SPECIAL ISSUE PAPER

WILEY

Extending the hyper-logistic model to the random setting: New theoretical results with real-world applications

Juan Carlos Cortés[1] | Ana Navarro-Quiles[2] | Sorina Madalina Sferle[1]

[1]Instituto Universitario de Matemática Multidisciplinar, Universitat Politècnica de València, Valencia, Spain

[2]Department of Statistics and Operational Research, Universitat de València, Burjassot, Spain

Correspondence

Sorina Madailna Sferle, Instituto Universitario de Matemática Multidisciplinar, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain.
Email: smsferle@doctor.upv.es

Communicated by: D. Zeidan

We develop a full randomization of the classical hyper-logistic growth model by obtaining closed-form expressions for relevant quantities of interest, such as the first probability density function of its solution, the time until a given fixed population is reached, and the population at the inflection point. These results are obtained under very general hypotheses on the distributions of the random model parameters by taking extensive advantage of the so-called random variable transformation method. To illustrate the practical implications of our findings, we apply them to model the growth of multicellular tumor spheroids using empirical data. In this context, we explore two methodologies—the Bayesian approach and the random least mean square method—aimed at effectively addressing the challenge of assigning appropriate distributions to model parameters. This ensures that probabilistic fits accurately capture the inherent uncertainties of tumor growth dynamics. Finally, we notably show that the results obtained using both approaches in the randomized hyper-logistic model align closely with each other, surpassing those yielded by the randomized logistic model.

KEYWORDS

hyper-logistic model, random differential equation, real-world application, uncertainty quantification

MSC CLASSIFICATION

34F05

1 | INTRODUCTION AND MOTIVATION

Questions related to the concept of growth and its patterns have always been the subject of numerous studies in completely different areas, such as forest growth in Ecology [1], microbial and bacterial growth in food in Biology [2], growth of tumors in Medicine [3, 4], or national economic growth in Economics [5], just to mention a few ones. The great interest in describing the growth dynamics in different settings has led to a number of distinctive growth curves, which, depending on the evolution of the quantity of interest over time, have linear, logarithmic, exponential, sigmoidal (S-shaped), and so forth, forms or a combination of some of them.

This study delves into randomized growth models, specifically those resulting in a sigmoidal curve with a horizontal asymptote. The curve exhibits three phases:

• Lag phase: Population growth is slow due to a scarcity of reproductive individuals.

- Log or exponential phase: Growth accelerates as resources abound and environmental resistance is low.
- Stationary phase: Growth decelerates due to survival competition, reaching the environment's carrying capacity.

A common feature in population dynamics is the exponential growth phase when the highest reproductive rate of the species is reached, which is called the inflection point. Therefore, the growth rate plays an important role in these models, as it measures the percentage of change of a given metric over a specific period. One of the first to develop a mathematical formula to estimate this rate and predict exponential growth was Malthus, which gave rise to the so-called Malthusian growth model, also named the exponential growth model. Malthus's model is the simplest population dynamics model. Despite its mathematical simplicity, Malthus's model is considered the cornerstone of growth models because it establishes the basis for more complex and refined models, such as the Verhulst, Gompertz, and Bertalanffy models (see the original contributions in [6, 7], and [8], respectively). Shortly after Bertalanffy, Richards [9] proposed a new growth model from which Verhulst, Gompertz, and Bertalanffy models can be obtained as special cases. Richard's model was designed to provide greater flexibility to the growth models that had been proposed up to that time. In this paper, we address the study of the hyper-logistic model, which has also been considered when dealing with growth modeling in different scenarios. As it shall be seen later, we will study this relevant growth model in case its parameters are random variables, which is more realistic in real-world applications.

For completeness, we shall first summarize the main features of relevant growth models to better motivate the hyper-logistic model. In the Malthus model, the growth rate, $\frac{dp(t)}{dt}$, is proportional to the size of the existing population, $p(t)$, which leads to the following initial value problem (IVP):

$$\begin{cases} \frac{dp(t)}{dt} = rp(t), \\ p(0) = p_0, \end{cases} \tag{1}$$

where r is the constant of proportionality and p_0 is the initial population size. Apart from its utility in modeling early-stage growth, this model often fails to accurately describe overall population dynamics as it generates unbounded growth, neglecting both resource limitations and competition factors.

Verhulst introduced the concept that populations reach a growth limit due to resource scarcity and intra-species competition. This is represented by the logistic growth equation, an extension of the exponential model. Here, the rate of population change ($\frac{dn(t)}{dt}$) is proportional to the existing population ($n(t)$) and the available resources ($1 - \frac{n(t)}{k}$), where $k > 0$ denotes the carrying capacity or the stable population level. The model is then formulated via the following IVP:

$$\begin{cases} \frac{dn(t)}{dt} = rn(t)\left(1 - \frac{n(t)}{k}\right), \\ n(0) = n_0, \end{cases} \tag{2}$$

where r is the intrinsic growth rate and n_0 is the population size at time $t = 0$. As we are interested in growth curves with a sigmoidal shape, we will limit ourselves to the case that $r > 0$. In addition, it is worth mentioning that the higher this rate is, the sooner the carrying capacity will be reached since the slope of the curve in the exponential phase is very high. However, if $r < 0$, the growth curve is asymptotic to 0, leading to population extinction, while if $r = 0$, the population remains constant and equal to n_0 over time. A key aspect of this model is that the population at the inflection point is exactly half the carrying capacity, that is, $n_{inf} = \frac{k}{2}$. This constraint imposes symmetry on the curve, which could be undesirable. However, despite this limitation, numerous studies have utilized the logistic curve to model various biological systems and disease dynamics [10–12]. Various modifications of the logistic model have emerged as parameterized versions, offering relaxed constraints or limitations inherent in the original logistic model [13].

In 1968, Blumberg proposed a modified logistic differential equation to address a key limitation of the Verhulst model: the requirement for the inflection point to align exactly with half of the carrying capacity. Previous efforts attempted to circumvent this by treating the growth rate constant as a time-dependent polynomial, but often resulted in underestimation of future values. Blumberg introduced the hyper-logistic function, also known as the generalized Verhulst model, by weighting each proportionality term with a certain power. This model is also referred to as the generalized Verhulst model [2, 14], as a generalization of Richard's model,

$$\begin{cases} \frac{dx(t)}{dt} = rx(t)^\alpha \left(1 - \frac{x(t)}{k}\right)^\gamma, \\ x(0) = x_0, \end{cases} \tag{3}$$

where r is the time-independent growth constant, k is the limiting value of the growth variable $x = x(t)$, and $\alpha, \gamma > 0$ are the shape parameters. An $\alpha > 1$ means that the initial population growth rate exceeds that implied by the original model, and an $\gamma > 1$ implies that the population is more sensitive to resource decline than the original model indicates. Otherwise, if $\alpha, \gamma < 1$, it does not exercise its full growth potential and is less sensitive, respectively [2]. These additional parameters highlight the significance of the Blumberg growth model, allowing for flexibility in placing the sigmoid curve's inflection point between the minimum and carrying capacity [15]. When differential equation in (3) is expressed as an integral equation, it may lack an analytical solution except for specific values of the shape parameters, as noted by Blumberg [16]. Additionally, the inflection point can be determined as

$$x_{\inf} = \frac{\alpha}{\alpha + \gamma} k. \tag{4}$$

For $\alpha \gg \gamma$, the inflection occurs near k, while for $\alpha \ll \gamma$, $x_{\inf}$ approaches 0 and inflection occurs only if $x_0 < x_{\inf}$ [13]. Note that if the exponents are equal, $\alpha = \gamma = 1$, both expressions (3) and (4) reduce to the Verhulst case. It is worth mentioning that Tsoularis and Wallace defined in 2002 the generalized population logistic model, from which the Blumberg equation is obtained as a particular case [13].

In 1976, Turner et al. introduced a general growth model with specific parameter constraints, leading to a series of nested empirical models [17]. These models exhibit a unique characteristic: The rate of size change is proportional to the product of one increasing function and another decreasing function with size. A particular case stemming from this equation is the hyper-logistic equation, which takes the following form:

$$\begin{cases} \frac{dv(t)}{dt} = \frac{b}{k} v(t)^{1-p} (k - v(t))^{1+p}, \ 0 < p < 1, \ b, k, v_0 > 0, \\ v(t_0) = v_0. \end{cases} \tag{5}$$

This reparametrizes Equation (3), where $v(t)$ represents the population or size of the organism at time t, and $k - v(t)$ indicates the difference between the size at time t and the carrying capacity k. The authors set $\frac{b}{k}$ as the proportionality constant, with b denoting the intrinsic growth constant. In their growth postulate, the exponents of the functions satisfy certain constraints: The first exponent is $1 - p$, and the second is $1 + p$, summing to 2 with $-1 < p < 1$ [17]. As we focus on growth models, p is confined to the interval $(0, 1)$. Consequently, Equation (5) relies on a single shape parameter, p, possessing one more parameter than Equation (2). Lastly, we assume that the population or size of the organism at the initial time, t_0, is v_0.

These shape parameters determine the balance between the two terms of the growth rate, thereby shaping the growth curve. Initially, the second term carries more weight due to the greater distance to the carrying capacity, rising to a power exceeding 1. Conversely, toward the end, the first term dominates as the second term diminishes near 0 when $v = v(t)$ approaches k, leading to almost zero growth during the stationary phase. Notably, both terms are comparable in magnitude during the exponential phase. Additionally, the constant velocity dictates the speed of transitioning between phases. When p approaches 0, we obtain a reparametrization of the logistic equation, expressed as

$$\begin{cases} \frac{dv(t)}{dt} = \frac{b}{k} v(t)(k - v(t)), \\ v(t_0) = v_0, \end{cases} \tag{6}$$

and for which the solution is

$$v(t) = \frac{k}{1 + e^{-b(t - t_0)} \left(\frac{k}{v_0} - 1 \right)}. \tag{7}$$

In this contribution, we focus on the study of growth model (5) whose solution is given by [17]

$$v(t) = \frac{k}{1 + \left(bp(t - t_0) + \left(\frac{k}{v_0} - 1 \right)^{-p} \right)^{-\frac{1}{p}}}. \tag{8}$$

TABLE 1 Formulation of some growth models via the relative instantaneous growth function $g(t) = \frac{\dot{x}(t)}{x(t)}$.

Model	$g(t) = \frac{\dot{x}(t)}{x(t)}$
Malthus	r
Logistic	$r\left(1 - \frac{x(t)}{k}\right)$
Richard	$rx(t)^{\alpha-1}\left(1 - \frac{x(t)}{k}\right)^{\gamma}$
Hyper-logistic	$\frac{b}{k}\frac{(k-x(t))^{1+p}}{x(t)^p}$

In order for this expression to have a sigmoid shape (corresponding to a growth curve), it is easy to check that $0 < v_0 < v(t) < k$ for $t > t_0$. Finally, notice that Blumberg or hyper-logistic model (3) is consistent with Turner's general model (5) when $\alpha = 1 - p > 0$ and $\gamma = 1 + p > 0$.

Remark 1. Although in the previous presentation, different growth models have been introduced with reference to the explicit expression of the function, say $f(t)$, defining the instantaneous variation of the population, $\frac{dx(t)}{dt} = \dot{x}(t)$, one can also motivate the formulation of every growth model in terms of the function, say $g(t)$, defining the relative instantaneous growth function $\frac{\dot{x}(t)}{x(t)}$. In Table 1, one summarizes both formulations for each of the growth models previously mentioned.

It is essential to point out that when growth models are applied to real-world data, their parameters must be estimated from samples containing uncertainties collected from sampling processes. Apart from this source of randomness, one should add the lack of intrinsic knowledge of the growth process, which may often depend on complex factors such as environment, genetics, and birth/death rates, that rarely are known in a deterministic manner. This approach encourages the use of random or stochastic growth models formulated through differential equations to enhance realism compared to deterministic models. To account for uncertainties, in the setting of differential equations, one mainly distinguishes two approaches, namely, stochastic and random differential equations, in short, SDEs and RDEs, respectively. SDEs are those where uncertainties are driven by an irregular process (nowhere differentiable), typically a Wiener process or a Brownian motion, leading to Itô-type SDEs [18, p. 96]. The rigorous treatment of SDEs requires special stochastic calculus such as Itô calculus [19]. From the point of view of uncertainty quantification, SDE can often be obtained by means of Gaussian perturbation (via the white noise process) of their deterministic differential equations counterparts. Although the celebrated central limit theorem from probability theory supports this approximation in many cases, it could limit the most proper pattern to describe uncertainties in real-world problems. Indeed, as highlighted in [20], the Wiener process is sample continuous but with unbounded variation paths, which does not suit the idea of modeling real situations since, in many situations, the real-world phenomenon is subjected to fluctuations that are known to be bounded. Apart from these issues, in the case of the stochastic formulation of the logistic model using the SDE approach, one can obtain two main classes of SDEs, namely, additive or multiplicative logistic SDE, depending on the way the intrinsic noise or perturbation, via the white noise, is done from its deterministic formulation. The results provided by these two logistic models are generally different, and, as addressed in [21], the approximation of parameter estimation could become challenging.

