Modeling of adulthood obesity in Spain using Itô-type stochastic differential equations

Julia Calatayud[a], Marc Jornet[a,*]

[a] *Instituto Universitario de Matemática Multidisciplinar,*
Universitat Politècnica de València,
Camino de Vera s/n, 46022, Valencia, Spain

Abstract

Obesity is growing riskily in developed and developing countries. This should pose major concerns for the countries, not only from the health point, but also from the economic perspective. Our case study relies on the excess weight dynamics in Spain. The Spanish National Health Survey (ENSE) 2017 collects the percentage of overweight and obese adults in Spain from the year 1987 to 2017. A recent contribution proposed a nonautonomous compartmental system of ordinary differential equations to calibrate the incidence of excess weight in the Spanish adulthood population. Essentially, three principles were followed: the total adulthood population is time-dependent, the subpopulations interact homogeneously along the country, and excess weight plays the role of an infectious disease that is transmitted through contact by social pressure. Accounting for both data and model errors, frequentist nonlinear regression and Bayesian inference were conducted. The methods agreed well in terms of fit, prediction, bands and sensitivity analysis. In the present paper, the deterministic compartmental system of ordinary differential equations is randomized in a different manner, by employing Itô-type stochastic differential equations. The derivatives of the compartments are perturbed by Gaussian white noise-type pure errors that have a rough and unpredictable structure. From the Euler-Maruyama discretization, several strategies are utilized for estimating the parameters, based on the moments method and maximum likelihood estimation. Comparison is performed numerically by assessing the fit to the data.

Keywords: adulthood obesity epidemic, mathematical modeling, Itô stochastic differential equation, estimation of parameters, simulation

2010 MSC: 60H10, 60H35, 65C30, 92D30

1. Introduction

Overweight and obesity are defined as an excessive fat accumulation. For adults, the classification into overweight or obese depends on the Body Mass Index (BMI = weight/height^2), where the weight is measured in kilograms and the height in meters. Overweight individuals

*Corresponding author
Email addresses: jucagre@doctor.upv.es (Julia Calatayud), marjorsa@doctor.upv.es (Marc Jornet)

have a BMI in the range $[25, 30)$ and obese individuals have it greater than or equal to 30. According to the World Health Organization (WHO) [1, 2], obesity is increasing in developed and developing countries at alarming rate, paradoxically coexisting with malnutrition. It has nearly tripled since 1975. In 2016, the prevalence of overweight and obesity among adults aged 18 or more years old were 39% and 13%, respectively. Yet most neglected, obesity entails serious health consequences and poses major risk for noncommunicable diseases such as diabetes, cardiovascular diseases, hypertension, musculoskeletal disorders and certain forms of cancer. This affects life quality and decreases life expectancy, augmenting the probability of premature death. Today excess weight causes more than 2.8 million deaths each year [3]. It has been formally considered a global pandemic [4].

The economic burden of obesity is significant. There are at least four major categories of economic impact [5]: direct medical costs, productivity costs, transportation costs, and human capital costs. In the United States, each obese adult raises the medical care costs by an average of 3429 dollars (year 2013) [6]. In Spain, it was estimated by an average of 7610 euro [7] (year 1999).

The magnitude of the health and economic impact of obesity, and its likely higher impact in the future [8, 9], highlight the importance of addressing the obesity epidemic by public authorities. In this regard, the use of mathematical models is an effective tool to describe, explain and predict the evolution of the epidemic and propose targeted measures [10].

A lot of approaches in mathematical modeling rely on the use of compartmental models of coupled differential equations [11, 12, 13, 14, 15]. The population is divided into distinct groups (compartments) according to the disease state. These models have been employed to analyze the incidence of obesity in the region of Valencia, Spain, [16] (for children) and [17] (for adults). Our focus is put on the recent study [18]. It used data on adulthood excess weight along thirty years, collected by the Spanish National Health Survey (ENSE). The ENSE is a periodic study conducted by the Spanish Ministry of Health, Consumption and Social Welfare, with the collaboration of the National Statistics Institute of Spain (INE), since 1987. It gathers transversal data on the health of the resident population in Spain. The last survey considered data from 2017 and was published in mid-2018 [19]. To design the compartmental model, paper [18] followed the following principles: the population is divided into normal weight, overweight and obese individuals according to BMI, the total adulthood population is time-dependent (INE data from [20]) and tends to a constant value asymptotically, the subpopulations interact homogeneously along the country [11], and excess weight plays the role of an infectious disease that is transmitted through contact by social influence and imitation. With regard to the first principle, the ENSE data is focused on excess weight and there is no information concerning underweight individuals (too low BMI); all persons with BMI less than 25 are joined in a single group. With regard to the last principle, human-to-human transmission of obesity is justified by medical and sociological studies [21, 22]. In the seminal contribution [21], the authors evaluated a densely interconnected social network and explicitly concluded that obesity appears to spread through social ties. For instance, a person's probability of becoming obese increased by 57% if he or she had a friend who became obese in a given interval; among pairs of adult siblings, if one sibling became obese, the chance that the other would become obese increased by 40%; if one spouse became obese, the likelihood that the other spouse would become obese augmented by 37%. Paper [22] and its references from the introduction defend that gains in weight are spread through the population by social influence, in a way reminiscent of a contagious disease of social transmission. Obviously, excess weight is not an infectious disease. But sociologically, taking into account human behavior, overweight and obese persons may be considered as some sort of "infectives", who transmit excess weight to healthy individuals through contact (social pressure for taking unhealthy habits). The importance of contacts in non-infectious phenomena, where the bad-habit group is a "transmitter", has also been revealed within alcohol and tobacco consumption, for example [15, 23]. For phenomena not related to health, for example the mobile telecommunications market share [24], contacts and imitation are key too.

The formulation of the compartmental model in [18] was, firstly, deterministic. The parameters were estimated by minimizing an objective function defined by the mean square error of the response (optimization). But excess weight dynamics, as most of the phenomena, inherently involves a vast set of complexities and uncertainties. Thus, due to both data and model errors, the incorporation of some sort of randomness into the model becomes indispensable. Frequentist nonlinear regression and Bayesian inference (brute-force and Adaptive Metropolis algorithms) [25, 26] were conducted. These methods agreed well in terms of fit, prediction, bands and sensitivity analysis. They suggested that prevention strategies should take priority over treatment strategies to manage adulthood obesity. Specially, the treatment of the transition from the obese to the overweight states is the least recommendable for mitigating the obesity epidemic.

Frequentist nonlinear regression and Bayesian inference present some drawbacks. The former applies linearization through a Jacobian, which entails errors, clearly observable when extrapolation for future years is carried out. The latter, although based on exact formulas, resorts to algorithms on Markov Chains for sampling, which may be a complex and time-consuming procedure. See [18, 25]. In this paper, the deterministic compartmental model from [18] is randomized in a different manner, via noise. The derivatives of the compartments are perturbed by Gaussian white noise-type pure errors that have a rough and unpredictable structure. The model becomes a stochastic differential equation of Itô-type driven by a standard Brownian motion (Wiener process), whose mathematical formalization in terms of integrals is due to to the development of a new operational calculus by Kiyoshi Itô [27, 28, 29]. The number of individuals at each compartment is thus a nowhere differentiable, continuous stochastic process. Existence and uniqueness of mean square solution relies on a local Lipschitz and a monotone condition due to nonlinearities, instead of the usual global Lipschitz and linear growth conditions [27, Ch. 2]. The stochastic differential equations are discretized by the Euler-Maruyama scheme. From the stochastic difference equations and the ENSE data, the parameters are estimated by means of different techniques, based on the moments method and maximum likelihood estimation [28, pp. 118–122]. For the implementation of the stochastic model, the code available in [30] for the software Mathematica® [31] can be used. Forward uncertainty quantification is carried out via Monte Carlo simulation. A detailed comparison between the parameter estimations is performed numerically by assessing the fit to the data.

