West Nile Prevention:

Pesticide recommendations using predictive modeling

James Hampton, Allen Stedman, Caitlin Streamer

Agenda

1. **Introduction**
2. Data Analysis
3. Modeling
4. Cost-Benefit Analysis
5. Recommendations

Chicago has a history of West Nile, a potentially deadly disease spread through infected mosquitos

West Nile Virus

- Spread to humans via infected mosquitos
- 20% of those infected develop mild symptoms such as fever and vomiting
- 0.7% of those infected develop serious symptoms that may result in death

West Nile in Chicago

- First human cases appeared in 2002
- Surveillance and control program established in 2004
- Mosquito traps throughout the city are tested for the virus
- Test results influence airborne pesticide usage

There is value in accurately **predicting West Nile outbreaks** to help the city **allocate resources** for **disease prevention**

Utilized Chicago weather, testing, and spray data from 2007 - 2014 to predict West Nile outbreaks in mosquitos

Data

- NOAA weather conditions from 2007 - 2014
- GIS data from 2011 and 2013 pesticide spray efforts
- Mosquito trap West Nile surveillance test results from 2007 - 2014

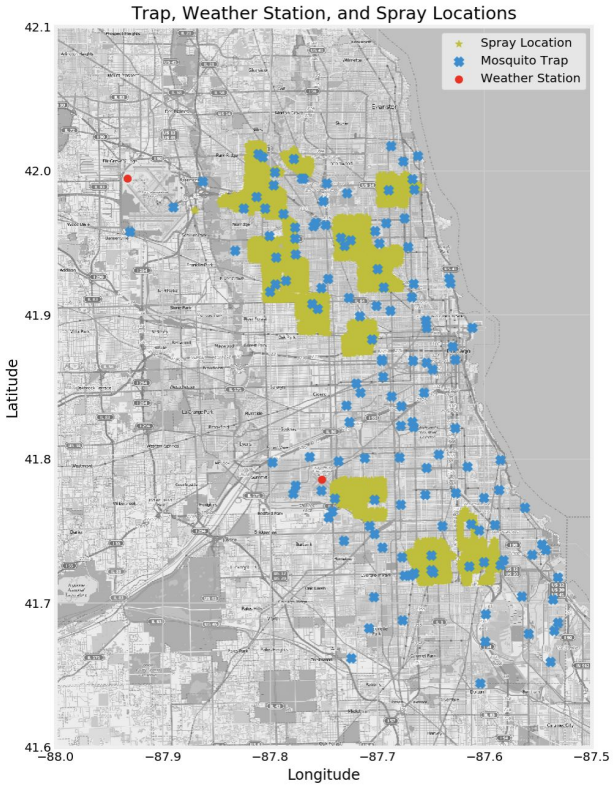
Goals

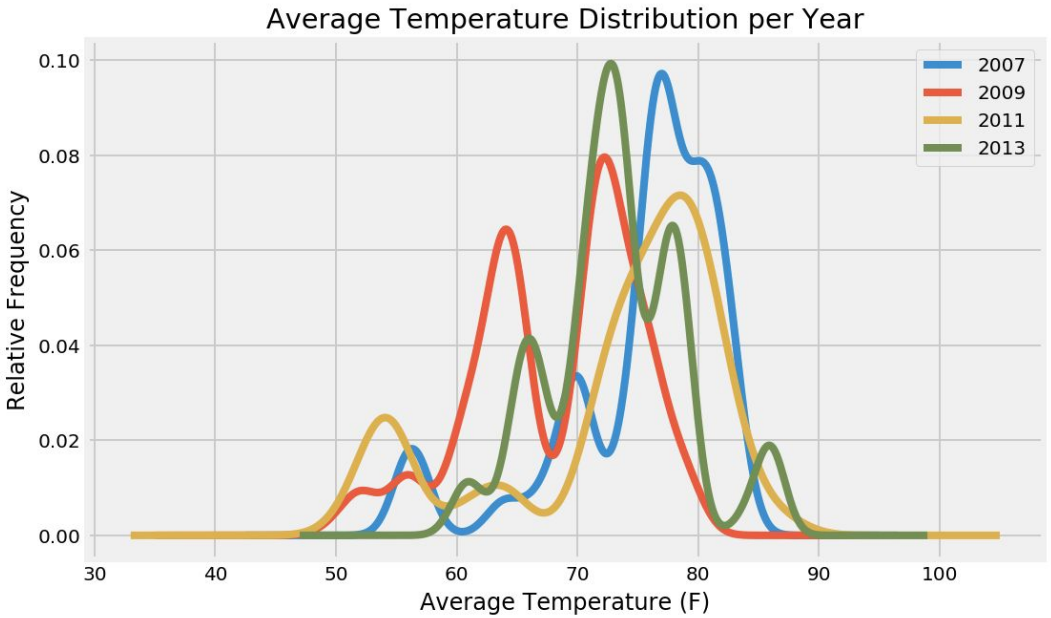
- Build machine learning model to predict presence of West Nile in mosquitoes in traps at specified locations throughout the city
- Conduct pesticide cost-benefit analysis to maximize outcomes for disease prevention

