

QUANTIFYING FUNCTIONALS OF AGE DISTRIBUTIONS IN THE WILD BY SOLVING AN OPERATOR EQUATION

January 2017

Hao Ji¹, Hans-Georg Müller^{1,2}, Nikos T. Papadopoulos³ and James R. Carey⁴

ABSTRACT

Residual demography is a recent concept that has proved to be a useful tool to gain insights about the age distributions of wild populations, especially insects. We develop an operator equation that permits the derivation of functionals of the age distribution in wild populations, such as mean age, within the framework of residual demography. Our method combines information from an observed captive cohort, which consists of subjects that are sampled from the wild with unknown ages and then raised in the laboratory until death, and from a reference cohort that consists of subjects raised in the laboratory since birth of the same population. Targeting functionals such as the mean of the wild age distribution has the advantage of avoiding strong assumptions such as stationarity and stability of the population that one would need when targeting the entire survival distribution in the wild. Our main result characterizes the existence of a solution of the operator equation that yields the functional of interest. The proposed method also enjoys straightforward and easy implementation. A data example is included illustrating an application, where one aims to attain the mean age of mosquitoes in the wild, based on seasonal captive cohorts from Greece and a simulated reference cohort, separately for various summer and fall months.

KEY WORDS: Aging in the wild · *Culex pipens* · Functional singular representation · Existence of solution · Inverse problem · Operator Equation · Residual demography

¹Department of Statistics, University of California, One Shields Avenue, Davis, CA 95616, USA

²Research supported by NSF grants DMS-1104426 and DMS-1407852.

³Department of Agriculture, Crop Production and Rural Environment, University of Thessaly, Fytokou Street, N. Ionia, GR-384 46, Volos, Greece

⁴Department of Entomology and Nematology, University of California, One Shields Avenue, Davis, CA 95616, USA

1 Introduction

Information about the age distribution and survival of insect and animal populations is of much interest in ecology and biodemography. In particular, for insect populations, many of the existing methods, such as the use of mechanical damage, chemical analysis and gene expression, are expensive, require major training and calibration efforts, and do not guarantee accuracy at older ages (Rao and Carey (2015)). Analysis of captive cohorts of insects has emerged as a useful method to reveal the characteristics of wild aging populations (see for example Müller et al. (2004); Carey et al. (2008); Carey et al. (2012); Müller et al. (2007); Goldstein (2009); Vaupel (2009)). In this paper, we develop a method to obtain certain functionals of the wild age distribution with observed data from residual demographical profiles and information from laboratory reference cohorts which involves solving an operator equation. Our main result below provides sufficient criteria for the existence of a solution of this equation.

The approach and operator equation are motivated by data that consist of residual demographical profiles of 1504 West Nile virus mosquitoes (*Culex pipens*). These mosquitoes were randomly and repeatedly sampled from the wild population at an unknown age. They were then transferred to captive cohorts that were reared out in the laboratory with the date of transfer and the time from transfer to death recorded for each of the mosquitoes in the cohort. Data collection began in June 2013 and lasted six months. Further background on this study can be found in Papadopoulos et al. (2015). Of interest are accompanying changes in the residual lifespans, see the exploratory plot in Figure 1, which shows the data from this study, overlaid with a smooth nonparametric local least squares fit for the conditional mean function $E(y|X = x)$, where x is the capture date and y is the captive lifetime. It is of great interest to gain insights on aging in the wild by utilizing these or similar data, which requires the development of appropriate methodological tools.

One approach to gain such insight is residual demography, where a reference cohort of mosquitoes is raised independently from the captive cohort in the laboratory under identical laboratory conditions from birth. Under basic assumptions on the dependency of mortality on age and current environment, no age-bias of captive cohort sampling, and stationarity and stability of the population (Preston et al. (2001)), a demographic convolution equation relates the unknown wild age distribution, the observed survival function from the observed captive cohort and the observed survival function of the reference cohort. This convolution equation has been derived under suitable assumptions and is related to

renewal theory (see Lotka (1939); Cox (1962); Feller (2008)). It suggests to adopt a deconvolution method (Müller et al. 2007) to extract the survival distribution of the wild population.

