Stochastic Models to Estimate Population Dynamics

Saba Infante [1,2, *], Luis Sánchez [3,4], Aracelis Hernández [1,2]

[1] *School of Mathematical Sciences and Information Technology, Yachay Tech University, Ecuador.*
[2] *Department of Mathematics, Faculty of Science and Technology, University of Carabobo, Venezuela*
[3] *Departamento de Matemática y Estadística, Instituto de Ciencias Básicas, Universidad Técnica de Manabí, Ecuador*
[4] *Instituto de Matemática y Cálculo Aplicado, Facultad de Ingeniería, Universidad de Carabobo, Venezuela*

Abstract The growth dynamics that a population follows is mainly due to births, deaths or migrations, each of these phenomena is affected by other factors such as public health, birth control, work sources, economy, safety and conditions of quality of life in neighboring countries, among many others. In this paper is proposed two statistical models based on a system of stochastic differential equations (SDE) that model the dynamics of population growth, and three computational algorithms that allow the generation of probability distribution samples in high dimensions, in models that have non-linear structures and that are useful for making inferences. The algorithms allow to estimate simultaneously states solutions and parameters in SDE models. The interpretation of the parameters is important because they are related to the variables of growth, mortality, migration, physical-chemical conditions of the environment, among other factors. The algorithms are illustrated using real data from a sector of the population of the Republic of Ecuador, and are compared with the results obtained with the models used by the World Bank for the same data, which shows that stochastic models Proposals based on an SDE more adequately and reliably adjust the dynamics of demographic randomness, sampling errors and environmental randomness in comparison with the deterministic models used by the World Bank. It is observed that the population grows year by year and seems to have a definite tendency; that is, a clearly growing behavior is seen. To measure the relative success of the algorithms, the relative error was estimated, obtaining small percentage errors.

Keywords Markov Chain Monte Carlo, Secuential Monte Carlo, Particle Markov Chain Monte Carlo, Stochastic Differential Equations.

AMS 2010 subject classifications 62L10; 62L12; 62L20.

DOI: 10.19139/soic.v7i2.538

1. Introduction

Biological models are a multidisciplinary research field that studies the interactions between the internal and external elements of a real phenomenon. The process can be represented with a dynamic mathematical system; that is, by means of systems of ordinary differential equations, where it is assumed that the observed dynamics are exclusively determined by internal deterministic mechanisms. However, real biological systems are always exposed to external factors that are not completely controlled or not possible to model explicitly. Any biological process can be studied from the perspective of systemic biology, such as the growth of a cell, the interaction between two bacteria or the blood circulation in an organism. Therefore, the study of population dynamics as a system of interactions is part of the field of systemic biology. All the ecological theories regarding populations and communities have to do, in one way or another, with various aspects such as their sizes, how they grow or decrease, whether they show stable growth or fluctuate, how they respond to environmental changes and how they influence the environment of other populations. This suggests that an approximate way to include environmental

variability, without modeling additional environmental factors, is to vary the rates of births and deaths per capita at random. Therefore, is hypothesized that changes in the environment produce random changes in the birth and mortality rates per capita of a population that are independent of the changes due to demographic variability. These hypotheses can provide only an approximation to the actual biological situation.

However, accepting these hypotheses leads to manageable mathematical models that can give an idea of the effects of environmental variability on the dynamics of a population. However, the environment also varies and is likely to affect otherwise random populations. Therefore, one possible way to model environmental effects would be to include additional variables in the mathematical model, such as rainfall, predator populations, competing populations and food supply. In this sense, including additional variables would quickly complicate the model and destroy the simple nature of the process. The models that explain population dynamics are based on a structure of differential equations of the type:

$$\frac{dy_t}{dt} = f(y_t) - h(t) \tag{1}$$

where y_t represents the size of the population in t, $f(y_t)$ indicates the population growth rate and $h(t)$ represents the capture rate. The dynamic behavior of the population depends on the relationship between both rates $f(y_t)$ and $h(t)$. The population level declines yes $f(y_t) < h(t)$, then $\frac{dy_t}{dt} < 0$. Whereas, if $f(y_t) = h(t)$, the population remains at a constant level, because $\frac{dy_t}{dt} = 0$, which means that the natural growth rate $f(y_t)$ equals the value of sustainable yield, maintaining a constant population level. In this sense, [28] established that population dynamics are exponential growth, where a population without restrictions would increase geometrically. However, populations can not maintain this exponential growth indefinitely, which gave rise to models of restricted population growth or logistic models, deduced by [38]. Then followed extensions that include competition between different species, and interaction type predator-prey ([39], [27], [40], [12], [23]). Based on the works of [27], models were developed that consider the population classified by age groups, in order to solve the limitation of the models that treat all the individuals of the population in an identical way. However, very few models allow to predict the dynamics of the populations in an effective way. Then, the need to extend these deterministic processes to models that include more complex variations of the dynamic system, when a deterministic model is used to understand the dynamics of any of the aforementioned processes, there is a risk of obtaining deficient models, with residual correlated, errors on estimates, and with results of incorrectly interpreted inferences. One way to model the dynamics of the real world is by including a random noise in the model, and consequently a natural extension of the deterministic models is to move from an ordinary differential equation to an SDE, where the relevant parameters are modeled through of a probability distribution. Under these assumptions, it is assumed that the dynamic system of the process is partially driven by noise. The model includes a random noise in the measurement of dynamic system observations, and also includes factors that can not be controlled such as epidemiological diseases, variations in blood pressure, hormonal oscillations, respiration, neuronal control of variables that model activity muscle, enzymatic processes, cellular metabolism, nerve activity, molecule interactions, or individual characteristics such as body mass index, genes, smoking, stress impact ([8]).

Statistical models based on stochastic differential equations type Itô are defined through a continuous time Markov process, these models have come using recently to model stochastic population dynamics ([1], [2], [3], [4], [5], [10], [22] and [29]). The attractiveness of these models lies in their flexibility to model problems with complex dynamics, that have non-linear and erratic behaviors, that cause the system to become chaotic as the size of the variables increases making it unpredictable. On the other hand, the application of these SDE-based models becomes feasible, because efficient computational algorithms are available to estimate the solutions and parameters under all these circumstances.

The main contribution of this research is related to the proposed methodology based on the Markov chain Monte Carlo and sequential Monte Carlo algorithms, that allow us to make Bayesian inference on a large number of statistical models in which it was not possible to perform before; in particular, are considered mathematical models, that incorporate randomness and that allow to describe the dynamics of population growth. Stochastic effects can influence the dynamic system of the population, i.e. it may increase, decrease or even completely change the dynamic behaviour of birth and mortality per capita in a population. The results obtained by our proposal are

contrasted with the estimates obtained by the World Bank models, unlike of the models, these estimates are based on censuses that are interpolations or extrapolations based on demographic models.

The two stochastic models proposed in this work are: stochastic growth suggested by [32], and second stochastic model that includes environmental variability according to [6]. The structures of these models are widely used in many applications, such as: economics and finance ([9], [11], [14], [15], [17], [18] among others), in robotics, [24], biological systems, [14], epidemiology, [20], [21] and [26], and for precipitation models, [37], [18], [19], [35] and [36], among others. The implementation of these models is done following a Bayesian methodology under which three algorithms are proposed: Metropolis-Hasting (MH), Monte Carlo Sequential (SMC) and Particle Markov Chain Monte Carlo (PMCMC).The first two algorithms are widely known in the literature, however the method based on the PMCMC algorithm is more recent and was proposed by [7], and is used to estimate the solutions and parameters of a stochastic differential equation of Itô, this algorithm is based on a combination of the Markov chains Monte Carlo methods (MCMC) and SMC methods. The idea is to divide the original sampling problem into smaller and simpler sampling problems, by focusing on some of the subcomponents of the target distribution $q(.)$.

The rest of the article is organized as follows. In Section 2, the problem is introduced and the two models to be studied are proposed, in Section 3 the three computational algorithms used to estimate the growth projection of the population are developed, in Section 4 the results of illustrate the proposed methodology with real data and establish comparisons with standard methods, in Section 5 conclusions and discussions are established.

