10-2019

# Forecasting Subnational Demographic Data using Seasonal Time Series Methods

Jorge M. Bravo

Edviges Coelho

# Forecasting Subnational Demographic Data using Seasonal Time Series Methods

Jorge M. Bravo, Universidade Nova de Lisboa NOVA IMS & MagIC & CEFAGE-UE, Portugal, jbravo@novaims.unl.pt

Edviges Coelho, Statistics Portugal & Universidade Lusófona (ECEO-UHLT), Portugal, edviges.coelho@ine.pt

## Abstract

Forecasts of monthly demographic data are a critical input in the computation of infra-annual estimates of resident population since they determine, together with international net migration, the dynamics of both the population size and its age distribution. The empirical time series of demographic data exhibits strong evidence of the presence of seasonality patterns at both national and subnational levels. In this paper, we evaluate the short-term forecasting performance of alternative linear and non-linear time series methods (seasonal ARIMA, Holt-Winters and State Space models) to birth and death monthly forecasting at the sub-national level. Additionally, we investigate how well the models perform in terms of predicting the uncertainty of future monthly birth and death counts. We use the series of monthly birth and death data from 2000 to 2018 disaggregated by sex for the 25 Portuguese NUTS3 regions to compare the model's short-term (one-year) forecasting accuracy using a backtesting time series cross-validation approach.

**Keywords:** Time series methods; seasonality; population forecasts; ARIMA; Backtesting.

## 1 INTRODUCTION

Population forecasts are widely used for analytical, planning and policy purposes (e.g., education, health, housing, pensions, security, spatial planning, transportation, public infrastructure and social policy planning) at national, regional and local levels (Smith, Tayman, & Swanson, 2001; Bravo, 2016, Bravo et al., 2018; Ayuso, Bravo & Holzmann, 2019). Concerns about the possible long-term effects of ageing or about the likely impact on population structure of significant internal and international migration flows have been increasingly attracting more attention to the accuracy of population projections. Forecasts of monthly births and deaths are a critical input in the computation of monthly estimates of resident population (MERP) since together with international net migration, they determine, the dynamics of both the population size and its age distribution. Statistical Offices and researchers typically produce MERP using the cohort-component method, a standard demographic tool that requires credible assessments about the future behaviour of age-specific fertility rates, sex and age-specific mortality rates and international and sub-national migrations, together with detailed information about a base year population. To perform this exercise, for each subpopulation and gender it is necessary to (Smith et al., 2001; Bravo, 2007; Bravo et al., 2010): (i) obtain monthly forecasts of the total number of births and deaths, (ii) estimate age-specific mortality rates considering period/cohort life tables derived from stochastic mortality models, eventually considering for heterogeneity in longevity (Ayuso, Bravo & Holzmann, 2017a,b), (iii) estimate the level and age pattern of net international migration, and (iv) consider a number of assumptions such as the distribution of age-specific fertility rates or the sex ratio a birth.

Birth and death forecasts can be produced using, among others, statistical time series methods (univariate or multivariate), structural models (e.g., vector autoregressive models) or machine learning methods (e.g., Artificial Neural Network (ANN), Support Vector Machines (SVM)). To generate reliable estimates, these methods must be consistent with the annual and intra-annual observed patterns in birth and mortality data, offer forecast accuracy and provide measures for the uncertainty in population forecasts. Empirical time series data for births and deaths exhibits strong evidence of the presence of seasonality patterns at both national and subnational (NUTS 2, 3) levels. These time series are typically non-stationary time series and contain trend and seasonal variations. For vital events computed for small populations on monthly time intervals, the need to uncover complex structures of temporal interdependence in time series data is critically challenged in the presence of seasonal variability.

