Prediction of Mosquitos with West Nile Virus

Reviewing Hanpu Yao’s Capstone

Mid-Point Work in Progress Review

by Alexander Nelms – CPLN 680 – 3/25/22

# General Comments

I was able to review Hanpu Yao’s Capstone by looking at his Mid-Term [*Draft and Code* Jupyter Notebook](https://github.com/CPLN-680-Spring-2022/Yao_Hanpu_Chicago_WestNileVirus/blob/main/Midterm%20draft%20and%20code.ipynb) as well as hearing his March 25th presentation ([Powerpoint](https://github.com/CPLN-680-Spring-2022/Yao_Hanpu_Chicago_WestNileVirus/blob/main/midterm%20pre.pptx)). Overall, I am very impressed by the state of Hanpu’s project – especially considering that the project was formulated in early February. The strong point of Hanpu’s Mid-term draft and presentation was his ability to use mosquito trap data to form kernel density maps of West Nile Virus mosquitos – partially answering his question of “Where is the high risk area of West Nile Virus in Chicago?” Particularly, I enjoyed Hanpu’s exploration of using ‘spatial lag with wind’ to add to his eventual prediction model.

Overall, I am interested in seeing Hanpu’s prediction model and how he specifically answers his question. To be honest, I think this Mid-term draft needs to a fleshed-out, formalized report – with detailed writing, a literature review and a methodology section. I will add more detailed remarks to Hanpu’s project in the sections below.

# Question & Goals

Hanpu’s current goals and questions are fairly straightforward and answerable. It would be more beneficial for the report to specifically define what he means by:

1. High Risk Areas – *e.g.* *where there are more WNV infected mosquitos? Or areas that have more mosquito-to-human infections?*
2. Finding where to allocate pesticide accordingly
3. Which time scales? – *e.g. hourly, time of day (morning), weekly, seasonly*
4. Which spatial scales? *– e.g. neighborhood, census tract, fishnet*

I think having a more detailed report would better communicate Hanpu’s methodology and goals.

# Literature Review

I don’t believe I found Hanpu’s Literature Review. It does seem time intensive to get deeper into the literature of West Nile Virus and how it spreads. But I believe that developing a Literature Review will help guide the project further and make the eventual findings stronger.

# Datasets

It appears as if Hanpu has already found a fair amount of data to turn into features for his prediction model. As Hanpu mentions, it would be nice to have spray data for more than 2 years. With more spray data, it will allow him to critique the city’s current pesticide treatment of mosquitoes and find the optimal treatment.

Another dataset that could help transform the study would be the ‘locations of human WNV infections’. It will likely be difficult to find, but it could be fruitful to understand how many humans are infected.

# Current Status & Problems

I think it is difficult to answers some of the draft’s stated problems as I do not know the intended methodology to answer the questions. I believe it would be best to narrow down the initial goal/question to a quantifiable answer. For example,

* *Which neighborhood in Chicago has the highest rate of mosquito infections?*
* *Which temporal & spatial factors have the highest correlation to infected mosquitos being trapped?*
* *Which areas in Chicago have the highest difference in infected mosquito rate and spray data?*