Due to the necessary strong assumptions when targeting the survival distribution, such as stationarity of the population, we aim here at the simpler target to obtain functionals of the age distribution of the wild population. Since targeting of the entire age distribution requires deconvolution, which is sensitive to the choice of various tuning parameters (Müller et al. 2007), we target here instead functionals of the age distribution in the wild, such as mean age.

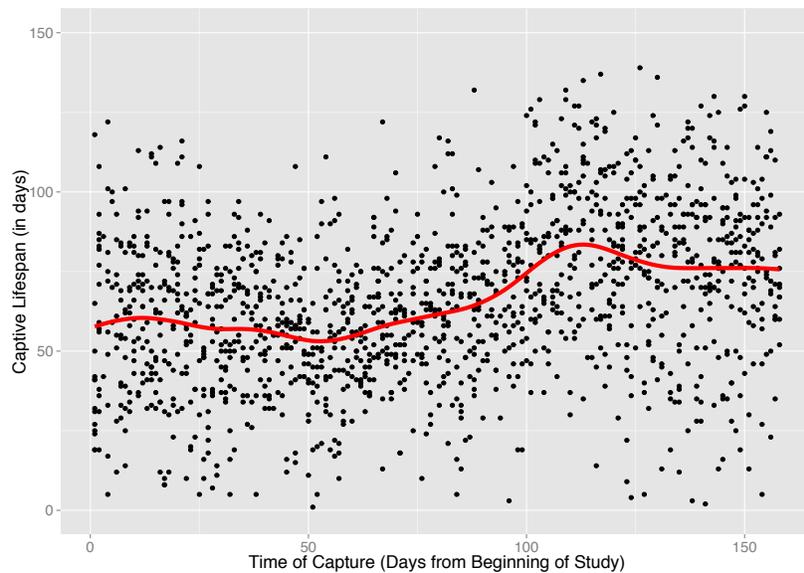


Figure 1: Remaining lifespan in the captive cohort versus time of capture (in days) for mosquitoes captured from June to November. The overlaid curve is the local least squares estimate of the conditional mean of remaining lifespan, given day of capture.

2 Functionals of the Age Distribution in the Wild

For an individual which is captured at unknown age W , then enters the captive cohort and dies after an observed residual life time in captivity C , age-at-death D can be written as

$$D = W + C.$$

Throughout the paper, we assume that $D \in [0, T]$ for some sufficiently large T , which is plausible since lifespan is a bounded quantity in biological applications. Let $F_X(t) = P(X \leq t)$ denote the distribution function, $\bar{F}_X(t) = 1 - F_X(t)$ the survival function, $f_X(t) = \frac{d}{dt}F_X(t)$ the probability density function, $\lambda_X(t) = f_X(t)/\bar{F}_X(t)$ the hazard rate, and $\Lambda_X(t) = \int_0^t \lambda_X(s) ds$ the cumulative hazard rate of a non-negative random variable X . Here the choice $X = W$ will label these quantities for subjects in the wild, $X = C$ for subjects in the captive cohort, and $X = R$ for subjects in the reference cohort. Furthermore, the density of the distribution of age-at-capture, which is the same as the density of age-in-the-wild, under the unbiased sampling condition that we impose, is denoted by f_W . The corresponding cumulative distribution function is F_W .

To further motivate our approach, we illustrate the estimated densities f_C of the distribution of age-at-death or equivalently survival in the captive cohorts that were obtained for the months June–November from the data in Figure 1. These densities were obtained with kernel estimators (Silverman 1986).

