Group E: Final Project Working Document

Simon and Emilie's notes from last week:

need to rewrite predictions lolz

big picture	small picture (snapshot)
more fed = more transmission more unfed = less transmission more gravid = more mosquito = more transmission	more fed = less transmission (cause they need time to digest) more unfed = they will bite so more transmission more gravid = more potential transmission (indicates that they had feeding activity)

Fed= more transmission (means that the climatic conditions are favourable for mosquitos to feed, so they are biting more, so more transmission)

Unfed= less transmission because less mosquitos are biting (indicated by the empty stomach) (because conditions may not be favourable to feed)

- unfed- no evidence for if it was biting people or not

Gravid= more transmission (have to have resources to take care of the eggs- population is more well fed)

Gravid or fed is evidence that it was engaging in feeding activity (biting activity and possible malaria transmission)

GAME PLAN:

Report Outline:

Dec 8th: 20%
* stylized as a journal article

* abstract
    - will write last
    - should be about ~150-200 words

* introduction
    - ~600 words

Inverted Triangle Introduction
- start broad (general ecological concepts) and narrow in on specific research (like specific species etc)
- citation style does not have to be anything specific (but has to stay consistent- ecology preferred)
- **make clear what the gap in knowledge is** (in our case it has something to do with the gravid state- have to do a bit more research as to what this means and how it is biologically relevant)
- is there a logical flow- things you cite should be **relevant** to understand your hypotheses and predictions
- Cite research that was done in the same region and how it is relevant to malaria transmission
- don't have to formulate the null and alternative (like the mid project update), just state the alternative (was only stated like that to help us in statistical analysis)

* methods
    - Can reuse and clean up methods from mid project report (should be about a page, ~400-500 words)
    - have to to cite any packages that we use for statistical analysis in our methods section

* results
    - ~500 words
    - Report what the results are for each test (this can be done by each of us individually- whichever test we do we will write up the results section for it)
    - NO INTERPRETATION YET

How to report test statistics

- summary () can help report t-test and ANOVA
- Model selection: create a table like her's
- Have to give the table a title (goes above the table) (a description as to what the table describes)
- figure captions have to be very detailed
- have to say if there is an error bar, standard error, confidence interval etc. if we are doing a boxplot

* discussion
  - ~750 words

Triangle Discussions
- say if hypothesis is supported or refuted
- this is where you interpret your results (in the results you are just STATING what the patterns and results are, no interpretation as to what biologically this means)
- here is where we get to try and interpret biologically what this means

* references
  - doesn't matter what style (let's go with CSE since that is what we use in all our bio classes)
  - not included in word count

* supplementary material - reproducible code
  - have to upload everything to GitHub

~ 2500 words, 6-8 figures/tables

Presentation Outline

December 6th (10%)

* 10 minute presentation, 2 min of questions
* content (background and methods) (3)
* content (results and conclusion) (3)
* delivery (3)

Presentation Tips
- title should be broad and catchy (key takeaway)
- use slide numbers

- 10 minute presentation is about 10 slides
- must define technical terms (even if they seem simple), can do this verbally
- dont put everything on the slide
- dont say you "proved" anything, just results were supported or not
- Key findings can be cited in the slides (Authors, Year, Journal) can go in the bottom right slide, but we do not HAVE to cite everything in the presentation
- end with a summary slide (can end with an important figure)
- Include as little text as possible
- "I dont have the answer to this, but I think"
- Copy paste with new things that will appear (can add way more slides), is helpful with figures- slowly build the figure up (can be really helpful when describing what it shows) in this case we would have way more than 10 slides, but still 10 slides worth of content
- describe the axes
- do not put tables up they are harder to read (draw attention as to what is important- circles, arrows)- if you are doing model selection than this table should be shown

- make sure to annotate code as we go (helps us understand what we wrote and why we did things the way we did)

Splitting up the work:

Abstract (to be written last): Simon (Sarah can help edit)
Introduction: Julia + Sarah **(Emilie will write hypotheses +predictions)**
Methods: Emilie will clean them up
Results: Test results will be explained by the person performing the test (Julia/Simon/Emilie)
Discussion: Sarah Can be similar to the results, but we should have a dedicated person editing this to streamline it

- Everyone will make bullet point notes on what they are trying to covey (their interpretation of the tests)

Figures/tables: Sarah
GitHib: My guy simon

Regression 1 (Hypothesis 1):

- Julia

Regression 2 (Hypothesis 2):

- Simon

Mixed model (hypothesis 2/Extra):

- Emilie will do
- Will probably have a plot and a table associated with it

Figures/graphs/tables 6-8

- Sarah
- look at presentation slides from last week about graph tables figures captions requirements
- **2 plots and two tables from Em,**
- Make a few ggplots based on the results, but choose also based on relevant gaps in knowledge

Timeline

Tuesday:
- preliminary statistical analysis
- planning

Thursday:
- continue working on stats
- Methods should be cleaned up and drafted
- rough draft (bullet points) of the introduction

Saturday/Sunday:
- rough draft (bullet points) of the introduction (sat)
- Code should all be finished
- Methods are finalized
- Results should be written
- Discussion should be drafted
- Things should be on slides!!

Monday
- finalize presentation
- Monday at 4pm we will practice it and edit anything if necessary

Tuesday
- presentation

Wednesday
- edit and finalize the final report - SHOULD BE FINISHED BY THIS DAY!!!

Thursday
- Submit on GitHub

Emilie

Methods Draft:

Field Work

The research was conducted and data were collected in the Kilombero Valley, a global malaria hotspot in Tanzania. Since 2007, entomologists from the University of Glasgow have trapped malaria vectors from four households across four villages (KID for Kidugalo, MIN for Minepa, LUP for Lupiro and SAG for Sagamaganga) within the valley (note: vectors were trapped using MET and CDC trap designs). This dataset ranges from June 2016 to September 2017. The individual trapped mosquitoes were sexed, identified as *An. gambiae* or *An. funestus*, and classified as either fed, unfed and gravid. A number of climatic factors --including maximum, minimum, and mean humidity and temperature-- were tracked daily at each vector trapping site. Standard CDC traps were left overnight inside the households within each village. Each village was visited at least 5 days per month over the project.

This dataset was collected by the University of Glasgow, in collaboration with the NERC Environmental Information Data Centre; this was not open-source community data. Due to sampling difficulties with the MET equipment, the CDC trap dataset was used for analysis. Six rows of data consisted of NAs due to house changes or issues encountered with the traps overnight. These rows were removed as they were missing abundance values. As a part of the data manipulation, we removed the columns that we were not investigating. Agml (number of male *An. gambiae)*, Afml (number of male *An. funestus)* were removed because we are investigating female mosquitoes since they are the only ones who can transmit malaria. We deleted the columns Acml (number of male *An.coustani*) and Acfem (number of female *An.coustani*) since there was no data on their fed, unfed or gravid states. OthAn (number of other *Anopheles species)* was also removed because we are specifically looking at two distinct species. MaxTemp, MaxHum, MinTemp, and MinHum were all removed because we will be using the MeanTemp and MeanHum to investigate our hypotheses.

