Adaptation and Variability of Aedes Aegypti in Relation to Food in the Southernmost Fringe of its Distribution - An Interdisciplinary Approach

Lovisa Björnberg, Josefin Persson

Master's thesis
2019:E57

Lund Institute of Technology
Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematics

Centrum Scientiarum Mathematicarum

Adaptation and variability of Aedes aegypti in relation to food in the southernmost fringe of its distribution

An interdisciplinary approach

Lovisa Björnberg & Josefin Persson

Centre for Mathematical Sciences
Lund University

Supervisor: Mario Natiello

Examiner: Claus Führer

August 29, 2019

Abstract

Aedes aegypti is the well-spread main vector of dengue and other viral diseases and in means of mosquito control, it is important to understand adaptation and variability within the species. This study has considered eggs of Ae. aegypti from two close regions in the area of Buenos Aires. The aim was to increase understanding in environmental and genetic characteristics in food intake. From a long-term perspective, this is a contribution to solving the issues of how modelling could take these factors into account, in order to understand or predict adaptive trends and subsequently, permanent modifications in mosquito populations. The experiment was performed in collaboration with the Buenos Aires Mosquito Study Group and was outlined to obtain data of development times and wing lengths of the subpopulations, when exposed to different diets. The subpopulations showed small differences in individuals all the way from egg to adult depending on origin and differences in growth strategies. Hence, assumptions about homogeneity or uniformity within species, that are often made in mosquito control contexts, should be reconsidered.

Sammanfattning

Myggburna virussjukdomar är ett stort problem i många tropiska och tempererade delar av världen. I olika delar av Sydamerika sprider myggarten Aedes aegypti bland annat denguefeber och artens utbredning i området har länge studerats av en särskild arbetsgrupp på Buenos Aires universitet.

Idag finns det många metoder för att förhindra spridningen av mygg. Tyvärr innebär de flesta stora problem – bekämpningsmedel riskerar att förstöra allt i sin väg och utsläpp av sterila, genetiskt modifierade myggor, förlitar sig på en myggtyp uppvuxen i laboratorium. Skillnader inom myggpopulationen i naturen kan vara tillräckliga för att överleva kontrollförfarandet. Ett sätt att teoretiskt undersöka metodens effektivitet vore att använda matematisk modellering – vilket gör variabilitet och anpassning till utmanande problem när det handlar om modellering av naturliga populationer.

Den här uppsatsen undersöker Aedes aegyptis naturliga anpassning. Experiment har genomförts i samarbete med arbetsgruppen på Buenos Aires universitet för att mäta och jämföra utvecklingstid och vinglängd på myggor från två olika platser i Buenos Aires, utsatta för olika dieter.

Resultaten från experimentet har analyserats och visar små skillnader i individer med olika ursprung, hela vägen från ägg till vuxen, beroende på ursprung. Skillnaderna är tydliga, oavsett vilken typ av näring myggorna intagit, om än i varierande grad. Resultaten är indikationer som varnar för att de antaganden om enhetlighet inom arter som ofta görs i myggkontrollsammanhang, bör ifrågasättas.

Acknowledgements

First, and most importantly, we would like to express our sincere gratitude to our supervisor Mario Natiello, at the Centre for Mathematical Sciences at Lund University, for his continuous support of our work towards this master's thesis and to Hernán Solari and Sylvia Fischer at the Faculty of Sciences of Buenos Aires University. Their motivation and support during the experiments in Buenos Aires have been of the utmost importance to our progress and results.

Second, the guidance and support from Pedro Montini and the colleagues of the Mosquito Study Group at the Buenos Aires University has been very useful to us. Their welcoming attitude, deep knowledge of biology and patience with our scarce Spanish skills have been indispensable. No matter if our questions were in English or Spanish, clever or even stupid, they always offered guidance and fine answers, always with smiles and sincere interest.

It is relevant to this project to notice that the advancements of this study, as well as many others, have been possible because of the development of a creative, long term, interdisciplinary collaboration between biologists in Buenos Aires and mathematicians in Buenos Aires and Lund. To sustain the necessary dialog between the disciplines, mathematically oriented students participate in field and laboratory work alongside with biologists and biology-oriented students participate in the mathematical developments. For that we are grateful, and we hope for the successful cooperation to continue in the future.

We would also like to thank Sida for the opportunity that came with the Minor Field Studies Scholarship to go to Buenos Aires and perform the field work and experiments. It has truly been a wonderful experience.

Finally, a big thank you to all our friends and family for the moral support we have received during the process of this master's thesis.

Table of Contents

Chapter 1
Introduction

Initially follows the background to the main question of this thesis, the purpose, objective and a brief description of the structure.

1.1 Background

Aedes aegypti (Ae. aegypti) is an epidemiologically important mosquito especially known to be the main vector of dengue fever [10] but also of several other viral diseases such as the yellow fever and zika virus [9, 13]. This mosquito species origins from tropical Africa but has spread across the world due to its highly anthropophilic[1] characteristics. It can now be found in tropical, subtropical and temperate urban ecosystems [7]. In temperate climate regions the epidemic outbreaks of dengue occur seasonally and as it may present an irregular outbreak pattern it has an important impact on social health. In order to act against the outbreaks, not only knowledge of the seasonality of the outbreaks is important but also knowledge of the seasonal behaviour of the mosquito itself.

Many methods to help prevent the spread of mosquito populations have proven serious disadvantages, for example pesticides pose a risk to other species as well as to human health [33]. As an alternative method, researchers have been developing genetically modified mosquitoes since the beginning of this century [26]. One of the approaches that have gained the most attention is modification of the male mosquito making them incapable of producing viable offspring [33]. The released males are expected to mate with the wild-type females and out compete the wild males. Over time and after multiple releases the mosquito population in the release area is expected to decrease, subsequently leading to a decrease in the spread of the disease as well [19].

Apart from the ethical issues associated with this approach, it is also a matter of ecological, epidemiological and economic uncertainties. Since the approach involves major financial investments,[2] and potentially ecological consequences, it is

[1] The species is attracted to human blood as a source of food.

[2] The cost of using genetically modified mosquitoes from a company called Oxitec has been estimated to be approximately US$1.9 million in the first year and US$384,000 each year thereafter for an urban population of 50,000 [1].

crucial to be aware of any varieties between and within populations. Otherwise, there is a risk that the genetically modified males that are released are not approved by the wild females and therefore do not get the opportunity to mate. This would ruin the trial and result in financial losses and to date unknown ecological consequences.

Mathematical models can be used as predictive tools to describe the population dynamics and potentially analyse the use of genetically modified mosquitoes. Since the models integrate biological knowledge they are not only useful for prediction but also for detecting areas where better biological knowledge is needed as well as areas where previous knowledge might have to be reconsidered. The mathematical models to date struggle with variability. Therefore, this is an essential area to investigate.

It has already been shown that the development of Ae. aegypti depends on the environmental conditions. So far, temperature [4, 14, 31], food scarcity [30] and food abundance [32, 3, 23, 30] are three factors that have been identified. The need to integrate the dependence with temperature and food of Ae. aegypti development, into mathematical models has promoted a new form of conceiving larval development and its variability [29, 30]. It has also been observed that eggs and larvae originated in different breeding areas present genetic differences [15, 24]. The current biological suspicions indicate that Ae. aegypti could be adapting also to the temperate climate [15, 21] and, despite the close distances, to the different environments of the Buenos Aires region.

There are several issues around the problem of fighting mosquitoes to prevent disease-spreading. In vector-borne diseases, see Section 2.0.1 for details, the mosquito is not the "enemy", it is the messenger. Rather than destroying the messenger, it would be interesting to focus on disrupting the message. Environmental impact and variability within the mosquito species are often forgotten characteristics involved in mosquito control. For example, pesticides are usually generic, and they destroy everything they meet, not only disease-bearing mosquitoes. Also, targeted sterile-insect techniques rely on one mosquito type, raised under laboratory conditions and to some extent different from field individuals. Internal variability in natural populations speaks against it, since it cannot be ruled out that the population could be flexible enough to circumvent the control procedure. Variability and adaptation are challenging issues when modelling natural populations.

1.2 Objective

It is fundamental to achieve a deep understanding of the biology of the vector, see Section 2.0.1, and to incorporate this knowledge in the modelling tools in order to introduce and adopt effective health policies. Although adaptation is a known property of living systems, most existing models and methods lack the ability to adapt as a consequence of genetic and environmental changes.

The goal of this project is to advance in the understanding of environmental and genetic characteristics in food intake as well as their consequences for the adult mosquito population. Variability and adaptation are to be determined against environmental differences. The differences are expected to be measured by comparing the wing length of mosquitoes from different breeding sites exposed to different diets. The data has been obtained through experiments in close collaboration with the Mosquito Study Group at Buenos Aires University.

The question addressed is to understand variability and adaptation in terms of the interplay between environment and genetics. In particular, could eggs with different origin display different adaptation response to food resources? Which would be the visible ecological consequences?

From a long-term perspective, this thesis aims to contribute to solving the issues of how modelling could take these factors into account in order to understand or predict adaptive trends and subsequently, permanent modifications in mosquito populations. One step to accomplish this is by discovering differences to map the current situation.

1.3 Limitations

Even though the ultimate endpoint of interest is the transmission of infection, attention is restricted to population dynamics of Ae. aegypti in this study. The Aedes aegypti is a species under the subgenus Stegomyia, which comes from the family of Culiciade. The species can be divided into two subspecies - Ae. aegypti formosus and Ae. aegypti aegypti (Ae. aegypti). The former can be found in forests and ecotones in sub-Saharan Africa and is the ancestor to the latter, domesticated, subspecies which is the object of study in this work. The studies of related species and subspecies are left for others.

Differences between the sexes of Ae. aegypti are known [20], but when separating the experimental data into males and females a problem with the events of death was encountered. Since the individuals die before the sex has been determined there is no way of knowing whether the dead individuals would have been male or female. It is impossible to know for certain how many events of death happen for the male versus female populations and thereof also impossible to know the initial number of individuals in each population. Assumptions about the origin of deaths would preferably be avoided to tamper with the experimental data as little as possible. Hence, all analysis with simulated data, separated by sex, was performed without considering mortality.

1.4 Structure

In the next chapter, the biology of Ae. aegypti is outlined, providing the essential information needed for understanding the experimental work as well as the biological aspects of the analysis. Then, a presentation of the experiment follows

to give a clear outline of how the experiment proceeded and on which assumptions decisions have been made. An overview of the results is given to provide the reader with a first glimpse of what is to come. Following in the next chapter is the theory essential for understanding the analysis of the collected data. After this, the outcome of the experiments and the results from the analysis are presented. Finally, the last two chapters contain the main discussion, conclusions and recommendations of further work on the topic. Some extra material can be found in the appendix.

Chapter 2

Biology of Ae. aegypti

On location in Buenos Aires, the authors education about the biology around Ae. aegypti started, supported by the expertise of the Mosquito Study Group at Buenos Aires University. The biology of Ae. aegypti is also stated in detail in volume 8 of OECS:s Safety Assessment of Transgenic Organisms in the Environment [20].

