Article

On a Retarded Nonlocal Ordinary Differential System with Discrete Diffusion Modeling Life Tables

Francisco Morillas [1,*] **and José Valero** [2,*]

1   Departament d'Economia Aplicada, Facultat d'Economia, Campus dels Tarongers s/n, Universitat de València, 46022 València, Spain
2   Centro de Investigación Operativa, Avda. Universidad s/n, Universidad Miguel Hernández de Elche, 03202 Elche, Spain
*   Correspondence: Francisco.Morillas@uv.es (F.M.); jvalero@umh.es (J.V.)

Abstract: In this paper, we consider a system of ordinary differential equations with non-local discrete diffusion and finite delay and with either a finite or an infinite number of equations. We prove several properties of solutions such as comparison, stability and symmetry. We create a numerical simulation showing that this model can be appropriate to model dynamical life tables in actuarial or demographic sciences. In this way, some indicators of goodness and smoothness are improved when comparing with classical techniques.

Keywords: lattice dynamical systems; ordinary differential equations; discrete nonlocal diffusion problems; retarded equations; life tables

1. Introduction

Nonlocal diffusion problems generated by convolution operators have been used widely in the literature in order to model diffusion processes. As explained in [1], a convolution operator can describe transport due to long-range dispersal mechanisms. Another application of such models is the approximation of reaction-diffusion equations, as done in [2] in the continuous case and in [3] in the discrete case. They have been also used to model phase transitions (see [4–7] among others). The long-time behavior of the solutions of such systems has been considered for example in [7,8]. In [9], the relation between a continuous nonlocal diffusion equation and its discrete counterpart was studied.

In our previous paper [10], we studied analytically and numerically the nonlocal ordinary differential system with discrete diffusion

$$\frac{d}{dt}u^i(t) = \sum_{r \in D} j_{i-r}u^r(t) - u^i(t) + \sum_{r \in \mathbb{Z} \setminus D} j_{i-r}g_r(t), i \in D, t > 0, \tag{1}$$

$$u^i(0) = u^{i,0}, i \in D,$$

where D is a suitable subset of $\mathbb{Z}$, showing that it is appropriate for modeling life tables. The main downside is that the predictions are only valid for a short term of about two or three years.

It seems, however, that some memory effects are present on life tables, so it would be better to consider a model with delay. This allows us to extend the validity of the predictions for larger periods of about five to ten years. Therefore, in this paper, we modify system (1) by means of the introduction of some delay terms. For life tables, the delay pattern is established via improvement rates [11]. We point out that our model is deterministic, non-parametric and predicts the evolution of the probability of death by age over time.

In Section 2, we study some properties of the solutions of the retarded nonlocal model such as a comparison principle, the stability of stationary points or the symmetry of

solutions under certain assumptions. The analysis is carried out for a system with either a finite or an infinite number of equations, extending to the model with delay similar results from [10].

In Section 3, we apply the new model to life tables in Spain. We use the observed values in the period from 1908 to 2012 to fit the model and then validate it by making predictions in the period 2013–2018. We compare the predicted values with the observed ones by using two indicators of goodness. In both cases, the conclusion is that the model is appropriate for making quality predictions in the middle-term. We repeat the same calculations but using the Lee-Carter model, which is seminal and paradigmatic for this theory, obtaining that in almost all cases, our model improves the aforementioned indicators when comparing it with this classical model.

2. Properties of Solutions of Non-Local Diffusion Problems with Finite Delay

Let D be a subset of $\mathbb{Z}$ of one of the following types:

$$D = \{m_1, m_1 + 1, \ldots, m_2\}, \ -\infty < m_1 < m_2 < +\infty, \tag{2}$$

$$D = \mathbb{Z}, \tag{3}$$

$$D = \{i \in \mathbb{Z} : i \leq m\}, \tag{4}$$

$$D = \{i \in \mathbb{Z} : i \geq m\}, \ m \in \mathbb{Z}. \tag{5}$$

For the cases (3)–(5), we define the Banach spaces

$$l_1^\infty = \{(u_i)_{i \in D} : \sup_{i \in D} |u_i| < \infty\},$$

$$l_2^\infty = \{(u_i)_{i \in \mathbb{Z} \setminus D} : \sup_{i \in \mathbb{Z} \setminus D} |u_i| < \infty\},$$

$$l_1^p = \{(u_i)_{i \in D} : \sum_{i \in D} |u_i|^p < \infty\},$$

$$l_2^p = \{(u_i)_{i \in \mathbb{Z} \setminus D} : \sum_{i \in \mathbb{Z} \setminus D} |u_i|^p < \infty\},$$

with the usual norms $\|u\|_{l_1^\infty} = \sup_{i \in D} |u_i|$, $\|u\|_{l_2^\infty} = \sup_{i \in \mathbb{Z} \setminus D} |u_i|$, $\|u\|_{l_1^p} = \left(\sum_{i \in D} |u_i|^p \right)^{\frac{1}{p}}$, $\|u\|_{l_2^p} = \left(\sum_{i \in \mathbb{Z} \setminus D} |u_i|^p \right)^{\frac{1}{p}}$. If $D = \mathbb{Z}$, then $l_1^\infty = l^\infty$, $l_1^p = l^p$. In the case (2), the phase space is $\mathbb{R}^m$, $m = m_2 - m_1 + 1$, with its norm denoted by $\|\cdot\|_{\mathbb{R}^m}$. The scalar products in $\mathbb{R}^m$ and $l_j^2, j = 1, 2$, will be given by $(\cdot, \cdot)_{\mathbb{R}^m}$ and $(\cdot, \cdot)_{l_j^2}$, respectively.

We consider the following non-autonomous retarded problem with non-local diffusion:

$$\begin{cases} \dfrac{d}{dt} u^i(t) = [J(t, u_t)]_i - u^i(t), i \in D, t > \tau, \\ u^i(\tau + s) \equiv \phi^i(s), i \in D, s \in [-h, 0], \end{cases} \tag{6}$$

where τ is the initial moment of time, and $J : \mathbb{R}^+ \times C([-h, 0], X) \to X$, with either $X = \mathbb{R}^m$, $X = l_1^\infty$ or $X = l_1^1$, is the non-autonomous convolution operator defined by

$$[J(t, u_t)]_i = \sum_{r \in D} \int_{-h}^0 j_{i-r} u^r(t + s) \alpha_i(s) d\mu(s) + \sum_{r \in \mathbb{Z} \setminus D} \int_{-h}^0 j_{i-r} g_r(t + s) \alpha_i(s) d\mu(s), \text{ if } i \in D,$$

where $u(t) = \left(u^i(s) \right)_{i \in D}$, $u_t = u(t + s)$ for $s \in [-h, 0]$, $j : \mathbb{Z} \to \mathbb{R}$, $g : \mathbb{R} \times \mathbb{Z} \setminus D \to \mathbb{R}$ and $d\mu(s) = \xi(s)ds$ being $\xi(\cdot)$ a probability density. We assume the following conditions:

(H1) $j_k \geq 0$ for all $k \in \mathbb{Z}$.
(H2) $\sum_{i \in \mathbb{Z}} j_i = 1$.
(H3) $g \in C(\mathbb{R}, l_2^\infty)$.
(H4) $\alpha_i \in C([-h, 0], \mathbb{R})$, $\alpha_i(s) \geq 0$ for $i \in D$, $s \in [-h, 0]$.

Throughout the paper, the space $C([-h,0], X)$ will be denoted for short by C_h and its supremum norm by $\|\cdot\|_{C_h}$ when no confusion arises. All the results in this section extend analogous ones from [10] in the same problem but without delay, that is, when $h = 0$. In the case with delay we need to add conditions concerning the functions α_i (see $(H4)$, $(H5)$, $(H7)$).

2.1. A Finite Number of Equations

We assume now that D is given by (2), so $X = \mathbb{R}^m$.

Let us prove first the existence and uniqueness of solutions by applying the abstract theory for differential equations with delay [12].

Lemma 1. *Let $(H1)$–$(H4)$ hold. For any $\phi \in C([-h,0], \mathbb{R}^m)$ and the initial moment τ, there exists a unique solution $u(\cdot) \in C^1([\tau, +\infty), \mathbb{R}^m)$ of problem (6).*

Proof. We write (6) in the form

$$\frac{du}{dt} = f(u_t) + \overline{h}(t), \tag{7}$$

where $f(u_t) = (f_{m_1}(u_t), ..., f_{m_2}(u_t))^T$, $\overline{h}(t) = (h_{m_1}(t), ..., h_{m_2}(t))^T$ with

$$f_i(\phi) = \sum_{r \in D} \int_{-h}^0 j_{i-r} \phi^r(s) \alpha_i(s) d\mu(s) - \phi^i(0), \quad \phi \in C_h,$$

$$h_i(t) = \sum_{r \in \mathbb{Z} \backslash D} \int_{-h}^0 j_{i-r} g_r(t+s) \alpha_i(s) d\mu(s).$$

The map $f : C_h \to \mathbb{R}^m$ is globally Lipschitz continuous, which follows from

$$\|f(\phi_1) - f(\phi_2)\|_{\mathbb{R}^m}^2 \leq \sum_{i \in D} \left(\sum_{r \in D} \int_{-h}^0 j_{i-r} |\phi_1^r(s) - \phi_2^r(s)| \alpha_i(s) d\mu(s) + \left| \phi_1^i(0) - \phi_2^i(0) \right| \right)^2$$

$$\leq \|\phi_1 - \phi_2\|_{C_h}^2 \sum_{i \in D} \left(\int_{-h}^0 \alpha_i(s) d\mu(s) + 1 \right)^2$$

$$= C \|\phi_1 - \phi_2\|_{C_h}^2.$$

On the other hand, the map $\overline{h} : [\tau, +\infty) \to \mathbb{R}^m$ is obviously continuous as $g \in C(\mathbb{R}, l_2^\infty)$.

Therefore, the map $J : \mathbb{R} \times C_h \to \mathbb{R}^m$, $J(t, \phi) = f(\phi) + \overline{h}(t)$, is jointly continuous in both variables and Lipschitz continuous on the second variable. Hence, Theorem 2.3 in [12] implies the existence of a unique solution $u \in C([\tau - h, t_0), \mathbb{R}^m)$, defined at least for some $t_0 > \tau$.

Since f is globally Lipschitz, it has sublinear growth, that is, there are constants $C_1, C_2 > 0$ such that

$$\|f(\phi)\|_{\mathbb{R}^m} \leq C_1 + C_2 \|\phi\|_{C_h} \text{ for all } \phi \in C_h. \tag{8}$$

Thus, it takes any bounded set of C_h into a bounded set of $\mathbb{R}_h$. Since h is continuous, the map J takes closed bounded subsets of $\mathbb{R} \times C_h$ into bounded subsets of $\mathbb{R}^m$.

Let us prove that the solution exists for any $t \geq \tau$. If a solution $u(\cdot)$ were non-continuable on $[\tau - h, b)$, for any closed bounded set U in $\mathbb{R} \times C_h$ there would exists a time t_U such that $u(t) \notin U$ for any $t_U \leq t < b$ ([12], Theorem 3.2). We will obtain an estimate showing that the last is impossible for any solution.

Multiplying (7) by u and using (8) and $g \in C(\mathbb{R}, l_2^\infty)$ we have

$$\frac{1}{2}\frac{d}{dt}\|u\|_{\mathbb{R}^m}^2 \leq \left(C_1 + C_2\|u_t\|_{C_h}\right)\|u(t)\|_{\mathbb{R}^m} + \left\|\overline{h}(t)\right\|_{\mathbb{R}^m}\|u(t)\|_{\mathbb{R}^m}$$
$$\leq C_3 + C_4\|u_t\|_{C_h}^2 \text{ for all } \tau \leq t < b.$$

Integrating between τ and $t + \theta$, $\tau \leq t < b$, $\theta \in [-h, 0]$, $t + \theta \geq \tau$, we obtain

$$\|u(t+\theta)\|_{\mathbb{R}^m}^2 \leq \|u(\tau)\|_{\mathbb{R}^m}^2 + 2C_3(b-\tau) + 2C_4\int_\tau^{t+\theta}\|u_s\|_{C_h}^2 ds.$$

Taking the supremum over $\theta \in [-h, 0]$ and taking into account that $\|u(t+\theta)\|_{\mathbb{R}^m} \leq \|\phi\|_{C_h}$ if $t + \theta < \tau$, we get

$$\|u_t\|_{C_h}^2 \leq \|\phi\|_{C_h}^2 + C_5 + C_6\int_\tau^t\|u_s\|_{C_h}^2 ds.$$

Gronwall's lemma implies that

$$\|u_t\|_{C_h}^2 \leq \left(\|\phi\|_{C_h}^2 + C_5\right)e^{C_6(t-\tau)} \text{ for all } \tau \leq t < b. \tag{9}$$

Let B_0 be the closed ball in C_h centered at 0 with radius $R = 2\left(\|\phi\|_{C_h}^2 + C_5\right)e^{C_6(b-\tau)}$ and put $U = [\tau, b] \times B_0$. Then, if $u(\cdot)$ were non-continuable, there would exist t_U such that $u(t) \notin U$ for any $t_U \leq t < b$. This contradicts (9).

