the data continued...

Created two modeling datasets

- West Nile Virus data + daily weather
- West Nile Virus data + monthly averages

Daily set smaller (~1000). Observations with Trace and/or Missing were dropped.

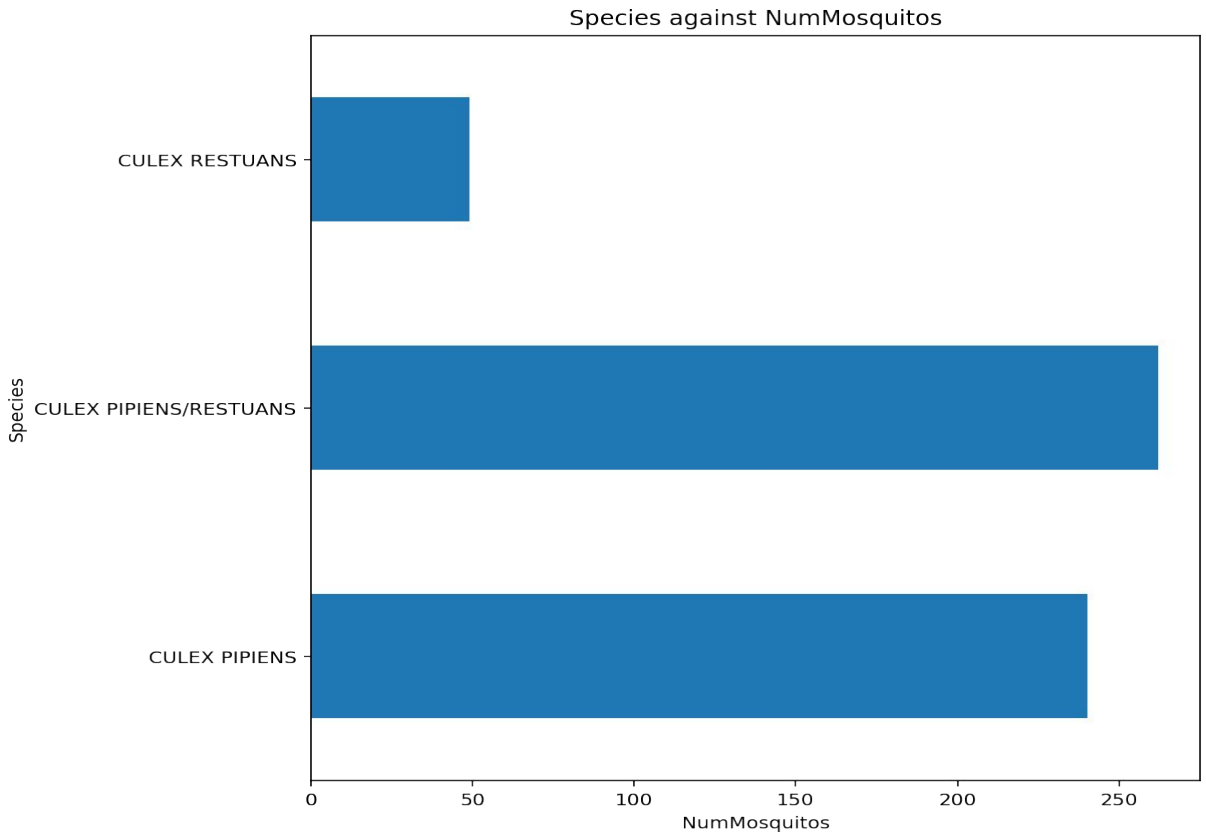
Monthly average set maintained all West Nile Virus observations.

No real difference in model success. Daily performed slightly better.

