Predicting the West Nile Virus Outbreak in Chicago

DALE WAHL, DIEGO RODRIGUEZ, MATT BOLLINGER,
NATALIE OLIVO, KYLE SANTANA

Overview

- Problem Statement
- Visualizing Data
- Our Model
- Our Proposal
- Cost-Benefit Analysis
- Recap and Next Steps

What We Know About WNV

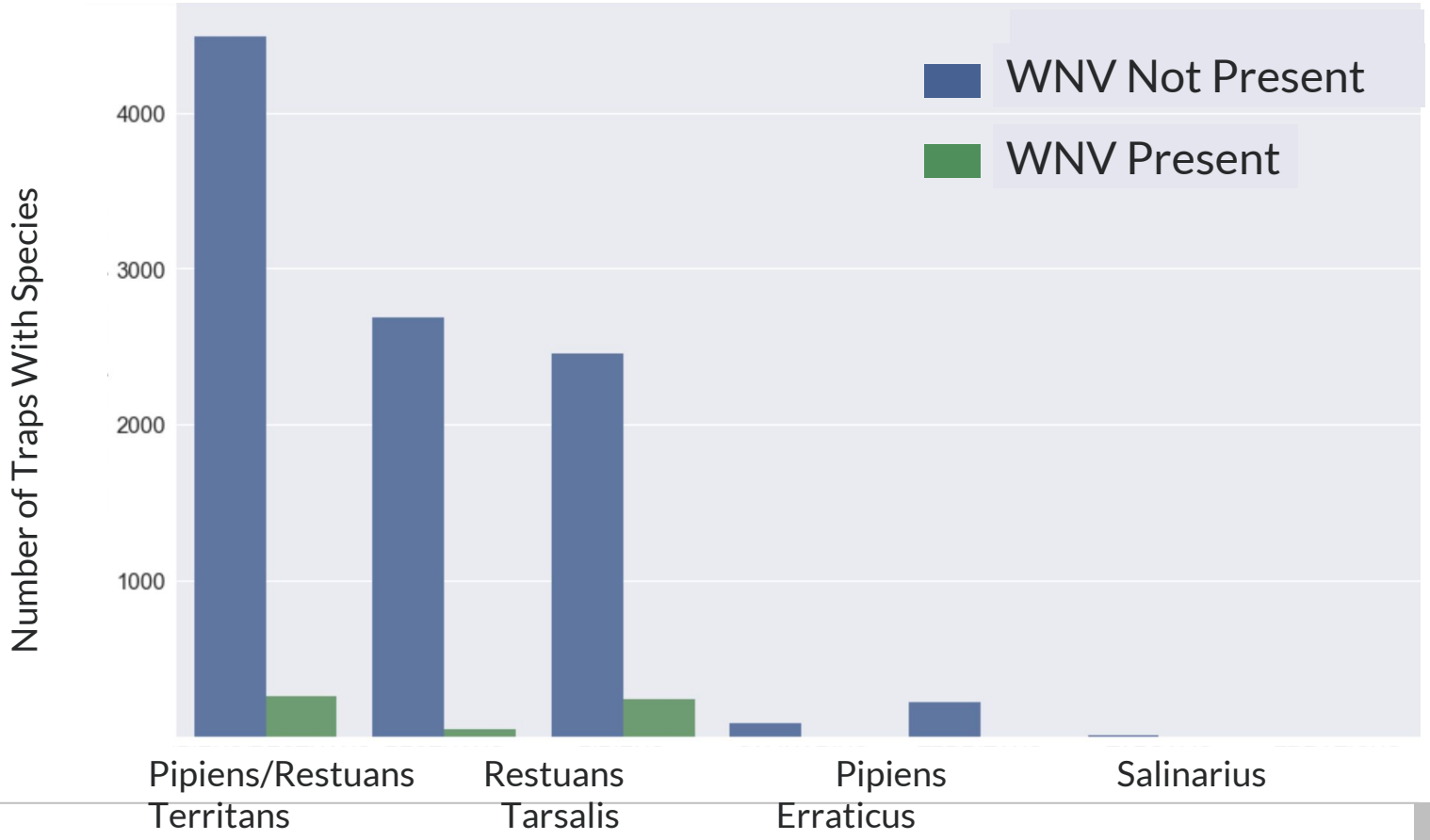
- It's most commonly spread to humans through infected mosquitos
- First cases in Chicago were reported in 2002
- By the end of 2002, Illinois had counted 884 cases and 67 deaths

