Additional File 1 - Uncovering mechanisms behind mosquito seasonality by integrating mathematical models and daily empirical population data: *Culex pipiens* in the UK

**S1 DDE Model Framework**

This section gives a summary of the DDE model framework. Further details of the model framework are given by Ewing et al. [1]. The four state equations which correspond to eggs, $E(t)$, larvae, $L(t)$, pupae, $P(t)$ and adults $A(t)$ at time $t$, are

\[
\begin{align*}
\frac{dE}{dt} &= R_E(t) - M_E(t) - \delta_E(T(t))E(t), \\
\frac{dL}{dt} &= R_L(t) - M_L(t) - (\delta_L(T(t)) + \delta_{DD}(L(t)))L(t), \\
\frac{dP}{dt} &= R_P(t) - M_P(t) - \delta_P(T(t))P(t), \\
\frac{dA}{dt} &= R_A(t) - \delta_A(T(t))A(t),
\end{align*}
\]

(S1)

where $T(t)$ gives the temperature at time $t$, $\delta_i(T)$ ($i = E, L, P, A$) represents the stage-specific, density-independent, temperature-driven, mortality rate, and $R_i(t)$ and $M_i(t)$ represent the rate of recruitment to and maturation from stage $i$ respectively. The density-dependent mortality term, $\delta_{DD}(L(t))$, incorporates mortality through predation, $\delta_\pi(L(t))$, as in the Ewing et al. [1] model, but also allows for the addition of larval competition, $\delta_{LC}(L(t))$, and can be written as $\delta_{DD}(L(t)) = \delta_\pi(L(t)) + \delta_{LC}(L(t))$.

The maturation equations are defined as

\[
\begin{align*}
R_E(t) &= b(t, T)A(t), \\
M_E(t) &= R_L(t) = R_E(t - \tau_E(t))S_E(t) \frac{g_E(T(t))}{g_E(T(t - \tau_E(t)))}, \\
M_L(t) &= R_P(t) = R_L(t - \tau_L(t))S_L(t) \frac{g_L(T(t))}{g_L(T(t - \tau_L(t)))}, \\
M_P(t) &= R_A(t) = R_P(t - \tau_P(t))S_P(t) \frac{g_P(T(t))}{g_P(T(t - \tau_P(t)))},
\end{align*}
\]

(S2)

with $g_i(T(t))$ as the development rate at temperature $T(t)$, $b(t, T)$ as the egg-laying rate, $\tau_i(t)$ and $S_i(t)$ as the survival of individuals in stage $i$ ($i = E, L, P$) at time $t$ respectively. The proportion of individuals which survive from recruitment into one class, to maturation to the next, is defined by the following sequence of DDEs.
\[
\frac{dS_E}{dt} = S_E(t) \left( \frac{g_E(T(t)) \delta_E(T(t) - \tau_E(t))}{g_E(T(t) - \tau_E(t))} - \delta_E(T(t)) \right),
\]
\[
\frac{dS_L}{dt} = S_L(t) \left[ \delta_{DD}(t - \tau_L(t)) + \delta_L(T(t) - \tau_L(t)) \right] \left( \frac{g_L(T(t))}{g_L(T(t) - \tau_L(t))} \right) - \delta_{DD}(L(t)) - \delta_L(T(t)),
\]
\[
\frac{dS_P}{dt} = S_P(t) \left[ \frac{g_P(T(t)) \delta_P(T(t) - \tau_P(t))}{g_P(T(t) - \tau_P(t))} - \delta_P(T(t)) \right].
\]

Lastly, the rate of change of the duration of each life stage is given by
\[
\frac{d\tau_i(t)}{dt} = 1 - \frac{g_i(T(t))}{g_i(T(t) - \tau_i(t))},
\]
and the duration of the gonotrophic cycle, \( \tau_G(t) \), is given by
\[
\frac{d\tau_G(t)}{dt} = 1 - \frac{g_G(T(t))}{g_G(T(t) - \tau_G(t))}.
\]