The first step in our analysis plan was to separate the species and status (fed, unfed or gravid) fixed effects as they were originally reported together. To test hypothesis one, linear regressions were run to investigate if mean temperature and mean humidity differently affect the abundance of mosquitos in fed, unfed and gravid states. Test for assumptions, talk about any transformations that were made to the data? Linear models were made … To test hypothesis two, linear regressions were run to take species into consideration to see if there is a difference in abundances of the mosquito status between species *An. gambiae* and *An. funestus*. This is nested data, with the random effects of village and households. We chose to look at the random effect of villages opposed to households because xyz. Mixed effect models were used to compare if mosquito abundances differ by villages. An AICc test was run to determine the best model across all linear and mixed models to investigate xyz

DO NOT DELETE OR RESOLVE ANY OF MY COMMENTS PLEASE.

some code interrogations (Simon)

- just to be 100% sure: can you suppress column in a dataset (I thought we always were obliged to create a new dataset when deleting stuff)
- to make things more understandable for analysis 2, maybe plot differences between the 2 species in some ways(which species prefers which temperatures etc..
- je pense qu'il faut faire croire qu'on à bosser dans github donc on va tous faire des edits et pull push request dans github pour montrer qu'on sait l'utiliser

Still we need to remember:

- Put a correct code intro (that can work each time)
- Annotate the code better, make it clean
- Are we putting our name on the part of code we were assigned?
- Add Sarah plots
- (maybe add in the end of the report or conclusion that a next step to add could be the mapping of areas mosquito density in each villages)

Methods Final Result:

Field Work

Data were collected in the Kilombero Valley, a global malaria hotspot in Tanzania. Since 2007, entomologists from the University of Glasgow have trapped malaria vectors from four households across four villages (KID for Kidugalo, MIN for Minepa, LUP for Lupiro and SAG for Sagamaganga) within the valley (see Figure 1). Vectors were trapped using MET and CDC trap designs. The data collected ranges from June 2016 to September 2017. The individual trapped mosquitoes were sexed, identified as *An. gambiae* or *An. funestus*, and classified into three mutually exclusive states: fed, unfed, and gravid. A number of climatic factors including mean temperature and humidity were recorded daily at each vector trapping site. Standard CDC traps were left overnight inside the households within each village. Each household within each village was visited at least 5 days per month over the project.

Data Manipulation

This dataset was collected by the University of Glasgow, in collaboration with the NERC Environmental Information Data Centre; this was not open-source community data. Due to sampling difficulties with the MET equipment, the CDC trap dataset was used for analysis. Six rows of data consisted of NAs due to house changes or issues encountered with the traps overnight. These rows were removed as they were missing abundance values.Columns regarding information on abundance of male mosquitoes were removed because female mosquitoes are only able to transmit malaria. We separated the species and status (fed, unfed or gravid) as they were originally reported together.

Data Analysis

Version 4.2.1 of R and R Studio was used for all data analysis. The packages lme4 (Bates et al., 2015), lmerTest (Kuznetsova et al., 2017), MuMIn (Barton, 2020), sjmisc (Lüdecke, 2018), tidyverse (Wickham et al., 2019) and lsmeans (Lenth, 2016) were used in our analysis. To test hypothesis one, we generated a linear model to investigate if mean temperature and mean humidity differently affect the abundance of mosquitoes in fed, unfed and gravid states. Assumptions of homogeneity and normality were met. To test hypothesis two, another linear model was generated to take species into consideration to see if there is a difference in abundances of the mosquito status between species *An. gambiae* and *An. funestus*. These candidate linear models included the same fixed effects as the models created for hypothesis one, but species was added as a fixed effect.

This is nested data, which has random effects of village and household (see Figure 1). We chose to look at the random effect of villages opposed to households because mosquitos may go from one village to another because of biological factors such as breeding location proximity (Al-Thukair et al., 2022). Mixed effect models were used to compare if mosquito abundances differ by villages. In all models (linear and mixed), we chose to look at the interaction between mean temperature and mean humidity because we believed that differences captured in humidity and temperature are biologically important. We found they were highly correlated, so they were systematically dropped as models were being created. In all models we kept status as a fixed effect as we are specifically investigating the abundance of mosquitoes in the three different states. We used an AIC (Akaike, 1974) to rank our twelve candidate models and identified the top model based on the AICc (Burnham & Anderson, 2002) value to investigate which predictors explain the variation in abundance in mosquitos. We considered results significant if $p \leq 0.05$.

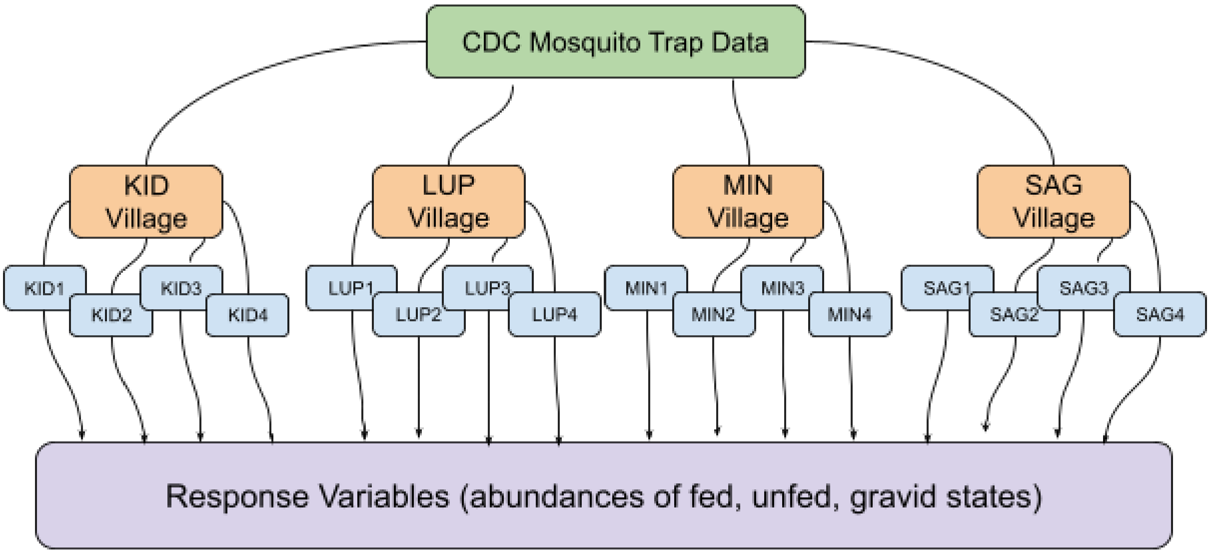

Figure 1. Method collection design using CDC traps in four villages, with four specific households within each village. Abundances of fed, unfed, and gravid mosquitoes were recorded, as well as their species and sex.