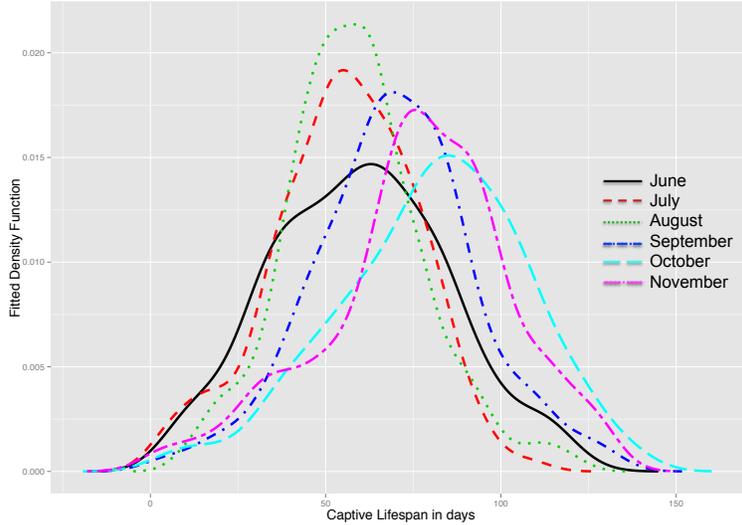


Figure 2: Fitted density functions of captive lifespan for mosquitoes captured in different months.

Our target in general are functionals of the wild age distribution. Given a specific function $h \in L^2([0, T]) = \{f(t) : \int_0^T |f(t)|^2 dt < \infty\}$, that is chosen based on a specific feature of interest

inherent in the wild age distribution F_W , we define a linear functional

$$H(F_W) : L^2([0, T]) \rightarrow \mathbb{R}, \quad H(F_W) = \int_0^T h(w) dF_W(w) = \int_0^T h(w) f_W(w) dw. \quad (1)$$

An important example is the expected value of age in the wild population,

$$EW = \int_0^T w f_W(w) dw,$$

where $h(w) = w$. Examples of other functionals of interest include the variance of the age distribution in the wild, which is related to the second moment,

$$EW^2 = \int_0^T w^2 dF_W(w) = \text{Var}(W) + (EW)^2,$$

from which we can obtain $\text{Var}(W) = EW^2 - (EW)^2$. Another example is the population proportion of subjects that are younger than a specific age,

$$P(W \leq t) = \int_0^T \mathbf{1}\{w \leq t\} dF_W(w),$$

where $\mathbf{1}\{w \leq t\}$ is the indicator function. Targeting higher order moments or proportions for tails of the distribution is more challenging than targeting the mean due to decreased numerical stability, which means that one needs larger sample sizes for the reference cohort and associated more precise estimates of the survival schedule of subjects under laboratory conditions to target such more complex functionals.

We adopt the following assumptions from Müller et al. (2007), where also the biological motivation for these assumptions is given:

1. The force of mortality depends only on the age of an individual and the current environment, and does not depend on past environmental exposure.
2. Age-bias of captures in the wild is negligible.
3. The population is stable, and birth rates are stationary throughout the observation period, or capture period.

A fourth assumption that was made in Müller et al. (2007) concerns the stationarity of the population, meaning that hazard rates depend on age of an individual but not on calendar time. As survival may depend on changing environmental conditions such as variations in food supply, this assumption

may be violated in certain situations. However, this assumption is only needed in order to target survival rates in the wild but is not needed for inference on the age distribution, which is our goal here. Specifically, we note that the third assumption above is not the same as requiring a stationary age distribution. In fact, since we target functionals of the age distribution that is present at the time of capture of flies from the wild, this age distribution is allowed to change in dependence on the capture time (referring to its calendar time). To illustrate our methods, in Section 5 we obtain estimates of the mean age in the wild for cohorts that are sampled at different calendar times and have different age distributions.

