Contents

1 Introduction

1.1 What are Asian Giant Hornets

The Asian giant hornet, *Vespa mandarinia*, is a species of hornet that is natural to the temperate and tropical climates of Eastern Asia, most commonly Japan. However, even though they are native to Eastern Asia, this hornet species was seen and identified in North America in 2019 and 2020. The impacts of the potential introduction are still being determined but will be discussed in this paper.

1.2 Why Are They an Issue

The introduction of the Asian giant hornet into the North American ecosystem poses numerous environmental and health risks. The Asian giant hornet is considered an invasive species by the United States Department of Agriculture. Besides food consumption and competition with other species, its most significant environmental impact is its effect on honey bees.

Honey bees are the natural prey of the Asian giant hornet. The Asian giant hornet will invade a honey beehive to eat the bees and their hive for protein. In Japan, the Japanese honey bee, which has co-evolved with the Asian giant hornet, has developed a defense for this attack. The Japanese honey bees will form a ball around the invading hornet and buzz their wings. This will create heat and raise CO2 levels to kill the invading hornet. [1]

However, the western honey bee, the bee present in North America, has not co-evolved with the Asian giant hornet, leaving them defenseless against an attack. Although a single Asian giant hornet would not destroy a colony, a number upwards of 2-3 hornets have been known to decimate beehives, leaving only a few bees alive.[1] The impact on honey bees in North America is not completely clear. However, the spread of hornets across North America will most likely have a large effect on honey bees' colonies.

An additional impact of the introduction of the Asian giant hornet in North America is the health concerns. Like other wasps, the Asian giant hornet can be incredibly defensive. The sting of an Asian giant hornet can be incredibly dangerous and painful. Evidence has shown that the stings of an Asian giant hornet, although rare, can cause skin necrosis and hemorrhaging. It is shown that Asian giant hornets kill around 50 people a year, which is not a lot compared to the 62 people killed by bees and wasps per year.[1] Although Asian giant hornets have a dangerous sting and will create a new hazard, their total impact is still not known.

1.3 Characteristics and Life Cycle

Vespa mandarinia are characteristically found nesting in the ground. Generally, the nests are built in abandoned animal burrows or a tree's cavity, still within 3 feet of the ground, or in the roots. Starting in spring, approximately from March to June, hibernating queens that had been fertilized before winter will search for an ideal location to find a nest. This nest will be within 30km of the original nest they emerged. Once the queen has found an ideal nest location, she will begin laying eggs. Once the nest's population has reached 40 workers, the queen will not leave the nest and will continue to lay eggs. A queen can lay over 1000 eggs in this time period. During the summer, approximately from July to September, the nest will reach a peak population of 100 workers. [2] At this time, workers generally stay 2km of the nest, but depending on food access, they can go up to 8km away from the nest. Around the end of October, new queens and males will begin to emerge. The male hornets exit the nest and wait at its entrance for the reproductive females to emerge and mate. After the females mate, with typically only a single male, they head out searching for an ideal overwintering spot. Generally, this spot will be in the soil, and they will remain there until they emerge in spring and found their colony. The original queen will die off, and the colony, subsequently, will weaken and die off as workers are dying and not being replaced.

2 Our Assumptions

For our purposes and abilities with the model, we had to make several assumptions. These assumptions are listed and discussed below:

- All positive identifications are distinct and separate wasps.
  - There is a possibility that multiple sightings could be of the same wasp. However, we make the assumption that each sighting is of a different wasp.
- Hornets stay in presumed area.

- Research shows that wasps will most likely stay within 2km of the nest and will not leave an 8km radius of the nest. However, outliers exist, and there is a possibility that a wasp went farther than that 8 km. For the purposes of our model, however, we assume that all wasps stay within an 8 km distance of the nest.

- For the data we have, we consider unverified and unprocessed reports as the same.
  - There were 15 unprocessed reports and 2342 unverified reports. However, because none of these reports are confirmed to be either negative or positive, we will consider them to be the same.