In recent decades a substantial amount of research has focused on the development and application of time series models in population forecasts, focusing either on total population growth or on individual components of growth (see, e.g., Saboia 1974; Lee 1974, 1992; Alho and Spencer 1985; Ahlburg 1992; Pflaumer 1992, Lee and Tuljapurkar 1994; McNown and Rogers 1989; Keilman, Pham & Hetland, 2002; Tayman, Smith, and Lin 2007; Alho, Bravo and Palmer, 2012; Abel et al. 2013; Bravo and Freitas, 2018). The main focus of these studies is largely on the identification and measurement of uncertainty in population forecasts, with little interest in the assessment of the models forecasting accuracy or the out-of-sample validity of the prediction intervals. Much of the research concerning the evaluation of time series models for birth and death forecasting has been focused on univariate time series ARIMA models at the national level, with little research on the predictive accuracy of these models at the sub-national level, particularly in small population areas (see, e.g., Land and Cantor, 1983). Fewer still have explored the use of the Holt-Winters exponential smoothing and State Space time series models in small population exercises. Additionally, despite the increasing interest in short-term trends and variability in mortality and fertility patterns, accessing up-to-date statistics is sometimes difficult since detailed information on birth and deaths counts are made available to researchers with a relative time lag. Also, researchers often need information on the present and near future, when data on birth and deaths counts could only be predicted.

In this paper, we address this gap and investigate and compare the predictive power of alternative linear and non-linear time series methods (seasonal ARIMA, Holt-Winters and State Space models) to birth and death monthly forecasting at the sub-national level using up-to-date demographic data. Using a series of monthly birth and death data from 2000 to 2018 disaggregated by sex for the 25 Portuguese NUTS3 regions, we compare the short-term (one year) forecasting accuracy of Seasonal ARIMA, Seasonal Holt-Winters and Seasonal State Space time series models. We adopt a backtesting time series cross-validation approach, i.e., we consider a multi-step forecasting approach with re-estimation in which the training data or base period (the interval between the month of the earliest and the latest demographic data used to make a forecast) is extended before re-selecting and re-estimating the model at each iteration and computing forecasts.

The main contributions of this paper are the following. First, we summarise and analyse the out-of-sample error performance of commonly used Seasonal ARIMA forecasting models together with alternative methods (Seasonal Holt-Winters and Seasonal State Space models), using a rich and large set of subpopulations and two different demographic events with different dynamics over time. Second, we evaluate the out-of-sample performance of the prediction intervals produced by these models. Third, we assess the consistency of the predictive performance of these methods in populations of different size and nature. Fourth, we evaluate the existence of significant differences in the model's forecasting accuracy between subpopulations of different sex. Fifth, we investigate how well the models perform in terms of predicting the uncertainty of future monthly birth and death counts. To evaluate forecast accuracy, we compare the resulting forecasts with observed data and measure forecast errors using different performance criteria (e.g., RMSE, MAPE, MAD). To assess forecast uncertainty, we compute the proportion of times observed values fall outside 95% confidence intervals computed for the mean. The selection of the appropriate forecasting method depends on several factors, including the past behaviour pattern of the time series, previous knowledge about the nature of the phenomenon being studied, the availability of statistical data and the predictive capacity of the model. Our results show that these simulations provide valuable insights regarding the forecasting performance of alternative time series models in small population forecasting exercises and on the validity of using such models as predictors of population forecast uncertainty and, thus, have significant practical implications. The remaining part of the paper is organised as follows. Section 2 describes the seasonal time series methods used in this paper. Section 3 details the research methods used to produce forecasts and assess model performance and the data features. Section 4 presents and discusses the results. Section 5 concludes this research.

## 2 MODELLING TREND AND SEASONAL TIME SERIES

Modelling the trend and seasonal components of demographic time series is a challenging endeavour. Following earlier work on decomposing a seasonal time series, Holt (1957) extended simple exponential smoothing methods to linear exponential smoothing to allow forecasting of data with time trends. The method was later extended by Winters (1960) to capture seasonality. Box and Jenkins (1970, 1976) developed a coherent and flexible three-stage iterative cycle for time series identification, estimation, and verification (commonly known as the Box-Jenkins approach) and popularised the use of autoregressive integrated moving average (ARIMA) models and its extensions (including some to handle seasonality in time series) in many areas of science. Ord et al. (1997), Hyndman et al. (2002) developed a class of state space models which incorporate some of the exponential smoothing methods. The ability of these methods to model complex structures of temporal interdependence observed in the data has been tested, but their capability for modelling demographic seasonal time series has not yet been fully and systematically investigated. In this section, we briefly review the forecasting methods used in this study for forecasting demographic time series showing seasonality.