Under assumptions 1-3, the following *convolution equation* (Müller et al. (2007)) holds, which characterizes the relationship between F_C , F_W and F_R , and is given by:

$$\bar{F}_C(t) = \int_0^T e^{-(\Lambda_R(t+w)-\Lambda_R(w))} f_W(w) dw = \int_0^T \frac{\bar{F}_R(t+w)}{\bar{F}_R(w)} f_W(w) dw, \quad (2)$$

based on the relationship $\bar{F}_R(t) = e^{-\Lambda_R(t)}$. Now define the kernel function,

$$g(t, w) = \frac{\bar{F}_R(t+w)}{\bar{F}_R(w)}, \quad (3)$$

which can be deduced from the reference cohort, since this cohort provides information about \bar{F}_R . For a function $\beta \in L^2([0, T])$, we then introduce the function

$$\gamma_\beta(w) = \int_0^T g(t, w)\beta(t)dt, \quad (4)$$

and observe that $\gamma_\beta \in L^2([0, T])$. Now we consider a linear functional of type (1), where we make the dependency on β explicit,

$$\Gamma_\beta(F_W) = \int_0^T \gamma_\beta(w)dF_W(w) = \int_0^T \int_0^T g(t, w)\beta(t)f_W(w)dt dw = \int_0^T \bar{F}_C(t)\beta(t)dt, \quad (5)$$

using equation (2). Importantly, for any given $\beta \in L^2([0, T])$, the r.h.s of (5) and therefore the functional Γ_β can be evaluated, using information available from the captive cohort.

Now, going back to (1), we have specified a function $h \in L^2([0, T])$ of interest. Since we can evaluate functionals of type (5), our goal is therefore to find a function β such that

$$h(w) = \gamma_\beta(w), \quad (6)$$

whence we then obtain

$$H(F_W) = \Gamma_\beta(F_W) = \int_0^T \bar{F}_C(t)\beta(t)dt. \quad (7)$$

So, if equation (6) can be solved, we have devised a tool to evaluate the bio-demographically relevant quantity $H(F_W)$.

3 Existence of a Solution

Since existence of a function β satisfying (6) allows us to obtain the target functional, it is of interest to study conditions under which existence is guaranteed. Define linear operators $\mathcal{G} : L^2([0, T]) \rightarrow L^2([0, T])$ as

$$\mathcal{G}(r)(w) = s(w) = \int_0^T g(t, w)r(t)dt. \quad (8)$$

Following Yang et al. (2011) (see also Kato (1995); Baker (1973)) we define the adjoint operator $\mathcal{G}^* : L^2([0, T]) \rightarrow L^2([0, T])$ of \mathcal{G} ,

$$\mathcal{G}^*(r)(w) = \int_0^T g(w, t)r(t)dt. \quad (9)$$

We next consider the two compound operators $\mathcal{A}_1 = \mathcal{G} \circ \mathcal{G}^*$ and $\mathcal{A}_2 = \mathcal{G}^* \circ \mathcal{G}$

$$\mathcal{A}_1(r)(w) = \int_0^T A_1(w, t)r(t)dt \quad (10)$$

$$\mathcal{A}_2(r)(w) = \int_0^T A_2(w, t)r(t)dt, \quad (11)$$

where

$$A_1(w, t) = \int_0^T g(u, w)g(u, t)du$$

$$A_2(w, t) = \int_0^T g(w, u)g(t, u)du$$

are the kernels of these linear operators. Note that \mathcal{A}_1 and \mathcal{A}_2 are linear Hilbert-Schmidt operators with L^2 kernels (see section V.3 of Kato (1995)). According to the spectral theorem for Hilbert-Schmidt operators, the kernels of \mathcal{A}_1 and \mathcal{A}_2 can be decomposed in the following way

$$A_1(w, t) = \sum_{k=1}^{\infty} \sigma_k^2 \psi_k(w) \psi_k(t),$$

$$A_2(w, t) = \sum_{k=1}^{\infty} \sigma_k^2 \phi_k(w) \phi_k(t),$$

where $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq 0$ are the shared eigenvalues of the two compound operators and ψ_1, ψ_2, \dots and ϕ_1, ϕ_2, \dots are orthonormal eigenfunctions for \mathcal{A}_1 and for \mathcal{A}_2 , respectively, with

$$\mathcal{A}_1(\psi_k) = \sigma_k^2 \psi_k, \quad \langle \psi_i, \psi_j \rangle = \int_0^T \psi_i(t) \psi_j(t) dt = \delta_{ij},$$

$$\mathcal{A}_2(\phi_k) = \sigma_k^2 \phi_k, \quad \langle \phi_i, \phi_j \rangle = \int_0^T \phi_i(t) \phi_j(t) dt = \delta_{ij},$$

where δ_{ij} is the Kronecker delta. Furthermore, one has $\psi_k = \frac{1}{\sigma_k} \mathcal{A}_1(\phi_k)$ for positive $\sigma_k, k = 1, 2, \dots$