3 Spread of Pest

3.1 Determining the Number of Nests

An important aspect of identifying and modeling the spread of *Vespa mandarinia* in the Whatcom County and surrounding areas is the ability to identify how many and where the nests are located. In Figure 1 each point represents a positive ID report of *Vespa mandarinia*. We know that typically *Vespa mandarinia* stays within 2km of the nest; the red circles represent this. With this information on the graph, we can begin the process of determining the likelihood of a nest being in a particular area. To determine the likely locations of a nest in a given area, numerous radii intersections are identified. Looking at the two clusters on the graph between latitudes 48.95 and 49.0, we see two potential nest sites. Looking strictly at this graph, each point should be within 2km of a nest, so there are eight potential nest sites. Since we know that, although the hornets generally stay within 2km of a nest, they can go as far as 8km away.

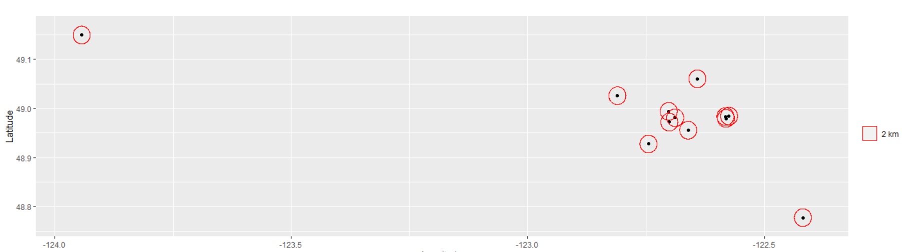

Figure 1: Representation of the points where positive ID's of *Vespa mandarinia* have been reported. The red lines represent a radius of 2km around the point of identification.

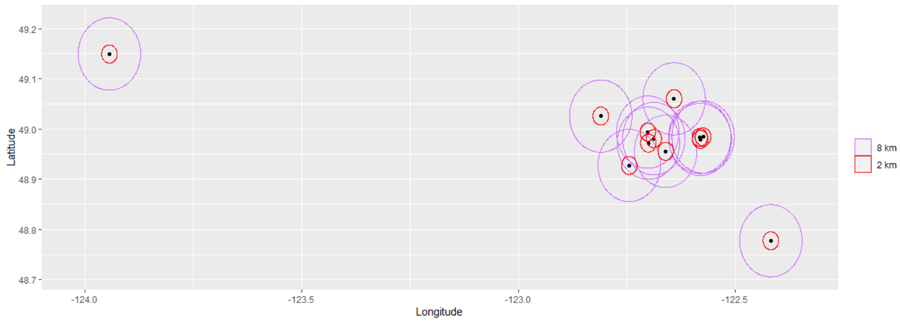

Figure 2: Representation of the points where positive ID's of *Vespa mandarinia* have been reported. The red lines represent a radius of 2km around the point of identification and the purple lines represent a radius of 8km.

Thus, Figure 2 shows the 8km radius in purple, so we can better determine potential nest sites. Having the 8km radius gives a better view of potential nest sites, as it is unrealistic and unlikely that there are eight nests, especially when there are such a low number of positive identifications. Using Figure 2, it is possible to determine that there are likely a total of 4 nests. From the data, it was reported and verified that the location of one nest was found. Although this nest was destroyed, it was reported to have had 600 unhatched eggs larvae and 150 live hornets. It is unknown, from the report, whether or not the queen was killed, so it must be assumed that a nest could be at that location. This nest is included to give an accurate representation of the potential spread of the pest. Figure 3 shows the location of this verified nest. In order to determine the potential locations for the other three nests, Figure 4 is used. By analyzing the locations of the positive IDs and both the 2km and 8km radii around them, we can determine the likely location of a nest. These are represented by X's on Figure 4.