Finally, we prove that $u(\cdot) \in C^1([\tau, +\infty), \mathbb{R}^m)$. We have seen that $\overline{h} : [\tau, +\infty) \to \mathbb{R}^m$ is a continuous function. Additionally, $f : C_h \to \mathbb{R}^m$ is globally Lipschitz and $u \in C([-h, +\infty), \mathbb{R}^m)$, so $u_t \in C([0, +\infty), C_h)$. Hence, the application $t \mapsto f(u_t) + \overline{h}(t)$, $t \geq \tau$, is a continuous function, so using the equality (7), we obtain that $\dfrac{du}{dt}$ is also continuous on $[\tau, +\infty)$. □

Further, we will prove that the solutions are non-negative if the initial condition is non-negative. This result implies in turn a comparison principle.

For an arbitrary (real) constant v, we denote $v^+ = \max\{0, v\}$, $v^- = \max\{0, -v\}$. We observe that $v = v^+ - v^-$ and that $v \cdot v^+ = (v^+)^2$.

Proposition 1. *Let* $(H1) - (H4)$ *hold. Assume that* $g_r(t) \geq 0$ *for any* $(t, r) \in \mathbb{R} \times \mathbb{Z} \backslash D$. *Let* $\phi \in C([-h, 0], \mathbb{R}^n)$ *be such that* $\phi^i(s) \geq 0$ *for* $s \in [-h, 0]$, $i \in D$. *Then, the solution* $u(\cdot)$ *of problem* (6) *satisfies that* $u^i(t) \geq 0$ *for all* $i \in D$, $t \geq \tau$.

Proof. Multiplying the equation in (6) by $(-u)^+$ and using the equality ([13], Lemma 2.2)

$$\left(-\frac{du}{dt}, (-u)^+\right)_{\mathbb{R}^m} = \frac{1}{2}\frac{d}{dt}\left\|(-u)^+\right\|_{\mathbb{R}^m}^2$$

we have

$$\frac{1}{2}\frac{d}{dt}\left\|(-u(t))^+\right\|_{\mathbb{R}^m}^2$$
$$= \sum_{i\in D}\sum_{r\in D}\int_{-h}^0 (-u^r(t+s))(-u^i(t))^+ j_{i-r}\alpha_i(s)d\mu(s)$$
$$- \sum_{r\in\mathbb{Z}\backslash D}\int_{-h}^0 j_{i-r}g_r(t+s)(-u^i(t))^+\alpha_i(s)d\mu(s) - \left\|(-u(t))^+\right\|_{\mathbb{R}^m}^2,$$

and discarding the negative terms

$$
\frac{1}{2}\frac{d}{dt}\left\|(-u(t))^+\right\|^2_{\mathbb{R}^m}
$$

$$
\leq \sum_{i \in D}\sum_{r \in D}\int_{-h}^0 (-u^r(t+s))(-u^i(t))^+\, j_{i-r}\alpha_i(s)d\mu(s)
$$

$$
= \sum_{i \in D}\sum_{r \in D}\int_{-h}^0 ((-u^r(t+s))^+ - (-u^r(t+s))^-)(-u^i(t))^+\, j_{i-r}\alpha_i(s)d\mu(s)
$$

$$
\leq \sum_{i \in D}\sum_{r \in D}\int_{-h}^0 (-u^r(t+s))^+(-u^i(t))^+\, j_{i-r}\alpha_i(s)d\mu(s)
$$

$$
\leq \frac{1}{2}\sum_{i \in D}\sum_{r \in D}\int_{-h}^0 \left(((-u^r(t+s))^+)^2 + ((-u^i(t))^+)^2\right) j_{i-r}\alpha_i(s)d\mu(s).
$$

We deduce by $(H2)$ that

$$
\frac{1}{2}\frac{d}{dt}\left\|(-u(t))^+\right\|^2_{\mathbb{R}^m} \leq \left\|(-u_t)^+\right\|^2_{C_h}\sum_{i \in D}\int_{-h}^0 \sum_{r \in D} j_{i-r}\alpha_i(s)d\mu(s)
$$

$$
\leq M\left\|(-u_t)^+\right\|^2_{C_h},
$$

where $M = \sum_{i \in D}\int_{-h}^0 \alpha_i(s)d\mu(s)$. Integrating over (τ, t) we find

$$
\left\|(-u(t))^+\right\|^2_{\mathbb{R}^m} \leq \left\|(-u(\tau))^+\right\|^2_{\mathbb{R}^m} + 2M\int_\tau^t \left\|(-u_s)^+\right\|^2_{C_h}ds.
$$

Taking $t = t + \theta$ for $\theta \in [-h, 0]$, $t + \theta \geq \tau$, and using that $\left\|(-u(t+\theta))^+\right\|_{\mathbb{R}^m} = \left\|(-\phi(t+\theta))^+\right\|_{\mathbb{R}^m}$ if $t + \theta < \tau$ the last expression reads as

$$
\left\|(-u(t+\theta))^+\right\|^2_{\mathbb{R}^m} \leq \left\|(-\phi)^+\right\|^2_{C_h} + 2M\int_\tau^{t+\theta} \left\|(-u_s)^+\right\|^2_{C_h}ds,
$$

and taking the supremum over $[-h, 0]$,

$$
\left\|(-u_t)^+\right\|^2_{C_h} \leq \left\|(-\phi)^+\right\|^2_{C_h} + 2M\int_\tau^t \left\|(-u_s)^+\right\|^2_{C_h}ds.
$$

Finally, Gronwall's lemma gives

$$
\left\|(-u_t)^+\right\|^2_{C_h} \leq \left\|(-\phi)^+\right\|^2_{C_h}e^{2M(t-\tau)},
$$

so $u(t) \geq 0$ for all $t \geq \tau$. $\square$

Corollary 1. *Let* $(H1) - (H4)$ *hold. Additionally, let* ϕ_1, ϕ_2 *be two initial conditions such that* $\phi_1 \geq \phi_2$ *in* $C([-h, 0], \mathbb{R}^n)$, *and let* $g_1, g_2 \in C(\mathbb{R}, l_2^\infty)$ *be two functions such that* $g_1(t) \geq g_2(t)$ *for all* $t \geq \tau$. *We consider the solutions* $u(\cdot), v(\cdot)$ *of problem* (6) *with* $\phi = \phi_1, g = g_1$ *and* $\phi = \phi_2$, $g = g_2$, *respectively. Then,* $u(t) \geq v(t)$ *for all* $t \geq \tau$.

Let us consider the autonomous case, that is, the function g does not depend on t. Assume also that α does not depend on i and denote

$$
M_1 = M_1(h) = \int_{-h}^0 \alpha(s)d\mu(s).
$$

A fixed point $u \in \mathbb{R}^m$ has to satisfy the following system:

$$-M_1 \sum_{r \in D} j_{i-r} u^r + u^i = M_1 \sum_{r \in \mathbb{Z} \setminus D} j_{i-r} g_r = b_i, i \in D.$$

We can rewrite it in the matrix form

$$Au = b, \tag{10}$$

where

$$A = I - M_1 \begin{pmatrix} j_0 & j_{-1} & j_{-2} & \cdots & \cdots & j_{m_1 - m_2} \\ j_1 & j_0 & j_{-1} & j_{-2} & \cdots & j_{m_1 - m_2 + 1} \\ j_2 & j_1 & j_0 & j_{-1} & \cdots & j_{m_1 - m_2 + 2} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ j_{m_2 - m_1 - 1} & \cdots & j_2 & j_1 & j_0 & j_{-1} \\ j_{m_2 - m_1} & \cdots & \cdots & j_2 & j_1 & j_0 \end{pmatrix}.$$

We consider the following condition:

$$M_1 \sum_{r \in D} j_{i-r} < 1, \forall i \in D. \tag{11}$$

Let r_i be the sum of the elements of the row i of the matrix A. Then, (11) implies that

$$r_i = 1 - M_1 \sum_{r \in D} j_{i-r} > 0.$$

Hence, A is diagonal dominant, so that the system has a unique solution, i.e., there exists a unique fixed point $\overline{u}$.

We will show that it is stable. For this aim, we need to study the properties of the solutions of the homogeneous system

$$\begin{cases} \frac{d}{dt} v_i(t) = \sum_{r \in D} \int_{-h}^0 j_{i-r} v^r(t+s) \alpha(s) d\mu(s) - v^i(t), i \in D, t > 0, \\ v^i(s) \equiv \phi_i(s), i \in D, s \in [-h, 0]. \end{cases} \tag{12}$$

We recall the definition of stability of stationary points.

Definition 1. *The zero solution of (12) is stable in the sense of Lyapunov if for each $\epsilon > 0$ there exists $\gamma(\epsilon) > 0$ such that if $\|v(0)\|_{\mathbb{R}^m} < \gamma$, then $\|v(t)\|_{\mathbb{R}^m} < \epsilon \ \forall t \geq 0$. Additionally, we say that 0 is asymptotically stable if it is stable and there is $\gamma > 0$ such that if $\|v(0)\|_{\mathbb{R}^m} < \gamma$, then $\lim_{t \to +\infty} \|v(t)\|_{\mathbb{R}^m} = 0$. It is exponentially stable if it is stable and, moreover, there exist $\gamma, c_0, \lambda > 0$ such that if $\|v(0)\|_{\mathbb{R}^m} < \gamma$, then $\|v(t)\|_{\mathbb{R}^m} \leq c_0 e^{-\lambda t}$ for all $t \geq 0$.*

Lemma 2. *Let $(H1) - (H4)$ hold. Assume that (11) holds and that $h < -\frac{1}{2} \log(1 - \delta(h))$, where*

$$\delta(h) = \min_{i \in D} \{1 - M_1(h) \sum_{r \in D} j_{i-r}\} = \min_{r \in D} \{1 - M_1(h) \sum_{i \in D} j_{i-r}\}.$$

Then, the zero solution of problem (12) is exponentially stable (hence, asymptotically stable). Moreover, it attracts uniformly as $t \to +\infty$ any solution starting on a bounded set.

Remark 1. *Although the condition $h < -\frac{1}{2} \log(1 - \delta(h))$ is an inequation, it is easy to see that for h small enough it holds true. Indeed, we can choose first $h_0 > 0$ such that $M_1(h_0) \sum_{r \in D} j_{i-r} < 1$ for any $i \in D$, which implies that $M_1(h) \sum_{r \in D} j_{i-r} < 1$ for all $h \leq h_0$, $i \in D$. Then, we choose $h \leq h_0$ such that $h < -\frac{1}{2} \log(1 - \delta(h_0)) \leq -\frac{1}{2} \log(1 - \delta(h))$, where we have used that $\delta(h)$ is a non-increasing function.*

Proof. Let $v(t)$ be an arbitrary solution to problem (12) with $\|\phi\|_{C_h} \leq d$.

Let $0 < \lambda < 2$. Then

$$\frac{d}{dt}\left(e^{\lambda t}\|v\|_{\mathbb{R}^m}^2\right)$$

$$= \lambda e^{\lambda t}\|v(t)\|_{\mathbb{R}^m}^2 + 2e^{\lambda t}\sum_{i \in D} v_i(t)\left(\sum_{r \in D}\int_{-h}^0 j_{i-r}v^r(t+s)\alpha(s)d\mu(s) - v_i(t)\right)$$

$$= (\lambda - 2)e^{\lambda t}\|v(t)\|_{\mathbb{R}^m}^2 + 2e^{\lambda t}\sum_{i \in D} v_i(t)\sum_{r \in D}\int_{-h}^0 j_{i-r}v^r(t+s)\alpha(s)d\mu(s).$$

We estimate the last term as follows:

$$2e^{\lambda t}\sum_{i \in D} v_i(t)\sum_{r \in D}\int_{-h}^0 j_{i-r}v^r(t+s)\alpha(s)d\mu(s)$$

$$\leq 2e^{\lambda t}\left(\frac{M_1(h)}{2}\sum_{i \in D}\sum_{r \in D} j_{i-r}\left(v^i(t)\right)^2 + \frac{1}{2}\sum_{i \in D}\sum_{r \in D}\int_{-h}^0 j_{i-r}|v^r(t+\theta)|^2\alpha(s)d\mu(s)\right)$$

$$\leq 2e^{\lambda t}\left(\frac{M_1(h)}{2}\sum_{i \in D}\left(v^i(t)\right)^2\sum_{r \in D} j_{i-r} + \frac{1-\delta(h)}{2M_1(h)}\int_{-h}^0\sum_{r \in D}|v^r(t+\theta)|^2\alpha(s)d\mu(s)\right)$$

$$= (1-\delta(h))e^{\lambda t}\left(\sum_{i \in D}\left(v^i(t)\right)^2 + \sup_{\theta \in [-h,0]}\sum_{r \in D}|v^r(t+\theta)|^2\right).$$

Hence,

$$\frac{d}{dt}\left(e^{\lambda t}\|v\|_{\mathbb{R}^m}^2\right) \leq (\lambda - 2)e^{\lambda t}\|v(t)\|_{\mathbb{R}^m}^2 + 2e^{\lambda t}(1-\delta(h))\|v_t\|_{C_h}^2$$

$$\leq 2e^{\lambda t}(1-\delta(h))\|v_t\|_{C_h}^2.$$

Integrating over $(0,t)$, we have

$$e^{\lambda t}\|v(t)\|_{\mathbb{R}^m}^2 \leq \|v(0)\|_{\mathbb{R}^m}^2 + 2(1-\delta(h))\int_0^t e^{\lambda s}\|v_s\|_{C_h}^2 ds.$$