Ae. aegypti is widely spread and can be found in both tropical and subtropical regions. It mainly uses artificial containers as breeding sites, such as vases, bowls and tires but it will also use tree holes or other places storing water. In difference to the other subspecies, which feeds on wild animals, the Ae. aegypti has a preference for human blood [2].

Using morphologic features, variation among populations within the same species can be described. The Ae. aegypti is a holometabolous insect and thereby has four stages of life - egg, larva, pupa and adult.

Eggs of Ae. aegypti are around one millimetre long and are smooth and ovoid shaped. They are white when newly ovipositioned and will turn black within two hours. They are deposited as single units on moist surfaces above the water line with varying distances. The eggs can be preserved for several months if kept in a dry environment but when immersed in water the hatching begins.

The Ae. aegypti larvae requires aquatic habitats and feed on microorganisms in the water. They hang almost vertically from the surface and use the siphon to breath the air above water level. They moult three times before going into pupal stage and the four stages between moulting are called instars.

The pupal stage is the non-feeding stage that evolves after the fourth instar and is the stage where the adult mosquito develops. In optimal conditions, the mosquito stays in the pupal stage between 2.0 to 3.6 days [20]. When the adult mosquito is ready to emerge, it will split the pupal skin and enter the aerial stage. In the case of females, the adult requires at least one blood meal before they lay eggs to begin the next gonotrophic cycle[1].

[1]Gonotrophic is defined as the life cycle including feeding and laying eggs.

Male and female adults can be identified by three features that differentiate them. Males have many more antennal cilia than females and makes the antennas similar to feathers. The sexes are also different in their mouth parts. The mouth parts of the females are short, constructed for the proboscises to be able to pierce the human skin, while the mouth parts of the males are as long as the proboscises. Finally, males have longer cercus segments. The identifications can be seen in Figure 2.1. The body size of the mosquitoes varies with food density and environmental conditions, but generally, females are larger than males and require longer time to pupation. Larger body size is also related to better reproduction abilities – large males tend to live longer and older and larger males transfer more sperm to females, while large females more successfully mates with males and lay a higher number of eggs per batch.

Earlier studies have shown that there is an optimal range of food availability. Mosquitoes with scarce food access have shown to be smaller, have higher mortality in larval and pupal stages and their development times are longer than for mosquitoes with good food access. Not only food scarcity have shown to have a negative effect on the mosquito, food abundance have shown similar results. Higher mortality in aquatic stages and a small decrease in adult size have been detected [30].

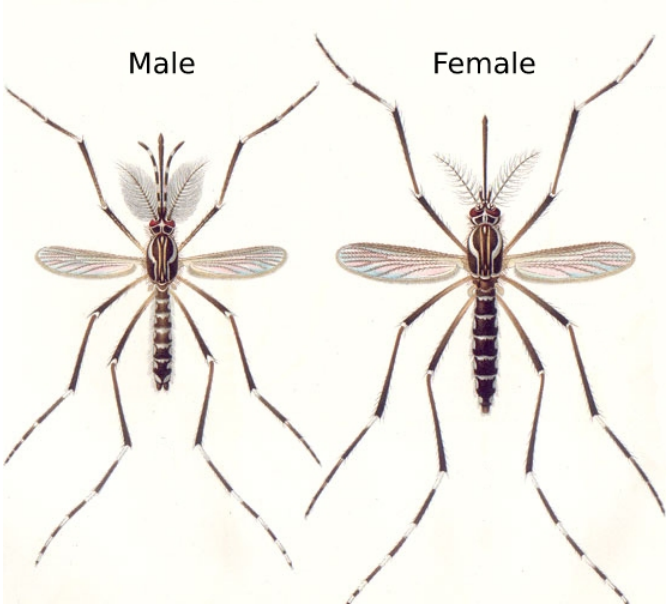


Figure 2.1: Identification of the sexes of Ae. aegypti. In comparison to females, males have more antennal cilia, their mouth parts are as long as the proboscises and they have longer cercus segments.

2.0.1 Vector borne diseases

Ae. aegypti is a vector[2] that can transmit infectious diseases such as dengue, zika, chikugunya and yellow fever. The diseases are arthropod-borne viruses or arboviruses, which are defined as viruses that are preserved in nature through biological transmission between hosts. The viruses transmit by arthropods – such as Ae. aegypti, other mosquitoes or other vectors such as ticks and flies. The arboviruses produce and multiply in susceptible vertebrate hosts. It also multiplies in the tissue of the arthropods and are passed on to other vertebrates by bites from an infected arthropod, for example when the Ae. aegypti mosquito is having a blood meal from a human. The necessary time for the virus to develop in the host is called the extrinsic incubation period. If the female mosquito has an adult life shorter than the extrinsic incubation period the vector transmission is reduced.

Ae. aegypti usually transmit viruses through a human-to-mosquito-to-human cycle. Transmission through vector is a quite fragile way of transmitting diseases and it is demanding for the virus, which has to be adapted to both the mosquito and the human host. However, vertical transmission of dengue has been recorded for Ae. aegypti which allows the virus to persist during time periods where environmental conditions limit the mosquito reproduction. Vertical transmission means that the infected females transmit the virus naturally directly to their eggs, i.e. skipping the intervening human host in the transmission cycle [20].

[2]In this case vector means messenger or intermediary.

Chapter 3

Outline of the experiment

The experiment was performed at the University of Buenos Aires with start in March. Ae. aegypti in Buenos Aires city has been studied by the Buenos Aires Mosquito Study Group for years and several advancements have been made in order to learn more about the biology of the species. Suspicions that Ae. aegypti found in the city and in a more rural environment could show different characteristics due to adaptation has come up. This experiment focuses on finding differences arising from heritage by investigating how the development times and sizes of adult mosquitoes change within the populations for different types and amounts of larval food. The different food types were constructed to simulate the food of their origin areas and the different amounts of food to enable visibility of the differences, i.e. how the subpopulations react within the range scarce-optimal food availability. Both populations started with the same number of larvae exposed to all different food types and levels. Constant conditions, such as food availability, temperature and light, were held for the larvae within the different food types and food levels. Due to adaptation, the populations were assumed to show differences when they were treated with food of their own environment in comparison to when they were treated with food of the other.

3.1 Choice of populations

The rural population was found at a horse farm frequented by one of the colleagues of the Mosquito Study Group. The farm is located in Villa Elisa, a city approximately 40 km south of the urban location in Buenos Aires. The eggs were found in the drinking basins of the horses both in and outdoors and the assumption was that they have fed on the bacteria developing from the horse's grass that have fallen into the drinking basins. It was eggs from this particular site as well as eggs from the urban area of Buenos Aires that were examined in this work.

3.2 Collection and measurements of eggs

Eggs of the two populations were collected using ovitraps [1] and since Ae. aegypti is the only container breeding mosquito species in the area, all the collected eggs were assumed to belong to the desired species. From now on, the eggs of the urban population will be referred to as the city eggs (CE) and the eggs from the rural population as the farm eggs (FE).

Before the beginning of the experiment, the eggs were photographed and measured using a dissecting microscope equipped with a digital camera. Width and length of the eggs were measured on digital photos using the Leica Application Suite V 4.0.0. An illustration of how the measures were taken is shown in Figure 3.1. Two eggs per spatula were measured resulting in a total of 60 measurements per population.

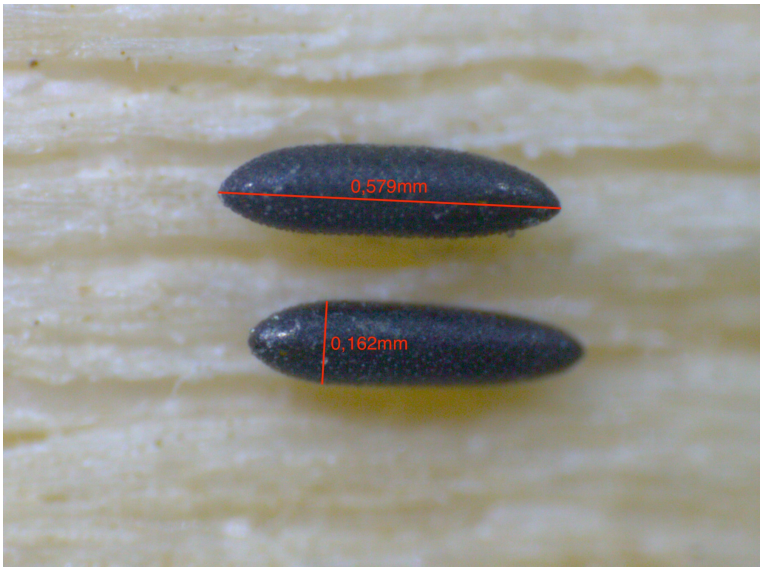


Figure 3.1: Two city eggs photographed and measured with use of Leica Application Suite V 4.0.0. The eggs show how the measurements of length(top egg) and width(bottom egg) were taken. Length and width were measured for every egg.

3.3 Egg hatching and breeding of larvae

After the measurements, the eggs were immersed in dechlorinated tap water to induce hatching and the hatched larvae were separated in cohorts of 10 individuals within 24 hours. Each cohort of the populations was assigned to one food treatment

[1] Ovitraps are small, dark glass cans containing water and a spatula. The female mosquito sees the breeding site and deposits her eggs. The spatulas are collected once every week and the eggs are stored.

with a total of eight different food treatments. Every treatment had three replicate cohorts, resulting in a total of 30 larvae exposed to the same treatment. The total number of larvae in the beginning of the experiment was 480. See Table 3.1 for an overview of the replicas, food treatments and populations.

Table 3.1: Table to display the divisions of populations and food treatments. Each population has eight food treatments and every treatment has 30 individuals, divided into three replicas with 10 in each.

Food treatment	City population	Farm population
City, High	(10) (10) (10)	(10) (10) (10)
City, Mid	(10) (10) (10)	(10) (10) (10)
City, Low	(10) (10) (10)	(10) (10) (10)
Farm, High	(10) (10) (10)	(10) (10) (10)
Farm, Mid	(10) (10) (10)	(10) (10) (10)
Farm, Low	(10) (10) (10)	(10) (10) (10)
Yeast, High	(10) (10) (10)	(10) (10) (10)
Yeast, Low	(10) (10) (10)	(10) (10) (10)

A treatment is defined by a food type and a food level. The food types are City, Farm and Yeast and the food levels are High, Mid and Low. To see how the types and levels were constructed, see section 3.4. Food type Yeast was only assigned levels High and Low because earlier studies have already stated optimal and scarce nutrition levels of food. Experiments using simulated food, as for food types City and Farm, are to date new and they were hence assigned all three food levels to make sure the right range within scarce-optimal-abundance food access was covered. To simplify the reading, the following notations are introduced: (T, L) where $T \in \{\text{City, Farm, Yeast}\}$ and $L \in \{\text{High, Mid, Low}\}$.