2. The problem

Natural phenomena are modeled generally in a suitable way by some Markov process which in turn can be written in the form of space state model. These models are constructed according to physical, chemical, environmental, or economic criteria, and on the bas is of incomplete observations or with noisy measurements. Markov processes are continuous-time processes with continuous trajectories but with measurements taken at a discrete time, and such events define the so-called partially observed diffusion processes ([20]). To define the general problem, the notation given in [32], and suppose that $x_t \in \mathbf{R}^m$ is a continuous time signal denoting the solution of the stochastic differential equation of Itô, given by:

$$dx_t = \Theta A(x)dt + CdB_t$$
$$x(0) = x_0 \tag{2}$$

Where

- $\{B_t, t > 0\}$ is a standard Brownian movement $m-$dimensional.
- $A : \mathbf{R}^k \to \mathbf{R}^l$ is a set of (possibly nonlinear) functions of Lipschitz globally known forces.
- The parameters that are estimated, the last m rows of the drift matrix (the first $k - m$ rows are known)
- $\Theta \in \mathbf{R}^{k \times 1}$ is a set of unknown parameters.
- $C \in \mathbf{R}^{k \times m}$ is a diffusion matrix.

The diffusion matrix $C \in \mathbf{R}^{k \times m}$ is of the form:

$$C = \begin{pmatrix} 0 \\ \Gamma \end{pmatrix} \tag{3}$$

Where $\Gamma \in \mathbf{R}^{m \times m}$ is a non-singular constant array. The estimated parameters in the drift and diffusion correspond to the coordinates, that are driven directly by the white noise.

Assuming, the standard regularity conditions of existence and uniqueness of the solution x_t of the stochastic differential equation given in (2) ([31]); it is also assumed that the process x_t is hypo-elliptical, this means that the noise propagates through the drift in all the components of the system given in (2).

The structure of the C matrix restricts noise in such a way, that it only acts on a subset of the variables that can be called rough variables; these variables can evolve through the coupling in the drift. The remaining variables of the system are called smooth variables. To distinguish between rough variables and smooth variables, is introduced the notation $x(t) = (u(t), v(t))^T$, where $u(t) \in \mathbf{R}^{k-m}$ is the smooth variable, and $v(t) \in \mathbf{R}^m$ is the rough variable.

The problem of interest is to consider u as an observed variable, and v as an unobserved variable, and estimate the parameters Γ and the entries of the Θ rows that correspond to the trajectories of the non-observed variables v, considering that $\{u_n\}_{n=0}^{N}$ is a known solution of the stochastic differential equation given in (2).

Under the Bayesian paradigm, a priori distribution is needed for the parameters Θ and $\Gamma\Gamma^T$, denoted as $p(\Theta, \Gamma\Gamma^T)$, and a likelihood for the data, denoted as $L(u, v|\Theta, \Gamma\Gamma^T)$, then the posterior distribution for the parameters can be estimated as:

$$p(\Theta, \Gamma\Gamma^T | u, v) \propto L(u, v|\Theta, \Gamma\Gamma^T) p(\Theta, \Gamma\Gamma^T) \tag{4}$$

In this type of problem, usually the exact likelihood of the transition density is unknown, so it is natural to use approximations to be able to make inference. Suppose that for the Markov sequence $\{x_t\}_{t=0}^{T}$, the transition map of $t \to t + 1$ is determined by:

$$x_{t+1} = x_t + \int_{t\Delta t}^{(t+1)\Delta t} \Theta A \{x_s\} ds + \int_{t\Delta t}^{(t+1)\Delta t} C dB_s \tag{5}$$

Then, the Euler-Maruyama approximation of this map is given by:

$$x_{t+1} \approx x_t + \Delta t \Theta A(x_t) + \sqrt{\Delta t} R(0, \Theta) \xi_t \tag{6}$$

Where $x_t \in \mathbf{R}^k$, $\xi_t \in \mathbf{R}^k$, $\xi_t \sim N(0, I)$, I is the identity matrix, and:

$$R(0, \Theta) = \begin{pmatrix} 0 & 0 \\ 0 & \Gamma \end{pmatrix} \in \mathbf{R}^{k \times k} \tag{7}$$

It is a non-invertible matrix. To prevent $R(0, \Theta)$ from being invertible, the following approximation is used:

$$x_{t+1} \approx x_t + \Delta t \Theta A(x_t) + \sqrt{\Delta t} R(\Delta t, \Theta) \xi_t \tag{8}$$

Where $R(\Delta t, \Theta)$ is invertible, and is given by:

$$R(\Delta t, \Theta) = \begin{pmatrix} \Delta t & \Delta t \\ \Delta t & \Gamma \end{pmatrix} \in \mathbf{R}^{k \times k} \tag{9}$$

In general, the solution x_t is observed with errors through a continuous observation process y_t, that is related to the process x_t through the following expression:

$$y_t = g(x_t, \Theta) + \epsilon_t \quad , \quad \epsilon_t \sim N(0, \sigma_\epsilon^2) \tag{10}$$

Where $g(.)$ is a vector function of known real value. Then, the model under study can be represented by:

$$dx_t = \Theta A(x)dt + C dB_t$$

$$y_t = g(x_t, \Theta) + \epsilon_t \quad , \quad \epsilon_t \sim N(0, \sigma_\epsilon^2) \tag{11}$$

When the dynamic system of continuous time is approximated by a discrete time system, the derivatives in the continuous time domain are approximated by an equation in difference in discrete time, that is:

$$x_t = f_1(x_{t-1}, \Theta) + \zeta_{1t} \quad , \quad \zeta_{1t} \sim N(0, \sigma_\zeta^2) \tag{12}$$

$$y_t = g_1(x_t, \Theta, t) + \epsilon_{1t} \quad , \quad \epsilon_{1t} \sim N(0, \sigma_\epsilon^2) \tag{13}$$

The models given by (12) and (13) are known as the equation of state and the equation of observation, respectively. Using Bayesian algorithms is estimated the a posteriori marginal distribution $p(x_t|y_{1:t})$ in time t in two steps, prediction and update, where $x_{0:t} = (x_0, \ldots, x_t)$ are the unknown solutions states and $y_{1:t} = (y_1, \ldots, y_k)$ are the observed solutions states.

To carry out the recursive procedure, it is assumed that the a posteriori marginal density $p(x_{t-1}|y_{1:t-1})$ is known at a time t. Then, using the equation (12) is obtained the updated density $p(x_t|y_{1:t-1})$ of the state in time t, and using the equation of Chapman-Kolmogorov transition:

$$p(x_t|y_{1:t-1}) = \int p(x_{t-1}|y_{1:t-1})p(x_t|x_{t-1})dx_{t-1} \tag{14}$$

When a new observation y_t is recorded, the posterior density is updated using the equation (13) and the Bayes Theorem as follows:

$$p(x_t|y_{1:k}) = \frac{p(x_{t-1}|y_{1:t-1})p(y_t|x_t)}{\int p(x_{t-1}|y_{1:t-1})p(y_t|x_t)dx_t} \tag{15}$$

In the [32] is considered the solution $x(t) = (u(t), v(t))^T$ divided into two parts, a known solution u_t, and a solution v_t unknown, so that this stochastic process can be seen as a partially observed latent process. The general objective is to estimate the missing solution and the parameters of the dynamic system proposed in the equation (2). The representation in the form of state space is given by:

$$v_1 \sim p(.)$$
$$\Theta_t = \Theta_{t-1} + \omega_t \quad , \quad \omega_t \sim N(0, \sigma_\omega^2)$$
$$v_t = f_2(v_{t-1}, \Theta_t) + \zeta_{2t} \quad , \quad \zeta_{2t} \sim N(0, Q)$$
$$u_t = g_2(v_t, \Theta_t) + \epsilon_{2t} \quad , \quad \epsilon_{2t} \sim N(0, R) \tag{16}$$

Where f_2 is a function, that maps the states within the same state space, g_2 is a function, that maps the state variables within the observed variables, and v_0 is a variable random with probability distribution given by $p(.)$.