Our Problem Statement

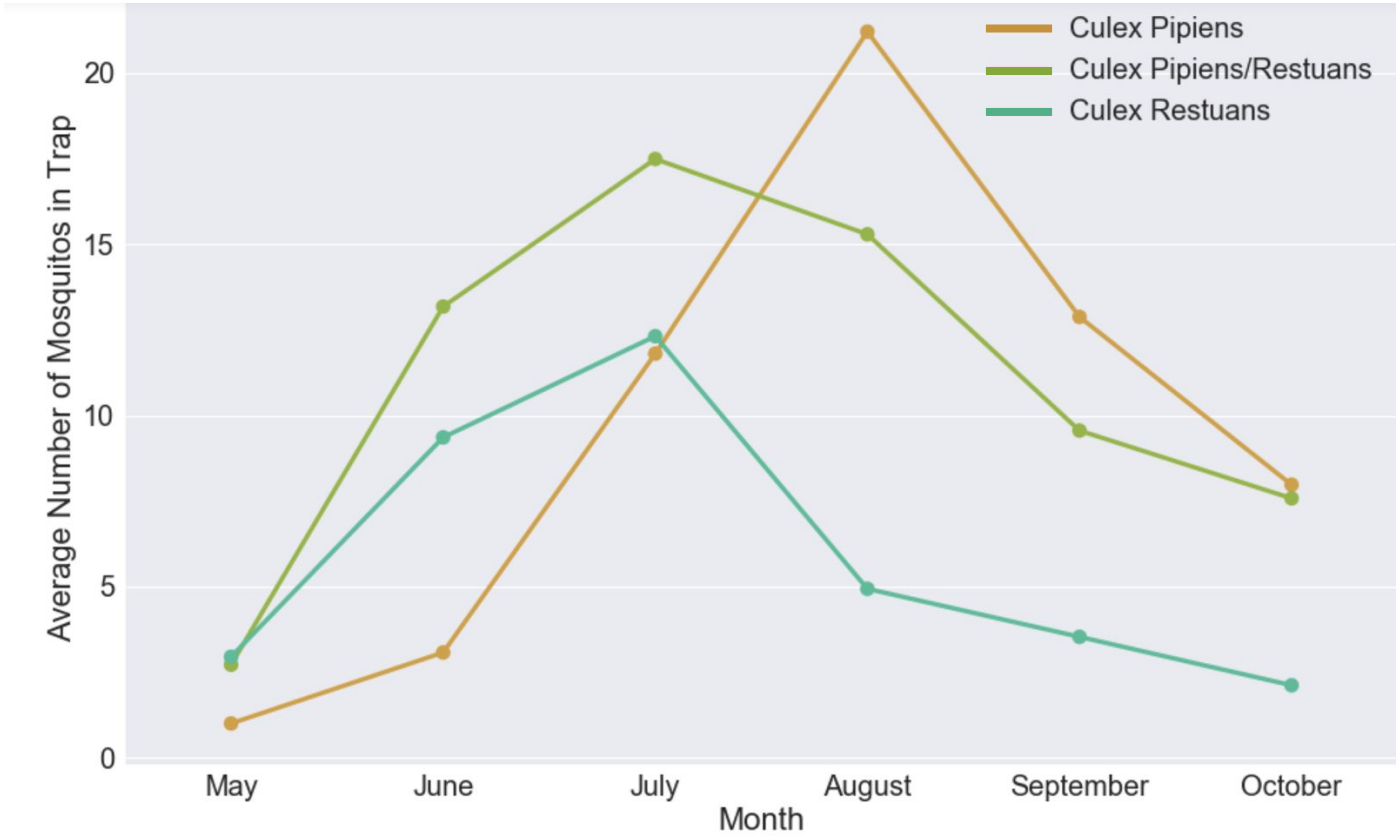
- We want to reduce West Nile Virus in Chicago
- We want to understand where, when and how WNV appears
- Pesticide is an essential tool for mosquito population control –we want to know how best to utilize it