We set $t + \theta$, $t + \theta \geq 0$, instead of t and obtain that

$$\|v(t+\theta)\|_{\mathbb{R}^m}^2 \leq e^{-\lambda(t+\theta)}\|v(0)\|_{\mathbb{R}^m}^2 + 2(1-\delta(h))e^{-\lambda(t+\theta)}\int_0^{t+\theta} e^{\lambda s}\|v_s\|_{C_h}^2 ds$$

$$\leq e^{-\lambda t}e^{\lambda h}\|v(0)\|_{\mathbb{R}^m}^2 + 2(1-\delta(h))e^{-\lambda t}e^{\lambda h}\int_0^t e^{\lambda s}\|v_s\|_{C_h}^2 ds.$$

Thus, as $\|v(t+\theta)\|_{\mathbb{R}^m} \leq \|\phi\|_{C_h}^2$ for $t + \theta < 0$, we have

$$e^{\lambda t}\|v_t\|_{C_h}^2 \leq e^{\lambda h}\|\phi\|_{C_h}^2 + 2(1-\delta(h))e^{\lambda h}\int_0^t e^{\lambda s}\|v_s\|_{C_h}^2 ds.$$

By Gronwall's lemma, we get

$$e^{\lambda t}\|v_t\|_{C_h}^2 \leq e^{\lambda h}\|\phi\|_{C_h}^2 e^{2(1-\delta(h))e^{\lambda h}t},$$

so

$$\|v_t\|_{C_h}^2 \leq e^{\lambda h}\|\phi\|_{C_h}^2 e^{\left(2(1-\delta(h))e^{\lambda h} - \lambda\right)t}. \tag{13}$$

Condition $h < -\frac{1}{2}\log(1-\delta(h))$ implies that $2(1-\delta(h))e^{2h} < 2$, so there exists $0 < \lambda^* < 2$ such that $2(1-\delta(h))e^{\lambda^* h} = \lambda^*$. Therefore, choosing $\lambda \in (\lambda^*, 2)$ we obtain that $2(1-\delta(h))e^{\lambda h} - \lambda < 0$. It follows from inequality (13) that 0 is exponentially stable and that it attracts uniformly as $t \to +\infty$ any solution starting on a bounded set. $\quad\square$

Corollary 2. *Let* $(H1)$–$(H4)$ *hold. We consider equation* (6) *in the autonomous case, that is, the function g does not depend on t and* $\tau = 0$. *Assume that* (11) *holds and that* $h < -\frac{1}{2}\log(1 - \delta(h))$. *Then, the fixed point* $\bar{u}$ *is exponentially stable for the autonomous problem* (6). *Moreover, it uniformly attracts every solution starting on a bounded set.*

We finish this subsection by showing that symmetry of the initial condition and the functions g, α and j imply symmetry of the solution.

Lemma 3. *Let* $(H1)$–$(H4)$ *hold. Let* $D = \{-d, -d+1, ..., 0, ..., d-1, d\}$, $d \in \mathbb{N}$. *Assume also that* $g_i(t) = g_{-i}(t)$, *for all* $i \in \mathbb{Z} \backslash D$, $t \geq \tau$, $j_i = j_{-i}$, $\alpha_i(s) = \alpha_{-i}(s)$, $\phi^i(s) = \phi^{-i}(s)$ *for any* $i \in \mathbb{Z}$, $s \in [-h, 0]$. *Then, the unique solution to problem* (6) *is symmetric, that is,*

$$u^i(t) = u^{-i}(t) \text{ for all } i \in D, t \geq \tau - h.$$

Proof. Let $w^i(t) = u^{-i}(t)$. Then, the hypothesis of the lemma implies that

$$
\begin{aligned}
\frac{dw^i}{dt} &= \frac{du^{-i}}{dt} \\
&= \sum_{r \in D} \int_{-h}^{0} j_{-i-r} u^r(t+s) \alpha_{-i}(s) d\mu(s) + \sum_{r \in \mathbb{Z} \backslash D} \int_{-h}^{0} j_{-i-r} g_r(t+s) \alpha_{-i}(s) d\mu(s) - u^{-i}(t) \\
&= \sum_{r \in D} \int_{-h}^{0} j_{i+r} w^{-r}(t+s) \alpha_i(s) d\mu(s) + \sum_{r \in \mathbb{Z} \backslash D} \int_{-h}^{0} j_{i+r} g_{-r}(t+s) \alpha_i(s) d\mu(s) - w^i(t) \\
&= \sum_{r \in D} \int_{-h}^{0} j_{i-\bar{r}} w^{\bar{r}}(t+s) \alpha_i(s) d\mu(s) + \sum_{r \in \mathbb{Z} \backslash D} \int_{-h}^{0} j_{i-\bar{r}} g_{\bar{r}}(t+s) \alpha_i(s) d\mu(s) - w^i(t)
\end{aligned}
$$

and $w^i(\tau + s) = u^{-i}(\tau + s) = u^i(\tau + s) = \phi^i(s)$ for $s \in [-h, 0]$. It follows that $w(\cdot)$ is the unique solution to problem (6) with initial condition ϕ. Hence, $w^i(t) = u^i(t) = u^{-i}(t)$ for all $i \in D$ and $t \geq \tau - h$. $\square$

2.2. An Infinite Number of Equations

We consider now the case where the system can admit an infinite number of equations. For this we define D as a subset of $\mathbb{Z}$ of one of the types given in (3)–(5).

First, we study the Equation (6) with $X = l_1^{\infty}$, so $J : \mathbb{R} \times l_1^{\infty} \to l_1^{\infty}$.

We will need the following extra assumption:

$(H5) \sup_{i \in D} \int_{-h}^{0} \alpha_i(s) d\mu(s) < \infty.$

Remark 2. *Conceptually, in life tables, it can be of interest to use an infinite number of equations. For example, domains of type* (5) *can be used for the biometric function* Life Expectancy, *which is defined (at year t) as* $\sum_{i=0}^{\infty} p_i(t) \cdot i$, *where* $p_i(t)$ *represents the survival probability at age i at year t. We note that in this work we use the function* $q_i = (1 - p_i)$, *where* q_i *represents the risk of death at age i. In Section* 3.1.2, *we apply model* (6) *with an infinite number of equations to the variable* q_i.

In order to obtain the existence and uniqueness of solutions we use once again the abstract theory for differential equations with delay given in ([12], Chapter 2). Although the results are stated in the phase space $\mathbb{R}^n$, the theorems are valid for a Banach space as well.

Lemma 4. *Assume that, in addition to* $(H1)$–$(H4)$, $(H5)$ *is satisfied. Then, for any* $\phi \in C([-h, 0], l_1^{\infty})$ *there exists a unique solution* $u(\cdot) \in C^1([\tau, +\infty), l_1^{\infty})$ *of problem* (6).

Proof. We proceed as in Lemma 1. System (6) is considered in the form (7), where in this case $(f_i(u_t))_{i \in D}$ and $(h_i(t))_{i \in D}$ are column vectors with infinite elements.

We prove first that the map $f : l_1^\infty \to l_1^\infty$ is globally Lipschitz continuous. Indeed, by (H2) and (H5), we have

$$\|f(\phi_1) - f(\phi_2)\|_{l_1^\infty}^2$$

$$\leq \sup_{i \in D} \left(\sum_{r \in D} \int_{-h}^0 j_{i-r} |\phi_1^r(s) - \phi_2^r(s)| \alpha_i(s) d\mu(s) + \left| \phi_1^i(0) - \phi_2^i(0) \right| \right)$$

$$\leq \sup_{s \in [-h,0],\, i \in D} \left| \phi_1^i(s) - \phi_2^i(s) \right| \sup_{i \in D} \left(\sum_{r \in D} \int_{-h}^0 j_{i-r} \alpha_i(s) d\mu(s) + 1 \right)$$

$$\leq \|\phi_1 - \phi_2\|_{C([-h,0],l_1^\infty)} \left(\sup_{i \in D} \sum_{r \in D} j_{i-r} \int_{-h}^0 \alpha_i(s) d\mu(s) + 1 \right)$$

$$\leq \|\phi_1 - \phi_2\|_{C([-h,0],l_1^\infty)} \left(\sup_{i \in D} \int_{-h}^0 \alpha_i(s) d\mu(s) + 1 \right).$$

Second, we check that $\overline{h}(t)$ is a continuous map for the cases (4) and (5). If $t_n \to t$ in $\mathbb{R}$, then by $(H3)$ we obtain that

$$\left\| \overline{h}(t_n) - \overline{h}(t) \right\|_{l_1^\infty}$$

$$= \sup_{i \in D} \left| \sum_{r \in \mathbb{Z} \setminus D} \int_{-h}^0 j_{i-r} g_r(t_n + s) \alpha_i(s) d\mu(s) - \sum_{r \in \mathbb{Z} \setminus D} \int_{-h}^0 j_{i-r} g_r(t + s) \alpha_i(s) d\mu(s) \right|$$

$$\leq \sup_{i \in D} \sum_{r \in \mathbb{Z} \setminus D} \int_{-h}^0 j_{i-r} |g_r(t_n + s) - g_r(t + s)| \alpha_i(s) d\mu(s)$$

$$\leq \sup_{s \in [-h,0],\, i \in D} |g_i(t_n + s) - g_i(t + s)| \sup_{i \in D} \sum_{r \in \mathbb{Z} \setminus D} j_{i-r} \int_{-h}^0 \alpha_i(s) d\mu(s)$$

$$\leq \|g(t_n) - g(t)\|_{C([-h,0],l_2^\infty)} \sup_{i \in D} \int_{-h}^0 \alpha_i(s) d\mu(s) \to 0 \text{ as } n \to \infty.$$

In order to prove that the solution is globally defined for $t \geq \tau$, we need to obtain an estimate of its growth. Let $b > \tau$ be the maximal time of existence of the solution $u(\cdot)$, so it is not continuable in $[\tau - h, b)$. Since f is Lipschitz, it has sublinear growth, and $h(\cdot)$ is bounded in $[\tau - h, b]$. Hence,

$$\left\| \frac{du}{dt} \right\|_{l_1^\infty} \leq K_1 + K_2 \|u_t\|_{C([-h,0],l_1^\infty)}, \text{ for } t \in [\tau - h, b),$$

so

$$\|u(t)\|_{l_1^\infty} \leq \|u(\tau)\|_{l_1^\infty} + \int_\tau^t \left\| \frac{du}{dr} \right\|_{l_1^\infty} dr$$

$$\leq \|u(\tau)\|_{l_1^\infty} + K_1(b - \tau) + K_2 \int_\tau^t \|u_r\|_{C([-h,0],l_1^\infty)} dr,$$

for $t \in [\tau - h, b)$. Arguing as in the proof of Lemma 1 we deduce that

$$\|u_t\|_{C([-h,0],l_1^\infty)} \leq \left(\|\phi\|_{C([-h,0],l_1^\infty)} + K_1(b - \tau) \right) e^{K_2(t-\tau)}.$$

Again by the same arguments of the proof of Lemma 1, we arrive at a contradiction. $\square$

Further, we will obtain a comparison principle as in Proposition 1. For this aim, we need two additional assumptions:

$(H6)$ There exist $K > 0$ and $a \in l_1^2$ such that $a_i > 0$, for all $i \in D$, and

$$\sum_{i \in D} \frac{a_i^2}{a_r^2} j_{i-r} \leq K \text{ for every } r \in D.$$

$(H7)$ There exists $\alpha(\cdot) \in C([-h,0], \mathbb{R})$ such that $\alpha_i(s) \leq \alpha(s)$ for all $s \in [-h, 0]$.

It is obvious that $(H7)$ is stronger than $(H5)$. Condition (H6) is satisfied for any kernel j_i with subexponential growth, that is, $j_i \leq ce^{-\alpha|i|}$ for some $\alpha, c > 0$ [10].

Proposition 2. *Let* $(H1) - (H4)$, $(H6) - (H7)$ *hold. Assume that* $g_r(t) \geq 0$ *for any* $(t, r) \in \mathbb{R} \times \mathbb{Z} \backslash D$. *Let* $\phi \in C([-h, 0], l_1^\infty)$ *be such that* $\phi^i(s) \geq 0$ *for* $s \in [-h, 0]$ *and* $i \in D$. *Then,* $u^i(t) \geq 0$ *for all* $i \in D, t \geq \tau$.