The organization of this paper is the following. In Section 2, the main content from [18] is exposed. Specifically, the formulations of the deterministic, frequentist and Bayesian models, and their results. In Section 3, the deterministic model from [18] is randomized by adding white noise perturbations. Existence and uniqueness of mean square solution are proved and the different methods for parameter estimation are detailed. In Section 4, the parameters are estimated numerically and the correspondence of the models to the data is assessed. Finally, Section 5 discusses the results obtained.

2. Previous mathematical model

In this section, the main content from [18] is exposed, and the reader is referred to that paper for further details. The formulation of the deterministic model is shown, since it is the basis for the Itô-type stochastic differential equations model that will be proposed later in Section 3. The randomization, based on frequentist and Bayesian approaches, is briefly explained, for later comparison with the Itô approach.

2.1. Excess weight model

The adulthood population is divided into normal weight, overweight and obese subpopulations, according to BMI. By adults, individuals aged 18 or more years old are considered. Let $S(t)$ and $O(t)$ denote the number of overweight and obese adults at time (year) $t \geq 0$. The initial time $t = 0$ corresponds to the year 1987. The ENSE has collected data from 1987 to 2017, at 11 periods: 1987, 1993, 1995, 1997, 2001, 2003, 2006, 2009, 2011, 2014 and 2017. These years correspond to times t equal to 0, 6, 8, 10, 14, 16, 19, 22, 24, 27, 30, which are denoted by $t_1, \ldots, t_{11}$. The recorded observations on total adulthood population (by INE), overweight and obese individuals, scaled by ten million by dividing by 10^7, are

$$(2.82, 3.04, 3.12, 3.2, 3.34, 3.47, 3.66, 3.82, 3.84, 3.82, 3.82),$$

$$(s_1, \ldots, s_{11}) = (0.9024, 1.09136, 1.10916, 1.1248, 1.23413, 1.27522, 1.37982, 1.41913, 1.40928, 1.36756, 1.42104),$$

$$(o_1, \ldots, o_{11}) = (0.20868, 0.27816, 0.34008, 0.3968, 0.44255, 0.47192, 0.56181, 0.61884, 0.6528, 0.64558, 0.66659).$$

As can be seen, excess weight has been worryingly rising in the last thirty years. The prevalence of obesity has been multiplied by 2.4 (in 1987, the prevalence was $0.20868/2.82 = 0.074\%$; in 2017, it was $0.66659/3.82 = 0.1745\%$; and $0.1745/0.074 \approx 2.4$). Up to the year 2014, it seemed that the rate of increase was slowing down at last. However, the year 2017 did not confirm that expected deceleration.

The adulthood population in Spain since 1987, scaled by ten million, was modeled as

$$T'(t) = \mu T(t)\left(1 - \frac{T(t)}{K}\right), \quad t \geq 0, \tag{1}$$

where $\mu, K > 0$. By using data from the INE, these parameters were estimated as

$$\hat{\mu} = 0.0491843, \; \hat{K} = 4.47700. \tag{2}$$

Model (1) became

$$T(t) = \frac{12.3654}{2.76197 + 1.71503\,\mathrm{e}^{-0.0491843t}}, \quad t \geq 0. \tag{3}$$

To construct the compartmental model, homogeneous mixing was assumed [11], i.e. the subpopulations interact along the country with equal probability. Contagion of obesity by social pressure and imitation was considered [21, 22], where excess weight individuals transmit fatness to normal weight peers by contact (see the Introduction Section for further justification). Accounting on these suppositions, the following nonautonomous model for excess weight was defined:

$$\begin{cases} S'(t) = \mu S(t)\left(1 - \frac{T(t)}{K}\right) + \beta \frac{T(t)-S(t)-O(t)}{T(t)}\left[S(t) + O(t)\right] - (\rho + \gamma)S(t) + \epsilon O(t), & t \geq 0, \\ O'(t) = \mu O(t)\left(1 - \frac{T(t)}{K}\right) + \gamma S(t) - \epsilon O(t), & t \geq 0, \\ S(0) = S_0, \\ O(0) = O_0. \end{cases} \tag{4}$$

The parameters, measured in year^{-1} [11, p. 27], were the following: $\beta > 0$ is the force of "infection" (transmission rate because of social pressure), $\rho > 0$ is the rate at which an overweight individual becomes normal weight, $\gamma > 0$ is the rate at which an overweight person moves to the obese compartment, and $\epsilon > 0$ is the rate at which an obese adult becomes overweight. These coefficients are the most important of the model, since variations of them determine the flow between compartments and the future evolution of the disease. The initial conditions S_0 and O_0 were the scaled number of overweight and obese adults in 1987, $s_1 = 0.902400$ and $o_1 = 0.208680$, respectively. The function $T(t)$ was given by (3).

Model (4) is well-posed. Given the data by ENSE, the minimization of the mean square error,

$$\sum_{i=1}^{11} \left(S(t_i|\beta,\gamma,\epsilon,\rho) - s_i\right)^2 + \sum_{i=1}^{11} \left(O(t_i|\beta,\gamma,\epsilon,\rho) - o_i\right)^2, \tag{5}$$

gave the following parameter estimates:

$$\hat{\beta} = 0.368989, \; \hat{\gamma} = 0.0222886, \; \hat{\epsilon} = 0.0344076, \; \hat{\rho} = 0.240838. \tag{6}$$

The MAPE (mean absolute percentage error),

$$\frac{100}{22}\left(\sum_{i=1}^{11} \frac{|S(t_i|\hat{\beta},\hat{\gamma},\hat{\epsilon},\hat{\rho}) - s_i|}{s_i} + \sum_{i=1}^{11} \frac{|O(t_i|\hat{\beta},\hat{\gamma},\hat{\epsilon},\hat{\rho}) - o_i|}{o_i}\right), \tag{7}$$

was 3.17. This MAPE is considered low by Lewis' scale and the forecast accuracy is high [32, p. 40].

According to this model, 36.65% and 23.74% of Spanish adults will be overweight and obese in the long run, respectively, assuming that the parameters values keep time-invariant. Sensitivity analyses, based on MAPE comparisons and differentiation, indicated that prevention strategies are more important than treatment strategies to control adulthood obesity.

In Mathematica®, the numerical solution to (4) was obtained by using the standard *ParametricNDSolveValue* built-in function, with no specified options. The execution lasted at most milliseconds. The minimization of (5) was performed by means of the standard *NMinimize* built-in routine, with no options. The region of minimization was $(0, 1)^4$.

2.2. Nonlinear regression

The nonlinear regression model [25, Ch. 7] was the following:

$$S_i = S(t_i|\beta,\gamma,\epsilon,\rho) + \mathcal{E}_i^S, \tag{8}$$

$$O_i = O(t_i|\beta, \gamma, \epsilon, \rho) + \mathcal{E}_i^O, \tag{9}$$

for $i = 1, \ldots, 11$. The responses S_i and O_i were the random variables corresponding to the overweight and obese adults (scaled by ten million). The terms $\mathcal{E}_i^S$ and $\mathcal{E}_i^O$ were random errors with zero expectation and constant variance σ^2, all of them identically distributed and independent. The parameters β, γ, ϵ and ρ were constant.

From the ENSE data, the estimation of the parameters was (2) and (6). The variance σ^2 was estimated as the mean square error (5) divided by $22 - 4$, where 22 is the number of observations (11 periods for overweight and obesity) and 4 is the number of parameters that control the flow between compartments:

$$\hat{\sigma} = 0.0321734. \tag{10}$$

Confidence-based prediction intervals were constructed through these estimates and the Jacobian matrix of the response with respect to the parameters. These intervals are useful to locate measures from 1987 to 2017 and to predict beyond 2017. Extrapolation in the long run gave the intervals $[1.532, 1.750]$ and $[0.6126, 1.513]$ for the number of overweight and obese adults, respectively. The conclusions on sensitivity analyses, based on MAPE comparisons and t-values, coincided with the deterministic model.

