



# An Independent Cascade Propagation Model of *Vespa mandarinia* Population with Propagation Index

Tong Zhou, Jin Gao, Tingzhen Liu, Yafei Jiang

College of Information Science and Engineering, Shenyang University of Technology, Shenyang, China  
Email: 1391725470@qq.com

**How to cite this paper:** Zhou, T., Gao, J., Liu, T.Z. and Jiang, Y.F. (2021) An Independent Cascade Propagation Model of *Vespa mandarinia* Population with Propagation Index. *Open Access Library Journal*, 8: e7345.

<https://doi.org/10.4236/oalib.1107345>

**Received:** March 24, 2021

**Accepted:** April 17, 2021

**Published:** April 20, 2021

Copyright © 2021 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

The *Vespa mandarinia* disaster has been a problem all over the world. In the prevention and control of *Vespa mandarinia*, the efficiency is often low and a lot of social resources are wasted because of the inability to effectively predict the spread of *Vespa mandarinia*. Therefore, this paper proposes a propagation model based on independent cascade. Based on the traditional propagation model, the habit factor and propagation index are added, which makes the habit and reproduction of wasps fully considered. *Vespa mandarinia* tend to migrate to the water edge, and the nodes closer to the water edge in the propagation model are more likely to be activated. The longer the *Vespa mandarinia* nests exist, the higher the probability of having more *Vespa mandarinia*, and the longer the nodes exist in the propagation model, the stronger the propagation ability. The model is validated by the real *Vespa mandarinia* disaster data in the United States in 2019, which proves that the propagation model considering *Vespa mandarinia* habits and reproduction will have better performance.

## Subject Areas

Biophysics, Biotechnology

## Keywords

Independent Cascade Model, Network Modeling and Simulation, Propagation Model, Ecological Informatics, Species Invasion

## 1. Introduction

*Vespa mandarinia* is the largest species of hornet in the world, and the occur-

rence of the nest was alarming. Additionally, the giant hornet is a predator of European honeybees, invading and destroying their nests. A small number of the hornets are capable of destroying a whole colony of European honeybees in a short time. At the same time, they are voracious predators of other insects that are considered agricultural pests [1]. When *Vespa mandarinia* was first discovered in Washington State, the United States, it had a serious impact on the local agriculture. When *Vespa mandarinia* were first found in Washington State, they had a serious impact on local agriculture [2]. So, it is very important to predict the reproduction and spread of *Vespa mandarinia* accurately for the prevention and control work, which is also the problem to be solved in this paper. Due to the small amount of data in such tasks, it is difficult for large-scale spatiotemporal prediction models such as graph neural network model [3] [4] to converge. Therefore, we introduce a traditional propagation model which only needs a small amount of data to calculation the parameters. The habits and nests of *Vespa mandarinia* were also considered.

## 2. Model

Classical propagation models include linear threshold model [5] and independent cascade model [6]. The linear threshold model considers that when an activated node cannot activate a neighbor node, its impact on the neighbor node is cumulative rather than discarded, and this contribution continues until the node is activated or the propagation process ends. The independent cascade model considers that an active node attempts to activate a neighbor node through a system variable: the probability of success, if it fails, the impact will be discarded [7]. However, it is a probability problem that *Vespa mandarinia* propagate from one location to another in this task, and there is no need for cumulative impact. Therefore, we use independent cascade propagation model to predict propagation.

Place 1000 randomly distributed nodes within 11,000 square kilometers of the *Vespa mandarinia* infestation area as checkpoints, and place these nodes in the node graph along with the nodes that have been identified as having *Vespa mandarinia*. When the distance between two nodes is less than the set threshold, an edge is established between the random node and the random node, the confirmed node and the random node. The weight on the edge is obtained by the weight formula between nodes.

In predicting *Vespa mandarinia* propagation, we use a directed graph to record the nodes that have been propagated and their propagation directions. When initializing that digraph, we put all the nodes that have been confirmed to appear *Vespa mandarinia* into a digraph. Searched for nodes that appeared in *Vespa mandarinia* on a directed graph, and conducted propagation operations for each of their neighboring nodes that did not appear in *Vespa mandarinia*.

