



STOCHASTIC DIFFERENTIAL EQUATION: AN APPLICATION TO MORTALITY DATA



Rajani P. Agadi ^{*1}✉, A. S. Talawar ¹

¹ Department of Statistics, Karnatak University, Dharwad, India

DOI: <https://doi.org/10.29121/granthaalayah.v8.i6.2020.538>



Article Type: Research Article

Article Citation: Rajani P. Agadi, and A. S. Talawar. (2020). STOCHASTIC DIFFERENTIAL EQUATION: AN APPLICATION TO MORTALITY DATA. International Journal of Research - GRANTHAALAYAH, 8(6), 229 – 235. <https://doi.org/10.29121/granthaalayah.v8.i6.2020.538>

Received Date: 08 May 2020

Accepted Date: 29 June 2020

Keywords:

Stochastic Differential Equation
Mortality
Estimation
Confidence Interval and Prediction

ABSTRACT

In the present paper we consider an application of stochastic differential equation to model age-specific mortalities. We use New Zealand mortality data for the period 1948–2015 to fit the model. The point predictions of mortality rates at ages 40, 60 and 80 are quite good, almost undistinguishable from the true mortality rates observed.

1. INTRODUCTION

Human *mortality rates* are very important for prediction purposes in social security systems, life insurance, public and private pension plans, etc. The mortality rates of all age and sex groups have a tendency to decline over time, but they clearly show random fluctuations. Many authors have contributed towards the modelling of mortality data through stochastic differential equations (Braumann, 1993; Braumann, 1999a; Braumann, 1999b; Kessler et al., 2012 and Ait-Sahalia, (2002, 2008)). These authors also have contributed towards estimation and other statistical inferential issues on stochastic differential equations. Braumann (1993) and Braumann (1999a) use Black–Scholes model for estimation, testing and prediction, including comparison tests among average return/growth rates of different stocks or populations. Ordinary differential equation (ODE) models have been used to study the growth of individual living beings, like farm animals, trees or fish. When the growth occurs in environments with random variations, stochastic differential equation (SDE) models have been proposed. Braumann (1999b) shows that qualitative results very similar to those obtained for specific models also hold for the general stochastic population growth model.

Lagarto and Braumann (2014) used SDE model (stochastic Malthusian) for mortality data of Portugal and considered precisely the joint evolution of the crude death rates of two age-sex groups. Lagarto (2014) has

considered a study of alternative structures and has determined, among those, which ones have a good performance. These can then be used in predictions and in applications.

2. STOCHASTIC MORTALITY MODEL

The random fluctuations are mostly explained by environmental stochasticity and use a stochastic differential equation (SDE) model (Black–Scholes or stochastic Malthusian model) to represent $D(t)$, the death rate at time t of a given age-sex group. In a random environment, the rationale is that the effect of environmental random fluctuations on the growth rate can be described approximately by a white noise, so that its accumulated effect up to time t can be approximated by a Wiener process $\sigma W(t)$. This gives to $\Delta D(t) = (R\Delta t + \sigma\Delta W(t))D(t)$, with $\Delta W(t) = W(t + \Delta t) - W(t)$, and, letting $\Delta t \rightarrow 0$, one gets the stochastic Malthusian growth model. We denote $Y(t) = D^m(t) - D^f(t)$, where $D^m(t)$ and $D^f(t)$ are the age specific mortality rates for males and females.

Consider a Stochastic differential equation (SDE)

$$\frac{dY_t}{dt} = b(t, Y_t) + \sigma(t, Y_t)W_t, \quad b(t, y) \in R \quad \sigma(t, y) \in R \quad (1)$$

Where W_t is one-dimensional 'white noise'

By Itô (1951) calculus, the Y_t satisfies the following Stochastic integral equation

$$Y_t = Y_0 + \int_0^t b(s, Y_s) ds + \int_0^t \sigma(s, Y_s) dW_s$$

Or in differential form

$$dY_t = b(t, Y_t)dt + \sigma(t, Y_t)dW_t \quad (2)$$

It is the Itô formula that is the key to the solution of many stochastic differential equations.

