Predicting the Spatio-temporal Reproductive Potential of *Aedes aegypti*

Tarek Alrefae [1] Túlio de Oliveira [2] José Lourenço [3]

9 August 2023

[1]Nuffield Department of Medicine, University of Oxford
[2]Director, Centre for Epidemic Response and Innovation
[3]Biosystems and Integrative Sciences Institute, Faculty of Sciences, University of Lisboa

Abstract

We extend an existing climate-dependent measure of *Aedes aegypti* reproductive potential, Index $\mathcal{Q}$, and use approximate Bayesian computation and mosquito abundance data to fit two of its unknown parameters. We consider local solutions for four urban areas where abundance data is available; we also propose an approximate global solution for making projections in cases where local data is unavailable. We specifically consider a set of six geographically and climatically diverse urban areas with varying risk exposures to mosquito-borne virus transmission. Spatial explorations of Brazil and global spatial estimates of $\mathcal{Q}$ complete our investigation.

Contents

1 Introduction **1**

2 Methods **2**

2.1 Index $\mathcal{Q}$ 2

2.2 Approximate Bayesian Computation 4

2.3 Data 7

3 Results & Discussion **9**

3.1 Showcase $\mathcal{Q}$ 9

3.2 Sensitivity of $\mathcal{Q}$ 11

3.3 Local $\mathcal{Q}$ 12

3.4 Approximate Global $\mathcal{Q}$ 14

3.5 Local vs. Global $\mathcal{Q}$ 19

3.6 Spatial $\mathcal{Q}$ 22

4 Conclusions **25**

5 Limitations & Future Directions **25**

References **27**

S1 Methods **I**

S2 Results **V**

1 Introduction

The abundance of mosquito-vectors of public health importance is highly correlated with the timing of outbreaks of mosquito-borne viruses (MBVs) (H. M. Yang et al. 2009; Obolski et al. 2019; Faridah et al. 2022). This correlation necessitates an understanding of abundance dynamics and motivates predictions in both space and time. To make these predictions, in this study we extend a previously proposed theoretical model of mosquito reproductive potential, to be introduced shortly as *Index* $\mathcal{Q}$ (Lourenço et al. 2017). This reproductive measure is a function of local temperature, humidity, and precipitation. The choice to pursue this investigation through the lens of a climate-driven metric reflects the growing impact of rising global temperatures and climate change on the abundance and population dynamics of mosquitoes (Groen et al. 2017). This invariably impacts the spread of MBVs (ECDC 2018) – the ultimate goal of our investigation via our reproductive index is to identify areas with the greatest risk of moving into the mosquito reproductive 'Goldilocks zone' and therefore increasing the risk of MBV transmission. The identification of these areas and any trends therein with respect to their spatial distribution can help scientists and public health professionals to model and efficiently implement coherent, evidence-based mosquito control and outbreak mitigation measures.

The focus of our investigation is limited to the *Aedes aegypti* species given its wide geographical range spanning tropical/subtropical/temperate areas across Africa, Asia, and the Americas. It is also the primary vector for a plethora of MBVs (Zika, dengue, chikungunya, etc.) (ECDC 2023). The global burden of dengue alone is estimated to be 400 million yearly infections, with 80% of the world's population at risk for MBV transmission (Franklinos et al. 2019; WMP 2023).

In this study we perform an extensive analysis of Index $\mathcal{Q}$, showcasing its explanatory potential and utilising empirical mosquito abundance data in four cities in the United States and India to calibrate two of its key parameters.

2 Methods

2.1 Index $\mathcal{Q}$

Consider the ODE system describing mosquito population dynamics

$$\begin{cases} \dfrac{\mathrm{d}A}{\mathrm{d}t} &= \dot{c}\, f\, \dot{\theta}^{\nu} \left(1 - \dfrac{A}{K}\right) V \;-\; (\dot{\epsilon}_A^{\nu} \;+\; \dot{\mu}_A^{\nu})\, A \\ \dfrac{\mathrm{d}V}{\mathrm{d}t} &= \dot{\epsilon}_A^{\nu} A \;-\; \dot{\mu}_V^{\nu} V \end{cases} \tag{1}$$

for aquatic (A) and viable (i.e. female) adult (V) mosquito populations with carrying capacity K. This system is parameterised by the *aquatic-to-adult transition rate* ($\dot{\epsilon}_A^{\nu}$), the *aquatic mortality rate* ($\dot{\mu}_A^{\nu}$), the *intrinsic oviposition rate* ($\dot{\theta}^{\nu}$), the *sex ratio* (f), the *egg-hatching success probability* ($\dot{c}$), and the *adult mortality rate* ($\dot{\mu}_V^{\nu}$). All dotted parameters are strictly dependent on some combination of temperature/humidity/precipitation. The sex ratio f is fixed at 0.5 .

Then, Index $\mathcal{Q}$ is defined as the *basic offspring number*:

$$\mathcal{Q}\left(\boldsymbol{T},\, \boldsymbol{H},\, \boldsymbol{R};\;\rho,\, \eta\right) \;=\; \frac{\dot{\theta}^{\nu}}{\dot{\mu}_V^{\nu}} \;\; \frac{\dot{c}\, \dot{\epsilon}_A^{\nu}\, f}{\dot{\epsilon}_A^{\nu} \;+\; \dot{\mu}_A^{\nu}} \quad , \tag{2}$$

$$\begin{cases} \dot{\mu}_V^{\nu}\left(\boldsymbol{T},\, \boldsymbol{H};\;\rho,\, \eta\right) &= \eta \left(\sum_{i=0}^{4} a_i\, T^i\right) \left(1 \;+\; \bar{H} \;-\; \dfrac{\Delta_H}{\sqrt{1 \;+\; \Delta_H^2}}\right)^{\rho} & (3) \\ \dot{c}\left(\boldsymbol{T},\, \boldsymbol{R};\;\rho\right) &= \dfrac{1}{100} \left(\sum_{j=0}^{3} b_j\, T^j\right) \left(1 + \dfrac{\Delta_R}{\sqrt{1 \;+\; \Delta_R^2}}\right)^{\rho} & (4) \\ \dot{\epsilon}_A^{\nu}\left(\boldsymbol{T};\;\eta\right) &= \eta \sum_{k=0}^{7} c_k\, T^k & (5) \\ \dot{\mu}_A^{\nu}\left(\boldsymbol{T};\;\eta\right) &= \eta \sum_{l=0}^{4} d_l\, T^l & (6) \\ \dot{\theta}^{\nu}\left(\boldsymbol{T}\right) &= \sum_{n=0}^{4} w_n\, T^n & (7) \\ f &= \dfrac{1}{2} & (8) \end{cases}$$

for a vector of temperatures $\boldsymbol{T}$ given in °C, relative humidity vector $\boldsymbol{H}$, and precipitation vector $\boldsymbol{R}$. Both $\boldsymbol{H}$ and $\boldsymbol{R}$ are restricted to the unit interval – the elements of $\boldsymbol{R}$ are simply normalised while relative humidity percentage is divided by 100. $\bar{H}$ represents mean humidity and $\Delta_X := X - \bar{X}$.

The dependence of these entomological parameters on their respective climatic inputs are the result of laboratory experiments (Equations (3), (5) – (7); H. M. Yang et al. 2009) and empirical data fitting (Equation (4); Lourenço et al. 2017). Thus, the temperature polynomial coefficients a_i, b_j, c_k, d_l, and w_n in Equations (3) – (7) are known (see S1 Methods).

Parameter	**Description**	**Source**
$\dot{c}$	Egg-hatching success probability	Lourenço et al. 2017
$\dot{\mu}_V^\nu$	Adult mortality rate	H. M. Yang et al. 2009
$\dot{\epsilon}_A^\nu$	Aquatic-to-adult transition rate	Ibid.
$\dot{\mu}_A^\nu$	Aquatic mortality rate	Ibid.
$\dot{\theta}^\nu$	Intrinsic oviposition rate	Ibid.
f	Sex ratio	(Assumed)

Table 1: Known entomological parameters of Index $\mathcal{Q}$ and their sources.

The remaining undetermined parameters ρ and η are the focus of our investigation. η is a linear scaling coefficient which modulates the effect of temperature on the aquatic-to-adult transition rate (ϵ_A) and the aquatic (μ_A) and adult (μ_V) mortality rates. When $\eta = 1$, this can be interpreted as 'the effects of temperature on these parameters in the field and in laboratory settings are identical'; the effects are presumed stronger (weaker) when $\eta > 1$ ($\eta < 1$). Similarly for the nonlinear exponential scaling factor ρ, which determines the impact of both precipitation and humidity on the egg-hatching success probability (c) and adult mortality rate (μ_V), respectively: when $\rho = 0$, these impacts are assumed non-existent; when $0 < \rho < 1$, they are assumed weaker than in laboratory settings; when $\rho = 1$, they are assumed identical; these impacts are considered increasingly great as $\rho \to \infty$.

Index $\mathcal{Q}$ has a precise biological interpretation: it is the mean number of viable female offspring that a single female adult mosquito produces in her lifetime (Lourenço and Recker 2014; Lourenço et al. 2017; Pliego et al 2017). The first expression (θ/μ_V) gives the mean number of eggs laid by one female mosquito; the quantity $1/(\epsilon_A + \mu_A)$ is the mean survival period for laid eggs, implying that the second expression $(\epsilon_A c f)/(\epsilon_A + \mu_A)$ is the probability that a single laid egg survives and successfully develops into a female

adult mosquito. Thus the combination of the two expressions yields $\mathcal{Q}$ as defined. When $\mathcal{Q} > 1$, the implication of reproductive potential on subsequent abundance suggests the mosquito population will grow; when $0 < \mathcal{Q} \leq 1$, it suggests decay; and when $\mathcal{Q} = 0$, there is no reproductive potential (see S1 Methods).

Implicit in our use of $\mathcal{Q}$ to model *Aedes* reproductive potential are some key assumptions inherited from the ODE system (1): we assume homogeneous spatial mixing of mosquitoes, and we explicitly collapse the biologically distinct egg, larval, and pupal mosquito life stages into a single 'aquatic' stage. Though system (1) accounts for carrying capacity K, our formulation of $\mathcal{Q}$ ignores this and other known external pressures such as predation, competition, host availability, etc.

The spatio-temporal heterogeneity of $\mathcal{Q}$ is directly moderated by its input climate data set. For a given set of location-specific time-series of temperature, humidity, and precipitation, the corresponding $\mathcal{Q}$ time-series represents the effect of that locale's climate on the reproductive potential of *Aedes aegypti* over the specified time period.

The third and final parameter is the time-lag parameter δ. Though not an argument of $\mathcal{Q}$, we include δ to account for the theoretical lag between reproductive potential and actual mosquito abundance. The need to account for this potential lag is a function of the monthly aggregation of our climate and mosquito data (discussed further in Section 2.3). Given that *aegypti* larvae are known to develop into adults between 7-10 days (CDC 2014; Pliego et al 2017) and given the monthly resolution of available climate data, we restrict $\delta \in \{0, 1\}$ in months. When $\delta = 0$, $\mathcal{Q}$ is fully-defined by Equation (2); when $\delta = 1$, the resulting $\mathcal{Q}$ time-series is simply shifted to the right by one month (discarding the final value of the corresponding $\mathcal{Q}_0$). Our restriction of δ assumes that $\mathcal{Q}$ is either in sync with mosquito abundance ($\delta = 0$) or precedes it by one month ($\delta = 1$).

2.2 Approximate Bayesian Computation

An approach via *Markov chain Monte Carlo* (MCMC) methods was not feasible due to the absence of an analytical likelihood function. Such a likelihood is undefinable as the statistical variation of the abundance data is currently unknown. We instead leverage the power of *approximate Bayesian computation* (ABC), which allows for the estimation of approximate posterior distributions while foregoing the necessity of formulating an explicit likelihood altogether. To make full use of our Bayesian approach we implement indirect priors for ρ and η via the aquatic-to-adult development period (ϵ_A^{-1}) and the

adult lifespan (μ_V^{-1}):

$$\epsilon_A^{-1} \sim \Gamma(42.5,\ 5) \tag{9}$$

$$\mu_V^{-1} \sim \mathcal{N}^+(12,\ 2) \tag{10}$$

where Γ denotes the gamma distribution with shape 42.5 and rate 5 (i.e. mean 8.5 and variance 1.7) and $\mathcal{N}^+$ denotes the truncated normal distribution (μ_V must be strictly positive) with mean 12 and standard deviation 2. These priors are informed by *biological plausibility*: the aquatic-to-adult development period (ϵ_A^{-1}) is known to lie within $7-10$ days (CDC 2014; Pliego et al 2017); adult lifespan (μ_V^{-1}) has been shown to follow Equation (10) in laboratory settings (H. M. Yang et al. 2009). Combined with our direct uniform priors

$$\rho \sim \mathcal{U}(0, 10) \tag{11}$$

$$\eta \sim \mathcal{U}(0, 5) \tag{12}$$

our hierarchical Bayesian model is now fully specified. A preliminary coarse grid-search where $\rho, \eta \in [0.1, 20]$ with a 0.05 spacing informed Equations (11) - (12).

Algorithm 1 showcases our particular ABC implementation. The `corr` in line 12 is the *Pearson correlation coefficient* between normalised values of both $\mathcal{Q}_\delta$ and $\mathcal{D}$. Forgoing the likelihood function allowed the vectorisation of lines $8-15$, greatly enhancing computational performance. We chose the number of samples N to be 10^7. The concept of ensuring 'equal representation' in line 17 will be explicated shortly in Section 2.3.

Given approximate posterior distributions for ρ, η, and δ, we implement a final methodological step for making projections. Line 3 of Algorithm 2 was achieved via the `fitdistrplus` package in `R` (Delignette-Muller and Dutang 2015), considering multiple candidate distributions, and choosing that which minimised the *Akaike information criteria* (AIC).