**S2 New and updated functional forms**

**S2.1 Adult mortality rate**

Motivated by the field observations, we hypothesise that old, post-diapause females experience an increased mortality due to the negative costs which diapause can exert on fitness [2]. The functional form for the adult death rate, \( \delta_A(t) \), has been modified from Ewing et al. [1] to incorporate an additional post-diapause death term, which is supported by the decreased adult abundances observed in May, following the initial peak in abundances upon diapause emergence. This term was added to the Gaussian function used in Ewing et al. [1]. The updated adult death rate function is given by

\[
\delta_A(t, T(t)) = \begin{cases} 
\alpha_A T(t)^{\eta_A} + \frac{\Gamma}{\sqrt{2\pi}\sigma^2} \exp \left( \frac{-(t - \tau_G(t) - D)^2}{2\sigma^2} \right), & T(t) > \left( \frac{b_{da}}{\alpha_A} \right) \frac{1}{\eta_A} \\
\text{updated parameter} & \text{otherwise} 
\end{cases}
\]

Here \( \alpha_A \) and \( \eta_A \) are constants fitted to data from the literature regarding the temperature dependence of adult longevity, as in Ewing et al. [1]. \( \Gamma \) is a scaling parameter defining the strength of the post-diapause mortality effect, \( \sigma^2 \) controls the length of time over which this post-diapause mortality acts and \( D \) is the day of the year on which an arbitrary threshold value (80%) of adults have exited diapause. The death rate was constrained not to drop below a base death rate of \( b_{da} \), which determines the mortality of diapausing females. The value for \( \Gamma \) was chosen such that the post-diapause death rate was sufficient to wipe out the adult population surviving from the previous year. The mortality duration parameter, \( \sigma^2 \), was chosen to maximise the correlation between the adult field data and the model-predicted adult abundance over the duration which the adult mortality acted (in late May) by increasing \( \sigma^2 \) in increments of 0.1 and choosing the best fit. The diapause exit parameter, \( D \), was chosen to coincide with the end of the first adult abundance peak in the field data, when it was assumed that the majority of the population had left diapause.
S2.2 Gonotrophic cycle

A logistic functional form was chosen for the gonotrophic cycle development rate because this gave a better fit to the data [3–5] than the logarithmic functional form originally presented by Ewing et al. [1] (adjusted increased from $R^2 = 0.89$ to $R^2 = 0.90$), whilst including fewer parameters. The logistic functional form is also differentiable at all points, where the original function was not. The logistic functional form is given by,

$$g_G(T(t)) = \frac{q_1}{1 + q_2 \exp(-q_3 T(t))}, \quad (S7)$$

where $g_G(T(t))$ is the rate of progression of the gonotrophic cycle at temperature $T(t)$, with $q_1$, $q_2$ and $q_3$ as fitted constants.

S2.3 Larval competition

Mortality due to intra-specific larval competition has been shown to occur in a range of mosquito species [3, 6–8]. Linear, exponential, quadratic and log-linear functional forms were considered for the mortality rate due to larval competition, with the exponential and quadratic forms being shown to give the best fit to Cx. pipiens data presented by Madder et al. [3] (Residual standard errors (RSE) for each form were used as the $R^2$ statistic is not valid for the nonlinear exponential model: exponential RSE= 0.0077, quadratic RSE= 0.0077, linear RSE= 0.015, log-linear RSE= 0.020). The exponential form was chosen over the quadratic to prevent negative mortality rates at low larval densities. Consequently, larval competition is represented by

$$\delta_{LC}(L(t)) = c_0 \exp \left( c_1 \frac{L(t)}{V} \right), \quad (S8)$$

where $c_0$ and $c_1$ are constants fitted by nonlinear least squares fitting. The exponential function fit to the Madder et al. [3] data is shown in Figure S1. Note that the underlying temperature dependent mortality rate was accounted for before fitting the death rate due to competition.

S2.4 Seasonally forced predation

The strength of seasonal predation is affected by the ratio of predators to prey, the attack rate of predators and the predators’ handling time. The three parameters have high uncertainty, as attack rates and handling times will vary greatly between predator species [9] and the ratio of predators to prey will vary seasonally, by location, and by species [10]. Similar to mosquito development and survival, predator attack rates and handling times may also be affected by temperature. As used by Ewing et al. [1], the larval death rate due to predation is given by the Holling type II function [11, 12]

$$\delta_\pi(L(t)) = \frac{aP(t)}{V + ahL(t)}, \quad (S9)$$

where $a$ is the attack rate, $h$ is the handling time, $V$ is the volume of habitat and $P(t)$ is the predator density at time $t$. Under a “constant predation” scenario it is assumed that $P(t) = rL(t)$, such that the predator density is equal to a constant proportion of the larval density. In considering seasonally forced predation, we make the extension that the proportion of predators to larvae, $r$, varies seasonally. Consequently, $r$ is replaced by $r(t)$ and predator density is related to larval density, according to
Mortality due to larval competition