The joint probability distribution of the parameters Θ, the latent process $v_{1:t} = (v_1, \ldots, v_t)$, and the observed solutions $u_{1:t} = (u_1, \ldots, u_t)$ is given by:

$$p(u_{1:t}, v_{1:t}, \Theta) = p(v_{1:t}|\Theta)p(u_{1:t}|v_{1:t}, \Theta)p(\Theta|u_{1:t})$$

$$p(u_{1:t}, v_{1:t}, \Theta) = p(v_1) \left(\prod_{i=2}^n p(v_i|v_{i-1}, \Theta) \right) \left(\prod_{i=1}^t p(u_i|v_i, \Theta) \right) p(\Theta|u_{1:t}) \tag{17}$$

The latent process is not observed, and the interest is to estimate the marginal posterior distribution:

$$p(\Theta|u_{1:t}) \propto p(u_{1:t}, \Theta)$$

$$= p(u_t|u_{1:t-1}, \Theta)p(u_{1:t-1}, \Theta)$$

$$= p(u_1, \Theta) \prod_{k=1}^t p(u_k|u_{1:k-1}, \Theta) \tag{18}$$

Where:

$$p(u_t|u_{1:t-1}, \Theta) = \int p(u_t|v_t, \Theta)p(v_t|u_{1:t-1}, \Theta)dv_t$$

$$= \int p(u_t|v_{t-1}, \Theta)p(v_{t-1}|u_{1:t-1}, \Theta)dv_{t-1} \tag{19}$$

And:

$$p(u_1, \Theta) = \int p(u_1|v_1, \Theta)p(v_1)dv_1 \tag{20}$$

In general, the latent process can not be integrated, so MCMC and SMC algorithms are required to explore the space in parameter set Θ, and latent process, $v_{1:t}$. The ideal would be to integrate the latent process by calculating the marginal verisimilitude given in the equation (19), then the MCMC algorithm could be run only in the parameter space.

In particular, the problem that is proposed to be solved in this work, consists in the adjustment of two stochastic models to project the dynamics that population growth follows:

A. Langivin model of second order for stochastic growth

In this model, the problem of systems of second-order hypoeliptical equations discussed in [32] and that have a physical meaning where u is considered to be the position and v is the momentum, and the general form of Langevin's second order equation given by:

$$dv = udt \tag{21}$$

$$du = (-\gamma u + f(v, D)) \, dt + \sigma dB \tag{22}$$

Where f is a force function (possibly nonlinear) parameterized by D, dB is a Brownian Movement, and the variables v and u are scalars, while the parameters γ, D, σ and the states must be estimated.

In this paper is considered a particular case of this model known as a stochastic growth model for which $x = (v, u)^T$ satisfies:

$$dv = udt \tag{23}$$

$$du = \sigma dB \tag{24}$$

In this case, the process has a single diffusion parameter, σ, which describes the size of the fluctuations, and which must be estimated in conjunction with the unknown solutions states, v_t. Expressing the model (24) in terms of the equation (2), is obtained that:

$$A(x) = x \quad ; \quad \Theta = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad ; \quad C = \begin{pmatrix} 0 \\ \sigma \end{pmatrix} \tag{25}$$

Then, the stochastic growth model in its discretized form uses a order first approximation of Euler-type with a step Δt for the equation (23), and for the second equation (24) the usual approximation of a stochastic integral is used, obtaining the following model:

$$\sigma_t = \sigma_{t-1} + \varepsilon_1 \quad ; \quad \varepsilon_1 \sim N(0, \sigma_{\varepsilon_1}^2)$$
$$v_{t+1} = v_t + u_t \Delta t + \varepsilon_2 \quad ; \quad \varepsilon_2 \sim N(0, \sigma_{\varepsilon_2}^2)) \tag{26}$$
$$u_{t+1} = u_t + \sigma_t (B_{t+\Delta t} - B_t) \quad ; \quad B_{t+\Delta t} - B_t \sim N(0, \sigma_{(B_{t+\Delta t} - B_t)}^2)$$

The parameters and states to be estimated σ_t and $v_{1:T} = (v_1, \ldots, v_T)$, respectively.

B. Differential equations system

Following [6], a system of differential equations is considered to model the stochastic growth of a population, but including environmental variability.

This system of differential equations is given:

$$dy_t = (b(t)y(t) - d(t)y(t))dt + \sqrt{(b(t)y(t) - d(t)y(t))}dW_1(t)$$
$$db_t = \beta_1(b_e - b(t))dt + \alpha_1 dW_2(t) \tag{27}$$
$$dd_t = \beta_2(d_e - d(t))dt + \alpha_2 dW_3(t)$$

With $(y_t, b_t, d_t) \in [0, \infty) \times \mathbf{R} \times \mathbf{R}$, y_t represents the size of the population, b_t represents the birth rate per capita, d_t represents the death rate per capita, b_e represents the average per capita birth rate, d_e represents the average death rate per capita, $W_i(t), i = 1, 2, 3$, are independent standard Wiener processes. The parameters are β_1, β_2, α_1 and α_2, but if environmental variability does not occur, then the value of the parameters is considered as $\beta_1 = \beta_2 = \alpha_1 = \alpha_2 = 0$.

Discretizing the model using the first-order Euler approximation, is obtained:

$$y_{t+1} = y_t + (b_t y_t - d_t y_t)\Delta_t + \sqrt{(b_t y_t - d_t y_t)}\sqrt{\Delta_t}\varepsilon_1 \quad ; \quad \varepsilon_1 \sim N(0, \sigma_{\varepsilon_1}^2)$$
$$b_{t+1} = b_t + \beta_1(b_e - b_t)\Delta_t + \alpha_1\sqrt{\Delta_t}\varepsilon_2 \quad ; \quad \varepsilon_2 \sim N(0, \sigma_{\varepsilon_2}^2) \tag{28}$$
$$d_{t+1} = d_t + \beta_2(d_e - d_t)\Delta_t + \alpha_2\sqrt{\Delta_t}\varepsilon_3 \quad ; \quad \varepsilon_3 \sim N(0, \sigma_{\varepsilon_3}^2)$$

Where the parameters and states to be estimated are β_1, β_2, α_1, α_2 and $y_{1:T} = (y_1, \ldots, y_T)$, respectively.

3. Methodology

In this paper, is proposed to use three computational algorithms based on the MCMC methods and the SMC methods, to estimate the projection of the growth dynamics of a population. The algorithms proposed are:

- Metropolis-Hastings (MH)
- Monte Carlo sequential (SMC)
- Particle Markov chain Monte Carlo (PMCMC)

3.1. Markov chains Monte Carlo Algorithm

The MCMC algorithm used in this work, the Metropolis-Hastings (MH) algorithm is proposed by [30] and [16] and dealt extensively with [33], [25], [13] and [34], among others.

Suppose is wanted to design an MH algorithm, whose stationary distribution is $p\left(v_{1:t}, \Theta | u_{1:t}\right)$, where $v_{1:t}$ are latencies unknown and $u_{1:t}$ are the observations. Then in iteration $l + 1$, is proposed a candidate $(v_{1:t}^c, \Theta^c)$ with proposed distribution $q\left(v_{1:t}^c, \Theta^c | v_{1:t}^{(l)}, \Theta^{(l)}\right)$ and accept this with probability:

$$p\left(v_{1:t}^c, \Theta^c | v_{1:t}^{(l)}, \Theta^{(l)}\right) = \min\left\{1, \frac{q(v_{1:t}^{(l)}, \Theta^{(l)} | v_{1:t}^c, \Theta^c)p(u_{1:t}|v_{1:t}^c, \sigma)p(v_{1:t}^c|\Theta^c, \gamma)p(\Theta^c, \beta)}{q(v_{1:t}^c, \Theta^c | v_{1:t}^{(l)}, \Theta^{(l)})p(u_{1:t}|v_{1:t}^{(l)}, \sigma)p(v_{1:t}^{(l)}|\Theta^{(l)}, \gamma)p(\Theta^l, \beta)}\right\} \tag{29}$$

Assuming that the following distribution can be used as the proposed distribution:

$$q\left(v_{1:t}^c, \Theta^c | v_{1:t}^{(l)}, \Theta^{(l)}\right) = q(\Theta^c|\Theta^{(l)})q(v_{1:t}^c|u_{1:t}, \Theta^c, \kappa) \tag{30}$$

Then, the probability of acceptance is simplified as follows:

$$\alpha = p\left(v_{1:t}^c, \Theta^c | v_{1:t}^{(l)}, \Theta^{(l)}\right) = \min\left\{1, \frac{q(\Theta^{(l)}|\Theta^c)p(u_{1:t}|\Theta^c, \gamma, \sigma)p(\Theta^c, \beta)}{q(\Theta^c|\Theta)p(u_{1:t}|\Theta^{(l)}, \gamma, \sigma)p(\Theta, \beta)}\right\} \tag{31}$$

However, two difficulties arise here. To approximate the equations (29) and (30), both the conditional distribution or filtered distribution $p(v_{1:t}|u_{1:t}, \Theta, \kappa)$, and the marginal distribution $p(u_{1:t}|\Theta, \gamma.\sigma)$, but both distributions are generally unknown, and in turn are required for can calculate probability of acceptance α.