Proof. We define the function $v(t) = (v^i(t))_{i \in D} = (a_i u^i(t))_{i \in D}$, which belongs to $C^1([\tau, \infty), l_1^2)$ as $a \in l_1^2$. We know by ([13], Lemma 2.2) that

$$\left(-\frac{dv}{dt}, (-v)^+ \right)_{l_1^2} = \frac{1}{2} \frac{d}{dt} \left\| (-v)^+ \right\|_{l_1^2}^2.$$

The function v satisfies the equality

$$-\frac{dv_i}{dt} = -a_i \sum_{r \in D} \int_{-h}^0 j_{i-r} u^r(t+s) \alpha_i(s) d\mu(s) - a_i \sum_{r \in \mathbb{Z} \backslash D} \int_{-h}^0 j_{i-r} g_r(t+s) \alpha_i(s) d\mu(s) + v_i(t).$$

Multiplying it by $(-v)^+$, we have

$$\frac{1}{2} \frac{d}{dt} \left\| (-v(t))^+ \right\|_{l_1^2}^2$$

$$= \sum_{i \in D} \sum_{r \in D} \int_{-h}^0 (-u^r(t+s))(-v^i(t))^+ a_i \, j_{i-r} \alpha_i(s) d\mu(s)$$

$$- \sum_{r \in \mathbb{Z} \backslash D} \int_{-h}^0 a_i j_{i-r} g_r(t+s)(-v^i(t))^+ \alpha_i(s) d\mu(s) - \left\| (-v(t))^+ \right\|^2$$

$$\leq \sum_{i \in D} \sum_{r \in D} \int_{-h}^0 (-u^r(t+s))^+ (-v^i(t))^+ a_i \, j_{i-r} \alpha_i(s) d\mu(s)$$

$$\leq \frac{1}{2} \sum_{i \in D} \sum_{r \in D} \int_{-h}^0 ((-v^r(t+s))^+)^2 \frac{a_i^2}{a_r^2} \, j_{i-r} \alpha_i(s) d\mu(s)$$

$$+ \frac{1}{2} \sum_{i \in D} \sum_{r \in D} \int_{-h}^0 ((-v^i(t))^+)^2 j_{i-r} \alpha_i(s) d\mu(s).$$

Making use of (H7), the last term is estimated by

$$\frac{1}{2} \sum_{i \in D} \sum_{r \in D} \int_{-h}^0 ((-v^i(t))^+)^2 j_{i-r} \alpha_i(s) d\mu(s)$$

$$\leq \frac{1}{2} \sum_{i \in D} ((-v^i(t))^+)^2 \int_{-h}^0 \alpha_i(s) d\mu(s) \tag{14}$$

$$\leq \frac{1}{2} \int_{-h}^0 \alpha(s) d\mu(s) \left\| (-v(t))^+ \right\|_{l_1^2}^2.$$

By (H6) and (H7), for the penultimate term, we obtain

$$
\begin{aligned}
\frac{1}{2} &\sum_{i \in D} \sum_{r \in D} \int_{-h}^{0} \left((-v^r(t+s))^+ \right)^2 \frac{a_i^2}{a_r^2} \, j_{i-r} \alpha_i(s) d\mu(s) \\
&\leq \frac{1}{2} \sum_{r \in D} \sum_{i \in D} \frac{a_i^2}{a_r^2} \, j_{i-r} \int_{-h}^{0} \left((-v^r(t+s))^+ \right)^2 \alpha(s) d\mu(s) \\
&\leq \frac{K}{2} \int_{-h}^{0} \sum_{r \in D} \left((-v^r(t+s))^+ \right)^2 \alpha(s) d\mu(s) \\
&\leq \frac{K}{2} \left\| (-v_t)^+ \right\|_{C([-h,0], l_1^2)}^2 \int_{-h}^{0} \alpha(s) d\mu(s).
\end{aligned}
\tag{15}
$$

Since $\sum_{i \in D} \frac{a_i^2}{a_r^2} j_{i-r} \int_{-h}^{0} \left((-v^r(t+s))^+ \right)^2 \alpha(s) d\mu(s)$ is convergent for each $r \in D$ and the iterated limit

$$
\sum_{r \in D} \sum_{i \in D} \frac{a_i^2}{a_r^2} \, j_{i-r} \int_{-h}^{0} \left((-v^r(t+s))^+ \right)^2 \alpha(s) d\mu(s)
$$

is convergent as well, the change of order in the summatories is justified. Furthermore, $v_t \in C\left([-h,0], l_1^2\right)$ implies that $S_k = \sum_{r \in D, \, |r| \leq k} \left((-v^r(t+s))^+ \right)^2 \alpha(s)$ converges to $\sum_{r \in D} \left((-v^r(t+s))^+ \right)^2 \alpha(s)$ uniformly in $s \in [-h,0]$, so it is correct to change the order of the summatory and the integral in the penultimate inequality of (15).

Thus, there exists a constant $M > 0$ satisfying

$$
\left\| (-v(t))^+ \right\|_{l_1^2}^2 \leq \left\| (-v(\tau))^+ \right\|_{l_1^2}^2 + M \int_{\tau}^{t} \left\| (-v_s)^+ \right\|_{C([-h,0], l_1^2)}^2 ds.
$$

Arguing as in the proof of Proposition 1 it follows that $u(t) \geq 0$ for all $t \geq \tau$. $\qquad\square$

Corollary 3. *Assume the conditions of Proposition 2. Let ϕ_1, ϕ_2 be two initial conditions such that $\phi_1 \geq \phi_2$ in $C\left([-h,0], l_1^\infty\right)$, and let $g_1, g_2 \in C(\mathbb{R}, l_2^\infty)$ be two functions such that $g_1(t) \geq g_2(t)$ for all $t \geq \tau$. We consider the solutions $u(\cdot), v(\cdot)$ of the problem (6) with $\phi = \phi_1, g = g_1$ and $\phi = \phi_2, g = g_2$, respectively. Then, $u(t) \geq v(t)$ for all $t \geq \tau$.*

Second, we study the equation (6) with $X = l_1^1$, so $J : \mathbb{R} \times l_1^1 \to l_1^1$ if we replace condition (H3) by

$(H3b)\, g \in C(\mathbb{R}, l_2^1)$.

Lemma 5. *Assume that conditions (H1), (H2), (H3b), (H4) and (H7) are satisfied. Then, for any $\phi \in C\left([-h,0], l_1^1\right)$ there exists a unique solution $u(\cdot) \in C^1\left([\tau, +\infty), l_1^1\right)$ of problem (6).*

Proof. As in Lemma 4, we only need to check that $f : l_1^1 \to l_1^1$ is globally Lipschitz continuous and that $\bar{h} : \mathbb{R} \to l_1^1$ is a continuous map for the cases (4) and (5).

By (H2) and (H7), we have

$$\|f(\phi_1) - f(\phi_2)\|^2_{l^1_1}$$

$$\leq \sum_{i \in D} \left(\sum_{r \in D} \int_{-h}^0 j_{i-r} |\phi_1^r(s) - \phi_2^r(s)| \alpha_i(s) d\mu(s) + \left| \phi_1^i(0) - \phi_2^i(0) \right| \right)$$

$$\leq \sum_{r \in D} \sum_{i \in D} j_{i-r} \int_{-h}^0 |\phi_1^r(s) - \phi_2^r(s)| \alpha(s) d\mu(s) + \|\phi_1(0) - \phi_2(0)\|_{l^1_1}$$

$$\leq \int_{-h}^0 \sum_{r \in D} |\phi_1^r(s) - \phi_2^r(s)| \alpha(s) d\mu(s) + \|\phi_1(0) - \phi_2(0)\|_{l^1_1}$$

$$\leq \|\phi_1 - \phi_2\|_{C([-h,0],l^1_1)} \left(\int_{-h}^0 \alpha(s) d\mu(s) + 1 \right).$$

If $t_n \to t$ in $\mathbb{R}$, then by (H3b) and (H7), we obtain that

$$\left\| \overline{h}(t_n) - \overline{h}(t) \right\|_{l^1_1}$$

$$= \sum_{i \in D} \left| \sum_{r \in \mathbb{Z} \setminus D} \int_{-h}^0 j_{i-r} g_r(t_n + s) \alpha_i(s) d\mu(s) - \sum_{r \in \mathbb{Z} \setminus D} \int_{-h}^0 j_{i-r} g_r(t + s) \alpha_i(s) d\mu(s) \right|$$

$$\leq \sum_{r \in \mathbb{Z} \setminus D} \sum_{i \in D} j_{i-r} \int_{-h}^0 |g_r(t_n + s) - g_r(t + s)| \alpha(s) d\mu(s)$$

$$\leq \int_{-h}^0 \sum_{r \in \mathbb{Z} \setminus D} |g_r(t_n + s) - g_r(t + s)| \alpha(s) d\mu(s)$$

$$\leq \|g(t_n) - g(t)\|_{C([-h,0],l^1_2)} \int_{-h}^0 \alpha(s) d\mu(s) \to 0 \text{ as } n \to \infty.$$

The change of the order in the summatories is justified in the same way as in Lemma 2. On the other hand, since the functions $l_1(\cdot) = \{|g_r(t_n + \cdot) - g_r(t + \cdot)| \alpha(s)\}_{r \in D}$, $l_2(\cdot) = \{|\phi_1^r(s) - \phi_2^r(s)| \alpha(s)\}_{r \in D}$ belong to $C([-h, 0], l^1_2)$ and $C([-h, 0], l^1_1)$ respectively, by a standard compactness argument, we obtain that the sequences $S_k^1 = \sum_{r \in \mathbb{Z} \setminus D, |r| > k} |g_r(t_n + s) - g_r(t + s)| \alpha(s)$ converge uniformly in $s \in [-s, 0]$ to $\sum_{r \in \mathbb{Z} \setminus D} |g_r(t_n + s) - g_r(t + s)| \alpha(s)$, as $k \to \infty$, and the sequence $S_k^2 = \sum_{r \in \mathbb{Z} \setminus D, |r| > k} |\phi_1^r(s) - \phi_2^r(s)| \alpha(s)$ converges uniformly to $\sum_{r \in \mathbb{Z} \setminus D} |\phi_1^r(s) - \phi_2^r(s)| \alpha(s)$. Thus, in both cases, the limit function is integrable and it is legitimate to change the order of the summatory and the integral. $\square$

We study the comparison principle also in this case.

Proposition 3. *Suppose that conditions (H1), (H2), (H3b), (H4) and (H7) are satisfied. Assume that $g_r(t) \geq 0$ for any $(t, r) \in \mathbb{R} \times \mathbb{Z} \setminus D$. Let $\phi \in C([-h, 0], l^1_1)$ be such that $\phi_i(s) \geq 0$ for $s \in [-h, 0]$ and $i \in D$. Then, $u^i(t) \geq 0$ for all $i \in D$, $t \geq \tau$.*

Proof. As $u \in C^1\left([\tau, +\infty), l_1^1\right) \subset C^1\left([\tau, +\infty), l_1^2\right)$, we obtain arguing as in Proposition 2 that

$$
\frac{1}{2}\frac{d}{dt}\left\|(-u(t))^+\right\|_{l_1^2}^2
$$

$$
\leq \sum_{i \in D}\sum_{r \in D}\int_{-h}^0 (-u^r(t+s))^+(-u^i(t))^+ j_{i-r}\alpha_i(s)d\mu(s)
$$

$$
\leq \frac{1}{2}\sum_{i \in D}\sum_{r \in D}\int_{-h}^0 \left((-u^r(t+s))^+\right)^2 j_{i-r}\alpha_i(s)d\mu(s)
$$

$$
+ \frac{1}{2}\sum_{i \in D}\sum_{r \in D}\int_{-h}^0 \left((-u^i(t))^+\right)^2 j_{i-r}\alpha_i(s)d\mu(s).
$$

Condition (H7) implies as in (14) and (15) that

$$
\frac{1}{2}\sum_{i \in D}\sum_{r \in D}\int_{-h}^0 \left((-u^i(t))^+\right)^2 j_{i-r}\alpha_i(s)d\mu(s) \leq \frac{1}{2}\int_{-h}^0 \alpha(s)d\mu(s)\|u(t)\|_{l_1^2}^2,
$$

$$
\frac{1}{2}\sum_{i \in D}\sum_{r \in D}\int_{-h}^0 \left((-u^r(t+s))^+\right)^2 j_{i-r}\alpha_i(s)d\mu(s)
$$

$$
\leq \frac{1}{2}\sum_{r \in D}\sum_{i \in D} j_{i-r}\int_{-h}^0 \left((-u^r(t+s))^+\right)^2\alpha(s)d\mu(s)
$$

$$
\leq \frac{1}{2}\int_{-h}^0 \sum_{r \in D} \left((-u^r(t+s))^+\right)^2\alpha(s)d\mu(s)
$$

$$
\leq \frac{1}{2}\left\|(-u_t)^+\right\|_{C_h}^2 \int_{-h}^0 \alpha(s)d\mu(s).
$$

The proof is finished in the same way as in Proposition 2. $\quad\square$

Corollary 4. *Assume the conditions of Proposition 3. Let ϕ_1, ϕ_2 be two initial conditions such that $\phi_1 \geq \phi_2$ in $C\left([-h,0], l_1^1\right)$, and let $g_1, g_2 \in C(\mathbb{R}, l_2^\infty)$ be two functions such that $g_1(t) \geq g_2(t)$ for all $t \geq \tau$. We consider the solutions $u(\cdot), v(\cdot)$ of the problem (6) with $\phi = \phi_1, g = g_1$ and $\phi = \phi_2, g = g_2$, respectively. Then, $u(t) \geq v(t)$ for all $t \geq \tau$.*

We finish this section by showing that if $D = \mathbb{Z}$ and the initial condition and the functions j_i, α_i and the initial condition are symmetric, then the solution is symmetric well.

Lemma 6. *Let either the conditions of Lemma 4 or 5 hold and let $D = \mathbb{Z}$. Assume that $j_k = j_{-k}, \phi^k(s) = \phi^{-k}(s), \alpha_k(s) = \alpha_{-k}(s)$ for all $k \in D, s \in [-h, 0]$. Then, the unique solution given in Lemmas 4 and 5 satisfies:*

$$
u^k(t) = u^{-k}(t) \text{ for all } k \in D, \ t \geq \tau - h.
$$

Proof. The proof follows the same lines as in Lemma 3 except that the term containing the function g is now missing. $\quad\square$

3. Application to Life Tables

Life tables are a widely used tool in demographic sciences or in actuarial field to study the behavior of a population in relation to the survival, or mortality, in a region or in a period of time. The implications of the results derived from them have a great importance, for example, for the analysis of the sustainability of the pension system; for the comparison of populations via indicators as expectation of life; for the estimation of the insurance premium; or for the calculation of the mathematical reserves to guarantee the pay of the claims [14].