Algorithm 1 ABC Algorithm

Input: Observed mosquito abundance data $\mathcal{D}$ and climate data $\mathcal{C}$
Output: Approximation of the posterior distribution of parameters ρ, η, δ
Normalise mosquito abundance data $\mathcal{D}$
Set the number of random points to be drawn N
Define the ranges for ρ, η, and δ
Randomly draw N points from the $\rho\eta$-plane (subject to their respective ranges)
Initialise an output data frame $\mathcal{O}$
for each locale **do**
  for all N points **do**
    Calculate μ_V^{-1}, ϵ_A^{-1} based on $\mathcal{C}$, ρ, and η
    Calculate the corresponding $\mathcal{Q}$ time-series via Equation (2) and normalise
    $\forall!$ δ, calculate $\texttt{corr}(\mathcal{D},\ \mathcal{Q}_\delta)$
    Store all calculations in $\mathcal{O}$
  end for
end for
Filter for biologically plausible (ρ, η) points subject to Equations (9) and (10)
Ensure equal representation between all locale's considered (removes sampling bias)
Choose those (ρ, η, δ) points which maximise the $\texttt{corr}(\mathcal{D},\ \mathcal{Q}_\delta)$ values (top 50%)
Output: Approximate posterior distributions for ρ, η, and δ

Algorithm 2 Index $\mathcal{Q}$ Projection Algorithm

Input: Approximate posteriors of ρ, η, and δ and climate data $\mathcal{C}_{\text{proj}}$
Output: Projections of Index $\mathcal{Q}$ in location corresponding to $\mathcal{C}_{\text{proj}}$
Find suitable analytical distributions for each of ρ, η, and δ
Draw M random samples from each analytical distribution
Calculate $\mathcal{Q}_\delta\left(\mathcal{C}_{\text{proj}};\ \rho,\ \eta\right)$ for each random sample and take mean of M samples
Output: M-sample mean Index $\mathcal{Q}$ projection for location of interest

2.3 Data

All climate data was sourced from the Copernicus data set provided by the EU (Hersbach et. al 2018) and was originally aggregated by monthly means; all subsequent investigations were carried out at a monthly temporal resolution.

Table 2 is a collection of four cities for which *aegypti* abundance data was available. These were used in Algorithm 1 to calibrate our ABC approach and obtain approximate posteriors of ρ, η, and δ.

Urban Area	**Time Frame (Mos.)**	**Climate (Köppen)**	**Source**
Hillsborough Co., FL, USA	12/2008 – 12/2017 (109)	Humid Subtropical (*Cfa*)	B. Yang et al. 2021
Miami-Dade Co., FL, USA	07/2016 – 02/2018 (20)	Tropical Monsoon (*Am*)	B. Yang et al. 2021
Maricopa Co., AZ, USA	03/2013 – 12/2017 (58)	Hot Desert (*BWh*)	JL[1]
Surat, Gujarat, India	01/1993 – 12/2012 (240)	Tropical Savannah (*Aw*)	JL[1]

Table 2

Once our posterior distributions for ρ, η, and δ were approximated, we applied Algorithm 2 to six urban areas of interest where MBVs spread by *Aedes* mosquitoes range from endemic to negligible. These are summarised in Table 3.

Urban Area	**MBV Endemicity**	**Climate (Köppen)**
Belém, Brazil	(i) Endemic	Tropical Rainforest (*Af*)
Darwin, Australia	(ii) Periodic	Tropical Savannah (*Aw*)
Istanbul, Türkiye	(ii) Periodic	Mediterranean (*Csa*)
Johannesburg, South Africa	(ii) Periodic	Subtropical Highland (*Cwb*)
Tokyo, Japan	(ii) Periodic	Humid Subtropical (*Cfa*)
Vancouver, Canada	(iii) Negligible	Temperate Oceanic (*Cfb*)

Table 3

[1]Please see the Data Availability Statement on page 26.

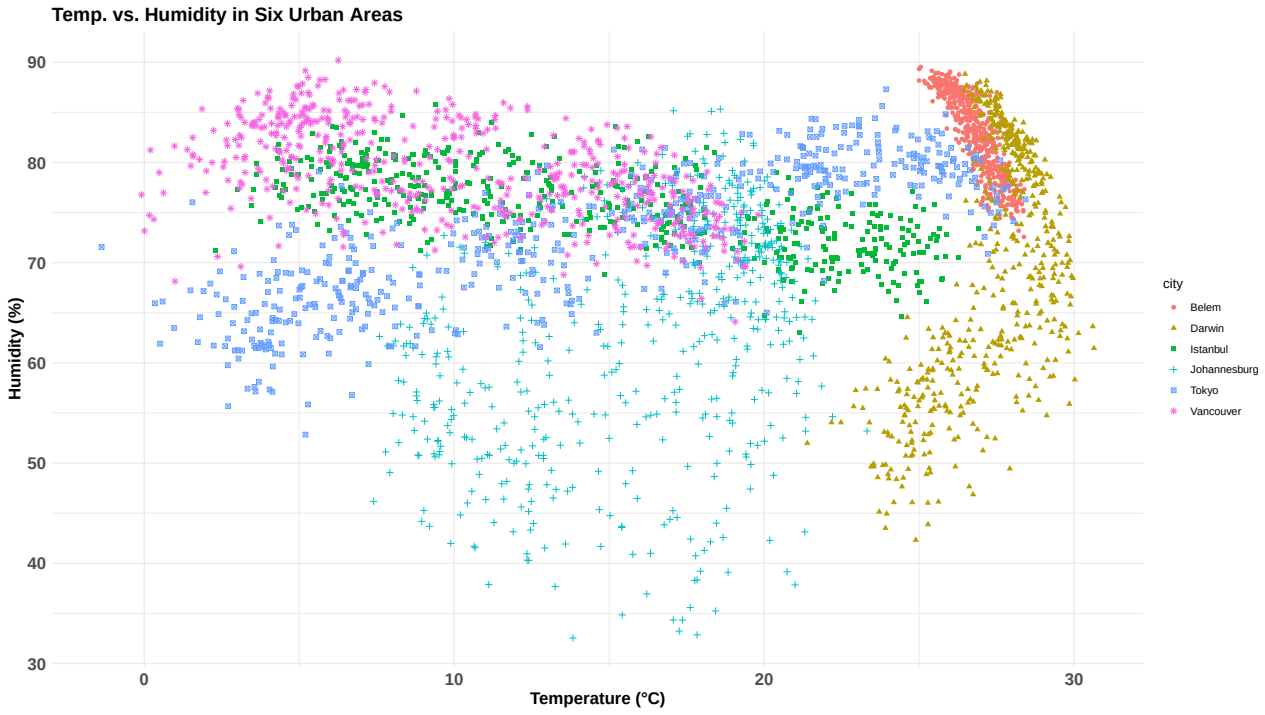


Figure 1: A plot of monthly mean relative humidity vs. temperature in °C for our six showcase cities from January 1979 to March 2023. This helps to contextualise their choice as a set of climatically diverse, representative urban areas which are suitable for eventual Index $\mathcal{Q}$ projections.

This distribution of urban areas is geographically and climatically diverse (Figure 1). Representing major cities on each continent save Antarctica, their diverse climates range from rainforest (Belém) to oceanic (Vancouver) to subtropical highland (Johannesburg). While dengue, chikungunya, and Zika virus are all currently significantly present only in Belém, sporadic and periodically imported dengue cases have been reported in all other cities save Vancouver (Schaffner and Mathis 2014; Quam et al. 2016; Msimang et al. 2018; NTGov 2023). Groups (ii) and (iii) in Table 3 represent cities at risk of increasing their MBV disease burden as climate change render them more and more suitable for mosquito reproduction for longer periods of time.

3 Results & Discussion

3.1 Showcase $\mathcal{Q}$

We begin by exploring the behavior of Index $\mathcal{Q}$ in each of our six showcase cities. We proceed with the 'baseline' set of parameter values

$$\Phi_0 \;:=\; \{\rho = 1, \quad \eta = 1, \quad \delta = 0\}.$$

That is, we assume that the climate-dependence of the various entomological parameters which combine to form $\mathcal{Q}$ as given in Equations (3) – (8) are identical to those dependencies derived from the laboratory experiments and data fitting processes from (H. M. Yang et al. 2009; Lourenço et al. 2017); we also assume no lag exists between mosquito reproductive potential and subsequent abundance.

Figure 2 displays the baseline behavior of Index $\mathcal{Q}$ in each of our six showcase cities for $\mathcal{Q}(\Phi_0)$. These six cities are climatically diverse, and this diversity informs their heterogeneity in $\mathcal{Q}$. Belém and Darwin are the only cities for which baseline $\mathcal{Q}$ never drops below the critical threshold $\mathcal{Q} = 1$; these are also the only two cities endemic to *Aedes*-spread MBVs [Zika, dengue, and chikungunya in Belém (WHO 2022a; WHO 2022b; WHO 2023); Ross River virus in Darwin (NTGov 2021)].

The tropical climates of Belém and Darwin make them very suitable for year-round successful *Aedes* reproduction. While other cities presented such as Tokyo, Istanbul and Johannesburg all reach relatively high levels of $\mathcal{Q}$ at baseline, they also crucially experience regular yearly periods in which $\mathcal{Q} = 0$, hampering the over-wintering of any existing *Aedes* populations there. These cities' respective winter climates form a significant barrier to MBV endemicity.

The top panel of Figure 2 shows that baseline $\mathcal{Q}$ is increasing with statistical significance between $1979-2023$ for four of the six showcase cities. (The two exceptions are Darwin, which displays a statistically insignificant negative slope, and Tokyo, which displays a statistically insignificant positive slope; $p > 0.05$). Projections for these cities using approximate global solutions for ρ, η, and δ are presented and discussed further in later sections.

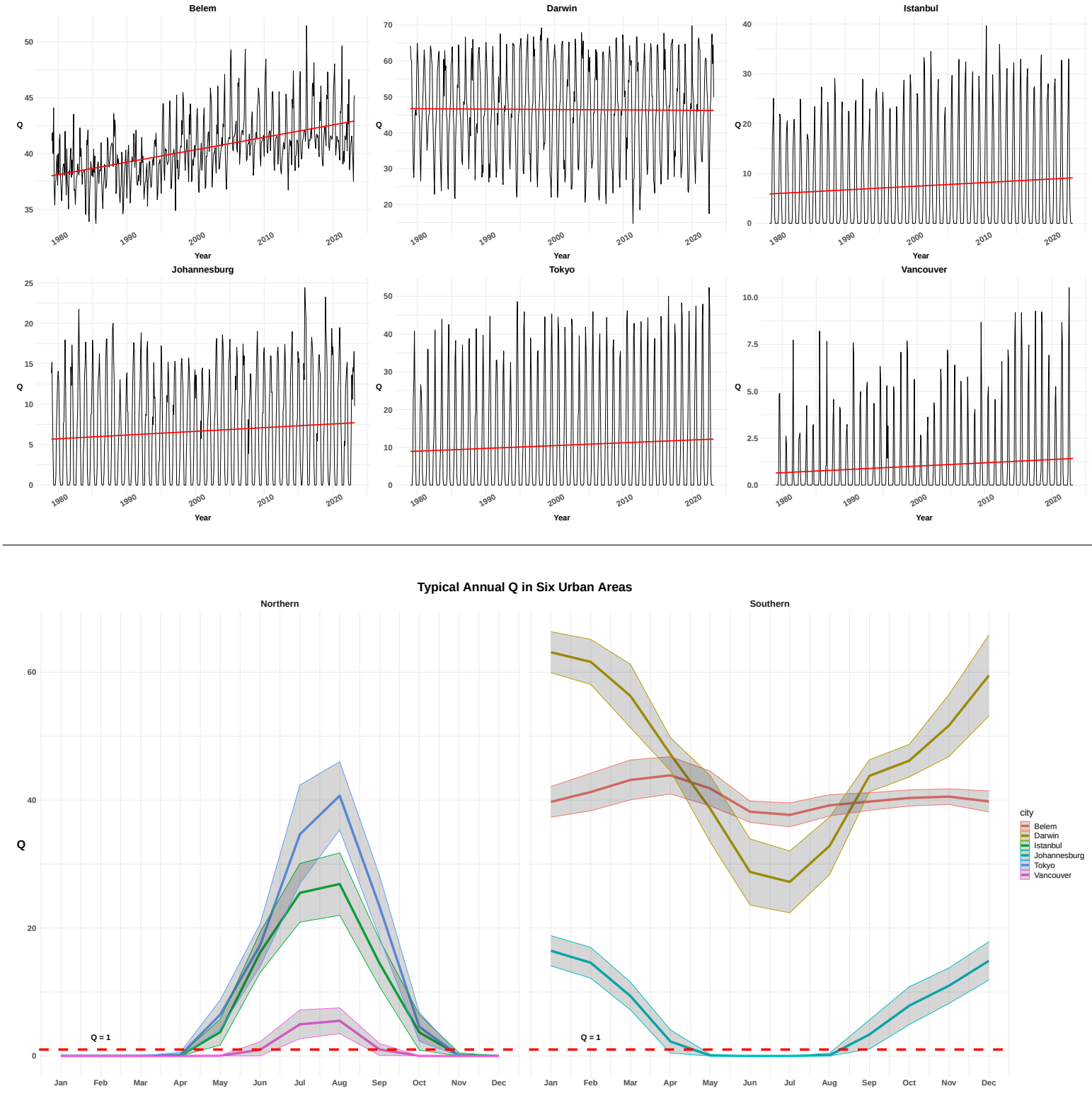


Figure 2: [Top] A multiplot of Index $\mathcal{Q}(\Phi_0)$ in each of the six cities between Jan. 1979 – March 2023. Note the differing scales of the y-axes in each city. The red lines are linear regressions of baseline $\mathcal{Q}$ in that city during that time period. [Bottom] Two plots of typical annual baseline Index $\mathcal{Q} \pm 1$ std. dev. across the six urban areas, separated by hemisphere. A typical year is an aggregation of a multi-year time-series by monthly averages. The red dashed line at $\mathcal{Q} = 1$ indicates the threshold level above which the mosquito-free equilibrium is unstable.

3.2 Sensitivity of $\mathcal{Q}$

It is unlikely that the climate-dependence of the various components of $\mathcal{Q}$ originally determined in controlled laboratory settings are perfectly representative of these relationships outside the laboratory. The parameters ρ and η allow for deviations from these original calculations, and we contend that these deviations are not spatially uniform, but rather particular to the climate of the area under consideration.