Various models describe the pattern of human mortality, the one such model that we consider is a stochastic Gompertz model also called Malthusian growth model.

$$dY(t) = RY(t)dt + \sigma Y(t)dW(t), \quad Y(0) = Y_0 > 0 \quad (3)$$

Where R is the arithmetic average growth rate, which is assumed constant over time and $W(t)$ is the white noise. The parameter R was interpreted in the literature as an average growth rate. We have used Itô's calculus to obtain the solution. The solution is the geometric Brownian motion process.

The solution of equation (3) is obtained by Itô calculus (see also Øksendal 2003)

$$Y(t) = y_0 \exp\left\{\left(R - \frac{\sigma^2}{2}\right)t + \sigma W(t)\right\} \quad (4)$$

or

$$Y(t) = y_0 \exp\{rt + \sigma W(t)\} \quad (5)$$

Where $r = R - \frac{1}{2}\sigma^2$ is called the geometric average growth rate.

We conclude that

- 1) If $R > \frac{1}{2}\sigma^2$, then $Y(t) \rightarrow \infty$ as $t \rightarrow \infty$ a.s
- 2) If $R < \frac{1}{2}\sigma^2$, then $Y(t) \rightarrow 0$ as $t \rightarrow \infty$ a.s
- 3) If $R = \frac{1}{2}\sigma^2$, then $Y(t)$ will fluctuate between arbitrary large and small values as $t \rightarrow \infty$ a.s

To fit model (4), we made it age-specific by writing

$$Y_x(t) = y_{x0} \exp \left\{ \left(R_x - \frac{\sigma_x^2}{2} \right) t + \sigma_x W(t) \right\} \quad (6)$$

Where $R_x = r_x + \frac{\sigma_x^2}{2}$ and hence, $r_x = R_x - \frac{\sigma_x^2}{2}$, which called as geometric average growth rate.

We apply the methodology used in (Lagarto and Braumann, 2014; Braumann, 2019; Talawar and Agadi, 2020) for parameters estimation. The value of r_x of the rate of decay of $Y_x(t)$ would be negative since the tendency is for $Y_x(t)$ to decrease exponentially. The parameter σ_x measures the intensity of the effect of environmental fluctuations (weather conditions, epidemic diseases, social conditions, etc.) on the rate of change. In this case of mortality rates are time equidistant observations with $\Delta = 1$ year and $t_k = k\Delta$ years ($k = 0, 1, 2 \dots n = 60$), that is, the mortality rates are age-specific with single year. Denoting by $Y_{x,k} = Y_x(t_k)$ and by $L_{x,k} = \ln(Y_{x,k}/Y_{x,(k-1)})$, the log-returns, one gets for model (6) the maximum likelihood estimates of the parameters:

$$\hat{r}_x = \frac{1}{n\Delta} \sum_{k=1}^n L_{x,k} \text{ /year and } \hat{\sigma}_x^2 = \frac{1}{n\Delta} \sum_{k=1}^n (L_{x,k} - \hat{r}_x \Delta)^2 \text{ /year}$$

Therefore the 95% confidence intervals of the parameters r_x and σ_x^2 are

$$\hat{r}_x \pm 1.96 \sqrt{\hat{\sigma}_x^2 / n\Delta} \text{ and } \hat{\sigma}_x^2 \pm \sqrt{2 \hat{\sigma}_x^4 / n}$$

3. APPLICATIONS OF THE MODEL

For the application of model we considered the mortality data of New Zealand from 1948-2015. In **Figure 1** it shows that age-specific death rates of 60-year-old females and males of the New Zealand population for each year of the period 1948–2015 (Source: <https://www.mortality.org/>). For convenience, we start counting time in 1948, so initial time, $t = 0$ corresponds to 1948.