The algorithm is summarized below:

Algorithm 1: Metropolis-Hastings algorithm

- **Step 1.** Initialize:

$$v_0 = \left(v_1^{(0)}, \ldots, v_N^{(0)}\right) \tag{32}$$

   And

$$\Theta_0 = \left(\Theta_1^{(0)}, \ldots, \Theta_N^{(0)}\right) \tag{33}$$

- **Step 2.** For a time: $l = 1$ until N, generate the candidate values:

$$\Theta^c \sim q(\Theta^c|\Theta_{(l-1)}) \tag{34}$$

   And

$$v_{1:t}^c \sim q(v_{1:t}^c|v_{1:t}^{(l-1)}, \Theta^c, \kappa) \tag{35}$$

   The current parameters $v_{1:t}^c$ and Θ^c are changed to $v_{1:t}^{(l-1)}$ and $\Theta_{(l-1)}$, respectively, with probability of acceptance α, and in another case the string remains in the current parameters.
- **Step 3.** Generate $u_l \sim U(0, 1)$.
- **Step 4.** If $u_l < \alpha$, then it is done $v_{1:t}^{(l)} = v_{1:t}^c$ and $\Theta_l = \Theta^c$, otherwise it is done $v_{1:t}^{(l)} = v_{1:t}^{(l-1)}$ and $\Theta_l = \Theta_{l-1}$.
- **Step 5.** It increases from l to $l + 1$, and is returned 2 step by step.

3.2. Sequential Monte Carlo Algorithm

The SMC procedures are a class of algorithms, that are used to sequentially densities approximate of the posterior densities $\{p(v_{1:t}|u_{1:t}); n \geq 1\}$, as well as marginal likelihood sequences $\{p_\theta(u_{1:t}|u_{1:t}); n \geq 1\}$ for $\theta \in \Theta$ given. The SMC algorithm generates a set of K particles $\{v_{1:t}^{(l)}\}_{l=1}^{K}$ and their respective weights $\{W_J^{(l)}\}_{l=1}^{K}$, to approximate the conditional distribution $p(v_{1:t}|u_{1:t})$, using an empirical measure:

$$\tilde{p}_\theta(dv_{1:t}) := \sum_{l=1}^{K} W_J^{(l)} \delta_{v_{1:t}^{(l)}}(dv_{1:t}) \tag{36}$$

Where $W_J^{(l)}$ is called weight of importance associated with the particles $v_{1:t}^{(l)}$, and δ is the delta function of Dirac. The SMC algorithm is summarized below:

Algorithm 2: SMC algorithm

- **Step 1.** All time $j = 0$,
  - Sample $v_0^{(l)} \sim \Pi_0(.|\Theta)$, for all $l = 1, \ldots, K$.
  - Compute and normalize the weights:

$$w_0^{(l)} = p(u_0, v_0^{(l)}|\Theta, \gamma, \sigma) \tag{37}$$

   And

$$W_0^{(l)} = \frac{w_0^{(l)}}{\sum_{l=1}^{K} w_0^{(l)}} \tag{38}$$

- **Step 2.** Time $j = 2, \ldots, J$:
  - Sample K iid variables $v_{0:j-1}^{(l)}$ according to the distribution:

$$\tilde{p}_\theta(dv_{1:j-1}) := \sum_{l=1}^{K} W_{j-1}^{(l)} \delta_{v_{1:j-1}^{(l)}}(dv_{1:j-1}) \tag{39}$$

  - Then, for each particle $l = 1, \ldots, K$, sample:

$$v_j^{(l)} \sim q_{SMC}(.|u_j, v_{0:j-1}^{(l)}, \Theta, \gamma, \sigma) \tag{40}$$

   (i.e propagate the particle), and set $v_{0:j}^{(l)} = (v_{0:j-1}^{(l)}, v_j^{(l)})$
  - Finally compute and normalize the weights:

$$w_j^{(l)} = \frac{p(u_j|v_j^{(l)}, \sigma)p(v_j^{(l)}|v_{j-1}^{(l)}, \Theta, \gamma)}{q_{SMC}(v_j^{(l)}|u_j, v_{0:j-1}^{(l)}, \Theta, \gamma, \sigma)} \tag{41}$$

   And

$$W_j^{(l)} = \frac{w_j^{(l)}}{\sum_{l=1}^{K} w_j^{(l)}} \tag{42}$$

The simulation of one trajectory $v_{1:J} = (v_1, \ldots, v_J)$ called run of SMC algorithm under the approximation of $p(v_{1:t}|u_{1:t}, \Theta, \gamma, \sigma)$, it is directly achieved by randomly one particle $v^{(l)}$ among the K particles with weights $\{W_j^{(l)}\}_{l=1}^{K}$. Besides, the marginal distribution $p(u_{1:t}|\Theta, \gamma, \sigma)$ can be estimated through the weights:

$$\tilde{p}(u_{1:t}|\Theta, \gamma, \sigma) = \Pi_{j=0}^{J}\left(\frac{1}{K}\sum_{l=1}^{K} w_j^{(l)}\right) \tag{43}$$

3.3. Particle Markov Chain Monte Carlo Algorithm

The PMCMC algorithms are exact approximations of a posteriori distributions $p(v_{1:t}|u_{1:t})$. These algorithms combine the SMC algorithms with the MCMC methods to approximate both $p(v_{1:t}|u_{1:t})$ or $p(\Theta, v_{1:t}|u_{1:t})$, in the sense that for any fixed number of particles $N \geq 1$ transition densities leave invariant target density. The SMC algorithm is an approximate simulation procedure for the target density $p(v_{1:t}|u_{1:t})$, and the latter is used as the proposed distribution for the MH algorithm.

The PMCMC methods are used to simultaneously estimate the parameters and states in non-linear spatial state models. This procedure can be interpreted as a standard update of the MH method, which leads to convergent algorithms.

The key to this PMCMC algorithm is to approximate the target density $p(v_{1:t}|u_{1:t})$, by choosing an optimal proposed density $q(v_{1:t}|u_{1:t}) = \tilde{p}(v_{1:t}|u_{1:t})$ and implementing a $\tilde{p}(v_{1:t}|u_{1:t})$ within a step of the MH algorithm, which leads to a familiar and simple method. Previously, an SMC simulation procedure is run to approximate $p(v_{1:t}|u_{1:t})$ but this procedure can not be implemented directly, since when calculating the acceptance rate of the MCMC algorithm, this it is required evaluating the marginal density, that is generated from an SMC algorithm.

Consider $v_{1:t}^*$ a candidate sample that updates the state $v_{1:t}$, with probability of acceptance given by:

$$\alpha(v_{1:t}, v_{1:t}^*) = \min \left\{ 1, \frac{p(v_{1:t}^*|u_{1:t})q(v_{1:t}|u_{1:t})}{p(v_{1:t}|u_{1:t})q(v_{1:t}^*|u_{1:t})} \right\} \tag{44}$$

The optimal choice of $q(v_{1:t}|u_{1:t})$ is:

$$q(v_{1:t}|u_{1:t}) = p(v_{1:t}|u_{1:t}) \tag{45}$$

However, in practice this choice becomes impossible. The developments obtained for the SMC methods ([7]), suggest the idea of using the approximation of $p(v_{1:t}|u_{1:t})$ as the proposed density, this means, that a sample is extracted as follows:

$$\hat{p}(dv_{1:t}|u_{1:t}) := \sum_{l=1}^{N} w_j^{(l)} \delta_{v_{1:t}^{(*l)}}(dv_{1:t}) \tag{46}$$

Where δ is the Dirac delta function. Given the samples and weights $\left\{ W_j^{(l)}, v_{1:t}^{(l)}, l = 1, \ldots, N \right\}$, for calculating the probability of acceptance is required the expression of the marginal distribution of $v_{1:t}^*$, which is difficult to estimate, however the SMC algorithm allows is obtained an unbiased estimator of marginal likelihood:

$$\hat{p}(v_{1:t}^*|u_{1:t}) = \prod_{j=0}^{J} \left(\frac{1}{K} \sum_{l=1}^{N} w_j^{(l)} \right) \tag{47}$$

This is a low cost algorithm because it requires very little input information, and provides satisfactory results, is needed to specify densities of importance in one dimension:

$$q(v_1|u_1) = p(v_1)$$

$$q(v_n|u_n, v_{n-1}) = p(v_n|v_{n-1}) \tag{48}$$

For $N \geq 2$, but when T is too large the algorithm has degeneration problems due to successive resampling steps, that decrease the number of different values for v_t^*.