Figure 3 displays six heat maps where the sensitivity of $\mathcal{Q}$ (in particular its tendency to be above the critical threshold $\mathcal{Q} = 1$ throughout the year) to ρ and η is shown to be related with that area's climate. While simultaneously maximising ρ and η drives this proportion toward 0, each city has a unique curve in the $\rho\eta$-plane across which this transition occurs. More extreme values of ρ and η must be input in order to minimise the proportion of time that $\mathcal{Q} > 1$ for Belém and Darwin. Meanwhile in Vancouver, where *Aedes* mosquitoes are not epidemiologically significant, relaxing climatic effects by minimising ρ and η can not yield $\mathcal{Q} > 1$ for more than 50% of the time.

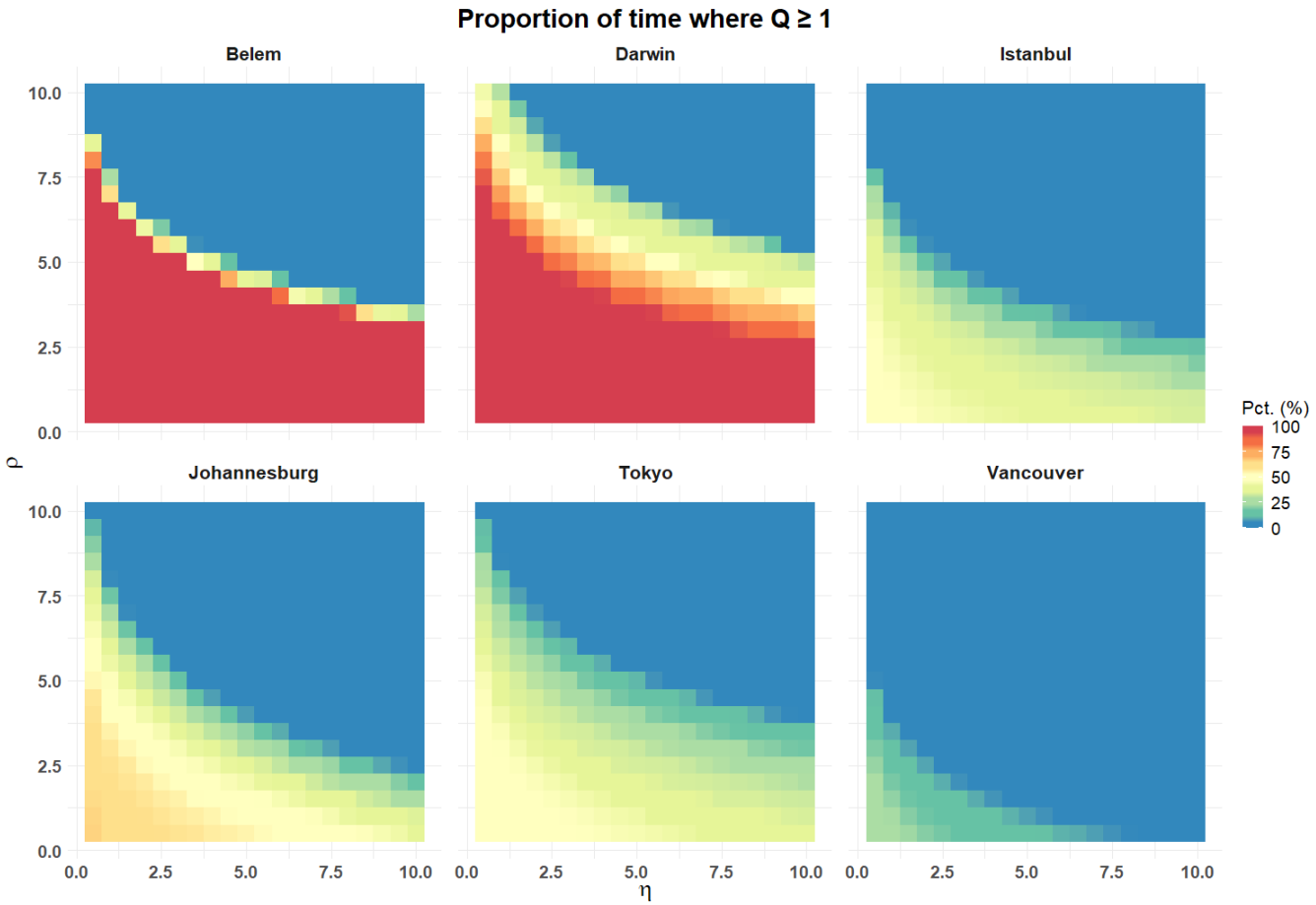


Figure 3: A set of six heat maps displaying the sensitivity of the proportion of time during the year where $\mathcal{Q} > 1$ to ρ and η for each showcase city described in Section 2.3. Here $\rho, \eta \in (0, 10]$ while $\delta = 0$. The time-series analysed here are the full $1979 - 2023$ time-series seen in the top panel of Figure 2.

This heterogeneity in the behaviour of $\mathcal{Q}$ and its sensitivity to ρ and η suggests that these laboratory-deviation measures are not only necessary but highly location- and climate-specific. The role that rainfall plays in mosquito development in a tropical rainforest is likely to be different than its role in a Mediterranean climate like Istanbul.

3.3 Local $\mathcal{Q}$

Applying Algorithm 1 to each of the four calibration locales from Section 2.3 independently further emphasises the notion that $\mathcal{Q}$ is critically dependent on local climate. (Applying ABC to each location individually removes the necessity of ensuring equal representation between sites as in line 17, but this will be re-introduced shortly in Section 3.4 when we collate our mosquito abundance and climate data).

Figure 4 displays the result of choosing those $\Phi_{\max}$ parameter configurations which resulted in the maximum correlation between normalised *Aedes* counts and associated $\mathcal{Q}(\Phi_{\max})$ in each city. These maximal correlations ranged from 0.64 (Surat) – 0.94 (Miami/Hillsborough). The elements of $\Phi_{\max}$ for each city are given in Table 4.

City	$\boldsymbol{\rho}_{\max}$	$\boldsymbol{\eta}_{\max}$	$\boldsymbol{\delta}_{\max}$
Maricopa	5.63	1.49	1
Surat	5.76	0.85	1
Miami	0.05	1.88	0
Hillsborough	0.25	1.93	0

Table 4: A summary of the elements of $\Phi_{\max}$ when Algorithm 1 is applied to each of the four cities independently. These values returned the maximal local correlation between normalised *Aedes* counts and associated $\mathcal{Q}(\Phi_{\max})$. For each city, a total of 10^7 randomly sampled points were considered.

While the goodness-of-fit of these Index $\mathcal{Q}$ time-series varies across the four sites, certain trends are present in all. When $\mathcal{Q}$ is parameterised locally, it coincides relatively well with *aegypti* abundance counts. This is most striking in Miami and Hillsborough, where the correlations are on the order of 0.94, but they are still very high in Maricopa (0.78) and Surat (0.64). Even in the 'worst' case presented here (Surat), Index $\mathcal{Q}$ is able to consistently match the timing (if not the relative magnitude) of the abundance peaks/troughs across a 20-year time frame.

The values in Table 4 further elucidate the reasoning behind the overarching goal of searching for meaningful ρ and η values. The elements of $\Phi_{\max}$ which produce the best fits in Maricopa and Surat involve $\rho_{\max}$ values above 5.5. That is, in order for Index $\mathcal{Q}$ to most closely match the observed *aegypti* abundance counts in these two cities,

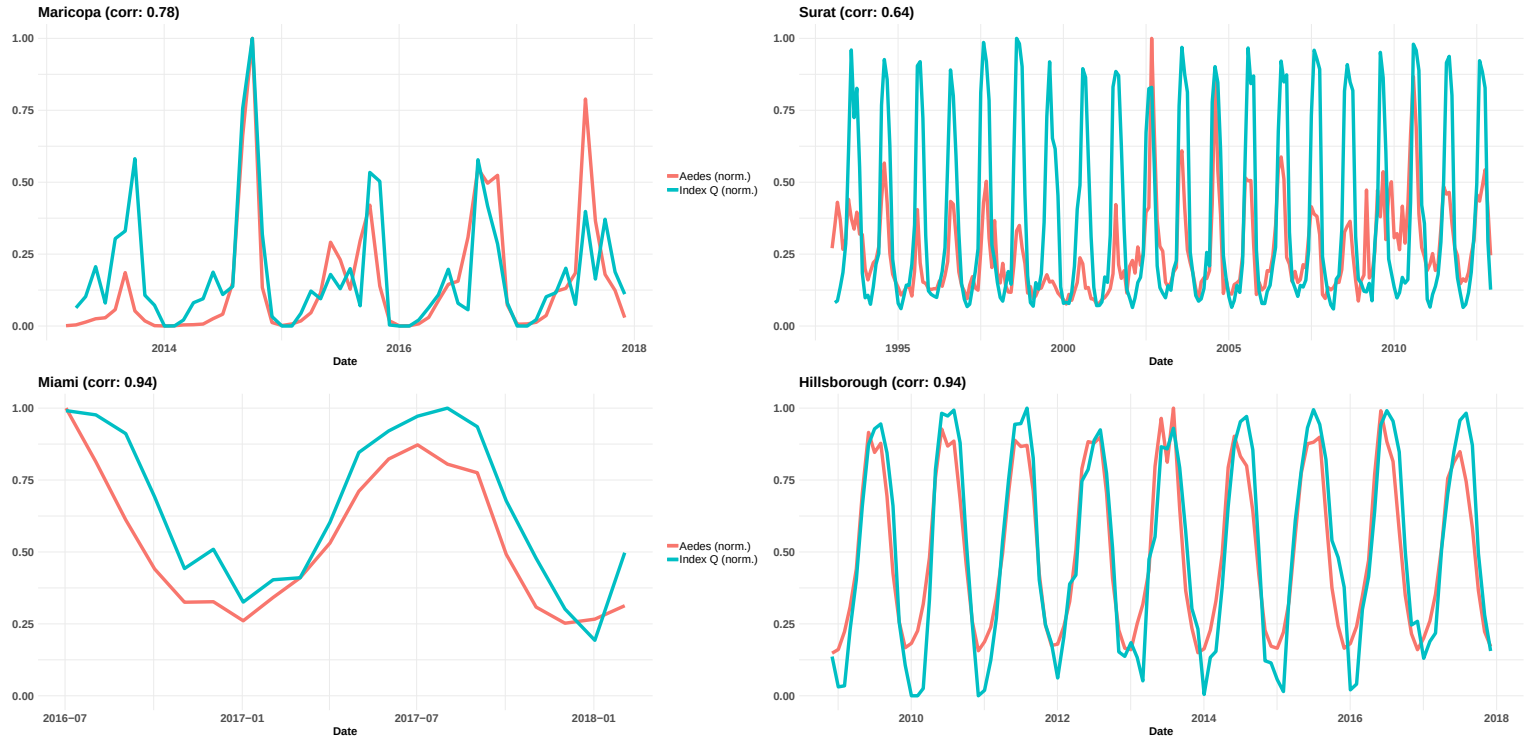


Figure 4: A multiplot showing the best local fit between Index $\mathcal{Q}$ (blue) and *Aedes* abundance data (red) in (clockwise from top-left) Maricopa, Surat, Hillsborough, and Miami. Both $\mathcal{Q}$ and the local *Aedes* counts were normalised for visualisation purposes. Note the differing time periods in each panel (associated with the availability of *Aedes* abundance data in that locale).

the effects of humidity and precipitation on the egg-hatching success probability and the adult mosquito mortality rate, respectively, must be more than 5.5 times higher than the original lab-based formulations would suggest. Meanwhile, in the cities with the highest correlation values (Miami and Hillsborough), the maximising ρ values were significantly less than 1, suggesting a severe *discounting* of these effects in these two cities. Similarly for η: the $\eta_{\max}$ value in Surat amounts to saying that the effect of temperature on the aquatic-to-adult transition rate and the aquatic and adult mortality rates is only 85% of the original dependencies given in (H. M. Yang et al. 2009); in the other three cities, this effect is instead amplified by $49-93\%$. The maximising δ values are evenly split between Maricopa/Surat ($\delta = 1$) and Miami/Hillsborough ($\delta = 0$).

The discrepancy between the values of $\Phi_{\max}$ in each of the four cities is a function of their unique climates. Maricopa County, with its hot desert climate, exhibits the least variability across all three climatic variables. Its consistently high temperatures and extremely low humidity and precipitation levels mean that the emphasis placed on these last two climatic indicators is paramount – any deviations in humidity and/or precipitation levels are much more relevant for mosquito reproductive potential here

than temperature (though the effect of temperature is still nearly 50% more pronounced here than in the lab). Similarly in Surat: the relatively consistent temperatures and infrequent precipitation are reflected in $\Phi_{\max}$ by setting $\eta < 1$ and $\rho > 1$. The tropical monsoon climate of Miami and the fully-humid subtropical climate of Hillsborough in contrast experience relatively consistently high levels of humidity and precipitation throughout the year – thus the emphasis is inverted, with $\eta > 1$ and $\rho < 1$.

We note that each of these cities' respective $\Phi_{\max}$ values represent one of many possible parameter configurations which yield similarly high correlation values. To rigorously examine these maximising sets of parameter configurations, we move towards finding an approximate global solution.

3.4 Approximate Global $\mathcal{Q}$

We generalise Index $\mathcal{Q}$ to account for those situations where *aegypti* abundance data is unavailable by collating the climate and mosquito abundance data of the four calibration cities, and performing Algorithm 1 on them collectively. Since the time-frames of the four cities' *Aedes* abundance data did not overlap, we implemented line 18 of Algorithm 1 on a per-city basis. Consider our parameter space $\Omega \subset \mathbb{R}^3$:

$$\Omega \;:=\; \{\underbrace{(\rho,\, \eta,\, \delta)}_{:=\Phi} \mid \rho \in [0, 10] \,\cap\, \eta \in (0, 5] \,\cap\, \delta \in \{0, 1\}\} \quad . \tag{13}$$

The biologically plausible subspace of Ω (Ω_{bp}) is unique to each city's climate, and is determined by ensuring that Equations (9) and (10) are satisfied. Once Ω_{bp} is determined, each city's final contribution towards the 'approximate global $\mathcal{Q}$' posterior distributions of the elements of Φ is given by

$$\Phi_{\max} \;:=\; \{\Phi \,\in\, \Omega_{\text{bp}} \mid \text{corr}\,(\mathcal{D},\, \mathcal{Q}(\Phi)) \,\geq\, \mathcal{M}_i\} \quad , \tag{14}$$

where $\mathcal{M}_i$ is the median correlation $\forall \Phi \in \Omega_{\text{bp}}$ in city i. This process highlights the necessity of ensuring equal representation between cities – sampling bias could otherwise skew the approximate posterior distributions in favor of one locale with particularly high correlation values. Proceeding in this fashion yields the approximate posterior distributions seen in Figures 5-7.

