Charles B. Sither

School of Health Sciences

Western Carolina University

Cullowhee, NC 28723

Email: cbsither@gmail.com

Phone: (828) 734-0682

**Mid-Infrared Spectroscopy Detects Gonotrophic Status and Age in *Aedes triseriatus* (Say)**

Charles B. Sither1\*, Bradley F. Guilliams2, Mark A. Rothermund2, Brian D. Byrd1, and Scott W. Huffman2

1. School of Health Sciences, Western Carolina University, USA
2. Department of Chemistry and Physics, Western Carolina University, USA

\*Corresponding Author

**Abstract**

**Introduction**

Mosquitoes remain a constant source of global communicable disease burden ( ).

The adult mosquito age plays a critical role in understanding and predicting mosquito-borne disease risk ( ) since many mosquito-transmitted pathogens are horizontally acquired (mosquito to host to mosquito) and require an extrinsic incubation period (EIP) because the mosquito is consider infectious (figure XX). Determining whether a mosquito has laid a batch of eggs – indicating that she has taken blood (in most mosquito species) – is an integral piece of information when investigating arboviral disease transmission. Current methods for determining mosquito age include morphological assessment ( ), ovary dissections [1], gene expression [2], mark-recapture [3], mass spectroscopy [4], and infrared spectroscopic methods [5–7]. Many of these methods are time intensive, require specialized labor, and/or expensive. All of which delay or inhibit collecting critical information when researching or controlling arboviral diseases ( ).

Mid-infrared spectroscopy (MIRS) has shown promise as a relatively inexpensive, low labor, and quick method for estimating mosquito age and species determination in common African malaria mosquitoes ( ),

Outside of *Aedes aegypti* (L.) few age grading methods have been attempted

Given that spectroscopy is non-destructive, less labor intensive, and does not require specialized knowledge compared to some other age grading methods, we sought to assess its applicability for determining gonotrophic status and age in *Ae. triseriatus* (Say) mosquitoes. Here, we measured the mid-infrared region (4,000 – 400 cm-1) employing a

1. Why mosquito matter
   1. Importance to public and veterinary health
   2. Important factors for determining disease transmission
2. Current methods in mosquito species identification
3. Current methods in mosquito physiological identification
4. Introduction to mid-infrared spectroscopy
   1. What is infrared spectroscopy
   2. Why is the mid-infrared region useful compared to near- and far-infrared
   3. Classifying MIRS data
      1. Partial least squares regression
      2. Neural networks
      3. Naïve Bayes
5. Previous work using mid-infrared spectroscopy
   1. Previous studies
   2. Advantages compared to current methods
6. Purpose of this study

**Methods**

*Specimen Rearing and Storage*

We reared *Aedes triseriatus* mosquitoes (MSU strain) at standard laboratory rearing conditions (27oC, 75% RH, 16:8 day/night photoperiod). Emerged adults (>500) were placed in a single rearing cage and held for 7 days allowing for mating and maturation. Approximately 250 mosquitoes were removed from the original cage and fed to repletion on human blood. We provided ovipositioning substrate 72 hours post blood feeding. After ovipositioning, mosquitoes were concurrently removed from both the original cage (non-bloodfed; nulliparous) and the secondary cage (bloodfed; suspected parous). Parous and nulliparous mosquitoes were killed by freezing and stored at -20oC. A total of 170 mosquitoes were then selected for MIRS analysis: 99 previously blood-fed, and 71 not blood-fed.

*Mid-Infrared Spectroscopy*

We measured the mid-infrared region (4000–650 cm−1) using a NicoletTM Centaurμs™ housing a Nicolet™ IS™ 10 FTIR spectrophotometer equipped with a potassium bromide (KBr) beamsplitter at 4 cm-1 resolution.