Life tables are structured in biometric functions and the relations between them (see, for example, [15–17]). Some of these important functions, among others, can be: the survival probability at exact age x, p_x, or its complement, the probability of death, q_x (the focus of this work); the expectation of life at concrete age x, e_x; the central rate of mortality by age, m_x, etc. Usually, we do not know the true values of the biometric functions and, then, they are treated as underlying values, which are unknown. In this situation, these values can be approximated using techniques or methodologies from this field of knowledge (see, [18], https://www.mortality.org/Public/Docs/MethodsProtocol.pdf or [19], among others). In this work, as in our previous paper [10], the aim is to estimate and predict the true probabilities of death (at each age x). In general, in order to estimate the values of a biometric function, different points of view and several types of techniques can be used. We consider three important paradigms: a first classification between the stochastic and the non-stochastic paradigms; second, we distinguish the situations when the model is based on the parametric point of view or, on the contrary, it does not consider parameters; finally, in the third level, we differentiate between the dynamic point of view (if the model considers the evolution with time) or the static point of view (the changes with time are not considered). The difference between the stochastic and non-stochastic frameworks is based, in actuarial practice, on the use or the interpretation of the results of the technique. For example, the Gompertz law [20] can be used in a deterministic framework (the results of the fit of the Gompertz law are used as unique values and as input in other steps of the process as, for example, in tarification) or it can be used in a stochastic framework, for example, by considering the fit as a realization of the random process of interest. The literature on this topic is extensive. The papers [11,15] contain a good introduction to the stochastic framework. The works of [21,22] introduce us to parametric and non parametric models. The works [16,23–26] allow us to know examples of stochastic, parametric and statics models. Concerning stochastic, parametric and dynamical models, the works [27–31] are interesting. We can find examples of stochastic, nonparametric and static models in [32–35]; for non-parametric and dynamical models, see [36].

As in the previous work [10], in this paper, we consider a non-stochastic paradigm, with a non-parametric model and a dynamical point of view. However, in this work, we consider the previous information as well. Thus, the difference between these two works is the use of the past information, via improvement rates. In the paper [10], history is not considered, and only the current (actual) data are used to estimate the predictions for the future. That procedure has the important limitation that the predictions are only valid for a short term, for periods of two or three years. The technique that is proposed in this work avoids this limitation and, by considering the history of the mortality rates, improves the previous results because it allows us to make predictions for periods of five to ten years.

In order to introduce the previous information, we establish a delay pattern via improvement rates [11].

3.1. Dynamical Kernel Graduation with Delay

In order to improve the dynamical model (1), which was proposed in [10], in this work, we introduce the history using the concept of Improvement Rate (see [37–40]). These rates characterize the evolution of the mortality year-to-year or between two arbitrary moments, t_0 and t_1. In this sense, we use the classical definition of improvement rate between t_0 and t_1 ($t_0 < t_1$) at age x, given by $r_x^{t_0,t_1} = q_x^{t_1}/q_x^{t_0}$. If we consider as delay the value $d = t_1 - t_0 > 0$, then, $r_x^{t_0,t_1} \equiv r_x^{t_0,t_0+d}$. Similarly, we define the global improvement rate between t_0 and t_1, $\bar{r}^{t_0,t_1}$, as the coefficient of the linear model without constant term of the observed death rates, that is, we fit the linear model

$$q_{\cdot}^{t_1} = \bar{r}^{t_0,t_1} \cdot q_{\cdot}^{t_0}$$

with the data $\left\{ q_0^{t_0}, \ldots, q_\omega^{t_0} \right\}$ and $\left\{ q_0^{t_1}, \ldots, q_\omega^{t_1} \right\}$.

3.1.1. Procedure by Steps

The proposed method can be summarized as follows:

1. We consider the observed mortality rates at each age (from 0 to 100) and each year in the considered period (from 1908 to 2019), and denote them by q_x^t, $x \in D = \{0, \dots, 100\}$, $t \in \{1908, \dots, 2019\}$. Additionally, we consider the values of $g_x(t)$, which is the rate of death either at "negative ages" or after the *actuarial infinite* (we have chosen it to be equal to 100).

2. We estimate the improvement rates by age and delays, that is:

$\frac{year}{delay}$	1	2	3	4	...		
	$\bar{r}_{1908}^{1909} \sim \frac{q_{\cdot}^{1909}}{q_{\cdot}^{1908}}$	$\bar{r}_{1908}^{1910} \sim \frac{q_{\cdot}^{1910}}{q_{\cdot}^{1908}}$	$\bar{r}_{1908}^{1911} \sim \frac{q_{\cdot}^{1911}}{q_{\cdot}^{1908}}$	$\bar{r}_{1908}^{1912} \sim \frac{q_{\cdot}^{1912}}{q_{\cdot}^{1908}}$	...	$\bar{r}_{1908}^{2017} \sim \frac{q_{\cdot}^{2017}}{q_{\cdot}^{1908}}$	$\bar{r}_{1908}^{2018} \sim \frac{q_{\cdot}^{2018}}{q_{\cdot}^{1908}}$
	$\bar{r}_{1909}^{1910} \sim \frac{q_{\cdot}^{1910}}{q_{\cdot}^{1909}}$	$\bar{r}_{1909}^{1911} \sim \frac{q_{\cdot}^{1911}}{q_{\cdot}^{1909}}$	$\bar{r}_{1909}^{1912} \sim \frac{q_{\cdot}^{1912}}{q_{\cdot}^{1909}}$	$\bar{r}_{1909}^{1913} \sim \frac{q_{\cdot}^{1913}}{q_{\cdot}^{1909}}$	...	$\bar{r}_{1909}^{2017} \sim \frac{q_{\cdot}^{2017}}{q_{\cdot}^{1909}}$	—
	$\bar{r}_{1910}^{1911} \sim \frac{q_{\cdot}^{1911}}{q_{\cdot}^{1910}}$	$\bar{r}_{1910}^{1912} \sim \frac{q_{\cdot}^{1912}}{q_{\cdot}^{1910}}$	$\bar{r}_{1910}^{1913} \sim \frac{q_{\cdot}^{1913}}{q_{\cdot}^{1910}}$	$\bar{r}_{1910}^{1914} \sim \frac{q_{\cdot}^{1914}}{q_{\cdot}^{1910}}$	...	—	—
	$\bar{r}_{1911}^{1912} \sim \frac{q_{\cdot}^{1912}}{q_{\cdot}^{1911}}$	$\bar{r}_{1911}^{1913} \sim \frac{q_{\cdot}^{1913}}{q_{\cdot}^{1911}}$	$\bar{r}_{1911}^{1914} \sim \frac{q_{\cdot}^{1914}}{q_{\cdot}^{1911}}$	$\bar{r}_{1911}^{1915} \sim \frac{q_{\cdot}^{1915}}{q_{\cdot}^{1911}}$	...	—	—
	$\vdots$	$\vdots$	...	—	—	—	—
	$\bar{r}_{2016}^{2017} \sim \frac{q_{\cdot}^{2017}}{q_{\cdot}^{2016}}$	$\bar{r}_{2016}^{2018} \sim \frac{q_{\cdot}^{2018}}{q_{\cdot}^{2016}}$	—	—	—	—	—
	$\bar{r}_{2017}^{2018} \sim \frac{q_{\cdot}^{2018}}{q_{\cdot}^{2017}}$	—	—	—	—	—	—

3. We estimate the mean, with respect to the delay, of the improvement rates:

$$\bar{\bar{r}}^d = \frac{1}{\#(d)} \sum_{t_0, t_1: d = |t_1 - t_0|} \bar{r}_{t_0}^{t_1}, \tag{16}$$

being $\#(d)$ the number of couples (t_0, t_1) such that $d = |t_1 - t_0|$. We will refer to these values as the *global improvement rates for the delay d*.

4. The annual improvement rates by delay, $\bar{\bar{r}}^d$, play an important role in the procedure because they contain previous information of the mortality process. However, the experience of studying this phenomenon allows us to assure that the importance of these rates are not the same for all delays, and it is clear (under usual conditions such as non-pandemic ones) that the behavior of the improvement rates become more important if they are close to the time of prediction. Using this realistic assumption, we assume that the importance of these rates is modulated by a probability distribution function. In particular, we consider a modified exponential function. To do this, we consider the exponential probability density function with the form

$$f_\lambda(s) = \begin{cases} \lambda e^{-\lambda s} & \text{if } s \geq 0, \\ 0 & \text{if } s < 0, \end{cases}$$

and obtain a density function in the interval $[0, h]$ by putting

$$\widehat{f}_\lambda(s) = \frac{\lambda e^{-\lambda s}}{\int_0^h \lambda e^{-\lambda s} ds}, \quad s \in [0, h]. \tag{17}$$

Then, we define the density function $\xi(\cdot)$ from (6) by $\xi(s) = \widehat{f}_\lambda(-s)$ for $s \in [-h, 0]$. In the particular case when in the numerical approximations we consider only integer delays, we can discretize the interval $[0, h]$ by using a finite number of integer delays $s = \{0, 1, \dots, d_{\max}\}$, where $d_{\max} = h \in \mathbb{N}$ is the maximum delay to be considered. Thus, instead of (17), we will use the discretized probability function

$$f^*(s) = \frac{\widehat{f}_\lambda(s)}{\sum_{s=0}^{d_{\max}} \widehat{f}_\lambda(s)}, \tag{18}$$

which approximates (17) at $s = 0, 1, ..., d_{\max}$.

Remark 3. *We could also use the values of the function $\widehat{f}_\lambda(s)$ at $s = 0, 1, ..., d_{\max}$ to define the approximative function $f^*(s)$, but we have used (18) because the results in the numerical simulations were quite good. The use of the exponential function to derive f^* is arbitrary. Thus, it is reasonable to think that the selected function is not optimal. In this sense, we consider that this topic could be a possible new line of investigation to improve the work.*

5.   Using the annual improvement rates, and the exponential distribution, we define the *weighted improvement rates* as

$$R^d_{weight} = \overline{\overline{r}}^d \cdot f^*(d) \tag{19}$$

6.   Using the improvement rates by delay, $\overline{\overline{r}}^d$, we can define the functions $\alpha(s)$ from (6), by putting

$$\alpha(-d) = \overline{\overline{r}}^d \text{ for } d \in 0, 1, ..., d_{\max},$$

where $\overline{\overline{r}}^0 = 1$, and then defininig $\alpha(s)$ at other points by linear interpolation, that is,

$$\alpha(s) = \alpha(-i-1) + (\alpha(-i) - \alpha(-i-1))(s+i+1) \text{ for } s \in (-i-1, -i), \ i = 0, ..., d_{\max} - 1. \tag{20}$$

This procedure can be implemented for a non-integer step as well. Namely, let b be a divisor of $d_{\max}$. Then, we define in a similar way as above the improvement rates $\overline{\overline{r}}^d$ for $d = b, 2b, ..., d_{\max}$. We observe that in this particular situation, the functions $\alpha(s)$ are independent of $i \in D$.

It is usual to assume some hypothesis in the behavior of the mortality. Namely,

H1.   For each age x, for an arbitrary moment t and for a small time step $\tau > 0$, the graduate value at $t + \tau$, denoted by $\overset{\circ}{q}_x(t + \tau)$, depends on:

*   all graduate values at moment t, $\overset{\circ}{q}_z(t)$, $z \in D \subset \mathbb{Z}$ (via Gaussian kernel graduation, see [17]).
*   all previous moments of time $t + s$, $s \in [-h, 0]$ (via the exponential probability density function and the improvement rates).

H2.   The relation between $\overset{\circ}{q}_x(t + \tau)$ and $\left\{ \overset{\circ}{q}_z(t) \right\}$, $\left\{ \overset{\circ}{q}_z(t + s) \right\}_{s \in [-h, 0[}$ is:

$$\overset{\circ}{q}_x(t + \tau) = \frac{1}{1+\tau} \overset{\circ}{q}_x(t) + \frac{\tau}{1+\tau} \sum_{z \in D} \int_{-h}^0 j_{x-z} \overset{\circ}{q}_z(t+s)\alpha(s)d\mu(s) \tag{21}$$

$$+ \frac{\tau}{1+\tau} \sum_{z \in \mathbb{Z} \setminus D} \int_{-h}^0 j_{x-z} g_z(t+s)\alpha(s)d\mu(s)$$

where $j.$ is a suitable kernel (in this work a Gaussian kernel), $g_z(\cdot)$ is the rate of death either at "negative ages" or after the *actuarial infinite*, $d\mu(s) = \xi(s)ds$ and $\alpha(s)$ is defined above using the annual improvement rates from the past.