Emilie's Citations

AIC:

Akaike, H. 1974. "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* 19: 716–23.

AICc:

Burnham, K. P., and D. R. Anderson. 2002. *Model Selection and Multimodel Inference*. New York, NY: Springer.

MuMIN:

Barton, K. 2020. " Multi-Model Inference." R Package Version 1.43.17. https://cran.r-project.org/web/packages/MuMIn/MuMIn.pdf.

lme4:

Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, **67**(1), 1–48. doi:10.18637/jss.v067.i01.

lmerTest:

Kuznetsova A, Brockhoff PB, Christensen RHB (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, **82**(13), 1–26. doi:10.18637/jss.v082.i13.

sjmisc:

Lüdecke, D. (2018). sjmisc: Data and Variable Transformation Functions. *Journal of Open Source Software*, *3*(26), 754.

Tidyverse

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to the tidyverse." *Journal of Open Source Software*, **4**(43), 1686. doi:10.21105/joss.01686.

lsmeans

Lenth RV (2016). "Least-Squares Means: The R Package lsmeans." *Journal of Statistical Software*, **69**(1), 1–33. doi:10.18637/jss.v069.i01.

(Al-Thukair et al., 2022) (Hypotheses and predictions section)

https://www.intechopen.com/online-first/82058

DOI: 10.5772/intechopen.104615

Abstract:

Context, method, results, discussion, conclusion, purpose

Malaria prevention is based on studies on its vectors. Understanding factors influencing mosquito bite rates is one of the interesting approaches to this topic. In four villages of the Kilombero Valley in Tanzania, mosquitoes were trapped, identified, sexed, and classified as fed, unfed or gravid. At each trapping site, humidity and temperature data were daily collected. This paper intends to investigate the effect of temperature and humidity on mosquito abundances in this African Sub-Saharan region. Use of statistical linear model and mixed effect model with R and RStudio allowed to interpret the data. The best fitted model was the saturated mixed model, considering the effect of both temperature and humidity on abundance, while taking into account status (fed, unfed, gravid) and species (*An. Funestus* and *An. Gambiae*) as well as adding the villages as random effect to correct for the nested data. Results showed temperature and abundance as well as their interaction to have a significant effect on mosquito abundances for the fed and unfed states but not for the gravid state. (Are we saying why we think this is due to?). Results also reveal a significantly higher abundance of *An. Gambiae* in each state recorded (Are we also saying why, the dependance to climatic shift etc..?). These results, further enhanced by other studies assessing the directionality of the interactions, other climatic factors, or other sampling locations, can help predict mosquito feeding habits and malaria transmission rates.

Discussion Topics (For Sarah - working on a diff doc bc no wifi while out)

- Talk about what the differences in states means in terms of malaria transmission in terms of humans
- Talk about differences in species
- Talk about models and if they make sense biologically
- Limitations to our models
- Possible future work

Notes from class: Triangle Discussion (FOLLOW THIS! Have to state if we reject or fail to reject our null hypotheses.)

- say if hypothesis is supported or refuted
- this is where you interpret your results (in the results you are just STATING what the patterns and results are, no interpretation as to what biologically this means)
- here is where we get to try and interpret biologically what this means

Julia's Introduction Draft

Version 1

Responsible for hundreds of thousands of deaths world-wide, malaria is a devastating disease that predominantly impacts children. Caused by parasites of the *Plasmodium* genus, malaria is transmitted by female mosquitoes of the genus *Anopheles* by biting humans while feeding (WHO 2022). Since the *Plasmodium* parasite requires 7 to 30 days to develop and become mature enough to infect humans (CDC 2022), it is typically only older mosquitoes that transfer the disease. Hundreds of species in the genus *Anopheles* are capable of transmitting malaria, and so the key malaria vector tends to vary with different geographical regions, as well as different factors such as temperature, humidity, and altitude (Daygena et al. 2017). Some of the key species involved in transmission include *An. arabiensis, An. gambiae,* and *An. funestus.* For the purposes of this paper, we will only investigate *An. gambiae* and *An. funestus.*

It is important to note that mosquitoes also possess a unique reproductive strategy, in which prior to oogenesis, a female mosquito will consume a large blood meal and digest it. The mosquito will then undergo gestation while digesting, and move very little during this time (OECD 2018). Thus, if a mosquito is gravid, this indicates that they consumed a blood meal at least somewhat recently.

In addition to previous research supporting a relationship between environmental factors and the key vector in a given area, studies demonstrate that there is also a correlation between temperature and humidity and whether a mosquito is fed or unfed (Agyekum et al. 2021). However, there is currently no research regarding whether temperature and/or humidity are correlated with whether or not a mosquito is gravid (carrying eggs). Thus, the purpose of this paper is to further investigate this relationship between climatic factors and gravidity.

Thus, we will investigate two alternative hypotheses regarding the abundance of mosquitoes in the Kilombero Valley in Tanzania. Hypothesis one posits that temperature and humidity have different effects on the abundances of mosquitos between the fed, unfed, and gravid states of mosquitos. Hypothesis two postulates that temperature and humidity have different effects on the abundances of the two different mosquitoes species (*An. gambiae* and *An. funestus*) between the fed, unfed and gravid states. We will also determine if villages have a statistically-significant random effect on mosquito abundance.

For hypothesis one, we predict that temperature and humidity will affect the fed and gravid abundances since prior research indicates that increased fluctuations in temperature and humidity results in more efficient digestion, causing more frequent feedings (Need a citation, emilie will find it). Literature also states that the abundance of *An. gambiae* is positively associated with moisture index and *An. fuestus* abundance is higher in the dry than rainy season in Sub-Saharan Africa (Minakawa et al., 2002). For hypothesis two, we predict that temperature and humidity will affect the abundance of *An. gambiae*, and will not affect the abundance of *An. funestus,* since *An. gambiae* abundance is known to rise dramatically in response to small temperature increases, likely due to accelerating the larval life stage and gonotrophic cycles (CITATION? Ask Sarah). Ultimately, this research can serve to help predict feeding habits of mosquitoes, allowing for better prevention measures against malaria.

OKAY VERSION WE HAND IN STARTS HERE

References

Afrane YA, Lawson BW, Githeko AK, Yan G. 2005. Effects of microclimatic changes caused by land use and land cover on duration of gonotrophic cycles of *anopheles gambiae*(diptera: Culicidae) in western Kenya Highlands. *Journal of Medical Entomology*. *42*(6): 974–980.

Agyekum TP, Botwe PK, Arko-Mensah J, Issah I, Acquah AA, Hogarh JN, Dwomoh D, Robins TG, Fobil JN. 2021. A Systematic Review of the Effects of Temperature on Anopheles Mosquito Development and Survival: Implications for Malaria Control in a Future Warmer Climate. *Int. J. Environ. Res. Public Health.* 18(14): 7255.

Akaike H. 1974. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control.* 19: 716–23.