### *2.1. Seasonal ARIMA Model*

The seasonal ARIMA model is an extension to the classical ARIMA model that supports the direct modelling of both the trend and seasonal components of a time series and it is widely used for forecasting. The model includes new parameters to specify the autoregression (AR), differencing (I) and moving average (MA) for the seasonal component of the series, as well as an additional parameter for the period of the seasonality (Hyndman and Athanasopoulos, 2013). The model's mathematical and statistical properties allow us to derive not only point forecasts but also probabilistic confidence intervals (Box and Jenkins 1976).

In this paper, we combine the seasonal and non-seasonal components into a multiplicative seasonal autoregressive moving average model, or SARIMA model, given by

$$\Phi_P(B^s)\phi(B)\nabla_s^D\nabla^d x_t = \delta + \Theta_Q(B^s)\theta(B)w_t \tag{1}$$

where $w_t$ denotes the Gaussian white noise process. The general model can be expressed as $ARIMA(p,d,q)\times(P,D,Q)_s$, where the ordinary autoregressive (AR) and moving average (MA) components are represented by polynomials $\phi(B)$ and $\theta(B)$ of orders $p$ and $q$, respectively, the seasonal AR and MA components are denoted by $\Phi_P(B^s)$ and $\Theta_Q(B^s)$ of orders $P$ and $Q$, respectively. The non-seasonal and seasonal difference components are represented by $\nabla^d = (1-B)^d$ and $\nabla^D = (1-B^s)^D$, respectively. The seasonal period $s$ defines the number of observations that make up a seasonal cycle (e.g., $s = 12$ for monthly observations).

The estimation process for the parameters in (1) for each of the 100 time series follows the standard Box-Jenkins (1976) methodology in an iterative 3-step procedure comprising the identification, estimation and evaluation and diagnostic analysis stages. Configuring the SARIMA model requires selecting the hyperparameters for both the trend and seasonal elements of the series. First, we analyse the stationary of the series and check whether or not a seasonal and/or non-seasonal difference is needed to produce a roughly stationary series. For this purpose, we analyse the patterns of the autocorrelation and partial autocorrelation function and conduct unit root differencing tests (Kwiatkowski–Phillips–Schmidt–Shin, 1992; Canova-Hansen, 1995) to determine the optimal order of differencing, $d$, and of seasonal differencing, $D$. We then identify the optimal $p, q, P$ and $Q$ hyper-parameters by fitting models within pre-specified maximum ranges and find the best model by optimizing a stepwise algorithm for the Akaike Information Criterion (AIC). Given the extensive number of experiments conducted in this paper (500 for each of the models tested), we limited the maximum value of $(p, q, P, Q)$ to 5. Each series was tested for the white noise with Bartlett's version of the Kolmogorov-Smirnov test. When the data suggest the inexistence of seasonal unit roots in the series and the seasonality is deterministic, we can express it as a function of seasonal dummy variables (and time eventually). In this case, an ARIMA model if fitted to the residuals of the equation:

$$Y_t = \alpha + \sum_{i=1}^{s-1} \gamma_{i,t} D_{i,t} + \beta t + \epsilon_t \tag{2}$$

where $Y_t$ is the variable of interest, $D_{i,t}$ are seasonal dummies, $t$ denotes time and $\epsilon_t$ is a white-noise error term.

Additionally, we examined the residuals of the selected model and formally examined the null hypothesis of independence of the residuals using the Box-Pierce/Ljung-Box test (also known as "portmanteau" tests). We also tested the normality of the residuals using the Jarque-Bera Test. After examining different models, the best SARIMA model was selected, parameters were estimated using the nonlinear least squares method, and the model was used for forecasting monthly births and deaths.

### *2.2. Holt-Winters' seasonal method*

The Holt-Winters method is a univariate automatic forecasting method that uses simple exponential smoothing (Holt 1957; Winters 1960). The forecast is obtained as a weighted average of past observed values in which the weight function declines exponentially with time, i.e., recent observations contribute more to the forecast than earlier observations. Forecasted values are dependent on the level, slope and seasonal components of the series being forecast. The Holt-Winters method is based on three smoothing equations - one for the level, one for the trend and one for the seasonality.