Visualizing
The WNV
Problem

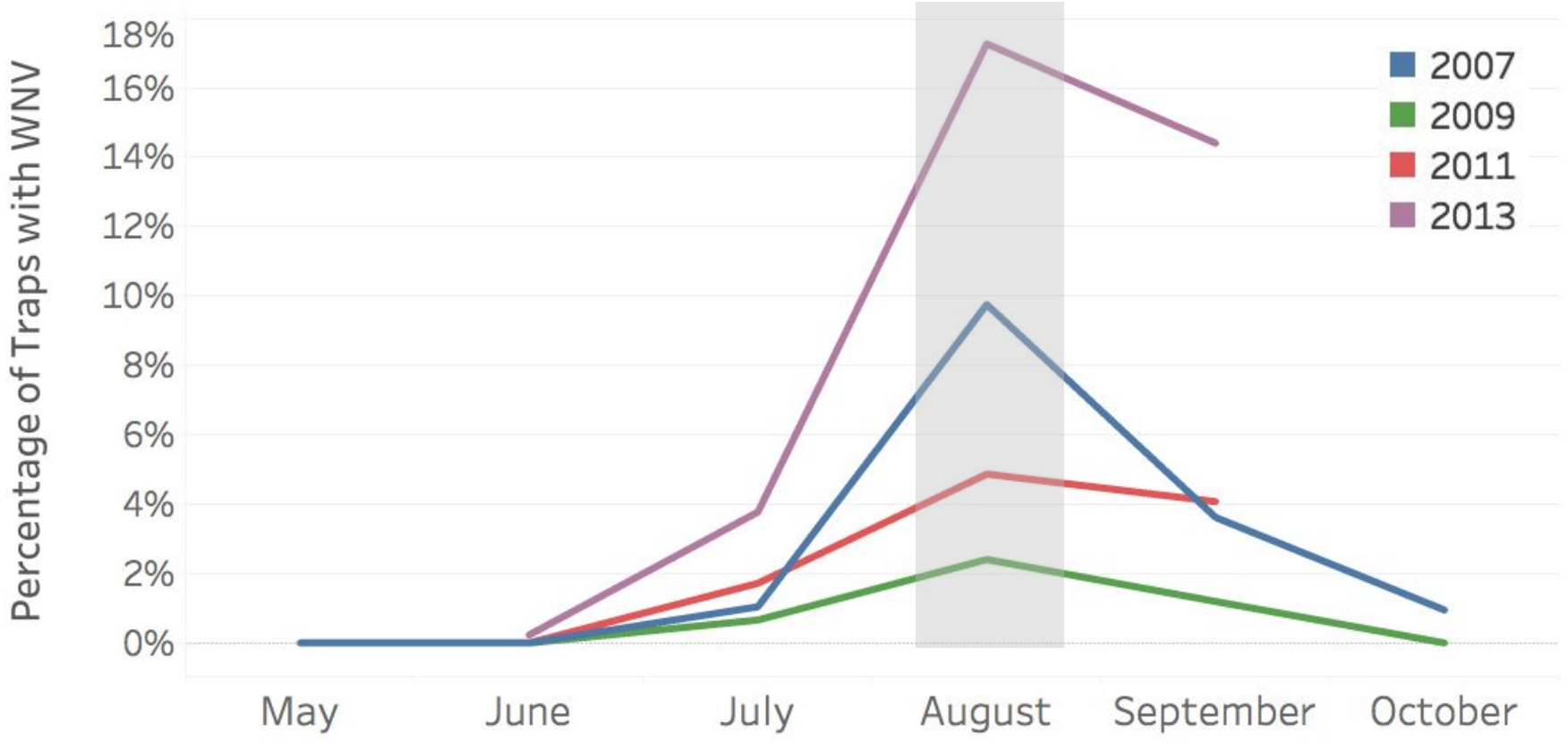
Pipiens and Restuans Carry WNV



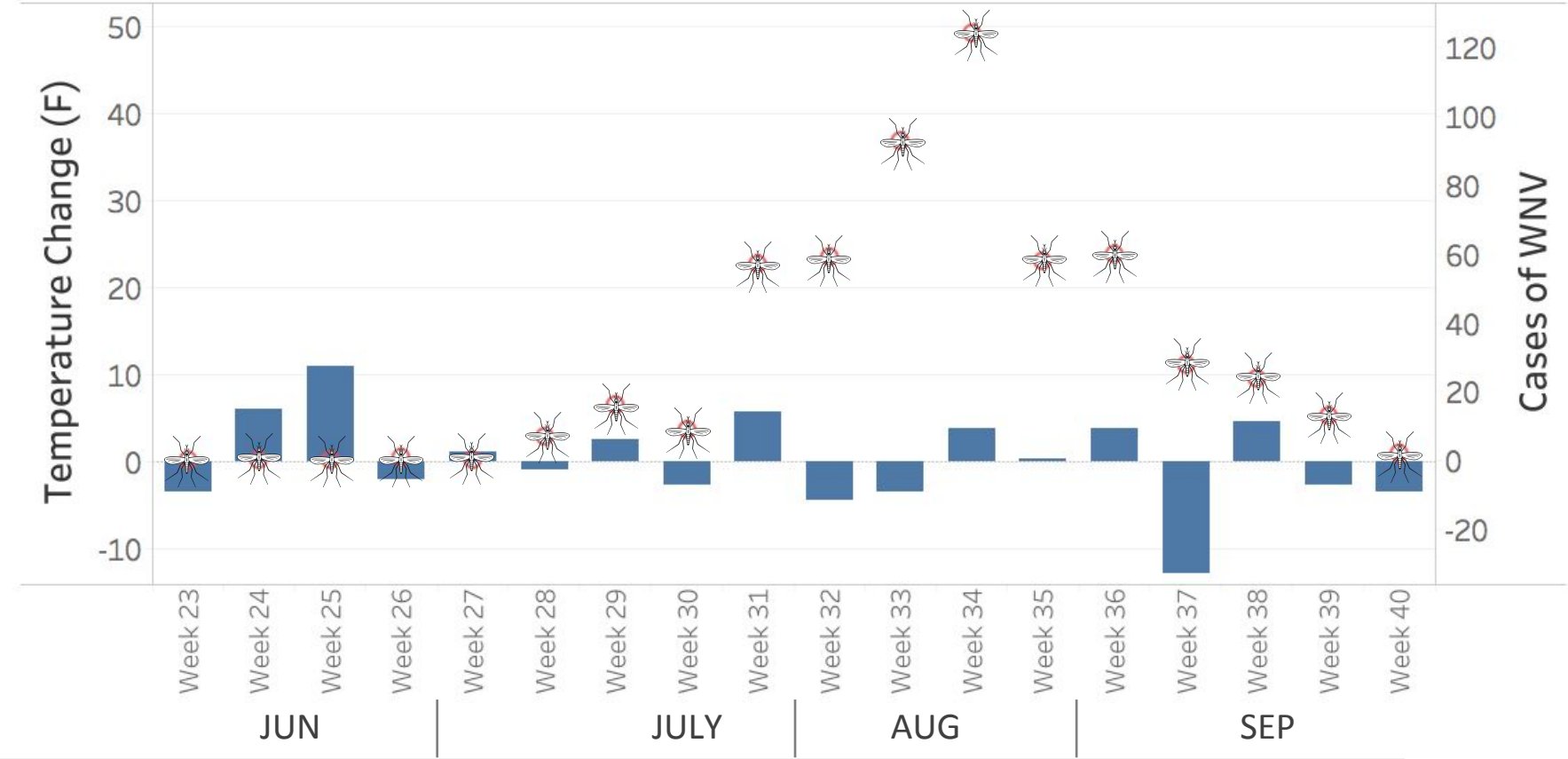
Tracking Mosquitoes Over Months