Al-Thukair A, Jemal Y, Nzila A. 2022. Influence of Climatic Factors on the Abundance and Profusion of Mosquitoes in Eastern Province, Saudi Arabia. *In: Mosquito Research - Recent Advances in Pathogen Interactions, Immunity, and Vector Control Strategies. IntechOpen.*

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*. 67(1): 1–48.

Barton K. 2020. Multi-Model Inference. *R Package Version 1.43.17*. https://cran.r-project.org/web/packages/MuMIn/MuMIn.pdf.

Burnham KP & Anderson DR. 2002. *Model Selection and Multimodel Inference*. New York, NY: Springer.

[CDC]. Centers for Disease Control and Prevention (USA). 2022. About Malaria. Atlanta (USA): Centers for Disease Control and Prevention

Charlwood JD. 2017. Some like it hot: A differential response to changing temperatures by the malaria vectors *anopheles funestus* and *an. gambiae* s.l. *PeerJ*. 5.

Daygena TY, Massebo F, Lindtjørn B. 2017. Variation in Species Composition and Infection Rates of Anopheles Mosquitoes at Differential Altitudinal Transects and the Risk of Malaria in the Highland of Dirasge Woreda, South Ethiopia. *Parasites and Vectors*. 10(1): 343.

Day JF. 2016. Mosquito oviposition behavior and vector control. *Insects*. 7(4): 65.

Genoud AP, Gao Y, Williams GM, Thomas BP. 2019. Identification of gravid mosquitoes from changes in spectral and polarimetric backscatter cross sections. *Journal of Biophotonics*. 12(10).

Kelly-Hope LA, Hemingway J, McKenzie FE. 2009. Environmental factors associated with the malaria vectors Anopheles gambiae and Anopheles funestus in Kenya. *Malaria Journal*. 8(1): 268.

Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*. 82(13): 1–26.

Lenth RV. 2016. Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*. 69(1): 1–33.

Lüdecke, D. 2018. sjmisc: Data and Variable Transformation Functions. *Journal of Open Source Software*. 3(26): 754.

Minakawa N, Sonye G, Mogi M, Githeko A, Yan G. 2002. The Effects of Climatic Factors on the Distribution and Abundance of Malaria Vectors in Kenya. *Journal of Medical Entomology*. 39(6): 833–841.

Mwingira V, Mboera L, Dicke M, Takken W. 2020. Exploiting the chemical ecology of mosquito oviposition behavior in Mosquito Surveillance and control: A Review. *Journal of Vector Ecology*. *45*(2): 155–179.

Ndoen E, Wild C, Dale P, Sipe N, Dale M. 2010. Relationships between Anopheline Mosquitoes and topography in West Timor and Java, Indonesia. *Malaria Journal*. *9*(1).

[OECD] Organization for Economic Cooperation and Development (France). 2018. Safety Assessment of Transgenic Organisms in the Environment: Volume 8. Paris (France): Organization for Economic Cooperation and Development

Soleimani-Ahmadi M, Vatandoost H, Zare M, Turki H, & Alizadeh A. 2015. Topographical distribution of Anopheline Mosquitoes in an area under elimination programme in the south of Iran. *Malaria Journal*. *14*(1).

Suh E, Grossman MK, Waite JL, Dennington NL, Sherrard-Smith E, Churcher TS, Thomas MB. 2020. The influence of feeding behaviour and temperature on the capacity of mosquitoes to transmit malaria. *Nat Ecol Evol*. 4(7): 940-951.

[WHO] World Health Organization (Switzerland). 2022. Malaria. Geneva (Switzerland): World Health Organization.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the tidyverse. *Journal of Open Source Software*. 4(43): 1686.

Yan G, Githeko A, Zhou G, Minakawa N, Omukunda E. 2006. Malaria Vector Productivity in Relation to the Highland Environment in Kenya. *The American Journal of Tropical Medicine and Hygiene*. 75(3): 448–453.

Investigating the effects of temperature and humidity on mosquito abundances within Kilombero Valley, Tanzania

Julia Marques[1], Sarah Martinez-Fuller[1], Emilie Nero[1], Simon Sevenier[1]
[1]*Department of Ecology Evolutionary Biology, University of Toronto*

EEB313, December 8th, 2022

Abstract

Malaria prevention is based on research performed on the vectors that transmit it-mosquitoes. Environmental factors that influence mosquito biting rates and abundances can help understand population control, which in turn can help control malaria transmission. Within four villages of the Kilombero Valley in Tanzania, mosquitoes were trapped, identified, sexed, and classified as fed, unfed, or gravid. At each trapping site, humidity and temperature data were daily collected. This paper intends to investigate the effects of temperature and humidity on mosquito abundances in this Sub-Saharan African region. Linear models and mixed effect models were generated using R and RStudio to look at which predictors captured variation within abundance data. The best fit model is a saturated mixed model, which considers the effect of both temperature and humidity on abundance, while taking into account status (fed, unfed, gravid), species (*An. funestus* and *An. gambiae*) and village as a random effect. Results show the interaction of temperature and humidity to have a significant effect on mosquito abundances for the fed and unfed states but not for the gravid state. These results, which can be further enhanced by other studies assessing the directionality of the interactions or other climatic factors can help predict mosquito feeding habits and malaria transmission rates.

Introduction

Malaria is a significant disease of global concern in many underdeveloped regions of the world, as it kills approximately 429,000 people annually (WHO, 2016). Malaria is a disease caused by parasites of the *Plasmodium* genus, which is transmitted through vectors. Vectors are living organisms that transmit infectious agents, such as parasites, to another animal or human. Malaria can only be transmitted by female mosquitoes of the genus *Anopheles* by biting humans while feeding (WHO, 2022). Since the *Plasmodium* parasite requires 7 to 30 days to develop and become mature enough to infect humans (WHO, 2022), it is typically only older mosquitoes that transfer the disease. Hundreds of species in the genus *Anopheles* are capable of transmitting malaria, and so the key malaria vector tends to vary with different geographical regions, as well as different factors such as temperature, humidity, and altitude (Daygena et al., 2017). The key species involved in transmission include *An. arabiensis, An. gambiae,* and *An. funestus* (Kelly-Hope et al., 2009).

Mosquitoes possess a unique reproductive strategy in which prior to oogenesis; a female mosquito will consume a large blood meal and digest it (Genoud et al., 2019). The mosquito will then undergo gestation while digesting, and move very little during this time (OECD, 2018). Thus, if a mosquito is gravid, this indicates that they consumed a blood meal at least within the past 69 to 97 hours (Genoud et al., 2019). Examining the abundances of gravid mosquitoes and climatic factors can provide information of when new mosquitos are likely to become active which can help anticipate a risk for exposure to malaria (Genoud et al., 2019). Previous research has been conducted on the biting behaviours in response to environmental conditions such as temperature where they found lower vector productivity in regions of lower temperature (Yan et al., 2006). Environmental factors such as temperature and humidity are important to understand as they can create suitable breeding sites for mosquitoes, which then affect their abundance and biting patterns (Kelly-Hope et al., 2009).

