ISSN: 2473-4901
v9 n5826

West Nile Virus in Colorado: Analytic and Geospatial Models of the Virus in Colorado

Johnny Snyder
josnyder@coloradomesa.edu
Davis School of Business
Colorado Mesa University
Grand Junction, Colorado 81506

Abstract

West Nile Virus found its way to North America in 1999, starting with the diagnosis of two cases of encephalitis in the Queens borough of New York City. WNV had found its way to Colorado by 2002. The main vector for West Nile Virus is the mosquito, primarily the *Culex* species. This research shows, from historical data along with qualitative, quantitative, and geospatial methods, that the primary variables behind West Nile Virus cases by county in Colorado are the county's urban/rural classification, water area in the county (in square miles), and if it is an El Niño year or not. Other variables, including population density in the county, and the average precipitation and temperature over the period July to October, are discussed and their merit in a model presented. Mapping tools are used to illustrate the presence of West Nile Virus as well as its spread, over time, through the counties in Colorado. The data set for this study covers 2005 to 2021 for the 64 counties of Colorado.

Keywords: West Nile Virus, analytic model, virus model, geospatial analysis

1. INTRODUCTION

West Nile Virus (WNV), a mosquito transmitted virus, was first identified in 1937 in the West Nile province of Uganda (Kramer, Li, & Shi, 2007). WNV emerged into North America, and in particular, New York in 1999 and has since spread throughout the USA, Canada, and Mexico (Petersen, 2019; Kramer, Li, & Shi, 2007). Colorado recorded its first case of WNV in 2002 (CDPHE, 2022) and saw a large number of cases in 2003 due, it is believed, to the initial large-scale testing for WNV, and the lack of immunity in the population (Marzec, 2022). Starting in 2005, the pattern of cases of WNV stabilized, to include high and low years, but nothing as severe as the initial highs of 2002 and 2003.

WNV can infect birds, which also serve as a reservoir and diffuser for the virus, horses, and humans, with the latter two being the most negatively affected by the virus (Peterson, 2019; Kramer, Li, & Shi, 2007). WNV is spread, primarily, by the enzootic cycle between birds and mosquitoes. There are over three hundred species of birds that have been identified as viable WNV hosts and generally the mosquito genus *Culex* is the transmission vector to humans (Ciota, 2017). Most humans show no symptoms upon being infected with WNV (80%), while approximately 20% exhibit flu-like symptoms, and less than 1% become seriously ill with arboviral encephalitis (Ciota, 2017; Wildlife Futures Team, 2021). In Colorado, there have been over 5800 cases of WNV identified in humans since WNV arrived to the state in 2002 (CDPHE, 2022). Extrapolating, this implies that there could have been over 23,000 undetected cases with up to 290 cases of encephalitis, making this a public health issue involving hospital stays, transplantation medicine, and transfusion concerns (Petersen, 2019; Ronca, Ruff, & Murray, 2021). While there are vaccines against WNV for horses, first approved in 2005, there are no human vaccines that have gone beyond Phase I/II trials, indicating a need for

 ISSN: 2473-4901
v9 n5826

other mitigation strategies for the virus and its vectors (Ronca, Ruff, & Murray, 2021).

Prevention of WNV through target mitigation strategies can assist the state of Colorado in preserving the health of the population, as well as avoidance of medical costs and lost productivity costs which can run into the tens of millions of dollars (Ronca, Ruff, & Murray, 2021). This research attempts to build a model that identifies when and where WNV is most likely to occur in the state. This model, along with appropriate visualizations, can assist decision makers in the public health arena in mitigation strategies for WNV in Colorado.

Observations and studies from Chicago to Texas have suggested that "enhancement of surveillance and vector control in limited geographic areas could produce an outsized impact on WNV incidence nationwide" (Petersen, 2019, p. 1457); further, Hadfield et al. (2019) state that a coordinated effort between state health departments could be used as vector control by "strategically focusing resources at a precise time and location to limit potential outbreaks" (p.10). This research addresses the "when and where" WNV might appear so that proactive mitigation strategies can be deployed.

The main variables for WNV models come under three general headings:

- Environmental (precipitation, humidity, and temperature as examples)
- Physical (elevation, water area, and land use as examples)
- Population (human, mosquito, and bird census data as examples).

Hadfield, et al. (2019) states "Dynamic extrinsic factors, such as rainfall and temperature, that influence mosquito and bird populations can be predictive of WNV intensity, yet the contributions of these extrinsic factors vary across the United States due to differences in regional ecology." (p. 7) This statement leads one to conclude that more localized models containing extrinsic factors are necessary to predict WNV occurrences. This is also the conclusion reached by Paz (2015) who states that although rainfall has a pattern of positive association with outbreaks of WNV, the literature shows mixed results, and that the response (to rainfall) "might change over large geographical regions, depending on differences in the ecology of mosquito vectors" (p. 2). Garcia-Carrasco et al. (2021) found that temperature and elevation were significant factors in the occurrence of WNV. They also suggest that mountains "could be barriers or at least filters for the spread of the disease" (p. 6). This implies that elevation could have an effect on WNV occurrences.