Algorithm 3: Particle Independent Metropolis-Hasting (PIMH) algorithm

- **Step 1.** Initialize $i = 0$ and run an SMC algorithm choosing a proposed density $q(v_{1:t}|u_{1:t}) = \hat{p}(v_{1:t}|u_{1:t})$, sampling; $v_{1:t} \sim \hat{p}(v_{1:t}|u_{1:t})$, and be $\hat{p}(v_{1:t})(0)$ the marginal likelihood estimator.
- **Step 2.** For $i \geq 1$,

– Run an SMC algorithm to estimate $p(v_{1:t}|u_{1:t})$, sampling $v_{1:t}^* \sim \hat{p}(v_{1:t}|u_{1:t})$, and let $\hat{p}(v_{1:t})^*$ be the marginal likelihood estimator.

– Then with probability:

$$\alpha = \min \left\{ 1, \frac{\hat{p}(v_{1:t})^*}{\hat{p}(v_{1:t})(i-1)} \right\} \tag{49}$$

It becomes $v_{1:t}(i) = v_{1:t}^*$, and $\hat{p}(v_{1:t})(i) = \hat{p}(v_{1:t})^*$, otherwise it becomes $v_{1:t}(i) = v_{1:t}(i-1)$, and $\hat{p}(v_{1:t})(i) = \hat{p}(v_{1:t})(i-1)$.

In [7] it was proved that the update of the PIMH sampler keeps the distribution afterwards $p(v_{1:t}|u_{1:t})$ invariant, and that under weak criteria such as the partial marginal $\hat{p}(v_{1:t})^*$, and $\hat{p}(v_{1:t})(i-1)$ represent consistent estimators of the marginal $p(v_{1:t})$, then the PIMH sampler is ergodic and converges to 1 when $N \to \infty$.

If the target distribution to be approximated is $p(\Theta, v_{1:t}|u_{1:t})$, it is considered:

$$p(\Theta, v_{1:t}|u_{1:t}) = p(\Theta|u_{1:t})p(v_{1:t}|u_{1:t}) \tag{50}$$

Then, a natural choice of the proposed density to update the MH is:

$$q(\Theta^*, v_{1:t}^*|\Theta, u_{1:t}) = q(\Theta^*|\Theta)q(v_{1:t}^*|u_{1:t}) \tag{51}$$

Then the probability of acceptance is:

$$\alpha = p \left(\frac{p(\Theta^*, v_{1:t}^*|u_{1:t})q(\Theta, v_{1:t}^*|\Theta^*, u_{1:t})}{p(\Theta, v_{1:t}^*|u_{1:t})q(\Theta^*, v_{1:t}^*|\Theta, u_{1:t})} \right) \tag{52}$$

$$\alpha = \frac{p(v_{1:t})^*p(\Theta^*)q(\Theta|\Theta^*)}{p(v_{1:t})p(\Theta)q(\Theta^*|\Theta)} \tag{53}$$

The PMCMC algorithm is detailed below:

Algorithm 4: PMCMC algorithm

- **Step 1.** It is initialized with $i = 0$, be $\Theta(0)$ an arbitrary distribution.
    – Run an SMC algorithm targeting $p_{\Theta(0)}(v_{1:t}|u_{1:t})$, sampling $v_{1:t}(0) \sim \hat{p}_{\Theta(0)}(v_{1:t}|u_{1:t})$, and defining $\hat{p}_{\Theta(0)}(u_{1:t})$ as the marginal likelihood estimator.

- **Step 2.** For $i \geq 1$,
    – Sample $\Theta^* \sim q(\Theta^*|\Theta(i-1))$.
    – Run an SMC algorithm targeting $p_{\Theta^*}(v_{1:t}|u_{1:t})$, sample $v_{1:t}^* \sim \hat{p}_{\Theta^*}(v_{1:t}^*|u_{1:t})$, and let $\hat{p}_{\Theta^*}(u_{1:t})$ the marginal likelihood estimator.
    – Calculate the probability of acceptance:

$$\alpha = \min \left\{ 1, \frac{\hat{p}_{\Theta^*}(u_{1:t})p(\Theta^*)q(\Theta(i-1)|\Theta^*)}{\hat{p}_{\Theta(i-1)}(u_{1:t})p(\Theta(i-1))q(\Theta^*|\Theta(i-1))} \right\} \tag{54}$$

It becomes $\Theta(i) = \Theta^*$, $v_{1:t}(i) = v_{1:t}^*$, and $\hat{p}_{\Theta(i)}(u_{1:t})(i) = \hat{p}_{\Theta^*}(u_{1:t})$; otherwise, $\Theta(i) = \Theta(i-1)$, $v_{1:t}(i) = v_{1:t}(i-1)$ and $\hat{p}_{\Theta(i)}(u_{1:t})(i) = \hat{p}_{\Theta(i-1)}(v_{1:t})(i-1)$.

4. Results

To illustrate the methodology proposed in this paper, real data is used. As discussed in the previous sections, two stochastic population growth models defined in [32] and [6], respectively, with the objective of forecasting the population growth of male and female aged between 35 and 39 years of the Republic of Ecuador, for a range that

ranges from year 1960 up to 2035. These results are contrasted with the estimates obtained by the World Bank models, which are generally based on national censuses of the population. The estimates for the years before and after the censuses are interpolations or extrapolations based on demographic models. The data is used property of the World Bank and is available on the website: https://data.worldbank.org/. To compare the estimates of the models, the percentage error (EP), which is a measure of the relative error expressed in percentage, with the estimated states $\hat{y}$ and the true values observed y_i. The percentage of the error can be used to compare each error in terms of 100%. The EP is defined as:

$$EP = |\frac{y_i - \hat{y}_i}{y_i}|100\% \tag{55}$$

4.1. Langiven stochastic growth model

In order to adjust Langiven's stochastic growth model, the data from the population growth of male and female in the Republic of Ecuador were considered. In the Table (1), the a priori distributions for the adjustment of the stochastic population growth model are indicated.

In the Figures (1) and (2), the estimates of the states (growth of the considered population of Ecuador) are shown

Parameters	MH	SMC	PMMH
σ_0	0	1	1
ϵ_1	100	1	10
ϵ_2	1000	10	1
ϵ_3	1000	1000	10
Δt	0.03	0.03	0.04

Table 1. A priori values to initialize the parameters of the MH, SMC, and PMMH algorithms, for the female and male data.

by the algorithms proposed in this work: MH, SMC and PMMH, together with the projections made by the World Bank model. In these graphs and tables (2) and (3), it can be observed that the algorithms proposed have a good performance in the prediction of the population up to the year 2035, taking as reference the model of the World Bank but from there onwards one can observe instability in the projections. The estimates provided by the World Bank come from demographic models and, therefore, are susceptible to biases and errors due to deficiencies in both the model and the data. For various reasons, a stochastic model is preferable to a deterministic one to simulate the growth of real populations, since there may be an implicit demographic randomness because the population is composed of a finite number of individuals, that are subject to random events whose effects may be relevant in small populations due to the risk of extinction. On the other hand, there is always a randomness due to sampling since populations are estimated by random sampling subject to measurement errors. In addition, there is an environmental randomness that is given by external factors such as climate, epidemics that can affect populations, and these factors have a random behavior. Consequently, stochastic models allow to better understand and investigate the influence of noise on the dynamics of population growth and make more reliable predictions because the modeling structure is the most appropriate.

In the Figures (3) and (4) show the percentage errors for the estimated values of the population of female and male in ages between 35 and 39 from 2015 up to 2035, obtained using the algorithms MH, SMC, and PMMH, with respect to the World Bank model. The results indicate that the percentage errors for the estimated states are much lower with respect to the values provided from the model of the World Bank model, with the exception of the results obtained using the PMMH algorithm, that show a very high percentage error from the year 2030, which implies a divergence. These estimated values validate the results shown in the Figures (1) and (2) and in the tables (2) and (3), observing good prediction regarding the population growth of female and male up to the year 2035, which implies an increase in the human population, either by immigration or by the increase in the birth rate with respect to the mortality rate. All this can affect natural resources and social infrastructure, affecting the government's commitment to maintain services and infrastructure.