A node that has been propagated at  $t$  has only a single opportunity to propagate its neighbors at  $t + 1$ .

Assuming that node  $V$  is activated at  $t$ , for any neighbor  $w$  of  $V$ , the probability that  $w$  is activated at  $t+1$  is  $P_{vm}$ .

The propagation diagram is shown below (Figure 1):

In the design of  $P_{vm}$  probabilities, We first need to consider the effect of distance on *Vespa mandarinia* migration [8]. The difficulty of migration of *Vespa mandarinia* continues to rise as the distance increases. This in turn leads to propagation that tends to be accomplished first at more distant nodes.  $(x_1, y_1)$   $(x_2, y_2)$  are latitude and longitude of two nodes, respectively. This was measured with the formula we formulated below:

$$L = 2R \arcsin \left( \sqrt{\sin^2 \left( \frac{x_1 - x_2}{2} \right) + \cos x_1 \cos x_2 \sin^2 \left( \frac{y_1 - y_2}{2} \right)} \right)$$

$$distScore = \frac{L_{max} - L}{L_{max}}$$

We need not only to consider the *distScore* obtained by the distance  $L$  between two nodes, but also the life habit of the *Vespa mandarinia* [9] [10]. Design considerations incorporating this knowledge into  $P_{vm}$  probabilities are highly advantageous for diffusion predictions. Through an in-depth exploration of the habit of *Vespa mandarinia*, it was found to prefer activities at the water side. This would be to bias *Vespa mandarinia* away from the water side more toward migrating toward the water side. Therefore, we introduce the straight-line distance of the nodes from the water edge. *Bias* is a constant. The weight formula for this term is shown below:

$$habitScore = \frac{S_v + Bias}{S_w + Bias}$$

We believe that the areas where *Vespa mandarinia* have been identified earlier will have more *Vespa mandarinia* than the areas where *Vespa mandarinia* have been identified later due to the continuous reproduction of *Vespa mandarinia* [11] [12]. And the advantage of quantity makes the areas that have been confirmed earlier have stronger communication ability. Therefore, we set the propagation index *Eff* for each node that has been propagated, and increase the propagation index of these nodes at the end of each round. The  $P_{vm}$  formula is shown below:

$$P_{vm} = \min \left( \left( \frac{S_v + Bias}{S_w + Bias} \times \frac{L_{max} - L}{L_{max}} Eff \right), 1 \right)$$

Propagation prediction for each round was performed as above, with nodes

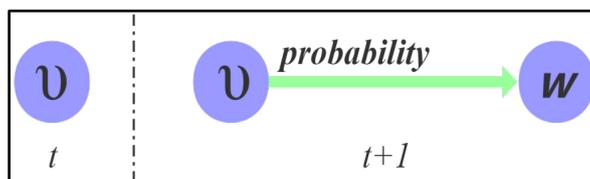


Figure 1. Probability of node  $v$  propagating to node  $w$ .

that were successfully propagated each round placed into a directed graph. In the next round, these nodes that were propagated successfully also performed a propagation operation on neighboring nodes to which they were not propagated.

### 3. Experiments

Our data set comes from the real *Vespa mandarinia* disaster in the United States in 2020 [2]. In order to verify the correctness of the propagation model, we divided fourteen confirmed *Vespa mandarinia* locations according to the time sequence. The first seven positions of *Vespa mandarinia* were used to predict the position of *Vespa mandarinia*, and the last seven positions were used to verify the prediction results. The quality of prediction results is measured by distance. The time difference between the seventh node and the fourteenth node in the dataset is about 150 days, while our model represents 50 days in a round of propagation. Therefore, we take the result of three rounds of model propagation as the prediction result. After three rounds, eight nodes are propagated, and we compare the first seven with the verification set. The measurement results are shown in the table below:

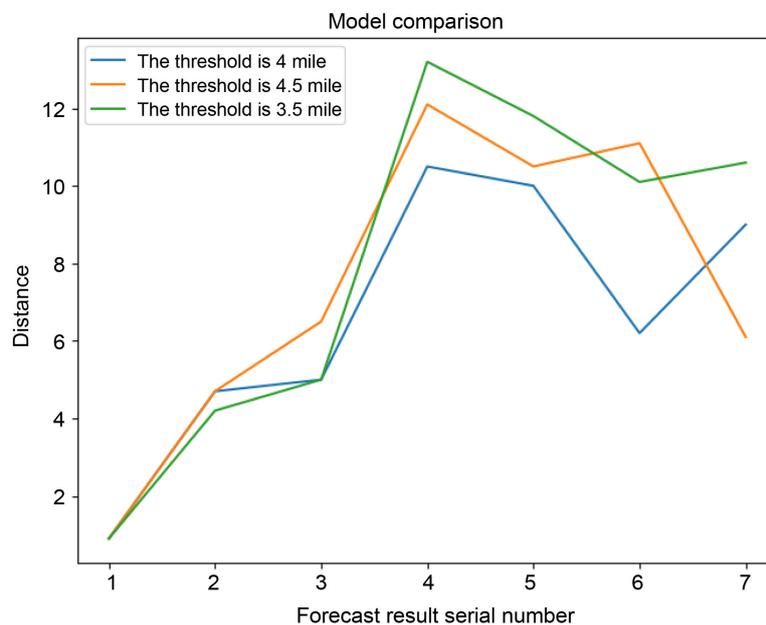
It can be seen from the **Table 1** that the model has a good performance in predicting the first few *Vespa mandarinia* occurrence sites. The first prediction is only 0.903 miles away from the real data, and the average distance between the first three predictions and the real data is 3.567 miles. This is because at the beginning, the data obtained by the model are real and reliable. However, with the increase of the number of propagation rounds, the proportion of the data generated by the model in all data increases. This leads to large data error, and constantly in the accumulation. This brings the average distance between the seven predictions and the real data to 6.657 miles.

We do sensitivity analysis on the threshold of node connection distance. By changing the threshold value, we can see whether the model is stable. The comparison is shown below:

As can be seen from **Figure 2**, after changing the threshold, the prediction results of the model are similar to those before changing the threshold, but the effect decreases slightly. Therefore, it can be considered that our model is

**Table 1.** Comparison between predicted propagation point and real propagation point.

| No. | Label     |             | Prediction |           | Distance/mile |
|-----|-----------|-------------|------------|-----------|---------------|
|     | Latitude  | Longitude   | Latitude   | Longitude |               |
| 1   | 48.993892 | -122.702242 | 48.98      | -122.7    | 0.903         |
| 2   | 48.927519 | -122.745016 | 48.99      | -122.7    | 4.775         |
| 3   | 48.984269 | -122.574809 | 48.93      | -122.65   | 5.069         |
| 4   | 48.979497 | -122.581335 | 49.03      | -122.8    | 10.507        |
| 5   | 48.983375 | -122.582465 | 49.01      | -122.8    | 10.032        |
| 6   | 48.984172 | -122.57472  | 49.07      | -122.62   | 6.286         |
| 7   | 48.98422  | -122.574726 | 49.01      | -122.77   | 9.029         |



**Figure 2.** Sensitivity analysis of *Vespa mandarinia* propagation model for distance threshold parameters.

relatively stable. And the parameters we choose are almost optimal.