2.3. Bayesian inference

Bayesian inference [25, Ch. 8] [26], in contrast to frequentist nonlinear regression, considers the parameters as random quantities. The coefficients β, γ, ϵ and ρ were random variables, with a certain probability distribution called prior distribution. They were given uniform laws on $(0, 1)$. The error variance σ^2 may also have a prior probability distribution. The two cases, constant and random error variance, were tackled. In the former case, the prior of σ^2 was a Dirac delta function centered at (10) squared. In the latter situation, a conditional distribution of σ^2 followed an inverse gamma distribution. The joint response, conditioned to the parameters (i.e. the likelihood), followed a product of independent Gaussian distributions centered at the deterministic solution and error variance σ^2:

$$S_i|\beta, \gamma, \epsilon, \rho, \sigma^2 \sim \text{Normal}\left(S(t_i|\beta, \gamma, \epsilon, \rho), \sigma^2\right),$$
$$O_i|\beta, \gamma, \epsilon, \rho, \sigma^2 \sim \text{Normal}\left(O(t_i|\beta, \gamma, \epsilon, \rho), \sigma^2\right).$$

From the data and Bayes' formula, the parameters, conditioned to data, followed posterior distributions. Forward uncertainty quantification (mean values and probabilistic prediction intervals) was performed through the posterior predictive distribution (S_i and O_i conditioned to data). The MAPE (7), calculated from mean values, was 3.23. Asymptotic extrapolation gave the pointwise expected values 1.64 and 1.07 and the intervals $[1.57, 1.70]$ and $[0.97, 1.19]$ for the number of overweight and obese adults, respectively. The conclusions on sensitivity analyses, based on Bayes' factors and the Savage-Dickey density ratio, agreed with the deterministic model.

The Metropolis algorithm was implemented in Mathematica®. It generates a Markov chain whose stationary distribution is the desired posterior distribution of the parameters. The standard and the Adaptive versions were considered, the latter providing a substantial reduction in computational time. It was checked that the use of generalized polynomial chaos expansions did not yield computational improvements.

3. New stochastic model

In this section, we present the Itô-type stochastic differential equation model. It is proved that the stochastic system is well-posed. Different methods for parameter estimation are suggested.

3.1. Formulation

Let us start from the deterministic formulation (4). The frequentist approach considered the parameters as constants and added a random error to the response. The Bayesian approach considered the parameters as random quantities, together with a Gaussian random error for the response. The approach based on Itô stochastic differential equations [27, 28] is different. The parameters are regarded as constants. The random error is not introduced into the response, but into the derivative of the response instead. The derivative of the response is perturbed by an idealized stochastic process called white noise, which is Gaussian, has zero mean, independent components, and infinite dispersion. This process is viewed as the formal derivative of a standard Brownian motion (Wiener process). Hence the response derivative has an unpredictable and rough behavior. Fixed a realizable path of the response, it is continuous and nowhere differentiable.

From (4), the following stochastic model is proposed:

$$\begin{cases} S'(t) = \mu S(t)\left(1 - \frac{T(t)}{K}\right) + \beta\frac{T(t)-S(t)-O(t)}{T(t)}\left[S(t) + O(t)\right] - (\rho+\gamma)S(t) + \epsilon O(t) + \sigma_1\xi_1(t), \\ O'(t) = \mu O(t)\left(1 - \frac{T(t)}{K}\right) + \gamma S(t) - \epsilon O(t) + \sigma_2\xi_2(t), \\ S(0) = S_0, \\ O(0) = O_0. \end{cases} \tag{11}$$

Here σ_1 and σ_2 are the positive diffusion coefficients. The terms $\xi_1(t)$ and $\xi_2(t)$ represent independent white noise stochastic processes, defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. They may be written as $\mathrm{d}B_1(t) = \xi_1(t)\mathrm{d}t$ and $\mathrm{d}B_2(t) = \xi_2(t)\mathrm{d}t$, where $B_1(t)$ and $B_2(t)$ are independent standard Brownian motions. By standard, we refer to the Gaussian distribution centered at zero with variance equal to t. Model (11) may be rewritten in differential form:

$$\begin{cases} \mathrm{d}S(t) = \left[\mu S(t)\left(1 - \frac{T(t)}{K}\right) + \beta\frac{T(t)-S(t)-O(t)}{T(t)}\left[S(t) + O(t)\right] - (\rho+\gamma)S(t) + \epsilon O(t)\right]\mathrm{d}t + \sigma_1\mathrm{d}B_1(t), \\ \mathrm{d}O(t) = \left[\mu O(t)\left(1 - \frac{T(t)}{K}\right) + \gamma S(t) - \epsilon O(t)\right]\mathrm{d}t + \sigma_2\mathrm{d}B_2(t), \\ S(0) = S_0, \\ O(0) = O_0. \end{cases} \tag{12}$$

These equations should be interpreted by integration. The solutions $S(t)$ and $O(t)$ to (12) are stochastic processes, whose trajectories are continuous and nowhere differentiable. They should be measurable, adapted with respect to the natural filtration of the Brownian motion $B = (B_1, B_2)$, and mean square integrable.

Notice that this new model uses two different infinitesimal standard deviations σ_1 and σ_2 for the overweight and the obese classes. It is reasonable that the overweight and the obese subpopulations, being different groups, may have different dispersion. This is different to the nonlinear regression and the Bayesian models from [18], which considered the same error variance for facility.

3.2. Existence and uniqueness of solution

For notational convenience, the previous stochastic model (12) is rewritten in generic form. Let

$$X(t) = \begin{pmatrix} S(t) \\ O(t) \end{pmatrix}, \quad X_0 = \begin{pmatrix} S_0 \\ O_0 \end{pmatrix}, \quad B(t) = \begin{pmatrix} B_1(t) \\ B_2(t) \end{pmatrix},$$

$$b_1(X(t), t) = \mu S(t) \left(1 - \frac{T(t)}{K}\right) + \beta \frac{T(t) - S(t) - O(t)}{T(t)} [S(t) + O(t)] - (\rho + \gamma) S(t) + \epsilon O(t),$$

$$b_2(X(t), t) = \mu O(t) \left(1 - \frac{T(t)}{K}\right) + \gamma S(t) - \epsilon O(t),$$

$$b(X(t), t) = \begin{pmatrix} b_1(X(t), t) \\ b_2(X(t), t) \end{pmatrix}, \quad \sigma(X(t), t) = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}.$$

Then (12) is

$$\mathrm{d}X(t) = b(X(t), t)\mathrm{d}t + \sigma(X(t), t)\mathrm{d}B(t), \quad t \geq 0, \quad X(0) = X_0. \tag{13}$$

Let $\|\cdot\|$ denote the Euclidean vector norm and the Fröbenius (or trace) matrix norm. It is said that $X(t)$ solves (13) if, for all time horizon $\tau > 0$, X is jointly measurable on $([0, \tau] \times \Omega, \mathcal{B}[0, \tau] \otimes \mathcal{F})$, it is adapted with respect to the natural filtration of B, and $\int_0^\tau \mathbb{E}[\|X(t)\|^2]\mathrm{d}t = \mathbb{E}[\int_0^\tau \|X(t)\|^2 \mathrm{d}t] < \infty$ ($\mathbb{E}$ is the expectation). There is uniqueness if $\mathbb{P}[X(t) = Y(t), \forall t \geq 0] = 1$ for any two solutions X and Y.