Figure 5 indicates that the approximate posterior distributions for ρ and η are positively-skewed in each of the four locations. This skewness is not spatially uniform – Maricopa

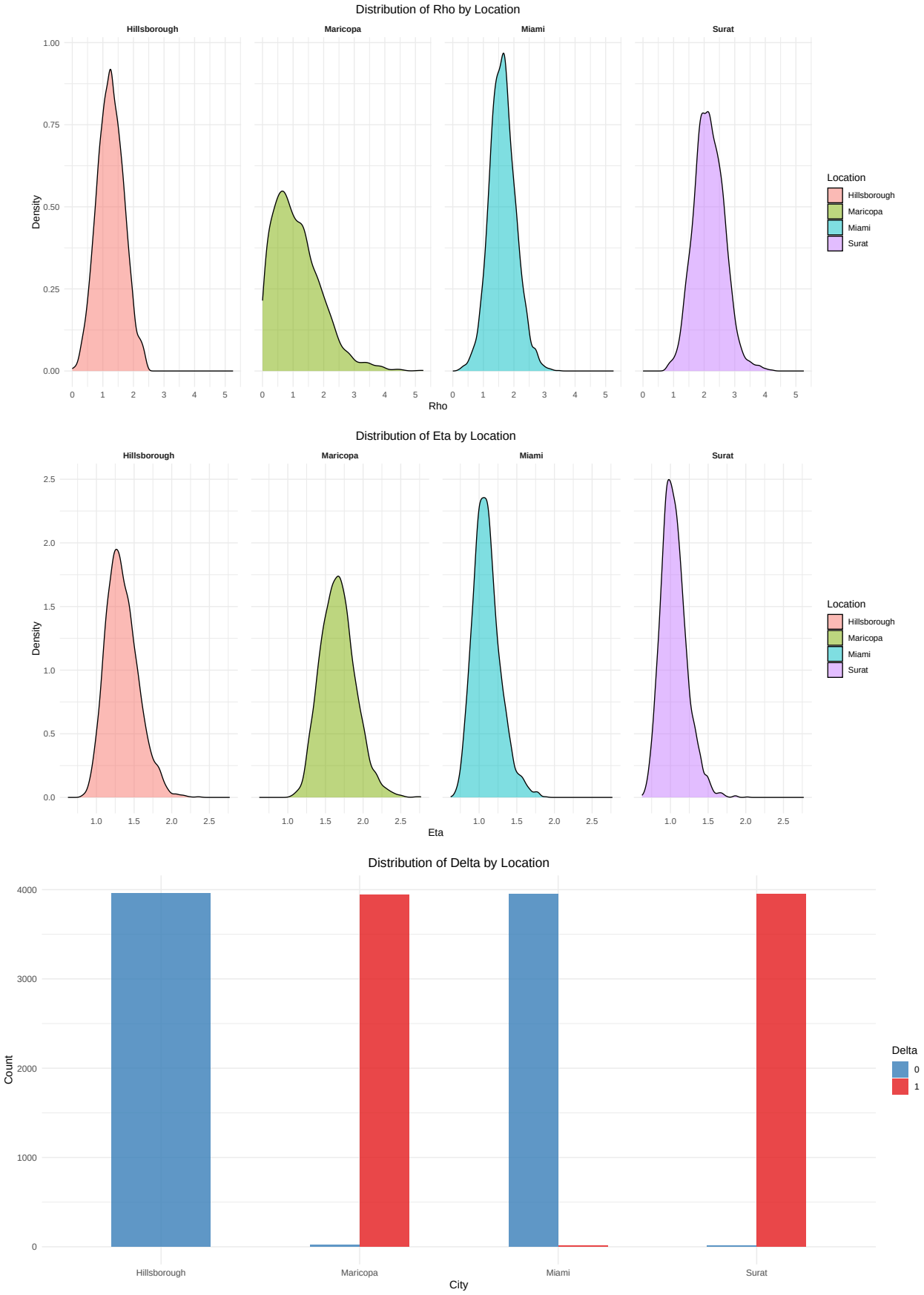


Figure 5: Approximate posterior distributions for ρ (top), η (middle), and δ (bottom), stratified by location.

and Surat appear more positively skewed than the other two cities in ρ, while Surat appears to be the least positively skewed in η. These stratified approximate distributions suggest that those parameter values Φ which maximise the correlation between Index $\mathcal{Q}$ and observed mosquito abundance in Maricopa and Surat have a tendency to include higher ρ values more often. In other words, the relationship between mosquito reproductive potential and actual mosquito abundance is more sensitive to precipitation and humidity in Maricopa and Surat than in Miami and Hillsborough.

Figure 6 combines the approximate posteriors of ρ and η (Figure 5) into a single 2-dimensional point-and-whisker plot (still stratified by location). Across all four cities, the mean (ρ, η) point was $(1.42, 1.40)$. In other words, to maximise the correlation between Index $\mathcal{Q}$ and our mosquito abundance data $\mathcal{D}$, we must consider the effects of temperature on aquatic development and aquatic and adult mortality and the effects of precipitation and humidity on egg hatching success and adult mortality, respectively, to each be roughly 40% more pronounced than in laboratory settings.

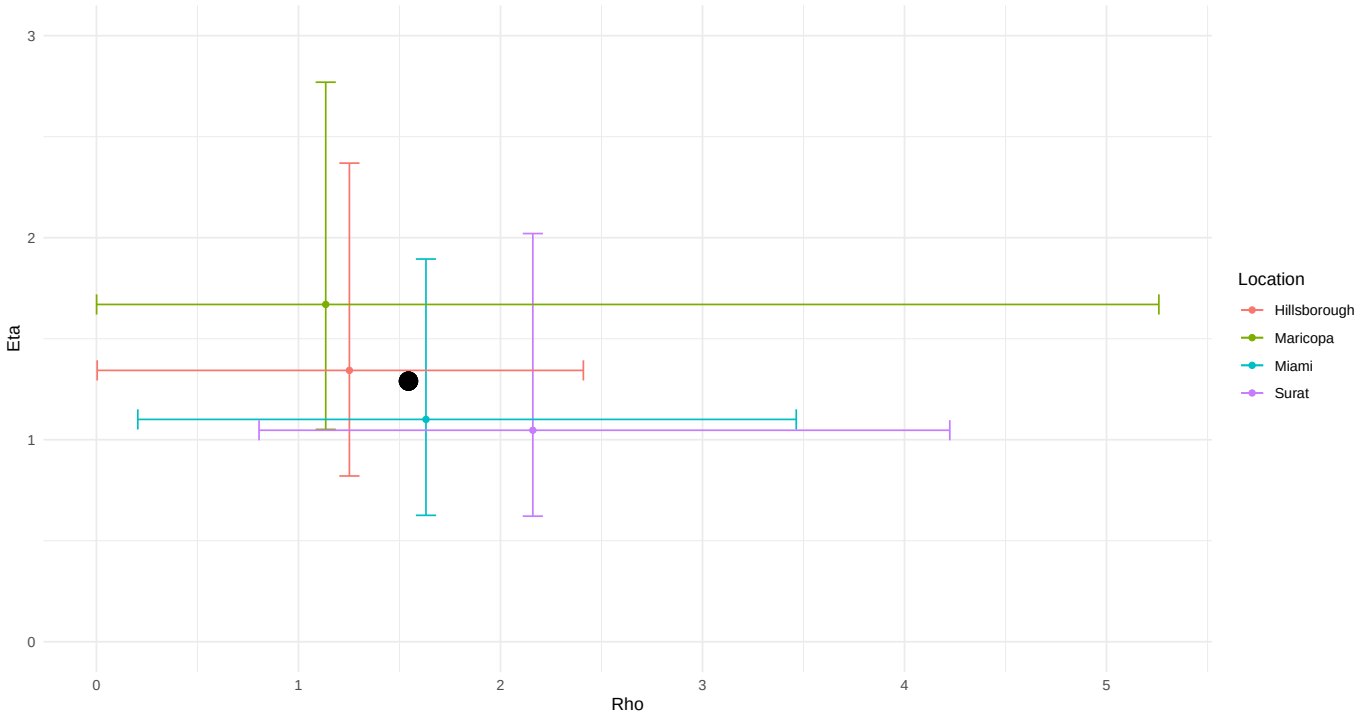


Figure 6: A combined point-and-whiskers plot showcasing the distribution of ρ and η for each city. For each city, the plotted point is the average (ρ, η) value while the whiskers represent the ranges for each parameter in that location. The black dot represents the mean (ρ, η) point across all sites.

Figure 6 suggests that the parameter values Φ which maximise the correlation between Index $\mathcal{Q}$ and *Aedes* abundance data in Maricopa do not contain any values of $\eta < 1$.

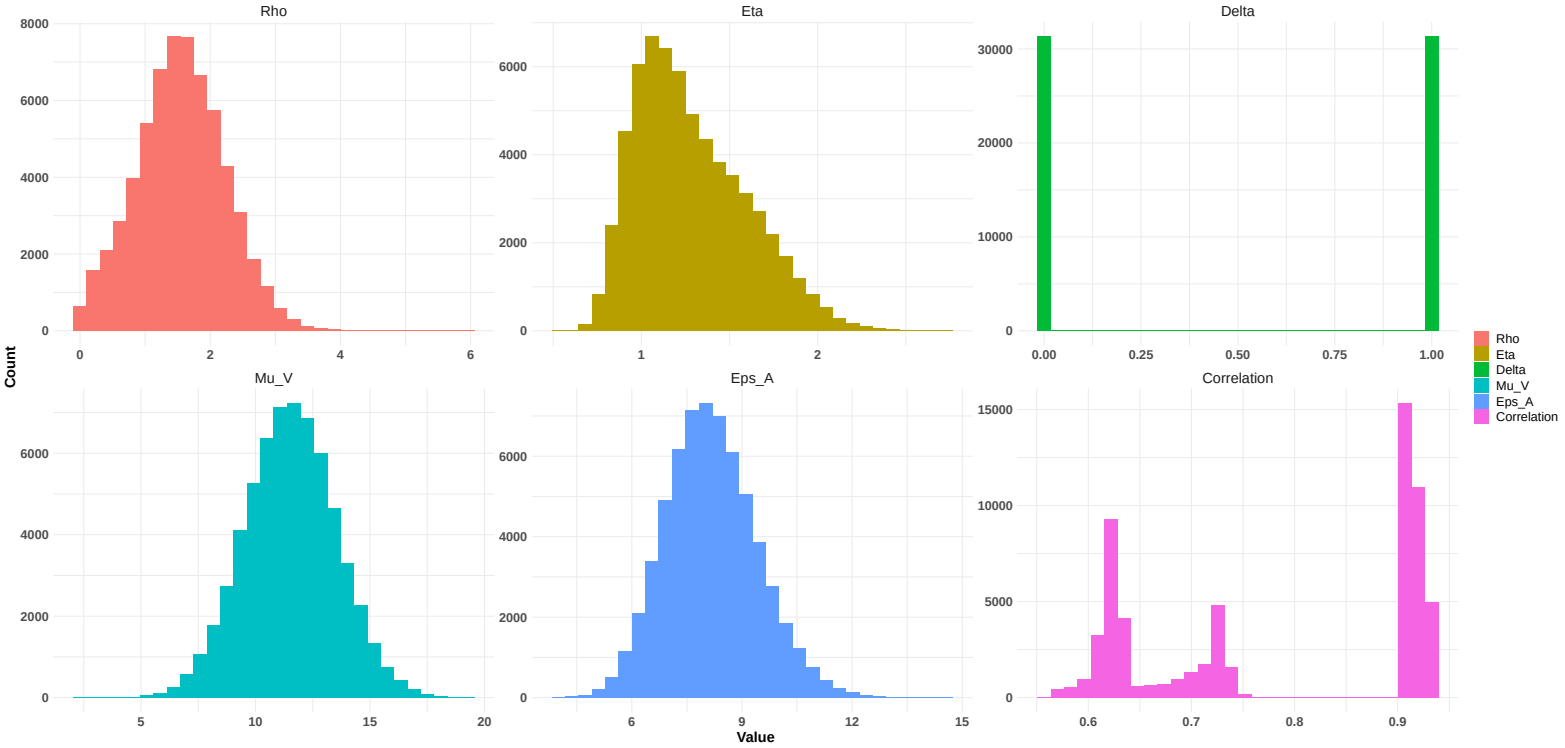


Figure 7: Approximate posterior distributions for (from left to right; top row): ρ, η, and δ, and (from left to right; bottom row) μ_V^{-1}, ϵ_A^{-1}, and correlation coefficients between corresponding $\mathcal{Q}$'s and mosquito abundance data in each of the four cities. This is the result of Algorithm 1.

That is, $\forall \Phi \in \Phi_{\max}$, the effect of temperature in Maricopa on ϵ_A, μ_A, and μ_V is strictly greater than that observed in laboratory settings. That the range of ρ in both Maricopa and Hillsborough contain values very close to 0 indicates that, in some instances, it is possible for Index $\mathcal{Q}$ to very closely approximate $\mathcal{D}$ while completely ignoring the effects of precipitation and humidity.

Aggregating these approximate posterior distributions for ρ, η, and δ yields the top row of Figure 7. In accordance with their location-specific distributions, the full approximate posteriors of both ρ and η remain positively-skewed. This skewness is noticeably more pronounced in η than in ρ. These approximate posteriors are accompanied by those of $\mu_V^{-1}, \epsilon_A^{-1}$, and the correlation coefficient values on the bottom row. The correlation coefficients are relatively high, ranging from 0.52 to 0.94 .