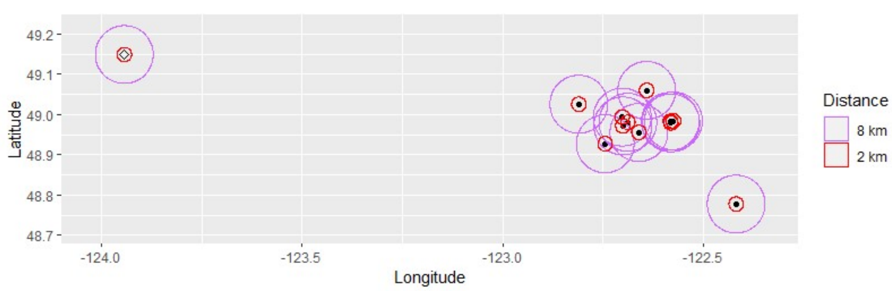

Figure 3: Representation of the points where positive ID's of *Vespa mandarinia* have been reported. The red lines represent a radius of 2km around the point of identification and the purple lines represent a radius of 8km.The open diamond on the graph representis the verified location of a *Vespa mandarinia* nest.

Figure 4: Representation of the points where positive ID's of *Vespa mandarinia* have been reported. The red lines represent a radius of 2km around the point of identification, the purple lines represent a radius of 8km, and X's represent the potential nest sites.

3.2 Showing the Spread Each Year

Given the location of a nest, a queen will move on average 30 km from its starting nest to the location of each year. This means that one can draw a 30 km circle around a nest's location to give an estimated idea of where new nests could be located. In our model, we only have confirmation of one nest's location and have estimated the locations of three additional nests. Because of this, we placed 30 km circles around each location of a positive identification. This can be seen in Figure 5, where the green circle gives the radius of the distance a queen will travel in one year. This will give researchers an area to search for potential nests in the following year. Additionally, these estimated areas can be used in conjunction with future data to determine where nests are located and where nests originated.

Additionally, using this model, researchers could estimate the area of potential nests after two years. Using the positive identified locations, we can place an additional circle around those points that extends up to 60 km. This can be seen in Figure 6, where the blue circle represents a 60 km radius. This would give us areas of where future queens could spread to after two years.

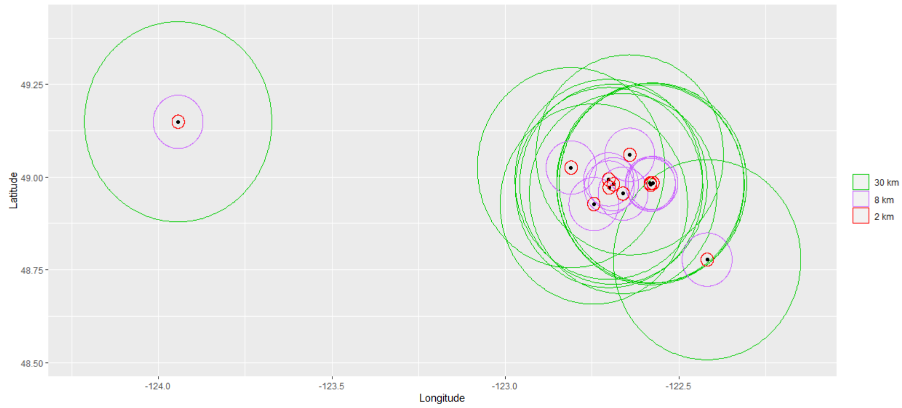

Figure 5: Locations where queens could potential start a new hive at within one year, using an average distance of 30 km.