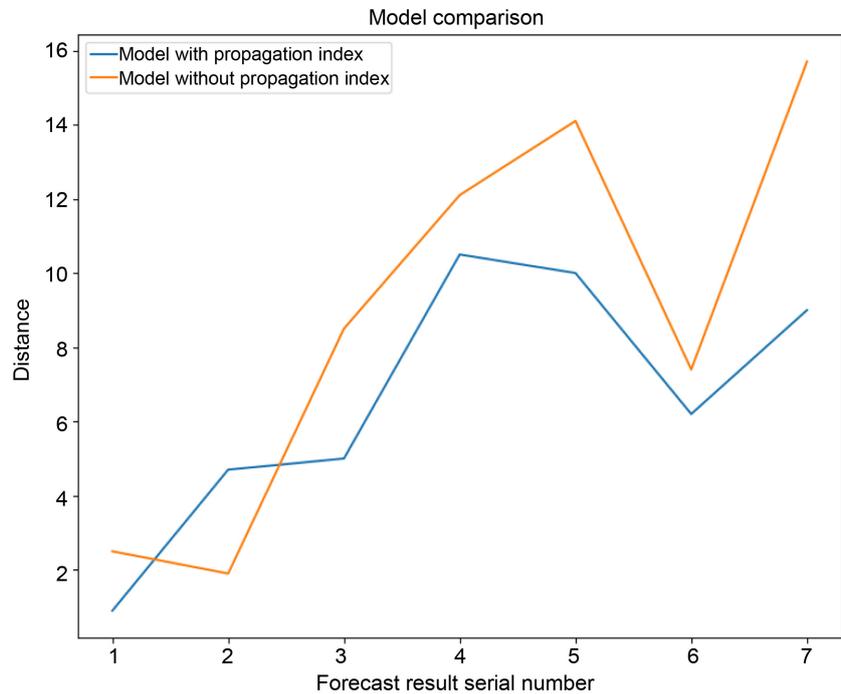
We also explore whether the communication index has a positive effect on the model prediction. The blue curve in the figure below is the actual distance between the predicted result and the real result of the model considering the propagation index, while the orange curve is the actual distance between the predicted result and the real result of the model not considering the propagation index.

From **Figure 3**, we can see that the model considering the propagation index performs better than the model not considering the propagation index for almost every predicted location. Therefore, the introduction of propagation index is beneficial to the model.

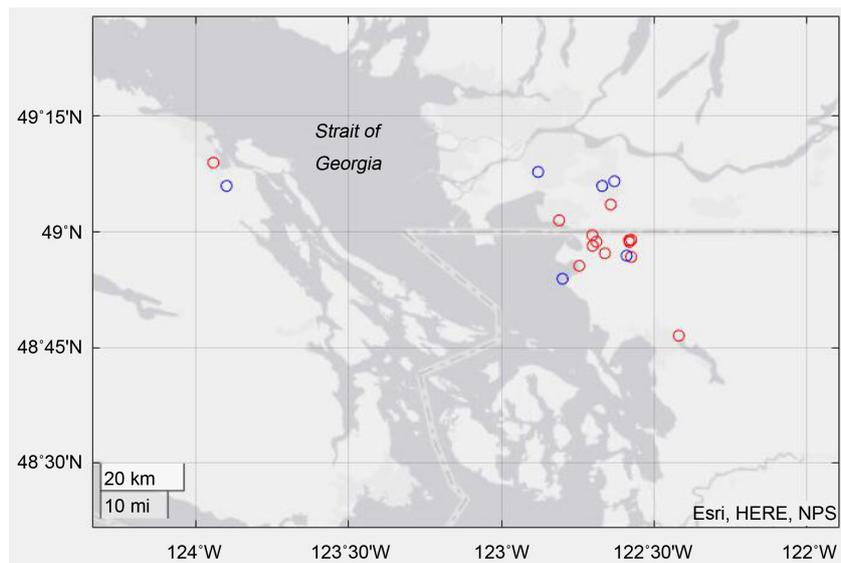
In our independent cascade propagation model. After the deduction and prediction of the model, the prediction results after 100 days and 200 days are obtained respectively. The results are shown in the **Figure 4**.

As can be seen from the above result diagram, *Vespa mandarinia* spread and migrated to the nearby areas to a certain extent. There are some clusters of *Vespa mandarinia* in the picture. This is because if there are many *Vespa mandarinia* in the vicinity of an area, it will have more chances to be spread. Because the life habits of *Vespa mandarinia* are used to guide the propagation model, we can see that the *Vespa mandarinia* migrate slowly to the water.