The three replicate cohorts for each food treatment and population were reared in cylindrical plastic containers (200 ml) containing 180 ml of food liquid, resulting in 48 different containers. Plastic caps were used to avoid contamination from the outside environment. The containers were randomly positioned in a 4 by 12 matrix and their positions remained constant during the complete duration of the experiment, with exceptions only for the time of inspection. Each container was inspected daily, Monday to Friday, and transferred to a new container each Monday, Wednesday and Friday to maintain constant food availability. Bacterial growth in the containers was expected to be minimal since the infusions were strained and no further nutrients were added. During the daily inspections, larvae, pupae and adults were counted and new pupae were transferred to individual containers fit for adult emergence. Temperature conditions were recorded and a $12:12$ (light:dark) photoperiod was maintained.

3.4 Food types and food levels

The two food types, City and Farm, used in the experiment were prepared to represent the natural food from the areas of the two populations. The food type Yeast was used as a reference to earlier experiments in order to detect any extreme deviations. There existed no completed data of experiments conducted with natural food representations to date, but there was an experiment with City food ongoing in Buenos Aires. The choices of food concentrations in the experiment of this thesis were based on the concentrations used in that experiment and by the intuition and expertise of the colleagues from the Buenos Aires Mosquito Study Group. All liquids were prepared in a large container before portioning it to the small ones used in the experiment. The levels were constructed to differ by 1/3 of the contiguous concentrations within the same food type.

City

Dried, fallen leaves in the urban region were collected and weighted into portions of 20 g. One portion was put into 1500 ml dechlorinated tap water every Monday, Wednesday and Friday and was left soaking for 7 days to increase the number of bacteria. The infusion was strained before use to keep a constant environment and added to dechlorinated tap water to achieve the desired food levels. Table 3.2 shows the concentrations for the different levels of City food.

Table 3.2: Concentrations of City food for the three levels High, Mid and Low.

	High	Mid	Low
Food infusion (ml)	1400	467	156
Concentration (mg/ml)	1.037	0.3457	0.1152

Farm

The Farm food consisted of dried grass, commonly used as horse food. It was also weighted into portions of 20 g and put into 1500 ml dechlorinated tap water one week before use. At the beginning of the experiment, the same concentrations were used for the Farm infusions as for City, but the High food level showed itself to be too high ending up in the optimal-abundance range. Since the optimal-scarcity range between High and Low levels was of interest, new concentrations were constructed. New cohorts were added to the existing experiment five days later with the same external conditions. The resulting concentrations can be found in Table 3.3.

Table 3.3: Concentrations of Farm food for the three levels High, Mid and Low.

	High	Mid	Low
Food infusion (ml)	156	52	17
Concentration (mg/ml)	0.1152	0.0384	0.0128

Yeast

The food used for the Yeast treatments were two different amounts of dry baker's yeast (Levex®) diluted in 1200 ml of dechlorinated tap water and portioned into the containers. The concentrations and corresponding amounts of yeast were decided using data from a previous experiment conducted by Victoria Romeo and colleagues [30], see Table 3.4.

Table 3.4: Concentrations of Yeast food for the levels High and Low.

	High	Low
Yeast (mg)	150	15
Concentration (mg/ml)	0.125	0.0125

3.5 Collection of data

Time to pupation, time to adult emergence and sex were recorded for each individual. The right wing was removed from all well-preserved adults and measured from the alular notch to the distal margin, excluding the fringe scales, to the nearest 0.001 mm by using the same dissecting microscope as for the eggs. An example of the measurement is shown in Figure 3.2. With no access to the experiment on weekends, some data were lost and had to be interpolated. Since the mosquito stays in the pupal stage between 2.0 to 3.6 days in optimal conditions [20], either the date of pupation or the date of adult emergence was correctly discovered at investigation. The other stage was assigned a date 2.0 to 3.6 days before or after the correct pupation or emergence. Estimation of the date was based on the corresponding data of the other individuals within the treatment. The same method was used for the treatments with scarce food conditions.

For a group of larvae reared under equal conditions, the body weight is roughly proportional to the cube of the wing length [7, 14]. Hence, the cube of wing length has been used in the analysis as an indicator of the adult size of emerging adults.

(a) Female **(b)** Male

Figure 3.2: Right wing of each sex, with corresponding measurements. The red line goes between the alular notch and the distal margin.

Chapter 4
Overview of results

To make it easier to follow the results and to understand where the basic knowledge of the next chapter will be applied, an overview of the results is given. The first part of the results provides a validation of the experimental data. This is to give an indication of how well the experiment was performed in relation to previous studies.

The next section presents a statistical method to determine if the replicas could be seen as a combined data set. The method was inspired by general linear models and bootstrapping and was made to improve sample sizes and to reduce the amounts of comparisons between the cohorts in the analysis. After this has been proven, each 3 replicas of 10 are handled as one sample of 30 individuals.

The following section analyses the eggs measured prior to the experiment. Results from a bivariate Kolmogorov-Smirnov test indicate differences in sample distributions, i.e. between the populations.

After the analysis of the eggs follows an analysis of characteristics of mosquitoes from city and farm populations. First, the food levels are verified to lie within the desired range. Then a graphical presentation of differences in development between populations are provided. With the use of synthetic data, Kolmogorov-Smirnov tests are made to identify where the differences can be found.

Span data of adult emergence times and adult weights are next to be outlined. Box and whisker plots are presented to show central values as well as shape and variability of the distributions of the data sets. One extreme deviation, not likely to arise from natural variability, is identified and noted in treatment (Farm, Mid) for males in the farm population. The data of this particular mosquito was removed from all prior analysis to avoid improper results. Hence, when reading the results, keep in mind that the outlier has not been part of the data.

Finally, a measure of efficiency is introduced to compare the amounts of biomass generated for the two populations.

Chapter 5
Basic knowledge

In the following chapter, some basic knowledge will be provided to facilitate the interpretation of subsequent results and conclusions. First comes theory about the distribution that has been relevant to the experiment. Then follows a couple of principles and methods that have played a central role in obtaining the results.

5.1 Multinomial distribution

In probability theory, the multinomial distribution is used in experiments where there are more than two outcomes. When there are only two outcomes the almost identical, but less general, binomial distribution is used. If the experiment is performed n times and there are k possible outcomes each time, the number of times that one obtains the i-th outcome is denoted by X_i. Then, the random vector X, defined as

$$X = [X_1, X_2, \ldots, X_k],$$

is a multinomial random vector [5].

Multinomial random vectors can formally be characterised by the following definition [27].

Definition (Multinomial distribution)**.** Let X be a $k \times 1$ discrete random vector, $n \in N$ and the support R_X of X be the set of $k \times 1$ vectors having non-negative integer entries summing up to n. Furthermore, let $p_1, \ldots, p_k$ be k strictly positive numbers such that

$$\sum_{i=1}^{k} p_i = 1.$$

It is true that X has a multinomial distribution with probabilities $p_1, \ldots, p_k$ and n number of trials if its joint probability mass function is given by

$$p_X(x_1, \ldots, x_k) = \begin{cases} \binom{n}{x_1, \ldots, x_k} \prod_{i=1}^{k} p_i^{x_i}, & if (x_1, \ldots, x_k) \in R_X \\ 0, & otherwise \end{cases}$$

where $\binom{n}{x_1, \ldots, x_k}$ is the multinomial coefficient.

5.1.1 Distribution of experimental data

Now consider the experiment of this thesis. Due to the assumption that individuals cannot skip steps in their life cycle, the life cycle of the mosquito causes two experiments. The first is based on the number of live larvae and the second on the number of live pupae. This essentially means that the mosquito cannot go from larva to adult without first turning into pupa. There are two or more outcomes in each experiment:

$$\textbf{Larva} \rightarrow \begin{cases} Pupate \\ Die \\ Remain\, larva \end{cases} \qquad \textbf{Pupa} \rightarrow \begin{cases} Emerge \\ Die \\ Remain\, pupa \end{cases}$$

For a single individual, exactly one outcome is possible each day and any outcome can occur with a probability $0 \leq p_i \leq 1$, such that $\sum_{i=1}^{3} p_i = 1$. Hence, the experiment of this thesis is recognised to be of multinomial character. The probabilities can be computed as the number of events on day j divided by the initial number of larvae.

5.2 Production of synthetic data

5.2.1 Global VS day-by-day

The mosquito population dynamics can be described equivalently by two setups; either from a global perspective or from a day-by-day point of view.

The global setup describes the occurrence of events, where for the larval events, only pupation and death are considered. The larval setup, where P is a vector of pupations and D is a vector of deaths, follows

$$(P, D) = \text{multinomial}(N_0, p_1^p, p_1^d, ..., p_k^p, p_k^d),$$

where N_0 is the number of larvae entering the experiment and p_i^p and p_i^d are the probabilities of pupation and larval death on day i, respectively, for $i = 1, ..., k$. Here, k is the last day of non-zero experimental events in the life cycle. The probability of pupation on day j is given by $p_j^p = \dfrac{n_j^p}{N_0}$, where n_j^p is the counted number of pupations of the 30 larvae on day j in the experiment. The probability of death is created similarly, where $p_j^d = \dfrac{n_j^d}{N_0}$, and where n_j^d is the number of deaths on day j. The experimental record guarantees that $\displaystyle\sum_{i=1}^{k}(p_i^p + p_i^d) = 1$.

As for the day-by-day approach, the probabilities are computed as for the global, i.e as the number of counted experimental events on a day divided by the total number of live larvae or pupae. In the example of pupation and larval death, let p_1^p be the probability of pupation and p_1^d of larval death on day 1. Then, let

$$(P_1, D_1, R_1) = \text{multinomial}(N_0, p_1^p, p_1^d, 1 - (p_1^p + p_1^d))$$

be a random multinomial deviate corresponding to the first day, with an outcome of P_1 pupations, D_1 larval deaths and R_1 remaining larvae. For day 2, there are only $N_1 = N_0 - (P_1 + D_1)$ larvae available for further development and the number is updated recursively with each passing day, i.e. $N_{j-1} = R_{j-1}$. Then, for day $j \in [1, k]$:

$$(P_j, D_j, R_j) = \text{multinomial}(N_{j-1}, p_j^p, p_j^d, 1 - (p_j^p + p_j^d)).$$

In particular, on the last pupation day:

$$(P_k, D_k, R_k) = \text{multinomial}(N_{k-1}, p_k^p, p_k^d, 0) = (N_k, R_{k-1} - N_k, 0),$$

If there are no deaths on the last pupation day, this is just

$$(P_k, D_k, R_k) = \text{multinomial}(N_{k-1}, 1, 0, 0) = (N_{k-1}, 0, 0)$$

with certainty. The same idea holds for the pupae setup. The equivalence between global and day-by-day approaches is discussed in M. Otero et. al (2011). In this thesis both approaches have proven useful, especially when compared with each other as a control of various implementations.

5.2.2 Producing synthetic data using the day-by-day approach

When producing synthetic data mainly the day-by-day approach was used. In this section, a more thoroughly explanation of how the synthetic data was produced will be handed.