Other authors (Barker, 2019; Ronca et al., 2021) have worked with surveillance data (of bird and mosquito populations) in an attempt to predict outbreaks of WNV. However, as Barker (2019) states "Human and veterinary disease cases are also tracked through surveillance systems for notifiable diseases, although notifications generally arrive too late to initiate mosquito control, sometimes weeks or even months after the infection has occurred" (p. 1511).

Variables that have the potential to affect the transmission of WNV in Colorado include:

- Temperature (Ciota, 2017; Hadfield, et al., 2019; Paz, 2015; Garcia-Carrasco, et al. 2021)
- Precipitation (Ciota, 2017; Hadfield, et al., 2019; Paz, 2015)
- Humidity (Paz, 2015)
- Seasonality (Paz, 2015)
- Wind (Paz, 2015)
- Urban/rural (Ciota, 2017)
- Land use (Hadfield, et al., 2019; Paz, 2015; Garcia-Carrasco, et al. 2021)
- Presence of water (Garcia-Carrasco, et al. 2021)
- Elevation (Hadfield, et al., 2019; Garcia-Carrasco, et al. 2021)
- Mosquito abundance data (Ronca, et al., 2021)
- Bird population data (Ronca, et al., 2021)

Other variables added to this data set for their perceived usefulness for multiple types of analysis as well as possible filters for visualizations include:

- Year
- County name
- County population
- Number of WNV cases per county
- El Niño, La Niña years
- Land area of county
- Severity of WNV in county, a qualitative variable constructed from historical data on the count of WNV cases per county

 ISSN: 2473-4901
v9 n5826

(primarily as a filter for geospatial analysis)

Variables which could be found are illustrated in Table 1, along with references or links to where to find the original data source.

Table 1: Variables for the Colorado WNV Models (see Appendix)

Some variables were difficult or otherwise unobtainable for this study. These include humidity, mosquito, bird, and wind data, all at a county level. Other variables that proved difficult to find include the number of irrigated acres per county (only found 2007 and 2012 data sets) and a reasonable elevation variable.

2. PROBLEM STATEMENT

This research seeks to determine what variables influence WNV presence in the counties of Colorado. By applying techniques such as:

- descriptive statistics
- correlation analysis
- hypothesis testing
- multiple linear regression analysis
- logistic regression analysis
- cluster analysis
- geospatial analysis
- graph analysis

salient variables will be selected for modeling purposes and mapping techniques will be used as an illustrative guide for deploying resources at a county level. This analysis will aid the decision maker (hospital, EMS, medical community, county/state level medical personnel) in when and where to apply mitigation measures for WNV, when and where to deploy limited personnel across the state of Colorado, where to locate sentinel devices for WNV, and when to notify public health decision makers of WNV activity.

3. METHODOLOGY AND RESULTS

Performing both multiple linear regression ($\hat{y}$= number of WNV cases in county) and logistic regression ($\hat{p}$= probability of WNV in county) with all possible variables, the best models, and their R^2 values are given in Table 2. These models were distilled using the backward elimination method, which uses all variables to construct a model, then removes the least significant variables, iteratively, until a model comprised solely of significant variables is left. (Camm, et al., 2024)

Table 2: Best models for WNV (see Appendix)

The first finding is that elevation played no significant part in the analysis, even though six different elevation variables (maximum, minimum, difference, qualitative over 7000 ft., elevation at center point of county, and high qualitative (maximum over 9000 ft.)) were tried in the models. Elevation has been shown to be significant in other models (Garcia-Carrasco, et al., 2021).

Many other variables were also not significant predictors of WNV in this county level model. They include La Nina, area of county, and an annual precipitation variable.

Evaluating the explanatory power of the models in Table 2 leads to the construction of Figure 1, a Venn diagram of the statistically significant variables from $\hat{y}$ and $\hat{p}$. Figure 1 shows the core variables of the model (the intersection) and the peripheral variables of the individual models. These variables will be used for the remainder of this work.

Figure 1: Significant variables for WNV models (see Appendix)

The variables identified in Figure 1 are defined here:

Core Variables (significant for both models)

- Urban/rural – a qualitative variable, 1 if the population of the county is greater than 150,000 people. This variable will be used to compare the two data sets for differences (urban counties versus rural counties) and as a filter for visualizations.
- Water area – a quantitative variable for the area (in mi^2) of each county that is covered by water.
- El Niño – a qualitative variable, 1 if it is an El Niño year (which occurred in 2005, 2007, 2010, 2015, 2016, and 2019). This variable will be used to compare the two data sets for differences (El Niño year versus non-El Niño years) and as a filter for visualizations.