Remark 4. *The hypothesis H2 is consistent in the sense that if $\tau \to 0$, then $\overset{\circ}{q}_x(t + \tau) \to \overset{\circ}{q}_x(t)$. Additionally, if $\tau \neq 0$, the new graduate value depends on the step τ and the delay.*

Remark 5. *We note that if the history is not important, then the annual improvement rates are equal to 1, so $\mu(s)$ can be interpreted as a degenerated probability function such that $d\mu(s) = 1$ at 0. In this situation, the expression (21) is analogous to (1).*

Remark 6. *The proposed relation (21) is based on the empirical experience and on the actuarial practice in a similar way as in [10].*

Next, we establish the relation between (21) and (6). From (21), we easily obtain that

$$\frac{(1+\tau)}{\tau}\mathring{q}_x(t+\tau) = \frac{\mathring{q}_x(t)}{\tau} + \sum_{z\in D}\int_{-h}^{0} j_{x-z}\mathring{q}_z(t+s)\alpha(s)d\mu(s) + \sum_{z\in\mathbb{Z}\setminus D}\int_{-h}^{0} j_{x-z}g_z(t+s)\alpha(s)d\mu(s),$$

$$\frac{\mathring{q}_x(t+\tau) - \mathring{q}_x(t)}{\tau} = \sum_{z\in D}\int_{-h}^{0} j_{x-z}\mathring{q}_z(t+s)\alpha(s)d\mu(s) - \mathring{q}_x(t+\tau) + \sum_{z\in\mathbb{Z}\setminus D}\int_{-h}^{0} j_{x-z}g_z(t+s)\alpha(s)d\mu(s).$$

Formally, if we consider a fixed time t, and we take the limits as $\tau \to 0$, then

$$\frac{d}{dt}\mathring{q}_x(t) = \sum_{z\in D}\int_{-h}^{0} j_{x-z}\mathring{q}_z(t+s)\alpha(s)d\mu(s) - \mathring{q}_x(t) + \sum_{z\in\mathbb{Z}\setminus D}\int_{-h}^{0} j_{x-z}g_z(t+s)\alpha(s)d\mu(s). \qquad (22)$$

Remark 7. *The validation of the model will be made using the observed values in the period 2013–2018 and the Lee-Carter model by taking the predicted values in both cases. This can be done showing that our predictions lie in the confidence interval of the predictions by the Lee-Carter model or even that they are near the expected value in this model.*

3.1.2. Application of Theoretical Results

Equation (22) fits into model (6) with $D = \{0, ..., 100\}$, $m = 101$ and $u^i(t) = \mathring{q}_i(t)$. The Gaussian kernel obviousy satisfies (H1) and (H2). The function $g_r(t) \equiv \overline{g}_r \in l_2^\infty$ is a constant function, where $\overline{g}_r$ is defined as follows:

$$\overline{g}_r = \begin{cases} 0 \text{ if } r < 0, \\ g_0 > 0 \text{ if } r > 100. \end{cases} \qquad (23)$$

Hence, (H3) is satisfied. Here, g_0 is the probability of death after the actuarial infinite. In this case, $\alpha_i(s) = \alpha(s)$, where $\alpha(\cdot) \in C([-h,0],\mathbb{R})$ is given in (20). Therefore, (H4) holds true. Since the function g does not depend on, the problem is autonomous, so we take $\tau = 0$.

Hence, Lemma 1 implies that for any initial condition $\phi \in C([-h,0],\mathbb{R}^m)$, there exists a unique solution $u(\cdot) \in C^1([0,+\infty),\mathbb{R}^m)$, where we take $\tau = 0$. Additionally, in view of Proposition 1 if $\phi \in C([-h,0],\mathbb{R}^n)$ is such that $\phi^i(s) \geq 0$ for $s \in [-h,0]$, $i \in D$, then $u^i(t) \geq 0$ for all $i \in D$, $t \geq 0$. This is an important property, because the functions $\mathring{q}_i(t)$ cannot be negative as they are probabilities.

We obtain that for small h there is a unique exponentially stable fixed point as well.

Lemma 7. *We choose $h_0 > 0$ such that*

$$M_0 = \int_{-h_0}^{0} \alpha(s)d\mu(s) \leq 1.$$

Then for any $h \leq h_0$ such that

$$h < -\frac{1}{2}\log(1-\delta_0),$$

where $\delta_0 = \min_{r\in D}\{1 - M_0\sum_{i\in D} j_{i-r}\}$, problem (22) possesses a unique fixed point which is exponentially stable.

Proof. Since j is a Gaussian kernel, $\sum_{i\in D} j_{i-r} < 1$ for any r, so $M_0 \leq 1$ implies that $\delta_0 > 0$. The result follows then from Remark 1 and Corollary 2. $\square$

Although in the actuarial practice, the set D is finite because the probability of death is considered to be constant after some age (the actuarial infinite), it is also possible to estimate the values of $\overset{\circ}{q}_x$ for any non-negative x. In this case, D is of the type (5); more precisely, $D = \{i \in \mathbb{Z} : i \geq 0\}$. As before, (H1), (H2), (H4) are satisfied and the function $g_r(t) \equiv \overline{g}_r$ is a constant function, where $\overline{g}_r$ is defined by: $\overline{g}_r = 0$ if $r < 0$.

If we consider the phase space $X = l_1^\infty$, then $\overline{g}_r \in l_1^\infty$, so (H3) is satisfied. As the Gaussian kernel j has subexponential growth, (H6) is true. Since $\alpha_i(s) = \alpha(s)$ for any $i \in D$, (H7) is obviously satisfied (which implies that (H5) also holds). Then, by Lemma 4 for any initial condition $\phi \in C([-h, 0], l_1^\infty)$ there exists a unique solution $u(\cdot) \in C^1([0, +\infty), l_1^\infty)$. Additionally, in view of Proposition 2, if $\phi \in C([-h, 0], l_1^\infty)$ is such that $\phi^i(s) \geq 0$ for $s \in [-h, 0], i \in D$, then $u^i(t) \geq 0$ for all $i \in D, t \geq 0$.

If we consider the phase space $X = l_1^1$, then $\overline{g}_r \in l_2^1$, so (H3b) is satisfied. Thus, Lemma 5 implies that for any initial condition $\phi \in C([-h, 0], l_1^1)$, there exists a unique solution $u(\cdot) \in C^1([0, +\infty), l_1^1)$. Finally, by Proposition 3 if $\phi \in C([-h, 0], l_1^1)$ is such that $\phi^i(s) \geq 0$ for $s \in [-h, 0], i \in D$, then $u^i(t) \geq 0$ for all $i \in D, t \geq 0$.

3.2. Details and Example

In this section, we will give the details of the numerical simulations.

3.2.1. Data, Period and Software

The data used in this study are downloaded from the Human Mortality Data Base [18]. It refers to the death rates of the Spanish population in the period 1908–2018. In this study, we combine several sub-periods to obtain the estimations. The period 1908–2012 is used to obtain some preliminary estimates; a sub-period of them, 1978–2012, is used to fit the significant improvement rates. We note that this last sub-period coincides with a period of significant changes in Spain. Among other events, the actual democratic status begins, which motivates the consolidation of the welfare system and implies the consolidation or changes of health protocols in Spain such as the transfer of home births to the hospital, significantly reducing the mortality at first ages [41,42]. The period 2013–2018 is used to validate the method.

The software used for all calculations and the construction of the figures is the R-Software [43]. Additionally, the data are downloaded from [18] and they are manipulated and visualized using the R-package [44]. This demographic package is also used to estimate the Lee-Carter Model.

3.2.2. Validation

The validation process has two parts. In the first part, we compare the forecasted data with the observed data in the same period. In the second part, to validate the model, we compare the proposed method with the classical Lee-Carter model [28].

In the first part, we fit the proposed model using the observed data in the period 1978–2012; next, we forecast the mortality in the years 2013–2018 to obtain the prediction of the mortality rates, $q_x^{t,f}$. Then, for each year in the validation period, we compare the observed death rates at each age, $q_x^{t,obs}$, with the forecasted values. The comparison is made using

$$I_{MSD}^t = \sum_{x=0}^{\omega} \frac{\left(q_x^{t,obs} - q_x^{t,f}\right)^2}{\left(q_x^{t,f}\right)^2} \text{ or } \chi_{MSD}^t = \sum_{x=0}^{\omega} \frac{\left(q_x^{t,obs} - q_x^{t,f}\right)^2}{q_x^{t,f}}. \tag{24}$$

It is remarkable that a modification of the previous indicators can be appropriate. In particular, we can use the smoothed values and not the observed or forecasted values directly. We denote this modified indicators as I_{SMSD}^t and χ_{SMSD}^t.

The second part of the validation consists in comparing the proposed method in this work with the classical Lee-Carter model [28]. The comparison is made in two ways. Firstly,

we forecast the mortality rates by the two methodologies in the period 2013–2018. Then, we compute I_{MSD}^t and χ_{MSD}^t (or its smoothed versions) and compare the results directly. Additionally, we project the mortality in the period 2019–2030 with both techniques and compare it to determine if they have a similar magnitude.

The Lee-Carter model was formulated to estimate the central mortality rates, m_x^t, at year t. This model can be formulated as

$$\ln m_x^t = a_x + k^t b_x + e_{x,t}, \tag{25}$$

where a_x, b_x are factors depending only on the age; k^t is a factor that describes the evolution of the mortality over the years; e_x^t is a realization of the white noise. It is remarkable that b_x modulates the factor k^t to adequate the behavior of the mortality due to social improvements such as the control of diseases, the improvement in health habits, nutrition... If we consider this model without other conditions, the number of parameters is $2 \times$ #(*number of ages*) + #(*number of years*) (in our case, $2 \times 101 + 111 = 313$ parameters). It is usual to model the k^t values as an autoregressive process of type ARIMA, and also that the fit provides us an autoregressive model of order 1 (one parameter); in this case, we reduce the number of parameters to 202. It is convenient to remark that the Lee-Carter model is a parametric model and the proposed method is nonparametric. In this sense, the assumptions of the proposed model are less restrictive.

Remark 8. *The election of the Lee-Carter model is not arbitrary and it is justified by the importance of that work. The paper of Lee-Carter model [28] is considered as a seminal model. It is the model with more influence in all ambits of this field of knowledge. We note that, in the actuality, there exist several versions of this model that improve it. However, we consider the comparison with this model by several reasonings: (i) It is usual in the scientific literature to use this model in comparisons; (ii) others models based on the Lee-Carter model either are not better in all conditions or the improvements are small; (iii) the Lee-Carter model has a great number of parameters, 200 or more; comparing this fact with our model, the difference is significant, as the number is less in our case.*

3.2.3. Estimations and Results

It is known that the mortality has a dynamic behavior in all country or regions. Figure 1 shows the evolution of the crude mortality rates in the period 1908–2018. This figure shows the evolution of the death rates over the last 111 years.

We note that the highest series in the graphic correspond to the first years of the study, at the beginning of the XX century. The lines at the bottom of the figure corresponds to the recent years. This indicates to us that there exists an evident improvement in the mortality rates.

In order to apply the expression (21) to the observed data and obtain the forecasted values, it is necessary to know the values of the function g_z after the actuarial infinite. In this work, we consider that these values correspond to the values of the mortality rates for ages greater than a characteristic age, when an effect known as "Plateau effect" begins [45,46]. This effect indicates to us that the risk of death does not increase to 1, but remains constant from certain age. Figure 1 shows this effect (simulated to facilitate interpretation), which can be observed in the final part of each curve and corresponds to the mortality risk values for ages in the range 90–130. In practice, the age at which the Plateau effect begins is unknown and it can be only estimated from the observed data. In this work, we consider that, for each year t greater than a *characteristic age*, $g_z^t = \max\{q_z^t\}_{z>70}$. Figure 2 shows the evolution of the death risk at last ages.
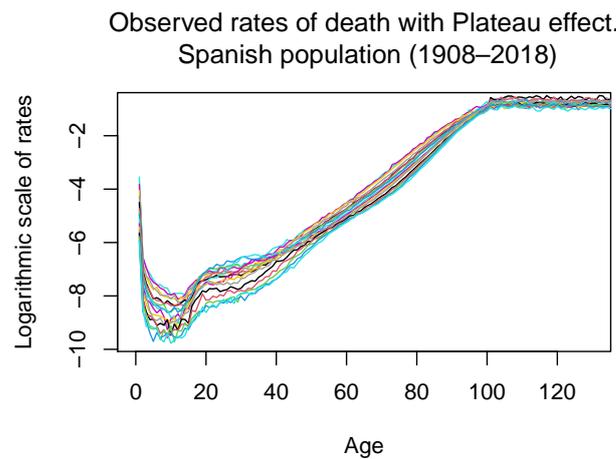
Figure 1. Rates of death.
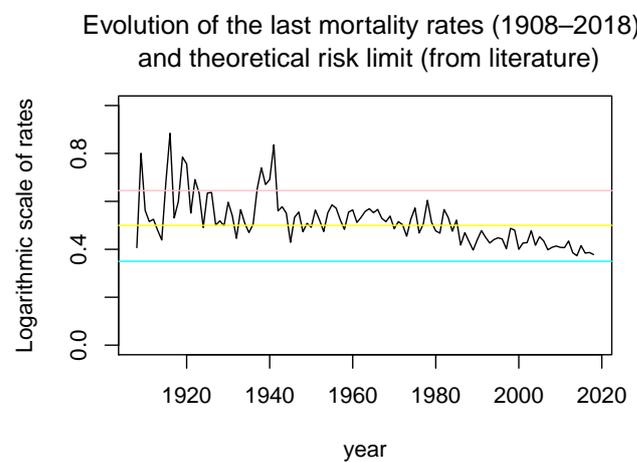
Figure 2. Last mortality rates.

