NIRS Model Trained on Lab-reared Mosquitoes Estimates Age of Wild Mosquitoes with 67% Similarity to Detinova

Masabho P. Milali1,2, Maggy Sikulu-Lord1,2, Samson S. Kiware1,2, Richard J. Povinelli3, and George F. Corliss1
1Marquette University USA, 2Hakara Health Institute Tanzania, 3QIMR Berghofer Medical Research Institute Australia

Introduction
No clear difference between spectra collected from lab-reared mosquitoes and those collected from wild mosquitoes (Milali, 2016). We applied NIRS model trained on lab-reared mosquitoes to estimate age of wild mosquitoes. We compared model estimates with Detinova estimates. The model estimated age with 67% similarity to Detinova. The ongoing practice of applying model trained on the lab-reared mosquitoes might be appropriate [3].

Method and Results
We applied a classification model trained on lab-reared An. arabiensis to classify the ages of wild An. arabiensis. Because we lack age labels of wild An. arabiensis, we indirectly validated our model by:
1. Analyzing the distribution of nulliparous and parous mosquitoes in each age-class from the model;
2. Comparing the number of mosquitoes in each age-class obtained when classification is done using a model and when done using Detinova ovary dissection; and
3. Relating with the historical studies conducted to determine the age structure of wild mosquito populations.

We thank Grand Challenge Canada for funding the study produced the data used in this study.

Results comparison one
Comparison two by Month
Results comparison two
Comparison three (Brownstein et al.)

We applied a classification model trained on lab-reared An. arabiensis to classify the ages of wild mosquitoes. Because we lack age labels of wild An. arabiensis, we indirectly validated our model by:
1. Analyzing the distribution of nulliparous and parous mosquitoes in each age-class from the model;
2. Comparing the number of mosquitoes in each age-class obtained when classification is done using a model and when done using Detinova ovary dissection; and
3. Relating with the historical studies conducted to determine the age structure of wild mosquito populations.

Results comparison one
Comparison two by Month
Results comparison two
Comparison three (Brownstein et al.)

We thank Grand Challenge Canada for funding the study produced the data used in this study.

Maths
Equation to calculate silhouette coefficient
Let
- \( s(o) = \text{Silhouette coefficient of a single object } 'o' \)
- \( b(o) = \text{Average distance of object } 'o' \) to the other objects in its cluster
- \( h(o) = \text{Average distance of object } 'o' \) to the other objects in the nearest cluster.

Then
\[
s(o) = \frac{b(o) - h(o)}{\max(b(o), h(o))}
\]

Jaccard Coefficient: Jaccard similarity coefficient is a static measure for comparing the similarity diversity of sample sets. Assume A and B are two different sets. Then
\[
J(A, B) = \frac{|A \cap B|}{|A \cup B|}
\]

Interpretation:
- Low JC = The sets are dissimilar (JC = 0, sets are different)
- High JC = The sets are similar (JC = 1, sets are same)

If we translate this to our problem, A can be Detinova and B can be NIRS. Therefore,
\[
J(\text{Detinova, NIRS}) = \frac{\text{Detinova\_NIRS}}{\text{Detinova\_NIRS}}
\]

Conclusion
While further studies may be required to explore a more appropriate way to estimate age of wild mosquitoes, these results strengthen the ongoing practice of training models to estimate age of wild mosquitoes using spectra collected from lab-reared mosquitoes [3]. The reliability of the age estimates from the model might still be questioned, which is acceptable as model estimates are not always expected to be accurate [2, 4, 5]. Despite of this known caveat, the most important advantage of using models is to give insights to situations where it is difficult to get the truth [4, 1]. Getting actual age in days of wild mosquitoes is almost impractical as it is very difficult, tedious, time inefficient, and expensive. Therefore, the ongoing practice of applying a model trained on laboratory-reared mosquitoes to estimate wild mosquitoes might not be an ideal, but the results from this study show that it might be reliable enough to give an insight on age structure of wild mosquito population, especially when complemented with other existing knowledge on age structure of wild mosquitoes.

References
2. Poiner E. 2007. All Models are Wrong, How do We Know Which is Useful? Ground Water. 45(4):393-9