Studies have demonstrated that there are correlations between temperature and humidity and whether a mosquito is fed or unfed (Agyekum et al., 2021). However, there is currently no research regarding whether temperature and/or humidity affect gravid (carrying eggs) abundances. The purpose of this paper is to further explore the relationship between climatic factors and gravidity within the *An. gambiae* and *An. funestus* species.

Thus, we will investigate two alternative hypotheses regarding the abundance of mosquitoes in the Kilombero Valley in Tanzania. Hypothesis one posits that temperature and humidity have different effects on the abundances of mosquitos between the fed, unfed, and gravid states of mosquitos. This will allow us to investigate if temperature and humidity affect biting behaviour by looking at the abundance differences in fed, unfed or gravid states. Hypothesis two postulates that temperature and humidity have different effects on the abundances of the two different mosquitoes species (*An. gambiae* and *An. funestus*) between the fed, unfed and gravid states. We will also determine if villages have a statistically-significant random effect on mosquito abundance.

For hypothesis one, we predict that temperature and humidity will affect the fed and gravid abundances since prior research indicates that increased fluctuations in temperature and humidity results in more efficient digestion, causing more frequent feedings (Suh et al., 2020. Literature also states that the abundance of *An. gambiae* is positively associated with moisture index and *An. funestus* abundance is higher in the dry than rainy season in Sub-Saharan Africa (Minakawa et al., 2002). For hypothesis two, we predict that temperature and humidity will affect the abundance of *An. gambiae*, and will not affect the abundance of *An. funestus,* since *An. gambiae* abundance is known to rise dramatically in response to small temperature increases, likely due to accelerating the larval life stage and gonotrophic cycles (Charlwood, 2017). Ultimately, this research can serve to help predict biting behaviours of mosquitoes, allowing for better prevention measures against malaria.

Methods

Field Work

The data was collected in the Kilombero Valley, a global malaria hotspot in Tanzania. Since 2007, entomologists from the University of Glasgow have trapped malaria vectors from four households across four villages (KID for Kidugalo, MIN for Minepa, LUP for Lupiro and SAG for Sagamaganga) within the valley (see Figure 1). Vectors were trapped using MET and CDC trap designs. The data collected ranges from June 2016 to September 2017. The individual trapped mosquitoes were sexed, identified as *An. gambiae* or *An. funestus*, and classified into three states: fed, unfed, and gravid. A number of climatic factors including mean temperature and humidity were recorded daily at each vector trapping site. Standard CDC traps were left overnight inside the households within each village. Each household within each village was visited at least 5 days per month over the project.

Data Manipulation

This dataset was collected by the University of Glasgow, in collaboration with the NERC Environmental Information Data Centre; this was not open-source community data. Due to sampling difficulties with the MET equipment, the CDC (Centre for Disease Control) trap dataset was used for analysis. Six rows of data consisted of NAs due to house changes or issues encountered with the traps overnight. These rows were removed as they were missing abundance values. Columns regarding information on abundance of male mosquitoes were removed as only female mosquitoes are able to transmit malaria, making them the population of interest in this case. The dataset was also manipulated to generate 2 new columns, one for species and one for status, in order to make statistical tests possible to run. The data originally listed abundance based on species and status as one variable.

Data Analysis

Version 4.2.1 of R and R Studio was used for all data analysis. The packages lme4 (Bates et al., 2015), lmerTest (Kuznetsova et al., 2017), MuMIn (Barton, 2020), sjmisc (Lüdecke, 2018), tidyverse (Wickham et al., 2019) and lsmeans (Lenth, 2016) were used in our analysis. To test hypothesis one, we generated a linear model to investigate if mean temperature and mean humidity differently affect the abundance of mosquitoes in fed, unfed and gravid states. Assumptions of homogeneity and normality were met. To test hypothesis two, another linear model was generated to take species into consideration to see if there is a difference in abundances of the mosquito status between species *An. gambiae* and *An. funestus*. These candidate linear models included the same fixed effects as the models created for hypothesis one, but species was added as a fixed effect.

This is nested data, which includes random effects of both village and household (see Figure 1). We chose to focus on the random effect of villages instead of the random effects of households as mosquitos may travel between villages due to biological factors such as breeding location proximity (Al-Thukair et al., 2022). Mixed effect models also took into account if mosquito abundances differ by villages. In all models (linear and mixed), we chose to look at the interaction between mean temperature and mean humidity because we believed that differences captured in humidity and temperature are biologically important. Temperature and humidity were found to be highly correlated, and so were systematically dropped during model generation. In all models we kept status as a fixed effect, due to the specific investigation of the abundances of mosquitoes across the three different states. We used an AICc (Akaike, 1974) to rank our twelve candidate models and identified the top model based on the AICc value (Burnham & Anderson, 2002) to investigate which predictors explain the variation in abundance in mosquitos. Results were considered significant if $p \leq 0.05$.

Results

The first linear model considers the effects of mean temperature and mean humidity on abundance, while taking into account status (fed, unfed, and gravid). The following linear model was fit: abundance ~ MeanTemp*MeanHum + status. Mean temperature and mean humidity were systematically dropped to create two more linear models. The results of the model (see Table 1) find that the interaction between mean temperature and mean humidity have a significant effect on fed and unfed states (t = 14.5, p = < 0.001). This model also finds that there is not a significant interaction between mean temperature and humidity on the gravid state (t = -0.43, p = 0.666). Mean temperature (t = -3.43, p = < 0.001) and mean humidity (t = -2.92, p = < 0.001) also has a significant effect on mosquito abundance. This linear model consists of: F(5, 5070 residuals) = 61.19, p = < 0.001. Across the timespan of the data collection, unfed mosquitoes were more abundant than the gravid or fed (see Figure 2).

The second linear model intends to observe differences between the two mosquito species *An. gambiae* and *An. funestus*. The following linear model was fit: abundance ~ MeanTemp*MeanHum + status + species. Similar to the first linear model investigating status abundance, the results show (see Table 2) the interaction of mean temperature and mean humidity to also have a significant effect on abundance between the fed and unfed state (t = 14.5, p = < 0.001) as well as a non-significant interaction affecting gravid abundance (t = -0.43, p = 0.666). We find a significant effect between species on mosquitos abundance (t = 5.06, p = < 0.001). This linear model consists of: F(6, 5069 residuals) = 55.51, p = < 0.001. For all three states, mosquitoes of the *An. gambiae* species had a higher mean abundance than mosquitoes of the *An. funestus* species (see Figure 3).

Six mixed models were generated, systematically dropping each fixed effect except status. The saturated model fit all predictors: abundance ~ MeanTemp*MeanHum + status + species + (1| village) (see Table 3). Similar to both linear model, we find the interaction of mean temperature and mean humidity to also have a significant effect on abundance between the fed and unfed state (t = 13.64, p = <0.001**)** as well as a non-significant interaction affecting gravid abundance (t = -0.44, p = 0.663). We find a significant effect between species on mosquitos abundance (t = 5.06, p = < 0.001). We also find a strong correlation between mean temperature and humidity (0.98). Village (n=4) accounts for 11.65 units of the variance, with a standard deviation of 3.41.