We refer to the sequences ψ_1, ψ_2, \dots and ϕ_1, ϕ_2, \dots as singular functions and to the sequence $\sigma_1, \sigma_2, \dots$ as singular values, where $\sigma_1 \geq \sigma_2 \geq \dots \geq 0$. We use the following singular representation of the operators \mathcal{G} and \mathcal{G}^* to characterize solutions β , which can be found in section V.3 of Kato (1995) or in Conway (1985).

Proposition. *Operators \mathcal{G} and \mathcal{G}^* can be represented as*

$$\mathcal{G}(f)(w) = \sum_{k=1}^{\infty} \sigma_k \langle f, \psi_k \rangle \phi_k(w), \quad \mathcal{G}^*(f)(w) = \sum_{k=1}^{\infty} \sigma_k \langle f, \phi_k \rangle \psi_k(w), \quad w \in (0, T). \quad (12)$$

Now assume that $\{\psi_k\}_{k=1}^{\infty}$ and $\{\phi_k\}_{k=1}^{\infty}$ form an orthonormal basis of $L^2([0, T])$, respectively, so that we may represent, in the L^2 sense, the functions β and h as follows:

$$\beta(t) = \sum_{k=1}^{\infty} \beta_k \psi_k(t), \quad h(w) = \sum_{k=1}^{\infty} h_k \phi_k(w).$$

With the singular representation in (12), (6) becomes

$$\sum_{k=1}^{\infty} h_k \phi_k(w) = h(w) = \gamma_{\beta}(w) = \mathcal{G}(\beta)(w) = \sum_{k=1}^{\infty} \sigma_k \beta_k \phi_k(w). \quad (13)$$

Now we are in a position to characterize the existence of a solution function β of (6).

Theorem. *Given a linear operator $\mathcal{G} : L^2([0, T]) \rightarrow L^2([0, T])$ with kernel $g(t, w)$ and its adjoint operator \mathcal{G}^* as defined in (8) and (9), and the singular representations described in (12), let $\{\psi_k\}_{k=1}^{\infty}$ and $\{\phi_k\}_{k=1}^{\infty}$ be the eigenfunctions of the compound operators \mathcal{A}_1 and \mathcal{A}_2 defined in (10) and (11), and σ_k the singular values. Assume the function h can be expanded as $h(w) = \sum_{k=1}^{\infty} h_k \phi_k(w)$ and the condition*

$$\sum_{k:\sigma_k>0} \left(\frac{h_k}{\sigma_k} \right)^2 < \infty$$

is satisfied and it holds that $\sigma_k = 0$ implies $h_k = 0$. Then there exists a function

$$\beta(t) = \sum_{k:\sigma_k>0} \frac{h_k}{\sigma_k} \psi_k(t) \quad (14)$$

that satisfies (6).

This result implies that the existence of β depends on properties of both the specified function $h(w)$ that characterizes the targeted functional and the kernel function $g(t, w)$. If the basic assumptions in the theorem are satisfied, (14) provides an explicit construction for the solution β of the

operator equation.

Proof. In order to show the existence of a solution, consider the construction

$$\beta(t) = \sum_{k=1}^{\infty} \beta_k \psi_k(t)$$

where $\beta_k = \frac{h_k}{\sigma_k}$ for $\sigma_k > 0$ and $\beta_k = 0$ for $\sigma_k = 0$.