ISSN: 2473-4901
v9 n5826

Peripheral Variables (significant for only one model)

- Population density – number of people per square mile. With any disease of the human, population density can play a factor.
- AvgTempJulyOct and AvgRainfallJulyOct – as named, these variables measure the average temperature in Fahrenheit and the average rainfall in inches over the four-month period of July to October. This is the period in which 80-90% of the WNV cases occur as can be seen in Figure 2 from the CDC. Due to the short life cycle of the mosquito (generally a couple of weeks) looking at the environmental variables during this time period seems reasonable.

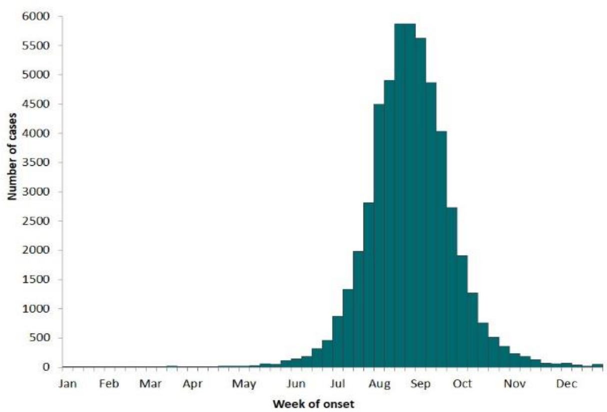
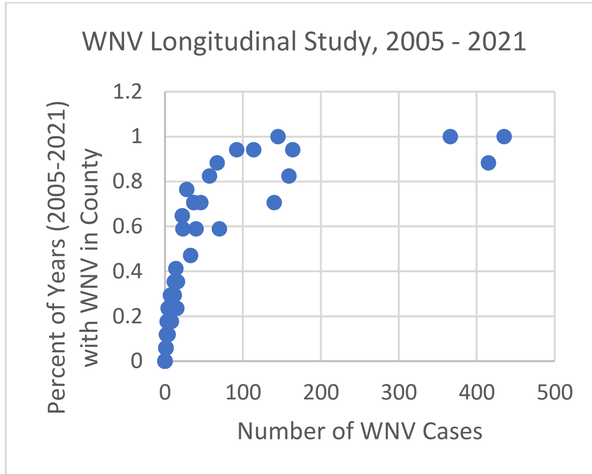


Figure 2: WNV in the USA during the year (CDC, 2021)

Time Period

- WNV arrived to New York State in 1999, then spread throughout the country. WNV made it to Colorado in 2002, and being a new virus there was a lot of testing and no natural immunity (Marzec, 2022), causing a large spike in WNV numbers for 2003 and 2004. In 2005, the pattern of WNV settled down, exhibiting peaks and valleys, but not as severe as in 2003 - 2004. This model uses the WNV case count, by Colorado county, for the time frame 2005 – 2021 assuming these years to be more representative of the WNV numbers in the state moving forward. (of course 2020 and 2021 might be skewed due to the presence of Covid in the state)

Analyzing the counties longitudinally for the percentage of years the county has had WNV cases present over the study period, Graph 1 is obtained.



Graph 1: Percent of years county has WNV, 2005-2021

Graph 1 points at a couple of interesting items, three counties have WNV every year (Denver, Boulder, Larimer) and three counties have a lot of WNV (Boulder, Larimar, Weld). Graph 1 also facilitates the development of a qualitative variable for a county's WNV status.

$$CountyLongitudinalWNVStatus = \left\{ \begin{array}{c} HighRiskCounty\ if\ Percent \geq 0.8 \\ MediumRiskCounty\ if\ 0.5 \leq Percent < 0.8 \\ LowRiskCounty\ if\ Percent < 0.5 \end{array} \right\}$$

It can be calculated that the HighRiskCounty group, when summed, represents 76.3% of the cases, leaving only 23.7% of the cases in the medium and low risk categories. This variable will be used in the visualizations as well as for separating the data set into multiple pieces for some descriptive analysis and interpretive guidance on differences between the risk groups. This analysis is presented in Table 3.

Table 3: Statistically significant differences when dividing the data set along qualitative lines (high, medium, low and El Niño dimensions) (see Appendix)

Overall, Table 3 tells us that there will be more WNV cases in El Niño years with warmer temperatures, counties with higher population,

ISSN: 2473-4901
v9 n5826

and lower elevation counties with more surface water. The data also showed more WNV cases in years with less rainfall, an unexpected result, as mosquitoes need water to reproduce.