Complementary to SDEs, RDEs are those where uncertainty is directly assigned in model inputs (initial/boundary conditions, source term, and/or coefficients) by assuming they are random variables (or stochastic processes) with regular sample behavior (e.g., continuous trajectories) [18, p. 97]. Assigning appropriate probability distributions to these model inputs is crucial under this approach. In recent years, several studies have adopted this methodology to investigate various growth models with uncertainties. For example, in [22], the logistic growth model is examined considering uncertainties in environmental carrying capacity and initial population. Similarly, in [23], Bertalanffy's non-autonomous growth model is extended to the random setting using the RDE approach.

A key advantage of RDEs lies in their flexibility to assign suitable probability distributions to model inputs, enabling the description of the dynamics of the growth phenomenon being studied. This is achieved while retaining the well-known advantages from classical calculus to obtain the corresponding RDEs' solution. Indeed, the solution of an RDE retains the same form as in the classical setting but with model parameters treated as random variables. Consequently, the resulting solution becomes a parametric stochastic process, requiring the determination of its probability distribution to describe the random dynamics accurately. However, it's worth noting that, compared to SDEs, the advancement of this approach, as indicated in various references [18, 24], is currently less developed.

In the case of the hyper-logistic model (5), the exponent, p, the proportionality constant, b, and the carrying capacity, k, that determine growth dynamics depend on environmental factors, especially temperature, so there is an intrinsic

uncertainty that in its deterministic formulation is neglected. Here, we shall follow the above-mentioned RDE approach, and for the sake of generality, we shall assume that all these parameters are random variables. Therefore, the formal randomization of model (5) writes

$$\begin{cases} \frac{dv(t,\omega)}{dt} = \frac{b(\omega)}{k(\omega)} v(t,\omega)^{1-p(\omega)} (k(\omega) - v(t,\omega))^{1+p(\omega)}, & \omega \in \Omega, \\ v(t_0) = v_0(\omega), \end{cases} \tag{9}$$

where $k = k(\omega)$, $b = b(\omega)$, $p(\omega)$, and $v_0 = v_0(\omega)$ are absolutely continuous random variables defined in a common complete probability space $(\Omega, \mathcal{F}_\Omega, \mathbb{P})$. As usual, the ω-notation highlights the dependence on events of the probability space. When evident, the ω-dependence will be hidden to alleviate the notation. For the sake of generality in our subsequent mathematical development, we shall assume a joint probability density function (PDF) for model parameters, say $f_0(b, k, p, v_0)$. When model parameters are independent, this function factorizes as the product of the marginal PDFs of each parameter, $f_0(b, k, p, v_0) = f_b(b) f_k(k) f_p(p) f_{v_0}(v_0)$. As previously indicated, now the solution is the following parametric stochastic process,

$$v(t,\omega) = \frac{k(\omega)}{1 + \left(b(\omega) p(\omega)(t - t_0) + \left(\frac{k(\omega)}{v_0(\omega)} - 1 \right)^{-p(\omega)} \right)^{-\frac{1}{p(\omega)}}}, \qquad \omega \in \Omega. \tag{10}$$

As specified in the introduction section, in our context, we shall assume that $0 < v_0 < v < k$ in order for the above solution to have a sigmoid shape, so increasing and then admitting an inverse. As invertibility of the solution is a condition that will be required later in the probabilistic analysis of the problem, hereinafter, we will assume that we work in the following conditional space $(\Omega, \mathcal{F}_\Omega, \mathbb{P}[\cdot | C])$, being $C = \{\omega \in \Omega : 0 < v_0(\omega) < v(\omega) < k(\omega)\} \in \mathcal{F}_\Omega$.

In the setting of RDEs, beyond solving the corresponding equation or model, it's vital to derive key probabilistic properties of the solution, including the expectation and variance functions. Moreover, computing its finite distributions (fidis), particularly focusing on the first PDF (1-PDF) denoted by $f_1(v, t)$, is highly desirable. Integration of this function enables the calculation of expectation, variance, and all one-dimensional statistical moments, that is,

$$\mathbb{E}[v(t, \omega)^n] = \int_{-\infty}^{\infty} v^n f_1(v, t) dv, \qquad n = 1, 2, \dots, \tag{11}$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator. Notably, for $n = 1$ and $n = 2$, we obtain the mean and the variance, respectively,

$$\mu_v(t) = \mathbb{E}[v(t, \omega)] = \int_{-\infty}^{\infty} v f_1(v, t) dv, \tag{12}$$

$$\sigma_v^2(t) = \mathbb{V}[v(t, \omega)] = \int_{-\infty}^{\infty} v^2 f_1(v, t) dv - (\mathbb{E}[v(t, \omega)])^2. \tag{13}$$

Additionally, the 1-PDF offers crucial probabilistic insights, such as the probability that the solution falls within a specific set of interest, denoted as $[v_1(\hat{t}), v_2(\hat{t})]$, for any fixed but arbitrary $\hat{t} \geq 0$,

$$\mathbb{P}(\{\omega \in \Omega : v(\hat{t}, \omega) \in [v_1(\hat{t}), v_2(\hat{t})]\}) = \int_{v_1(\hat{t})}^{v_2(\hat{t})} f_1(v, \hat{t}) dv,$$

and to construct $(1 - \alpha)\%$ probabilistic intervals (PIs), for $\alpha \in (0, 1)$,

$$\int_0^{v_1(\hat{t})} f_1(v, \hat{t}) dv = \frac{\alpha}{2} = \int_{v_2(\hat{t})}^1 f_1(v, \hat{t}) dv,$$

where

$$1 - \alpha = \mathbb{P}(\{\omega \in \Omega : v(\hat{t}, \omega) \in [v_1(\hat{t}), v_2(\hat{t})]\}).$$

This way, we obtain a complete probabilistic description of the stochastic solution at each arbitrary time instant $\hat{t}$.

The main contribution of this paper is the randomization of the hyper-logistic growth model using the so-called RDE approach and its subsequent comprehensive probabilistic analysis. This study includes the calculation of 1-PDF under very general hypotheses on the data. The study also incorporates, as a novelty, the application of the theoretical results to modeling the growth of multicellular tumor spheroids using real-world data. The application shows the superiority of the randomized hyper-logistic model with respect to the random logistic model.

The paper is organized as follows. Section 2 is divided into three parts. In Section 2.1, the 1-PDF of the solution stochastic process (10), corresponding to the randomized hyper-logistic model, is determined under very general hypotheses on the random inputs. In real-world applications of the hyper-logistic model, apart from the solution itself, there are other relevant quantities of interest, such as the time until a given population (in a wide sense) size is reached and the inflection point. In our setting, these quantities must be interpreted as random variables. This also aims to determine the PDF of these random variables in Sections 2.2 and 2.3, respectively. To face the above-mentioned goals within Section 2, we will extensively apply the random variable transformation (RVT) technique introduced at the beginning of this section. This probabilistic method has the advantage of providing a semi-explicit expression (in terms of integrals), which is very useful in practice. In Section 3, we present a real-world application of the randomized hyper-logistic model to study the dynamic growth of multicellular tumor spheroids. Subsequently, we apply the theoretical results obtained in the previous section to obtain a complete solution of the random model. Assigning adequate probability distributions to the model parameters, which are random variables, is a critical step when applying the model to real data, and two different approaches are used for this purpose. Specifically, we apply Bayesian and random least mean square (RLMS)-based techniques in Sections 3.1 and 3.2. In both subsections, we include a discussion about the obtained results via these approaches. In Section 3.3, we compare the results provided by the two aforementioned techniques for assigning the distributions of the parameters of the randomized hyper-logistic model. We finish this section by performing a comparison of the results obtained via the randomized versions of the hyper-logistic and logistic models in Section 3.4. The main conclusions of our study are drawn in Section 4.

2 | FULL PROBABILISTIC SOLUTION OF THE RANDOMIZED MODEL

The main objective of this section is to determine an exact expression to the 1-PDF of the stochastic solution of the randomized hyper-logistic model given in (10) as well as further relevant probabilistic information for its applications in real-world scenarios. Specifically, the PDFs of two interesting quantities in the analysis of population dynamics are also obtained: The PDF of the time until a given fixed population size is reached, particularly the meantime and the time of the inflection point, and the PDF of the population at the inflection point. For this purpose, the RVT technique is applied. This method enables us to obtain the joint PDF of a random vector, which is functionally dependent on another random vector whose PDF is given, in terms of the aforementioned known joint PDF, via an inverse transformation and the corresponding Jacobian. In Theorem 1, we state this theorem in its classical multidimensional form, although it should be noted that it can be adapted to many particular situations [25, 26].

Theorem 1 (RVT technique [25]). *Let $u(\omega) = (u_1(\omega), \dots, u_n(\omega))$ and $w(\omega) = (w_1(\omega), \dots, w_n(\omega))$ be n-dimensional absolutely continuous random vectors defined in a complete probability space $(\Omega, \mathcal{F}_\Omega, \mathbb{P})$, where $\omega \in \Omega$. Let $\mathbf{r} : \mathbb{R}^n \to \mathbb{R}^n$ be a one-to-one transformation of u into w, that is, $w = \mathbf{r}(u)$. Assume that $\mathbf{r}$ is continuous in u and has continuous partial derivatives with respect to u. Then, if $f_u(u_1, \dots, u_n)$ denotes the joint PDF of vector $u(\omega)$, and $\mathbf{s} = \mathbf{r}^{-1} = (s_1(w_1, \dots, w_n), \dots, s_n(w_1, \dots, w_n))$ represents the inverse mapping of the transformation $\mathbf{r} = (r_1(u_1, \dots, u_n), \dots, r_n(u_1, \dots, u_n))$, the joint PDF of vector $w(\omega)$ is given by*

$$f_w(w_1, \dots, w_n) = f_u(\mathbf{s}(w_1, \dots, w_n)) |J_n|,$$

where $|J_n|$ is the absolute value of the Jacobian, which is defined by

$$J_n = \det\left(\frac{\partial \mathbf{s}}{\partial w}\right) = \det \begin{pmatrix} \frac{\partial s_1(w_1,\ldots,w_n)}{\partial w_1} & \cdots & \frac{\partial s_n(w_1,\ldots,w_n)}{\partial w_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_1(w_1,\ldots,w_n)}{\partial w_n} & \cdots & \frac{\partial s_n(w_1,\ldots,w_n)}{\partial w_n} \end{pmatrix}.$$

As previously indicated, to alleviate the notation, we will omit the ω-dependence for random variables when evident within the context.