The tendency of the series in Figure 2 is decreasing and the values of this series converge to the limit 0.35 (blue line at the bottom of the figure), which is consistent with the works of other authors ([40], p. 41) in the Spanish case, for both genders and for ages up to 105 years. Other authors consider that this limit is 0.684 [46] (yellow-middle line in the figure), or 0.5 [45] (pink-top line in the figure). This type of assertion is not without controversy [47]. In our simulations, we have taken $g_z^{2012} \equiv 0.385638$ as the probability of death after the actuarial infinite. This is the value of g_0 in (23).

In step 2 of the method, the global improvement rates are estimated using (16). Figure 3 shows these rates for the first 21 lags; we find that the behavior is similar to the first 40 lags. We have analyzed these first delays and consider that they are enough to capture the dynamic of the improvements. Due to this, we have chosen the value of h equal to 21.

We fit a linear model for the global improvement rates (dashed line in Figure 3) without the constant coefficient and the obtained value of the parameter is $\hat{b} = -0.004171019$. In this method, we consider as global improvement rates the values derived from the expression $\overline{\overline{r}}^d = \overline{\overline{r}}(d) = 1 + \hat{b} \cdot d, d \in \{1, 2, \ldots, 21\}$. Hence, with such a procedure, the function $\alpha(s)$ is linear.

On other hand, we assume that the impact of the improvement rates is not equal for all of them. Its importance is modulated by another function. As we have said, the improvement rates are modulated by a probability function which is exponential, as in expression (18), and the weigthed improvement rates are obtained as in (19). The parameter λ in (18) is equal to $\frac{11}{12}$. Figure 4 shows the probability function which is used.
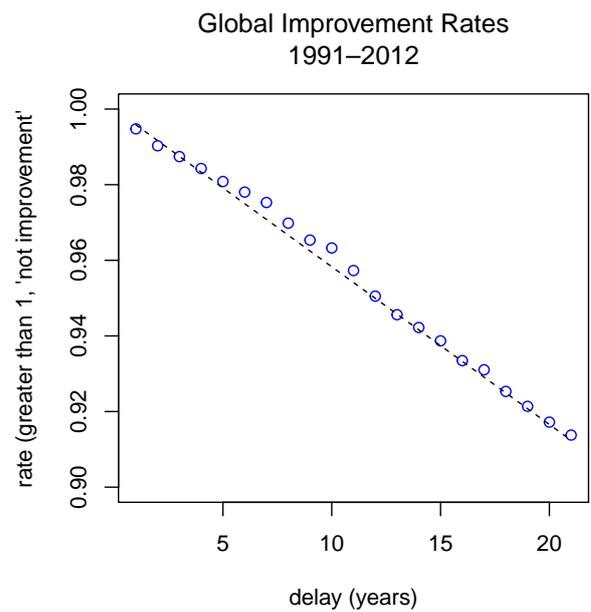
Figure 3. Improvement rates.
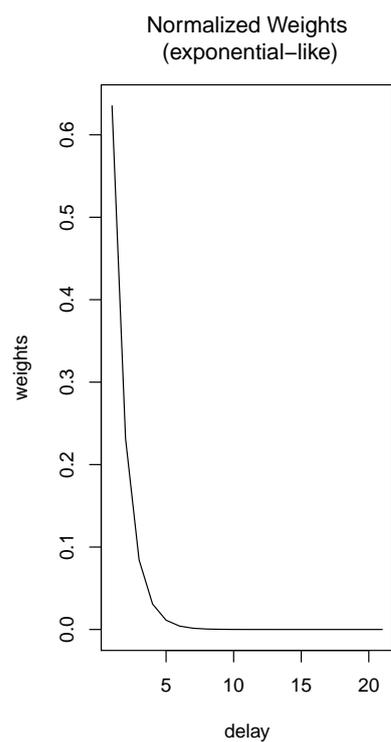
Figure 4. Normalized weights.

Now, we have defined all the elements which are necessary to apply the method for forecasting.

In order to estimate the solution, we use expression (21) with $\tau = 1$ and approximate the integral term with the Riemann summatory.

Firstly, and with the objective of using the observed data to measure the quality of the technique, we apply the method using the data in the period 1908–2012 and obtain two types of predictions for the period 2013–2018: (i) applying the proposed technique; (ii) applying the Lee-Carter method. Additionally, the indicators of quality defined by (24) are estimated.

Figure 5 shows the predictions with the two techniques, the proposed technique and the Lee-Carter method. We can observe that these two series of values are close to the observed data. The magnitude of the values are coherent in the two methods. Comparing them with a graphical analysis, we are not able to identify whether one method is better than the other.
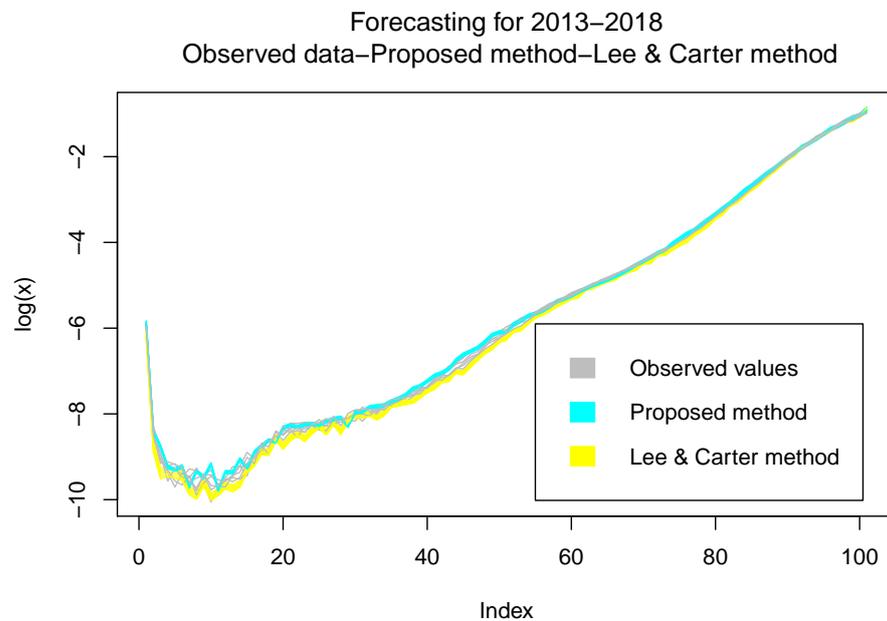
Figure 5. Forecasting.

In order to clarify if one technique is better than the other, we need to calculate a numerical indicator. In this sense, we consider the *quality indicators* defined in (24). Using the observed data as reference data and the estimation made with each technique as the assumption, we obtain the values in Table 1.

Table 1. Summary of the measures of goodness of the predictions.

	I_{MSD}^{2013}	I_{MSD}^{2014}	I_{MSD}^{2015}	I_{MSD}^{2016}	I_{MSD}^{2017}	I_{MSD}^{2018}
Model	0.05461045	0.06223897	0.02171756	0.03676833	0.02777047	0.03182172
Lee-Carter	0.08731890	0.11195858	0.02715306	0.09994482	0.07357846	0.10755673
%Improve.	62.5	55.6	80.0	36.8	37.7	29.6
	χ_{MSD}^{2013}	χ_{MSD}^{2014}	χ_{MSD}^{2015}	χ_{MSD}^{2016}	χ_{MSD}^{2017}	χ_{MSD}^{2018}
Model	0.1641217	0.2378390	0.1283401	0.2026115	0.1892803	0.2083615
Lee-Carter	0.1919128	0.2775853	0.0947843	0.2470046	0.2353773	0.3415634
%Improve.	85.5	85.7	135.4	82.0	80.4	61.0

We can observe that the proposed technique is more accurate that the Lee-Carter method in both indicators of goodness. In the case of the indicators I_{MSD}^t, the improvement is for all years and the values of the error of the proposed technique with respect to the observed values are between 30% and 80% of the error of the Lee-Carter model. In the case of the indicators χ_{MSD}^t, we observe that the improvement is not for all years, as in 2015, the Lee-Carter model has better results, but we think that this occurs because that year has a great oscillation at initial ages. For the rest of the years, the rate of improvement of the proposed method with respect to the Lee-Carter model was between 30% and 80%.

Now, as we said before, we will compare the proposed technique with the Lee-Carter method in forecasting. Figure 6 shows the predictions with these two methods for the period 2019–2028.
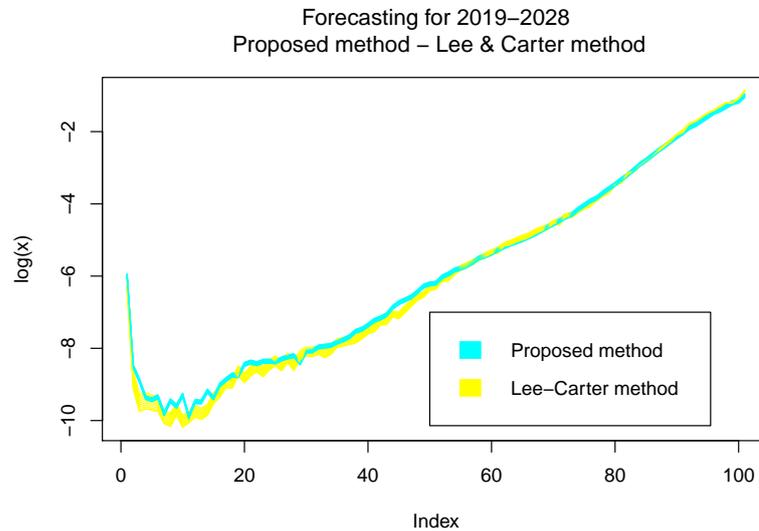
Figure 6. Long-term forecasting.

We observe that the Lee-Carter method has greater variability than the proposed method for the ages in the intervals 5–15 and 40–50. In addition, the mortality curves have similar values for the rest of the ages. As we do not have observed values of mortality for the period 2019–2028, it is difficult to determine which of the two techniques is better. In this situation, we will only compare the magnitude of the predictions with these two techniques, checking whether the values are similar or not. Since the Lee-Carter method is a reference method, we will verify if our results are within the confidence intervals of the Lee-Carter predictions. Figure 7 shows the values of the predictions with the proposed technique jointly with the upper and lower confidence curves of the Lee-Carter predictions (at a 95% confidence-level) for the years 2021–2028. It is clear that the values are between the two bounds of the Lee-Carter model, which indicates us that the method is appropriate. Additionally, considering that with the predictions between 2013 and 2018, the proposed method has turned out to be more accurate than the Lee-Carter method, it is convenient to reflect on which of these methods is appropriate.

We note that we can measure the goodness of the forecasting otherwise by using the quality indicators (24) but changing the reference values. First, we must assume as correct the values of the Lee-Carter values; second, we assume as correct the values of the proposed method. The comparison is made using the caution principle, depending of the application. For example, if we use the values of the proposed method when the true values are the Lee-Carter ones, the indicator I_{MSD}^{2019} was equal to 1.812869, which tells us that "globally" the discrepancies were close to 81%. If we use the Lee-Carter values when the true values are the ones of the proposed technique, then I_{MSD}^{2019} was 2.551214, which indicates to us that, "globally", discrepancies were close to 155%. The "possible error" was double if we used the Lee-Carter values. The same situation occurs for the rest of the years.
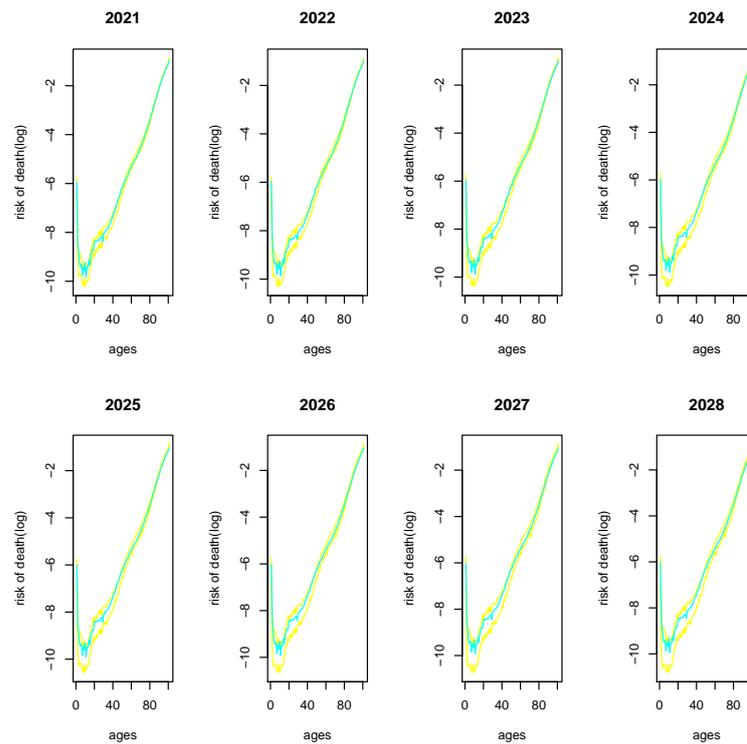
Figure 7. Comparison by using the confidence intervals of the Lee-Carter method.

These two ways of measuring the quality of the proposed method allow us to affirm that, in the middle-term, the proposed method is quite appropriate.