In the Table (4) the posterior estimate of the σ parameter of the population growth model is shown, for the three algorithms proposed up to the year 2035. The results indicate, a close stability in the estimated value of the
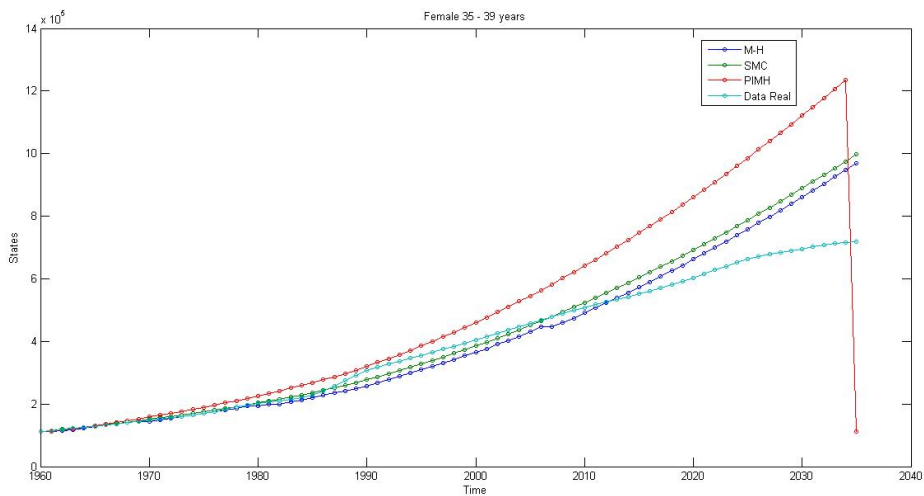
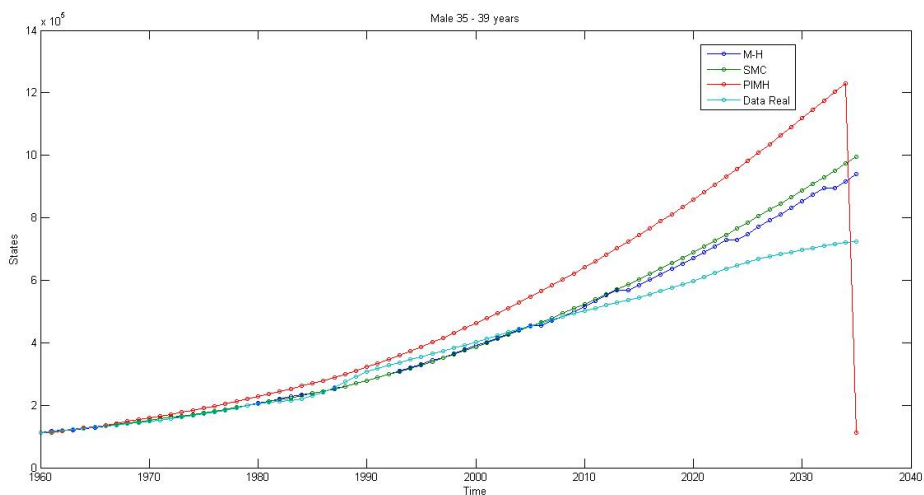Figure 1. Algorithm: MH , SMC and PMMH for Female

Figure 2. Algorithm: MH , SMC and PMMH for Male

Estimated BM	MH	SMC	PMMH	Years
545494	586744	603399	745339	2015
576000	634950	654313	811842	2018
598000	668458	689873	858770	2020
659000	764280	785255	983263	2025
697000	866836	887794	111846000000	2030
724000	972081	995009	126013000000	2035

Table 2. Estimated values of the population of male from 35 to 39 years old of the Republic of Ecuador by the World Bank (BM) and the algorithms MH, SMC and PMMH
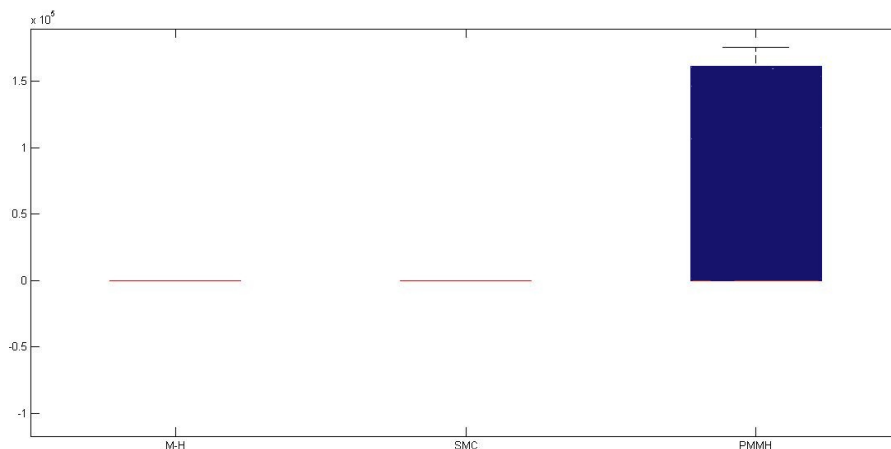
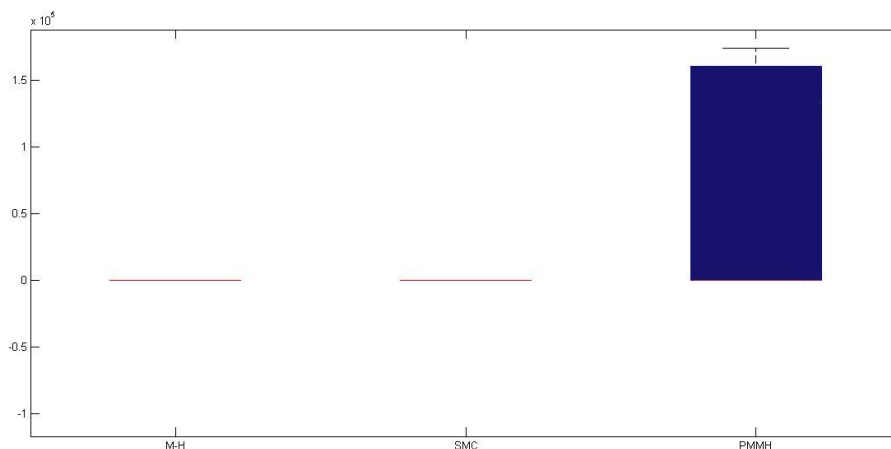Figure 3. Percentage error for the algorithms: MH, SMC and PMMH for Female

Figure 4. Percentage error for the algorithms: MH, SMC and PMMH for Male

Estimated BM	MH	SMC	PMMH	Years
551393	587363	604299	746172	2015
581000	636649	655714	813522	2018
603000	671152	691570	860486	2020
662000	764589	787527	985922	2025
696000	866741	890100	112135000000	2030
719000	949192	996820	126275000000	2035

Table 3. Estimated values of the population of female from 35 to 39 years old of the Republic of Ecuador by the World Bank (WB) and the algorithms MH, SMC and PMMH

parameter σ for each algorithms with the data of the female and male sex, but no common convergence to the value of the parameter with respect to the convergence of each algorithm.

Gender	MH	SMC	PMMH
Male	$\sigma = 28.12$	$\sigma = 1.74$	$\sigma = 2880$
Female	$\sigma = 27.9$	$\sigma = 1.64$	$\sigma = 2552$

Table 4. Parameter estimation σ

4.2. Stochastic model of population growth with variability in the environment

To adjust the stochastic model of population growth that experiences variability in the environment, data were considered that stem from the population growth of the Republic of Ecuador, both female and male. The Table (5) shows the a priori distributions that were considered for the adjustment of the stochastic model.

In the Figures (5) and (6), the estimates of the states (growth of the considered population of the Republic

Parameters	MH	SMC	PMMH
β_1	0.001	0.001	0.001
β_2	0.001	0.001	0.001
α_1	0.001	0.001	0.001
α_2	0.001	0.001	0.001
ϵ_1	0.001	0.001	0.001
ϵ_2	0.001	0.001	0.001
ϵ_3	0.001	0.001	0.001
Δt	0.000003	0.000003	0.000003
b_e	30000	30000	30000
d_e	8000	8000	8000

Table 5. A priori values to initialize the parameters of the MH, SMC and PMMH algorithms, for the female and male data

of Ecuador) are shown by the algorithms MH, SMC, and PMMH, together with the projections made by the World Bank models, and as commented above, these estimates come from demographic models and, therefore, are susceptible to biases and errors due to deficiencies in both the model and the data. In these Figures and in the Tables (6) and (7) it can be seen that the algorithms proposed have a good performance in the prediction of the population up to the year 2035, taking as reference the model of the World Bank.