2.1 | Computing the 1-PDF of the solution stochastic process $v(t,\omega)$ given by (10)

Let $t > t_0$ be fixed, and using the same notation as in Theorem 1, let us define the following deterministic mapping $\mathbf{r} : \mathbb{R}^4 \to \mathbb{R}^4$, that transforms the random vector $u = (b, k, p, v_0)$, whose PDF $f_u(u) = f_u(b, k, p, v_0)$ is assumed to be known (observe that it is the joint PDF of the random parameters also previously denoted by $f_0(b, k, p, v_0)$), in the random vector $w = (w_1, w_2, w_3, w_4) = r(b, k, p, v_0)$, as follows:

$$w_1 = r_1(b, k, p, v_0) = \frac{k}{1 + \left(bp(t-t_0) + \left(\frac{k}{v_0} - 1\right)^{-p}\right)^{-\frac{1}{p}}},$$

$$w_2 = r_2(b, k, p, v_0) = k,$$

$$w_3 = r_3(b, k, p, v_0) = p,$$

$$w_4 = r_4(b, k, p, v_0) = v_0.$$

It is worth pointing out that the above transformation is not unique; nonetheless, it represents the most elementary choice: considering that the parameter b can be isolated from the equation, we keep the rest as the identity transformation. Then, isolating b, the inverse mapping is

$$b = s_1(w_1, w_2, w_3, w_4) = \frac{\left(\frac{w_2}{w_1} - 1\right)^{-w_3} - \left(\frac{w_2}{w_4} - 1\right)^{-w_3}}{w_3(t-t_0)},$$

$$k = s_2(w_1, w_2, w_3, w_4) = w_2,$$

$$p = s_3(w_1, w_2, w_3, w_4) = w_3,$$

$$v_0 = s_4(w_1, w_2, w_3, w_4) = w_4.$$

The inverse transformation is well defined since, by hypothesis, $0 < v_0(\omega) < v(\omega) < k(\omega)$, $b(\omega) > 0$, and $0 < p(\omega) < 1$ with probability 1 (w.p. 1). Finally, we calculate the Jacobian

$$|J| = \frac{w_2}{w_1^2(t-t_0)}\left(\frac{w_2}{w_1} - 1\right)^{-w_3-1},$$

which is positive w.p. 1 since, as previously pointed out, in the conditional space, the following condition $w_2(\omega) = k(\omega) > v(\omega) = w_1(\omega)$ holds w.p. 1. So, applying Theorem 1, the PDF of the random vector w is given by

$$f_w(w) = f_u\left(\frac{\left(\frac{w_2}{w_1} - 1\right)^{-w_3} - \left(\frac{w_2}{w_4} - 1\right)^{-w_3}}{w_3(t-t_0)}, w_2, w_3, w_4\right)\frac{w_2}{w_1^2(t-t_0)}\left(\frac{w_2}{w_1} - 1\right)^{-w_3-1}.$$

Then, marginalizing with respect to the first component, which is the solution, and taking $t > t_0$ arbitrarily, the 1-PDF of the solution stochastic process is

$$f_1(v, t) = \int_0^\infty \int_0^1 \int_0^\infty f_u\left(\frac{\left(\frac{k}{v} - 1\right)^{-p} - \left(\frac{k}{v_0} - 1\right)^{-p}}{p(t - t_0)}, k, p, v_0\right) \frac{k}{v^2(t - t_0)}\left(\frac{k}{v} - 1\right)^{-p-1} dv_0 dp dk. \tag{14}$$

In the case of independence among the input parameters, as the joint PDF is equal to the product of the marginals, the expression for the 1-PDF would be

$$f_1(v, t) = \int_0^\infty \int_0^1 \int_0^\infty f_b\left(\frac{\left(\frac{k}{v} - 1\right)^{-p} - \left(\frac{k}{v_0} - 1\right)^{-p}}{p(t - t_0)}\right) f_k(k) f_p(p) f_{v_0}(v_0) \frac{k}{v^2(t - t_0)}\left(\frac{k}{v} - 1\right)^{-p-1} dv_0 dp dk. \tag{15}$$

Remark 2. From a computational standpoint, it is interesting to observe that expressions (14) and (15) can be represented via an expectation with respect to v_0, p, and k,

$$f_1(v, t) = \mathbb{E}_{v_0, p, k}\left[f_b\left(\frac{\left(\frac{k}{v} - 1\right)^{-p} - \left(\frac{k}{v_0} - 1\right)^{-p}}{p(t - t_0)}\right) \frac{k}{v^2(t - t_0)}\left(\frac{k}{v} - 1\right)^{-p-1}\right]. \tag{16}$$

In the first case, the weaker hypothesis that (v_0, p, k) has a joint PDF can be assumed, while in the second case, these random variables are independent. Therefore, in both cases, we can make use of Monte Carlo simulations to obtain the 1-PDF $f_1(v, t)$.

Remark 3. Later, we will deal with the case where v_0 is a deterministic value, say $v_0 = v_0^*$. Then, its PDF is defined by means of a Dirac Delta function $\delta(v_0 - v_0^*)$. In such a case, the expression (15) writes

$$f_1(v, t) = \int_0^\infty \int_0^1 \int_0^\infty f_b\left(\frac{\left(\frac{k}{v} - 1\right)^{-p} - \left(\frac{k}{v_0} - 1\right)^{-p}}{p(t - t_0)}\right) f_k(k) f_p(p) \delta(v_0 - v_0^*) \frac{k}{v^2(t - t_0)}\left(\frac{k}{v} - 1\right)^{-p-1} dv_0 dp dk. \tag{17}$$

2.2 | PDF of time until a given fixed population size is reached

A key magnitude in population dynamics is the time until a given fixed amount of the population, $\hat{v}$, is reached. An expression of this quantity can be determined from the deterministic solution (8), taking $v = \hat{v}$ and isolating the time t. Assuming that the input parameters are random variables, the random parameter that represents the time until a given population $\hat{v}$ is reached is

$$t(\omega) = \frac{\left(\frac{k(\omega)}{\hat{v}} - 1\right)^{-p(\omega)} - \left(\frac{k(\omega)}{v_0(\omega)} - 1\right)^{-p(\omega)} + b(\omega)p(\omega)t_0}{b(\omega)p(\omega)}.$$

Let $u = (b, k, p, v_0)$ be the random vector of input parameters, and $f_u(u) = f_u(b, k, p, v_0) = f_0(b, k, p, v_0)$ its PDF, which is assumed to be known. In this case, in agreement with the earlier reasoning, we define the following deterministic mapping to apply the RVT method to obtain the PDF of $t(\omega)$ in terms of $f_u(u)$,

$$w_1 = r_1(b, k, p, v_0) = t = \frac{\left(\frac{k}{\hat{v}} - 1\right)^{-p} - \left(\frac{k}{v_0} - 1\right)^{-p} + bpt_0}{bp},$$

$$w_2 = r_2(b, k, p, v_0) = k,$$

$$w_3 = r_3(b, k, p, v_0) = p,$$

$$w_4 = r_4(b, k, p, v_0) = v_0.$$

The inverse mapping and the Jacobian are, respectively,

$$b = s_1(w_1, w_2, w_3, w_4) = \frac{\left(\frac{w_2}{\hat{v}} - 1\right)^{-w_3} - \left(\frac{w_2}{w_4} - 1\right)^{-w_3}}{w_3(w_1 - t_0)},$$

$$k = s_2(w_1, w_2, w_3, w_4) = w_2,$$

$$p = s_3(w_1, w_2, w_3, w_4) = w_3,$$

$$v_0 = s_4(w_1, w_2, w_3, w_4) = w_4,$$

$$|J| = \left| \frac{-1}{(w_1 - t_0)^2 w_3} \left(\left(\frac{w_2}{\hat{v}} - 1\right)^{-w_3} - \left(\frac{w_2}{w_4} - 1\right)^{-w_3} \right) \right| = \frac{1}{(w_1 - t_0)^2 w_3} \left(\left(\frac{w_2}{\hat{v}} - 1\right)^{-w_3} - \left(\frac{w_2}{w_4} - 1\right)^{-w_3} \right) > 0,$$

which are well defined when $0 < v_0(\omega) < \hat{v} < k(\omega)$ w.p. 1. This relation is fulfilled by hypothesis, since as $0 < v_0(\omega) < v(t, \omega) < k(\omega)$ for $t > t_0$, then $0 < v_0(\omega) < \hat{v} < k(\omega)$. The joint PDF of the random vector w is

$$f_w(w) = f_u \left(\frac{\left(\frac{w_2}{\hat{v}} - 1\right)^{-w_3} - \left(\frac{w_2}{w_4} - 1\right)^{-w_3}}{w_3(w_1 - t_0)}, w_2, w_3, w_4 \right) \frac{1}{(w_1 - t_0)^2 w_3} \left(\left(\frac{w_2}{\hat{v}} - 1\right)^{-w_3} - \left(\frac{w_2}{w_4} - 1\right)^{-w_3} \right).$$

Calculating the marginal distribution of the random variable w_1, we directly obtain the PDF of the time until a fixed amount of the population $\hat{v}$ is reached

$$f(t; \hat{v}) = \int_0^\infty \int_0^1 \int_0^\infty f_u \left(\frac{1}{p(t - t_0)} \left(\left(\frac{k}{\hat{v}} - 1\right)^{-p} - \left(\frac{k}{v_0} - 1\right)^{-p} \right), k, p, v_0 \right)$$
$$\times \frac{1}{(t - t_0)^2 p} \left(\left(\frac{k}{\hat{v}} - 1\right)^{-p} - \left(\frac{k}{v_0} - 1\right)^{-p} \right) dv_0 dp dk. \tag{18}$$

As before, if the model parameters are independent random variables, the PDF expression becomes

$$f(t; \hat{v}) = \int_0^\infty \int_0^1 \int_0^\infty f_b \left(\frac{1}{p(t - t_0)} \left(\left(\frac{k}{\hat{v}} - 1\right)^{-p} - \left(\frac{k}{v_0} - 1\right)^{-p} \right) \right) f_k(k) f_p(p) f_{v_0}(v_0)$$
$$\times \frac{1}{(t - t_0)^2 p} \left(\left(\frac{k}{\hat{v}} - 1\right)^{-p} - \left(\frac{k}{v_0} - 1\right)^{-p} \right) dv_0 dp dk. \tag{19}$$

In models of this nature, the calculation of specific times, notably the meantime, $t_{\frac{1}{2}}$, and the time at the inflection point, t_{inf}, are of particular relevance. In the deterministic theory, the first time occurs when $v = \frac{k}{2}$, while the second when $v = v_{inf} = \frac{k(1-p)}{2}$. Both times can be computed by substituting the corresponding v in (8) and isolating t (see Remark 4). Now, randomizing the input parameters and applying the RVT method, we obtain, respectively,

$$f(t_{\frac{1}{2}}) = \int_0^\infty \int_0^1 \int_0^\infty f_u \left(\frac{1}{p(t - t_0)} \left(1 - \left(\frac{k}{v_0} - 1\right)^{-p} \right), k, p, v_0 \right) \frac{1}{(t - t_0)^2 p} \left(1 - \left(\frac{k}{v_0} - 1\right)^{-p} \right) dv_0 dp dk,$$

$$f(t_{inf}) = \int_0^\infty \int_0^1 \int_0^\infty f_u \left(\frac{1}{p(t - t_0)} \left(\left(\frac{2}{1-p} - 1 \right)^{-p} - \left(\frac{k}{v_0} - 1 \right)^{-p} \right), k, p, v_0 \right)$$
$$\times \frac{1}{(t - t_0)^2 p} \left(\left(\frac{2}{1-p} - 1 \right)^{-p} - \left(\frac{k}{v_0} - 1 \right)^{-p} \right) dv_0 dp dk,$$

which are well defined in the conditional space previously indicated.

Remark 4. It can be calculated exact expressions for $t_{\frac{1}{2}}$ and t_{inf}:

$$t_{\frac{1}{2}} = \frac{1 - \left(\frac{k}{v_0} - 1 \right)^{-p} + bpt_0}{bp}, \qquad t_{inf} = \frac{\left(\frac{2}{1-p} - 1 \right)^{-p} - \left(\frac{k}{v_0} - 1 \right)^{-p} + bpt_0}{bp}.$$

It is interesting to observe that, under our assumption $0 < p < 1$, one gets

$$t_{\frac{1}{2}} - t_{inf} = \frac{1 - \left(\frac{2}{1-p} - 1 \right)^{-p}}{bp} > 0.$$

This justifies that in the hyper-logistic model, the inflection point does not coincide with half of the carrying capacity; even more, the former is always smaller than the latter, which is a main difference with respect to the logistic model. This property also holds in our random setting since we assume $0 < p < 1$ w.p. 1.

2.3 | PDF of the population size at the inflection point, v_{inf}

We have seen that the inflection point, v_{inf}, is a distinctive feature of the hyper-logistic model with respect to the logistic model. We here characterize the inflection point of the hyper-logistic model in the random setting by computing its PDF. To this end, we first obtain its explicit expression.

$$\frac{d^2 v}{dt^2} = \frac{b^2}{k^2} v^{1-2p} (k - v)^{1+2p} (k(1 - p) - 2v) = 0 \Rightarrow v_{inf} = \frac{k(1 - p)}{2} > 0.$$

The positiveness of v_{inf} is obtained due to $k > 0$ and $0 < p < 1$.

Now, by applying the RVT method, one obtains the PDF of v_{inf},

$$f(v_{inf}) = \int_0^1 f_{k,p} \left(\frac{2v_{inf}}{1-p}, p \right) \frac{2}{1-p} dp, \qquad v_{inf} > 0, \tag{20}$$

being $f_{k,p}(k, p)$ the marginal distribution of the joint PDF $f_0(b, k, p, v_0)$ of the input parameters.

3 | APPLICATION

This section aims to show how the previous theoretical results developed in Section 2 can be applied to model a real-world problem. We shall deal with the modeling of a tumor growth assuming that its dynamics adjusts a sigmoid curve emanated from the RDE (9), where $v(t)$ denotes the size of the tumor at the time instant t. Specifically, we study the growth of multicellular tumor spheroids, that is, how the average spheroid volume (measured in $10^9 \, \mu m^3$) varies over time (in days). To conduct our analysis, we will use the data collected in Table A1 (see Appendix A). For each of the 45 observations, we have the mean volume of 50 individual spheroids at different time instants. The data set can be found in [3], which has been previously published in [27, 28]. It is worth pointing out that further information about how the spheroid measurements were obtained, their culture and so forth, can be found within these references.