Cluster analysis (k = 3 clusters) also indicated that counties with high WNV cases clustered around:

- counties with more water area
- counties that have denser population
- counties with a higher AvgTempJulyOct
- counties lower in elevation,

which helped to identify salient variables and well-designed filters with which to build visualizations to illustrate these primarily analytical results.

4. VISUALIZATIONS

Visualization 1 is a Tableau dashboard tracking WNV cases against four of the independent variables of the model, water area, temperature, precipitation, and population. Other variables, namely urban/rural and the CountyLongitudinalWNVStatus are used as filters for the dashboard.

Visualization 1: A Dashboard for Tracking WNV Variables (see Appendix)

Visualization 1 gives us some information not only about the status of the WNV model, but information on future directions as well. The Colorado WNV "Goldilocks Zone" for mosquitoes has been addressed in news articles (Bailey, 2022) where it is stated "Spring rain, summer drought, and heat created ideal conditions for mosquitoes to spread the West Nile Virus through Colorado last year" (p. 1). Visualization 1 yields information on the Goldilocks Zone for "High" WNV counties (this is the filter applied in Visualization 1) as:

- AvgTempJulyOct: 57 – 67 degrees Fahrenheit
- AvgRainfallJulyOct 5.5 – 7.0 inches
- Urban/Rural – 80% of the "high WNV" counties are classified as urban
- County water area in mi^2 – positive linear relationship between water area and WNV

The Goldilocks Zone might, according to the article by Bailey, also need a spring rain variable or a drought variable to identify this zone more accurately.

Visualization 1 also indicates that Denver County (formally the City and County of Denver) might be an outlier and could be removed in future analysis. It would also be prudent to address the other city/county designated area, City and County of Broomfield, for potential removal as well. These large urban areas could also be importing the virus from camping or other "out of county" activities by the county residents.

Visualization 1 also indicates that the number of WNV cases increase as the amount of surface water increases in the county (slope = 0.6 cases/mi^2 of surface water in county for counties with a "high" WNV status). This indicates that minimization of water area (particularly in urban counties) could lead to a lessening of the WNV impact in urban areas. This would also align with the West's increasing water conservation issues – reduction of water features in golf courses, urban parks, and other (arguably) unnecessary water features in the urban landscape.

While Visualization 1 focused on "High" WNV counties (counties that have had WNV present for more than 80% of the years in this study), Visualization 2 explores, geospatially, the impact of urban (greater than 150,000 population) counties in Colorado. It should be noted that over the time span of this study, two counties, Mesa, and Pueblo, changed from rural to urban (Mesa eclipsed 150,000 in 2015 and Pueblo in 2006), noting the dynamic nature of studies of this type.

Visualization 2: Geospatial analysis of the WNV load in Colorado (see Appendix)

Visualization 2 indicates that urban counties, which follow the I-25 and I-70 corridors (except in the mountains), and that WNV also follows the urban corridors particularly in the counties that have more surface water. Additionally, the population density choropleth verifies that Denver County could be a candidate for removal from this study. These maps could be utilized to effectively direct mosquito and WNV mitigation efforts.

5. FINDINGS

WNV follows the urban corridor throughout Colorado, occurring more in urban counties with large surface areas of water. WNV mosquitoes enjoy certain temperature ranges and

ISSN: 2473-4901
v9 n5826

precipitation conditions. WNV occurs during the July – October time frame and occurs more in an El Niño year. Counties can be classified as high, medium, or low depending on their WNV load, directing mitigation efforts to areas most in need. It appears that any major WNV events would be in one of Boulder, Larimar, or Weld counties, making these counties useful as sentinel counties, indicating when WNV is in the state.

WNV is in Colorado and will be with us for the foreseeable future. Major findings include the positive correlation of WNV with temperature, the positive correlation with precipitation, and the positive correlation with El Niño years. Current climate change predictions coupled with the results of this study indicate that WNV cases in select counties in Colorado can be expected to increase, indicating a need for mitigation strategies and WNV preparedness of medical personnel. These models and visualizations will assist public health decision makers in protecting the health of the citizens of Colorado.

6. BUSINESS BENEFIT

Hospitals, EMS, state, county, and local officials (health agencies) can all benefit from this analysis by using it to plan for WNV cases in the counties of Colorado.

Medical personnel can benefit by being aware of high-risk counties for WNV and being alert to the symptoms of WNV as the WNV season approaches.

State and county officials can direct limited resources to those counties most in need of mosquito mitigation strategies, surveillance activities, or water conservation efforts (in particular, surface water features). This can protect the citizens of Colorado and protect businesses from lost employee productivity.

7. SUMMARY AND FUTURE DIRECTIONS

Addressing WNV in Colorado, its causes and primary locations can protect the citizens and the labor force of Colorado. Mitigation strategies should be directed to identified high risk counties of the state. Early warning for medical personnel can be based on weather patterns observed and/or other indicator variables.