Plugging in the representation into right hand side of (6), it follows that

$$\gamma_\beta(w) = \int_0^T g(t, w) \beta(t) dt = \mathcal{G}(\beta)(w) = \sum_{k=1}^{\infty} \sigma_k \beta_k \phi_k(w) = \sum_{\sigma_k > 0} h_k \phi_k(w).$$

By the assumptions, $h_k = 0$ when $\sigma_k = 0$ for all k , so that

$$\sum_{\sigma_k > 0} h_k \phi_k(w) = \sum_{k=1}^{\infty} h_k \phi_k(w) = h(w).$$

Therefore, the proposed β is a solution of (6). □

4 Implementation

In this section, we introduce the numerical procedures to solve the operator equation (3), aiming to obtain the target $H(F_W)$ for a given $h(w)$. Based on the previous discussion, we propose to fit the function β according to (6), (14) by minimizing the squared L^2 distance

$$\beta^* = \operatorname{argmin}_\beta \int_0^T \left\{ \int_0^T g(t, w) \beta(t) dt - h(w) \right\}^2 dw. \quad (15)$$

To minimize (15) numerically, we discretize the variables w and t , followed by a least squares algorithm. Specifically, we first choose regular time grids for both age at capture w and captive lifetime t from 0 to the maximum remaining lifespan of individuals from the captive cohort, and denote these as $\{w_1, \dots, w_q\}$ and $\{t_1, \dots, t_p\}$. Let Δt and Δw be the constant time differences of the time grids.

Define the matrix \mathbf{G} and vector β and \mathbf{h} as

$$\mathbf{G}_{q \times p} = (g_{ij}) = g(t_j, w_i),$$

$$\beta = (\beta(t_1), \dots, \beta(t_p))^T,$$

$$\mathbf{h} = (h(w_1), \dots, h(w_q))^T.$$

Then the minimization simplifies to the following matrix form:

$$\boldsymbol{\beta}^* = \arg \min_{\boldsymbol{\beta}} \Delta w \mathbf{1}_q^T (\Delta t \mathbf{G} \boldsymbol{\beta} - \mathbf{h})^2, \quad (16)$$

where $\mathbf{1}_q$ is a q -dimensional vector with all entries being 1.

In (16), one encounters a least squares problem, where the corresponding responses are $h(w_i)$, $i = 1, 2, \dots, q$, and the predictors are $x_{i1} = g(t_1, w_i)$, $x_{i2} = g(t_2, w_i)$, \dots

Note that the survival function \bar{F}_R in $g(t, w)$ is estimated with the empirical version based on the available reference cohort. In the case where data are censored, i.e. individual flies are lost to follow-up, we can replace empirical distributions with empirical Kaplan-Meier estimators (Lo et al. (1989)). Assuming linear independence of predictors, we find $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \dots, \hat{\beta}_p)^T = (\hat{\beta}(t_1), \dots, \hat{\beta}(t_p))^T$ such that

$$\sum_{i=1}^q \left\{ \sum_{j=1}^p \Delta t x_{ij} \beta_j - h(w_i) \right\}^2 \rightarrow \min. \quad (17)$$

We then estimate the functional $H(F_W)$ of the wild lifetime distribution by numerical integration,

$$\widehat{H(F_W)} = \sum_{j=1}^p \widehat{F}_C(t_j) \hat{\beta}_j \Delta t. \quad (18)$$

5 Application to Captive Cohorts of Mosquitoes

In the mosquito captive cohort data, there are 1504 mosquitoes captured during one half year period starting from June 2013. Remaining lifespans of the captured individuals, along with other information such as their physiological stage and species were recorded. We aim at estimating the mean age of the species in the wild with uncertainty control via bootstrapping. For obtaining mean age, we choose $h(w) = w$ in (1).