A single IR spectrum of each mosquito tibia was acquired using a ## SPEC MODEL HERE Brads work: Nicolet model Centaurus infrared microscope ## FT-IR spectrophotometer equipped with ## MEASUREMENT APPARATUS AND SPEC MEASUREMENT TYPE (CAVITY RING DOWN?) ## and transformed with the ## SOFTWARE HERE ## software. To ensure consistency amongst measurement locations between mosquito specimens, the tibia from the midleg of each mosquito was removed and slide mounted prior to spectral acquisition. Background and MIR spectra were determined by averaging over >60 scans at a resolution of 4cm-1 with a spectral range 4000 to 650cm-1

*Experimental Design*

We conducted two separate experiments: predicting mosquito age and predicting gonotrophic status. In the first experiment, *Aedes triseriatus* mosquitoes were sampled at specific time intervals post-emergence (±1 day) from a single mosquito colony

*Gonotrophic Status Prediction*

We implemented a Naïve Bayes Classifier to differentiate between nulliparous and parous *Aedes triseriatus* mosquitoes. Prior to model training and assessment, we randomly assigned data to training (80% of all data) and testing (20% of all data) datasets. We used the Gaussian Naïve Bayes algorithm implemented in the scikit-learn package (v1.5.1) in Python3 (v3.12.0) for all model training and testing. As stated, we chose a Gaussian distribution as our prior model with prior probabilities of 50%. Meaning, the base model without any data assumes that there is a 50% chance of a mosquito being parous or nulliparous. A

We initially constructed a multinomial Naïve Bayes classifier to predict whether the data could predict the MIRS sampling date. This was done to ensure that even t

We implemented a Partial Least Squares regression (PLS) algorithm with Python3 (v3.12.0) in the scikit-learn package (v1.5.1) to distinguish between nulliparous and parous *Aedes triseriatus* mosquitoes. Prior to model training and assessment, we randomly assigned 80% of the data to a training dataset and 20% to a test dataset stratified by prediction status (nulliparous/parous) and sampling date. The PLS algorithm requires a single hyperparameter, number of components, that must be decided in advance of model training. To determine the optimal hyperparameter value, we performed a grid search over *n* components from *n* = 2 to *n* = 12 and selected the optimal components based on which component minimized mean squared error (MAE) over 10 shuffled K-folds. We attempted a second optimization step involving Variable Importance in Project (VIP) (CITE HERE) excluding variables with a VIP score < 1. Using these selected features, we trained a second model containing only regions with VIP scores ≥ 1. From here on out, Base Model refers to the non-VIP optimized model, while VIP Model refers to the model that underwent VIP feature selection.

*Age Prediction*

**Results**

**Discussion**

**References**

1. Detinova TS. Age-grouping methods in Diptera of medical importance, with special reference to some vectors of malaria. World Health Organization Monograph Series. 1962;13–191.

2. Joy TK, Gutierrez EHJ, Ernst K, Walker KR, Carriere Y, Torabi M, et al. Aging Field Collected Aedes aegypti to Determine Their Capacity for Dengue Transmission in the Southwestern United States. PLOS ONE. 2012;7:e46946.

3. Colchero F, Clark JS. Bayesian inference on age-specific survival for censored and truncated data. Journal of Animal Ecology. 2012;81:139–49.

4. Wagner I, Grigoraki L, Enevoldson P, Clarkson M, Jones S, Hurst JL, et al. Rapid identification of mosquito species and age by mass spectrometric analysis. BMC Biol. 2023;21:10.

5. Siria DJ, Sanou R, Mitton J, Mwanga EP, Niang A, Sare I, et al. Rapid age-grading and species identification of natural mosquitoes for malaria surveillance. Nat Commun. 2022;13:1501.

6. Mwanga EP, Siria DJ, Mshani IH, Mwinyi SH, Abbasi S, Jimenez MG, et al. Rapid classification of epidemiologically relevant age categories of the malaria vector, Anopheles funestus. Parasites Vectors. 2024;17:143.