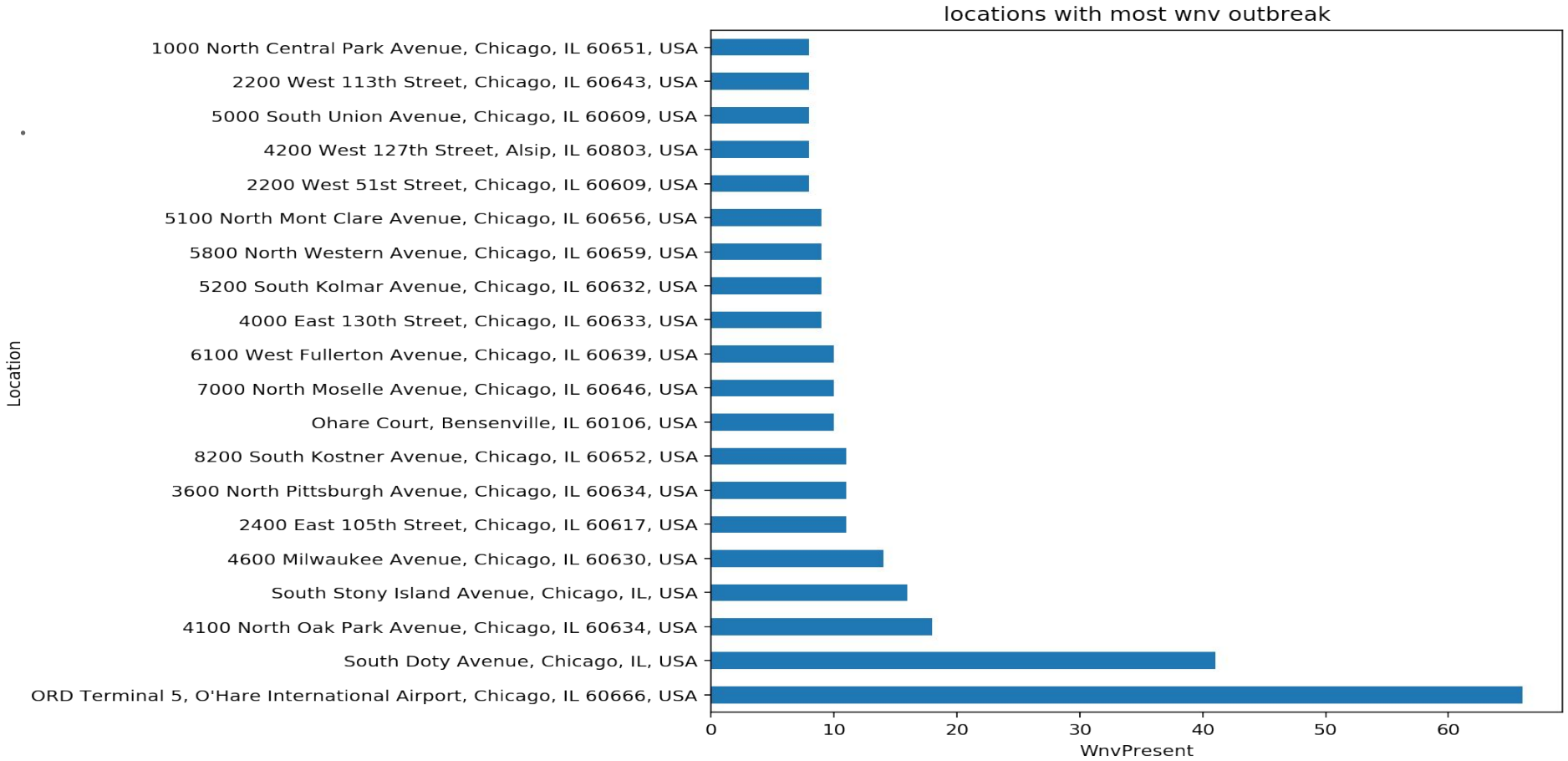
Mosquitoes most associated with West Nile Virus

Looking at the species number below you see that Cules Pipiens/Restuans is the most but i would also be careful with culex pipiens considering the ratio. So it will be of great benefit to concentrate more pesticide spraying or trapping in areas where Culex Pipiens are found also not ignoring the other areas.