In [27, Th. 3.6], two conditions are provided for the existence and uniqueness of solution on $[0, \infty)$ (we restrict to our two-dimensional situation):

H1: For every real number $\tau > 0$ and integer $n \geq 1$, there exists a positive constant $C_{\tau,n}$ such that, for all $t \in [0, \tau]$ and all $X, Y \in \mathbb{R}^2$ with $\|X\| \leq n$ and $\|Y\| \leq n$,

$$\|b(X, t) - b(Y, t)\|^2 + \|\sigma(X, t) - \sigma(Y, t)\|^2 \leq C_{\tau,n} \|X - Y\|^2.$$

H2: For every $\tau > 0$, there exists a positive constant C_τ such that, for all $(X, t) \in \mathbb{R}^2 \times [0, \tau]$,

$$X^\top b(X, t) + \frac{1}{2} \|\sigma(X, t)\|^2 \leq C_\tau (1 + \|X\|^2).$$

Let us check these two conditions for our particular model (12). The part concerning the diffusion σ is trivial. We focus on the drift part. The local Lipschitz condition from H1 for b is clear, because b has continuous partial derivatives of first order with respect to X. The monotone condition H2 needs some work. We have

$$X^\top b(X, t) = S b_1(S, O, t) + O b_2(S, O, t).$$

We first focus on the number of contacts $(T(t) - S - O)/T(t) \times (S + O)$. If $0 \leq S + O \leq T(t)$, then $(T(t) - S - O)/T(t) \times (S + O) \leq S + O \leq \|(S, O)\|^2/2$. Otherwise, it is negative. Hence

$$\begin{aligned} S b_1(S, O, t) &= \mu S^2 \left(1 - \frac{T(t)}{K}\right) + \beta S \frac{T(t) - S - O}{T(t)} [S + O] - (\rho + \gamma) S^2 + \epsilon S O \\ &\leq \mu S^2 + \beta S \frac{\|(S, O)\|^2}{2} + \epsilon \frac{\|(S, O)\|^2}{2} \lesssim \|(S, O)\|^2 \end{aligned}$$

and

$$Ob_2(S,O,t) = \mu O\left(1 - \frac{T(t)}{K}\right) + \gamma S - \epsilon O$$
$$\leq \mu O + \gamma S \lesssim \|(S,O)\|^2,$$

where $\lesssim$ denotes less than or equal to except a positive constant. As a consequence, H2 holds. It is concluded that (12), (13), possesses a unique solution on $[0,\infty)$.

It is to be noted that H1 and H2 are not the typical conditions for existence and uniqueness of solution, since the assumptions usually rely on a global Lipschitz and a linear growth condition (i.e. $n = \infty$ in H1, and $\|b(X,t)\|^2 + \|\sigma(X,t)\|^2 \leq C_\tau(1+\|X\|^2)$ in H2) [27, Ch. 2]. The nonlinearities that are present in (12) do not allow for those assumptions.

From [27, Th. 4.1, Cor. 4.5], the processes $S(t)$ and $O(t)$ have statistical moments of any order $p > 0$:

$$\mathbb{E}[\|(S(t),O(t))\|^p] \leq 2^{\frac{p-2}{2}}(1 + \|(S_0,O_0)\|^p)\mathrm{e}^{p\alpha_\tau t}, \quad \forall t \in [0,\tau], \quad \forall p \geq 2,$$
$$\mathbb{E}[\|(S(t),O(t))\|^p] \leq (1 + \|(S_0,O_0)\|^2)^{\frac{p}{2}}\mathrm{e}^{p\alpha_\tau t}, \quad \forall t \in [0,\tau], \quad \forall\, 0 < p < 2,$$

where α_τ is a positive constant related to H2.

Strictly speaking, the response processes may take negative values, which does not make sense. Nonetheless, the probability of such occurrence is, for the typical values of the parameters, negligible.

3.3. *Estimation of parameters: Method of moments*

In statistical inference, the method of moments consists in estimating population parameters by equating the population statistics to the sample statistics. This is based on the law of large numbers. Such philosophy is applied for stochastic differential equations.

The stochastic model (12) is discretized by the Euler-Maruyama scheme. Given a time t and a step size $\Delta t > 0$, the scheme reads as follows:

$$S(t+\Delta t) = S(t) + b_1(S(t),O(t),t)\Delta t + \sigma_1(B_1(t+\Delta t) - B_1(t)), \tag{14}$$

$$O(t+\Delta t) = O(t) + b_2(S(t),O(t),t)\Delta t + \sigma_2(B_2(t+\Delta t) - B_2(t)). \tag{15}$$

Taking into account that $B_i(t+\Delta t) - B_i(t)$ is Gaussian, centered at zero with standard deviation given by $\sigma_i\sqrt{\Delta t}$, it is obtained

$$\frac{S(t+\Delta t) - S(t)}{\sqrt{\Delta t}} - b_1(S(t),O(t),t)\sqrt{\Delta t} \sim \mathrm{Normal}(0,\sigma_1^2), \tag{16}$$

$$\frac{O(t+\Delta t) - O(t)}{\sqrt{\Delta t}} - b_2(S(t),O(t),t)\sqrt{\Delta t} \sim \mathrm{Normal}(0,\sigma_2^2) \tag{17}$$

Convergence of the Euler-Maruyama scheme under assumptions H1 and H2, in the sense of

$$\lim_{\Delta t\to 0}\mathbb{E}\left[\sup_{0\leq t\leq T}\|(\tilde{S}(t),\tilde{O}(t)) - (S(t),O(t))\|^2\right] = 0,$$

where $(\tilde{S}(t), \tilde{O}(t))$ is the approximate solution by Euler-Maruyama, is studied in [33]. Notice that, since the diffusion coefficients do not depend on S and O, Milstein scheme reduces to Euler-Maruyama[1].

Let

$$u_i = \frac{s_{i+1} - s_i}{\sqrt{t_{i+1} - t_i}} - b_1(s_i, o_i, t_i)\sqrt{t_{i+1} - t_i},$$
$$v_i = \frac{o_{i+1} - o_i}{\sqrt{t_{i+1} - t_i}} - b_2(s_i, o_i, t_i)\sqrt{t_{i+1} - t_i},$$

for $i = 1, \ldots, 10$. These realizations are independent, because the increments of Brownian motion are independent. If μ and K are given by (2) and the flow parameters β, γ, ϵ and ρ are assumed to be already fixed, then the instantaneous standard deviations σ_1 and σ_2 may be estimated by the standard deviations of the samples $\{u_1, \ldots, u_{10}\}$ and $\{v_1, \ldots, v_{10}\}$, respectively, by (16) and (17). Let $\hat{\sigma}_1$ and $\hat{\sigma}_2$ be such estimates.

The parameters β, γ, ϵ and ρ may have different sources of estimation:

M1: One may simply consider the deterministic fit (6).

M2: From (16) and (17), one may impose $\sum_{i=1}^{10} u_i = 0$ and $\sum_{i=1}^{10} v_i = 0$. But this approach has two problems: the estimation is overdetermined (two equations for four unknowns), and the realizations may not fluctuate around zero randomly (it might be possible that $u_1, \ldots, u_5$ are positive and $u_6, \ldots, u_{10}$ are negative, for instance). Hence one may add other equations such as $\sum_{i=1}^{5} u_i = 0$, $\sum_{i=6}^{10} u_i = 0$, $\sum_{i=1}^{5} v_i = 0$ and $\sum_{i=6}^{10} v_i = 0$. In general, one solves

$$\min_{\beta,\gamma,\epsilon,\rho} \left(\sum_{i=1}^{10} u_i\right)^2 + \left(\sum_{i=1}^{10} v_i\right)^2 + \left(\sum_{i=1}^{5} u_i\right)^2 + \left(\sum_{i=6}^{10} u_i\right)^2 + \left(\sum_{i=1}^{5} v_i\right)^2 + \left(\sum_{i=6}^{10} v_i\right)^2.$$

We do not recommend to use sums with less than five terms, because being normally distributed does not mean to alternate positive and negative values for each realization. Notice that the probability that a zero-mean normal random variable has constant sign for five consecutive realizations is $0.5^5 = 0.03125$, which is less than the possible threshold 0.05. That is why we partitioned at five terms. Let $\hat{\beta}$, $\hat{\gamma}$, $\hat{\epsilon}$ and $\hat{\rho}$ be the estimates.