7. Jiménez MG, Babayan SA, Khazaeli P, Doyle M, Walton F, Reedy E, et al. Prediction of mosquito species and population age structure using mid-infrared spectroscopy and supervised machine learning. Wellcome Open Research [Internet]. 2019 [cited 2024 Sep 18];4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6753605/

**Tables**

**Figures**

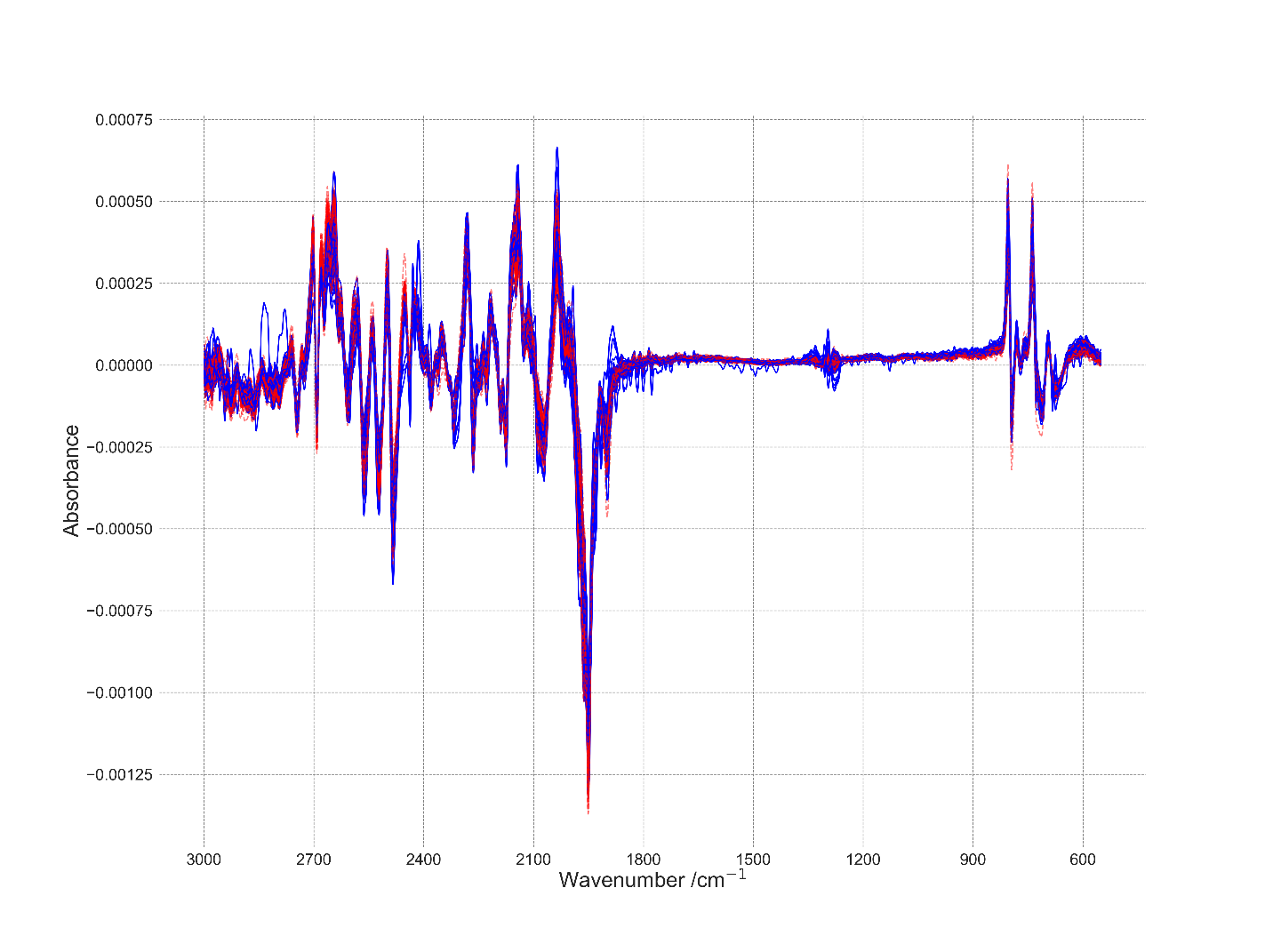
****

Figure 1. Absorbance

A screenshot of a computer

Description automatically generated

Figure 2. Entropy (log2) of the site-by-site probabilities for each spectrum. Lower entropy indicates greater informativeness.