Figure S1: Larval Competition: The exponential function fit to the Madder et al. [3] data is shown ($R^2 = 0.96$).

\[ P(t) = r(t)L(t) = r_{\text{max}} \left( \frac{1}{2} + \frac{1}{2} \sin \left( \frac{2\pi(t - \nu)}{365} \right) \right)^\chi L(t), \]  

(S10)

where $r(t)$ is the number of predators per larva at time $t$, $r_{\text{max}}$ is the maximum number of predators per larva, $\nu$ defines the time at which the predation peak occurs and $\chi$ defines the time period over which predation is high, as displayed in Figure S2. We assume a fixed volume of larval habitat as the volume of habitat in the field setup was fixed throughout the study. The attack rate and handling time of predators was chosen using studies on common UK predators of Cx. pipiens [13], whilst $r_{\text{max}}$, $\nu$ and $\chi$ were fitted to the field data using ABC rejection sampling.

S3 Parameter values

The full set of parameter values used in the simulations is given in Table S1.

S4 Initial history, inoculation and solver code

As in Ewing et al. [1], we solve the system of DDEs in Fortran 90 using the DDE solver (DDE_SOLVER) written by Thompson et al. [20]. The code for the model described here can be found at Ewing et al. [21]. The historical values for the system are as described in Ewing et al. [1], with all stages assumed to be empty for $t < 0$ and all temperatures assumed to be constant at 5°C. To initiate the system we assume that some inoculation takes place at $t = 0$. This consists of adding $I_0$ individuals into the adult class ($I_0 = 5000$ in this case). Simulations were begun on the 1st of January and run for 24 months, with the last 12 months of simulated values used for comparison with the field data.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_E$</td>
<td>Fit parameter in egg maturation</td>
<td>$2.20 \times 10^{-3}$</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$\beta_E$</td>
<td>Fit parameter in egg maturation</td>
<td>1.77</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$\alpha_L$</td>
<td>Fit parameter in larval maturation</td>
<td>$3.15 \times 10^{-3}$</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$\beta_L$</td>
<td>Fit parameter in larval maturation</td>
<td>1.12</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$\alpha_P$</td>
<td>Fit parameter in pupal maturation</td>
<td>$7.11 \times 10^{-4}$</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$\beta_P$</td>
<td>Fit parameter in pupal maturation</td>
<td>1.89</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$\mu_0E, \mu_0L, \mu_0P$</td>
<td>Baseline immature death rate</td>
<td>0.0157</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$\mu_1E, \mu_1L, \mu_1P$</td>
<td>Optimum temperature for immature survival</td>
<td>20.5</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$\mu_2E, \mu_2L, \mu_2P$</td>
<td>Width parameter for immature death rate</td>
<td>7</td>
<td>Estimated from laboratory data</td>
</tr>
<tr>
<td>$\alpha_A$</td>
<td>Fit parameter in adult death</td>
<td>$2.17 \times 10^{-8}$</td>
<td>Ciota et al. [14]</td>
</tr>
<tr>
<td>$\beta_A$</td>
<td>Fit parameter in adult death</td>
<td>4.48</td>
<td>Ciota et al. [14]</td>
</tr>
<tr>
<td>$b_{di}$</td>
<td>Threshold immature death rate</td>
<td>$\frac{1}{20}$</td>
<td>Almirón et al. [15] and Loetti et al. [16]</td>
</tr>
<tr>
<td>$b_{da}$</td>
<td>Baseline adult death rate</td>
<td>0.003</td>
<td>Sulaiman et al. [17], Onyeka et al. [18], and Bailey et al. [19]</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>Fit parameter in competition</td>
<td>0.00319</td>
<td>[3]</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>Fit parameter in competition</td>
<td>0.00469</td>
<td>[3]</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Attack rate of predators</td>
<td>1.03</td>
<td>Onyeka [13]</td>
</tr>
<tr>
<td>$h$</td>
<td>Handling time of predators</td>
<td>0.043</td>
<td>Onyeka [13]</td>
</tr>
<tr>
<td>$V$</td>
<td>Volume of larval habitat</td>
<td>20l</td>
<td>By calculation</td>
</tr>
<tr>
<td>$R$</td>
<td>Egg raft size</td>
<td>200</td>
<td>Vinogradova [4]</td>
</tr>
<tr>
<td>$q_1$</td>
<td>Gonotrophic cycle fit parameter</td>
<td>0.202</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$q_2$</td>
<td>Gonotrophic cycle fit parameter</td>
<td>74.5</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$q_3$</td>
<td>Gonotrophic cycle fit parameter</td>
<td>0.246</td>
<td>Ewing et al. [1]</td>
</tr>
<tr>
<td>$\xi_S$</td>
<td>Spring diapause threshold</td>
<td>13.7</td>
<td>From fieldwork</td>
</tr>
<tr>
<td>$\xi_A$</td>
<td>Autumn diapause threshold</td>
<td>15</td>
<td>From fieldwork</td>
</tr>
<tr>
<td>$\omega_S$</td>
<td>Spring diapause transition</td>
<td>5</td>
<td>From fieldwork</td>
</tr>
<tr>
<td>$\omega_A$</td>
<td>Autumn diapause transition</td>
<td>3.5</td>
<td>From fieldwork</td>
</tr>
<tr>
<td>$\Gamma$</td>
<td>Post-diapause mortality mult.</td>
<td>8</td>
<td>From simulation</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>Post-diapause mortality duration</td>
<td>4</td>
<td>From simulation</td>
</tr>
<tr>
<td>$D$</td>
<td>80% diapause exit threshold day of year</td>
<td>109</td>
<td>From fieldwork</td>
</tr>
<tr>
<td>$r_{max}$</td>
<td>Peak predators per larva</td>
<td>0.0214</td>
<td>From ABC fitting</td>
</tr>
<tr>
<td>$\upsilon$</td>
<td>Predation timing parameter</td>
<td>19.84</td>
<td>From ABC fitting</td>
</tr>
<tr>
<td>$\chi$</td>
<td>Predation sharpness parameter</td>
<td>2.45</td>
<td>From ABC fitting</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>Latitude</td>
<td>51.6</td>
<td>Latitude of Wallingford field site</td>
</tr>
</tbody>
</table>