Proposed Analytical Distributions

In order to move towards an approximate global solution for locations without available *Aedes* abundance data, we used the `fitdistrplus` package in `R` to compare multiple candidate analytical distributions for the approximate posterior distributions of ρ, η, and δ (Figure 7). Figure 8 displays a *Cullen and Frey* graph for the discrete approx-

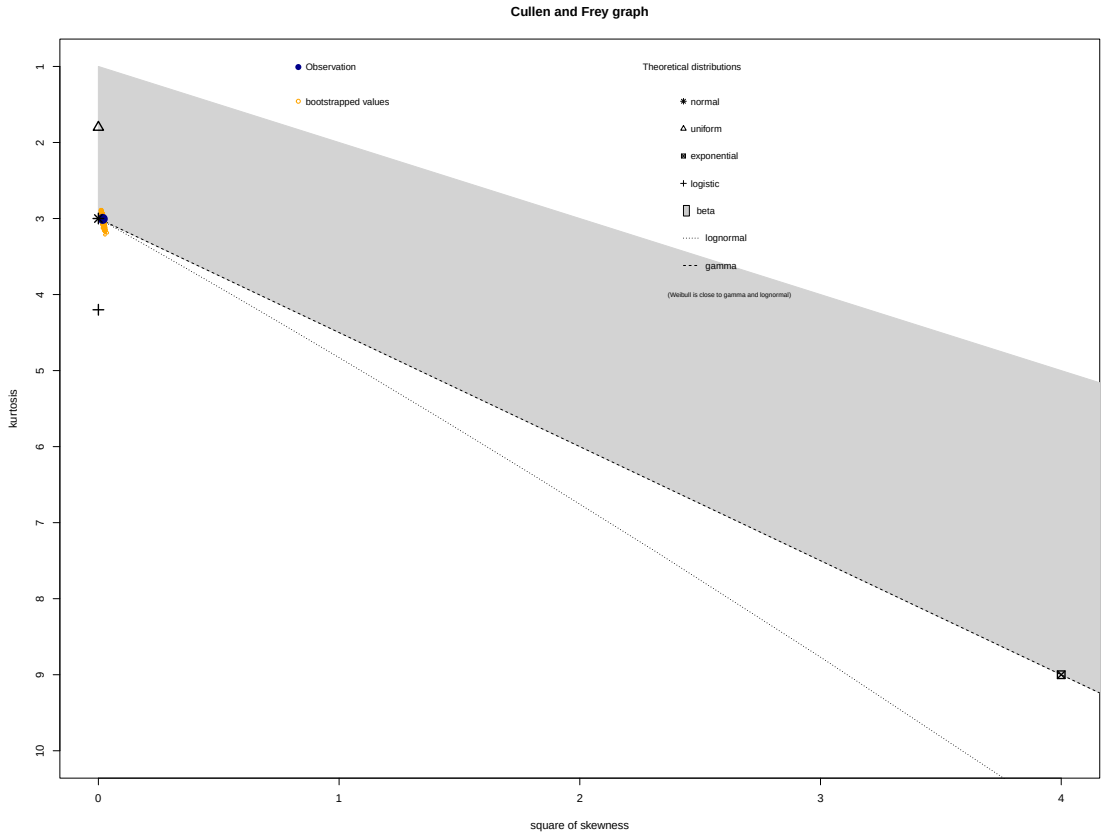


Figure 8: A Cullen and Frey graph showcasing the estimated kurtosis and squared skewness of our approximate posterior distribution for ρ. The blue and orange dots represent the original calculation and $10,000$ subsequent iterations, respectively.

imate posterior distribution of ρ seen in the top-left panel of Figure 7. This plots the *kurtosis* of the approximate distribution against its *square of skewness* in order to decide which candidate distributions are appropriate to consider. This assessment was bootstrapped $10,000$ times. Based on this output, we considered normal, log-normal, gamma, and Weibull distributions as possible analytic candidates for ρ.

In order to decide which of these four candidate analytical distributions most closely fit our approximate posterior distributions (Figure 7), we tested each and calculated the AIC. Then, we chose that fit which minimised the AIC value. (This process was repeated for η). The AIC values for each potential fit of ρ and η are seen in Table 5.

Candidate Distribution	$\mathbf{AIC}_\rho$	$\mathbf{AIC}_\eta$
Normal	128610	32358
Lognormal	161569	26007
Gamma	140218	27190
Weibull	130665	35896

Table 5: A table displaying the corresponding AIC values for our approximate posterior distributions of ρ and η when fitted to each of four considered candidate distributions. The minimum values in each column are highlighted (indicating best fit).

Thus, we propose analytical distributions for ρ and η such that

$$\rho \sim \mathcal{N}^{+}(1.55,\ 0.68) \tag{15}$$

$$\eta \sim \text{Lognormal}(0.22,\ 0.24) \tag{16}$$

where these values were determined via *maximum likelihood estimation*. Upon inspection we let

$$\delta \sim \text{Binom}(n = 1,\ p = 0.5) \tag{17}$$

Equations (15)–(17), taken with the known climate-dependent entomological parameters in Equations (3)–(8), fully specify our approximate global solution for Index $\mathcal{Q}$.

3.5 Local vs. Global $\mathcal{Q}$

Algorithm 2 was implemented to check how well our approximate global solution of Index $\mathcal{Q}$ correlates with the observed mosquito abundance data in each of the four cities. We sampled $M = 1,000$ values of Φ from the proposed analytical distributions in Equations (15)–(17). Figure 9 displays the correlation between our $1000-$sample mean $\mathcal{Q}$ and *aegypti* abundance data $\mathcal{D}$ in each city.

The correlations between $\mathcal{Q}$ and $\mathcal{D}$ in each of the four cities ranges from 0.59 (Surat) to 0.85 (Miami). These relatively high values indicate that our analytical distributions are appropriate in considering a global solution for Index $\mathcal{Q}$. The correlations in each of these cities is strictly lower than their local solution counterparts from Figure 4. The loss of local accuracy is the price of generalisation.

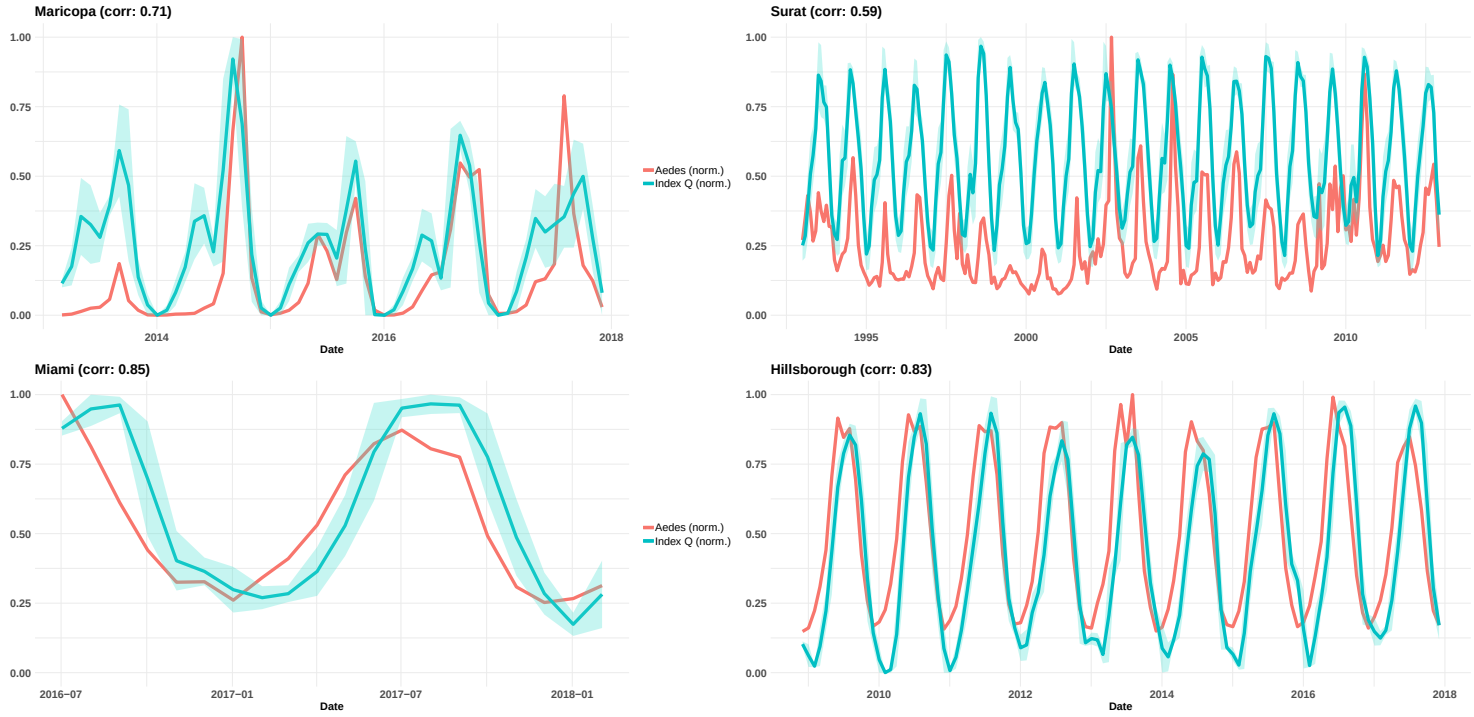


Figure 9: A multiplot showcasing the correlation between the mean of 1000 random samples of Index $\mathcal{Q}$ (blue line; ribbon is ± 1 std. dev.) and *Aedes* abundance data (red line) in each of the four cities we investigated. Both Index $\mathcal{Q}$ and *Aedes* abundance time-series are normalised to aid visualisation.

Figure 10 represents a return to the six showcase cities via 1000−sample mean $\mathcal{Q}$'s based on Equations (15) – (17). Despite Belém experiencing high variability in $\mathcal{Q}$ based on our proposed analytical distributions, it is still not enough to ensure $\mathcal{Q} < 1$ at any point. Though Darwin's projections are much less variable, the same holds true there. The remaining four cities continue to have annual periods in which $\mathcal{Q} = 0$. Their relatively narrow ribbon bands, especially in comparison to those of Belém and Darwin, suggest that the long-term climate of these cities is significantly more stable and annually consistent.

The linear regressions plotted in Figure 10 (purple) convey the effect of climate change on *aegypti* reproductive potential. In all cities save Darwin, $\mathcal{Q}$ is increasing between $1979 - 2023$ (all are statistically significant save Tokyo; $p > 0.05$). The consistency of these long-term environmental trends validates the concern that areas traditionally considered safe from local, non-imported MBV transmission could potentially increase their MBV burden as global climate changes.



Figure 10: A multiplot of 1000-sample means of Index $\mathcal{Q} \pm 1$ std. dev. in each of our six showcase cities (for which *Aedes* abundance data was not available). Note the differing $y-$axes in each of the panels. These projections are in line with the available climate data, ranging from Jan. 1979 to March 2023. The black dashed lines represent the critical threshold $\mathcal{Q} = 1$. The purple lines display a linear regression of $\mathcal{Q}$ in each city during this time period.

3.6 Spatial $\mathcal{Q}$

The final results examine Index $\mathcal{Q}$ beyond the scope of individual cities. Drawing 1000 random samples from the approximate global solution in Equations (15) – (17) yielded a mean $\mathcal{Q}$ for each latitude/longitude pixel in Figure 11.

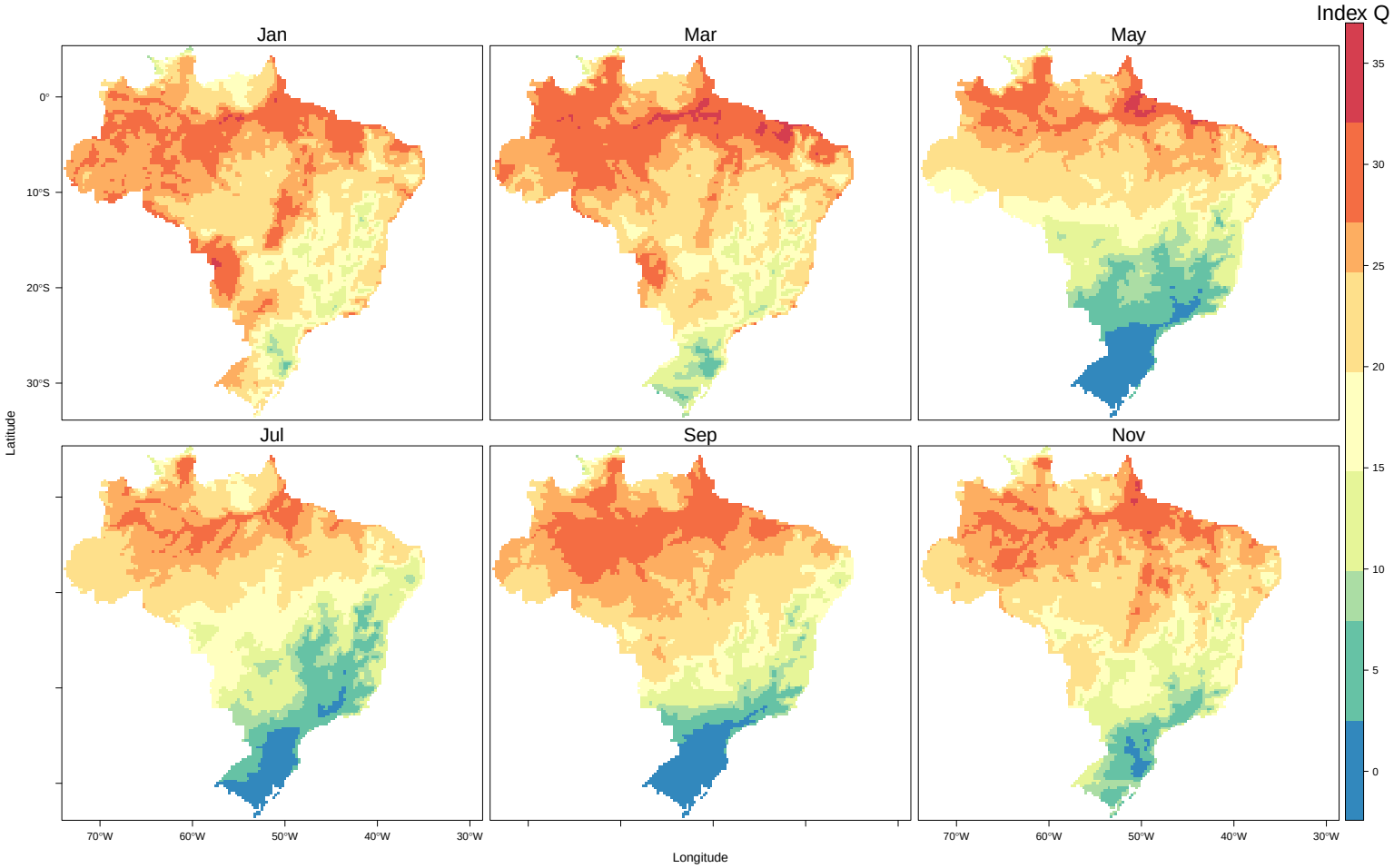


Figure 11: A bi-monthly evolution of Index $\mathcal{Q}$ across Brazil in 2022 (the last full year for which climate data was available). Note that the majority of Brazil's landmass lies in the Southern hemisphere, and thus the warmer summer period occurs between November and February.