Figures 2-4, show the results of using model (6) for the age-specific mortality rates of 60 year-old New Zealand males and females. Notice the decline from $Y_{60}(0) = 0.0138/\text{year}$ in 1948 to $Y_{60}(60) = 0.00489/\text{year}$ in 2008 and $Y_{60}(75) = 0.00377/\text{year}$ in 2023. For parameter estimation, we only use the data of the period 1948–2008, reserving the period 2009–2023 for prediction purposes.

Therefore the estimates of the parameters are

$$\hat{r}_{60} = \frac{1}{n\Delta} \sum_{k=1}^n L_{60,k} = \frac{1}{n\Delta} \ln \frac{Y_{60,n}}{Y_{60}(0)} \text{ /year}$$

$$\hat{\sigma}_{60}^2 = \frac{1}{n\Delta} \sum_{k=1}^n (L_{60,k} - \hat{r}_{60} \Delta)^2 \text{ /year}$$

The 95% confidence intervals of the parameters are

$$\hat{r}_{60} \pm 1.96 \sqrt{\hat{\sigma}_{60}^2 / n\Delta} = (-1.0373 \pm 0.3371/\text{year}) \text{ and}$$

$$\hat{\sigma}_{60}^2 \pm \sqrt{2 \hat{\sigma}_{60}^4 / n} = (0.0291 \pm 0.003756)/\text{year}.$$

The parameter value \hat{r}_{60} obtained is used in (6) to get the plots and putting $\sigma_x = 0$, as if we have a deterministic environment, we obtain an adjusted curve (the thin solid line in **Figures 1-7**) that gives an idea of the trend. If one has the data up to 2008 (corresponding to $n = 60$ and $t_n = 2008 - 1948 = 60$ years) and wants to make predictions for the future period 2009–2023, one uses model (6) with the estimated parameter values (based only on data up to 2009), starting with the observed mortality rate of 2009, $Y_{60}(t_n) = Y_{60}(60) = 0.00489/\text{year}$. It is easy to work with $Z_{60}(t) = \ln Y_{60}(t)$, since it is Gaussian. For $\tau > 0$ years, using the same technique as in Braumann (2019), one gets the point predictor

$$\hat{Z}_{60}(t_n + \tau) = \ln 0.00489 + \hat{r}_{60} \tau$$

$$\hat{Z}_{60}(t_n + \tau) = \exp(\hat{Z}_{60}(t_n + \tau))$$

and an approximate 95% confidence prediction interval for $\hat{Z}(t_n + \tau)$

$$\hat{Z}_{60}(t_n + \tau) \pm 1.96\sqrt{\hat{\sigma}^2\tau(1 + \tau/t_n)}$$

From which extremes, by taking exponentials, one gets an approximate confidence prediction interval for $\hat{Y}_{60}(t_n + \tau)$. The similar procedure can be used to estimate r_x and σ_x^2 at each age x . The three lines in **Figure 1-7** show the extremes of such a prediction interval (outer lines) and the point prediction (middle line). As we can see, the point predictions quite good, almost undistinguishable from the true mortality rates observed in 2009–2023.