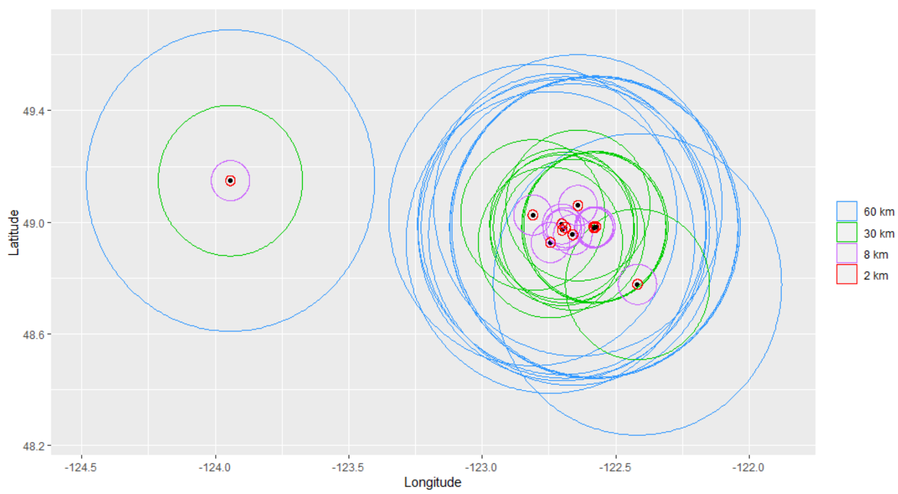

Figure 6: Locations where queens could potential start a new hive at within two year, using an average distance of 60 km.

3.3 What Level of Precision

It is crucial to notice that our model will not give precise locations of the nests, but it will provide us with an insight about where to look for nests and give us an idea of the spread. With the limited amount of data collected, it is challenging to determine the nests' locations with precision. However, as more data is collected, it is possible to improve the accuracy of our model.

Furthermore, as long as our data continues to be updated, we can identify with confidence areas in which we should be alert for *Vespa mandarinas*. Inside the 30km radius of each positive report, we know that there can be a nest; we see this in Figure 5. As a reference, we also have the 60km boundaries that we see in Figure 6. This 60km radii show the possible spread within two years. As mentioned before, this information becomes more accurate as more data is collected.

3.4 Weaknesses of the Model

One thing that must be taken into account when considering our model is that it is entirely dependent on positive identifications. The model acts as if those are the only wasps in the area when there could be unverified sightings in areas without any positive identifications. This weakens our spread prediction because there could be additional nests that would spread to new areas.

Additionally, our model does not consider the year or the time of year a positive identification took place. This is a weakness in our model because it completely disregards the part of the year that only queens are active. A potential improvement to our model would be to look at different parts of the year to see if we can determine where queens are making new nests. Also, some of the sightings were submitted in different years, and many changes could take place within a year. A future model could explore how sightings changed year to year.

Another weakness of our model is that our model does not consider the geography of the land. Our graphs act as if these hornets and nests exist on a flat plane when in fact, there are mountains, rivers, hills, bodies of water, and various other land formations. In reality, the land geography could have an effect on where queens choose to place their new hive. For example, if a queen is living on a peninsula surrounded by water, they will most likely follow the land rather than fly over the water. This weakness could be improved, however, with more in-depth graphing and map analysis.

A further weakness of our model is that we discuss the one positively identified nest and use that in our spread model. However, that nest was destroyed in 2019. However, we have no knowledge if the queen was killed and thus act as if she survived and could build a new hive. This is a weakness because there is a possibility that the queen was killed and will not be able to spread and build a new nest. This could be improved with further data about the destroyed nest.

A final weakness of our model is that it requires more data to make accurate predictions. We only have 14 positively identified sightings, which is not much data. This leaves many questions open about where the nests could be and how many nests we could expect to see. Even though we still estimated where nests could be, more data is necessary to make more accurate judgments.

4 Mistaken Classifications and the Model

4.1 Applying our model to unverified reports

An important part of our model is working with unverified reports. Unverified reports are reports that the lab did not confirm as negative or positive. There could be many reasons for not being able to identify a report. In the data for our model, there were 2357 unverified or unprocessed reports.

One of the problems with unverified reports is that we cannot know how accurate they are. To give us an idea of which of these reports we should take action on first, we ranked the reports based on how many positive reports were within 30km. If there are many positive reports around the same area, an unverified report is likely positive. There are some unverified reports with the same number of positive reports near. To know which of them goes in our ranking first, we see how close each report is from a positive report and put the one that is closest to a positive report first.