Current and future research could include the variables presented in Table 4.

Table 4: Current and Future Variables (see Appendix)

Analysis needed to enhance the model would include addition of identified variables, drill down analysis (instead of county, is there zip code level data available), and interaction effects (in particular between temperature, humidity, rainfall, and elevation – if available – in pursuit of the elusive "mosquito line" – a hypothesized but never derived "contour" line (iso-line) that would divide the state into two zones – mosquito and mosquito-free). A variable map dividing the variables into categories (environmental, physical, population, filters) which could be used to gage the impact of each of the categories would be useful.

8. REFERENCES

Bailey, M. (2022). Climate Change May Push the US Toward The 'Goldilocks Zone' For West Nile Virus. Retrieved from: https://techilive.in/climate-change-may-push-the-us-toward-the-goldilocks-zone-for-west-nile-virus/

Barker, C. (2019). Models and Surveillance Systems to Detect and Predict West Nile Virus Outbreaks. *Journal of Medical Entomology,* 56(10), p. 1508-1515.

Camm, J., Cochran, J., Ohlmann, J., and Fry, M. (2024). Business Analytics, 5th. Cengage Learning, Inc.

Centers for Disease Control (CDC). (2021). West Nile virus disease cases reported to CDC by week of illness onset, 1999-2020, graphical result. Retrieved from: https://www.cdc.gov/westnile/statsmaps/cummapsdata.html

Ciota, A. (2017). West Nile virus and its vectors. Current Opinion in Insect Science. 22, p. 28-36.

Colorado Department of Public Health and Environment (CDPHE). (2022). West Nile Virus Data. Retrieved from: https://cdphe.colorado.gov/animal-related-diseases/west-nile-virus/west-nile-virus-datahttps://cdphe.colorado.gov/animal-related-diseases/west-nile-virus/west-nile-virus-data

ISSN: 2473-4901
v9 n5826

Garcia-Carrasco, J., Muñoz, A., Olivero, J., Segura, M., and Real, R. (2021). Predicting the spatio-temporal spread of West Nile Virus in Europe. PLoS Neglected Tropical Disease, 15(1).

Hadfield, J., Brito, A., Swetnam, D., Vogels, C., Tokarz, R., Andersen, K., Smith, R., Bedford, T., and Grubaugh, N. (2019). Twenty years of West Nile virus spread and evolution in the Americas visualized by Nextstrain. PLOS Pathogens. 15(10). p. 1-18.

Kramer, L., Li, J., and Shi, P. (2007). West Nile Virus. The Lancet Neurology. 6(2), p. 171-181.

Marzec, N. (2022). Personal email communication. MPHZoonotic Disease Unit Manager Communicable Disease Branch, CDPHE.

Paz, S. (2015). Climate change impacts on West Nile Virus transmission in a global context. *Philosophical Transactions B* 370:20130561

Petersen, L. (2019). Epidemiology of West Nile Virus in the United States: Implications for Arbovirology and Public Health. Journal of Medical Entomology. 56(6). p. 1456-1462.

Ronca, S., Ruff, J. and Murray, K. (2021). A 20-year historical review of West Nile virus since its initial emergence in North America: Has West Nile virus become a neglected tropical disease? PLoS Neglected Tropical Diseases. 15(5). p. 1-14.

Tableau. (2022). Software package. Retrieved from: https://www.tableau.com

Wildlife Futures Team. (2021). West Nile Virus. Retrieved from: https://www.vet.upenn.edu/research/centers-laboratories/research-initiatives/wildlife-futures-program/resources/fact-sheets/fact-sheet-detail/west-nile-virus#:~:text=West%20Nile%20virus%20infects%20over,bite%20of%20an%20infected%20mosquito

Zaiontz, C. (2022). Real Statistics software add-in for Excel. Available from: https://www.real-statistics.com/