The Figures (7) and (8) show the percentage errors for the estimated values of the population of female and male between 35 and 39 years of age. The Republic of Ecuador from 2015 up to 2035, estimates obtained with the algorithms MH, SMC and PMMH with respect to the World Bank model. These results show that the percentage errors for the estimated states using the MH, SMC and PMMH algorithms are lower with respect to the values provided by the World Bank model. These estimated values validate the results shown in the Figures (5) and (6) and in the Tables (6) and (7), observing good prediction regarding the population growth of women and men up to the year 2035, which corroborates an increase in the human population, which as mentioned above may be due to immigration or an increase in the birth rate with respect to the mortality rate, which would affect the use of natural resources and the development of social infrastructure and in consequence affects government commitment for maintaining these services and this infrastructure properly.

In the Tables (8) and (9) the posterior estimation of the parameters β_1, β_2, α_1 and α_2 of the stochastic population growth model for the female and male data respectively, using the three proposed algorithms for the year 2035. In these results a stability is observed in the values of the parameters for each algorithm, that is closed to the data, in addition a common convergence to the value of the parameters with respect to the convergence of each algorithm is observed.
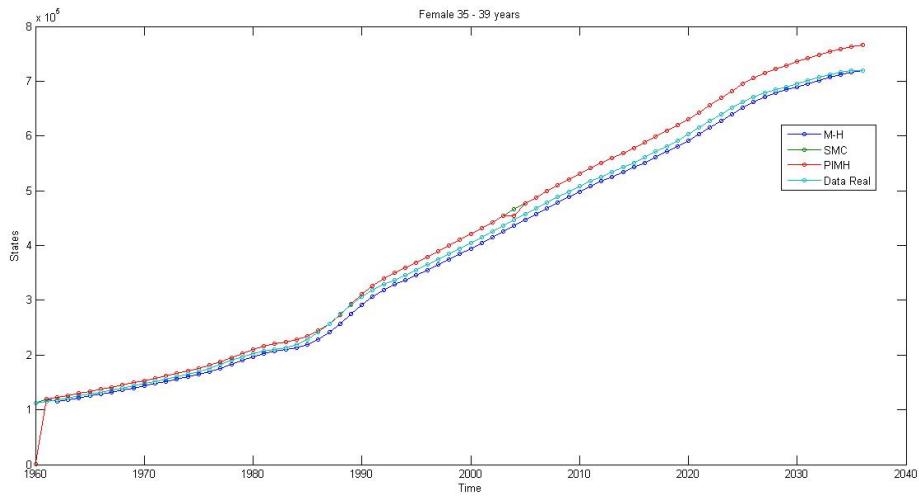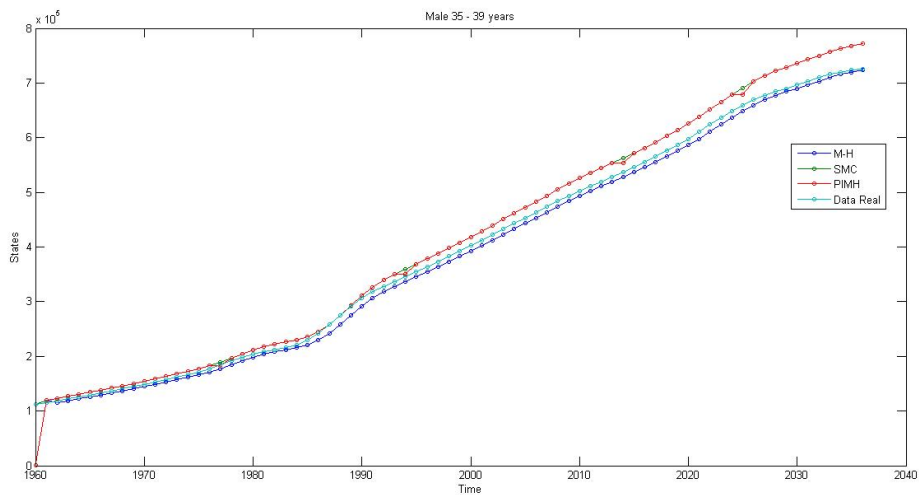
Figure 5. Algorithm: MH , SMC and PMMH for Female

Figure 6. Algorithm: MH , SMC and PMMH for Male

Estimated BM	MH	SMC	PMMH	Years
545494	536611	572027	572027	2015
576000	566000	603356	603356	2018
598000	587000	625742	625742	2020
659000	648000	690768	690768	2025
697000	690000	735540	735540	2030
724000	720000	767520	767521	2035

Table 6. Estimated values of the population of male from 35 to 39 years old of the Republic of Ecuador by the World Bank (BM) and the algorithms MH, SMC and PMMH
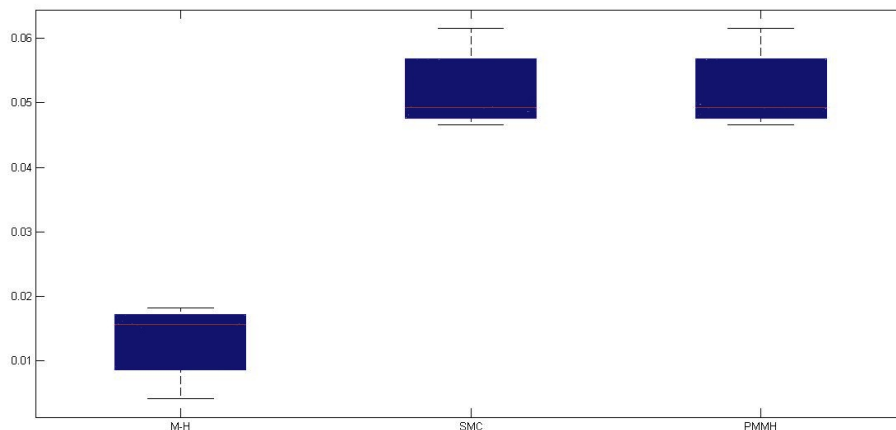
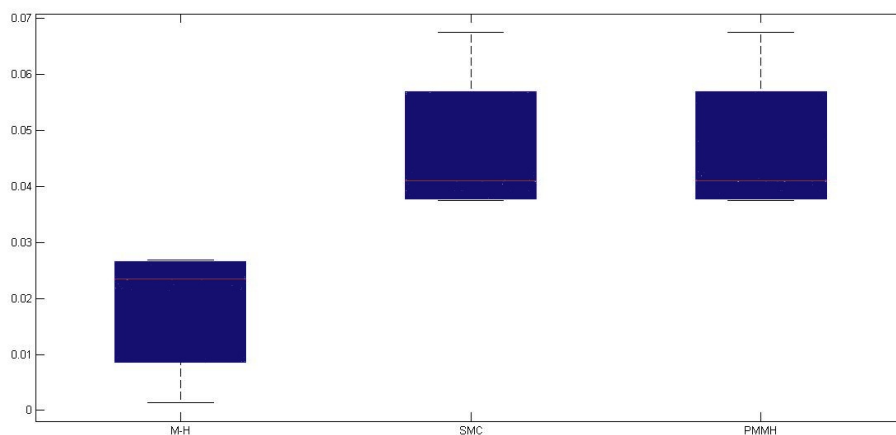Figure 7. Percentage error for the algorithms: MH, SMC and PMMH for Female

Figure 8. Percentage error for the algorithms: MH, SMC and PMMH for Male

Estimated BM	MH	SMC	PMMH	Years
551393	542488	578292	578292	2015
581000	571000	608686	608686	2018
603000	592000	631073	631072	2020
662000	652000	695032	695032	2025
696000	690000	735540	735540	2030
719000	716000	763256	763256	2035

Table 7. Estimated values of the population of female from 35 to 39 years old of the Republic of Ecuador by the World Bank (WB) and the algorithms MH, SMC and PMMH

Parameters	MH	SMC	PMMH
β_1	0.9972	1.023	0.9896
β_2	0.9824	1.002	1.001
α_1	0.4981	0.5041	0.4981
α_2	0.486	0.4924	0.4966

Table 8. Estimated values a posteriori of the parameters by means of the algorithms MH, SMC and PMMH, for the male data

Parameters	MH	SMC	PMMH
β_1	0.9944	0.9965	1.016
β_2	0.9952	1	1.017
α_1	0.5025	0.5131	0.501
α_2	0.4879	0.4925	0.5082

Table 9. Estimated values a posteriori of the parameters using the MH, SMC and PMMH algorithms, for female data