M3: From (16) and (17), one may impose $\sum_{i=1}^{10} u_i = 0$ and $\sum_{i=1}^{10} v_i = 0$. But one also may use the fact that the moments of third order of a zero-mean Gaussian law are zero. Therefore, $\sum_{i=1}^{10} u_i^3 = 0$ and $\sum_{i=1}^{10} v_i^3 = 0$. The four equalities may have no suitable solution due to nonlinearities. In general, one solves

$$\min_{\beta,\gamma,\epsilon,\rho} \left(\sum_{i=1}^{10} u_i\right)^2 + \left(\sum_{i=1}^{10} v_i\right)^2 + \left(\sum_{i=1}^{10} u_i^3\right)^2 + \left(\sum_{i=1}^{10} v_i^3\right)^2.$$

[1] Given a general stochastic differential equation problem (13), the Euler-Maruyama scheme is

$$X_{n+1} = X_n + b(X_n, t_n)(t_{n+1} - t_n) + \sigma(X_n, t_n)(B(t_{n+1}) - B(t_n)),$$

and Milstein scheme is

$$X_{n+1} = X_n + b(X_n, t_n)(t_{n+1} - t_n) + \sigma(X_n, t_n)(B(t_{n+1}) - B(t_n)) + \frac{1}{2}\sigma(X_n, t_n)\frac{\partial \sigma}{\partial X}(X_n, t_n)\left[(B(t_{n+1}) - B(t_n))^2 - (t_{n+1} - t_n)\right].$$

Milstein discretization presents higher order of convergence, in general. Precisely when σ does not depend on X, Milstein scheme reduces to Euler-Maruyama.

M4: From (16) and (17), one may impose $\sum_{i=1}^{10} u_i = 0$ and $\sum_{i=1}^{10} v_i = 0$. On the other hand, let $\{\mathcal{F}_t\}_{t\geq 0}$ be the natural filtration of Brownian motion. As $B(t+\Delta t) - B(t)$ is independent of $\mathcal{F}_t$, it holds $\mathbb{E}[B(t+\Delta t) - B(t)|\mathcal{F}_t] = \mathbb{E}[B(t+\Delta t) - B(t)] = 0$. As a consequence,

$$\mathbb{E}\left[\frac{S(t+\Delta t) - S(t)}{\sqrt{\Delta t}} - b_1(S(t), O(t), t)\sqrt{\Delta t} \middle| \mathcal{F}_t \right] = 0$$

(analogously for $O(t)$). Since $S(t)$ is $\mathcal{F}_t$-measurable,

$$\begin{aligned} 0 &= S(t)\mathbb{E}\left[\frac{S(t+\Delta t) - S(t)}{\sqrt{\Delta t}} - b_1(S(t), O(t), t)\sqrt{\Delta t} \middle| \mathcal{F}_t \right] \\ &= \mathbb{E}\left[\frac{S(t)(S(t+\Delta t) - S(t))}{\sqrt{\Delta t}} - S(t) b_1(S(t), O(t), t)\sqrt{\Delta t} \middle| \mathcal{F}_t \right]. \end{aligned}$$

By applying expectation,

$$\mathbb{E}\left[\frac{S(t)(S(t+\Delta t) - S(t))}{\sqrt{\Delta t}} - S(t) b_1(S(t), O(t), t)\sqrt{\Delta t} \right] = 0.$$

For $t_1, \ldots, t_{10}$, the random variables within the last expectation are uncorrelated, therefore two new equations are derived: $\sum_{i=1}^{10} s_i u_i = 0$ and $\sum_{i=1}^{10} o_i v_i = 0$.

It is by no means clear which of the four approaches gives rise to more satisfactory results. Especially in our case study, in which only a small amount of data is available at spaced times. The performance of the different estimators must be studied through Monte Carlo experiments.

3.4. Estimation of parameters: Maximum likelihood method

The maximum likelihood estimation chooses the parameters that are most likely to have generated the sample. Let

$$\mathcal{L}(\beta, \gamma, \epsilon, \rho, \sigma_1, \sigma_2 | s_1, \ldots, o_{11}) = \pi_{(S_1, \ldots, O_{11})}(s_1, \ldots, o_{11} | \beta, \gamma, \epsilon, \rho, \sigma_1, \sigma_2)$$

be the likelihood, where π is the probability density function. By maximizing it, it is maximized

$$\mathbb{P}[S_1 \in [s_1, s_1 + \mathrm{d}s_1], \ldots, O_{11} \in [o_{11}, o_{11} + \mathrm{d}o_{11}]] = \mathcal{L}\,\mathrm{d}s_1 \cdots \mathrm{d}o_{11}.$$

From the law of total probability and the Markovian nature of $(S(t), O(t))$, the likelihood is factorized as

$$\mathcal{L} = \pi_{(S_1, O_1)}(s_1, o_1) \times \prod_{i=1}^{10} \pi_{(S_{i+1}, O_{i+1})|(S_i, O_i)}(s_{i+1}, o_{i+1} | s_i, o_i).$$

The main difficulty here is the computation of these transition densities, which satisfy the Fokker-Planck (or forward Kolmogorov) partial differential equation.

We base on a different approach [28, pp. 118–121], [34]. The Euler-Maruyama scheme (14) and (15) is used, which gives rise to simple Gaussian transition densities, obviously at the expense of an error:

$$\begin{aligned} &\pi_{(S_{i+1}, O_{i+1})|(S_i, O_i)}(s_{i+1}, o_{i+1} | s_i, o_i) \\ &= \pi_{\mathrm{Normal}(s_i + b_1(s_i, o_i, t_i)(t_{i+1} - t_i), \sigma_1^2(t_{i-1} - t_i))}(s_{i+1}) \pi_{\mathrm{Normal}(o_i + b_2(s_i, o_i, t_i)(t_{i+1} - t_i), \sigma_2^2(t_{i-1} - t_i))}(o_{i+1}) \\ &= \frac{1}{\sigma_1 \sqrt{2\pi(t_{i+1} - t_i)}} \mathrm{e}^{-\frac{(s_{i+1} - s_i - b_1(s_i, o_i, t_i)(t_{i+1} - t_i))^2}{2\sigma_1^2(t_{i+1} - t_i)}} \times \frac{1}{\sigma_2 \sqrt{2\pi(t_{i+1} - t_i)}} \mathrm{e}^{-\frac{(o_{i+1} - o_i - b_2(s_i, o_i, t_i)(t_{i+1} - t_i))^2}{2\sigma_2^2(t_{i+1} - t_i)}}. \end{aligned}$$

The first density, $\pi_{(S_1,O_1)}(s_1, o_1)$, is the density of the two initial conditions S_0 and O_0. Since these conditions are constant, their density function is a Dirac delta function. As there is no dependence on the flow parameters, these delta functions are typically disregarded.

As a consequence, the log-likelihood is

$$\log \mathcal{L} \propto -10(\log\sigma_1 + \log\sigma_2) - \frac{1}{2\sigma_1^2}\sum_{i=1}^{10}\frac{(s_{i+1} - s_i - b_1(s_i, o_i, t_i)(t_{i+1} - t_i))^2}{t_{i+1} - t_i}$$
$$-\frac{1}{2\sigma_2^2}\sum_{i=1}^{10}\frac{(o_{i+1} - o_i - b_2(s_i, o_i, t_i)(t_{i+1} - t_i))^2}{t_{i+1} - t_i},$$

where $\propto$ denotes proportional, by omitting terms that do not depend on the flow parameters. By maximizing the log-likelihood (equivalently minimizing $-\log\mathcal{L}$) with respect to β, γ, ϵ, ρ, σ_1 and σ_2, with μ and K given by (2), the maximum likelihood estimates $\hat{\beta}$, $\hat{\gamma}$, $\hat{\epsilon}$, $\hat{\rho}$, $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are obtained. This approach is called MLE2. There is also the possibility of using the deterministic estimates (6), so that the log-likelihood is only maximized with respect to σ_1 and σ_2. This strategy is referred to as MLE1.

4. Numerical analysis

In this section, we deal with the stochastic model (12) computationally. After reviewing the main commands for working with stochastic differential equations in Mathematica®, the parameters are estimated by employing the methods described in the previous section. A comparison between the methods is performed by evaluating the correspondence between the real system and the mathematical model. A sensitivity analysis is also conducted.