For each treatment, the three replicas were merged into one and an event record was produced. The event record notes how many pupations, death of larva, death of pupa and adult emergences occur every day. For each day, the frequency of each event is given by the number of occurrences divided by the available population for that event, which will represent the probabilities of events in the multinomial distribution. Due to the assumption that individuals cannot skip steps in their life cycle, the life cycle of the mosquito causes two available populations each day, dividing the production of the synthetic data into two steps which will be explained below. For each day, a multinomially distributed random number of events are generated in each step, using Matlabs built-in function **mnrnd** [18]. The function takes the available population as input and the probabilities for each possible event. Other tools and functions work just as well. For example, R and its built-in function **multirand** were used for a number of examples and controls. The synthetic data was produced to make three replicas of ten individuals each.

In the first step, the available population is the number of living larvae and there are three possible events each day – pupation, death of larva and remain being larva. At the first day, the number of available population as input is 10. The random multinomial function generates random outputs for the synthetic data, according to the probabilities based on the event record, to represent the number of events occurring that day. The data is stored and the available population is updated as input for the next day. Note that the available population is decreasing

with the days, where $0 \leq n_i \leq 10$.

In the second step, the available population is the number of living pupae and there are three possible events each day – adult emergence, death of pupa and remain being pupa. The initial available population is 0 since there are no pupae the first day. The number is updated for each day based on pupation from the first step, and it is also dependent on the number of dead pupae and adult emergences which already has occurred and been stored.

Figure 5.1 shows an example of how the synthetic data was created for one replica to provide an overview of the process.

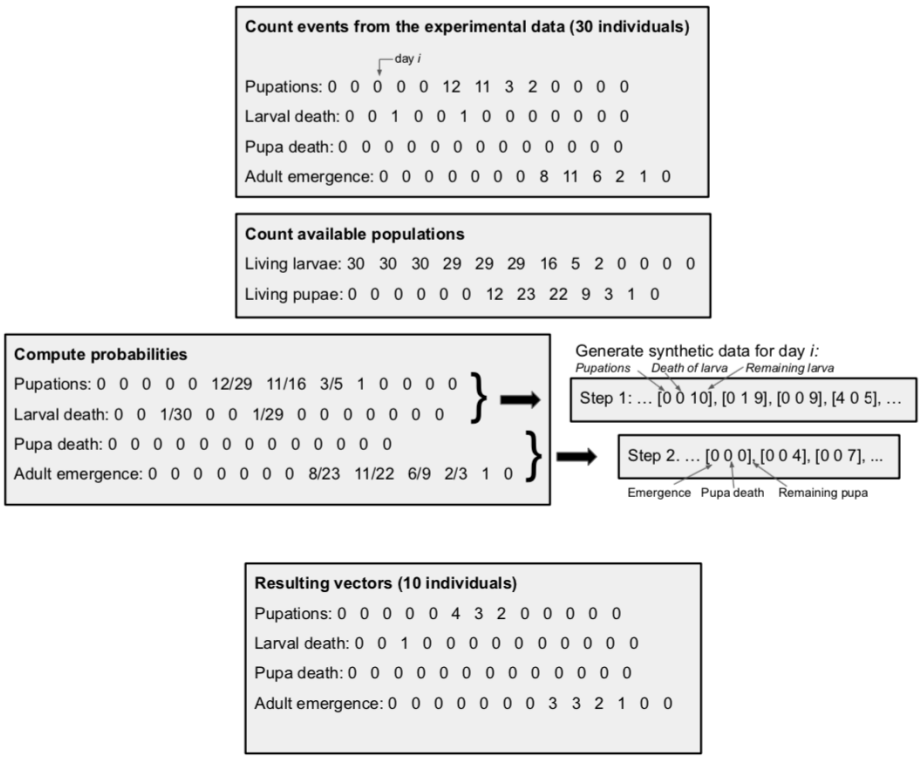


Figure 5.1: Overview of the process to generate synthetic data.

5.2.3 Adjustment of experimental frequencies

To invoke the use of experimental frequencies in the description of the process, a previous condition is that the frequencies are stable, i.e. that the outcomes f_d will not vary significantly with N_0. However, in any single experiment it cannot be ruled out that a given f_d is deviating largely from its (unknown) stabilised value. While this is possible for any f_d in any experiment, in this thesis only the "worst

case" will be considered, namely when many experimental frequencies are zero at intermediate days.

Reading a zero in a binomial or multinomial trial of size 30 is not enough to assume that there will be zero events even if N is large, say $N = 3000$. Assume that such a zero result was not a peculiar effect of chance, but rather a typical outcome. The least demanding formulation of this assumption is that the underlying probability in our process is $p < f_d$, where p is computed such that the actual experimental outcome has probability larger than $1/2$.

A minimal adjustment can be conceived as follows. Consider the periods with no pupations and their adjacent days as globally accurate but biased in the detail. Hence, instead of the frequencies $\cdots, \frac{n_a}{N}, \frac{0}{N}, \cdots, \frac{0}{N}, \frac{n_b}{N}, \cdots$ their values are proposed to be adjusted with two correction factors $\alpha, \beta \in (0, 1)$, obtaining $\cdots, \alpha\frac{n_a}{N}, \frac{\beta}{N}, \cdots, \frac{\beta}{N}, \alpha\frac{n_b}{N}, \cdots$. The factor α reduces slightly the empirical nonzero frequencies while β replaces the zero records. Assume that there is a total of s zero records, possibly including some day after the last registered event. These factors are chosen in such a way that:

$$\begin{aligned} s\frac{\beta}{N} + \alpha\left(\sum_i \frac{n_i}{N}\right) &= 1 \\ \alpha^N &= \frac{1}{2} \end{aligned}$$

Meaning that α, β are such that the same number of events occurs in the considered period and the probability of their occurrence in the same (amount of) days as observed days is at least $1/2$. However, there exist no longer any days where the probability of occurrence is zero.

5.3 Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test is a statistical non-parametric test. The test can be used on two sets of data to determine whether they could arise from the same distribution.

The Kolmogorov-Smirnov test is applied as follows.

1. Find the maximum absolute difference, D_n, in the two cumulative probability distributions being compared. The comparison considers all four possible ranking combinations.

2. Define the Z_n statistic by $Z_n = \sqrt{n}D_n$. Convert Z_n to Z_∞, by $1 - Z/Z_\infty = 0.53n^{-0.9}$, where $n = n_1 n_2/(n_1 + n_2)$ for the two-sample test.

3. Calculate the significance from $P(> Z_\infty) = 2\exp[-2(Z_\infty - 0.5)^2]$. At very most, this probability will be too large by a factor of ~ 1.5.

The probability $P(x < X, y < Y)$ needs to be evaluated only where X and Y are found in the data, since D_n will not be a maximum otherwise. No errors of practical importance will be made in $P(> Z)$ since the statistic distribution has been shown to be almost distribution-free. A more detailed report on the Kolmogorov-Smirnov test can be found in Peacock (1983).

To apply the Kolmogorov-Smirnov test on the data sets of this study, the function **kstest2** and **kstest_2s_2d** was used in Matlab. The algorithms of the functions are based on the theory above and **kstest_2s_2d** takes two, two-dimensional, sample matrices as arguments. The default significance level is 0.05 but can be set as an additional, optional argument if preferred. The null hypothesis is that both data sets were drawn from the same continuous distribution and the function returns 1 if the hypothesis should be rejected. Otherwise it returns 0.

Similar functions have been implemented and can be found online for Python and C. There is also a package in R, 'Peacock.test', where the function **peacock2** implements the original definition of the two-dimensional test by Peacock (1983).

5.4 The bootstrap principle

The idea of bootstrapping is to fit a model to data, use the fitted model to calculate the functional and generate new synthetic data from the model to get the sampling distribution, then repeat the estimation on the simulated output. Here, the experimental frequencies of occurrence of the various events are used to generate a computer-based repetition of the experiment, i.e. starting with N_0 larvae, the different number of deaths, pupations and adult emergences are obtained as random deviates from the model.

With the original data set x, let the parameter estimate from the data be θ. Synthetic data sets simulated from the fitted model will be $\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_B$ and the corresponding re-estimates of the parameters on the synthetic data are $\tilde{\theta}_1, \tilde{\theta}_2, ..., \tilde{\theta}_B$. The statistic T is used to estimate the functional, with sample value $\tilde{t} = T(x)$, and values of the synthetics of $\tilde{t}_1 = T(\tilde{X}_1), \tilde{t}_2 = T(\tilde{X}_2), ..., \tilde{t}_B = T(\tilde{X}_B)$. Succeeding applications follows without any modifications also when the functional of interest is the parameter, or a component of the parameter [6].

5.5 General linear models

An article by Haase (2014) describes the purpose of a general linear model as a representation of a variable y using a combination of variables $x_1, x_2, ..., x_p$. The representation is linearly given by

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_p x_{ip} + e_i,$$

or in matrix notation,

$$\mathbf{y} = \mathbf{Xb} + \mathbf{e}.$$

It can also be described geometrically by projecting y into the space V_x of linear combinations of $x_1, x_2, ..., x_p$, see Figure 5.2.

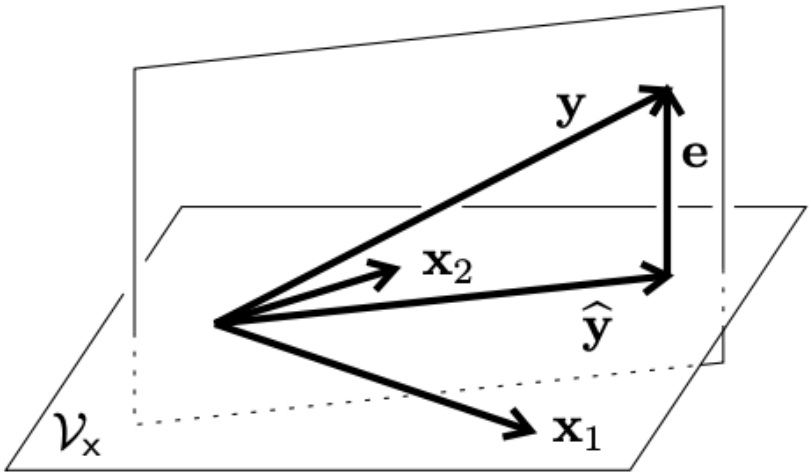


Figure 5.2: Geometrical interpretation of a general linear representation of $\mathbf{y}$ by a combination of $\mathbf{x}$ [11].

The coefficient vector $\mathbf{b}$ for the projection can be algebraically determined by

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

As Figure 5.2 shows, the representation decomposes $\mathbf{y}$ into two orthogonal vectors, $\hat{\mathbf{y}}$, which is the prediction of $\mathbf{y}$, and the error $\mathbf{e}$. The vector $\mathbf{b}$ should be chosen to minimise $|\mathbf{e}|$.

Chapter 6
Results

6.1 Relation to previous studies

To validate the data and to confirm that the experiment and its outcome follow the expected behaviour of Ae. aegypti, the data was analysed with similar methods as in the work of V. Romeo Aznar and colleagues in 2015 [29] and compared to their previous results.