Data Analysis

1. Introduction
2. **Data Analysis**
3. Modeling
4. Cost-Benefit Analysis
5. Recommendations

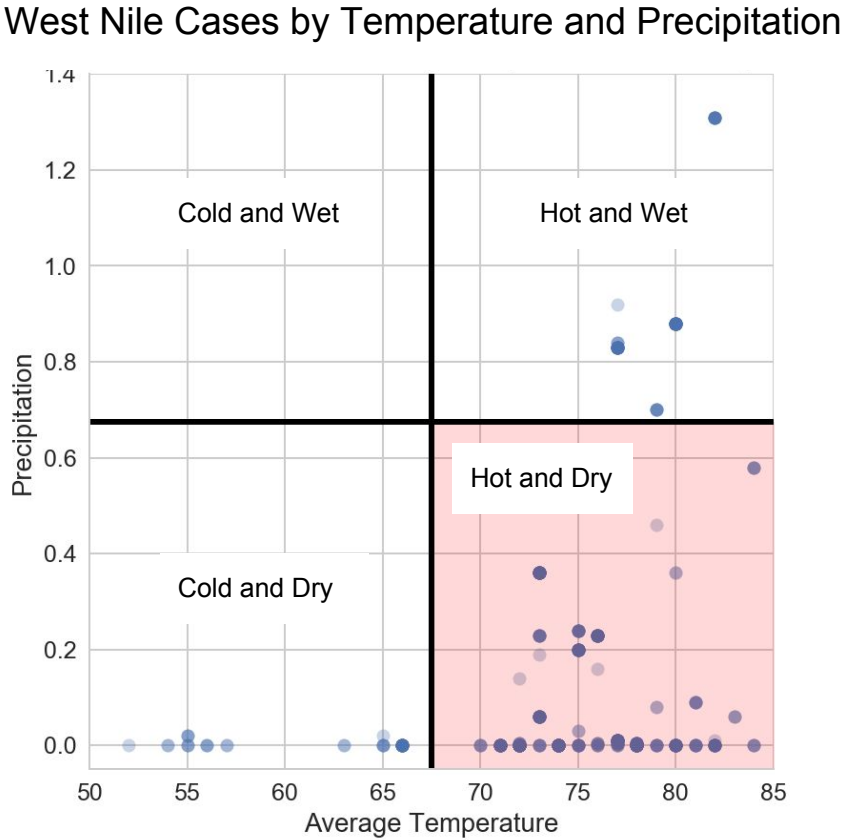
Traps are located throughout Chicago, weather stations at airports recorded detailed temperature data





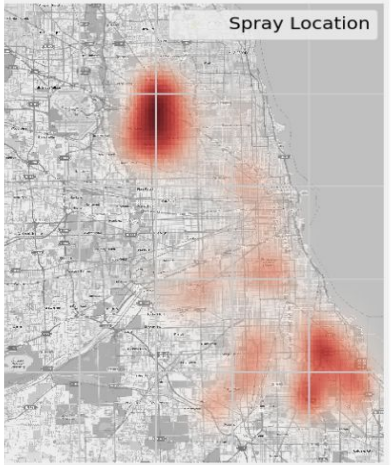
2007 and 2011 had the highest median temperature (77 F)