For computational details, the reader is referred to [30]. The function *ItoProcess* defines a stochastic process that satisfies a stochastic differential equation. The standard Brownian motion is defined through *WienerProcess*. The instruction *RandomFunction* defines a realizable path, where the discretization of the time domain is specified. This *RandomFunction* can be evaluated at a particular time by using *"SliceData"*. When several paths are required to calculate statistics, the number of Monte Carlo simulations can be specified within *RandomFunction*. The statistics *Mean* and *StandardDeviation* are applied to *RandomFunction*. Minimization of a real function is conducted by *NMinimize*.

Apart from providing code, [30] also estimates parameters for the FitzHugh-Nagumo model (two drift and one diffusion parameters), by minimizing the mean square error with respect to the deterministic fit. We checked that such approach does not give good results in our case study, since the deterministic fit does not seem to be significantly improvable in terms of mean square error and MAPE. Further, it is our opinion that the incorporation of diffusion terms does not necessarily seek pointwise improvements, but the inclusion of prediction intervals that capture the uncertainty and variability of data.

Methods M1–M4 (moments) and MLE1, MLE2 (maximum likelihood), are applied. Recall that M1–M4 have the same formulas for the diffusion coefficients, but the differences arise from the flow parameters. Method M1 uses the deterministic estimates. Method M2 uses a zero mean value and a grouping of times. Method M3 uses third-order moments. Procedure M4 employs another statistic. On the other hand, MLE1 uses the deterministic fit and estimates only the diffusion coefficients by maximum likelihood. Strategy MLE2 estimates all of the parameters (drift and diffusion) by maximum likelihood.

For forward uncertainty quantification, 1000 realizations are used for the Monte Carlo simulation. In Table 1, the results are reported. The estimates of the flow parameters and the diffusion coefficients are tabulated. Also shown is the MAPE with respect to the mean value. Asymptotic values are reported (PI stands for prediction interval), by keeping the parameters time-independent. To analyze whether the prediction intervals capture the eleven measurements, the probability of obtaining each measurement or beyond is determined, and the number (#) of those probabilities that are less than 0.05 is shown.

	M1	M2	M3	M4	MLE1	MLE2
$\hat{\beta}$	0.369	0.282	0.0591	0.427	0.369	0.345
$\hat{\gamma}$	0.0223	0.0259	0.0588	0.0329	0.0223	0.0261
$\hat{\epsilon}$	0.0344	0.0451	0.133	0.0638	0.0344	0.0469
$\hat{\rho}$	0.241	0.184	0.0302	0.283	0.241	0.227
$\hat{\sigma}_1$	0.0202	0.0209	0.0392	0.0218	0.0196	0.0196
$\hat{\sigma}_2$	0.0157	0.0151	0.0190	0.0147	0.0149	0.0142
MAPE	3.2	3.3	7.2	3.6	3.2	3.2
$\mathbb{E}[S(\infty)]$	1.64	1.67	2.00	1.66	1.64	1.66
$\mathbb{E}[O(\infty)]$	1.06	0.96	0.881	0.86	1.07	0.84
$\mathrm{PI}_{S(\infty)}$	[1.59, 1.69]	[1.61, 1.73]	[1.78, 2.22]	[1.61, 1.71]	[1.59, 1.69]	[1.57, 1.76]
$\mathrm{PI}_{O(\infty)}$	[0.95, 1.18]	[0.86, 1.06]	[0.76, 0.99]	[0.77, 0.94]	[0.96, 1.18]	[0.68, 1.00]
$\#\mathbb{P}[s_i] < 0.05$	3	2	0	3	3	3
$\#\mathbb{P}[o_i] < 0.05$	0	0	1	0	0	0

Table 1: Results of the different methods.

For M1 and MLE1, the flow parameters coincide with (6) by definition. The rest of strategies present different estimates. Method M3 is the one that deviates more from the deterministic fit. The flow parameters from MLE2 are the most similar to those from the deterministic fit. All methods have small diffusion coefficients, lower than (10). In terms of MAPE, M3 is the worst method. Procedures M1, M2, MLE1, MLE2, and those from [18], have comparable MAPEs, while the MAPE of M4 is slightly higher. All methods present a similar number of outliers, except M3, which precisely was the approach with higher MAPE by far. Ideally, the number of outliers for 11 observations should be 0 or 1, which is not the case. Asymptotically, M1, MLE1 and Bayesian inference obtain close predictions.

Figure 1 plots the model outputs. The solid line is the expected value. The shaded region is the prediction with 0.95 probability. The circles are the real measurements. The upper profile corresponds to the overweight class, while the lower profile to the obese group. All of the methods render similar performance, except M3. In Figures 2 and 3, the probabilities of obtaining each measurement or beyond are determined ($\mathbb{P}[s_i]$ means $\mathbb{P}[S_i > s_i]$ if it is less than 0.5 or $\mathbb{P}[S_i < s_i]$ otherwise). They allow for assessing the suitability of the prediction intervals.

Model selection should be based on MAPE, quality of the bands, and model simplicity. It is clear that method M3 is not the best. M4 exhibits a high MAPE measure as well. Strategies M1, M2, MLE1 and MLE2 present comparable fit, also with respect to [18]. But take into account that M1 and MLE1 are the simplest, because they employ the deterministic fit for the flow parameters and only estimate the diffusion coefficients. In performance, the only significant difference between these four methods is the asymptotic behavior. Procedures M1 and MLE1

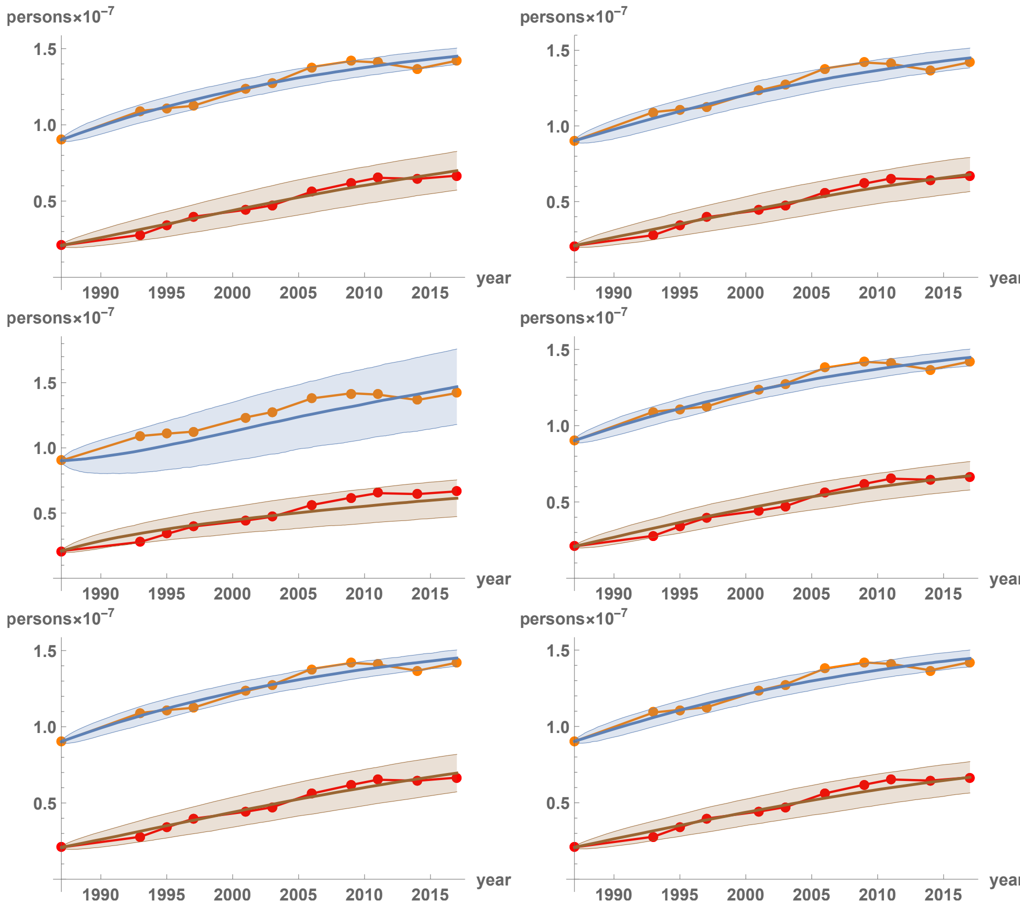

Figure 1: Predictions of the models (left-up panel is M1, right-up panel is M2, left-center panel is M3, right-center panel is M4, left-down panel is MLE1, right-down panel is MLE2). The solid line is the expected value. The shaded region is the prediction with 0.95 probability. The circles are the real measurements. The upper profile corresponds to the overweight class, while the lower profile to the obese group.