To obtain a complete probabilistic description of the solution of the randomized model given in (10) at any given point in time, we shall determine its 1-PDF. According to the results obtained in Section 2.1, we first need to assign an adequate
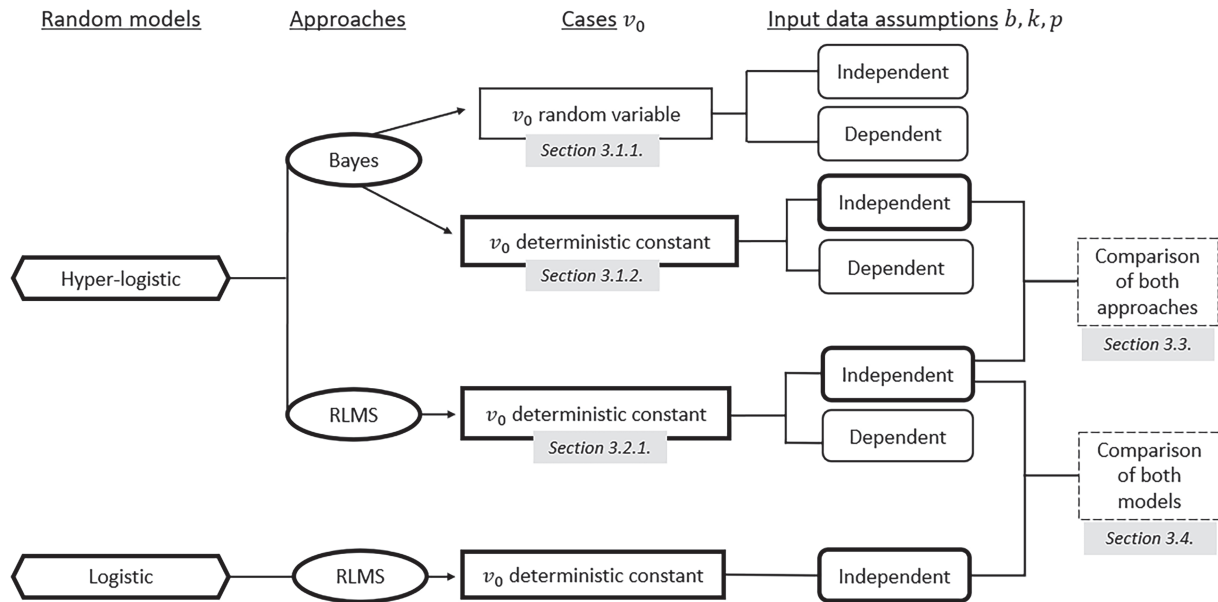
FIGURE 1 Scheme of Section 3. The boxes that we will use to make comparisons between methods and models are highlighted.

joint PDF for the model parameters, $f_0(b, k, p, v_0)$. This step is crucial in practice since such a distribution should capture the uncertainty of the data. In this paper, we will address this challenge using two approaches: the first one is based on Bayesian inference [18, 29], and the second one relies on the RLMS technique [18, 30]. As the application of our theoretical results to model the growth of multicellular tumor spheroids shall require distinguishing several cases when applying the Bayesian and RLMS approaches, for the sake of clarity, Figure 1 shows the organization of our subsequent analysis.

3.1 | First approach: Bayesian inference

As mentioned in Section 1, the observed data involve uncertainties. Therefore, let y_i be the observations of the mean volume of spheroids obtained for each time instant t_i, $i = 0, \dots, 44$. The variability or noise in the observations is defined as an additional component as follows:

$$y_i = v(t_i) + \epsilon_i, \qquad i = 0, \dots, 44,$$

where $\epsilon_i \sim N_T(0, \sigma^2)$ are independent and identically distributed (i.i.d.) random variables distributed as a truncated normal distribution with $T = [0, \infty[$ and $v(t_i) \equiv v(t_i, \omega)$ is the solution given in (10) at $t = t_i$, $i = 0, \dots, 44$.

The solution, being a stochastic process, yields a PDF at each time t. Thus, the mean or expectation of these random functions characterizes the probabilistic fit to the data, which captures its trend. Therefore, each observation is the mean plus a normal (Gaussian) perturbation centered on it with variability that takes positive values, hence the truncation. Alternatively, the Gamma distribution could serve as a suitable option due to its strictly positive values. Although the results are similar, the use of the normal distribution is computationally advantageous.

The model parameters are the values used in the model to represent and explain the patterns in the observed data. As a result, inherent uncertainty within the data manifests intrinsically during the model fitting procedure. The Bayesian approach provides a more complete way of quantifying uncertainty in model outcomes, as it takes into account not only the variability in the observed data but also that of the model parameters, treating them as random variables with associated probability distributions. Using Bayes' theorem, we obtain the posterior probability distribution of the model parameters given the available information, which is based on updating previous information about the parameters as new data are obtained, that is,

$$p(\theta|\mathbf{y}) \propto p(\mathbf{y}|\theta)p(\theta),$$

$$p(\mathbf{y}|\theta) = \prod_{i=0}^{44} N(y_i|v(t_i), \sigma^2), \tag{21}$$

where

$$\mu_i = v(t_i) = \frac{k}{1 + \left(bp(t_i - t_0) + \left(\frac{k}{v_0} - 1 \right)^{-p} \right)^{-\frac{1}{p}}}.$$

So far, for the sake of generality in our mathematical analysis, we have assumed that all model parameters are random variables. However, in practice, it might be more realistic for some of them to be deterministic. Our previous results are consistent since deterministic constants can be treated as degenerated random variables, that is, random variables whose probability is concentrated at the corresponding deterministic point, so being its distribution, the Dirac Delta function centered at that point. As discussed in detail in the next two subsections, this is the case for the initial condition v_0.

3.1.1 | v_0 as random variable

Given the generality of the theoretical results and their intended application, we begin by postulating that all model parameters are random variables.

To address the inferential process within the Bayesian framework, assigning suitable prior distributions to the parameters is crucial. We have opted for univariate prior distributions for each model parameter, since to the best of our knowledge, no literature is available about the parameters. Therefore, the prior distribution, say $p(\theta)$, will be determined by the product of the prior distributions of model parameters. Notably, parameters k, v_0, and b are defined as random variables with strictly positive domains, while p lies within the interval $(0, 1)$. Additionally, we have graphically explored the behavior of the solution of the differential equation for different combinations of parameter values so that the solution reflects the trend of the sigmoidal function. Consolidating our findings, we established the following non-informative prior distributions:

$$b \sim \text{Unif}(0.01, 10), \ k \sim \text{Unif}(1, 200), \ p \sim \text{Unif}(0.01, 1),$$
$$v_0 \sim \text{Unif}(0.001, 1), \ \tau \sim \text{Ga}(0.5, 0.5). \tag{22}$$

The parameter τ comes from how the Normal distribution is defined in WinBUGS software, $N(\mu, \tau)$, where $\tau = \frac{1}{\sigma^2}$, which was used to carry out the Bayesian analysis in this work. A commonly employed non-informative prior distribution for τ is a Gamma distribution, parameterized to maintain moderate variance values and prevent extreme assignments.

The posterior distribution of model parameters is estimated through Bayesian sampling techniques, such as Gibbs sampling, a type of Markov chain Monte Carlo (MCMC) algorithm. In this study, all analyses were conducted using RStudio software, utilizing the `R2WinBUGS` package to interface with WinBUGS software, where the mentioned algorithm is employed. The MCMC method involves building a Markov chain whose stationary distribution coincides with the target distribution. Indeed, the chain resulting from Gibbs sampling is a sample of the desired posterior distribution. The kernel density estimation technique was applied to obtain the densities. The Gaussian distribution was chosen for the kernel function, and Silverman's rule was used to determine the bandwidth.

We set three chains, whose initial values are randomly generated by the WinBUGS program itself, with 5M iterations and 2M for the burn-in period. We set 2000 for the `n.thin` argument of the RStudio program's `bugs` function, indicating that the algorithm saves one value every 2000 iterations and thus avoiding autocorrelation between values. The resulting sample size is $N = 1500$ for each chain. However, to obtain the densities, the samples of all three are combined, giving a total number of 4500. From the results obtained from these configurations, we obtain the marginal posterior distributions of each parameter. Finally, the `coda` package in RStudio was employed to conduct posterior analyses of the generated Markov chains, including convergence diagnostics, to ensure their convergence to their stationary distributions. In addition, these include the cross-correlations between different variables in MCMC output. A moderate and strong relationship is observed between them (0.65, -0.55, -0.75, and -0.8 between k and p, b and p, b and k, and p and v_0, respectively). This may be due to a combination of the intrinsic algebraic structure of the model and complex biological phenomena.

Once the posterior distribution of the parameters has been drawn, we can employ the theoretical outcomes since they all depend on the PDF $f_0(b, k, p, v_0)$. It is noteworthy that we can distinguish between the following two cases:

- *Independent*: assuming independence among the random variables b, k, p, and v_0. Therefore, the PDF $f_0(b, k, p, v_0)$ will be the product of the individual marginal distributions, that is, $f_B(b) f_K(k) f_P(p) f_{V_0}(v_0)$.
- *Dependent*: it occurs due to the dependencies between the variables, as highlighted above. This dependency structure is effectively captured using copulas, which offer a flexible framework for modeling complex dependency patterns. By
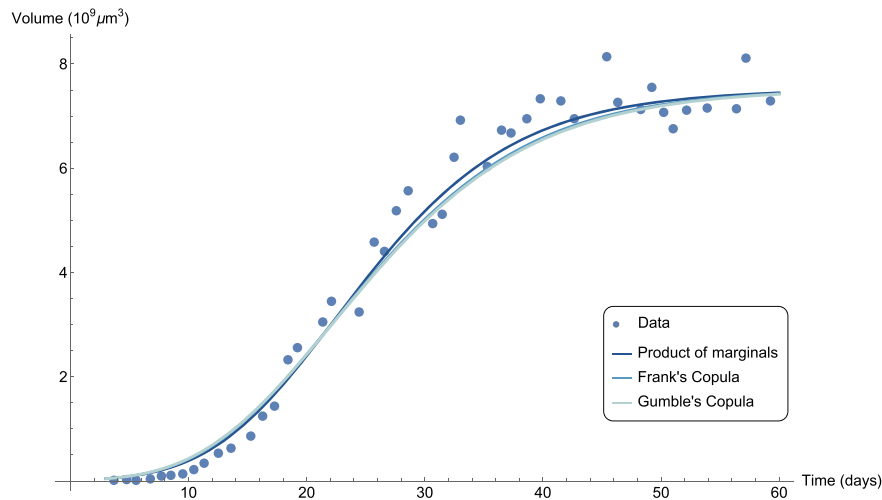
FIGURE 2 Probabilistic fits (solid lines) to the data (points) collected in Table A1, representing the expected value of the solution (10) of the randomized hyper-logistic model. We consider two cases: first, when the four parameters b, k, p, v_0 are considered independent random variables, giving rise to PDF derived from the product of marginal posterior distributions obtained using the Bayesian approach. Secondly, when these parameters present a dependence structure modeled via Frank's and Gumbel's copulas that determine the corresponding joint PDFs. [Colour figure can be viewed at wileyonlinelibrary.com]

encapsulating the joint behavior of the variables, the copula function yields the joint PDF of the random variables, $f_0(b, k, p, v_0)$. Among the options, Frank's and Gumbel's copulas emerge as the most appropriate for our context. The former is used to model both positive and negative dependencies and is capable of capturing a wide variety of dependency patterns. The latter is also commonly used to model both types of dependencies and possesses the ability to capture some asymmetry. Both copulas are characterized by a single parameter, denoted as a for Frank's copula and c for Gumbel's copula, which will be adjusted using optimization techniques as explained in Section 3.2. Specifically, we have obtained $a = 0.1$ and $c = 1.5$, respectively.

In this way, we illustrate that both cases can be implemented and applied in practice. First, we proceed to fit the data. In Figure 2, we can observe the probabilistic fit to the data for both cases, where a considerable similarity is perceived. However, it is remarkable that the mean does not effectively capture the data pattern. This observation is in line with the discussion on the variability of the sigmoidal curve's phases. In particular, and as also detailed in the paper [3], the controlled initial conditions of the first phase of cell culture experiments result in minimal variability in cell volume growth (i.e., researchers take cells at their initial growth phase, so their volume during the first phase of the curve has virtually no variability). If v_0 has variability, it negatively affects the quality of the fit since it is concentrated in larger values during the first phase. Therefore, considering it as a random variable does not contribute substantially to improving the model's ability to explain the data. Moreover, compared to the variability of the later phases, this variability is negligible, suggesting that v_0 can be considered deterministic without losing the quality of fit. This conclusion motivates the analysis presented in the next subsection.