Model selection was done using an AICc on all models, and the best performing model is the saturated mixed model: abundance ~ MeanTemp*MeanHum + status + species + (1| village) with an AICc of 48618.85 (see Table 4). All models which include species as a predictor perform better than models without species as a predictor.

Discussion

Hypothesis one postulates that temperature and humidity influence the abundances of fed-, unfed-, and gravid-state mosquitoes. This hypothesis is supported by prior research indicating that increased fluctuations in temperature and humidity results in more efficient digestion, causing more frequent feedings (Day, 2016; Daygena et al., 2017). Despite this, our results failed to reject the null hypothesis. While the results align with the known relationship between temperature and humidity and whether a mosquito is fed or unfed (Agyekum et al., 2021), we found no evidence for a relationship between temperature and humidity and gravid-state mosquitoes. This has interesting epidemiological implications, as it suggests that gravidity may not be an important factor to consider when looking at biting rates and malaria transmission when considering shifts in climatic factors, despite the state duration being temperature-dependent (Day, 2016).

As for hypothesis two, we postulated that climate-caused effects on state abundance are species specific. This too is supported by literature, specifically, that *An. gambiae* is positively associated with moisture index (recall: likely due to accelerated larval life stage and gonotrophic cycles) and *An. funestus* with higher abundances in the Sub-Saharan dry season (Minakawa et al., 2002). While we did expect interaction of mean temperature and mean humidity on fed versus unfed state abundances, our results found no significant interaction between the climatic factors and gravid abundance from any of the models (linear or mixed) created. The effects of climate on both species overall abundance as well as their fed and unfed abundances is unsurprising given *An. gambiae*'s success in moist climates and previous literature (Minakawa et al., 2002). For each state, *An. gambiae* had significantly higher abundances than *An. funestus*. Given the results of testing hypothesis one, it is unsurprising that species-specific gravidity is unaffected by mean temperature and humidity.

Conclusion

Limitations and Future Investigations

Mosquitoes are very particular when it comes to selecting a location for oviposition (Kelly-Hope et al., 2009), which often means oviposition is delayed. Thus, females may remain gravid for longer periods of time than what is developmentally necessary if they are not in a suitable environment (Day, 2016). For this reason, gravid mosquito abundance may somewhat reflect the suitability of the environment for oviposition, rather than direct climatic effects on gravid abundance. This is worth noting when considering our results, because although we assessed the effects of temperature and humidity on the abundances of gravid individuals, we could not break down the gravid category by the time spent in that state because that was not part of the dataset. Thus, we cannot be sure that there are no indirect climatic effects.

This unexplored factor has important biological and epidemiological relevance because the longer oviposition is delayed, the more likely that an infected mosquito will transmit the pathogen to the next generation, thus rapidly expanding the pathogen-carrying population (Day, 2016). Such delays are due to the time it takes for the pathogen to mature, replicate, and invade the saliva and tissues being temperature-dependent, so if the environmental and climatic conditions are non-optimal for oviposition, oviposition will be delayed, and the more likely the next generation will be born already carrying the pathogen (Day, 2016). Therefore, there is a chance that there exists a relationship between the abundance of gravid-state mosquitos and the climatic factors of interest, but it is obscured by the variable time spent gravid. Investigating how to further break down the gravid state category would make possible more thorough research.

In terms of epidemiological relevance, research like this is essential to expanding our understanding of vector behaviour and biology so that these characteristics can be exploited for affordable and more widely generalizable disease control strategies. Specifically, understanding gravidity, oviposition, and other reproductively relevant processes and behaviours is imperative to the development of improved surveillance and control of malaria vectors (Mwingira et al., 2020). This study further developed knowledge of the relationship between climate conditions --temperature and humidity-- and mosquito life-states, which is crucial to the future of malaria vector control efforts in an ever-warming environment.

A major limitation of our study is directionality (increase or decreases) on abundances. Further studies are required to determine whether abundances of these states increase or decrease based on climatic factors such as temperature and humidity. Also, these results may not be generalizable to non-malaria hotspots or differing geographical and environmental conditions. For example, topographical barriers such as hills, rivers, and valleys may impact mosquito population movement and feeding behaviour (Ndoen et al., 2010; Soleimani-Ahmadi et al., 2015; Afrane et al., 2005). This certainly impacted our data given that the study site is a valley. Thus distribution of states (fed, unfed, and of most interest, gravid) and consequent malaria transmission rates may be more topologically relevant than we accounted for in our analysis. In this way, future research investigating topological distributions of mosquito fed, unfed, and gravid states would be valuable.

Appendix A: Figures and Tables

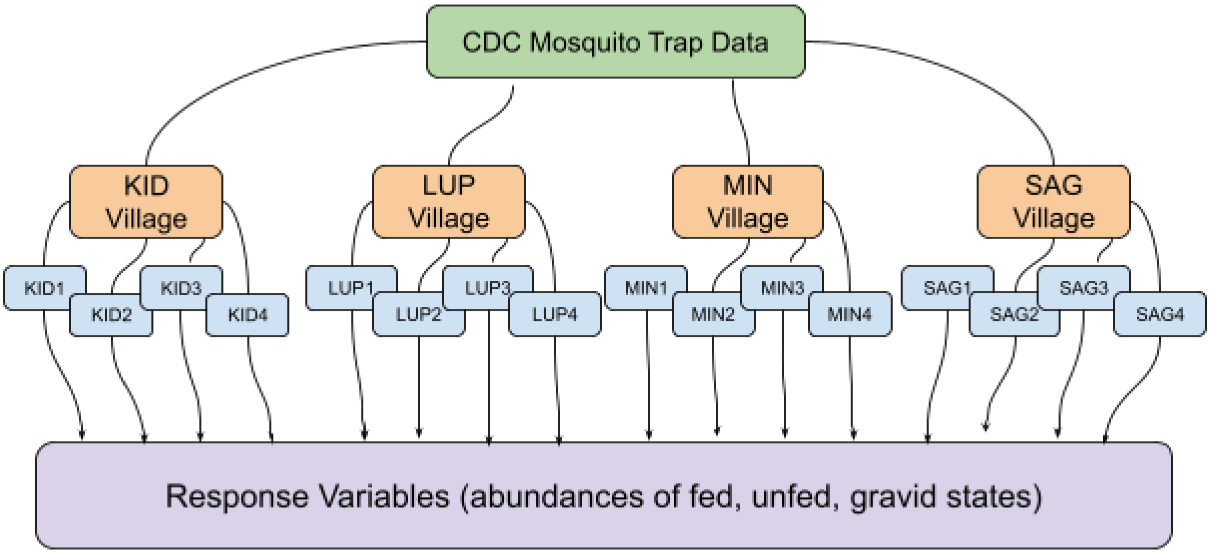

Figure 1. Nested method collection design using CDC traps from four villages, with four specific households within each village. Abundances of fed, unfed, and gravid mosquitoes were recorded, as well as their species and sex.