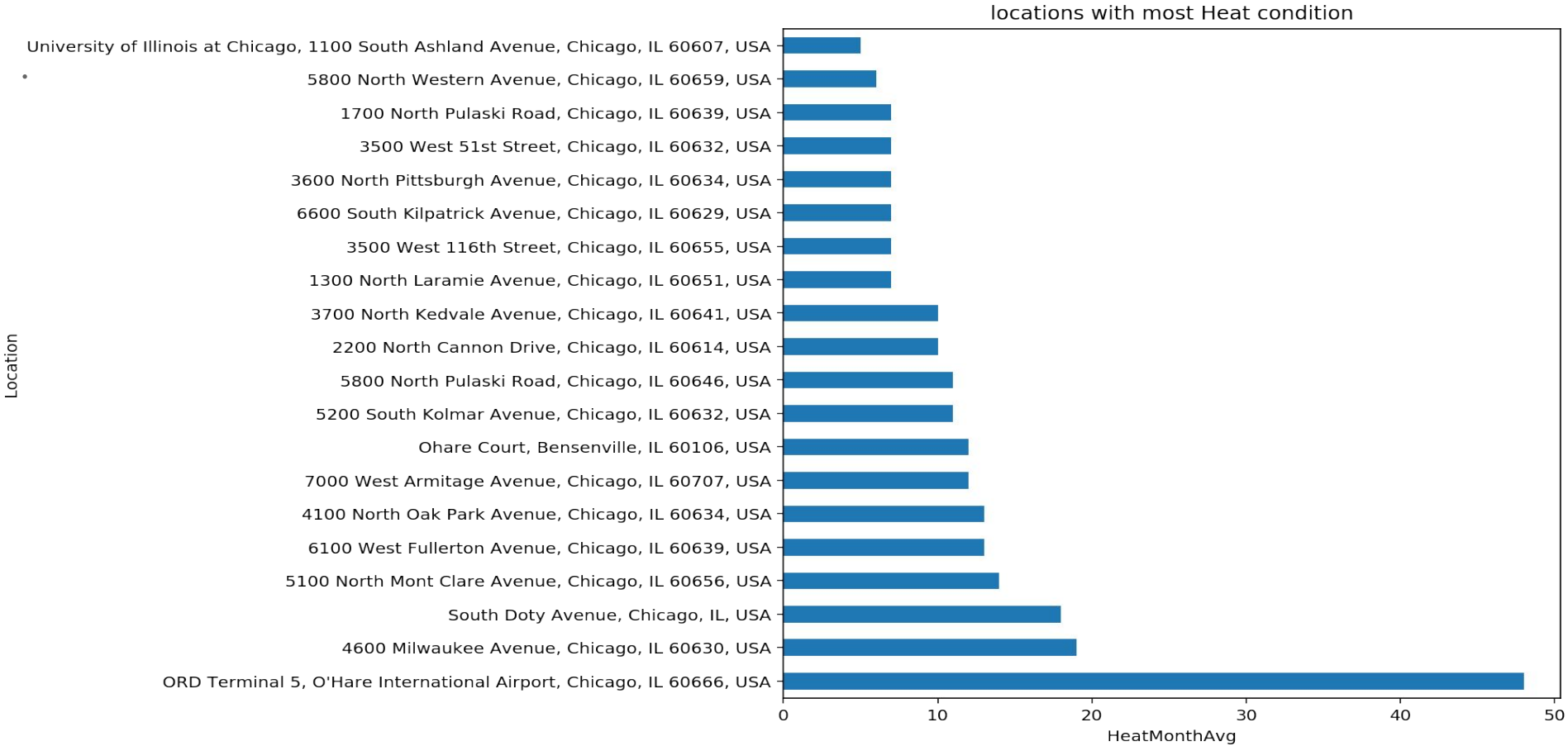
CULEX PIPIENS/RESTUANS	**262**
CULEX PIPIENS	**240**
CULEX RESTUANS	**49**



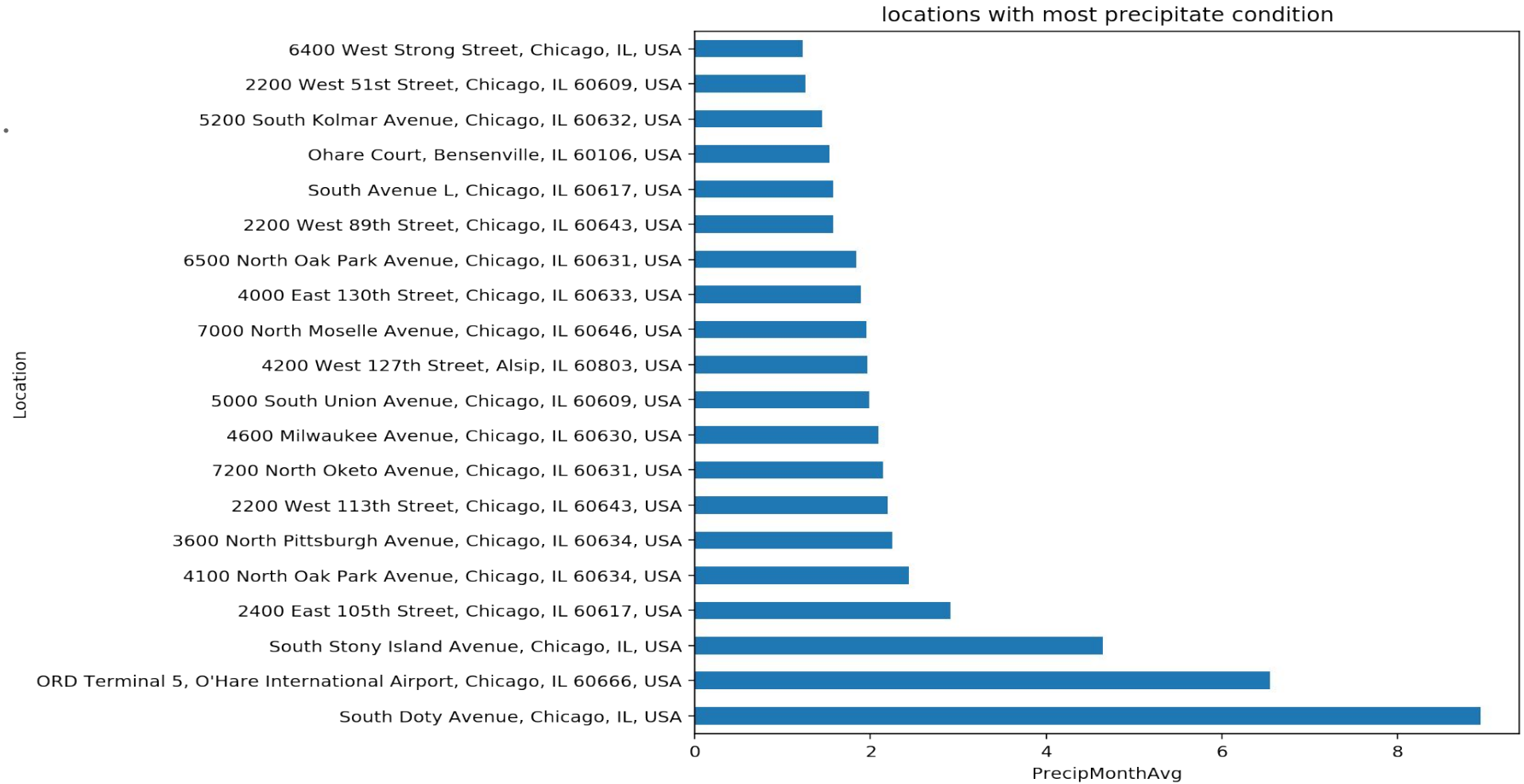
Now we need to know locations with the highest number of virus outbreak so as to know the prevalent specie, weather condition and time