Index $\mathcal{Q}$, based solely on widely available climate data and our proposed analytical distributions for ρ, η, and δ is able to uncover some striking dynamics. The consistently low values in the south and south-east coincide with those locales having the lowest national MBV burden (Brady et al. 2019; Codeco et al. 2022). Meanwhile, the consistently high values in the north, north-east and south-west of the country instead correlate with those areas with the highest MBV burden (Andrioli, Busato and Lutinski 2020; Silva Neto et al. 2022). Note that the seasonality of $\mathcal{Q}$ follows that of Brazil, whose warmer summer period is between November and February.

Following the same sampling process, Figure 12 displays the mean $\mathcal{Q}$ globally between January 1979 - June 2023. The red band in the tropics across northern South America, sub-Saharan Africa, India, Southeast Asia, and northern Australia is where reproductive potential is estimated to be maximal. This band is also where the burden of *aegypti*-related MBVs is the highest. Figure 13 conveys the effect of climate change on reproductive potential. A monthly linear regression of $\mathcal{Q}$ between 1979 and 2023 indicates that the global mean $\mathcal{Q}$ value has risen from approximately 4.5 to around 5, an 11% increase in that time. Figure 14 further illuminates the effect of over four decades of climate change on *aegypti* reproductive potential. While the tropics and extreme northern/southern reaches remain unaltered, $\mathcal{Q}$ has worryingly doubled across most of North America, Europe, the Caucasus, and Siberia. The very few places that exhibit a reduction in $\mathcal{Q}$ lie in the Arctic circle and the heart of the Arabian desert.

These spatio-temporal projections can aid not only in identifying which locales are most vulnerable to potential MBV transmission, but also *when* they are most vulnerable. This could make the allocation of resources designed for vector-control and outbreak-mitigation more efficient. Long-term shifts in these spatial distributions (e.g. as a result of climate change) can also help inform future public health measures.

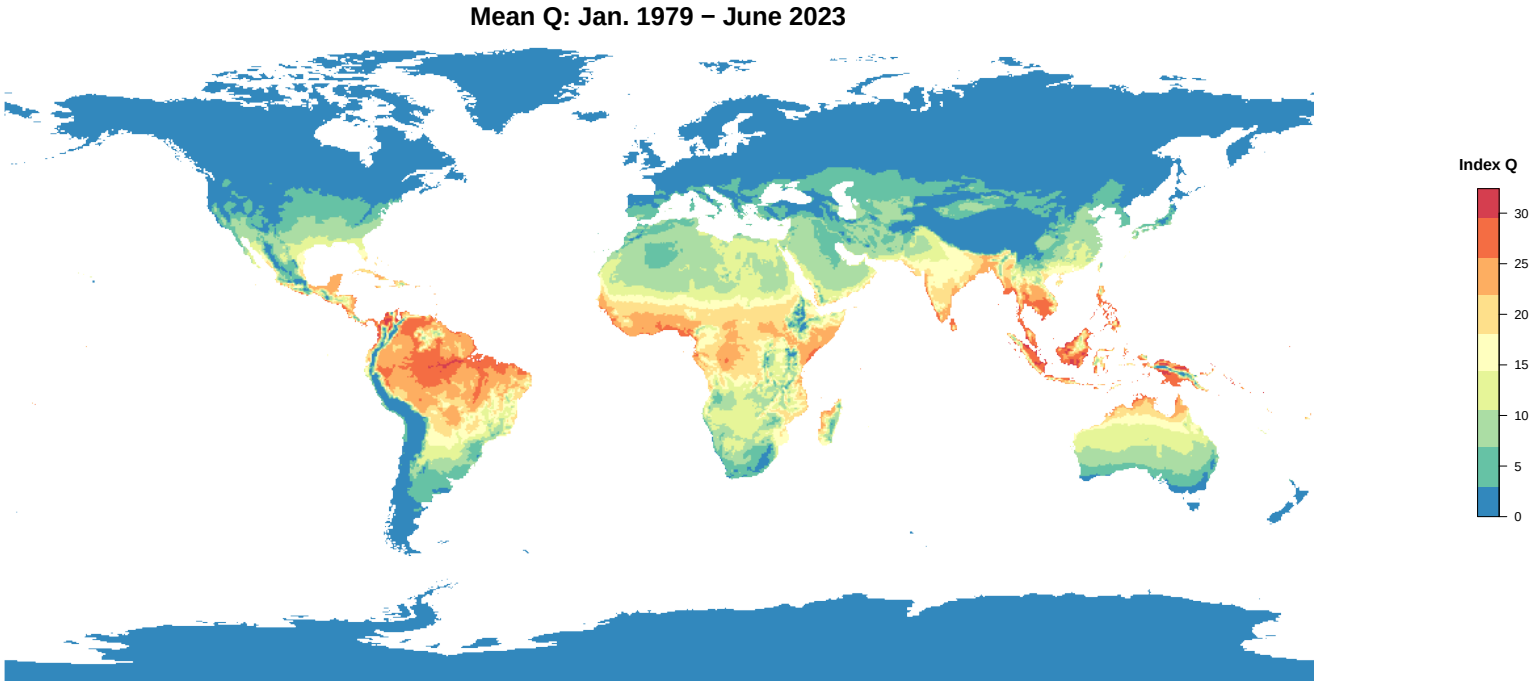


Figure 12: A map of mean global Index $\mathcal{Q}$ between Jan. 1979 – June 2023 at a 0.25° by 0.25° resolution. For each pixel, the mean $\mathcal{Q}$ value displayed here is the average of the 534 monthly $\mathcal{Q}$'s in this time period.

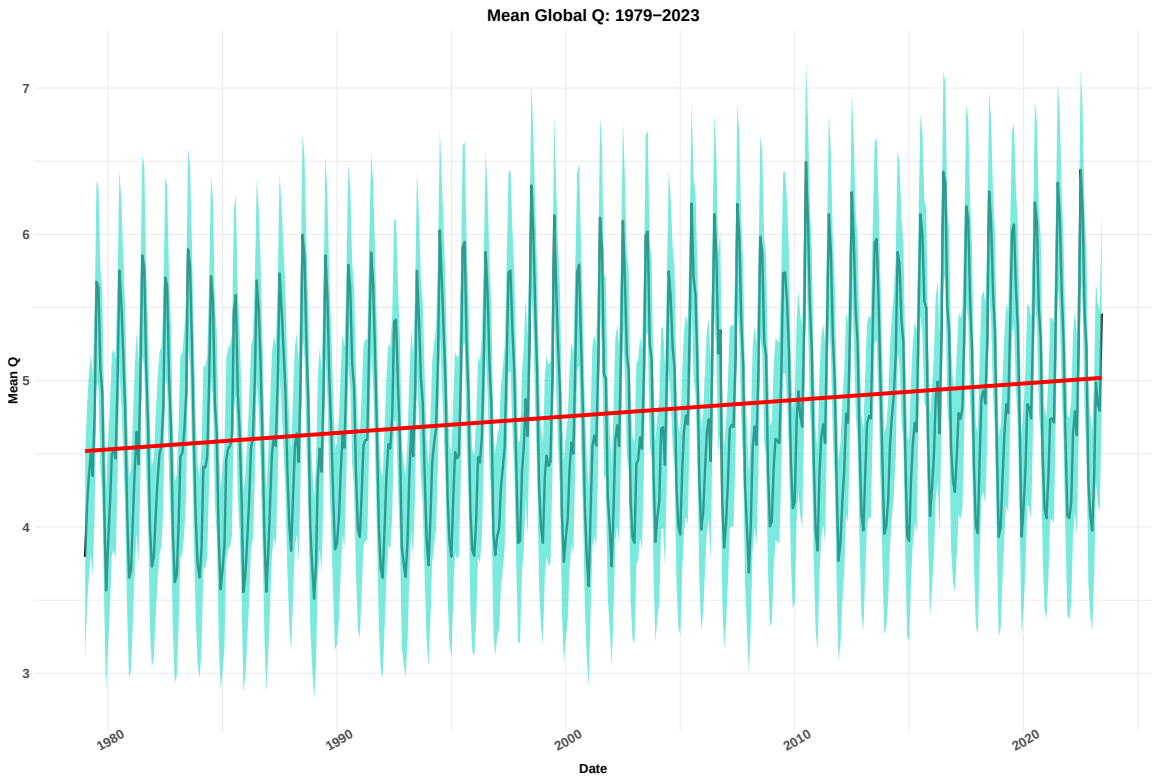


Figure 13: A time-series of the global mean $\mathcal{Q}\pm1$ std. dev. between Jan. 1979 – June 2023 (turquoise). Each time point corresponds to the mean $\mathcal{Q}$ across the globe for the corresponding month. A linear regression of the global mean $\mathcal{Q}$ value over time is given in red.

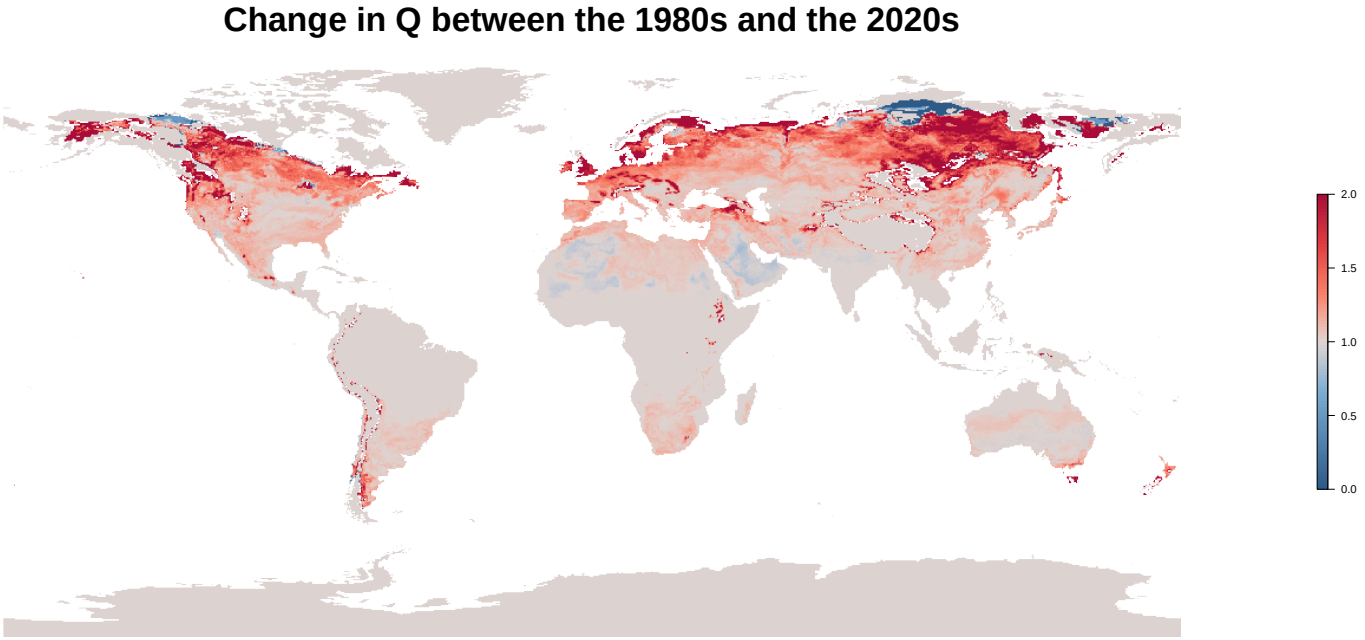


Figure 14: A map displaying the global change in $\mathcal{Q}$ over time. For each pixel, the displayed value is a ratio, where the numerator is the proportion of time where $\mathcal{Q} > 1$ between $1979 - 1989$ and the denominator is the proportion of time where $\mathcal{Q} > 1$ between $2013 - 2023$. A ratio above 1 (light/dark orange and red) indicates growing climatic accommodation for *aegypti* reproduction; a ratio below 1 (light/dark green and blue) indicates decreasing hospitability between $1979 - 2023$.

4 Conclusions

We used climate and *aegypti* abundance data from four cities to estimate two unknown scaling parameters of Index $\mathcal{Q}$. Using only temperature, humidity, and precipitation, we were able to use Index $\mathcal{Q}$ to predict with relatively high accuracy the mosquito abundance dynamics of four spatially and climatically heterogeneous locations across various time-frames. We were then able to generalise $\mathcal{Q}$ and project the mosquito reproductive potential and abundance dynamics across six diverse urban areas of interest across the globe. Climate change trends between $1979 - 2023$ were shown to be increasing *aegypti* reproductive potential across many regions worldwide, including those areas not presently endemic to MBVs of public health importance. The work presented here mostly functions as a 'proof of concept' showcasing that this data-driven mathematical and computational approach considering the effects of climate on mosquito population dynamics could be used to inform and optimise MBV transmission control in both endemic and non-endemic areas of the world.