WNV is prevalent in hot and dry conditions

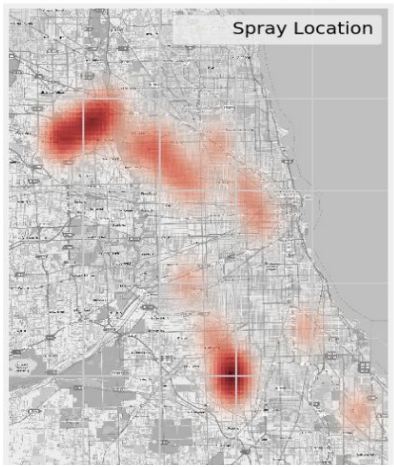


Data analysis revealed varying patterns of West Nile presence, with 2007 and 2013 having more positive tests

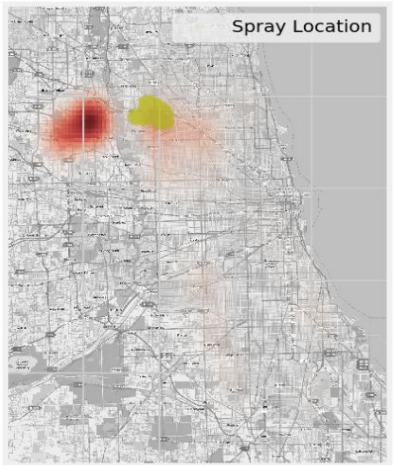
West Nile Virus in 2007



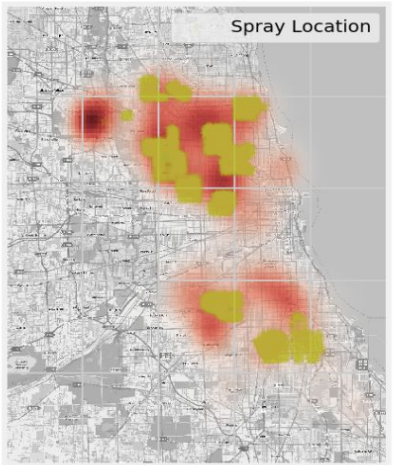
West Nile Virus in 2009



West Nile Virus in 2011



West Nile Virus in 2013



Pesticide spray efforts in 2011 and 2013 targeted some outbreak areas, but the largest areas were not addressed

Modeling

1. Introduction
2. Data Analysis
3. **Modeling**
4. Cost-Benefit Analysis
5. Recommendations

Selected Logistic Regression model to predict WNV after exploring other binary classification algorithms

Binary Classification Algorithms

- **Logistic Regression**
- Dense Neural Network
- Convolutional Neural Network
- Decision Tree
- Ensemble Tree Methods
- Support Vector Machine

Evaluation Metric

AUC-ROC

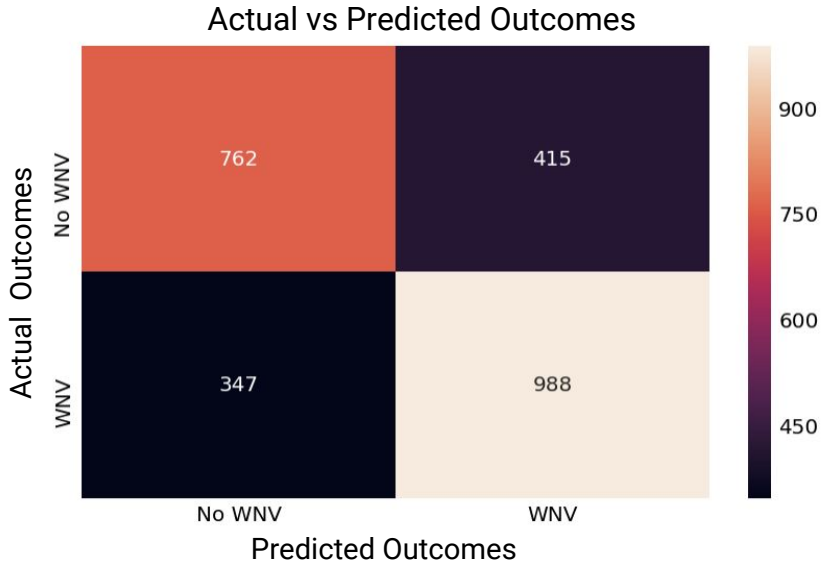
- Evaluates performance of our model
- Compares true positive rate against false positive rate
- Scores range from 0.5 to 1, with 1 being the best