It is important to point out that there exist several ways to improve this work: (i) introducing one improvement rate for each age instead of estimating a unique improvement rate for all ages, $q_.^{t_1} = \bar{r}^{t_0,t_1} \cdot q_.^{t_0}$ (obviously, this increases the number of parameters of the problem, but no more than in the Lee-Carter model); (ii) introducing a process that allows us to select the better probability function.

Finally, we would like to remark that the main differences in the numerical simulations of this paper with respect to the ones in [10] are the following:

- In the model of this paper, we have considered the influence of the values of the variable in the past and not only in the current moment of time, as in [10].
- The functions $\alpha(s)$ are new in this paper and we have approximated them via the improvement rates $\bar{\bar{r}}^d$.
- We have considered the probability functions (17), (18) in order to modulate the improvement rates $\bar{\bar{r}}^d$.
- In [10], the value of the function g was taken to be equal to 0 outside the domain D, whereas in the present work, the value of g for ages greater than 100 was taken to be equal to a positive number due to the "Plateau effect".
- The numerical simulations show that the predictions are good enough for at least a period of 8 years, whereas in the model without delay, the predictions were good for a period of 3 years.

4. Conclusions

In this work, we considered a system of ordinary differential equations with non-local discrete diffusion and finite delay. This paper extends the results in [10], incorporating prior information in the form of delay. With this, it is possible to improve the prediction window with respect to the previous work; in [10], the horizon of the prediction is adequate up to three years; however, in this work, it is verified (with the data in the considered

period and for the defined measures) that the prediction horizon can be extended up to 8 years, and that it gives coherent values, in magnitude, when we compare it with other classical techniques such as the Lee-Carter model, up to 18 years.

The first part of the paper is dedicated to theoretically analyzing the system: firstly, a system with a finite number of equations; secondly, a system with an infinite number of them. In both cases, we prove the existence of solutions and several properties of them such as comparison results, stability and symmetry.

In the second part of the work, we applied the model to life tables, defining the parameters of the model (for example, the delay is incorporated as improvement rates of the probability of death) and applying the theoretical results of the first part by estimating and projecting to the future the risk of mortality (by age). In order to calculate the numerical estimations, we considered the data of Spanish mortality in the period 1908–2018 from [18]; the choice of this source of information is due to the fact that this organization uses a methodology that unifies certain methodological aspects of the statistical bureaus of several countries, providing a common framework that allows us to assure that the proposed model can be applied to any region in the world. In order to validate the proposed model and compare it with the Lee-Carter model, we defined some indicators of goodness and smoothness. In this sense, these measures indicate to us that this model is more accurate than the classical Lee-Carter model in the considered period of validation (between 2013 and 2018). Additionally, the proposed model gives us coherent values up to 2028 (maybe even for a larger period of time) when we compare these values with the values of the Lee-Carter model.

Finally, we consider several aspects of this work that could be improved in future researches. For example, we could consider improvement rates which are dependent of age or use other density probability functions, and not only the exponential distribution. It is possible to introduce other aspects in the model as the cohort effect, which is usual in the demographic and actuarial fields.

Author Contributions: Conceptualization, F.M. and J.V.; methodology, F.M. and J.V.; software, F.M. and J.V.; validation, F.M. and J.V.; formal analysis, F.M. and J.V.; investigation: F.M. and J.V.; resources, F.M. and J.V.; writing—original draft preparation, F.M. and J.V.; writing—review and editing, F.M. and J.V.; visualization, F.M. and J.V.; supervision, F.M. and J.V.; project administration, F.M. and J.V.; funding acquisition, F.M. and J.V. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are openly available in Human Mortality Data Base at www.mortality.org or www.humanmortality.de.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fife, P. Some nonclassical trends in parabolic and parabolic-like evolutions. In *Trends in Nonlinear Analysis*; Springer: Berlin, Germany, 2003; pp. 153–191.
2. Cortazar, C.; Elgueta, M.; Rossi, J.D. Nonlocal diffusion problems that approximate the heat equations with Dirichlet boundary conditions. *Israel J. Math.* **2009**, *170*, 53–60. [CrossRef]
3. Pérez-Llanos, M.; Rossi, J.D. Numerical approximations for a nonlocal evolution equations. *SIAM J. Numer. Anal.* **2012**, *49*, 2103–2123. [CrossRef]

4. Bates, P.; Fife, P.; Ren, X.; Wang, X. Travelling waves in a convolution model for phase transitions. *Arch. Ration. Mech. Anal.* **1997**, *138*, 105–136. [CrossRef]

5. Bates, P.; Chmaj, A. A discrete convolution model for phase transitions. *Arch. Ration. Mech. Anal.* **1999**, *150*, 281–305. [CrossRef]

6. Bates, P.; Chmaj, A. An integrodifferential model for phase transitions: Stationary solutions in higher dimensions. *J. Stat. Phys.* **1999**, *95*, 1119–1139. [CrossRef]

7. Bates, P.; Han, J. The Dirichlet boundary problem for a nonlocal Cahn-Hilliard equation. *J. Math. Anal. Appl.* **2005**, *311*, 289–312. [CrossRef]

8. Chasseigne, E.; Chaves, M.; Rossi, J.D. Asymptotic behavior for nonlocal diffusion equations. *J. Math. Pures Appl.* **2006**, *86*, 271–291. [CrossRef]

9. Bogoya, M; Gómez S., C.A. Discrete model of a nonlocal diffusion equation. *Rev. Colomb. Mat.* **2013**, *47*, 83–94.

10. Morillas, F.G.; Valero, J. On a nonlocal discrete diffusion system modeling life tables. *Rev. R. Acad. Cienc. Exactas Fis. Nat. Ser. A Mat. RACSAM* **2014**, *108*, 935–955. [CrossRef]

11. Cairns, A.J.G.; Blake, D.; Dowd, K. Modelling and management of mortality risk: A review. *Scand. Actuar. J.* **2008**, *2008*, 79–113. [CrossRef]

12. Hale, J.K. *Theory of Functional Differential Equations*; Springer: New York, NY, USA, 1977.

13. Amigó, J.M.; Giménez, A.; Morillas, F.; Valero, J. Attractors for a lattice dynamical system generated by non-newtonian fluids modelling suspensions. *Internat. J. Bifur. Chaos* **2010**, *20*, 2681–2700. [CrossRef]

14. European Parliament and of the Council. Risk Management and Supervision of Insurance Companies (Solvency II)-Consolidated text: Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the Taking-Up and Pursuit of the Business of Insurance and Reinsurance. Available online: https://eur-lex.europa.eu/ (accessed on 6 December 2020).

15. Chiang, C.L. *The Life Tables and Its Applications*; Krieger Publishing Company: Malabar, FL, USA, 1984.

16. Benjamin, B.; Pollard, J.H. *The Analysis of Mortality and Other Actuarial Statistics*; Institute of Actuaries and the Faculty of Actuaries: Oxford, UK, 1993.

17. Ayuso, M.; Corrales, H.; Guillén, M.; Perez-Marín A.M.; Rojo, J.L. *Estadística Actuarial Vida*; Universitat de Barcelona, Edicions: Barcelona, Spain, 2007.

18. University of California, Berkeley (USA); Max Planck Institute for Demographic Research (Germany). Human Mortality Database. Series of Death Rates of Spain. Available online: www.mortality.org (accessed on 1 December 2020).

19. Instituto Nacional de Estadística (Statistics National Institute of Spain). Life Tables: Methodolgy. Available online: https://ine.es/en/metodologia/t20/t2020319a_en.pdf (accessed on 6 December 2020).

20. Gompertz, B. On the nature of the function of the law of human mortality and on a new mode of determining the value of life contingencies. *Trans. R. Soc.* **1825**, *115*, 513–585.

21. London, D. *Graduation: The Revision of Estimates*; ACTEX Publications: New Hartford, CT, USA, 1985.

22. Beer, J. Dealing with uncertainty in population forecasting (Statistics Netherlands). Available online: https://www.cbs.nl/-/media/imported/documents/2000/24/dealing-with-uncertainty.pdf (accessed on 21 December 2020)

23. Helligman, L.M.A.; Pollard, J.H. The age pattern of mortality. *J. Inst. Actuar.* **1980**, *107*, 49–82. [CrossRef]

24. Forfar, D.O.; McCutcheon, M.A.; Wilkie, M.A. On graduation by mathematical formula. *J. Inst. Actuar.* **1988**, *115*, 1–149. [CrossRef]

25. Gavin, J.; Haberman, S.; Verrall, R. Moving weighted average graduation using kernel estimation. *Insur. Math. Econ.* **1993**, *12*, 113–126. [CrossRef]

26. Debón, A.; Montes, F.; Sala, R. A comparison of parametric models for mortality graduation. Application to mortality data for the Valencia Region (Spain). *SORT Stat. Oper. Res.* **2005**, *29*, 269–288.

27. Villegas, A.M.; Kaishev, V.K.; Millossovich, P. StMoMo: A R package for stochastic mortality modeling. *J. Stat. Softw.* **2018**, *84*, 1–38. [CrossRef]

28. Lee, R.; Carter, L. Modelling and forecasting US mortality. *J. Am. Stat. Assoc.* **1992**, *87*, 659–671.

29. Debón, A.; Montes, F.; Sala, R. A comparison of models for dynamic life tables. Application to mortality data from the Valencia Region (Spain). *Lifetime Data Anal.* **2006**, *12*, 223–244. [CrossRef]

30. Debón, A.; Montes, F.; Puig, F. Modelling and forecasting mortality in Spain. *Eur. J. Oper. Res.* **2008**, *189*, 624–637. [CrossRef]

31. Haberman, S.; Renshaw, A. On age-period-cohort parametric mortality rate projections. *Insur. Math. Econ.* **2009**, *45*, 255–270. [CrossRef]

32. Dodd, E.; Forster, J.J.; Bijak, J.; Smith, P.W.F. Smoothing mortality data: Producing the English life table, 2010–2012. *J. R. Stat. Soc. Ser. A Stat. Soc.* **2016**, *181*, 1–19. [CrossRef]

33. Copas, J.B.; Haberman, S. Non-parametric graduation using kernel methods. *J. Inst. Actuar.* **1983**, *110*, 135–156. [CrossRef]

34. Morillas, F.G.; Baeza, I.; Pavia, J.M. Risk of death: A two-step method using wavelets and piecewise harmonic interpolation. *Estadística Española* **2016**, *58*, 245–264.

35. Morillas, F.G.; Sampere, I.B. Using wavelet techniques to approximate the subjacent risk of death. In *Modern Mathematics and Mechanics*; Sadovnichiy, V.A., Zgurovsky, M.Z., Eds.; Springer International Publishing: Cham, Switzerland, 2019; Chapter 28, pp. 541–557.

36. Brouhns, N.; Denuit, M.; Keilegom, I.V. Bootstrapping the Poisson log-bilinear model for mortality forecasting. *Scand. Actuar. J.* **2005**, *2005*, 212–224. [CrossRef]

37. Dodd, E.; Forster, J.J.; Bijak, J.; Smith, P.W.F. Stochastic modelling and projection of mortality improvements using a hybrid parametric/semi-parametric age–period–cohort model. *Scand. J.* **2020**. [CrossRef]
38. Haberman, S.; Renshaw, A. Parametric mortality improvement rate modelling and projecting. *Insur. Math. Econ.* **2012**, *50*, 309–333. [CrossRef]
39. Haberman, S.; Renshaw, A. Modelling and projecting mortality improvement rates using a cohort perspective. *Insur. Math. Econ.* **2013**, *53*, 150–168. [CrossRef]
40. Albarrán Lozano, I.; Ariza Rodríguez, F.; Cóbreces Juárez, V.M.; Durbán Reguera, M.L.; Rodríguez-Pardo del Castillo, J.M. Riesgo de Longevidad y su aplicación práctica a Solvencia II. VIII International Awards "Julio Castelo Matrán", Fundación Mapfre. Available online: https://www.fundacionmapfre.org/fundacion/es_es/publicaciones/diccionario-mapfre-seguros/r/riesgo-de-longevidad.jsp (accessed on 21 December 2020).
41. Montes, M.J. Las Culturas del Nacimiento. Ph.D. Thesis, Universitat Rovira i Virgili, Tarragona, Spain, 2007.
42. Lledó, J.; Pavía, J.M.; Morillas, F.G. Transformaciones en la distribución semanal de nacimientos. Un análisis temporal 1940–2010. *Revista Española Investigaciones Sociológicas* **2017**, 159, 151–162. [CrossRef]
43. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020. Available online: https://www.R-project.org/ (accessed on 21 December 2020).
44. Hyndman, R.J. (with Contributions from Heather Booth, Leonie Tickle and John Maindonald). Package Demography: Forecasting Mortality, Fertility, Migration and Population Data (R Package Version 1.22). 2019. Available online: https://CRAN.R-project.org/package=demography (accessed on 21 December 2020).
45. Dolgin, E. Longevity data hint at no natural limit on lifespan. *Nature* **2018**, *559*, 14–15. [CrossRef]
46. Barbi, E.; Lagona, F.; Marsili, M.; Vaupel, J.W.; Kenneth, W.W. The plateau of human mortality: Demography of longevity pioneers. *Science* **2018**, *360*, 1459–1461. [CrossRef] [PubMed]
47. Dong, X.; Milholland, B.; Vijg, J. Evidence for a limit to human lifespan. *Nature* **2016**, *538*, 257–259. [CrossRef] [PubMed]