To achieve this, we went through several steps. These steps would take a lot of time to calculate by hand; we created programs using Java to perform all of the steps. We first calculate the distance between the coordinates of the two reports. For this, we use the Haversine formula, which calculates the distance between two points on a sphere. This formula is perfect for our model because it uses the latitudes and longitudes of the report's location, which we obtain from the data set for this problem. The Haversine formula to find the distance, d, is the following:

$$d = 2r\sin^{-1}\left(\sqrt{\sin^2\left(\frac{\varphi_2-\varphi_1}{2}\right)+\cos\varphi_1\cos\varphi_2\sin^2\left(\frac{\lambda_2-\lambda_1}{2}\right)}\right) \tag{1}$$

Where r represents the radius of the earth, φ_1 and φ_2 are the latitudes of the locations of the reports, and λ_1 and λ_2 are the longitudes of the reports [3].

For each report that is considered unverified, we calculate the distance to every positive report. If the distance is less than 30 km, we put its corresponding coordinates in a list. While calculating all the distances between coordinates, our Java program updates the number of positive reports close to our unverified report. The program also keeps track of which is the closest positive report and how far it is, as seen in Figure 7.

Our program can also give us a list of all of the coordinates that are within 30km from an

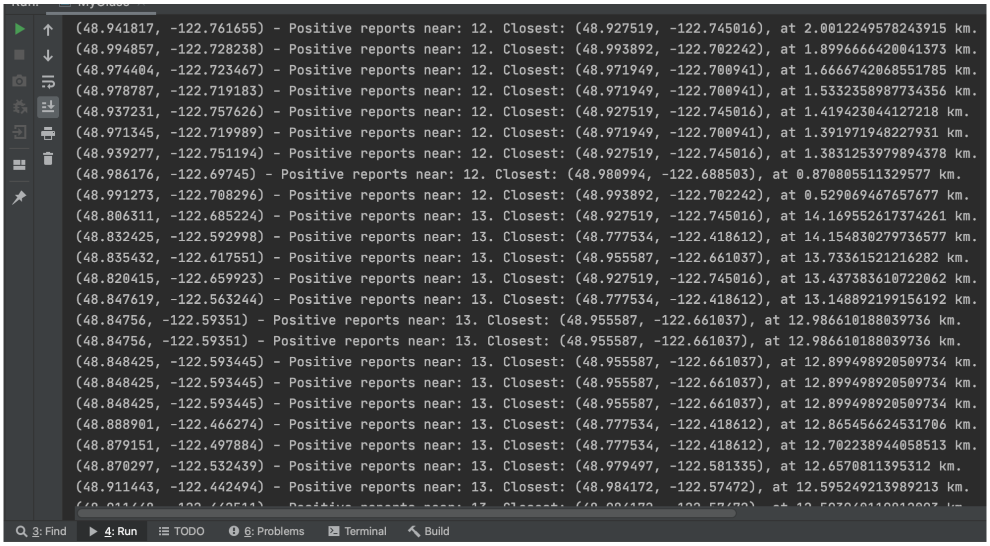

Figure 7: Program output that shows how many positive reports are near the unverified cases.