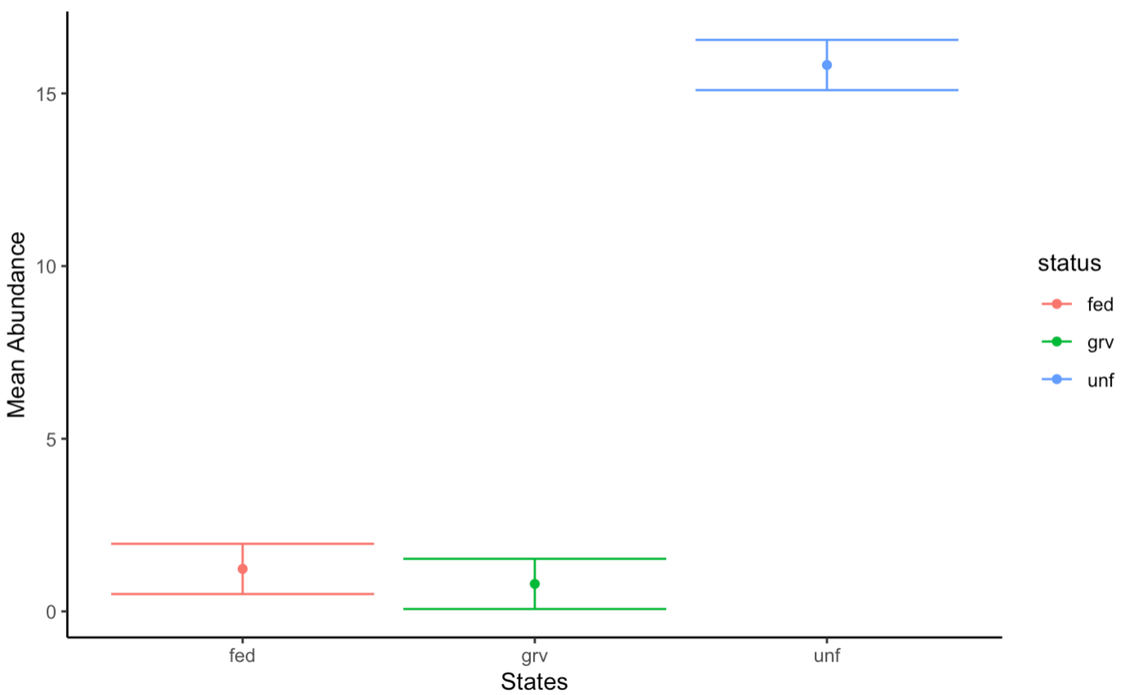

Figure 2. Mean abundance of mosquitoes separated into the three states. The error bars represent one standard error. We found that unfed mosquitoes were the most abundant, followed by fed and then gravid mosquitoes. There was a significant effect of mean temperature and humidity on the fed and unfed states, but an insignificant effect on the gravid state.

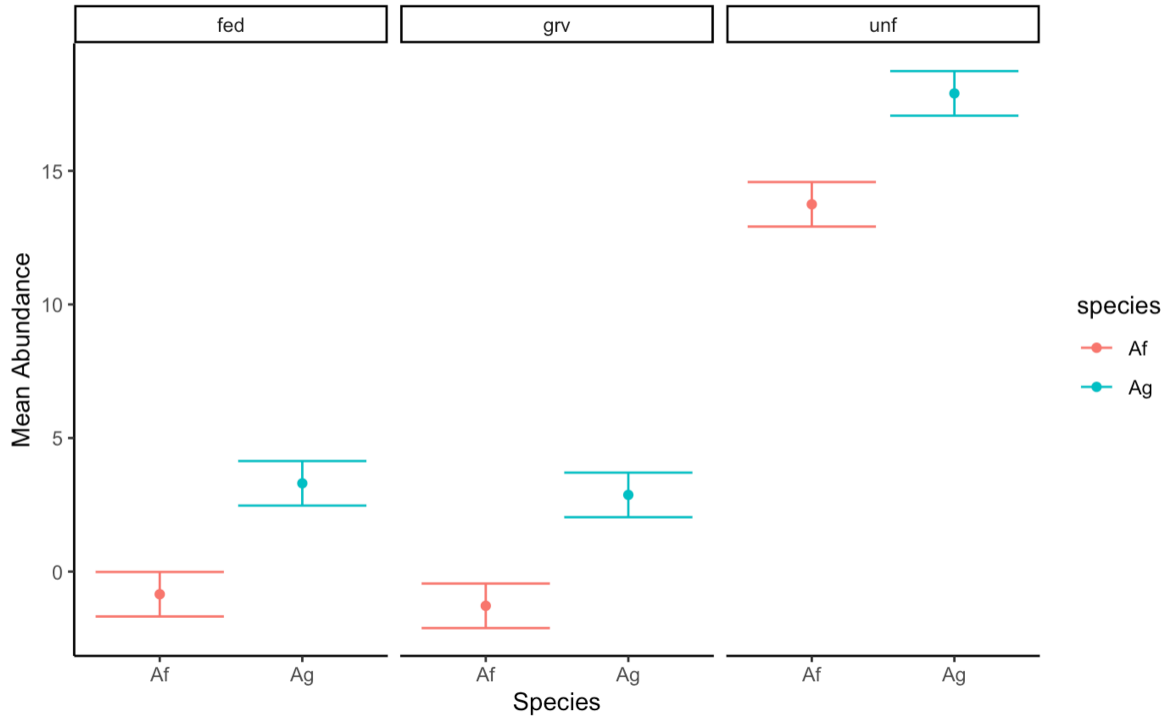

Figure 3. Mean abundance of mosquitoes across the two species separated into the three states. Af represents *An. Funestus* and Ag represents *An. gambiae*. The error bars represent one standard error. We found that unfed mosquitoes were the most abundant across both species, and that the abundance of gravid and fed mosquitoes were similar. There was a significant effect of mean temperature and humidity on the both species.

Table 1. Results from a linear model of the effect of temperature and humidity on mosquito status abundance.

	Estimate	Std error	*t*-value	*P*
Intercept	65.62	18.72	3.504	**<0.001**
Mean Temp	-2.44	0.71	-3.43	**<0.001**
Mean Hum	-0.82	0.28	-2.92	**0.003**
Gravid Status	-0.43	1.01	-0.43	0.666
Unfed Status	14.59	0.996	13.64	**<0.001**

Mean Temp*Mean Hum	0.03	0.01	2.86	**0.004**

Statistically significant results are in bold.

Table 2. Results from a linear model of the effect of temperature and humidity on mosquito status and species *An. gambiae* and *An. funestus* abundance.