ISSN: 2473-4901
v9 n5826

Appendix: Tables, Figures, and Visualizations

Variable	URL
Calendar year	https://data.colorado.gov/Demographics/Total-Population-by-County-by-Year/9dd2-kw29
County name in Colorado	
Population (2020 Census) of county	
Number of WNV cases in county for given year	https://cdphe.colorado.gov/animal-related-diseases/west-nile-virus/west-nile-virus-data
Relative measure of WNV cases against county population	
Rainfall in county	https://www.ncdc.noaa.gov/
Urban: 1 if population of county is greater than 150,000; 0 otherwise	https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/2010-urban-rural.html
Minimum county elevation, ft	https://en.wikipedia.org/wiki/List_of_counties_in_Colorado
Maximum county elevation, ft	
Difference in elevation, Maximum - Minimum, ft	
Elevation at the geographical center of the county, ft	https://www.usgs.gov/products/data/all-data
Water area of county, square miles	https://en.wikipedia.org/wiki/List_of_counties_in_Colorado#County_data
Percent coverage of county by water	
Number of people in county per square mile	derived from population and area variables
1 if El Nino year, 0 otherwise	https://ggweather.com/enso/oni.htm
1 if La Nina year, 0 otherwise	
Area of county, square miles	https://en.wikipedia.org/wiki/List_of_counties_in_Colorado#County_data
1 if mean elevation of county is over 7000 ft, 0 otherwise	http://www.cohp.org/records/mean_elevation/mean_elevations.html
Average temperature per county for the four-month period (July – Oct.), Fahrenheit	https://www.ncdc.noaa.gov/cag/county/time-series
Rainfall average per county for the four-month period (July – Oct.), inches	https://www.ncdc.noaa.gov/cag/county/time-series
Qualitative rating of WNV historical case count per county (High, Medium, Low)	derived from WNV case data from 2005 to 2021

Table 1: Variables for the Colorado WNV Models

Model	R^2
$\hat{y} = -2.86 + 7.61(urban/rural) + 0.29(water\ area\ mi^2) + 1.09(El\ Niño) + 0.22(AvgRainfallJulyOct)$	25.7%
$\hat{p} = 1/(1 + e^{\wedge}(-\hat{y}))$, where $\hat{y} = -9.56 + 2.14(urban/rural) + 0.03(water\ area\ mi^2) + 0.4(El\ Niño) + 0.001(population\ density) + 0.13(AvgTempJulyOct)$	***Pseudo R^2** $R^2_L = 30.3\%$ $R^2_{CS} = 31.8\%$ $R^2_N = 44.3\%$

Table 2: Best models for WNV

 ISSN: 2473-4901
v9 n5826

* See the Real Statistics help page at: https://www.real-statistics.com/logistic-regression/significance-testing-logistic-regression-model/ (Zaiontz, 2022)