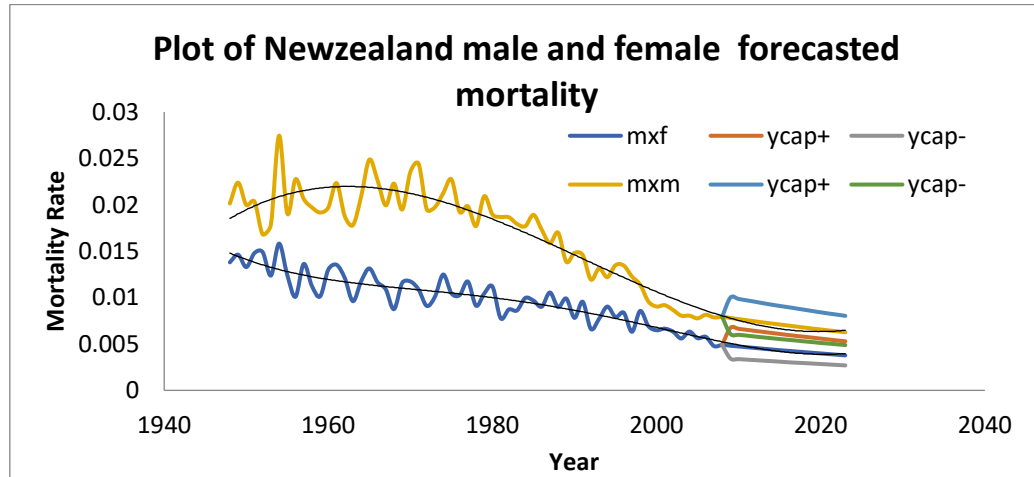


Figure 1: Forecasted mortality rate of New Zealand male and female

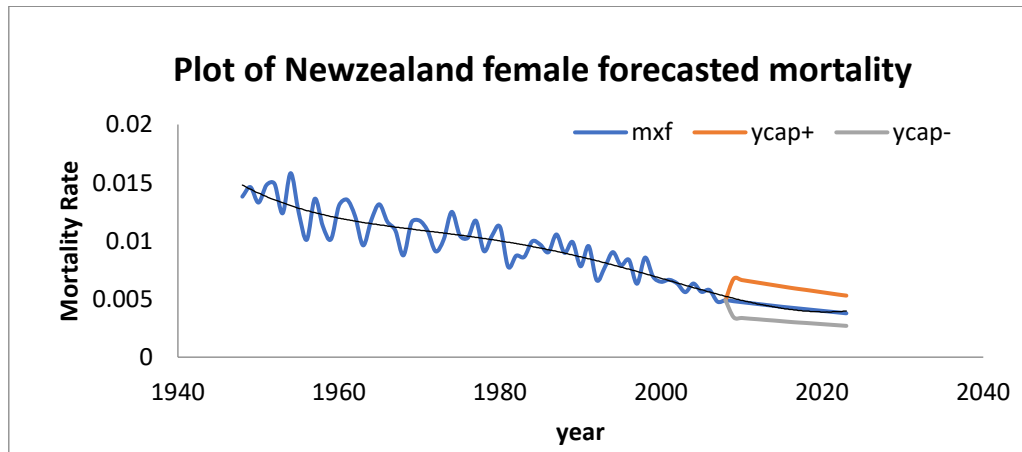


Figure 2: Forecasted mortality rate of New Zealand female

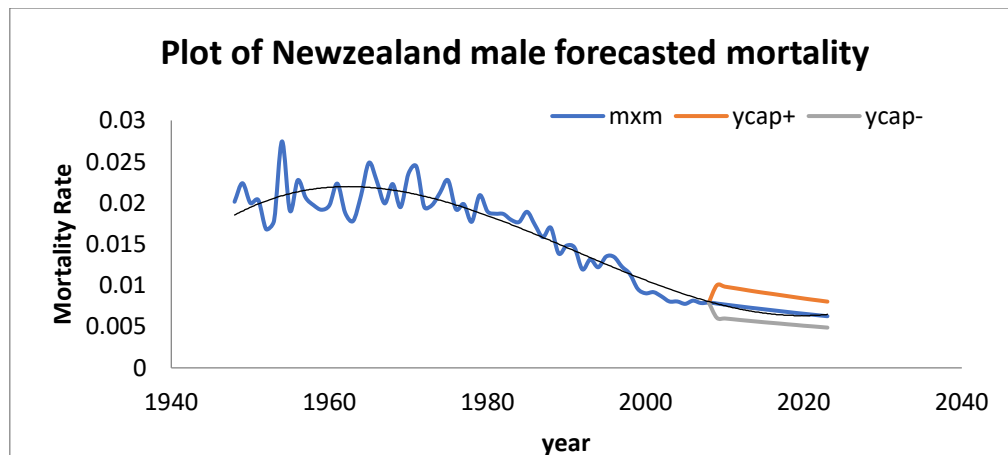


