2021 MCM

问题C：确认有关黄蜂的嗡嗡声

2019年9月，在加拿大不列颠哥伦比亚省的温哥华岛上发现了 中国大虎头蜂（也称为亚洲大黄蜂）的群体/殖民地。 黄蜂的巢被迅速摧毁，但事件的消息迅速在整个地区传播。 自那时以来，在邻近的华盛顿州已经发生了几次确认的害虫目击事件，以及许多错误的目击事件。 有关侦查，大黄蜂监视和公共目击的地图，请参见下面的图1。

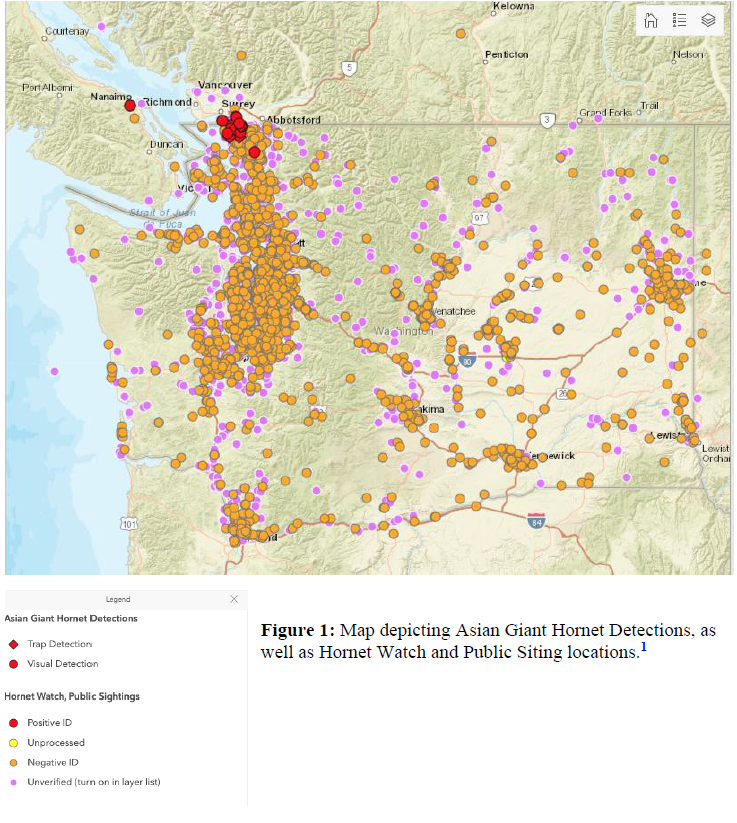


图1

图1：地图描绘了亚洲大黄蜂的检测情况，以及大黄蜂监视和公共选址的位置

亚洲大黄蜂是世界上最大的大黄蜂，其巢的发生令人震惊。此外，大黄蜂是欧洲蜜蜂的捕食者，入侵并摧毁了它们的巢穴。少量的大黄蜂能够在短时间内摧毁整个欧洲蜜蜂群落。同时，它们是被认为是农业害虫的其他昆虫的猛禽。

亚洲大黄蜂的生命周期与许多其他黄蜂相似。受精的女王在春天出现并开始新的殖民地。在秋天，新皇后离开巢穴，冬天将在土壤中度过冬天，等待春天。一位新女王预计将在30公里范围内筑巢。问题附件中包含有关亚洲大黄蜂的更多详细信息，也可以在网上找到。 由于可能会对当地蜜蜂种群造成严重影响，因此，亚洲大黄蜂的存在会引起很多焦虑。华盛顿州建立了求助热线和一个网站，供人们报告对这些黄蜂的目击情况。

根据公众的这些报告，纽约州必须决定如何优先分配有限的资源，以进行后续调查。虽然一些报道被确定为亚洲大黄蜂，但其他许多目击者却被证明是其他类型的昆虫。

这个问题的主要问题是“我们如何解释公共报告提供的数据？（How can we interpret the data provided by the public reports?）”以及“鉴于政府机构的资源有限，我们可以使用哪些策略对这些公共报告进行优先排序以进行进一步调查？（What strategies can we use to prioritize these public reports for additional investigation given the limited resources of government agencies?）” 您的论文应探讨并解决以下方面：