Location with high heat condition



Location with most Precipitation



Model Baseline and Predictors

Set up baseline by looking at the most occurring class: “0”

- Baseline Acc - 94.76%

- Model Predictors: Trap, Species, Heat, Cold, PrecipTotal, Wetbulb, DewPoint

Multinomial Naive Bayes

$R^2 = 0.9490$

Only predicted two cases of west nile virus

Kaggle score = 0.6327

Neural Network

Network Topology : 7 > 7 > 7 > 1

0.2 dropout in each hidden layer

R^2 = 0.9490

Predicted zero cases of west nile virus

Random Forest Classifer

R^2 = 0.9345

Kaggle score = 0.5944

Specificity = 0.9799

Precision = 0.2063

Predicted / Actual	0	1	All
0	2443	50	2493
1	121	13	134
All	2564	63	2627

Locations to Spray

Top 10 predicted locations of WNV incidents

4100 North Oak Park Avenue, Chicago, IL 60634, USA	25
4600 North Melvina Avenue, Chicago, IL 60630, USA	22
4000 North Tripp Avenue, Chicago, IL 60641, USA	21
3900 North Springfield Avenue, Chicago, IL 60618, USA	17
ORD Terminal 5, O'Hare International Airport, Chicago, IL 60666, USA	17
5800 North Ridge Avenue, Chicago, IL 60660, USA	17
1000 East 67th Street, Chicago, IL 60637, USA	15
South Wallace Street, Chicago, IL, USA	14
8000 South Kedzie Avenue, Chicago, IL 60652, USA	10
4900 West Sunnyside Avenue, Chicago, IL 60630, USA	10

Cost Benefit Analysis

Predicted a total of 69 locations to contain the virus

Cost for a season of coverage: $700/location (Quote from Mosquito Squad Chicago)

Cost for a single session spray: about $200

A single spray lasts 21 days

Cost Benefit Analysis cont.

- Tiers of Coverage:
  - Full season spray of all 69 locations where WNV present
    - $48,300
  - Two session spray of 69 locations where WNV present
    - $27,600
  - Two session spray of the top 20 locations where WNV present
    - $8000

Month	Street	WnvPresent
8	S DOTY AVE	21
8	W OHARE AIRPORT	16
8	S STONY ISLAND AVE	13
8	W OHARE AIRPORT	12
9	W OHARE AIRPORT	9
8	W OHARE AIRPORT	6
9	S DOTY AVE	6
8	E 105TH ST	6
9	W OHARE AIRPORT	5
8	N OAK PARK AVE	5
8	N MOSELL AVE	4
8	N PITTSBURGH AVE	4
8	N OKETO AVE	4
8	N OAK PARK AVE	4
8	N CENTRAL PARK DR	4
8	S DOTY AVE	4
8	E 130TH ST	4
8	S UNION AVE	3
8	W 127TH PL	3
8	E 105TH ST	3

Conclusion

Most WNV presence is detected in August.

Recommend a two session spray at all of the locations, once at the beginning of August and once at the end of August.

Month	Street	WnvPresent
8	S DOTY AVE	21
8	W OHARE AIRPORT	16
8	S STONY ISLAND AVE	13
8	W OHARE AIRPORT	12
9	W OHARE AIRPORT	9
8	W OHARE AIRPORT	6
9	S DOTY AVE	6
8	E 105TH ST	6
9	W OHARE AIRPORT	5
8	N OAK PARK AVE	5
8	N MOSELL AVE	4
8	N PITTSBURGH AVE	4
8	N OKETO AVE	4
8	N OAK PARK AVE	4
8	N CENTRAL PARK DR	4
8	S DOTY AVE	4
8	E 130TH ST	4
8	S UNION AVE	3
8	W 127TH PL	3
8	E 105TH ST	3

Recommendation and Next Steps

- More spraying data. How necessary it is to reducing risk?
- Zoning data to look at patterns between commercial and residential blocks.
- Health informatics. Pesticides are a concern for many.