Figure 3: Forecasted mortality rate of New Zealand male

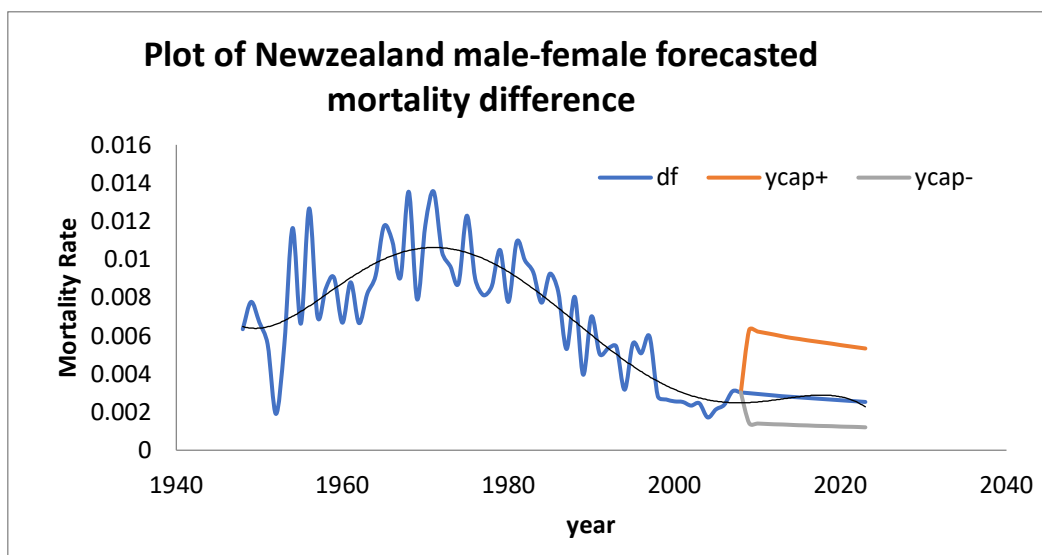


Figure 4: Forecasted mortality rate of New Zealand male-female difference

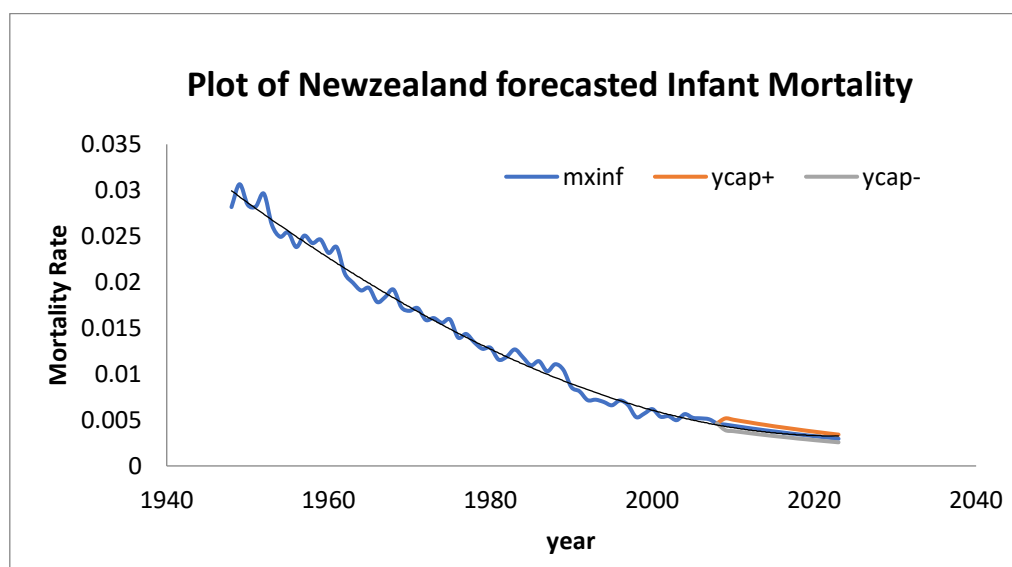


Figure 5: Forecasted mortality rate of New Zealand Infant