3.1.2 | v_0 as a deterministic constant

Hereinafter, we shall consider the initial condition, $v_0^* = 0.0160808$, as a fixed parameter. This value has been estimated via a deterministic fitting using the `NonlinearModelFit` function by Wolfram Mathematica©. Substituting this value into the likelihood defined in (21), keeping the same prior distributions for the remaining random variables as in (22), and employing the same settings for the algorithm as described in Section 3.1.1, we obtain the marginal posterior distributions for each model parameter shown in Figures B1 and B2 (right panel). The `deviance`, which appears in the results obtained in Appendix B, is provided automatically by the WinBUGS software as an integral component of the Bayesian analysis process. It is defined as $-2 \log(\text{likelihood})$. From these plots, we have obtained some statistical measures to help characterize the central location and dispersion of the data. We specifically have calculated the mean and standard error (see Table B1), providing a more complete understanding of the distributions.

As for the posterior analyses, convergence is confirmed by the Gelman–Rubin test, which indicates a potential scale reduction factor of 1 for all estimated parameters (see Figure B3). This convergence is visually corroborated by trace plots
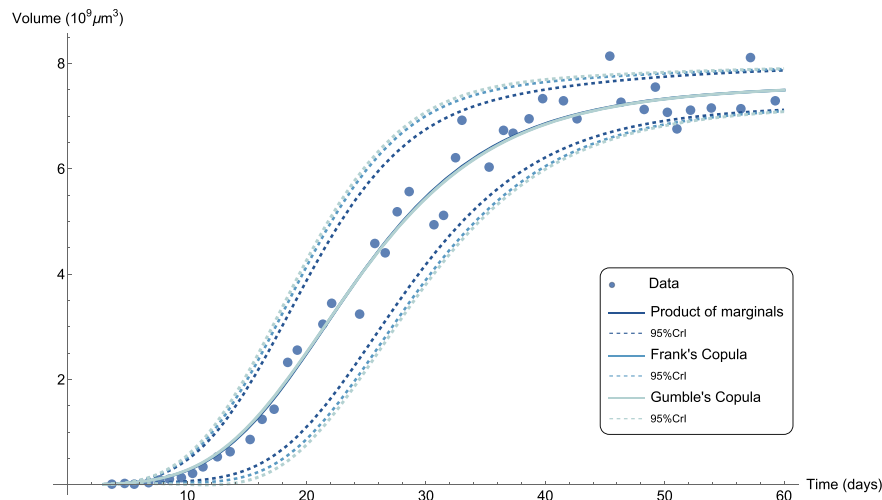
FIGURE 3 Probabilistic fit via expectations (solid lines) using the data in Table A1, together with the 95% CrIs (dashed lines) of the solution (10) derived from the random hyper-logistic model, in two different scenarios. First, when the parameters b, k and p are treated as independent random variables, and thus, the PDF is the product of their marginal posterior distributions obtained by Bayesian inference. Second, when these random parameters exhibit a dependence structure, and the joint PDF is determined using the Frank's and Gumbel's copulas. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Evaluation metrics are computed for the model parameters under two distinct scenarios.

	Product of marginals	**Frank's copula**	**Gumbel's copula**
MAPE	18.6729	19.3869	19.808
RMSE	0.357796	0.359216	0.359959

Note: First, when parameters are assumed to be independent (so, their density is the product of marginal posterior distributions). Second, when parameters are treated as dependent random variables (and their joint density is calculated via copulas). In both cases, the initial condition v_0 is treated as a deterministic constant within the Bayesian framework.

(left panel of Figures B1 and B2), which show well-mixed chains with random scatter around the mean value. As before, we also explored cross-correlation, whose strength and direction of the relationships between parameters are illustrated in Figure B4.

For the sake of consistency with the preceding section, we will address both scenarios, independence and dependence, in the application of the theoretical findings. Figure 3 illustrates the first results obtained. On the one hand, we have the probabilistic fit to the data (solid lines), which is consistent and identical for both scenarios, indicating its effectiveness. On the other hand, we determine the 95% credible intervals (CrIs) around the mean of the posterior distribution. It should be noted that similar results are obtained for the dependence case (see dashed lines for Frank's copula [$a = 0.1$] and Gumbel's copula [$c = 1.5$]); however, these CrIs are slightly wider than those obtained for the independence case (see dashed lines for the product of marginals).

In addition, we computed goodness-of-fit metrics, including the mean absolute percentage error (MAPE) and the root mean squared error (RMSE), whose values are collected in Table 2. Considering the table values and the fits with their respective 95% CrIs, we focus on the scenario where parameters are assumed independent. This decision was made to meet the primary objective of this section: to illustrate the application of the theoretical results in a real-world context. By opting for independence between parameters, we ensure a manageable computational burden, unlike the significant increase in complexity involved with copulas. Both the probabilistic fit and CrIs yield satisfactory results, effectively capturing the trend and uncertainty of the data.

3.1.3 | Results and discussion

After estimating the distributions of the model parameters with $v_0^* = 0.0160808$, we constructed the 1-PDF of the solution stochastic process for the randomized hyper-logistic model. The resulting plot is depicted on the left side of Figure 5, illustrating that the PDF is approximately concentrated around the mean volume quantity for each time instant and displays a Gaussian shape with varying degrees of kurtosis. This is due to the data's variability according to the

sigmoidal curve's different phases. In the initial phase, characterized by low variability, cells adapt to culture conditions until reaching a critical point that initiates the exponential phase. Variability in this phase, which contains the inflection point, may arise from environmental factors, intrinsic variability of the cells, and/or the response of the biological system to the experimental conditions. In the deceleration phase, cells reach the maximum carrying capacity, and/or there are factors limiting cell growth. The variability in this phase could be due to the complexity of the biological processes that regulate growth deceleration. It is also observed that the mass density shifts with time toward the PDF of the carrying capacity, k.

As an interesting supplement, we add some statistical characteristics of the 1-PDF obtained above, such as the expectation, variance, and 95% CrIs, which are listed in Table B2. Note how the expected value increases with time and stabilizes around the average carrying capacity 7.5. However, the variance increases until $t = 25$ and subsequently decreases, stabilizing at approximately 0.035. For instance, at $t = 45$ days, the expected mean volume of the spheroid is $7.15839\ 10^9\ \mu m^3$ with a 95% CrI of [6.6, 7.6].

In addition, as outlined in Section 2, the PDF of time has been determined. Figure 6, left panel, shows the time required for the spheroid to reach certain fixed sizes $\hat{v}$ that have been chosen to span the entire range of $\hat{v}$. It is worth noting that as the volume increases, the PDF tends to become more platykurtic and, therefore, less informative. This pattern is further highlighted in Table B3, where it is observed that both the mean and variance increase with volume. This phenomenon occurs because as the volume nears the carrying capacity, there are more instances where that size has been achieved.

We point out that all the computations performed to conduct the previous analysis, including the probabilistic fit, CrIs, and fit measures, were carried out using Wolfram Mathematica©.

In summary, the Bayesian approach captures the general trend and inherent variability present in the data, thereby addressing the complexity of the system. This is accomplished through posterior distributions, which encapsulate the uncertainty regarding parameter values after considering the observed data and prior information.

3.2 | Second approach: RLMS technique

As outlined earlier, to effectively apply the theoretical findings from Section 2.1, it is essential to assign suitable distributions to the model parameters. Unlike the Bayesian approach, which updates posterior distributions based on prior knowledge and observed data, the RLMS technique requires the assignment of parametric PDFs. These densities are characterized by a specific set of parameters, which are obtained by an optimization process using some fitting criterion on the sample, such as the minimization of the mean square error. The probability distributions for each parameter are chosen, taking into account their respective domains and all available information about them. Since there are several possible options for the choice of such distributions, this step constitutes a critical point of this technique but also an advantage in practice because of its flexibility.

3.2.1 | v_0 as a deterministic constant

In this work, the chosen option relies on the positivity of model parameters and validation through statistical tests. Unlike the Bayesian approach previously presented, which requires starting with a prior probability distribution reflecting pre-data knowledge about a model parameter, the RLMS technique utilizes the data itself to guide the selection of parametric probability distributions. The distribution selected for each model parameter, along with the justification for such a choice, is as follows:

- Based on various tests and the requirement for strict positivity, the parameter b typically assumes values between 0 and 1. Hence, we propose a Beta distribution for it, expressed as

$$b \sim Be(\alpha, \beta),\ \alpha > 0,\ \beta > 0.$$

- For the shape parameter, p, which is constrained within the known range $(0, 1)$, we opt for a uniform distribution, denoted as

$$p \sim Unif(\rho_1, \rho_2),\ 0 < \rho_1 < \rho_2 < 1.$$

- For the carrying capacity, k, we assume a log-normal distribution since it must be positive and asymmetric (longer right tail, and although k has a clear upper limit, if it varies at the upper end, this distribution can capture such variability).

Formally, we denote this distribution as

$$k \sim \mathrm{LN}(\mu, \sigma), \ \mu \in \mathbb{R}, \ \sigma > 0.$$

- For consistency with the Bayesian methodology explained earlier and considering the nature of our data, it is more advisable to consider the initial condition v_0 as a deterministic constant. Therefore, in this subsection, we take the same fixed numerical value as in the Bayesian approach, namely,

$$v_0^* = 0.0160808.$$

To determine the parameters for each proposed probability distribution, α, β, μ, σ, ρ_1, and ρ_2, the following steps have been carried out using Wolfram Mathematica©. Initially, a nonlinear deterministic fit of the solution of the hyper-logistic equation defined in (8) is applied to the data. This yielded point estimates of the parameters and their associated standard errors (see Table C1). Subsequently, we performed an initial search to estimate the parameters of the probability distributions assigned to the model's random inputs, α^0, β^0, μ^0, σ^0, ρ_1^0, and ρ_2^0. This involved ensuring that their expectations aligned with the point estimates in Table C1, and their variances matched the squared standard errors in the same table. Notably, for p, we considered a smaller variance to enhance fitting accuracy. Specifically,

$$b \sim \mathrm{Be}(\alpha^0, \beta^0) : \mathbb{E}[b] = 0.178912, \ \mathbb{V}[b] = 0.0162014^2,$$
$$k \sim \mathrm{LN}(\mu^0, \sigma^0) : \mathbb{E}[k] = 7.60358, \ \mathbb{V}[k] = 0.219763^2,$$
$$p \sim \mathrm{Unif}(\rho_1^0, \rho_2^0) : \mathbb{E}[p] = 0.192177, \ \mathbb{V}[p] = 0.0001^2.$$

We obtain the first estimate for the parameters from the exact expressions of the mean and variance of each distribution. From these starting values, the next step is to find the optimal parameter estimates, α, β, μ, σ, ρ_1, and ρ_2, by solving the following optimization program:

$$\min_{\alpha, \beta, \mu, \sigma, \rho_1, \rho_2} \sum_{i=0}^{44} (y_i - \mathbb{E}[v(t_i, \omega; \alpha, \beta, \mu, \sigma, \rho_1, \rho_2)])^2, \tag{23}$$

which consists of minimizing the squared errors between the observed values y_i, listed in Table A1, and the expectation of the solution at each time instant t_i. The expectation is calculated using (12), where $f_1(v, t)$ is defined in (17). Given the parametric nature of this approach, the PDFs involved in (17) are expressed as follows:

$$f_b\left(\frac{\left(\frac{k}{v}-1\right)^{-p} - \left(\frac{k}{v_0}-1\right)^{-p}}{p(t-t_0)}\right) = \frac{\left(\frac{\left(\frac{k}{v}-1\right)^{-p} - \left(\frac{k}{v_0}-1\right)^{-p}}{p(t-t_0)}\right)^{\alpha-1} \left(1 - \frac{\left(\frac{k}{v}-1\right)^{-p} - \left(\frac{k}{v_0}-1\right)^{-p}}{p(t-t_0)}\right)^{\beta-1}}{\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}},$$

$$f_k(k) = \frac{1}{k\sigma\sqrt{2\pi}} e^{-\frac{(\ln k - \mu)^2}{2\sigma^2}}, \qquad f_p(p) = \begin{cases} \frac{1}{\rho_2 - \rho_1}, & p \in (\rho_1, \rho_2), \\ 0, & \text{otherwise}, \end{cases}$$

where Γ is the Gamma function. Bear in mind that in (17), $f_{v_0}(v_0) = \delta(v_0 - v_0^*)$ represents a D irac Delta function centered at 0.0160808, reflecting our assumption that $v_0^* = 0.0160808$. To minimize the objective function (23) in the multidimensional space, a nonlinear optimization algorithm is applied. More specifically, we employed the Nelder–Mead method, which explores the solution space around the initial estimates of the parameters α^0, β^0, μ^0, σ^0, ρ_1^0, and ρ_2^0. Consequently, the following local optimum values are obtained:

$$\alpha = 105.518, \ \beta = 453.518, \ \mu = 2.01784, \ \sigma = 0.0299154, \ \rho_1 = 0.149667, \ \rho_2 = 0.197989.$$

The resulting quadratic error is 5.68149, and its square root is 2.38359. After determining the optimal parameters of the distributions, we computed their main statistical characteristics, such as mean, variance, and standard error, as detailed in Table C2.