5. Discussions and Conclusions

The stochastic models used to study population dynamics through stochastic differential equations are well known in the field of statistics. However, in the current literature there are not many efficient and reasonable estimation methods from the computational point of view, that allow dealing with these complex non-linear structures that in most cases do not have closed expressions for the approximation of the likelihood. In this paper two stochastic growth models were proposed, a first model determined by the second-order stochastic differential equation of Lagiven, and a second model determined by a system of stochastic differential equations that models the environmental variability of a population. In addition, three computational algorithms were provided: MH, SMC and PMCMC to estimate the solutions states and parameters. By combining the SMC and MCMC methods, the problem of large dimensions is divided into many low-dimensional problems, that simplify calculations without sacrificing the properties of the target distribution. The adjustment of the models was carried out using real data from the Republic of Ecuador, and the results obtained predict a population growth from the year 2018 to the year 2035. These results were compared with those estimated by the deterministic models of the World Bank, indicating that with the models proposed in this work under the Bayesian methodology, the estimated data more appropriately fits the real population growth. Is estimated later the parameters of the models, and it is shown that the proposal adequately and reliably models the dynamics of demographic randomness, sampling errors and environmental randomness better than with the deterministic models used by the World Bank. To measure the relative success of the estimation algorithms, the relative error expressed as a percentage was proposed as a measure of goodness of fit, obtaining small percentage errors for the proposed models. Know a priori the population growth of a country can prevent the social and economic impact; in addition, the results obtained by these models can serve as a reference for decision making of government agencies, that are responsible for planning urban development policies, infrastructure, health services, education, drinking water, and growth harmonized with the environment, among others.

REFERENCES

1. E.J. Allen, *Stochastic Differential Equations and persistence time of two interacting populations*, Dynamics of Continuous, Discrete, and Impulsive Systems, pp. 271–281, 1999.
2. E.J. Allen and H.D. Victory , *Modelling and simulation of a schistosomiasis infection with biological control*, Acta Tropical, vol. 87, pp. 251–261, 2003.
3. E.J. Allen, L.J.S. Allen and H. Schurz, *A comparison of persistence-time estimation for discrete and continuous population models that include demographic and environmental variability*, Mathematical Biosciences, vol. 196, pp. 14–38, 2005.

4. L.J.S. Allen, *An Introduction To Stochastic Processes With Applications to Biology*, Pearson Education Inc., Upper Saddle River, New Jersey, 2003.
5. L.J.S. Allen and E.J. Allen, *A comparison of three different stochastic population models with regard to persistence time*, Theoretical Population Biology, vol. 68, pp. 439–449, 2003.
6. L.J.S. Allen, *Modelling with* Itô *Stochastic Differential Equations*, Mathematical Modelling: Theory and Applications, Springer, 2007.
7. C. Andrieu, A. Doucet and R. Holenstein, *Particle Markov Chain Monte Carlo methods*, Journal of the Royal Statistical Society B, vol. 72, no. 3, pp. 269–342, 2010.
8. S. Ditlevsen and A. Samson , *Introduction to stochastic models in biology*, In Stochastic biomathematical models, Springer, 2013.
9. G. Evensen and P.van Leeuwen, *Assimilation of geosat altimeter data for the agulhas current using the ensemble Kalman filter with a quasi-geostrophic model*, Monthly Weather Review, vol. 24, pp. 85–96, 1996.
10. K.E. Emmertand and L.J.S. Allen, *Population extinction in deterministic and stochastic discrete-time epidemic models with periodic coefficients with applications to amphibians*, Natural Resource Modeling, vol. 19, pp. 117–164, 2006.
11. V. Fernandez and J. Rubio, *Estimating dynamic equilibrium economies: Linear versus nonlinear likelihood*, Journal of Applied Econometrics, vol. 20, pp. 891–910, 2005.
12. G. Gause, *The struggle for existence*, Baltimore, 1934.
13. J.B. Gelman, H.S. Stern, D.B. Dunson, A. Vehtari, and D.R. Rubin, *Bayesian Data Analysis*, Third Edition, Chapman and Hall/CRC, 2013.
14. A. Golightly and D. J. Wilkinson, *Bayesian parameter inference for stochastic biochemical network models using particle Markov chain Monte Carlo*, Interface Focus, vol. 1, pp. 807–820, 2012.
15. A. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, 1989.
16. W. K. Hastings, *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*, Biometrika, vol. 57, pp. 97–109, 1970.
17. S. Heston, *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, The Review of Financial Studies, vol. 6, pp. 327–343, 1993.
18. S. Infante, C. Luna, L. Sánchez and A. Hernández, *Approximations of the solutions of a stochastic differential equation using Dirichlet process mixtures and gaussian mixtures*, Statistics Optimization and Information Computing, vol. 4, pp. 289–307, 2016.
19. S. Infante, L. Sánchez and F. Cedeno, *Filtros para predecir incertidumbre de lluvia y clima*, Revista de Climatología, vol. 12, pp. 33–48, 2012.
20. E. Ionides, *Inference and filtering for partially observed diffusion processes via sequential Monte Carlo*, Working paper available from http://www.stat.lsa.umich.edu/?ionides/pubs/WorkingPaper-filters.pdf, 2003.
21. E. Ionides, C. Bretó and A. King, *Inference for Nonlinear Dynamical Systems*, Proceedings of the National Academy of Sciences of the United States of America, vol. 103, pp. 18438–18443, 2006.
22. N. Kirupaharan and L.J.S. Allen, *Coexistence of multiple pathogen strains in stochastic epidemic models with density-dependent mortality*, Bulletin of Mathematical Biology, vol. 66, pp. 841–864, 2004.
23. V. Kostitzin, *Mathematical biology*, Harrap, 1939.
24. E. Lazkano, B. Sierra, A. Astigarraga and J. Martnez, *On the use of bayesian networks to develop behaviours for mobile robots*, Robotics and Autonomous Systems, vol. 55, pp. 253–265, 2007.
25. J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer Science & Business Media, 2008.
26. F. Lindsten, M. Jordan and T. Schön, *Particle gibbs with ancestor sampling*, The Journal of Machine Learning Research, vol. 15, pp. 2145–2184, 2014.
27. A. Lotka, *Elements of physical biology*, Williams and Wilkins Company, Baltimore, 1925.
28. R. Malthus, *An Essay on the Principle of Population*, Penguin, Harmondsworth, England, 1798.
29. R.K. McCormack and L.J.S. Allen, *Disease emergence in deterministic and stochastic models for host and pathogen*, Applied Mathematics and Computation, vol. 168, pp. 1281–1305, 2005.
30. N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller and E. Teller, *Equations of state calculations by fast computing machines*, J. Chem. Phys., vol. 21, pp. 1087–1091, 1953.
31. B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, Springer, New York, 6th edition, 2003.
32. Y. Pokern, A. Stuart and P. Wiberg, *Parameter estimation for partially observed hypoelliptic diffusions*, Journal of the Royal Statistical Society Serie B., vol. 71, pp. 49–73, 2009.
33. C. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. Springer-Verlag, New York, 2004.
34. C. Robert, *The Metropolis-Hastings algorithm*, Statistics Reference Online, John Wiley & Sons, Ltd., pp. 1–15, 2016.
35. L. Sánchez, S. Infante, V. Grifin and D. Rey, *Spatio-temporal dynamic model and parallelized ensemble kalman filter for precipitation data*, Brazilian Journal of Probability and Statistics, vol. 30, pp. 653–675, 2016.
36. L. Sánchez, S. Infante, J. Marcano and V. Grifin, *Polynomial chaos based on the parallelized ensemble kalman filter to estimate precipitation states*, Statistics Optimization and Information Computing, vol. 3, pp. 79–95, 2015.
37. F. Sigrist, H. Künsch and W. Stahel, *A dynamic nonstationary spatio-temporal model for short term prediction of precipitation*, The Annals of Applied Statistics, vol. 6, pp. 1452–1477, 2012.
38. P. Verhulst, *Notice sur la loi que la population suit dans son accroissement*, Corres. Math. et Physique, vol. 10, pp. 113–121, 1838.
39. P. Verhulst, *Recherches mathématiques sur la loi d'accroissement de la population*, Nouveaux mémoires de l'Academic Royale des Science et Belles-Lettres de Bruxelles, vol. 18, pp. 1–41, 1845.
40. V. Volterra, *Fluctuations in the abundance of a species considered mathematically*, Nature, vol. 118, pp. 558–560, 1926.