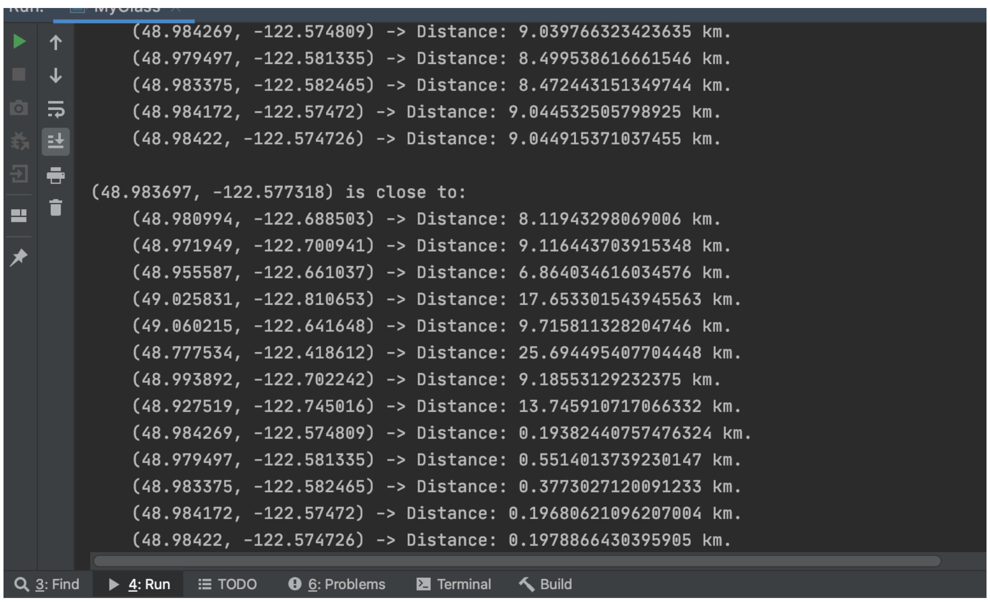

Figure 8: Program output that shows a list of every close positive report.

unverified report (See Figure 8).

With this, we can have an idea of how to proceed about the cases that could not be verified by the lab. We consider that the coordinates that come first in our ranking should be considered for further analysis.

4.2 Detailed Look at all Factors

4.2.1 Location in Respect to Positive IDs

A factor that has an impact in determining if an unverified report is likely a positive ID and should be further investigated or not is its proximity to a recorded positive ID. As mentioned in Section 2.1, the positive IDs were used to identify potential nest sites, and each hornet has to be within 8km of a nest. Therefore, proximity to numerous positive reports lends to the idea that the unverified report is close to a potential nest site and is another possible idea. This factor was taken into consideration significantly and applied to the model as previously discussed.

4.2.2 Location in Respect to Negative IDs

Another factor that can also play a role in determining if an unverified report is likely a positive Id and should be investigated further is its proximity to negative IDs. Suppose a particular area has a high number of negative identifications, then the likelihood of an unverified report being negative increases. One thing that could also be addressed is what the report had mistaken for the *Vespa mandarinia*. For example, suppose there is a cluster of negative IDs that are wood hornets all-around an unverified report. In that case, that report shouldn't be as much or a priority to investigate as it is most probable that it is also a wood hornet. Although this is not taken into account directly in our model, we focus mainly on the proximity to positive cases.

4.2.3 Report Description

Report description is a factor that could play a huge role in determining if an unverified report is likely a positive ID or not. Some reports describe the bug, such as the size and color of the insect. Using these descriptions, researchers could determine the likely hood of a report being a misidentification or not. Our model currently does take the description into account because of how varied they are. Further work on our model could take this into account, however.

4.2.4 Is There a Photo or Video Accompanying the Report

An additional factor that could help determine if a report should be investigated or not is if there is a photo submitted with the report. Researchers do not have the time nor the resources to go to every single physical location of a report and often rely on the photographic evidence provided to help determine if the report is a misidentification or not. If a report does not have a photo or video, it could make it incredibly difficult for the researchers to determine if the report is a hornet or not. Due to this, more attention should generally be given to reports that have a photo or video documenting the report's evidence.

4.3 Time

4.3.1 What Time of Year Was the Report

Time of year is something that should be considered due to the fact that depending on the time of year the hornets are more or less active and the population may be higher or lower. During the summer the population in a nest is at its peak and in the winter the only hornets alive are queens in the soil overwintering. With this in mind, looking at Figure 9 unverified reports in august have a higher probability of being positive than the unverified reports during December.