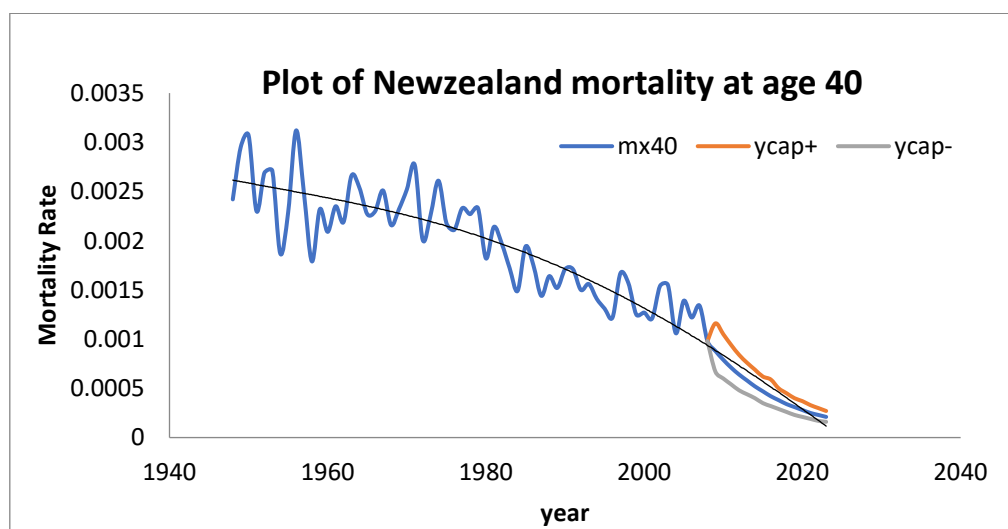


Figure 6: Forecasted mortality rate of New Zealand at age 40

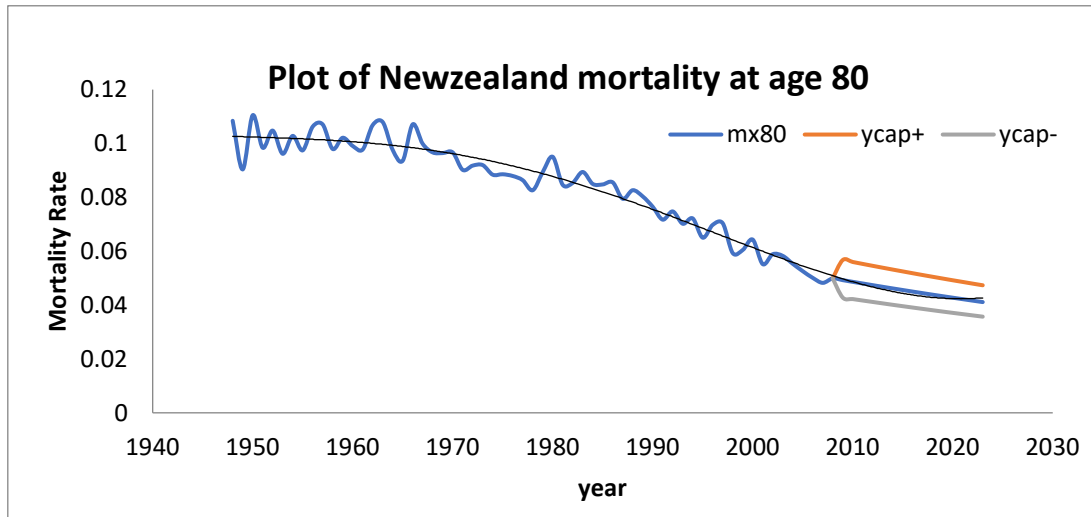


Figure 7: Forecasted mortality rate of New Zealand at age 80

Table 1: The parameter values for males and females of New Zealand mortality data.