* 讨论并讨论这种有害生物在一段时间内的传播是否可以预测以及其精确度如何。 
* 据报道，大多数目击者将其他大黄蜂误认为大黄蜂。 仅使用提供的数据集文件和（可能）提供的图像文件来创建，分析和讨论可预测错误分类可能性的模型。 
* 使用您的模型来讨论分类分析如何导致对最有可能成为正面目光的报告进行优先调查。 
* 解决随着时间的推移会在给定其他新报告的情况下如何更新模型以及应该多久进行一次更新。 
* 使用您的模型，什么构成证据证明该害虫已在华盛顿州被根除？

最后，您的报告应包括两页的备忘录，总结您对华盛顿州农业部的研究结果。

您的PDF解决方案（总共不超过25页）应包括：

* 一页的摘要表
* 目录
* 完整的解决方案
* 两页的备忘录
* 引用列表

注意：MCM竞赛现在限制为25页。 您提交的所有内容均计入25页的限制（摘要表，目录，参考列表和任何附录）。 您不得使用未经版权法限制使用的未经授权的图像和材料。 确保引用您的想法来源和报告中使用的材料。

**问题C的一般准则**

除了上面列出的特定要求外，请记住这是一个统计建模练习。 提交的内容应遵循与数据使用相关的最佳实践。 这些期望的一些示例包括但不限于以下内容： 

* 定义您使用的所有指标和成本函数。 
* 参数的任何估计都应包括间隔估计。 
* 任何结果都应包括对结果拟合优度的估计。 
* 应明确陈述所有假设，尤其是与数据或错误相关的分布。 
* 应检查与数据相关的所有假设，并应检查针对这些假设的技术的稳健性。 
* 应明确说明与方法或技术有关的所有假设。

**Attachments**

We provide the four following materials for this problem. THE DATA FILES PROVIDED CONTAIN THE ONLY DATA YOU SHOULD USE FOR THIS PROBLEM.

1. **2021MCM\_ProblemC\_Vespamandarinia.pdf**

Background information from Pennsylvania State University Extension that describes the insect.

2. **2021MCM\_ProblemC\_DataSet.xlsx**

A spreadsheet with 4440 reports of sightings with the following fields:

**GlobalID**: a unique label for each sighting record.

**Detection Date**: the reported date of the sighting.

**Notes**: comments provided by the person submitting the report. This can be a member of the public, or occasionally a state employee.

**Lab Status**: the official classification of the sighting by the state department of agriculture after analysis. *Positive ID* means it is confirmed as an Asian Giant Hornet. *Negative ID* means it is excluded. *Unprocessed* means it has not yet been classified. *Unverified* means no determination was made due to lack of information.

**Lab Comments**: what the state entomology lab added to the record after analysis.

**Submission Date**: the date the report was made to the state. This date can be significantly after the detection date.

**Latitude (of sighting)**: these data are provided by the state after converting the address provided by the report.

**Longitude (of sighting)**: these data are provided by the state after converting the address provided by the report.

3. **2021MCM\_ProblemC\_Files.rar**

A rar file with 3305 images submitted with the sighting reports.

The 662MB file can be downloaded from: **http://www.comapmath.com/MCMICM/2021MCM\_ProblemC\_Files.rar**

**A password is required to open the file: Af6SP7rdm33PxPJmDb4wZq7cw**

4. **2021MCM ProblemC\_ Images\_by\_GlobalID.xlsx**

A spreadsheet mapping the images to the sightings with the following fields:

**FileName**: the name of an image in the rar folder.

**GlobalID**: a unique label for each sighting record. This is consistent across the two spreadsheets.

**FileType**: Images arrive as .jpg, .pdf, .png, .jfif, octet-stream, xml open format, or .zip files. Videos arrive as .mp4 or quicktime files.

**Reference**

Washington State Department of Agriculture. 2020 Asian Giant Hornet Public Dashboard.

https://agr.wa.gov/departments/insects-pests-and-weeds/insects/hornets/data Accessed 11/5/2020.