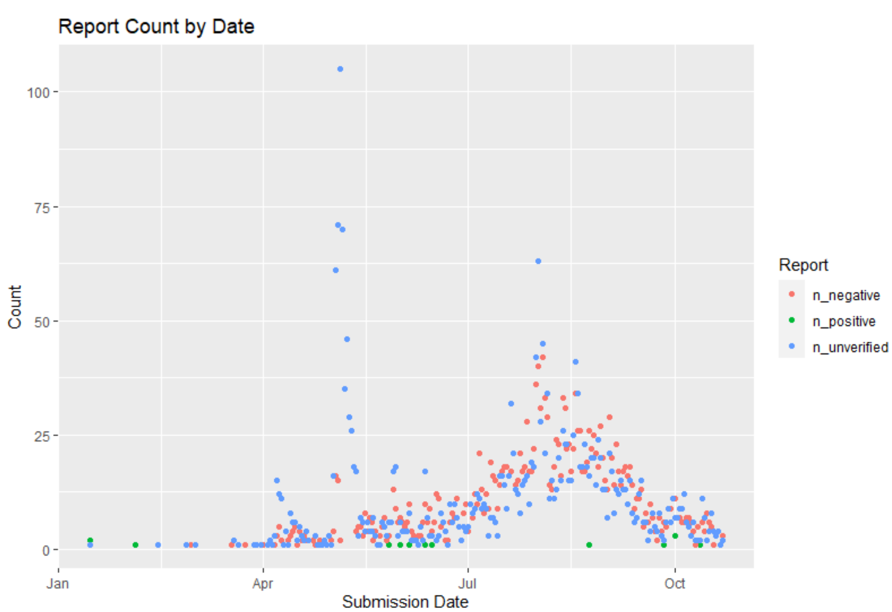

Figure 9: Caption

4.3.2 What Year Was the Report and Sighting

When determining the credibility and likelihood of an unverified report being a positive ID, the date of the sighting should be taken into account. It is known that *Vespa mandarinia* was first found in Washington State fall of 2019 [1]. Knowing this all reports of seeing *Vespa mandarinia* are very probable of being negative ID's if the predate 2019.

4.4 Weaknesses in our Model

This model shares the weakness outline for the first model as they rely on the same data. In tandem with those previously outlined weaknesses, another weakness of our model is that it does not take every factor that we have outlined. We determined and selected factors that have the greatest impact. Therefore, we focused on the proximity to a positive ID and home many different positive IDs the report is near. In the future, a way to strengthen the model more factors could be introduced into it.

5 Our Model Over Time

5.1 Adding New Reports to the Graph

Our model is able to be updated and become more accurate, and more reports are added to the graph. Each data point collected adds information and accuracy to our generated graphs. Adding more positive identifications, more negative identifications, and verifying the unverified reports that model will be more accurate and finding potential nests and mapping the spread of the hornets will become more precise.

5.2 How Often Updates Should Occur

Our model should receive an update of new data every 4 months to correspond with the changing seasons. This would allow researchers to focus on different seasons of the year, which could affect when wasps are out of the nest and when queens search for a new nest. This would also allow researchers to see how their predictions are coming true because they could watch the spread of hornets to new nests and areas.

6 Determining Eradication of the Pest

To determine the eradication of the pest, the incoming reports need to be evaluated. When, after a year since all the worker and male hornets from previous reports have died and only the queens remain, there are not positive reports, and there is a substantial decrease in unverified reports, meaning they are all able to be verified as negative ID's there is an implication that the pest has been eradicated. After two years, it can be said that the pest has been completely eradicated from the area.

7 Conclusion

The potential spread of the Asian giant hornet in North America is a danger that needs to be monitored and tracked. To do this, researchers are relying on the eye witness reports of people all across the north-eastern United States and Canada to help determine where these hornets could be. The model we have created can use these eye witness reports to help predict the spread of the pest. It allows researchers to identify which reports are most likely positive identifications, to allocate resources best while also allowing researchers to estimate where potential nests could be, and model the potential spread of new hives in new areas. Our model also can be constantly updated, taking in new data points and new factors that could make it more accurate. Overall, our model would be incredibly helpful in detecting, predicting, and eradicating this pest in North America.