5 Limitations & Future Directions

Among the limitations in our approach is the relatively small number of data sets we were able to utilise, as well as their diversity – we only explored the singular *Aedes aegypti* species, and only in four urban areas (three of which were in the USA). While *aegypti* is especially epidemiologically relevant and geographically diverse, the ability to develop similar measures of reproductive potential for *Culex pipiens, Aedes albopictus,* and *Anopheles albimanus* (to name a few) would be indispensable in combatting the spread of many other MBVs and neglected tropical diseases.

The discovery of an analytical likelihood for mosquito reproductive potential would allow for the pursuit of an analogous MCMC-based investigation, allowing for a more robust investigation into the role that ρ and η play in the evolution of Index $\mathcal{Q}$. The results could be used as validation for our own ABC-based investigation.

The integration of $\mathcal{Q}$ with MBV epidemiological data such as relevant case counts and reported infections could aid in the investigation of the relationships between *aegypti* reproductive potential, actual abundance, and associated MBV disease burdens.

A more rigorous consideration of the effects of climate change (locally and globally) on mosquito reproductive potential would also be a very worthy candidate for exploration.

Author Contributions

Conceptualization: JL, 1048500;
Data Curation: JL, 1048500;
Formal Analysis: JL, 1048500;
Investigation: JL, 1048500;
Methodology: JL;
Writing - Original Draft Preparation: 1048500;
Supervision: JL;
Writing - Review & Editing: JL.

Competing Interests

The authors declare that no competing interests exist.

Financial Disclosure

The authors received no specific funding for this work.

Supplementary Information Captions

S1 Methods. Mathematical justifications of the critical threshold $\mathcal{Q} = 1$.

S2 Results. Statistical details pertaining to Section 3.4.

Availability Statements

Data

The external data used in this project comprised the *Aedes aegypti* abundance data for Miami-Dade and Hillsborough Counties; this is publicly available and can be found in (B. Yang et al. 2021). The abundance data for Maricopa County and Surat are not publicly available and were made available to us by JL. Details about the Maricopa and Surat data sets can be obtained by contacting them directly. (Email omitted to maintain anonymity, but can be provided upon request). The global climate data used to generate all world maps is too large (> 1GB) to submit, but is publicly available from the Copernicus data set provided by the EU (Hersbach et. al 2018).

Code

The code used to analyse all data and generate all plots in this paper will be submitted with this dissertation, but pertaining to the non-public abundance data used, cannot be made fully public in an online code repository (e.g. GitHub). The version of `R` used throughout this paper is 4.3.1 (R Core Team 2023); the version of `RStudio` used is 2023.6.1.524 (Posit team 2023).

References

Andrioli, Denise Catarina, Maria Assunta Busato and Junir Antonio Lutinski (Feb. 2020). 'Spatial and temporal distribution of dengue in Brazil, 1990 - 2017'. In: *PLOS ONE*. URL: https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0228346#pone.0228346.ref005.

Brady, Oliver J. et al. (Mar. 2019). 'The association between zika virus infection and microcephaly in Brazil 2015–2017: An observational analysis of over 4 million births'. In: *PLOS Medicine*. URL: https://journals.plos.org/plosmedicine/article?id=10.1371%2Fjournal.pmed.1002755.

CDC (Feb. 2014). In: *Centers for Disease Control and Prevention*. URL: https://stacks.cdc.gov/view/cdc/24391.

Codeco, Claudia T. et al. (May 2022). 'Fast expansion of dengue in Brazil'. In: *The Lancet Regional Health - Americas*. URL: https://www.thelancet.com/journals/lanam/article/PIIS2667-193X(22)00091-6/fulltext.

Delignette-Muller, Marie Laure and Christophe Dutang (2015). 'fitdistrplus: An R Package for Fitting Distributions'. In: *Journal of Statistical Software* 64.4, pp. 1–34. DOI: 10.18637/jss.v064.i04.

ECDC (Nov. 2018). 'The importance of vector abundance and seasonality'. In: *ecdc.europa.eu*. URL: https://www.ecdc.europa.eu/en/publications-data/importance-vector-abundance-and-seasonality.

— (Jan. 2023). In: *ecdc.europa.eu*. URL: https://www.ecdc.europa.eu/en/disease-vectors/facts/mosquito-factsheets/aedes-aegypti.

Faridah, Lia et al. (Mar. 2022). 'Temporal Correlation Between Urban Microclimate, Vector Mosquito Abundance, and Dengue Cases'. In: *Journal of Medical Entomology* 59.3, pp. 1008–1018. ISSN: 0022-2585. URL: https://doi.org/10.1093/jme/tjac005.

Franklinos, Lydia H V et al. (June 2019). 'The effect of global change on mosquito-borne disease'. In: *thelancet.com*. URL: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(19)30161-6/fulltext.

Groen, Thomas A. et al. (Oct. 2017). 'Ecology of West Nile virus across four European countries: Empirical modelling of the Culex pipiens abundance dynamics as a function of weather'. In: *Parasites & Vectors*. URL: https://parasitesandvectors.biomedcentral.com/articles/10.1186/s13071-017-2484-y.

Hersbach et. al, H. (Nov. 2018). 'Copernicus Climate Data Store'. In: *Essential climate variables for assessment of climate variability from 1979 to present*. URL: https://cds.climate.copernicus.eu/cdsapp#!/dataset/ecv-for-climate-change?tab=overview.

Lourenço, José and Mario Recker (Aug. 2014). 'The 2012 Madeira Dengue Outbreak: Epidemiological determinants and future epidemic potential'. In: *PLoS neglected tropical diseases*. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4140668/.

Lourenço et al., José (Sept. 2017). 'Epidemiological and ecological determinants of zika virus transmission in an urban setting'. In: *eLife*. URL: https://pubmed.ncbi.nlm.nih.gov/28887877/.

Msimang, Veerle et al. (Apr. 2018). 'Increased importation of dengue cases into South Africa: A risk for establishment of local endemicity?' In: *NICD Bulletin*. URL: https://www.nicd.ac.za/wp-content/uploads/2017/03/Increased-importation-of-dengue-cases.pdf.

NTGov (Oct. 2021). 'Ross River virus'. In: *NT.GOV.AU*. URL: https://nt.gov.au/wellbeing/health-conditions-treatments/viral/ross-river-virus.

— (Apr. 2023). 'Dengue'. In: *NT Health*. URL: https://health.nt.gov.au/health-conditions-and-disease-information/dengue.

Obolski, Uri et al. (Apr. 2019). 'MVSE: An R-package that estimates a climate-driven mosquito-borne viral suitability index'. In: *Methods in ecology and evolution*. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7202302/.

Pliego et al, Emilene Pliego (June 2017). 'Seasonality on the life cycle of Aedes aegypti mosquito and its statistical relation with dengue outbreaks'. In: *Applied Mathematical Modelling*. URL: https://www.sciencedirect.com/science/article/pii/S0307904X17303955#sec0014.

Posit team (2023). *RStudio: Integrated Development Environment for R*. Posit Software, PBC. Boston, MA. URL: http://www.posit.co/.

Quam, Mikkel B et al. (Feb. 2016). 'Dissecting Japan's dengue outbreak in 2014'. In: *The American journal of tropical medicine and hygiene*. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4751952/.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Schaffner, Francis and Alexander Mathis (Aug. 2014). 'Dengue and dengue vectors in the who European region: Past, present, and scenarios for the future'. In: *The Lancet*. URL: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(14)70834-5/fulltext.

Silva Neto, Sebastião Rogério da et al. (May 2022). 'Arboviral disease record data - dengue and Chikungunya, Brazil, 2013–2020'. In: *Nature News*. URL: https://www.nature.com/articles/s41597-022-01312-7.

WHO (Dec. 2022a). 'Chikungunya fact sheet'. In: *World Health Organization*. URL: https://www.who.int/news-room/fact-sheets/detail/chikungunya.

— (Feb. 2022b). 'Zika Epidemiology Update - February 2022'. In: *World Health Organization*. URL: https://www.who.int/publications/m/item/zika-epidemiology-update---february-2022.

— (Mar. 2023). 'Dengue and severe dengue'. In: *World Health Organization*. URL: https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue.

WMP (July 2023). In: *worldmosquitoprogram.org*. URL: https://www.worldmosquitoprogram.org/en/learn/mosquito-borne-diseases/dengue.

Yang, Bingyi et al. (Mar. 2021). *Modelling distributions of Aedes aegypti and aedes albopictus using climate, host density and interspecies competition*. URL: https://journals.plos.org/plosntds/article?id=10.1371%2Fjournal.pntd.0009063.

Yang, H. M. et al. (Feb. 2009). 'Assessing the effects of temperature on the population of Aedes aegypti, the vector of dengue: Epidemiology amp; infection'. In: *Cambridge Core*. URL: https://www.cambridge.org/core/journals/epidemiology-and-infection/article/assessing-the-effects-of-temperature-on-the-population-of-aedes-aegypti-the-vector-of-dengue/E2FE126FB84D0DE97A94E68343B4649

S1 Methods

The coefficients a_i, b_j, c_k, d_l, and w_n in Equations (3) – (7) from Section 2.1 (Methods - Index $\mathcal{Q}$) are known. Their values as given in (H. M. Yang et al. 2009; Lourenço et al. 2017) and implemented in all calculations in this study are given below.

$$\begin{aligned}\sum_{i=0}^{4} a_i\, T^i \;=\; & 0.8692 \;-\; 0.159T \;+\; 0.01116T^2 \\ & -\, 0.0003408T^3 \;+\; 0.000003809T^4\end{aligned} \tag{S1.18}$$

$$\sum_{j=0}^{3} b_j\, T^j \;=\; -1.848 \;+\; 0.2794T \;-\; 0.009254T^2 \;+\; 0.00009226T^3 \tag{S1.19}$$

$$\begin{aligned}\sum_{k=0}^{7} c_k\, T^k \;=\; & 0.131 \;-\; 0.05723T \;+\; 0.01164T^2 \;-\; 0.001341T^3 \\ & +\, 0.00008723T^4 \;-\; 0.000003017T^5 \\ & +\, \left(5.153 \times 10^{-8}\right)\, T^6 \;-\; \left(3.42 \times 10^{-10}\right)\, T^7\end{aligned} \tag{S1.20}$$

$$\begin{aligned}\sum_{l=0}^{4} d_l\, T^l \;=\; & 2.13 \;-\; 0.3797T \;+\; 0.02457T^2 \\ & -\, 0.0006778T^3 \;+\; 0.000006794T^4\end{aligned} \tag{S1.21}$$

$$\sum_{n=0}^{4} w_n\, T^n = \; -5.4 \;+\; 1.8T \;-\; 0.2124T^2 \;+\; 0.01015T^3 \;-\; 0.0001515T^4 \tag{S1.22}$$

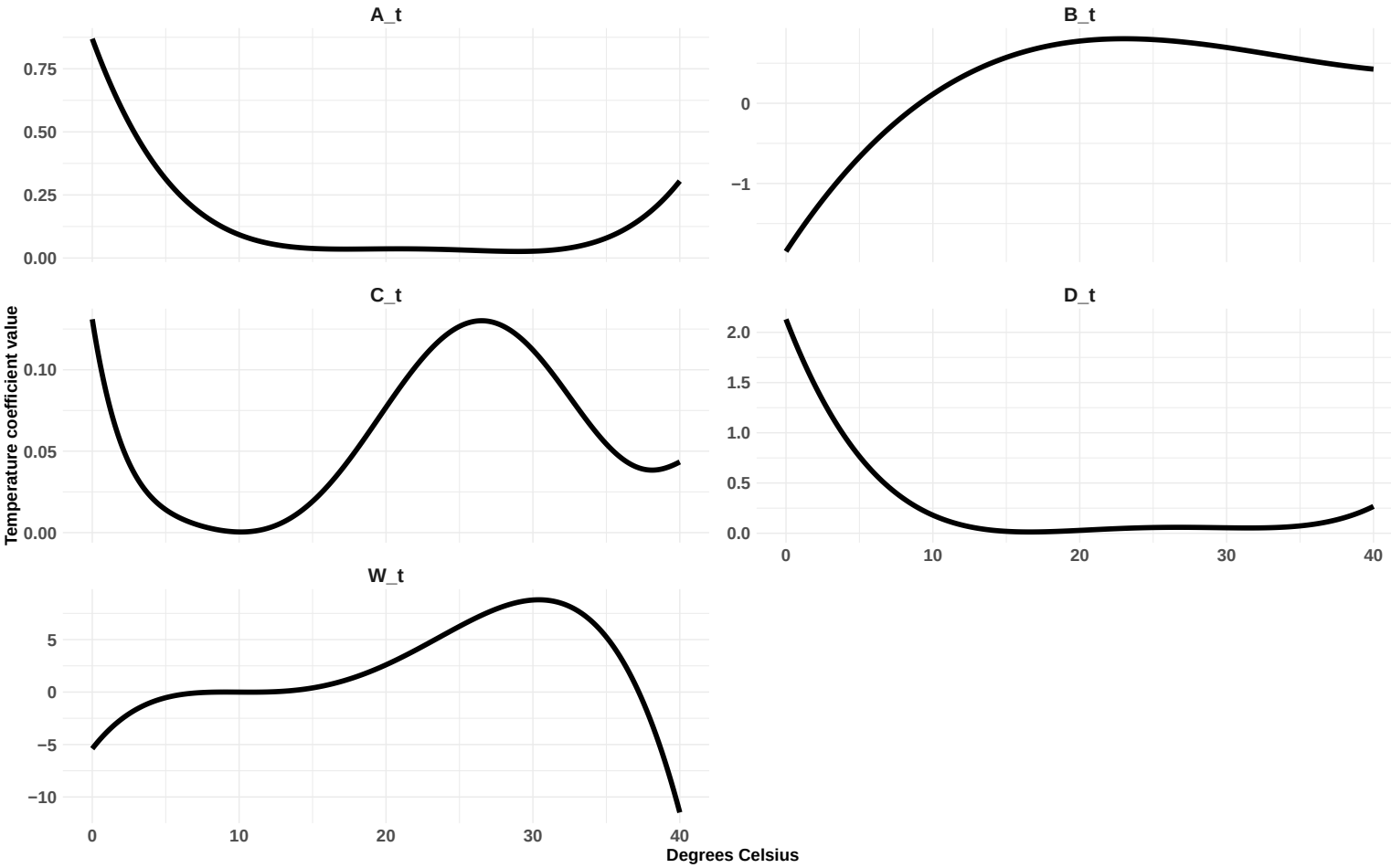


Figure S1.15: Five plots showcasing the relationship between temperature in °C and each of the temperature coefficient values given by Equations (S1.18) – (S1.22). The temperature range displayed here is between 0° and 40°C. Note the differing y-axes in each subplot.