Table S1: Parameter values used to run the DDE model simulations.
Figure S2: Predator seasonal forcing: The seasonal forcing function, $r(t)$, is shown, highlighting how changes to $\nu$ and $\chi$ affect the ratio of predators to larvae throughout the season.

S5 Dynamic time warping

Dynamic time warping [22] was used to compare the predicted and observed patterns of abundance as the timing of features in the data did not always align in time. Consequently, traditional measures like route mean squared error (RMSE) were not able to assess the ability of the model to capture features in the data when the stage duration was either over- or under-estimated. The warping applied in comparing the fitted egg-to-pupae procedure to the field data is shown in Figure S3.

S6 ABC fitting including competition

The priors and posterior distributions for the three predation parameters and both competition parameters are shown in Figure S4. In all cases the difference between the parameter estimates from fitting only predation parameters using ABC and fitting predation and competition parameters using ABC is small.

S7 Immature stage durations

The immature stage durations at all time points throughout the season are shown in Figure S5. Note that none of the immature stage durations are seen to steadily increase throughout the season, which suggests that the steadily increasing discrepancy between observed and predicted stage durations in Figure 6 (a) does not stem from one specific stage duration being increasingly overestimated through the year.
Figure S3: DTW mapping: The mappings applied by the DTW mapping algorithm are shown. Plot (a) shows the indices of the field observations on the $x$-axis and the model predictions on the $y$-axis and the red line shows the relationship if no time warping were applied. Plot (b) shows the points on each time series mapped onto their corresponding points on the other.

Figure S4: Prior and posteriors from ABC fitting: The priors and posterior distributions for the ABC fitting run with all three predation parameters, $r_{max}$, $\nu$ and $\chi$, and both competition parameters $c_0$ and $c_1$ are shown. The solid vertical lines show the median of the posterior distribution in each case. In the case of the predation parameters, the dotted vertical lines show the median of the posterior distribution fitted using the experimentally derived competition parameters. The vertical dotted lines on the plots for the competition parameters show the experimentally derived estimates.
Figure S5: Immature stage durations: The immature stage durations based on the hourly temperature data from butt 4 are shown.

S8 Seasonality under constant predator to prey ratio

The extent to which the addition of a seasonally forced predator-prey ratio was influencing the observed pattern of seasonal abundance was investigated by comparing the results of the DDE model (Figure 7) with the output when the ratio of predators to prey was held constant by setting $\chi = 0$. 
Figure S6: Model results under constant predation: The field data (black line) from butt 4 is shown against the full DDE model predictions (red line) with the ratio of predators to prey held constant by setting $\chi = 0$. The scaled abundances presented are 7-day moving averages of the field and model-predicted abundances.
References