*Best model was a **tuned Logistic Regression** with an **AUC-ROC score of 0.68** on testing data*

Model uses predictors that are favorable for the spread of the virus, confirming insights from data analysis

Key Takeaways:

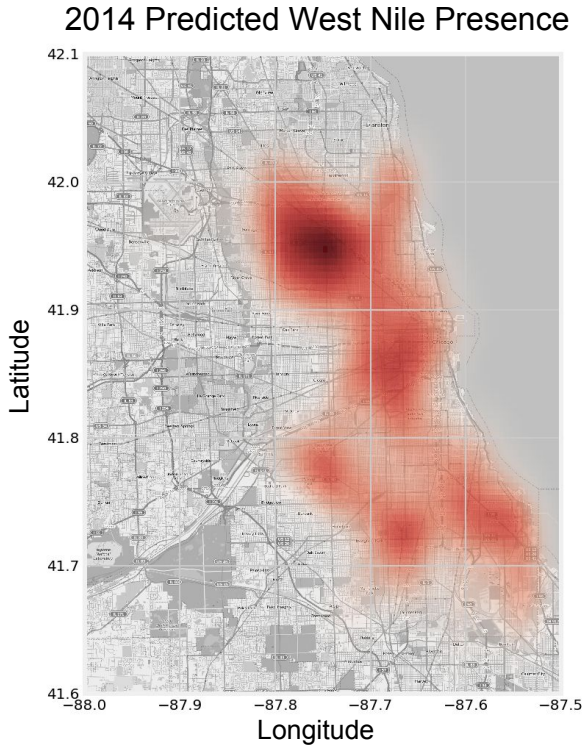
- **WetBulb temperature, Sunrise time, August, Year, and total Precipitation** are the strongest predictors of a mosquito testing positive for WNV
- Confirms assumption that **dry, hot conditions are favorable** for the spread of the virus
- **Long, hot summer months** are at the highest risk of a WNV outbreak



Cost-Benefit Analysis

1. Introduction
2. Data Analysis
3. Modeling
4. **Cost-Benefit Analysis**
5. Recommendations

Using the model's WNV predictions, we can recommend target areas for pesticide spray



Recommended Spray Area

- 120,000 acres predicted to test positive for West Nile Virus
  - Lincolnwood
  - Old Irving Park
  - North Mayfair
  - Portage Park
  - Albany Park
- Estimated spray cost for given acreage is $800,000

Costs of medical and economic loss due to WNV far outweigh costs of pesticide use

West Nile Invoice

Inpatient Costs:	$40,000
Economic Loss:	$12,000
Total:	**$52,000**

Bottom Line

Given our 2014 WNV predictions, **only 15 cases** of severe West Nile need to be **prevented** in order to make the $800,000 pesticide spray **cost effective**

We recommend to **always spray,** given the high medical costs, economic impact, and unpredictability of WNV outbreaks in people

Recommendations

1. Introduction
2. Data Analysis
3. Modeling
4. Cost-Benefit Analysis
5. **Recommendations**

Recommend spraying 120k targeted acres in Chicago for 2014 season in order to prevent West Nile outbreaks

Recommendations

- Use pesticide spray in neighborhoods predicted to test positive for WNV
  - Lincolnwood
  - Old Irving Park
  - North Mayfair
- Spraying entirety of predicted areas is cost-effective due to high costs of medical and economic loss of WNV cases in people

Next Steps

- Further improve predictive model
  - Consider additional features
  - Gather more data
  - Ensemble models
- Examine efficacy of spraying pesticides against WNV
- Built models to predict outbreaks more reliably in humans

Appendix

ROC plot and coefficients of final logistic regression model