The following three proofs serve to justify why the threshold $\mathcal{Q} = 1$ is a critical value, below which the mosquito-free equilibrium is globally asymptotically stable, and above which it is unstable. A caveat on potential identifiability issues is also presented.

Theorem S1.1. *The ODE system (1) has two equilibria.*

Proof. The trivial equilibrium $A^* = V^* = 0$ is accompanied by the non-trivial equilibrium obtained by solving the system

$$\begin{cases} c\, f\, \theta \left(1 - \dfrac{A}{K}\right) V & = (\epsilon_A \, + \, \mu_A)\, A \\ \mu_V V & = \epsilon_A A \end{cases}$$

which yields the non-trivial solution

$$\begin{cases} A^* & = K \left(1 \, - \, \dfrac{1}{\mathcal{Q}}\right) \\ V^* & = \dfrac{\epsilon_A}{\mu_V} K \left(1 \, - \, \dfrac{1}{\mathcal{Q}}\right) \end{cases}$$

□

Theorem S1.2. *The mosquito-free equilibrium is globally stable when* $\mathcal{Q} \leq 1$.

Proof. Let $\mathcal{L}$ be a Lyapunov function such that $\mathcal{L} : \mathbb{R}^2 \to \mathbb{R}$,

$$\mathcal{L} \;=\; \frac{\epsilon_A}{\epsilon_A \, + \, \mu_A} A \, + \, V$$

Then,

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} \;=\; -\mu_V V \left[1 \, - \, \mathcal{Q} \left(1 \, - \, \frac{A}{K}\right)\right] \quad .$$

When $\mathcal{Q} \leq 1$,

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} \;=\; - \underbrace{\mu_V}_{>0} \underbrace{V}_{\geq 0} \underbrace{\left[1 \, - \, \underbrace{\mathcal{Q}}_{\leq 1} \underbrace{\left(1 \, - \, \frac{A}{K}\right)}_{\in [0,1]}\right]}_{\in [0,1]} \leq 0 \quad .$$

Since the maximal invariant set contained in $\dfrac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} = 0$ is the trivial equilibrium, global asymptotic stability of the trivial (mosquito-free) equilibrium when $\mathcal{Q} \leq 1$ is ensured via LaSalle's invariance principle. □

A leading-order analysis of $\mathcal{Q}$ suggests an identifiability problem with respect to our two parameters of interest. At leading order,

$$\mathcal{Q} \;\sim\; \frac{1}{\eta}\, T^3\, \alpha^{\rho} \quad ; \quad \alpha \;:=\; \left(\frac{1 + \dfrac{\Delta_R}{\sqrt{1 \,+\, \Delta_R^2}}}{1 \,+\, \bar{H} \,-\, \dfrac{\Delta_H}{\sqrt{1 \,+\, \Delta_H^2}}} \right) \;>\; 0 \quad . \tag{S1.23}$$

Theorem S1.3. *$\mathcal{Q}$ is conditionally identifiable at leading order.*

Proof. Let $\Phi_1 = \{\rho_1, \eta_1\}$ and $\Phi_2 = \{\rho_2, \eta_2\}$. Then

$$\begin{aligned} \mathcal{Q}_{\Phi_1} \;=\; \mathcal{Q}_{\Phi_2} \;&\Longleftrightarrow\; \frac{1}{\eta_1}\, T^3\, \alpha^{\rho_1} \;=\; \frac{1}{\eta_2}\, T^3\, \alpha^{\rho_2} \\ &\Longleftrightarrow\; \frac{\eta_2}{\eta_1} \;=\; \alpha^{\rho_2 \,-\, \rho_1} \\ &\Longleftrightarrow\; \log\left(\frac{\eta_2}{\eta_1}\right) \;=\; (\rho_2 \,-\, \rho_1)\, \log(\alpha) \end{aligned}$$

Case (I): $\Phi_1 \;=\; \Phi_2$ is trivially true $\forall\, \alpha > 0$.

Case (II)a: $\rho_1 \neq \rho_2 \;\;\&\;\; \alpha \neq 1 \implies \log(1) = \underbrace{(\rho_2 - \rho_1)}_{\neq 0}\, \underbrace{\log(\alpha)}_{\neq 0} \quad \Rightarrow\Leftarrow$

Case (II)b: $\rho_1 \neq \rho_2 \;\;\&\;\; \alpha = 1 \implies \log(1) = (\rho_2 - \rho_1)\, \log(1)$ is true $\forall\, \rho_2, \rho_1$

Case (III): $\eta_1 \neq \eta_2 \implies \log\left(\dfrac{\eta_2}{\eta_1}\right) = 0 \quad \Rightarrow\Leftarrow$

Case (IV)a: $\eta_1 \neq \eta_2 \;\;\&\;\; \rho_1 \neq \rho_2 \;\;\&\;\; \alpha \neq 1 \implies \alpha = \left(\dfrac{\eta_2}{\eta_1}\right)^{(\rho_2 - \rho_1)^{-1}}$

Case (IV)b: $\eta_1 \neq \eta_2 \;\;\&\;\; \rho_1 \neq \rho_2 \;\;\&\;\; \alpha = 1 \implies \log\left(\dfrac{\eta_2}{\eta_1}\right) = 0 \quad \Rightarrow\Leftarrow$

Cases (II)b and (IV)a suggest that $\mathcal{Q}$ is identifiable at leading order subject to the *identifiability conditions*:

$$\begin{cases} \text{when } \rho_1 \neq \rho_2, \text{ then } \alpha \neq 1 \\ \text{when } \rho_1 \neq \rho_2 \;\;\&\;\; \eta_1 \neq \eta_2, \text{ then } \alpha \neq \left(\dfrac{\eta_2}{\eta_1}\right)^{(\rho_2 - \rho_1)^{-1}} \end{cases}$$

□

S2 Results

The following are supplementary figures for Section 3.4 (Approximate Global $\mathcal{Q}$), specifically the 'Proposed Analytical Distributions' subsection.

Figure S2.16 is an analogous Cullen and Frey graph for η (see Figure 8 in Section 3.4).

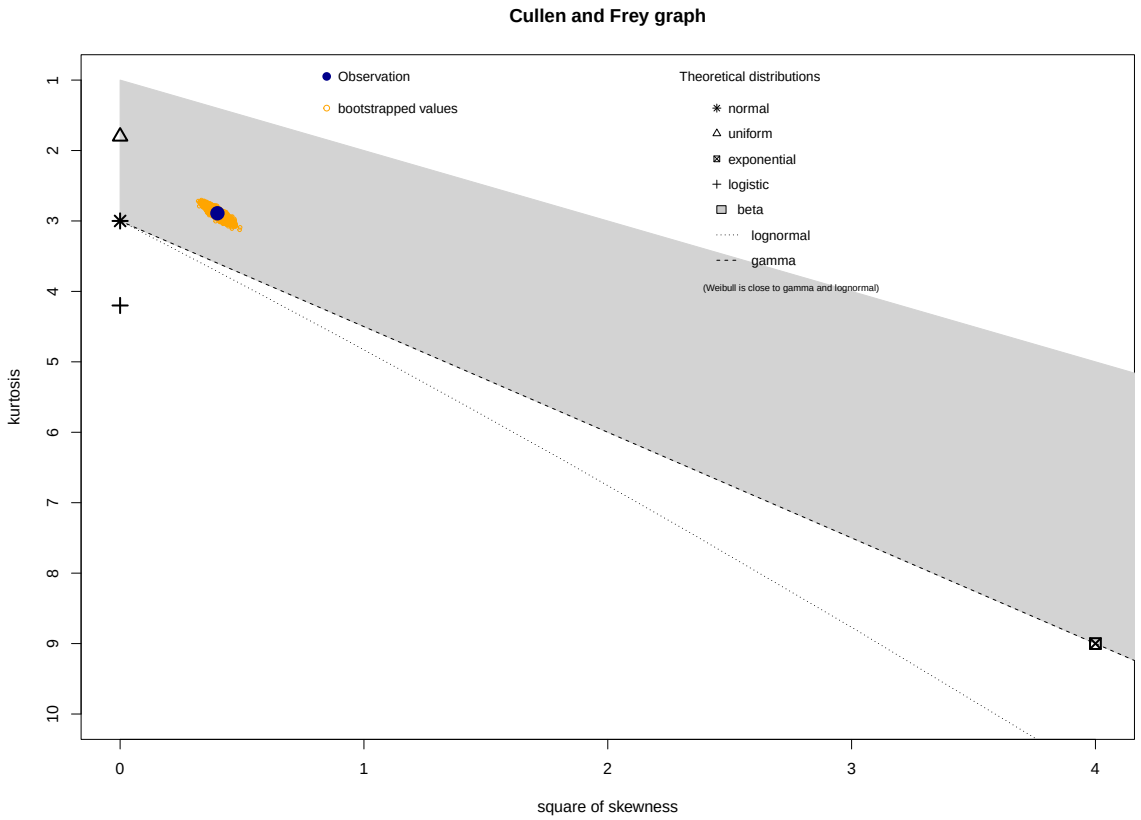


Figure S2.16: A Cullen and Frey graph showcasing the estimated kurtosis and squared skewness of our approximate posterior distribution for η. The blue and orange dots represent the original calculation and $10,000$ subsequent iterations, respectively.

Figure S2.17 showcases the best fits ultimately chosen as analytical distributions for η and ρ, respectively. (These correspond to Equations (15) and (16) in Section 3.4). These were outputs of the `fitdistrplus` package in `R`. They include a comparison of the empirical (discrete approximate distribution) and theoretical (proposed analytical distribution) PDFs and CDFs, as well as quantile-quantile (Q-Q) and percentile-percentile (P-P) comparison plots. Recall that the analytical distributions ultimately chosen for each of ρ and η were based on which distribution minimised the AIC (we considered normal, log-normal, gamma, and Weibull distributions for each parameter; see Table 5 for their respective AIC values).

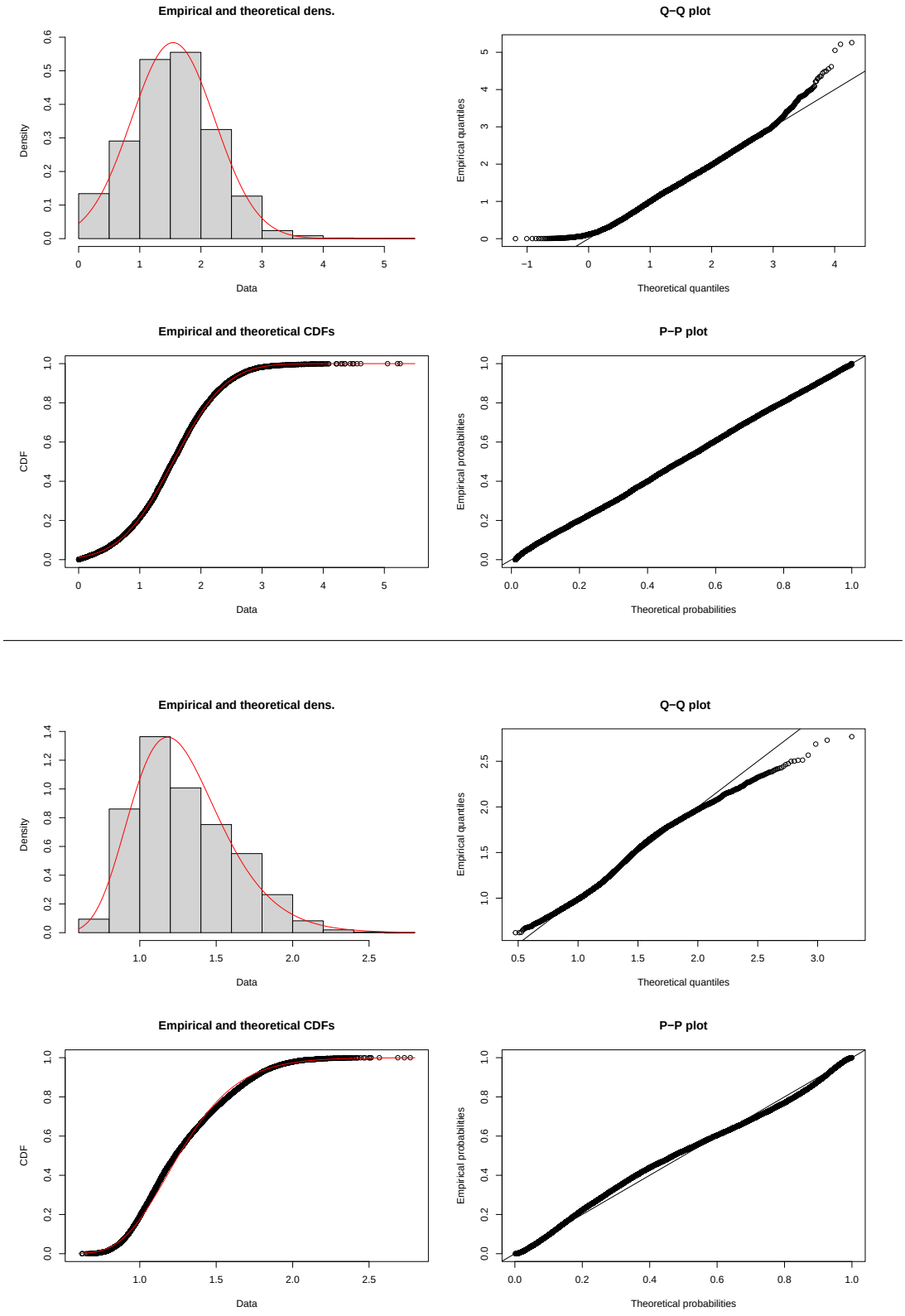


Figure S2.17: A collection of statistical comparison plots between the discrete approximate posterior distributions and proposed analytical distributions of ρ (top) and η (bottom).