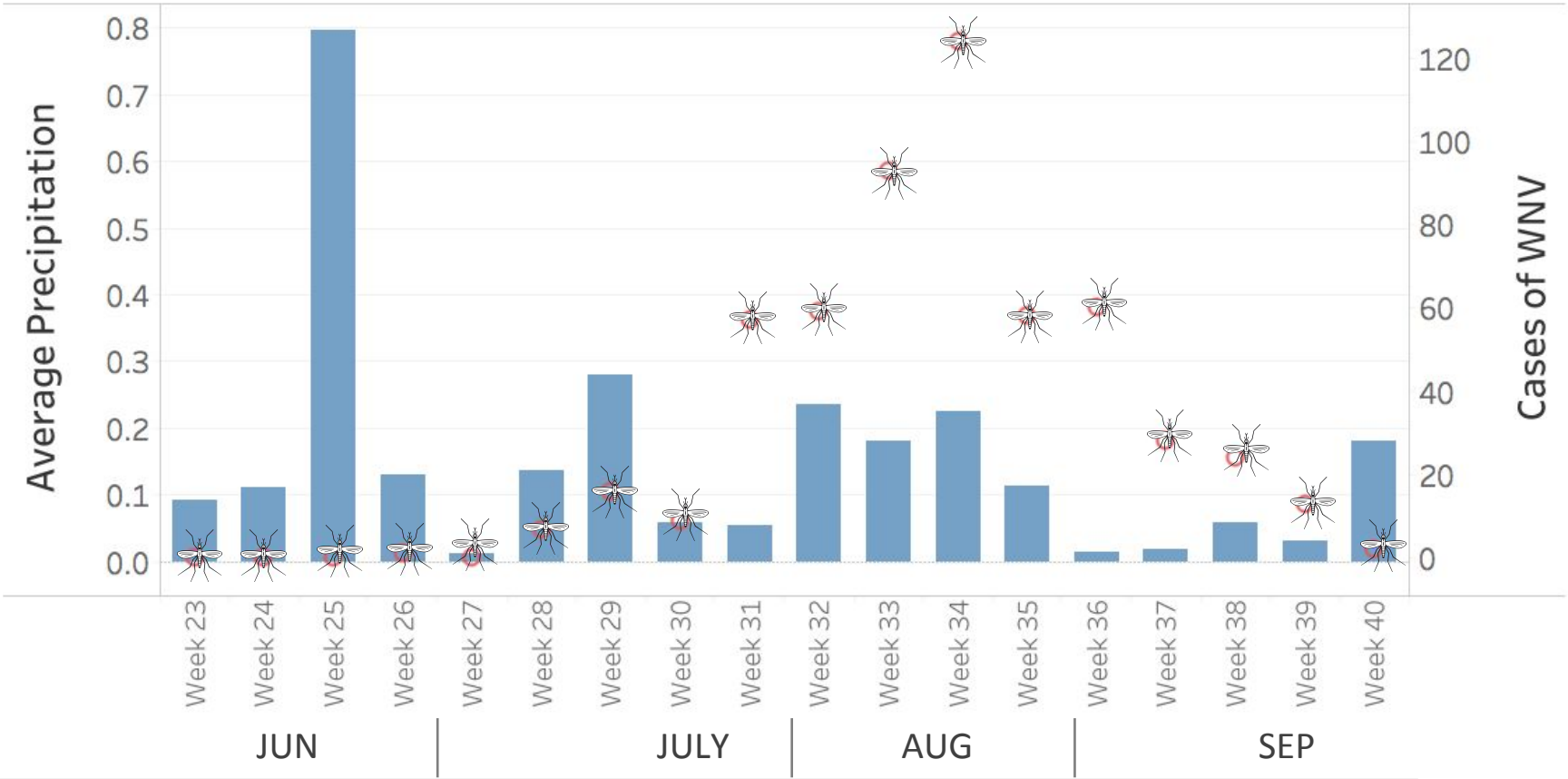
WNV Cases Peak in August



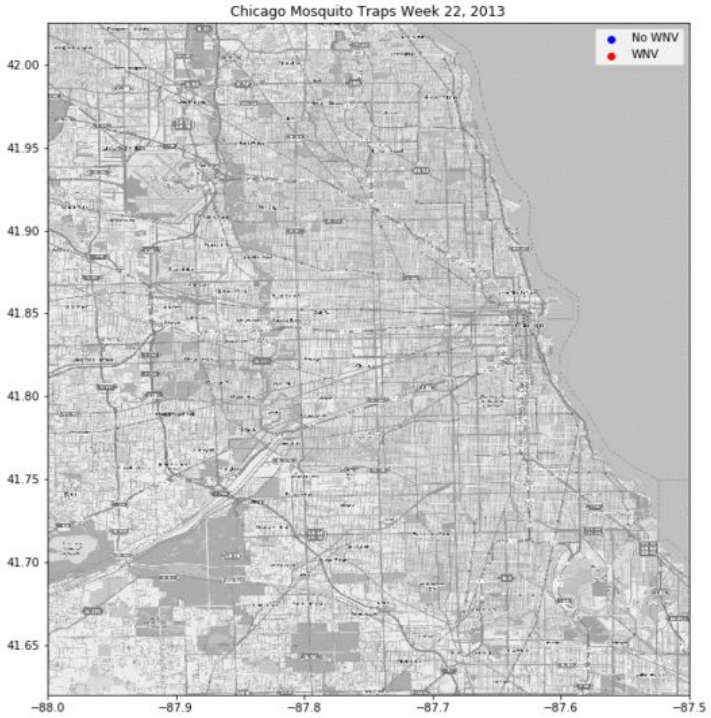
WNV Cases Rise After Elevated Temperatures