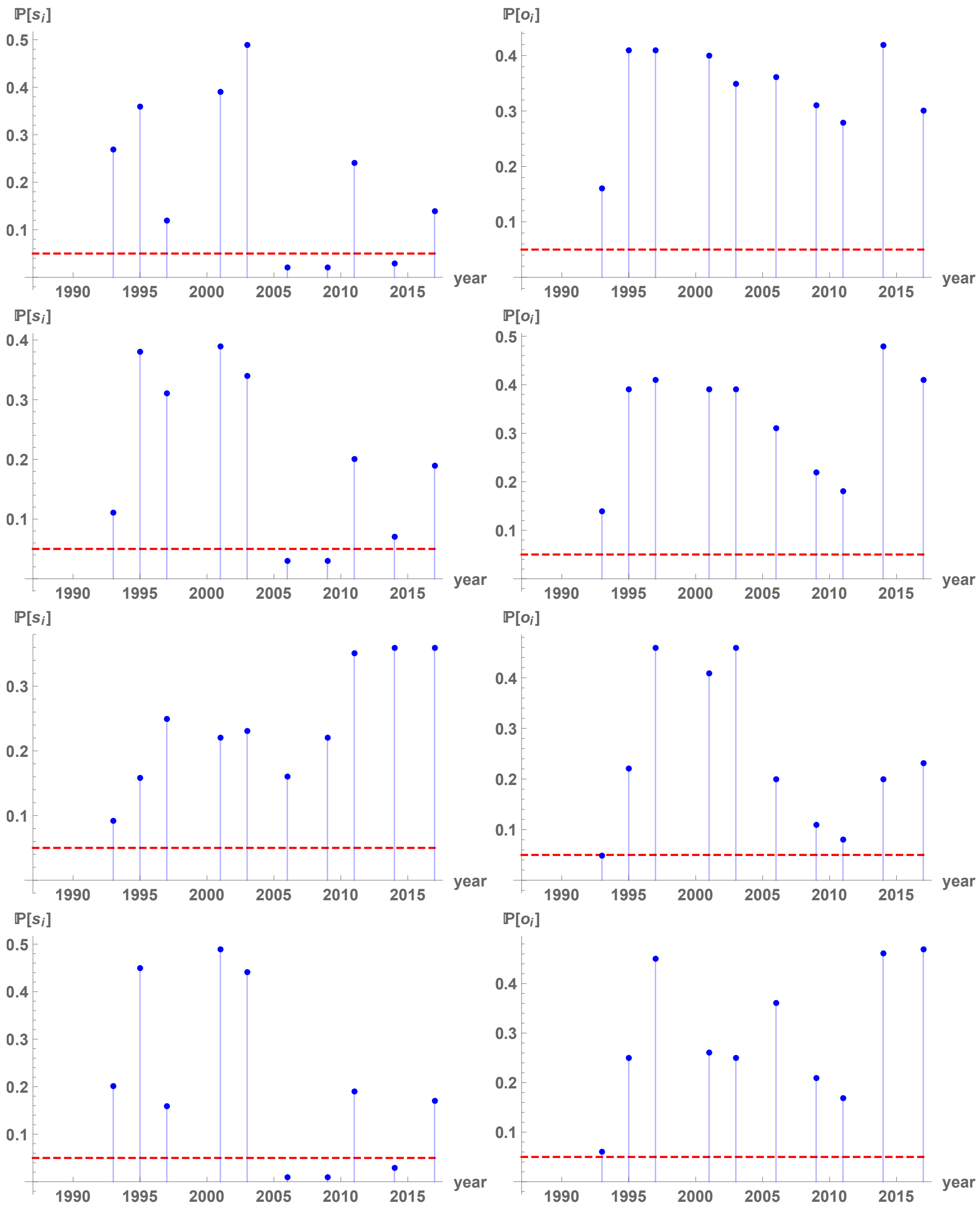

Figure 2: Probabilities of obtaining each measurement or beyond (M1–M4 from top to bottom). The threshold 0.05 is highlighted.

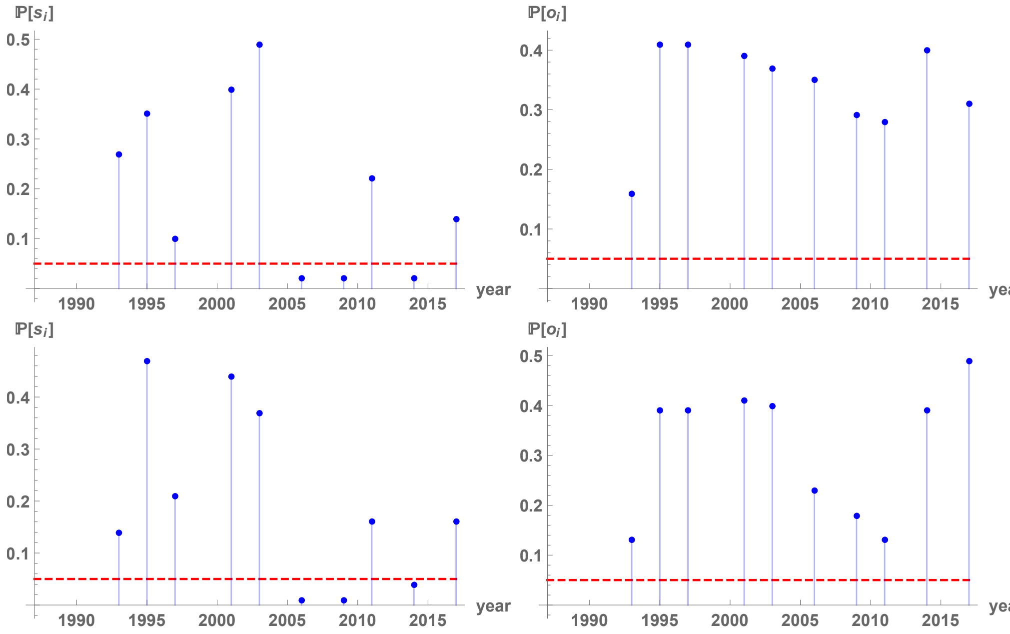

Figure 3: Probabilities of obtaining each measurement or beyond (MLE1 at the top and ML2 at the bottom). The threshold 0.05 is highlighted.

agree with Bayesian inference. Methods M2 and MLE2 are slightly different to them, especially regarding the asymptotic obese subpopulation which is rendered lower prevalence. Nonetheless, they are limited in the sense that the number of outliers is 2 (for M2) or 3, which might be too many for just 11 observations. The prediction intervals should ideally be wider.

A MAPE-based sensitivity analysis underscores, as in [18], the importance of prevention strategies. Each flow parameter is set to 0 and the inverse methods are applied. The higher the new MAPE is, the more important the removed parameter is. By order of higher influence, one has γ, β, ρ and ϵ. For example, within the framework of MLE2, the MAPEs are 21.9, 17.4, 7.1 and 5.8, respectively. The coefficient γ controls the flow from the overweight class to the obese class. The parameter β describes the movement from the normal weight group to the overweight group. Health-related communication campaigns [35, 36] should be implemented to prevent people from becoming unhealthier. This is the best approach to stop or at least alleviating the obesity epidemic.

5. Conclusion

Mathematical models are a useful tool to describe the evolution of diseases and assess the impact of control measures. Our case study has been the excess weight dynamics in the Spanish adulthood population. Data from 1987 to 2017, at eleven periods of time, have been available thanks to the ENSE and INE.