To illustrate the proposed method, we use a reasonably specified simulated reference cohort, with age-at-death times following the Gompertz distribution, denoted by $Gompertz(\eta, b)$, with the following probability density function,

$$f(x; \eta, b) = b \eta e^{bx} e^{\eta} e^{-\eta e^{bx}}, \quad (19)$$

where we specified the parameters η and b as follows: We had available an observed reference cohort of the same species raised in the lab, but the corresponding mean lifespan of this reference cohort was around 25 days, less than that of the observed captive cohort, which was 100 days. We therefore

used the reference cohort to estimate the parameter b in (19) via maximum likelihood estimation and then adjusted the value of the second parameter η so that the mean lifespan of the simulated reference cohort matched the longest mean lifespan of the captive cohort samples aggregated by month, for the entire observation period. This procedure led to the parameters $\eta = 0.00258$ and $b = 0.054$. The resulting reference cohort has a mean lifespan of 100 days, and covers the lifespan of the observed captive cohort.

To obtain the mean age in the wild for each month where the study was conducted, we sampled a reference cohort from the proposed Gompertz distribution, and applied the method described in the previous section. In order to construct confidence intervals for the expected age of mosquitoes in the wild for the selected month, we used the bootstrap. In the bootstrap approach, we sampled reference cohorts from the above Gompertz distribution repeatedly and sampled captive cohorts nonparametrically from the observed captive cohort for the selected month in each iteration, also repeatedly. We then estimated the expected age in the wild based on the sampled cohorts, repeating this resampling 2000 times. The 95% confidence intervals are constructed using lower and upper quantiles of the bootstrap estimates at levels 2.5% and 97.5%. The results with 2000 bootstrap iterations are listed in Table 1.

These results indicate that there are monthly trends in the mean age of the wild mosquito population. From June to August, the mean age in the wild hovers around a constant level, and then drops dramatically during September and October, to bounce back a bit in November. Compared with the residual lifetimes observed in the captive cohort as depicted in Figure 1, we find that mean age of the population in the wild shows an opposite trend. This is as expected, since flies that are on average younger when captured in October as compared to flies captured in June will live longer in the captive cohort than those captured in June. Since the results are based on a simulated reference cohort, it is worth noting that the actual mean age of the wild population might be off by a factor, depending on how well the simulated reference cohort mimics a real reference cohort. Nevertheless, the observed trends in the mean age of the wild population for different months are unlikely to strongly depend on the detailed shape of the reference cohort. From a biological perspective, a more youthful population in October is likely due to physiological shifts as mosquitoes prepare to enter hibernation mode.

Table 1: Estimated mean ages (expected values) of Mosquito populations in the wild for each month, obtained by solving operator equation (6), and 95% confidence intervals for each month, obtained by bootstrapping.

Month	Estimated Mean Age (in Days)	95% C.I.
June	89.10	(73.24, 105.77)
July	90.96	(78.63, 101.98)
August	86.00	(75.34, 98.63)
September	61.32	(46.91, 73.96)
October	35.18	(9.42, 51.97)
November	54.40	(38.42, 72.77)

6 Discussion and Conclusions

We developed an approach to infer properties of the wild age distribution that can be cast in the form of a functional of the age distribution within the residual demography framework. The method is applicable to evaluate finite moments of the wild age distribution, or proportions of individuals in the wild that fall into certain age brackets, as long as the underlying operator equation has a solution. Our theorem provides general conditions when such a solution exists.

By aiming at functionals of the wild age distribution instead of the distribution itself, fewer assumptions are needed and the approach is more robust. For the previous methods, it had to be assumed that the wild age density function and the population hazard rates solely depend on the age of an individual but do not depend on calendar time. In our illustrative example, dependence on month is of primary interest, and for the proposed approach the population stationarity assumption invoked in previous work is neither reasonable nor needed. Numerically, the proposed method is straightforward and easy to implement.

Since solving the operator equation (6) amounts to solving an inverse problem, numerical instability can be reflected by a low condition number of the matrix \mathbf{G} in (16), especially when the matrix has a large number of zero elements. This can be counteracted by choosing a sparser time grid in the discretization, and by choosing a smaller time upper bound for the time grid of the captive lifetimes t , rather than the maximum remaining lifespan in the observed captive cohort.