WNV Cases Rise After Elevated Rainfall

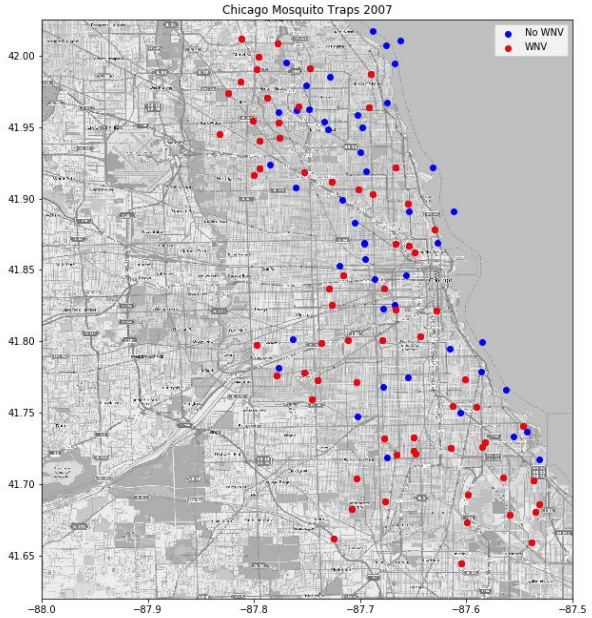


WNV Spreads From Suburbs To City Center

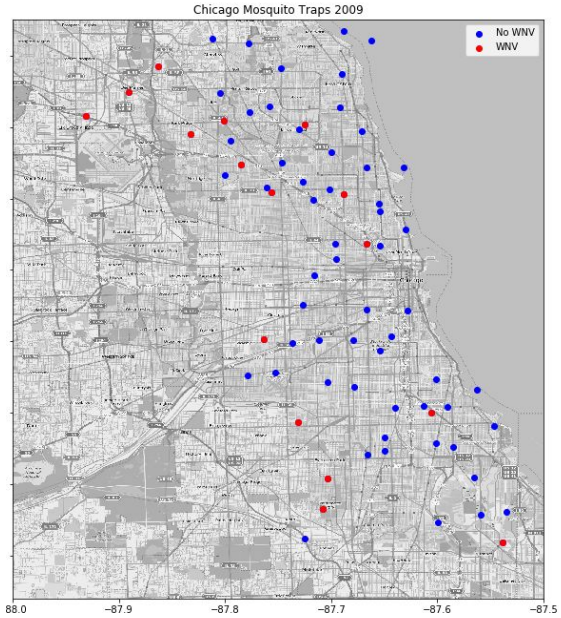


WNV Presence Over Time

2007



2009



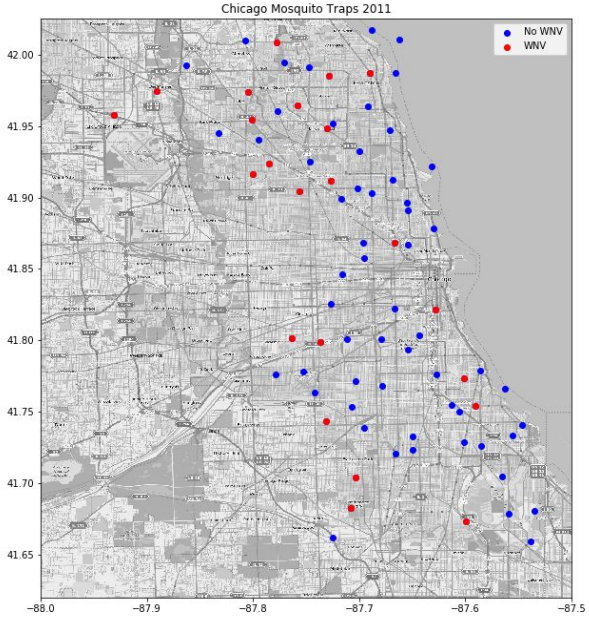
OCCURRENCE

2007: More cases of WNV than any other year

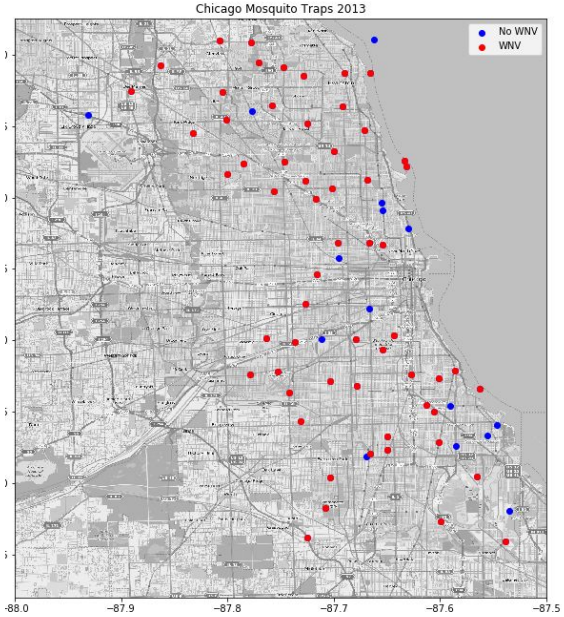