	Product of marginals	**Frank's copula**	**Gumbel's copula**
MAPE	14.9529	15.2872	15.216
RMSE	0.356311	0.355721	0.356895

TABLE 3 Model evaluation metrics are assessed under both independence (product of the proposed distributions with their respective optimal values obtained) and dependence (utilizing copulas) between parameters within the RLMS approach.
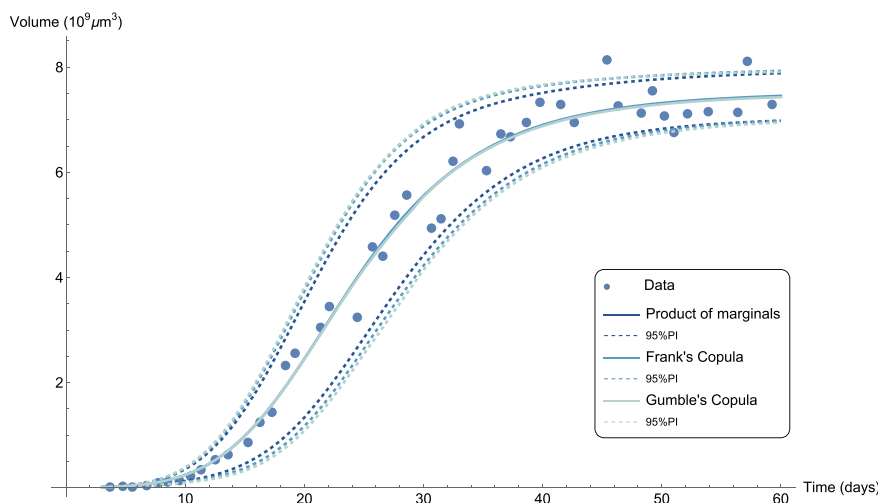

FIGURE 4 Probabilistic fit via expectations (solid lines) using the data in Table A1, together with the 95% PIs (dashed lines) of the solution (10) derived from the random hyper-logistic model, in two different scenarios. First, when the parameters b, k, and p are treated as independent random variables, and thus, the PDF is the product of the proposed distributions obtained by RLMS technique. Second, when these random parameters exhibit a dependence structure, and the joint PDF is determined using the Frank's and Gumbel's copulas. [Colour figure can be viewed at wileyonlinelibrary.com]

To complement the analysis conducted within the RLMS framework and maintain alignment with the study carried out using Bayesian methodology, we will also examine scenarios involving both independence and dependence.

Copula parameter estimation

The joint PDF using copulas is constructed with Wolfram Mathematica©'s `CopulaDistribution` function, incorporating Frank's and Gumbel's copulas, along with their respective parameters a and c, and the probability distributions of the model parameters, obtained using the Bayesian and the RLMS approaches, as arguments. By means of a copula, the dependence between model parameters is taken into account. The copula's parameter is determined in a similar way to the adjustment of the parameters of the distributions proposed above using the optimization program (23). However, in this fitting case, the expectation of the solution depends only on the copula's parameter, $\mathbb{E}[v(t_i, \omega; a)]$ for Frank's copula and $\mathbb{E}[v(t_i, \omega; c)]$ for Gumbel's copula, using $f_1(v, t)$ as given in (14).

3.2.2 | Results and discussion

Once we have confirmed that the RLMS technique yields satisfactory fits to the data in all cases, we will proceed with probabilistic fitting. Regarding the intervals, it is important to note that 95% PIs are established under the RLMS approach. Similar results to the Bayesian scenario are obtained for the case of copulas, as evidenced in Table 3. Furthermore, the PIs are slightly wider in the copula case, particularly during the exponential growth phase, compared to when independence is assumed (see Figure 4). However, it is worth pointing out that this difference is less pronounced compared to the Bayesian approach.

According to the results obtained under the Bayesian approach, we select the scenario where the model parameters are assumed to be independent random variables. On the right side of Figure 5, we show the 1-PDF then obtained. It can observed that it exhibits a behavior similar to the one described in the first approach. Additionally, Table C3 highlights the ascending trend of the expected value over time, approaching a mean value of the carrying capacity of approximately 7.43. As for the variance, it experiences an increase until $t = 25$, after which it decreases to stabilize around approximately 0.05. Everything is in full agreement with the results obtained throughout the probabilistic fitting process.
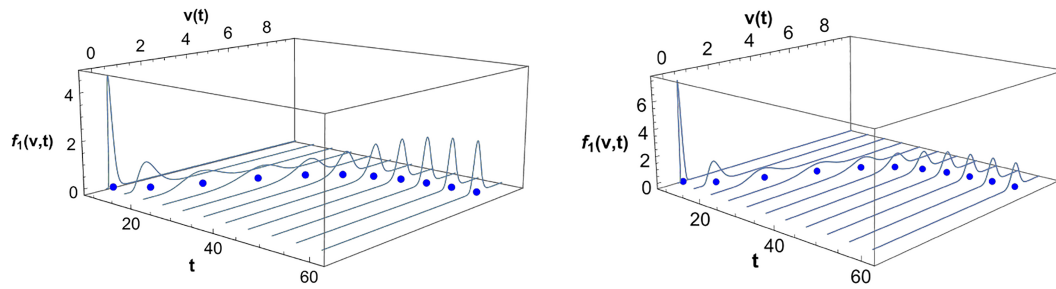
FIGURE 5 Left side: Graphical representation of the 1-PDF, $f_1(v, t)$, as defined in (17), derived from the solution of model (9) using the Bayesian approach. This is achieved by applying the product of marginal posterior distributions for the densities of the random model inputs. Right side: The same 1-PDF, now obtained using densities of the random model inputs with parameters estimated via the optimization program (23) through the RLMS approach. In both cases, $v_0^* = 0.0160808$, and the 1-PDF is shown at the time instants $t \in \{10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60\}$. [Colour figure can be viewed at wileyonlinelibrary.com]
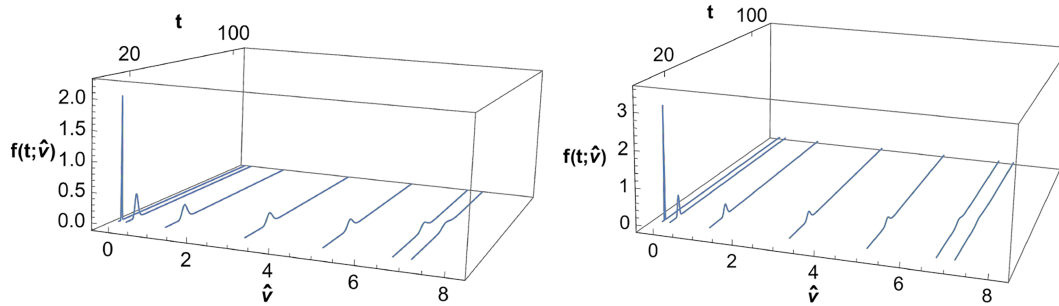

FIGURE 6 Left side: Graphical representation of the PDF, $f(t; \hat{v})$, given in (19), representing the time until different fixed values of volume $\hat{v} \in \{0.0273598, 0.218878, 1.24487, 3.24213, 5.11628, 6.73051, 7.15458\}$ are reached. The densities for the model parameters are obtained by the Bayesian approach. Right side: The same PDFs; now, we use densities for the model parameters obtained by the RLMS approach. In both cases, $v_0^* = 0.0160808$. [Colour figure can be viewed at wileyonlinelibrary.com]

Similarly, the PDF of time has been computed, which is plotted on the right side of Figure 6. As in the first approach, with the spheroid size increasing, the shape of the time density becomes more flattened. This trend is also evident in Table C4, where we observe an increase in both the mean and variance values as the fixed value of $\hat{v}$ increases.

3.3 | Comparison of both approaches

Examining the results obtained from both methodologies for estimating the densities of random model inputs when the initial condition v_0 is treated as a deterministic constant and b, k, and p are assumed to be independent variables, it is evident that both the Bayesian and RLMS approaches accurately capture the expected value, yielding closely matching results. However, a minor discrepancy exists in the 95% intervals, as the Bayesian approach yields CrIs slightly wider than the PI's obtained via the RLMS approach. This fact is particularly observed at the onset of the second phase of the sigmoidal curve. This fact could be explained because of the inherent difference between these two methodologies. Indeed, in the Bayesian approach, prior distributions are iteratively updated with observed data, resulting in a posterior distribution that consider both variability and dependency structure. In the RLMS technique, parametric probability distributions are selected based on the physical interpretation of the model inputs, and parameters are estimated using a criterion that ignores the presence of a dependence structure between variables.

When comparing the metrics, we see that the RMSE values are almost identical for both methods (Bayesian method: 0.3577 and RLMS method: 0.3563). However, there is a notable difference in the MAPE values (18.67 and 14.95, respectively). This variation is due to the higher sensitivity of the MAPE to values close to zero, which is consistent with the results obtained in both fits.

The results collected in Tables B2 and C3 outline the statistical attributes of the 1-PDFs. They reinforce the agreement between the Bayesian and RLMS methodologies. Notably, the mean values of the PDFs exhibit gradual increases over time, converging to 7.49 and 7.43, respectively. Similarly, the variances stabilize around 0.03 and 0.05, respectively, underscoring the consistency of the results provided by both approaches.

Concerning the time PDFs, consistent behavior is observed in both cases, resulting in similar mean values of the PDFs for each fixed quantity. However, it is noteworthy that the variance for larger quantities is notably higher in the second approach.

In summary, although both Bayesian and RLMS approaches effectively estimate expected values, they differ in their treatment of variability and dependence structure. Hence, the RLMS approach provides slightly smaller metrics and narrower intervals than the Bayesian technique.

3.4 | Comparing the results using hyper-logistic versus logistic models

As indicated in Section 1, the hyper-logistic model overcomes some drawbacks of the logistic model, particularly its less flexibility to describe the dynamics in certain growth phenomena. This issue is strongly related to the fact that in the logistic model, the inflection point must coincide with half of the carrying capacity, while in the hyper-logistic model, this point depends on the new parameter p, which provides greater flexibility. This section is devoted to illustrating this key point within the setting of the above biological application and considering the randomized version of both models. To ensure a fair comparison, we use the same parametric probability distributions as those employed in the RLMS approach for the b and k parameters of the randomized hyper-logistic model. Furthermore, we assume independence among the model's parameters and maintain the same value for $v_0^* = 0.0160808$. To carry out the probabilistic fitting using the random logistic model, we first calculated the 1-PDF of its solution using the RVT technique. This results in the following expression:

$$f_1(v, t) = \int_0^\infty \int_0^\infty f_B(b) f_K \left(\frac{e^{-b(t-t_0)} - 1}{\frac{e^{-b(t-t_0)}}{v_0} - \frac{1}{v}} \right) \delta(v_0 - v_0^*) \left| \frac{(1 - e^{-b(t-t_0)})v_0^2}{(ve^{-b(t-t_0)} - v_0)^2} \right| dv_0 db. \tag{24}$$