As an illustrating example, we applied the proposed method to the mosquito captive cohorts and simulated reference cohorts using a Gompertz distribution. We used bootstrap to provide confidence intervals for the estimates of the mean age of cohorts in the wild at different calendar times from June to November. In our application, the results are consistent with the survival schemes that are observed for the captive cohorts and reveal clear differences in the distributions of wild age for mosquitoes over different months. While of biological interest in itself, this may also have implications for the assessment of the potential for disease transmission by mosquitoes, where such transmittable diseases include Japanese encephalitis, meningitis or West Nile virus. Our approaches will also be useful to obtain information about wild age distributions and their dependency on seasonality or other features for other species.

References

- Baker, C.R.: Joint measures and cross-covariance operators. *Transactions of the American Mathematical Society* **186**, 273–289 (1973)
- Carey, J.R., Müller, H.G., Wang, J.L., Papadopoulos, N.T., Diamantidis, A., Koulousis, N.A.: Graphical and demographic synopsis of the captive cohort method for estimating population age structure in the wild. *Experimental Gerontology* **47**(10), 787–791 (2012)
- Carey, J.R., Papadopoulos, N.T., Müller, H.G., Katsoyannos, B.I., Kouloussis, N.A., Wang, J.L., Wachter, K., Yu, W., Liedo, P.: Age structure changes and extraordinary lifespan in wild medfly populations. *Aging Cell* **7**(3), 426 – 437 (2008)
- Conway, J.B.: *A Course in Functional Analysis*. Springer, New York (1985)
- Cox, D.R.: *Renewal theory*, vol. 4. Methuen and Co, London (1962)
- Feller, W.: *An introduction to probability theory and its applications*, vol. 2. Wiley & Sons (2008)
- Goldstein, J.R.: Life lived equals life left in stationary populations. *Demographic Research* **20**, 3–6 (2009)
- Kato, T.: *Perturbation theory for linear operators*. Berlin: Springer-Verlag (1995)

- Lo, S., Mack, Y., Wang, J.: Density and hazard rate estimation for censored data via strong representation of the Kaplan-Meier estimator. *Probability Theory and Related Fields* **80**(3), 461–473 (1989)
- Lotka, A.J.: A contribution to the theory of self-renewing aggregates, with special reference to industrial replacement. *The Annals of Mathematical Statistics* **10**, 1–25 (1939)
- Müller, H.G., Wang, J.L., Carey, J.R., Caswell-Chen, E.P., Chen, C., Papadopoulos, N., Yao, F.: Demographic window to aging in the wild: constructing life tables and estimating survival functions from marked individuals of unknown age. *Aging Cell* **3**(3), 125 – 131 (2004)
- Müller, H.G., Wang, J.L., Yu, W., Delaigle, A., Carey, J.R.: Survival and aging in the wild via residual demography. *Theor Popul Biol* **72**(4), 513–522 (2007)
- Papadopoulos, N., Carey, J.R., Ioannou, C., Wang, J.L., Müller, H.G., Ji, H., Luckhart, S., Lewis, E.: Seasonal patterns of post-capture longevity in wild mosquito populations (*Culex pipens*): Demographic phases and ecological implications (2015). Manuscript submitted for publication.
- Preston, S.H., Heuveline, P., Guillot, M.: *Demography: measuring and modeling population processes*, vol. 27. Blackwell Publishers, Malden (2001)
- Rao, A.S.S., Carey, J.R.: Generalization of Carey’s equality and a theorem on stationary population. *Journal of Mathematical Biology* **71**(3), 583–594 (2015)
- Silverman, B.W.: *Density estimation for statistics and data analysis*. Monographs on Statistics and Applied Probability. Chapman & Hall, London (1986)
- Vaupel, J.W.: Life lived and left: Carey’s equality. *Demographic Research* **20**, 7–10 (2009)
- Yang, W., Müller, H., Stadtmüller, U.: Functional singular component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**(3), 303–324 (2011)