2009: WNV reserved to the suburbs

WNV Presence Over Time

2011



2013

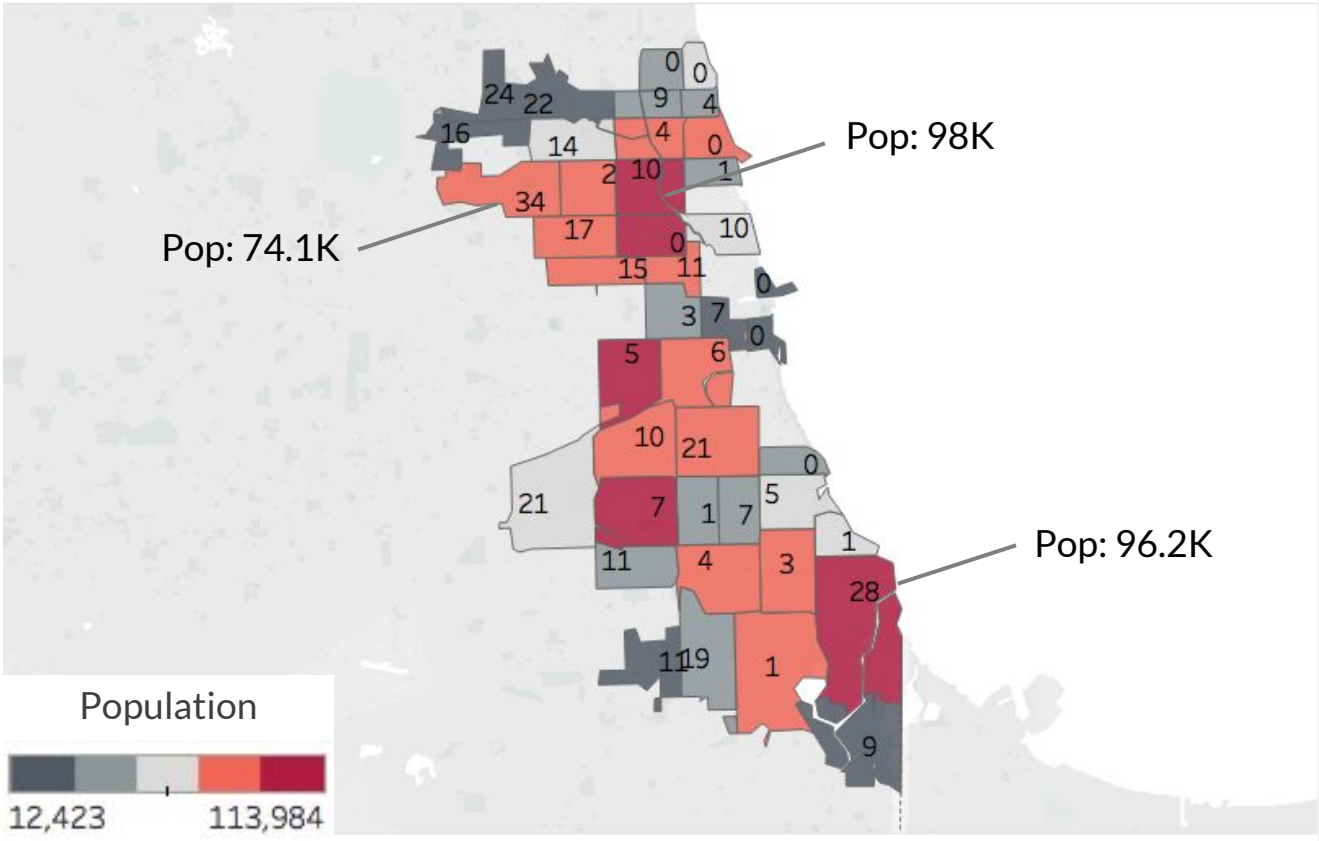


OCCURRENCE

2011: Limited WNV cases

2013: Severe WNV presence

WNV Cases per Zip Code



OCCURRENCE

Several high-population zip codes have reported high WNV cases; highlighted are three zips