In their report, a nonlinear relationship between accumulated normalised (wing length)^3 and accumulated normalised pupation was presented. The data had been fitted, using least squares, to a polynomial where the lowest grade was chosen. The fitted function was given as $f_{\text{prev}}(x) = x(1 + 0.1182(1 - x))$. To compare the results of this experiment to theirs, a nonlinear function for each population was fitted to the experimental data, using the same method as in the work of 2015. The resulting parameter values were $b_{\text{farm}} = 0.1790$ and $b_{\text{city}} = 0.1924$. These values differ on the third significant digits in comparison to each other as well as to the parameter $b_{\text{prev}} = 0.1182$, indicating a similarity between the experiments. Hence, the analysis can be proceeded with confidence in the outline of the experiment as well as in the assumption that the mosquitoes are somewhat representative to the overall species. Figure 6.1 shows the experimental accumulated normalised (wing length)^3 in relation to accumulated normalised pupation for each population and their corresponding fitted functions.

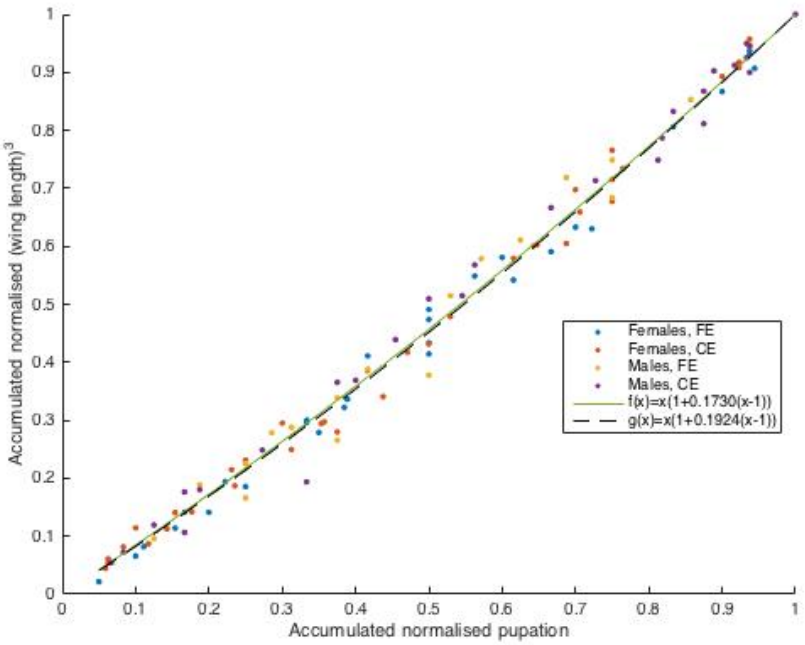
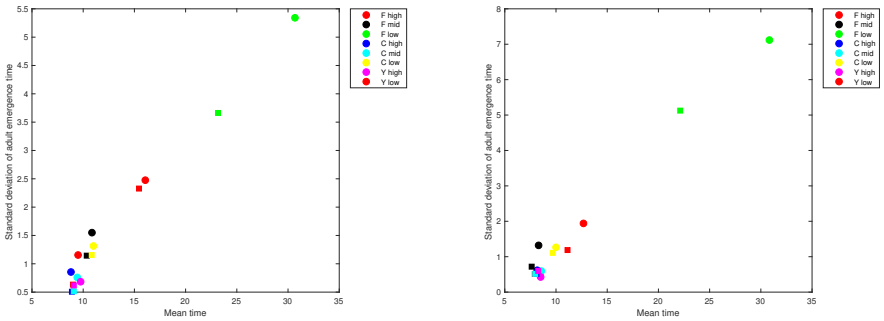


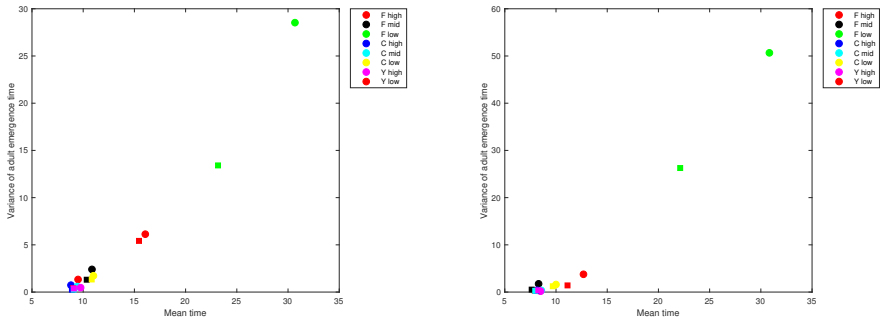
Figure 6.1: Accumulated normalised pupation in relation to accumulated normalised (wing length)^3. Both sexes and egg types have been separated and the accumulation and normalisation has been performed within each treatment and plotted together. A function for each population has been fitted where f(x) correspond to farm population and g(x) to city population.

To continue the comparison to previous studies, standard deviation and variance of adult emergence time as a function of average time to adult emergence of each treatment was computed and plotted, see Figure 6.2 and 6.3. The graph in 6.2 is noted to be linear and the graph in 6.3 to be of quadratic character. The variance and standard deviation from the data of this experiment is likely to be underestimated due to the interpolation of the weekend events. Nevertheless, the expected linear and quadratic behaviour is showing in the graphs which is in line with the corresponding results from V. Romeo Aznar et al [29]. However, no quantitative conclusions can be extracted merely from these results.



(a) Females **(b)** Males

Figure 6.2: Standard deviation of adult emergence time as a function of average time to adult emergence of each treatment. Circles mark farm and squares mark city populations.



(a) Females **(b)** Males

Figure 6.3: Variance of adult emergence time as a function of average time to adult emergence of each treatment. Circles mark farm and squares mark city populations.

6.2 Testing replicas to be considered a unity

The experiment was created with three replicas of each treatment and population as a security in the event of anything unexpected to happen, e.g. fungus contaminating the container. These replicas would preferably be considered a unity in the analysis in order to increase sample size and enable fewer comparisons between the cohorts. Each treatment had three replicas with 10 larvae each for the two different egg types. Recall that males and females differ in size and adult emergence time and since the identification of sex happens in the adult stage, the number of males and females could vary in the different cohorts. Hence, the sexes need to be separated not to affect the analysis. When separating the male and female data within the cohorts of 10 larvae, the sample size is too small to run the Kolmogorov-Smirnov test, i.e. a test to indicate if the sample data could arise from the same distribution. Hence, data was simulated, and a method was created to test if the replicas could be considered a unity. The method was inspired by general linear models and bootstrapping, which were explained more thoroughly in Chapter 5.

6.2.1 GLM and bootstrap inspired analysis

The problem consist of a daily 5-dimensional non-negative integer vector with population records: $(\#l, \#dl, \#p, \#dp, \#a)$, i.e. number of living larvae, dead larvae, living pupae, dead pupae and adults. The sum of the entries is always $N = 10$ for each of the three replicas[1], hence the entries are not independent. Let the four independent variables be: pupation, death of larva, death of pupa and emergence. These are the events that change the populations. It is important to note that the possible events are independent, not the populations. The null-hypothesis is that the three replicas are taken from the same population.

If all three replicas arise from the same distribution, in the limit of large numbers, they could be considered as one where all three replicas are combined. Hence, a linear combination of all three samples can be computed that is identically zero, with exception for the error. The combination coefficients A_i add up to zero. One may think of optimising the choice of coefficients, but here a simpler approach is taken.

Let the coefficients be integers and $|A_1| + |A_2| + |A_3| = 10$. Since they add to zero, fix $A_1 \leq 0$, $A_2 \geq 0$. Then there is a finite number of combinations, namely $(-5, 5, 0), (-5, 4, 1), (-5, 3, 2), (-5, 2, 3), (-5, 1, 4), (-5, 0, 5), (-4, 5, -1), (-3, 5, -2)$, $(-2, 5, -3), (-1, 5, -4), (0, 5, -5)$, by permuting the replicas we also obtain $(-4, -1, 5), (-3, -2, 5), (2, -3, 5), (-1, -4, 5)$. The proposed measure of zero is

$$E_\infty = \frac{1}{10d} \sum_i \left| A_1 v_1 + A_2 v_2 + A_3 v_3 \right|_i, \tag{6.1}$$

where i spans over the days where events take place, v_1, v_2, v_3 are the vectors with daily count of the independent variables for each replica and the A_1, A_2, A_3 should

[1] Except for those replicas containing individuals where the data could not be retrieved.

be chosen so that E_∞ is as small as possible. Vectors with the parameters A_1, A_2, A_3 minimising E_∞ for each treatment of the experimental data set is collected.

Synthetic data with 1000 samples times 3 replicas containing 10 initial larvae is produced and E_∞ is computed for the 1000 triples using the already collected vectors with A_1, A_2 and A_3. The outcomes are ordered by size and the value a is computed, such that 95% of the replicas have an outcome smaller or equal to a. If the experimental data has $E_\infty < a$ the assumption that the replicas come from the same distribution cannot be rejected with 95% confidence. The production of the synthetic data is explained in further detail in Chapter 5, and the results of the analysis is presented in Table 6.1.

Table 6.1: Results of the distributional analysis of the replicas.

Food type	Level	Egg type	E_∞	95% critic value
F	High	F	0.31	0.61
		C	0.08	0.61
	Mid	F	0.27	0.49
		C	0.29	0.60
	Low	F	0.18	0.26
		C	0.13	0.37
C	High	F	0.31	0.63
		C	0.18	0.64
	Mid	F	0.23	0.58
		C	0.06	0.76
	Low	F	0.23	0.52
		C	0.31	0.55
Y	High	F	0.23	0.54
		C	0.17	0.70
	Low	F	0.21	0.38
		C	0.18	0.44

From Table 6.1 it is clear that E_∞ is below the critical 95% value for all treatments, concluding that the three replicas can be considered to arise from the same distribution. Adjustments of the frequencies of events were made to allow for uncertainty of the experimental data, see details in Chapter 5.2.3. The repeated analysis with adjusted frequencies did not put any of the E_∞ above the 95 % critic values, i.e. it did not alter the conclusion. Hence, the replicas were treated as one complete sample, of 30 individuals each, in the continuing analysis.

6.3 Analysis of eggs

The first approach of identifying characteristics of the two populations was to examine the eggs measured in the beginning of the experiment. Expected values and standard deviations were calculated for width and length, using measurements of 60 eggs of each population. The results are presented in Table 6.2.

Table 6.2: Descriptive statistics of egg sizes from the two populations given in millimetres.