	$Y(0)/\text{year}$	$Y(t_n)/\text{year}$	$Y(75)/\text{year}$	\hat{r}/year	$\hat{\sigma}^2/\text{year}$
NZ_{60}^f	0.0138	0.00489	0.00377	-0.0173	0.0291
NZ_{60}^m	0.02014	0.00792	0.00626	-0.0156	0.01586
NZ_{lnf}^T	0.02817	0.00463	0.00296	-0.0301	0.0051
NZ_{40}^T	0.00242	0.00098	0.00021	-0.1051	0.0198
NZ_{80}^T	0.10842	0.0499	0.04111	-0.0129	0.0051
NZ_{60}^D	0.00634	0.00303	0.00253	-0.01231	0.1421

4. CONCLUSIONS

This stochastic mortality model gives good prediction intervals and the point prediction for each age. The point predictions of mortality rates at ages 40, 60 and 80 are quite good, almost undistinguishable from the true mortality rates observed. The similar procedure can be used to estimate parameters at each age x . Once all the age-specific mortalities are obtained, the different columns of a life table can be constructed using their interrelationships.

SOURCES OF FUNDING

None.

CONFLICT OF INTEREST

None.

ACKNOWLEDGMENT

None.

REFERENCES

- [1] Aït-Sahalia, Y. (2002). Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes: Comment. Journal of Business and Economic Statistics. 20(3):317-21 DOI: 10.1198/073500102288618405

- [2] Aït-Sahalia, Y. (2008). Closed-form likelihood expansions for multivariate diffusions. *Ann. Statist.* 36 (2008), no. 2, 906--937. doi:10.1214/009053607000000622. <https://projecteuclid.org/euclid.aos/1205420523>
- [3] 1205420523
- [4] Braumann, C.A. (1993) Model fitting and prediction in stochastic population growth models in random environments. *Bulletin of the International Statistical Institute*, LV (CP1), 163–164. Braumann, C.A. (1999a) Comparison of geometric Brownian motions and applications to population growth and finance. *Bulletin of the International Statistical Institute*, LVIII (CP1), 125–126.
- [5] Braumann, C. A., *Introduction to Stochastic Differential Equations with Applications to Modelling in Biology and Finance*, John Wiley & Sons Ltd, 2019.
- [6] Itô, K. (1951) On Stochastic Differential Equations, *American Mathematical Society Memoirs*, No. 4.
- [7] Kessler, M., Lindner A. and, Sorensen, M. *Statistical Methods for Stochastic Differential Equations*, Chapman and Hall/CRC, 2012.
- [8] Lagarto, S. and Braumann, C.A. Modeling human population death rates: A bi-dimensional stochastic Gompertz model with correlated Wiener processes, in *New Advances in Statistical Modeling and Applications* (eds, A. Pacheco, R. Santos, M.R. Oliveira, and C.D. Paulino), Springer, Heidelberg, pp. 95–103, doi:10.1007/978-3-319-05323-3_9, 2014.
- [9] Øksendal, B. *Stochastic Differential Equations. An Introduction with Applications*, 6th edition, Springer, New York, 2003.
- [10] Talawar, A. S. and Agadi, R.P. (2020). Novel corona virus pandemic disease (covid-19): Some applications to south asian countries. *International Journal of Research and Analytical Reviews*, 7(2), 905-912.