As can be seen from the diagram above (**Figure 5**), the *Vespa mandarinia* have made further migration to the nearby areas. We can see that due to the enhancement of aggregation effect, more nodes are propagated in this iteration than in the last one. That is, compared with the previous round, more *Vespa mandarinia* were found in this round. We can also see that the *Vespa mandarinia*



**Figure 3.** Error influence of propagation index on prediction propagation point.

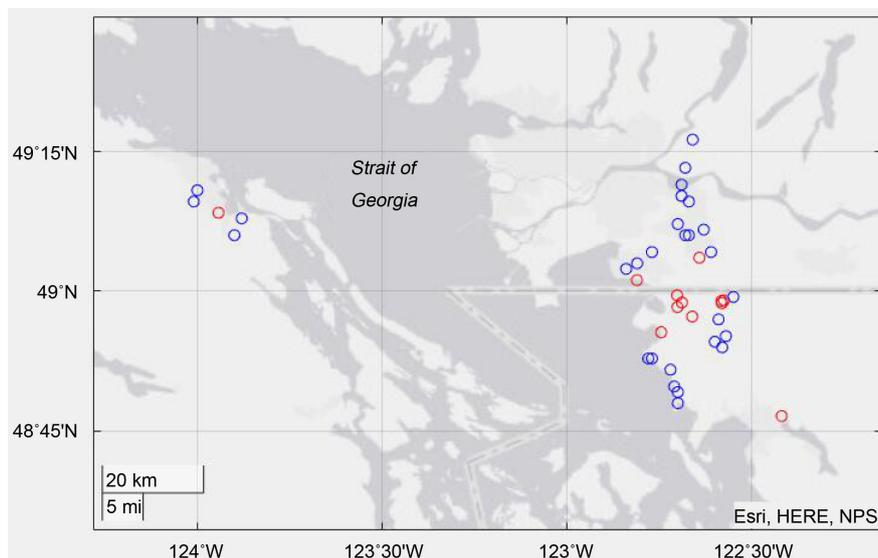


**Figure 4.** Prediction of spatial and temporal distribution of Bumblebee after 100 days.

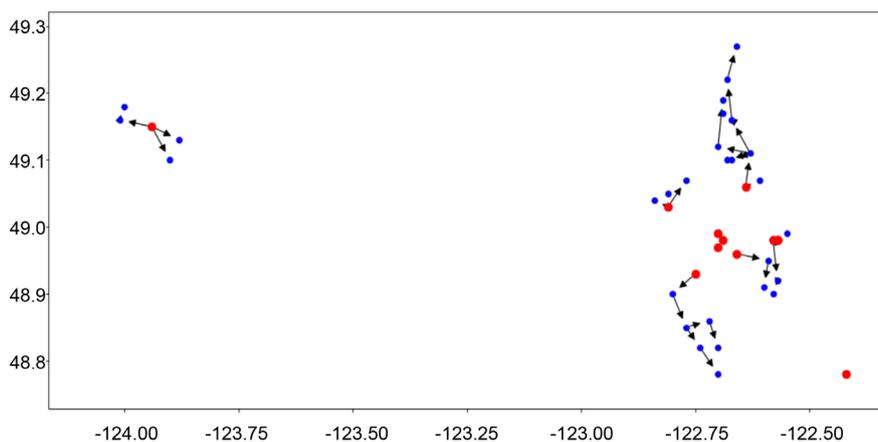
migrating to the water side continues to move to the water side, which makes its migration path form a line in the picture.

We also draw the spread trend chart, we can see the trend of *Vespa mandarinia* spread more intuitively.

From the **Figure 6**, we can see how *Vespa mandarinia* propagate around step by step, which can provide more targeted guidance for *Vespa mandarinia* control.