Our
Model

Feature Selection

- Utilized all features provided (Weather, Spray, Geo)
- Preliminary feature selection
  - Geography - Latitude and Longitude
  - Weather - Precipitation, Temperature, Dew-point, Wind speeds
- Spray data
  - Affected total mosquitos, but poor indicator of WNV
  - Missing for test years

Feature Engineering

- Weather improvements
  - Temperature (3 month average, 4 week lowest)
  - Dew-point (4 week average)
  - Precipitation (3 month total)
- Geospatial data
  - Distances from prior epicenters of outbreak added precision to model

Model Adjustment

Voting Classifier

Low False Positives, but higher False Negatives

XGBoost

82.5% Sensitivity (avoiding many more False Negatives)


Our
Proposal

Our Spray Recommendations

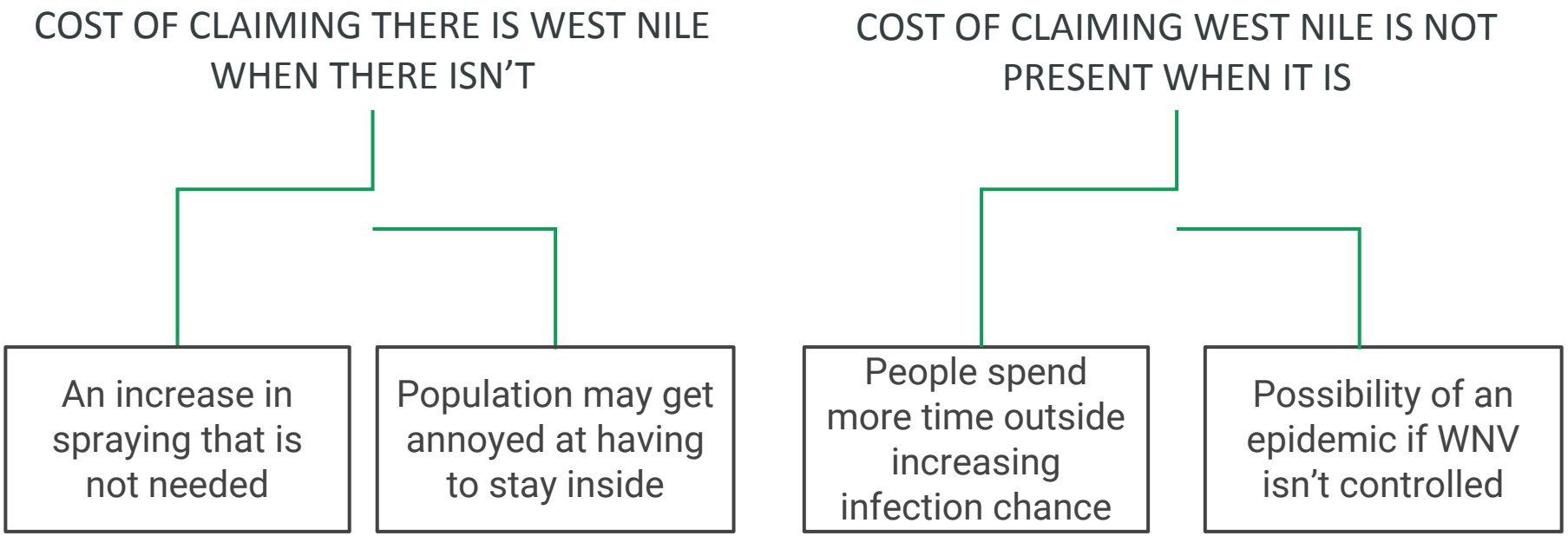
- Based on our model, we can predict when and where WNV occurs
- Spray in areas with WNV based on the following criteria:
  - Traps that showed repeated cases of WNV counts
  - Traps with a high percentage of WNV per mosquito count
  - Areas with high human population

Our Model Performance

Error Matrix

		Our Model Predictions: WNV Present	Our Model Predictions: WNV Not Present
Actual West Nile Virus	WNV Present	82.5% ✓	17.4% ✗ ✗
	WNV Not Present	28.8% ✗	71.2% ✓

Consequences of Model Error



Cost-Benefit Analysis

Spraying in Chicago	Not Spraying in Chicago
$1.1 M to spray the city	**$1.9 M a year in economic damages**
Spraying reduces WNV by ⅔ (In 2016 an estimated 180 cases were prevented through spraying)	**Each WNV infection costs the city $21 K per person**