A recent contribution, [18], studied these data through a compartmental system of ordinary differential equations. The compartments were normal weight, overweight, and obesity. The parameters that control the flow between the subpopulations were estimated by minimizing the mean square error. Random versions of the proposed model were studied by means of frequentist nonlinear regression and Bayesian inference.

Frequentist nonlinear regression and Bayesian inference exhibit some drawbacks [18, 25]. The former entails inaccuracies due to linearization, clearly visible when extrapolation for forthcoming years is carried out. The latter, by contrast, is based on exact formulas, but in practice it is simulated by Markov Chain Monte Carlo algorithms, which may be complex and time-consuming.

In the present paper, the deterministic formulation of the model from [18] has been modified by adding Gaussian white noise random perturbations into the response derivatives. This has given rise to Itô stochastic differential equations driven by Brownian motions. The solution, which gives the number of overweight and obese adults in Spain at each year, has been a stochastic process. Existence and uniqueness of solution has been established, as well as moment bounds. From the Euler-Maruyama discretization, different inverse strategies have been applied for estimating the parameters: M1–M4, based on the moments method, and MLE1 and MLE2, relying on maximum likelihood estimation. The former has been based on obtaining information concerning the statistics and equating them to the sample statistics. The latter has been based upon Gaussian transition densities and a minimization procedure.

The numerical experiments have been summarized in Table 1 and Figures 1, 2 and 3. Methods M1, M2, MLE1 and MLE2 have presented the best fit, in terms of MAPE and prediction intervals. Procedures M1 and MLE1, which employ the deterministic fit for the flow parameters and only estimate the diffusion coefficients, are the simplest and have provided very similar results to the Bayesian inference from [18], avoiding its computational complexity and running time. Nonetheless, the four methods are limited in the sense that the number of outliers is 2 (for M2) or 3, which might be too many for just 11 observations. The prediction intervals should ideally be wider.

The MAPE-based sensitivity analysis has agreed with [18], which gives priority to prevention interventions over treatment strategies.

Conflict of Interest Statement

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

References

[1] World Health Organization (WHO). Obesity and overweight. Available at https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight (Accessed 18 September 2020).

[2] World Health Organization (WHO). Controlling the global obesity epidemic (Nutrition topics). Available at https://www.who.int/nutrition/topics/obesity/en/ (Accessed 18 September 2020).

[3] World Health Organization (WHO). 10 facts on obesity. Available at https://www.who.int/features/factfiles/obesity/en/ (Accessed 18 September 2020).

[4] James PT, Leach R, Kalamara E, Shayeghi M. The worldwide obesity epidemic. Obesity 2001;9:228–33.

[5] Hammond RA, Levine R. The economic impact of obesity in the United States. Diabetes Metab Syndr Obes 2010;3:285–295.

[6] Biener A, Cawley J, Meyerhoefer C. The High and Rising Costs of Obesity to the US Health Care System. J Gen Intern Med 2017;32(1):6–8.

[7] Krief B. Costes sociales y económicos de la obesidad y de sus patologías asociadas. Madrid: Gabinete de estudios Bernard Krief; 1999.

[8] Serdula MK, Ivery D, Coates RJ, Freedman DS, Williamson DF, Byers T. Do obese children become obese adults? A review of the literature. Prev Med. 1993;22(2):167–177.

[9] Wang Y, Beydoun MA, Liang L, Caballero B, Kumanyika SK. Will all Americans become overweight or obese? Estimating the progression and cost of the US obesity epidemic. Obesity (Silver Spring) 2008;16(10):2323–2330.

[10] Chitnis N, Schpira A, Smith D, Hay SI, Smith T, Steketee R. Mathematical modelling to support malaria control and elimination. Roll Back Malar Prog Impact Ser (World Health Organization, Progress & impact series) 2010;5:1–48.

[11] Brauer F. Compartmental models in epidemiology. In: Mathematical epidemiology. Berlin: Springer; 2008. p. 19–79.

[12] Ma Z, Li J. Dynamical modeling and analysis of epidemics. London: World Scientific; 2009.

[13] Acedo L, Moraño JA, Santonja FJ, Villanueva RJ. A deterministic model for highly contagious diseases: the case of varicella. Physica A 2016;450:278–286.

[14] Acedo L, Díez-Domingo J, Moraño JA, Villanueva RJ. Mathematical modelling of respiratory syncytial virus (RSV): vaccination strategies and budget applications. Epidemiol Infect 2010;138:853–860.

[15] Santonja FJ, Sánchez E, Rubio M, Morera JL. Alcohol consumption in Spain and its economic cost: a mathematical modeling approach. Math Comput Model 2010;52(7-8):999–1003.

[16] Jódar L, Santonja FJ, González GC. Modeling dynamics of infant obesity in the region of Valencia, Spain. Comput Math Appl 2008;56:679–689.

[17] Santonja FJ, Villanueva RJ, Jódar L, Gonzalez-Parra G. Mathematical modelling of social obesity epidemic in the region of Valencia, Spain. Math Comp Model Dyn 2010;16(1):23–34.

[18] Calatayud J, Jornet M. Mathematical modeling of adulthood obesity epidemic in Spain using deterministic, frequentist and Bayesian approaches. Chaos Soliton Fract 2020;140(110179).

[19] Spanish National Health Survey (ENSE) 2017. Spanish Ministry of Health, Consumption and Social Welfare. Available at `https://www.mscbs.gob.es/estadEstudios/estadisticas/encuestaNacional/encuesta2017.htm` (Accessed 18 September 2020).

[20] National Statistics Institute of Spain (INE). Population figures. Available at `https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=ultiDatos&idp=1254735572981` (Accessed 18 September 2020).

[21] Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. N Engl J Med 2007;357:370–379.

[22] Blanchower DG, Oswald AJ, Landeghem BV. Imitative obesity and relative utility. J Eur Econ Assoc 2009;7:528–538.

[23] Casabán MC, Cortés JC, Navarro-Quiles A, Romero JV, Roselló MD, Villanueva RJ. A comprehensive probabilistic solution of random SIS-type epidemiological models using the Random Variable Transformation technique. Commun Nonlinear Sci Numer Simulat 2016;32:199–210.

[24] Cervelló R, Cortés JC, Santonja FJ, Villanueva RJ. The dynamics over the next few years of the Spanish mobile telecommunications market share: A mathematical modelling approach. Math Comp Model Dyn 2014;20(6):557–565.

[25] Smith RC. Uncertainty quantification: theory, implementation, and applications. Philadelphia: SIAM; 2013.

[26] Lesaffre E, Lawson AB. Bayesian biostatistics. New York: Wiley, Statistics in Practice; 2012.

[27] Mao X. Stochastic differential equations and applications. Elsevier; 2007.

[28] Allen E. Modeling with Itô stochastic differential equations. Dordrecht, Netherlands: Springer Science & Business Media; 2007.

[29] Yin W, Cao J. Nonlocal stochastic differential equations with time-varying delay driven by G-Brownian motion. Math Method Appl Sci 2020;43(2):600–612.

[30] Paláncz B. Stochastic simulation and parameter estimation of the FitzHugh-Nagumo model. The Mathematica Journal 2016;18. Doi: 10.3888/tmj.18-6.

[31] Wolfram Research. 2019. Inc. Mathematica. Version 12.0, Champaign, Illinois.

[32] Lewis CD. Industrial and business forecasting methods. London: Butterworths; 1982.

[33] Higham DJ, Mao X, Stuart AM. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. SIAM J Numer Anal 2002;40(3):1041–1063.

[34] Elerian O. A note on the existence of a closed form conditional transition density for the Milstein scheme. Economics Discussion Papers 1998;W18.

[35] Snyder LB. Health communication campaigns and their impact on behavior. J Nutr Educ Behav 2007;39(2):S32–S40

[36] Bauman A, Smith BJ, Maibach EW, Reger-Nash B. Evaluation of mass media campaigns for physical activity. Eval Program Plann 2006;29(3):312–322 .