**Figure 5.** Prediction of spatial and temporal distribution of Bumblebee after 200 days.



**Figure 6.** Diagram of *Vespa mandarinia* transmission after 200 days.

#### 4. Conclusion

In this paper, we propose an independent cascade propagation model to predict the propagation of *Vespa mandarinia*. It fully considered the habits and reproduction of *Vespa mandarinia*, and provided guidance for the prediction of propagation model. The effectiveness of this guidance is verified by experiments. This model can also be applied to the propagation of other pests to make the control measures more effective.

#### Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

#### References

- [1] Kwon, O. and Choi, M.B. (2020) Interspecific Hierarchies from Aggressiveness and

- Body Size among the Invasive Alien Hornet, *Vespa Velutina Nigrithorax*, and Five Native Hornets in South Korea. *PLoS ONE*, **15**, e0226934.  
<https://doi.org/10.1371/journal.pone.0226934>
- [2] Wilson, T.M., Takahashi, J., Spichiger, S.E., *et al.* (2020) First Reports of *Vespa mandarinia* (Hymenoptera: Vespidae) in North America Represent Two Separate Maternal Lineages in Washington State, United States, and British Columbia, Canada. *Annals of the Entomological Society of America*, **4**, 322-337.  
<https://doi.org/10.1093/aesa/saaa024>
- [3] Scarselli, F., Gori, M., Tsoi, A.C., *et al.* (2009) The Graph Neural Network Model. *IEEE Transactions on Neural Networks*, **20**, 61.  
<https://doi.org/10.1109/TNN.2008.2005605>
- [4] Baskararaja, G. and Manickavasagam, M. (2012) Subgraph Matching Using Graph Neural Network. *Journal of Intelligent Learning Systems and Applications*, **4**, 274-278.  
<https://doi.org/10.4236/jilsa.2012.44028>
- [5] Qiang, Z., Pasilio, E.L. and Zheng, Q.P. (2019) Model-Based Learning of Information Diffusion in Social Media Networks. *Applied Network Science*, **4**, 80-82.  
<https://doi.org/10.1007/s41109-019-0215-3>
- [6] Saito, K., Nakano, R. and Kimura, M. (2008) Prediction of Information Diffusion Probabilities for Independent Cascade Model. *International Conference on Knowledge-Based Intelligent Information & Engineering Systems*, Springer-Verlag.
- [7] Ho, T.K.T., Bui, Q.V. and Bui, M. (2019) Information Diffusion on Complex Networks: A Novel Approach Based on Topic Modeling and Pretopology Theory. *Vietnam Journal of Computer Science*, **6**, 285-309.  
<https://doi.org/10.1142/S2196888819500155>
- [8] Barrows, E. (2013) Habitat Abundances of a Cricket-Parasitizing Wasp *Rhopalosoma nearcticum* (Hymenoptera: Rhopalosomatidae) in a United States Mid-Atlantic park. *Open Journal of Animal Sciences*, **3**, 311-313.  
<https://doi.org/10.4236/ojas.2013.34046>
- [9] Kojima, J.I., Hartini, S., Kahono, S., *et al.* (2002) Males, Mature Larvae, and Nests of *Ropalidia plebeja*, a Nearly Solitary Paper Wasp Endemic to Sulawesi (Insecta: Hymenoptera: Vespidae), with Taxonomic Notes on the *R. plebeja* Group. *Species Diversity: An International Journal for Taxonomy Systematics Speciation Biogeography & Life History Research of Animals*, **7**, 1-28.  
<https://doi.org/10.12782/specdiv.7.1>
- [10] Phil, R. (1928) The Nesting Habits of the Wasp, *Chalybion caeruleum*. *Annals of the Entomological Society of America*, **1**, 25-35.
- [11] Makino. Post-Hibernation Ovary Development in Queens of the Japanese Giant Hornet *Vespa mandarinia* (Hymenoptera: Vespidae). *Entomol*, **19**, 440-443.  
<https://doi.org/10.1111/ens.12205>
- [12] Takahashi, Akimoto, Martin, *et al.* (2004) Mating Structure and Male Production in the Giant Hornet *Vespa mandarinia* (Hymenoptera: Vespidae). *Applied Entomology and Zoology*, **39**, 67-75.