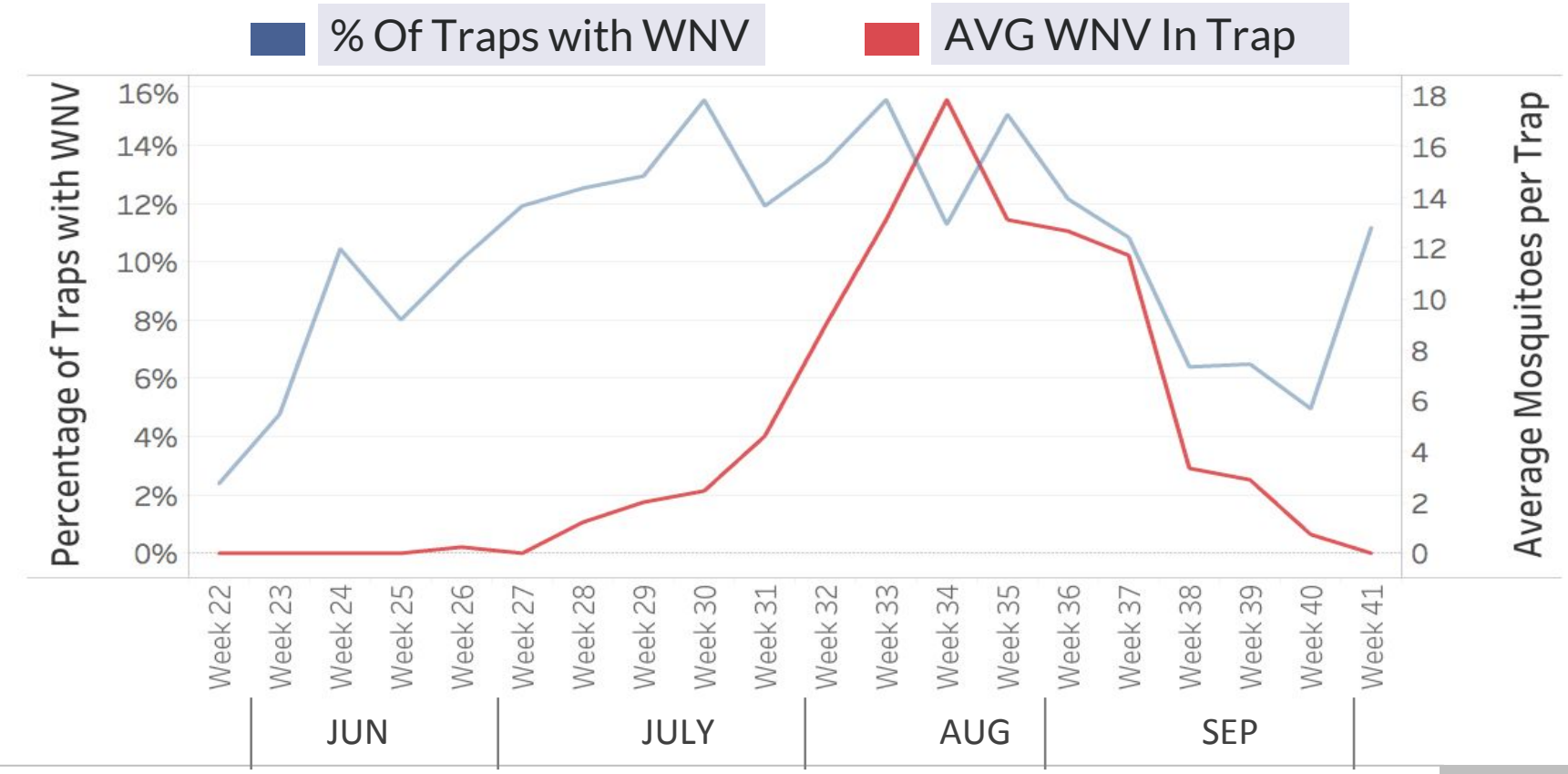
Recap

- West Nile Virus is highly seasonal, becoming most prominent in July and August
- Breakouts tend to originate in suburbs and spread to the center of the city
- West Nile outbreaks are more severe in summers with heavy precipitation
  - **Measures of humidity, such as dew point, also predict West Nile presence**

Areas of Note

- Areas with high percentage WNV per Mosquito:
  - 4600 N Milwaukee Ave - 14 cases, 1% of mosquitos
  - 6300 W 64TH St - 6 cases, 1.1% mosquitos
  - 1300 N Laramie Ave - 7 cases, 1% mosquitos
- Discontinued Traps (2007 only year):
  - 6000 N Avondale Ave - 3.7%
  - 6500 N Oak Park Ave - 2.6%
  - 7900 W Foster Ave - 2.4%

Our Window of Opportunity



Next Steps

- We recommend a mosquito vector control on July 14 and August 11th to hit mosquito populations before they grow
- Tracking mosquito population growth and incorporating it into our model will greatly improve precision
- Evaluate alternative tactics, such as clearing dead birds
- Investigate the effectiveness of spraying insecticide with complete spray data

QUESTIONS?

THANK YOU

Dale, Diego, Matt, Natalie, & Kyle


Additional
Insight and
Info

A Closer Look: Data Points

Original

- Longitude and Latitude
- Temperature Min, Max, and Average
- Average Wind Speed
- Total Precipitation
- Species
- Dew Point
- Sea Level (Station 1 and 2)
- Snowfall
- Wet Bulb
- Station Pressure

- Engineered
- Distance from Chicago O'Hare Airport (miles)
- Average Precipitation over 90 Days
- Average Precipitation over Four Weeks
- Average Temperature Over 90 Days
- Average Temperature Over Four Weeks
- Minimum Temperature Over Four Weeks
- Average Dew Point Over Four Weeks
- Average Precipitation over Four Weeks
- Month
- Species: Pipiens and Restuans (categorized)

Vector Control

- Source Reduction: the removal of mosquito breeding habitats
- Habitat Modification: manipulate habitat to reduce breeding
- Bio-control: introducing natural predators of mosquitoes
- Larvicide: using pesticides to reduce larval populations
- Adulticide: using pesticides to reduce adult populations