8 Memorandum

The invasive species, *Vespa mandarinia*, poses a significant risk to the North American ecosystem and the people living in it. Thus, measures must be taken to help predict where potential hives may be and how the insect may spread over the years. The data needed for these predictions must be taken from eyewitness accounts taken from the area where the insect may be. However, not all of these reports are of *Vespa mandarinia*, and researchers will need to develop of way to determine which reports deserve to be analyzed and studied first. We have developed a model with current results to help predict hive locations, hornet spread, and report reliability. Our model is not perfect, however, and must be updated over time. However, the simplicity of this model allows it to easily be updated with new information and factors to increase the reliability of the results.

First, let us look at the hive locations. Identifying the possible locations of nests will significantly help the ability to track and eradicate this invasive species. To help determine these locations, our model only focuses on positively identified reports. These reports are given a latitude and longitude coordinate that allows us to see where they are in respect to each other, as seen in Figure 2. In this Figure, we have plotted each positively identified sighting locations circled each of them with two rings, one red and one purple. The red circle represents a radius of 2 km, and the purple circle represents a radius of 8 km. The 2 km circle represents how far hornets usually leave the nest, and the 8 km circle represents how far a hornet will travel from a nest, max. However, each of these circles is a sighting, not a confirmed nest. Since we only have one nest's confirmed location, we use each sighting as a potential nest and then use the overlapping circles to determine where a nest would most likely be. Figure 4 we have looked at the clustering of the sighting locations and the distances of the sighting locations to determine that there are three potential nests, marked with Xs.

We can also use this clustering of data points to determine these insects' spread over one and two years. It is known that hornet queens will create a new nest location, on average, within a 30 km radius of the original nest every year. Because of this, we can place 30 km circles around the positively identified reports to get a general area of where queens could potentially spread to. Again, not every point is a nest but is considered a potential nest. This gives us Figure 5, where the green circles outline potential locations for new queens to create new hives. We recommend researchers focus their efforts in these areas when searching for nests and the Asian giant hornet because of our model. We can also extend these circles to 60 km two give us a potential two-year spread. This is less accurate but is still helpful for researchers to focus their future efforts on.

Lastly, we have created a second model to help researchers determine which reports are most likely positive identifications. When creating public submissions for eye witness reports, you would likely receive many misidentified insects. Our model will help researchers determine which reports are most likely misidentifications by going through several steps. The first is to determine how many positive reports are within a 30 km radius. We say that if there are many reports around then, that unverified report is more likely to be positive. Using the Haversine distance formula, we create a ranked list of unidentified reports with positive identifications within a 30 km radius, as seen in Figure, 7. Our model can also give you a more descriptive list and tell you which specific positive reports an unidentified report is close to, as seen in Figure 8. We consider the reports that come first in our rankings should be considered for further analysis.

Our model is also highly variable and can be continuously updated. More factors can be added in to help determine both the potential spread and the unverified ranking. Much can be done with our model, and it can significantly benefit researchers in their ability to help eradicate this new pest. The findings presented in this paper will help keep this insect from invading further into North America.

References

[1] Michael J. Skvarla. Asian giant hornets. https://extension.psu.edu/asian-giant-hornets, May 2020. [Online; posted 6-May-2020].

[2] Cameron Jack Caitlin Gill and Andrea Lucky. Vespa manarinia. http://entnemdept.ufl.edu/creatures/MISC/BEES/Vespa_mandarinia.html, 2020. [Online; accessed 6-Feb-2021].

[3] prakhar7. Haversine formula to find the distance between two points on a sphere. https://www.geeksforgeeks.org/haversine-formula-to-find-distance-between-two-points-on-a-sphere/, 2020. [Online; accessed 6-Feb-2021].