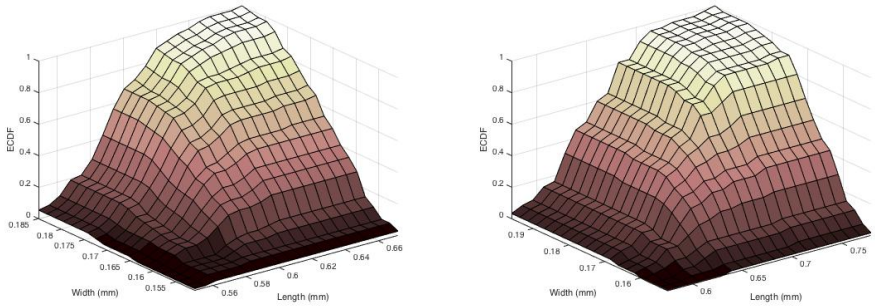
	City eggs	Farm eggs
Expected value (length)	0.6031	0.6431
Expected value (width)	0.1630	0.1687
Standard deviation (length)	0.0293	0.0398
Standard deviation (width)	0.0072	0.0094

The expected values are noted to be higher for the farm eggs, but the farm eggs also show higher standard deviations. To conclude whether or not farm eggs actually can be considered bigger than city eggs, confidence intervals were constructed for the lengths and widths of the two egg types, see Table 6.3. It is clear that the confidence intervals cover the expected values of the opposite population. Hence it is not enough to conclude that the eggs actually arise from different distributions, although the expected values are higher for the farm eggs.

Table 6.3: Confidence intervals for egg length and egg width of the two populations city (CE) and farm (FE).

	Lower bound	Upper bound
Length (CE)	0.5456	0.6606
Length (FE)	0.5650	0.72117
Width (CE)	0.1490	0.1770
Width (FE)	0.1503	0.1871

The analysis proceeded with plotting the empirical joint cumulative distribution functions, resulting in Figure 6.4. The figure indicates graphical differences between the two populations but no conclusions can be drawn solely from Figure 6.4.



(a) City **(b)** Farm

Figure 6.4: Empirical joint cumulative distribution functions of egg length and egg width of the two populations.

The analysis continued with a hypothesis test using the non-parametric bivariate Kolmogorov-Smirnov test. The null hypothesis was that the two samples could arise from the same distribution. The test included both length and width as input and concluded that the hypothesis of the eggs coming from the same distribution could be rejected on a significance level of 0.05.

6.4 Testing food levels by using univariate Kolmogorov-Smirnov test (0.05)

Previous studies show that food levels have a proven effect on development times and wing length within ranges from food scarcity to food abundance [30, 32, 3, 23]. To confirm that the food levels are within the desired range, scarce-optimal, a univariate Kolmogorov-Smirnov test was performed to test data sets of different levels within the same food type. Wing lengths and adult emergence times were tested to investigate the effect of the different food levels. Sexes were separated for two reasons, the difference in distributions within each treatment and the differences in adult emergence time and $(\text{wing length})^3$. The logical test results can be found in the appendix, Tables A.4 and A.3.

The test results indicate significant differences between food levels High and Low for all food types and for both sexes. Since emergence times are longer and $(\text{wing length})^3$ are smaller for food types with level Low, see confirmation in Table 6.4 and 6.5, the level can be considered within the scarce food range.

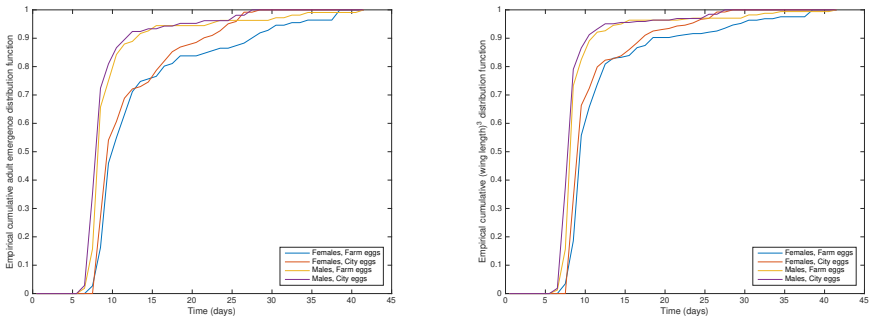
When comparing the logical values of adult emergence times and $(\text{wing length})^3$ for Mid-levels to both High and Low, the logical values are not always consistent in their indications. However, for all treatments, except for the upper levels of City food, the test shows significant differences at least in one of adult emergence

or (wing length)^3. These differences are always in advantage to the high levels, see confirmation in Tables 6.4 and 6.5 under Section 6.5, indicating that an appropriate range has been reached.

Whereas City food does not show any significant differences in this test, it will later be shown that the High levels provide larger biomass as outcome, i.e. the High level of food is preferable and can therefore be considered optimal. These results can be found in Table 6.6 in Section 6.6. Some graphical representation of the food levels can be seen in Figure 6.8 and 6.7, see Section 6.5.1.

6.5 Accumulated emergence times and adult weight

Having verified the appropriate range of food levels High and Low, the populations were separated to look for characteristics, according to their food treatments. To get a graphical idea of how the populations developed in relation to their food intake, sexes were separated within populations and the empirical cumulative distribution functions of adult emergence time and (wing length)^3 were plotted over time, see Figure 6.5. Note how females and males are distinct in their development time and adult weight.



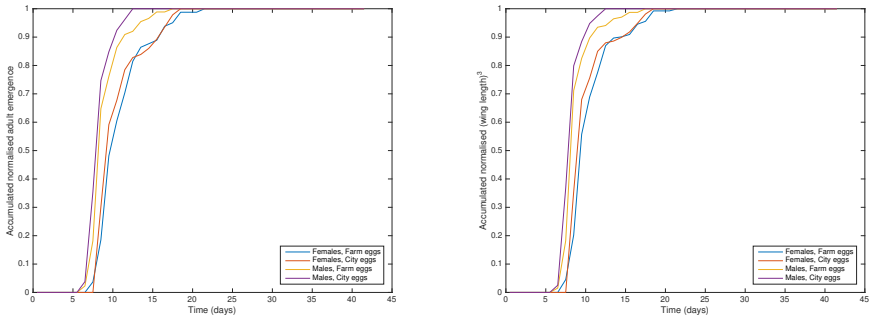
(a) Adult emergence **(b)** (Wing length)^3

Figure 6.5: Empirical cumulative distribution functions of adult emergence and (wing length)^3 separated for females and males as well as city and farm populations.

The graphs indicate that there are differences in how the populations develop. To see where differences lie, same sexes between populations were Kolmogorov-Smirnov tested. Adult emergence times for each of the treatments were cross-tested, farm against city population and vice versa for males and females respectively. Experimental data was used for one ($\sim$ 15 samples) and synthetic data (3000 samples) for the other. Evidently, the tests indicate differences between farm and city populations in treatments (Farm, Mid), (Farm, Low) and (Yeast,

High) for females and (Farm, High), (Farm, Mid), (Farm, Low), (City, Mid) and (Yeast, Low) for males. The detailed outcomes of the tests are presented in the appendix, Tabular A.5.

The treatments indicating differences were removed and accumulated adult emergence and (wing length)^3 of the remaining treatments were once again plotted, see Figure 6.6, to get a new visual understanding.



(a) Adult emergence **(b)** (Wing length)^3

Figure 6.6: Accumulated adult emergence and (wing length)^3 for remaining treatments after removing all treatments showing a significant difference in the Kolmogorov-Smirnov tests. The data have been separated into females and males as well as city and farm populations.

6.6 Span of data within treatments

Each treatment was separated based on both sex and population to investigate the range of the experimental data. The recorded values of first and last adult emergence as well as largest and smallest (wing length)^3 were compiled, see Table 6.4 and 6.5.

Table 6.4: Span of adult emergence within treatments, divided into sex and population.

Treatment	Males CE	Males FE	Females CE	Females FE
Farm, High	7.5-8.5	7.5-9.5	8.5-10.5	7.5-11.5
Farm, Mid	6.5-8.5	6.5-11.5	8.5-11.5	8.5-12.5
Farm, Low	14.5-27.5	22.5-41.5	15.5-28.5	22.5-38.5
City, High	7.5-8.5	7.5-9.5	8.5-9.5	7.5-10.5
City, Mid	7.5-8.5	7.5-10.5	8.5-9.5	8.5-11.5
City, Low	8.5-12.5	8.5-13.5	9.5-12.5	8.5-12.5
Yeast, High	7.5-9.5	7.5-9.5	8.5-10.5	8.5-11.5
Yeast, Low	9.5-12.5	10.5-15.5	10.5-18.5	12.5-21.5

In Table 6.4, it can be seen that the upper values of adult emergence in farm population are always higher or equal to the same values in the city population. The same holds almost true for the lower values, with exception for females in treatments (Farm, High), (City, High) and (City, Low). The difference between the upper and lower values are strictly higher for the farm population in all treatments except (Yeast, Low), where they have equal spans.

Table 6.5: Span of (Wing length)^3 within treatments, divided into sex and population.

Treatment	Males CE	Males FE	Females CE	Females FE
Farm, High	9.84-13.65	10.09-14.96	18.17-28.34	22.31-28.43
Farm, Mid	4.72-7.30	5.43-11.30	10.97-26.62	10.84-23.91
Farm, Low	3.31-8.68	3.45-12.47	7.01-17.96	8.82-16.41
City, High	8.22-12.93	9.70-14.28	21.39-30.37	22.86-33.54
City, Mid	10.30-13.46	9.16-14.23	22.79-26.68	21.32-30.40
City, Low	5.80-8.90	7.01-10.11	12.71-20.28	13.96-18.40
Yeast, High	9.97-15.38	11.44-16.66	22.24-30.08	19.53-32.61
Yeast, Low	5.05-7.44	5.13-8.49	5.77-18.99	7.27-23.76

Based on Table 6.5, the upper values for male (wing length)^3 in the farm population are larger than in city population, and the same trend can be seen in the lower values with only exception in (City, Mid). For females however, the trends are much less significant. The upper values of (wing length)^3 of farm population are

only larger for all treatments with High food levels as well as for Yeast treatments. The lower values are somewhat more consistent with the trend, only deviating with larger values for the city population in (Farm, Mid), (City, Mid) and (Yeast, High).

6.7 Analysis of deviating values

An early check of the experimental data was conducted to detect extreme deviations which might affect any of the analysis. Box and whisker plots were used to present the adult emergence times for each treatment in graphs, see Figure 6.7 for females and 6.8 for males. The graph summarises the data set in an explanatory way to show the central values, the shape of the distributions and its variability. In the graph, the median value is marked by the red line, the 25th and 75th percentiles by the top and bottom edges of the box and the whiskers extend to the most extreme data points not considered outliers. The outliers are plotted individually marked by + signs.

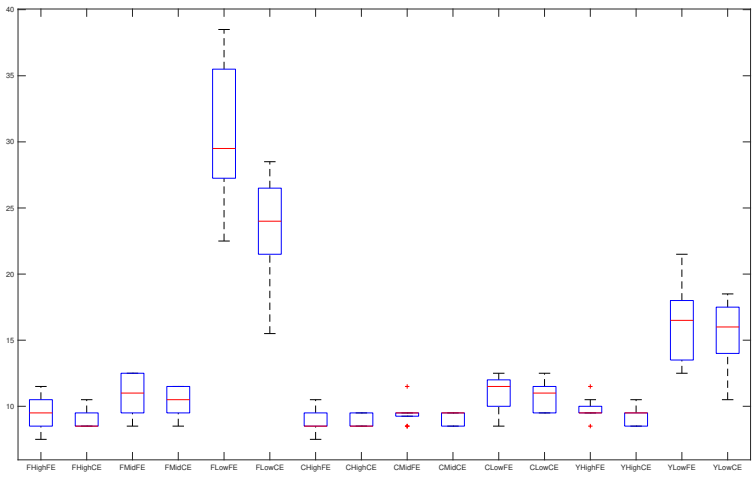


Figure 6.7: Box plot of time to adult emergence of female individuals of each treatment.