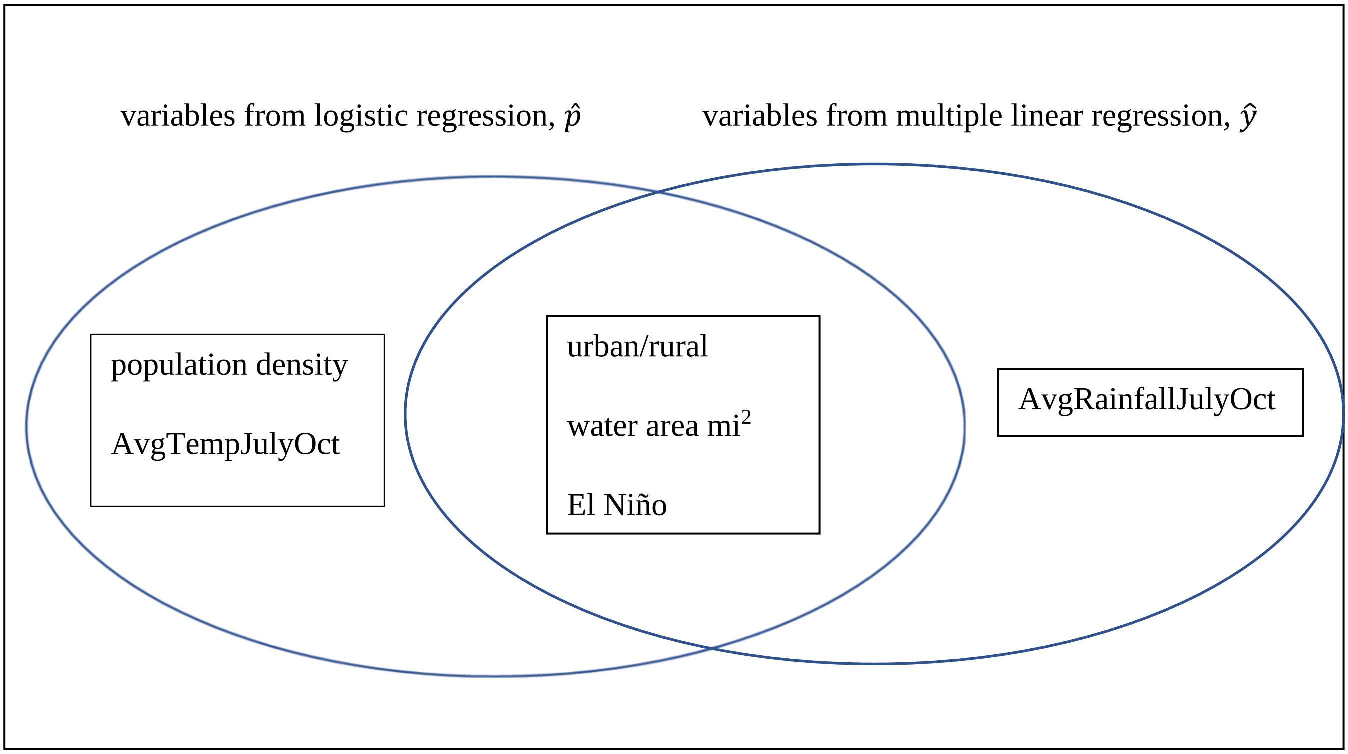


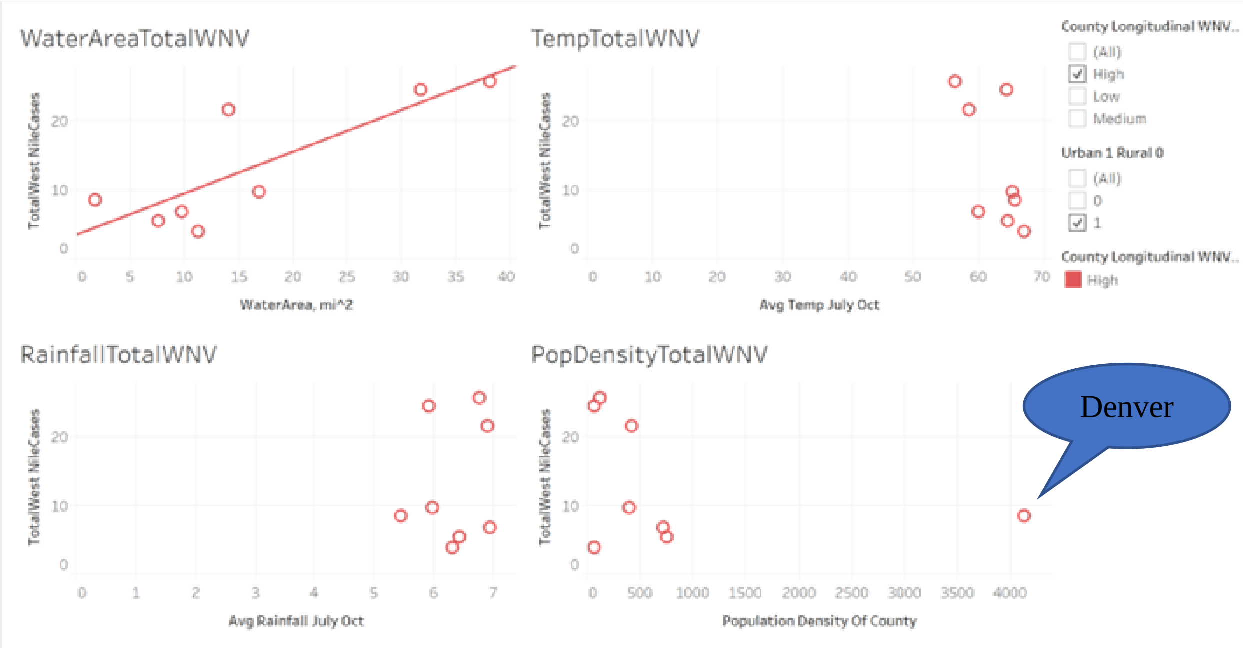
Figure 1: Significant variables for WNV models