In Figure 7, one observes that the logistic model does not fit the data well since it starts the growth phase from the beginning, unlike the hyper-logistic model, which is better suited to the first phase. The superiority of the hyper-logistic model over the logistic model can also be assessed by means of MAPE and RMSE

$$\text{MAPE (hyper-logistic)} = 14.9529 < 97.3157 = \text{MAPE (logistic)},$$

$$\text{RMSE (hyper-logistic)} = 0.356311 < 0.373154 = \text{RMSE (logistic)}.$$

Finally, we calculated the expectation of the PDF of the inflection point for both models and found that the expected value of the inflection point for the hyper-logistic model is smaller (approximately 3.11) than that for the logistic model (approximately 3.68). This discrepancy may explain why the hyper-logistic model fits better in the initial phase, as its curvature changes earlier than that of the logistic model.
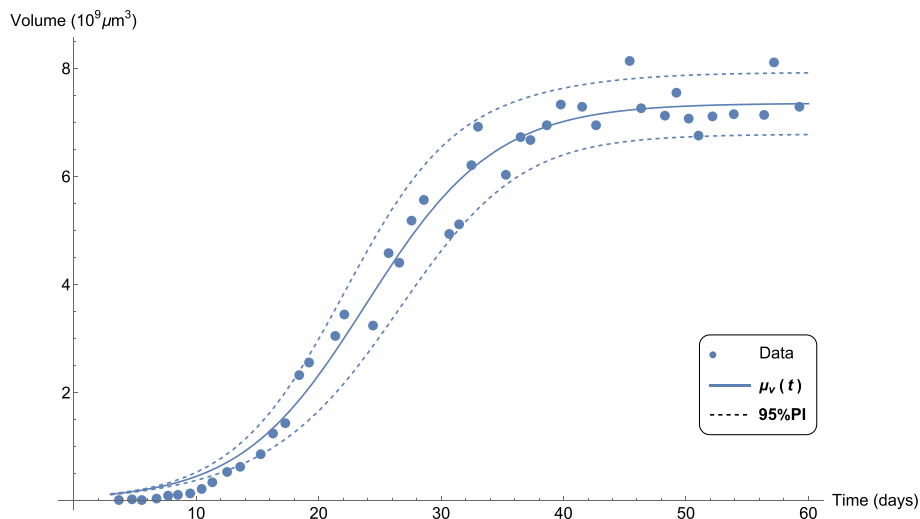

FIGURE 7 Probabilistic fit to the data collected in Table A1, representing the expectation, $\mu_v(t)$, along with the 95% PI's of the 1-PDF, expressed in (24), of the solution derived from the randomized logistic model using the RLMS method. [Colour figure can be viewed at wileyonlinelibrary.com]

4 | CONCLUSIONS

This paper presents a methodology for handling uncertainties within the hyper-logistic model, formulated as a RDE. Under this approach, all model inputs, including the initial condition and equation coefficients, are treated as random variables, resulting in a stochastic process for the solution. This stochastic modeling approach enhances realism in applications, as it considers uncertainties inherent to the studied phenomenon's complexity and measurement errors in sample data. Theoretical findings in this paper address these crucial aspects by deriving the 1-PDF of the solution to the randomized hyper-logistic model in terms of the density functions of model inputs. Our theoretical findings have broad applicability, as demonstrated through a comprehensive example that utilizes real-world data. Within this application, we discuss the critical task of assigning appropriate probability distributions to model inputs, exploring two approaches that yield consistent results while highlighting pertinent considerations for using the random hyper-logistic model. This underscores this open question within uncertainty quantification, which certainly is problem dependent. Through our present study, we aim to advance the field of differential equations with uncertainties, particularly in the mathematical modeling of growth phenomena, which finds extensive application across disciplines such as Biology, Ecology, Medicine, Economics, and beyond.

AUTHOR CONTRIBUTIONS

Juan Carlos Cortés: Conceptualization; methodology; investigation; validation; formal analysis; supervision; funding acquisition; visualization; project administration; resources. **Ana Navarro-Quiles:** Conceptualization; methodology; software; investigation; validation; formal analysis; supervision; visualization; resources. **Sorina Madalina Sferle:** Conceptualization; methodology; software; data curation; investigation; validation; formal analysis; funding acquisition; visualization; resources; writing—original draft.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this article.

ORCID

Juan Carlos Cortés https://orcid.org/0000-0002-6528-2155
Ana Navarro-Quiles https://orcid.org/0000-0003-3800-072X
Sorina Madalina Sferle https://orcid.org/0000-0002-2028-1191

REFERENCES

1. G. Petter, H. Kreft, Y. Ong, G. Zotz, and J. Sarmento Cabral, *Modelling the long-term dynamics of tropical forests: from leaf traits to whole-tree growth patterns*, Ecol. Model. **460** (2021), 109735.
2. M. Peleg and M. G. Corradini, *Microbial growth curves: what the models tell us and what they cannot*, Critical Rev. Food Sci. Nutrit. **51** (2011), 917–945.
3. M. Marusic, Z. Bajzer, J. P. Freyer, and S. Vukavlovic, *Analysis of growth of multicellular tumour spheroids by mathematical models*, Cell Prolif. **27** (1994), 73–94.
4. C. Vaghi, A. Rodallec, R. Fanciullino, J. Ciccolini, P. J. Mochel, M. Mastri, C. Poignard, J. Ebos, and S. Benzekry, *Population modeling of tumor growth curves and the reduced Gompertz model improve prediction of the age of experimental tumors*, PLOS Comput. Biol. **16** (2020), e1007178.
5. I. K. Dassios and M. T. Devine, *A macroeconomic mathematical model for the national income of a union of countries with interaction and trade*, J. Econ. Struct. **5** (2016), 18.
6. P. H. Verhulst, *Notice sur la loi que la population poursuit dans son accroissement*, Correspondance Mathámatique et Physique **10** (1838), 113–121.

7. B. Gompertz, *On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies*, Philosop. Trans. Royal Soc. London **115** (1825), 513–585.

8. V. F. Bertalanffy, *Quantitative laws in metabolism and growth*, Quart. Rev. Biol. **32** (1957), 217–231.

9. F. J. Richards, *A flexible growth function for empirical use*, J. Exper. Botany **10** (1959), 290–300.

10. M. Tong, Z. Yan, and L. Chao, *Research on a grey prediction model of population growth based on a logistic approach*, Discrete Dyn. Nature Soc. **2020** (2020), 2416840.

11. V. Mndez, M. Assaf, D. Campos, and W. Horsthemke, *Stochastic dynamics and logistic population growth*, Phys. Rev. E **91** (2015), 062133.

12. C. Y. Shen, *Logistic growth modelling of COVID-19 proliferation in China and its international implications*, Int. J. Infect. Dis. **96** (2020), 582–589.

13. A. Tsoularis and J. Wallace, *Analysis of logistic growth models*, Math. Biosci. **179** (2002), 21–55.

14. M. Peleg, M. G. Corradini, and M. D. Normand, *The logistic (Verhulst) model for sigmoid microbial growth curves revisited*, Food Res. Int. **40** (2007), 808–818.

15. J. L. Rocha and S. M. Aleixo, *Dynamical analysis in growth models: Blumberg's equation*, Discrete Contin. Dyn. Syst.-Ser. B **18** (2013), 783–795.

16. A. A. Blumberg, *Logistic growth rate functions*, J. Theor. Biol. **21** (1968), 42–44.

17. M. E. J. Turner, E. L. J. Bradley, K. A. Kirk, and K. M. Pruitt, *A theory of growth*, Math. Biosci. **29** (1976), 367–373.

18. R. C. Smith, *Uncertainty quantification: theory, implementation, and applications*, SIAM, New York, 2014.

19. E. Allen, *Modelling with Itô stochastic differential equations*, Springer, New York, 2007.

20. T. Caraballo, R. Colucci, J. López-de-la Cruz, and A. Rapaport, *A way to model stochastic perturbations in population dynamics models with bounded realizations*, Commun. Nonlinear Sci. Numer. Simul. **77** (2019), 239–257.

21. J. Heydari, C. Lawless, D. A. Lydall, and D. J. Wilkinson, *Fast Bayesian parameter estimation for stochastic logistic growth models*, Biosystems **122** (2014), 55–72.

22. F. A. Dorini, M. S. Cecconello, and L. B. Dorini, *On the logistic equation subject to uncertainties in the environmental carrying capacity and initial population density*, Commun. Nonlinear Sci. Numer. Simul. **33** (2016), 160–173.

23. J. Calatayud, T. Caraballo, J. C. Corts, and M. Jornet, *Mathemtical methods for the randomized non-autonomous Bertalanffy model*, Electron. J. Differ. Equ. **50** (2020), 1–19.

24. H. T. Banks, S. Hu, and W. Clayton Thompson, *Modeling and inverse problems in the presence of uncertainty*, Monographs and Research Notes in Mathematics, Chapman & Hall/CRC, New York, 2019.

25. T. T. Soong, *Random differential equations in science and engineering*, Academic Press, New York, 1973.

26. G. Casella and R. L. Berger, *Statistical inference*, The Duxbury Advanced Series in Statistics and Decision Sciences, Duxbury Thomson Learning, USA, 2002.

27. J. P. Freyer and R. M. Sutherland, *Regulation of growth saturation and development of necrosis in EMT6/Ro multicellular spheroids by the glucose and oxygen supply*, Cancer Res. **46** (1986), 3504–3512.

28. J. P. Freyer, *Role of necrosis in regulating the growth saturation of multicellular spheroids*, Cancer Res. **48** (1988), 2432–2439.

29. A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*, 3rd ed., CRC Press, New York, 1995.

30. J.-C. Corts, A. Navarro-Quiles, F. Santonja, and S. M. Sferle, *Statistical analysis of randomized pseudo-first/second order kinetic models. Application to study the adsorption on cadmium ions onto tree*, Chemom. Intell. Lab. Syst. **240** (2023), 104910.

APPENDIX A: DATA

TABLE A1 The data consist of 45 observations corresponding to the mean volume ($10^9 \, \mu m^3$) of 50 individual multicellular tumor spheroids at different time instants (days).

Time	3.68304	4.75446	5.55804	6.76339	7.70089	8.50446	9.50893	10.4464	11.317
Volume	0.0136799	0.0273598	0.0136799	0.0410397	0.0957592	0.109439	0.136799	0.218878	0.341997
Time	12.5223	13.5938	15.2679	16.2723	17.2768	18.4152	19.2188	21.3616	24.442
Volume	0.533516	0.629275	0.861833	1.24487	1.43639	2.32558	2.55814	3.05062	3.24213
Time	22.0982	26.5848	25.7143	27.5893	30.6696	31.4732	28.5938	35.2902	32.4777
Volume	3.44733	4.40492	4.58276	5.18468	4.93844	5.11628	5.56772	6.03283	6.21067
Time	33.0134	36.4955	37.2991	38.6384	39.7768	41.5179	42.6563	46.3393	45.4018
Volume	6.92202	6.73051	6.67579	6.94938	7.33242	7.29138	6.94938	7.26402	8.13953
Time	49.2188	48.2813	50.2232	51.0268	52.1652	53.9063	56.3839	59.2634	57.1875
Volume	7.5513	7.12722	7.0725	6.75787	7.11354	7.15458	7.1409	7.29138	8.11218

Source: Marusic et al. [3].

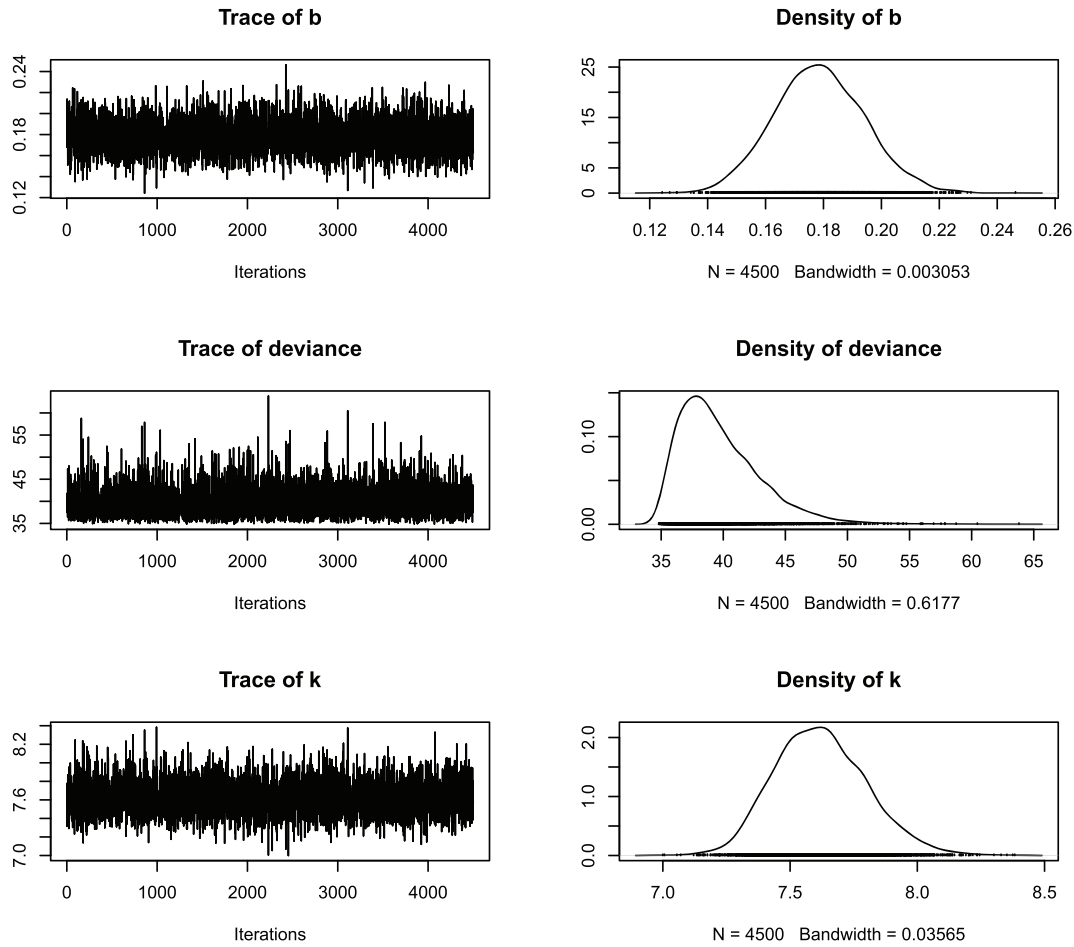## APPENDIX B: RESULTS OBTAINED FOR FIXED V0, FIRST APPROACH (BAYES)

FIGURE B1 First chart of the MCMC trace plots (left column) and marginal posterior distributions of the model parameters (right column), obtained using the Bayesian approach and considering v_0 as a deterministic constant.