A screenshot of a computer

Description automatically generated

Figure 3. Heatmap of the cross-correlation for spectrum with entropy (log2) < 0.9.

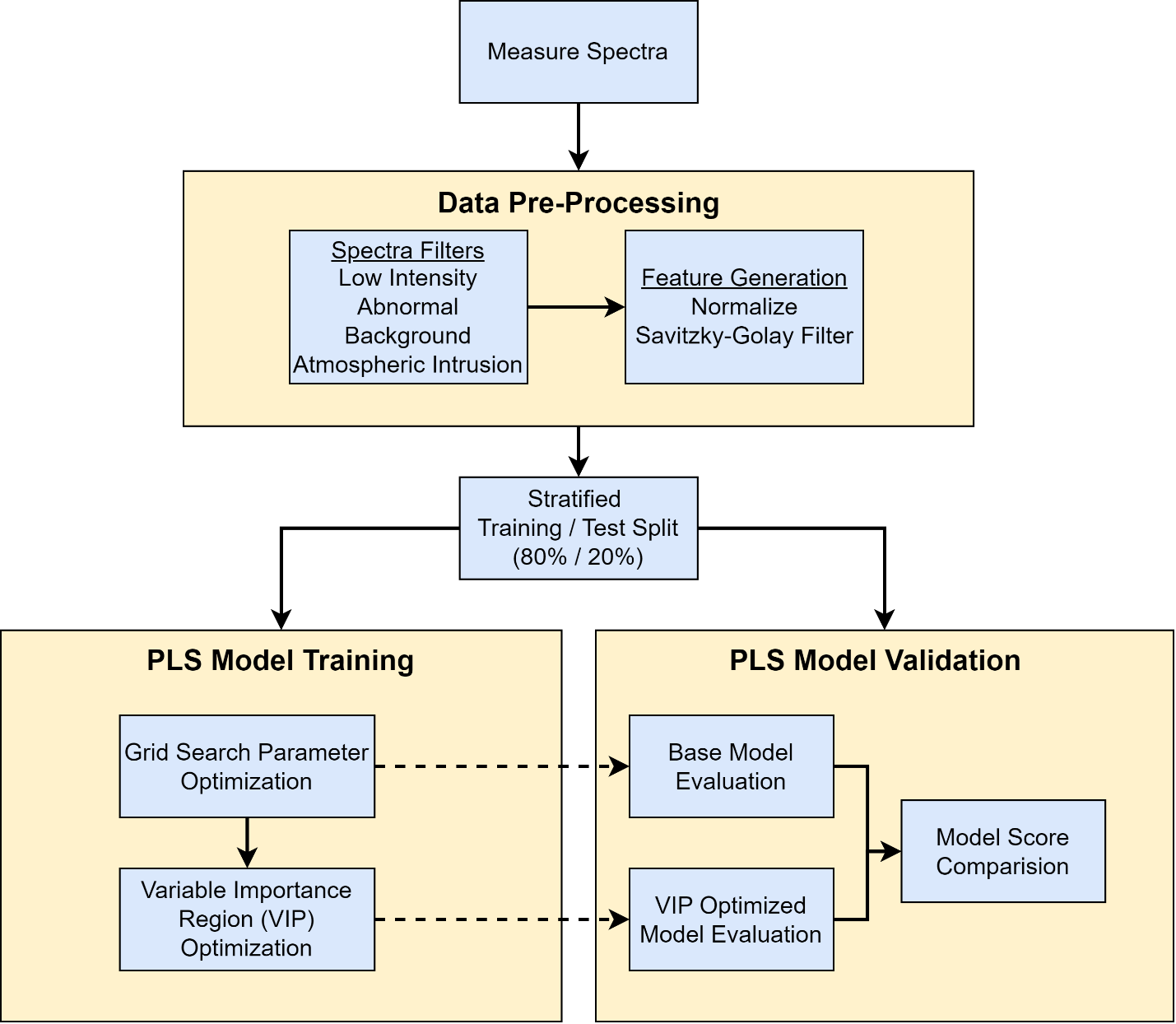


Figure XXX. Workflow diagram for PLS model training and evaluation.