ISSN: 2473-4901
v9 n5826

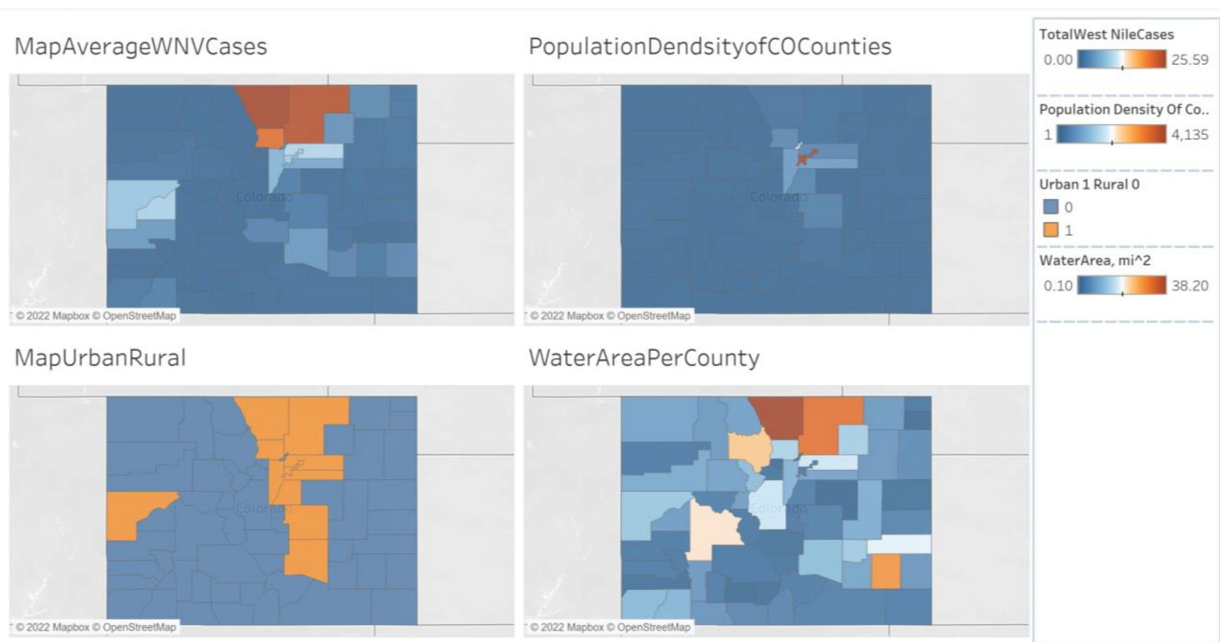
WNV Load Averages	Population	TotalWest NileCases	CasesPer 100000	Rainfall, in	Elevation, minimium, ft	Elevation, maximum, ft	WaterArea, mi^2	WaterPercent OfCounty	PopulationDensity OfCounty	Area, sq mi	MeanElevation Over7000ft	AvgTemp JulyOct	Elevation, ft CenterPoint	AvgRainfall JulyOct
Mean(medium,Low)	34873.67	0.68	4.08	1.25	5174.56	11000.00	5.68	0.40	57.85	1647.31	0.56	59.02	7069.00	7.30
Mean(high)	339314.56	11.85	7.54	0.66	4522.70	9206.80	15.11	1.08	672.03	1514.13	0.30	62.92	5756.60	6.33
t_test value	29.53	18.57	2.92	-6.06	-6.18	-5.56	15.96	18.33	13.97	-1.52	-6.23	7.51	-7.68	-4.63
Mean(low)	12484.47	0.24	3.43	1.34	5345.31	11388.58	5.70	0.42	10.46	1642.21	0.64	58.00	7344.44	7.40
Mean(high,medium)	248132.77	7.61	7.44	0.74	4427.05	9135.89	10.59	0.70	493.36	1589.30	0.21	63.49	5725.89	6.55
t_test value	-28.18	-14.71	-4.27	7.87	11.40	8.99	-9.77	-8.75	-13.80	0.76	14.24	-14.08	12.40	5.13
Qualitative comments	Higher WNV counties have a larger population	Higher WNV counties have more cases (the dependent variable)	Higher WNV counties have more cases per 100,000 also (another dependent variable)	Higher WNV counties have less rainfall	Higher WNV counties are lower in elevation	Higher WNV counties are lower here too	Higher WNV counties have more surface water	Higher WNV counties have more surface water	Higher WNV counties have greater population density	Higher WNV counties have "the same" area	Higher WNV counties are lower in elevation	Higher WNV counties are warmer	Higher WNV counties are lower in elevation	Higher WNV counties have less rainfall
Mean(El Nino)			6.30									60.26		6.45
Mean(not El Nino)			4.00									59.00		8.00
t_test value			2.88									2.41		-6.86
Qualitative comments			El Nino years have more WNV cases per county							With the areas being the same, this might make another nice grouping for agricultural use variables in future studies		El Nino years are warmer		El Nino years have less rainfall

Table 3: Statistically significant differences when dividing the data set along qualitative lines (high, medium, low and El Niño dimensions)

ISSN: 2473-4901
v9 n5826



Visualization 1: A Dashboard for Tracking WNV Variables



Visualization 2: Geospatial analysis of the WNV load in Colorado

Variable	Comments

ISSN: 2473-4901
v9 n5826

spring rainfall data	some authors have observed, on a country-wide basis, that these variables are indicators of WNV (a drought variable could also be a good filter variable for visualizations)
a drought variable	
WNV patient latitude and longitude	for more accurate accounting of WNV case locations
average humidity per county	difficult to obtain or compute
average elevation per county	
mosquito data (location, population density)	cuts across county boundaries
bird data (location, population density, migration patterns)	cuts across county boundaries
garden zone (planting guidelines, might be based on elevation)	cuts across county boundaries
monthly data rather than annual data (for drill down)	would be twelve times the data mining
agricultural variables (water usage)	these have shown some promise, percent of county irrigated (2007 and 2012) and number of cattle in county (2017-2021) were explored, and number of cattle was found to be significant in analytic models
interaction effects	these have shown some promise, temperature*rainfall was found to not be significant, but temperature*elevationCenterPoint was significant

Table 4: Current and Potential Future Variables