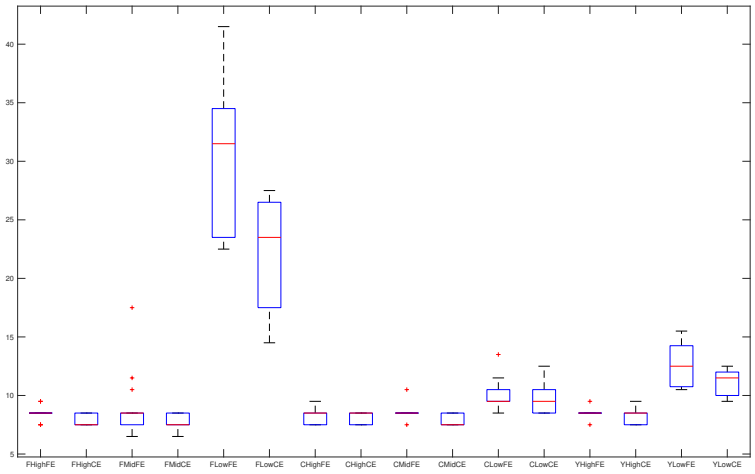


Figure 6.8: Box plot of time to adult emergence of male individuals of each treatment. Note the most deviant outlier as the + outside of the whiskers for (Farm, Mid).

From the graphs in Figure 6.7 and 6.8, one outlier in particular can be seen as significantly deviant, namely the upmost cross in males (Farm, Mid). It is clear that this value is extreme in its deviation and it is not likely that the deviation arises from natural variability within the distribution. This deviating value was removed from the data set prior any analysis to avoid disturbance.

The figures also show a graphical representation of the adult emergence times in relation to the food levels. As expected from previous results, levels High and Mid shows to be quite similar while level Low shows greater difference to the others, although to varying degree. The scarcity is most visible in the (Farm, Low) treatment. From Figure 6.8 and 6.7 it is also clear that the food levels are too different to be quantitatively comparable between treatments Farm, City and Yeast.

6.8 Measure of efficiency

An efficiency measure was introduced in order to investigate how much biomass each treatment generates on average.

Let P_i^x be the average weight per individual within sexes for treatment i $\in$ {(T,L); where T = {Farm, City, Yeast}, L = {High, Mid, Low}} and population x $\in$ {CE, FE}. Let A_i^x be the proportion of eggs[2] of each treatment surviving to adulthood, i.e. A=$1 - \frac{d}{N}$ where d is the number of deaths within a treatment and N is the total

[2]Note that the eggs naturally cannot yet be separated into males and females.

number of larvae for that treatment. Some treatments contain missing data, resulting in $N < 30$. The measure of efficiency for each treatment, population and sex is given by $\mathrm{Eff}_i^x = \mathrm{A}_i^x \cdot \mathrm{P}_i^x$. Note that the values of A are equal for both sexes in a treatment within a population.

The measure of efficiency for each treatment was computed and the results are presented in Table 6.6. It can be seen that the farm population has a higher measure of efficiency for both females and males with two exceptions only, (City, Mid) for females and (Farm, High) for males. The differences between the populations within these treatments are however very small.

Table 6.6: Measure of efficiency for females and males in city and farm population.

Treatment	Females CE	Females FE	Males CE	Males FE
Farm, High	23.07	23.20	11.87	11.68
Farm, Mid	16.92	17.23	5.75	7.06
Farm, Low	9.62	10.18	5.31	5.45
City, High	25.15	25.79	10.45	11.46
City, Mid	23.87	23.86	11.09	11.31
City, Low	14.80	15.03	6.90	7.54
Yeast, High	26.96	27.73	11.58	12.78
Yeast, Low	12.41	13.48	5.45	5.70

Chapter 7
Discussion

Comparisons with previous work, by V. Romeo Aznar (2015) [29], showed clear similarities in the graphs of Normalised accumulated weight vs. Normalised accumulated pupation and in the Variance vs. Mean of adult emergence time. This is a confirmation that the experiment in this thesis was well conducted and further that it is likely that the subpopulations involved in the experiment belong to the same population included in the previous study which provided confidence to move forward with the analysis.

The investigation of the replicas was a technical intermediate step that occupied a large portion of the thesis analysis time. However, the methodology developed is general, does not only apply to this study and account for much of the mathematical contribution. The mathematical knowledge and methods enable a robustness to the analysis and results that biology and statistics alone could not provide. With that in the luggage, it was possible to move on with the distributional analysis.

In the analysis of the eggs the first new result was discovered. It was seen that the eggs differ in shape between city and farm origin. The eggs did not arise from the same distribution and data was put together in order to discover where the differences could be found. After further analysis it was concluded that city individuals had a shorter development time and were smaller in adult size. The farm individuals had a longer development time and a larger adult size which could contribute to better reproduction opportunities since the amount of eggs a female lays is related to her size. This will be further discussed shortly.

It was seen that less food resulted in longer development time and less weight. The response to food scarcity follows in general terms the expected behaviour. The differences between subpopulations are evident in scarcity conditions, which is reasonable – when the mosquitoes are close to a crisis situation the strategic differences are relevant while in an abundance situation, the strategy differences are less important. The thesis provides indications that the food-levels are not quantitatively equal between food types. In practice, it does not matter for High and Mid, but the reaction to Low level food is systematically equal but to varying degrees.

It is also concluded that males are significantly smaller than females, which was expected. The measure of efficiency enables the discovery that the mosquitoes can adapt to different nutritional assets. Farm eggs are slightly more effective in converting nutrition into adult size. It is concluded that farm eggs have a somewhat longer development time and becomes a little larger than city eggs. It has little or no meaning to decide which strategy is "better". Larger adult weight is related to larger amount of laid eggs, which is good for the survival of the group, but to get there a longer development time is needed, which means a longer exposure to natural hazards, and a larger risk of never getting to adulthood. However, all discussed differences can be considered to be small.

During the course of writing this thesis a few questions were encountered and since they may have been encountered by the reader as well, they will now be addressed. The first question handled the replicas and whether or not the difference in number of females and males among replicas could be explained as fluctuations within the same distribution. The answer to this question is simply – yes. With a limited number of ~ 10 samples it is very likely that the ratio of females and males vary among replicas even though the overall ratio is roughly 50/50.

The second question was linked to the limited data collection. Is it possible that the limited number of 480 individuals could bias the results? This cannot be discarded – but to prevent this effect a new method was invented and used, namely the extension of the experimental frequencies for larval development. The extension introduced more possible events than the strictly experimental, which allowed some variation. It turned out that the additional possibilities did not alter the main results but only added a robustness to them.

The limited scope of the experiment did not allow for more than 30 individuals for each treatment which could be considered a small data sample. This was also addressed by the extension of the experimental frequencies, in combination with the synthetically produced data.

It cannot be ruled out that the small differences that were discovered arise by chance. This thesis shows indications that there exist differences between farm and city populations from the egg-stage throughout to adulthood. They are however indications. The more indications, the more it is likely that they are systematically true and the less they are likely to be random. It can only be claimed that suspicions about population differences are justified. Compared with other mosquito studies where results are considered relevant, experiments consist of $100 - 200$ individuals [29]. In this thesis the experiment has 480 individuals in support.

7.1 Conclusions

The question addressed was to understand variability and adaptation in terms of the interplay between environment and genetics. In particular, it has been proven that eggs with different origin display different adaptation response to food resources. Some of the visible ecological consequences have also been shown. The eggs are different in terms of size, where farm eggs are larger than city eggs. Individuals from rural origin have longer development time, their adult sizes are larger, and their measure of efficiency is higher.

In summary, the experiment shows small differences in individuals all the way from egg to adult depending on origin and differences in growth strategies. The differences are shown regardless of the type of nutrition the mosquitoes adopt, albeit to varying degrees. The thesis warns that assumptions about homogeneity or uniformity within species, that are often made in mosquito control contexts, should be called into question.

Chapter 8
Future work

There is much work to do in the area of modelling the dynamics of Ae. aegypti. As this thesis shows, a general model is not enough to explain the population dynamics of the species, it needs to be narrowed down and subpopulations need to be verified and modelled. It would be of interest to identify more subpopulations and to further investigate how different subpopulations vary between areas.

A general characteristic of models is that they in some sense reproduce a previously known mosquito. However, this thesis suggests that adaptation is permanently going on. The challenge for modelling techniques would be to incorporate properties such as adaptation and environmental pressure selection.

In this thesis, constant conditions were held to discover if any differences existed between subpopulations. After concluding that it actually does, further experiments can be carried out to investigate how the subpopulations react when outer conditions are varying.

To continue on this work, it would be of interest to verify the food levels of the treatments. This was a first attempt and could be used as a good reference of the different levels. Treatment (Farm, Low) showed the most significant differences between the subpopulations and it would be interesting to find a similar level of scarcity for City food. When investigating the optimal levels, a better way to outline the experiment would be to introduce more frequent inspections per day. At an optimal food level, the windows between first and last pupation and first and last adult emergence are very small. When only having one inspection per day, the risk of losing valuable information of differences between the subpopulations increase significantly. Hence, new experiments for checking differences between the populations at an optimal food level is recommended. Even though the experiment of this thesis did not show significant differences at the optimal levels, there might be differences hiding in said possible loss of data. The experiment of this thesis was limited to only investigating differences at the scarcity-optimal range. For future work, experiments on the optimal-abundance range could also be done.

When knowing the appropriate food levels, the natural next step would be to create models and adjust parameters similar to previous work by V. Romeo Aznar et al (2015).