TABLE B1 Statistical characteristics such as the mean and standard error of the marginal posterior density functions of the random variables of the model defined in (9), except v_0, which is considered deterministic.

	b	deviance	k	p	σ
Estimate	0.1786	39.71	7.62	0.1944	0.4017
Standard error	0.01549	3.308	0.1809	0.03176	0.04364

TABLE B2 Statistical characteristics, such as the mean, variance, and CrIs, of the 1-PDF obtained using the Bayes approach, whose graphical representation is shown on the left side of Figure 5.

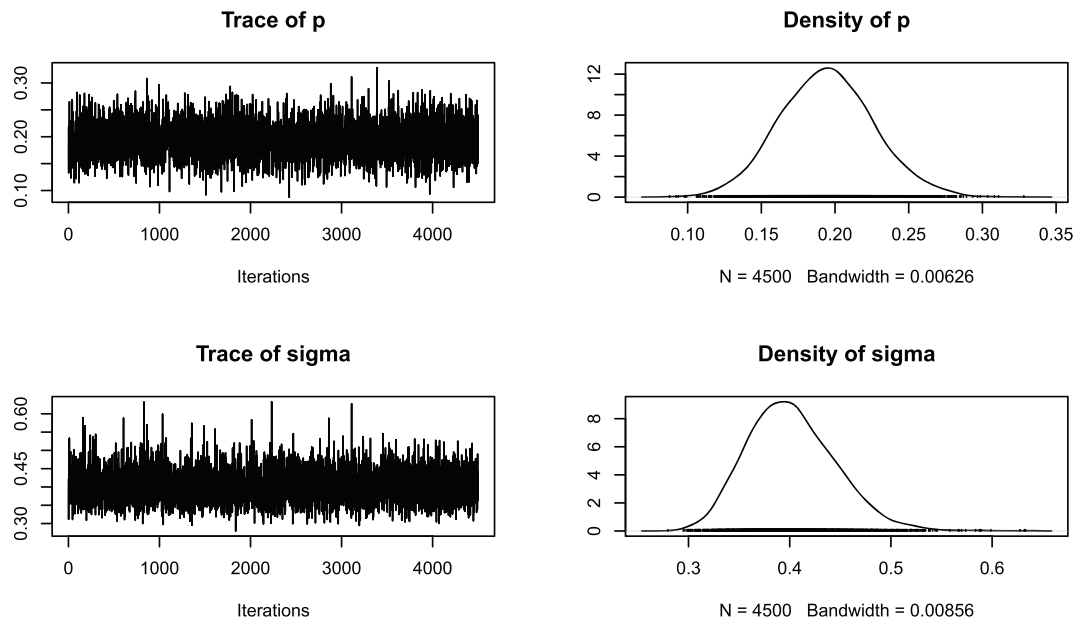| 1-PDF characteristics	$t = 10$	$t = 15$	$t = 20$	$t = 25$	$t = 30$	$t = 35$	$t = 40$	$t = 45$	$t = 50$	$t = 55$	$t = 60$
$\mathbb{E}[v(t, \omega)]$	0.273301	1.05565	2.49032	4.13588	5.46967	6.34385	6.86133	7.15839	7.33023	7.43207	7.49433
$\mathbb{V}[v(t, \omega)]$	0.010187	0.1475	0.484507	0.60241	0.410351	0.209336	0.103581	0.060179	0.043734	0.037545	0.035127
95% CrI, lower limit	0.13	0.5	1.3	2.6	4	5.4	6.1	6.6	6.9	7.1	7.2
95% CrI, upper limit	0.52	1.9	3.8	5.6	6.4	7	7.4	7.6	7.7	7.8	7.9

FIGURE B2   Second chart of the MCMC trace plots (left column) and marginal posterior distributions of the model parameters (right column), obtained using the Bayesian approach and considering v_0 as a deterministic constant.
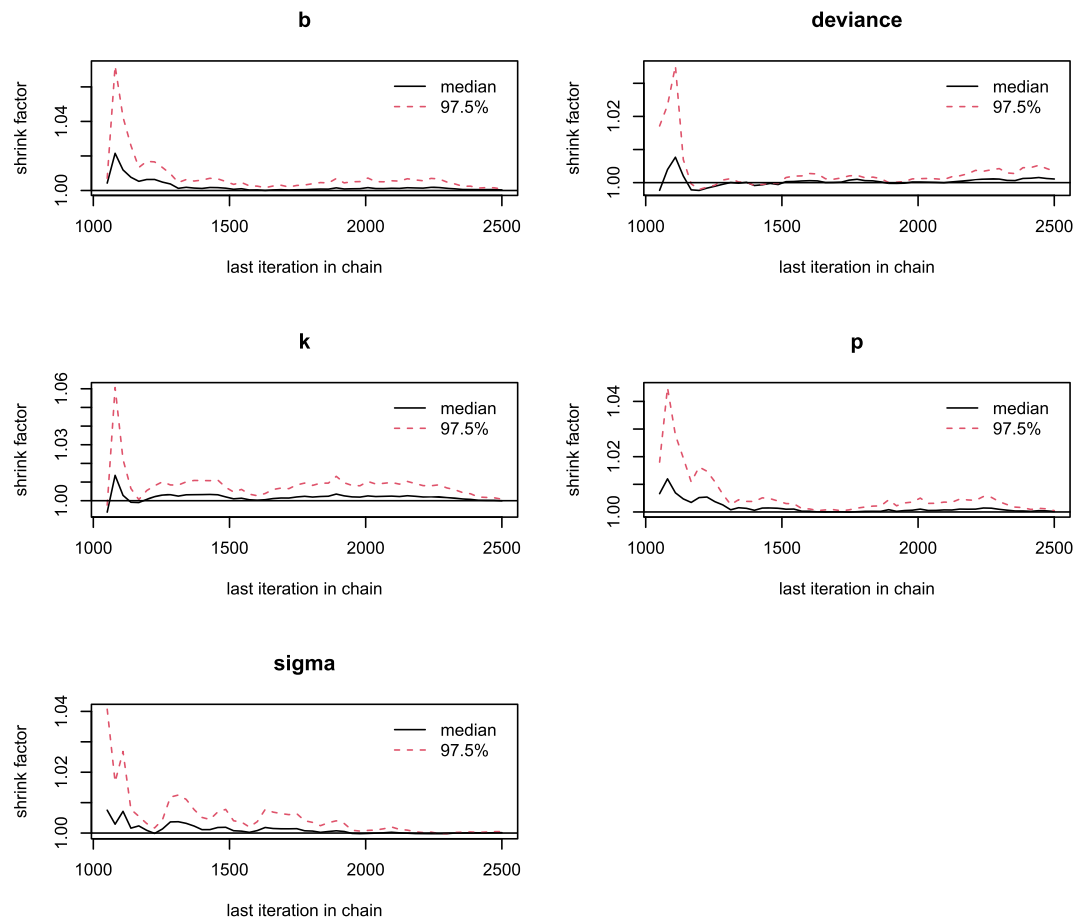

FIGURE B3   Plot illustrating the Gelman–Rubin diagnostic tool, used to evaluate the convergence of the three chains in the Bayesian approach, considering v_0 as a deterministic constant. The shrink factor compares the variance within each chain to the variance between chains. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE B3 Statistical characteristics, such as the mean and variance, of the time PDF obtained by the Bayesian approach, whose plot is shown on the left side of Figure 6.

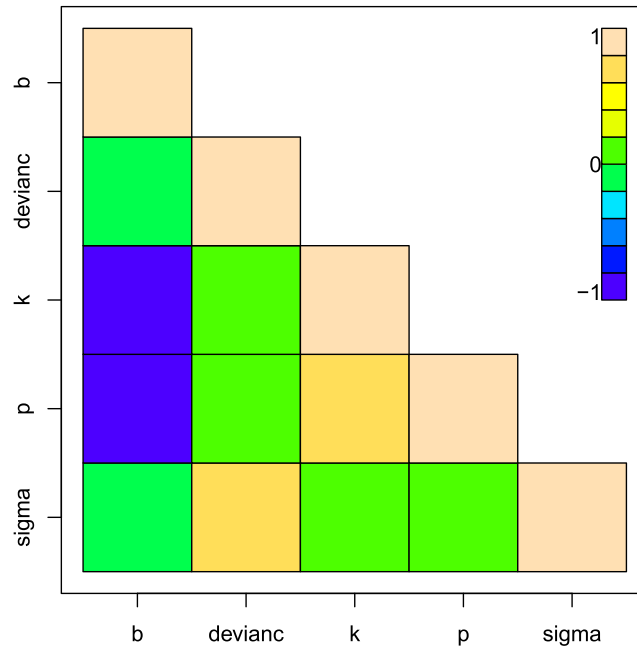| PDF features	$v = 0.027359$	$v = 0.218878$	$v = 1.24487$	$v = 3.24213$	$v = 5.11628$	$v = 6.73051$	$v = 7.15458$
$\mathbb{E}[t(v,\omega)]$	4.65729	9.62562	16.1362	22.3953	28.3725	38.2585	45.2899
$\mathbb{V}[t(v,\omega)]$	0.0416408	1.10746	3.34223	5.49184	7.72859	13.9343	45.2849

FIGURE B4 Image depicting the correlation matrix of the MCMC samples, generated from the Bayesian approach by combining the three chains. The columns represent the model parameters, and v_0 is assumed to be a deterministic constant. [Colour figure can be viewed at wileyonlinelibrary.com]

APPENDIX C: RESULTS OBTAINED WITH THE SECOND APPROACH (RLMS)

TABLE C1 Point estimates of model parameters defined in (5), together with their respective standard errors, derived from the nonlinear fit.

	b	k	p	v_0
Estimate	0.178912	7.60358	0.192177	0.0160808
Standard error	0.0162014	0.219763	0.0935527	0.0348665

TABLE C2 Statistical characteristics of the probability distributions proposed for the RLMS approach, utilizing the optimal parameters obtained from the optimization program given in (23).

	b	k	p
Mean	0.18875	7.52543	0.173828
Variance	0.000273417	0.0507045	0.000194585
Standard error	0.0165353	0.225177	0.0139494

TABLE C3 Statistical properties, including mean, variance, and 95% PIs, of the 1-PDF derived from the RLMS approach, the graph of which is shown on the right side of Figure 5.

1-PDF characteristics	$t = 10$	$t = 15$	$t = 20$	$t = 25$	$t = 30$	$t = 35$	$t = 40$	$t = 45$	$t = 50$	$t = 55$	$t = 60$
$\mathbb{E}[v(t,\omega)]$	0.241276	0.973678	2.43736	4.17358	5.54882	6.40668	6.8888	7.15319	7.30015	7.38428	7.434
$\mathbb{V}[v(t,\omega)]$	0.003141	0.067132	0.303296	0.438285	0.316627	0.172637	0.097491	0.067408	0.056449	0.052633	0.051463
95% CI, lower limit	0.15	0.6	1.5	2.8	4.4	5.5	6.2	6.6	6.8	6.9	7
95% CI, upper limit	0.37	1.6	3.6	5.4	6.6	7.1	7.4	7.6	7.8	7.8	7.9

TABLE C4 Statistical characteristics, such as mean and variance, of the time PDF acquired through the RLMS approach, the graph of which is shown on the right side of Figure 6.

PDF features	$v = 0.027359$	$v = 0.218878$	$v = 1.24487$	$v = 3.24213$	$v = 5.11628$	$v = 6.73051$	$v = 7.15458$
$\mathbb{E}[t(v,\omega)]$	4.71208	9.83904	16.3363	22.4222	28.1917	38.066	43.9726
$\mathbb{V}[t(v,\omega)]$	0.0104762	0.464555	1.66117	3.31346	5.48689	16.3314	166.013