	Estimate	Std error	*t*-value	*P*
Intercept	65.62	18.72	3.504	**<0.001**
Mean Temp	-2.44	0.71	-3.43	**<0.001**
Mean Hum	-0.82	0.28	-2.92	**0.003**
An. gambiae	4.15	0.82	5.06	**<0.001**
Gravid Status	-0.43	1.01	-0.43	0.666
Unfed Status	14.59	0.996	13.64	**<0.001**
Mean Temp*Mean Hum	0.03	0.01	2.86	**0.004**

Statistically significant results are in bold.

Table 3. Results from a mixed effects model on the effect of temperature and humidity on mosquito status, species *An. gambiae* and *An. funestus* abundance, using village as a random effect.

	Estimate	Std error	*t*-value	*P*
Intercept	49.19	18.76	2.62	**0.008**
Mean Temp	-1.87	0.71	-2.62	**0.008**
Mean Hum	-0.61	0.28	-2.197	**0.028**
An. gambiae	4.15	0.81	5.10	**<0.001**
Gravid Status	-0.43	0.996	-0.435	0.663
Unfed Status	14.59	0.996	13.64	**<0.001**
Mean Temp*Mean Hum	0.02	0.01	2.12	**0.034**

Statistically significant results are in bold.

Table 4. Model selection of the best fit model on the effect of temperature and humidity on mosquito abundance using all linear and mixed models.

Model	**AICc**	**ΔAICc**	***df***
abundance~MeanTemp*MeanHum + species + status + (1|village)	48618.85	0	9
abundance~MeanTemp + status + species + (1|village)	48620.20	1.35	7
abundance~MeanTemp+MeanHum + species + status + (1|village)	48621.31	2.46	8
abundance~MeanHum + status + species + (1|village)	48626.18	7.33	7
abundance~MeanTemp*MeanHum + status + (1|village)	48642.78	23.93	8
abundance~status + (1|village)	48648.08	29.23	5
abundance~MeanTemp*MeanHum + species + status	48669.86	51.01	8
abundance~MeanTemp + status + species	48674.54	55.69	6
abundance~MeanHum + status + species	48682.89	64.04	6
abundance~MeanTemp*MeanHum + status	48693.47	74.62	7
abundance~MeanTemp + status	48698.11	79.26	5
abundance~MeanHum + status	48706.41	87.56	5

References

Afrane YA, Lawson BW, Githeko AK, Yan G. 2005. Effects of microclimatic changes caused by land use and land cover on duration of gonotrophic cycles of *anopheles gambiae*(diptera: Culicidae) in western Kenya Highlands. *Journal of Medical Entomology*. *42*(6): 974–980.

Agyekum TP, Botwe PK, Arko-Mensah J, Issah I, Acquah AA, Hogarh JN, Dwomoh D, Robins TG, Fobil JN. 2021. A Systematic Review of the Effects of Temperature on Anopheles Mosquito Development and Survival: Implications for Malaria Control in a Future Warmer Climate. *Int. J. Environ. Res. Public Health.* 18(14): 7255.

Akaike H. 1974. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control.* 19: 716–23.

Al-Thukair A, Jemal Y, Nzila A. 2022. Influence of Climatic Factors on the Abundance and Profusion of Mosquitoes in Eastern Province, Saudi Arabia. *In: Mosquito Research - Recent Advances in Pathogen Interactions, Immunity, and Vector Control Strategies. IntechOpen.*

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*. 67(1): 1–48.

Barton K. 2020. Multi-Model Inference. *R Package Version 1.43.17*. https://cran.r-project.org/web/packages/MuMIn/MuMIn.pdf.

Burnham KP & Anderson DR. 2002. *Model Selection and Multimodel Inference*. New York, NY: Springer.

[CDC]. Centers for Disease Control and Prevention (USA). 2022. About Malaria. Atlanta (USA): Centers for Disease Control and Prevention

Charlwood JD. 2017. Some like it hot: A differential response to changing temperatures by the malaria vectors *anopheles funestus* and *an. gambiae* s.l. *PeerJ*. 5.

Daygena TY, Massebo F, Lindtjørn B. 2017. Variation in Species Composition and Infection Rates of Anopheles Mosquitoes at Differential Altitudinal Transects and the Risk of Malaria in the Highland of Dirasge Woreda, South Ethiopia. *Parasites and Vectors*. 10(1): 343.

Day JF. 2016. Mosquito oviposition behavior and vector control. *Insects*. 7(4): 65.

Genoud AP, Gao Y, Williams GM, Thomas BP. 2019. Identification of gravid mosquitoes from changes in spectral and polarimetric backscatter cross sections. *Journal of Biophotonics*. 12(10).

Kelly-Hope LA, Hemingway J, McKenzie FE. 2009. Environmental factors associated with the malaria vectors Anopheles gambiae and Anopheles funestus in Kenya. *Malaria Journal*. 8(1): 268.

Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*. 82(13): 1–26.

Lenth RV. 2016. Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*. 69(1): 1–33.

Lüdecke, D. 2018. sjmisc: Data and Variable Transformation Functions. *Journal of Open Source Software*. 3(26): 754.

Minakawa N, Sonye G, Mogi M, Githeko A, Yan G. 2002. The Effects of Climatic Factors on the Distribution and Abundance of Malaria Vectors in Kenya. *Journal of Medical Entomology*. 39(6): 833–841.

Mwingira V, Mboera L, Dicke M, Takken W. 2020. Exploiting the chemical ecology of mosquito oviposition behavior in Mosquito Surveillance and control: A Review. *Journal of Vector Ecology*. *45*(2): 155–179.

Ndoen E, Wild C, Dale P, Sipe N, Dale M. 2010. Relationships between Anopheline Mosquitoes and topography in West Timor and Java, Indonesia. *Malaria Journal*. *9*(1).

[OECD] Organization for Economic Cooperation and Development (France). 2018. Safety Assessment of Transgenic Organisms in the Environment: Volume 8. Paris (France): Organization for Economic Cooperation and Development

Soleimani-Ahmadi M, Vatandoost H, Zare M, Turki H, & Alizadeh A. 2015. Topographical distribution of Anopheline Mosquitoes in an area under elimination programme in the south of Iran. *Malaria Journal*. *14*(1).

Suh E, Grossman MK, Waite JL, Dennington NL, Sherrard-Smith E, Churcher TS, Thomas MB. 2020. The influence of feeding behaviour and temperature on the capacity of mosquitoes to transmit malaria. *Nat Ecol Evol*. 4(7): 940-951.

[WHO] World Health Organization (Switzerland). 2022. Malaria. Geneva (Switzerland): World Health Organization.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the tidyverse. *Journal of Open Source Software*. 4(43): 1686.

Yan G, Githeko A, Zhou G, Minakawa N, Omukunda E. 2006. Malaria Vector Productivity in Relation to the Highland Environment in Kenya. *The American Journal of Tropical Medicine and Hygiene*. 75(3): 448–453.