References

[1] Alfaro-Murillo J.A., Parpia A.S., Fitzpatrick M.C., Tamagnan J.A., Medlock J., Ndeffo-bah M.L., "A Cost- Effectiveness Tool for Informing policies on Zika Virus Control", *PLoS Negl Trop Dis*, 10(5), 2016, `https://doi.org/10.1371/journal.pntd.0004743`

[2] Ananya B. and Andrew J., "Morphology and Morphometry of Aedes aegypti Larvae", *Department of Zoology and School of Entomology, St. John's College*, 2013

[3] Arrivillaga J. and Barrera R., "Food as a limiting factor for Aedes aegypti in water- storage containers", *Vector Ecology*, 29, pp. 11–20, 2004

[4] Bar-Zeev M., "The effect of temperature on the growth rate and survival of the immature stages of Aedes aegypti", *Bull. Entomol*, 49, 1958

[5] Blom G., Enger J., Englund G., Grandell J., Holst L., "Sannolikhetsteori och statistikteori med tillämpningar", Studentlitteratur, 2005, 91-44-02442-8

[6] Carr J.L., Gibbs B., Madani H. and Allen N., "INR performance of the GOES-13 spacecraft measured in orbit", *Advances in the Astronautical Sciences*, 131, 2008

[7] Christophers R., "Aedes aegypti - the yellow fever mosquito", *Cambridge University Press*, 1960

[8] Craig Jr. G.B., Vandehey R.C., Hickey W.A., "Genetic variability in populations of Aedes aegypti", *Bulletin of the World Health Organization*, 24, pp. 527 – 539, 1961

[9] Dégallier P.N., Hervé J.P., A.F.A.T.D., Sa, G.C., "Aedes aegypti - importance de sa bioécologie dans la transmission de la dengue et des autres arbobirus", *Bulletin de la Societe de Pathologie Exotique*, 1, pp. 97–110, 1988

[10] Gubler D.J., "Dengue and dengue hemorrhagic fever", *Clinical Microbiology Review*, 11, pp. 480–496, 1998

[11] Haase R., "Introduction and Review of Univariate General Linear Models", *Multivariate General Linear Models*, pp. 1–26, 2014

[12] Hamilton M. B., "Population Genetics", Wiley-Blackwell, 2009

[13] Hennessey M., Fischer M. and Staples J.E., "Zika Virus spreads to new areas — region of the Americas", *MMWR Morb. Mortal. Wkly*, 65, 2016

[14] Heuvel M.J., "The effect of rearing temperature on the wing length thorax length, leg length, and ovariole number of the adult mosquito, Aedes aegypti (L.)", *Trans. R. Entomol*, 7, 1963

[15] Huber K., Le Loan L., Hoang T.H., Ravel S., Rodhain F. and Failloux A-B., "Genetic differentiation of the dengue vector, aedes aegypti (Ho Chi Minh City, Vietnam) using microsatellite markers", *Molecular Ecology*, 9, 2002

[16] Huber K., Le Loan L., Hoang T.H., Tien T.K., Rodhain F. and Failloux A-B., "Temporal genetic variation in aedes aegypti populations in Ho Chi Minh City (Vietnam)) using microsatellite markers", *Heredity*, 1, 2002

[17] Legros M., Otero M., Romeo Aznar V., Solari H., Gould F., and Lloyd A., "Comparison of two detailed models of Aedes aegypti population dynamics", *Ecosphere*, 7, pp. 1–19, 2016

[18] MathWorks, "Multinomial random numbers - MATLAB mnrnd - MathWorks Nordic", `https://se.mathworks.com/help/stats/mnrnd.html`, 2019, Accessed: 2019-06-01

[19] Meghani Z. and Boëte C., "Genetically engineered mosquitoes, Zika and other arboviruses, community engagement, costs, and patents: Ethical issues", *PLOS Neglected Tropical Diseases*, 12(7), 2018, `https://doi.org/10.1371/journal.pntd.0006501`

[20] OECD, "Safety Assessment of Transgenic Organisms in the Environment, Volume 8: OECD Consensus Document of the Biology of Mosquito Aedes aegypti", *Harmonisation of Regulatory Oversight in Biotechnology*, OECD Publishing, Paris, 2018, `http://dx.doi.org/10.1787/9789264302235-en`

[21] Orlando Giménez J., Fischer S., Zalazar L. and Stein M., "Cold season mortality under natural conditions and subsequent hatching response of aedes (stegomyia) aegypti (diptera: Culicidae) eggs in a sub-tropical city of Argentina", *Journal of medical entomology*, 5, pp. 879–885 2015

[22] Otero M., Barmak D.H., Dorso C.O., Solari H.G. and Natiello M.A., "Modeling dengue outbreaks", *Mathematical Biosciences*, 232(2), pp. 87–95, 2011

[23] Padmanabha H., Correa F., Legros M., Nijhout H.F., Lord C. and Lounibos L.P., "An eco-physiological model of the impact of temperature on Aedes aegypti life history traits", *Journal of Insect Physiology*, 12, pp. 1597–1608, 2011

[24] Paupy C., Ngan C., Huber K., Lecoz N., Reynes J-M., Roghain F. and Failloux A-B., "Influence of breeding sites features on genetic differentiation of Aedes aegypti populations analyzed on a local scale in Phnom Penh municipality of Cambodia", *The American journal of tropical medicine and hygiene*, 1, pp. 73–81, 2004

[25] Peacock J.A., "Two-dimensional goodness-of-fit testing in astronomy", *Chemistry Letters*, 1983

[26] Resnik D.B., "Field Trials of Genetically Modified Mosquitoes and Public Health Ethics", *American Journal of Bioethics*, 17(9), pp. 24–26, 2017, https://doi.org/10.1080/15265161.2017.1353170

[27] Roland H.E. and Moriarty B., "Multinomial Distribution", *System Safety Engineering and Management*, 2009, https://doi.org/10.1002/9780470172438.ch11, Accessed: 2019-07-05

[28] Romeo Aznar V., Otero M., De Majo M., Fischer S. and Solari H., "Modeling the complex hatching and development of Aedes aegypti in temperate climates", *Ecological Modelling*, 253, pp. 44–55, 2013, http://dx.doi.org/10.1016/j.ecolmodel.2012.12.004

[29] Romeo Aznar V., De Majo M., Fischer S., Francisco D., Natiello M. and Solari H., "A model for the development of Aedes (Stegomyia) aegypti as a function of the available food", *Journal of Theoretical Biology*, 365, pp. 311–324, 2015, http://dx.doi.org/10.1016/j.jtbi.2014.10.016

[30] Romeo Aznar V., Alem I., De Majo M., Byttebier B., Fischer S. and Solari H., "Effects of scarcity and excess of larval food on life history traits of Aedes aegypti (Diptera: Culicidae)", *Journal of Vector Ecology*, 43, pp. 117–124, 2018

[31] Rueda L.M., Patel K.J., Axtell R.C., Stinner R.E., "Temperature-dependent development and survival rates of Culex quinquefasciatus and Aedes aegypti (Diptera: Culicidae)", *Med. Entomol*, 27, 1990

[32] Subra R. and Mouchet J., "The regulation of preimaginal populations of Aedes aegypti (L.) (Diptera: Culicidae) on the Kenya Coast. II. Food as a main regulatory factor", *Ann. Trop. Med. Parasitol*, 78, 1984

[33] World Health Organization, "Guidance framework for testing of genetically modified mosquitoes", Geneva, Switzerland: World Health Organization, 2014

Appendix A

Some extra material

A.1 Males, females and number of deaths

Table A.1 shows the distribution of sexes within treatments and Table A.2 shows the number of deaths recorded in the experiment.

Table A.1: Number of individuals in each treatment, separated into sex and population. Data missing for two mosquitoes, treatment (Farm, Low, CE) and (Farm, Low, FE), hence a total of 29. Also note that the deviating value in males (Farm, Mid, FE) has not been counted.

Treatment	Males CE	Females CE	Males FE	Females FE
F, High	14	16	18	10
F, Mid	16	13	15	14
F, Low	8	18	7	17
C, High	13	15	15	13
C, Mid	17	12	15	13
C, Low	16	12	16	12
Y, High	14	16	12	16
Y, Low	8	20	11	16

Table A.2: Shows number of deaths within treatments, separated into population.

Treatment	CE	FE
F, High	0	2
F, Mid	1	0
F, Low	3	5
C, High	2	2
C, Mid	1	2
C, Low	2	2
Y, High	0	1
Y, Low	2	3

A.2 Results of Kolmogorov-Smirnov tests

Kolmogorov-Smirnov tests were conducted with males and females separated. Tables A.3 and A.4 show results from Kolmogorov-Smirnov tests conducted on both adult emergence time and (wing length)^3 between food levels within populations.

Table A.5 show results from Kolmogorov-Smirnov tests conducted on adult emergence times between city and farm populations within food treatments. Experimental data ($\sim$ 15 samples) is tested against synthetic data (3000 samples) and then again reversed.

The logical value 1 indicates differences within the tested food levels, whereas 0 means that the test does not prove differences. In Table A.5, the p-value $\in [0, 1]$ is also included as reference.

Table A.3: Results from Kolmogorov-Smirnov tests for males, conducted on adult emergence times as well as wing lengths, between food levels within populations. The logical value 1 indicates differences within the tested food levels, whereas 0 means that the test does not prove differences.

Food	Egg	Level	Adult emergence	Wing length
F	FE	High & Low	1	1
		High & Mid	0	1
		Mid & Low	1	0
	CE	High & Low	1	1
		High & Mid	0	1
		Mid & Low	1	0
C	FE	High & Low	1	1
		High & Mid	0	0
		Mid & Low	1	1
	CE	High & Low	1	1
		High & Mid	0	0
		Mid & Low	1	1
Y	FE	High & Low	1	1
	CE	High & Low	1	1

Table A.4: Results from Kolmogorov-Smirnov tests for females, conducted on adult emergence times as well as wing lengths, between food levels within populations. The logical value 1 indicates differences within the tested food levels, whereas 0 means that the test does not prove differences.

Food	Egg	Level	Adult emergence	Wing length
F	FE	High & Low	1	1
		High & Mid	0	1
		Mid & Low	1	1
	CE	High & Low	1	1
		High & Mid	1	1
		Mid & Low	1	1
C	FE	High & Low	1	1
		High & Mid	0	0
		Mid & Low	1	1
	CE	High & Low	1	1
		High & Mid	0	1
		Mid & Low	1	1
Y	FE	High & Low	1	1
	CE	High & Low	1	1

Table A.5: Results from Kolmogorov-Smirnov tests with males and females separated, conducted on adult emergence times, between city and farm populations within treatments. The logical value 1 indicates differences between populations within the tested food treatment and level, whereas 0 means that the test does not prove differences.

Food	Level	Experimental	Synthetic	p-value	Logical value
Males					
F	High	F	C	0.0033	1
		C	F	0.0180	1
	Mid	F	C	0.0912	0
		C	F	0.3981	0
	Low	F	C	0.0046	1
		C	F	$6.52 \cdot 10^{-4}$	1
C	High	F	C	1.0000	0
		C	F	1.0000	0
	Mid	F	C	0.0019	1
		C	F	$8.36 \cdot 10^{-4}$	1
	Low	F	C	0.4496	0
		C	F	0.9979	0
Y	High	F	C	0.5325	0
		C	F	0.5786	0
	Low	F	C	0.0138	1
		C	F	0.2698	0
Females					
F	High	F	C	0.4463	0
		C	F	0.4513	0
	Mid	F	C	0.0421	1
		C	F	0.3664	0
	Low	F	C	$7.16 \cdot 10^{-9}$	1
		C	F	$8.29 \cdot 10^{-7}$	1
C	High	F	C	0.8923	0
		C	F	0.7524	0
	Mid	F	C	0.7105	0
		C	F	0.8502	0
	Low	F	C	0.6221	0
		C	F	0.8574	0
Y	High	F	C	0.0027	1
		C	F	0.0032	1
	Low	F	C	0.2185	0
		C	F	0.3109	0

Master's Theses in Mathematical Sciences 2019:E57
ISSN 1404-6342

LUTFMA-3392-2019

Mathematics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lth.se/