The model-specific formulation depends on whether seasonality is modelled in an additive or multiplicative way. The additive method is selected when the seasonal variations are approximately constant through the series, whereas the multiplicative method is preferred when the seasonal variations change proportionally to the level of the series (Hyndman and Athanasopoulos, 2013). The **additive method** is specified as:

$$\begin{aligned} l_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} - b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\ s_t &= \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m} \\ y_{t+h|t} &= l_t + \ hb_t + s_{t-m+h} \end{aligned} \qquad (3)$$

where $l_t$, $b_t$ and $s_t$ denote the level, trend and seasonal components, respectively, with corresponding smoothing parameters $\alpha$, $\beta$ and $\gamma$; $y_{t+h|t}$ is the forecast for $h$ periods ahead at time $t$. The Holt-Winters' **multiplicative method** is defined as:

$$\begin{aligned} l_t &= \alpha\frac{y_t}{s_{t-m}} + (1-\alpha)(l_{t-1} - b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\ s_t &= \gamma\frac{y_t}{(l_{t-1} + b_{t-1})} + (1-\gamma)s_{t-m} \\ y_{t+h|t} &= (l_t + \ hb_t)s_{t-m+h} \end{aligned} \qquad (4)$$

We initialize the model's hyperparameters using the decomposition approach suggested by Hyndman et al. (2008) and implemented in the `forecast` package in R. The procedure involves first computing a moving average trend to the first 2 years of data, then subtracting (for additive HW) or dividing (for multiplicative HW) the smooth trend from the original data to get de-trended data. The initial seasonal values (e.g., December) are then obtained from the averaged de-trended data (Decembers). Next, the procedure involves subtracting (for additive HW) or dividing (for multiplicative HW) the seasonal values from the original data to get seasonally adjusted data. Finally, by fitting a linear trend to the seasonally adjusted data we get the initial values for the level and slope. After examining each time series for both the additive and multiplicative versions of the Holt-Winters' seasonal method, we finally selected the model showing lower residual sum of squares to produce forecasts of monthly births and deaths.

### 2.3. Exponential smoothing state space model

We investigated the use of State Space models underlying exponential smoothing methods in monthly births and deaths forecasting. State Space models consist of a measurement equation that describes the observed data, and some state equations that describe how the unobserved components or states (level, trend, seasonal) change over time (Hyndman and Athanasopoulos, 2013). We examined both the additive and multiplicative error versions of the model and automatically selected the best model using the procedure included in R forecast package.

The general Gaussian state space model involves a measurement equation relating the observed data to an unobserved state vector $x_t = (b_t, s_t, s_{t-1}, \dots, s_{t-(m-1)})$, an initial state distribution and a Markovian transition equation that describes the evolution of the state vector over time state. In this paper, we use State Space models that underlie the exponential smoothing methods of the form (Hyndman et al., 2002):

$$Y_t = \mu_t + k(x_{t-1})\varepsilon_t \quad (5)$$

$$x_t = f(x_{t-1}) + g(x_{t-1})\varepsilon_t \quad (6)$$

where $\varepsilon_t \sim N(0, \sigma^2)$, $\mu_t = Y_{t-1}$ and where, for additive error models $k(x_{t-1}) = 1$, such that $Y_t = \mu_t + \varepsilon_t$, whereas for multiplicative error models $k(x_{t-1}) = \mu_t$ such that $Y_t = \mu_t(1 + \varepsilon_t)$. Model estimation involves measuring the unobservable state (prediction, filtering and smoothing) and estimating the unknown parameters using MLE methods. We initialize the model's hyperparameters using the decomposition approach suggested by Hyndman et al. (2008) and implemented in the `forecast` package in R.

## 3 RESEARCH METHODOLOGY

The objective of this research is to empirically compare the forecasting performance of alternative trend and seasonal time series models over short-term horizons. To this end, we set out a backtesting framework and use monthly demographic data for the period 2000-2018. In this section, we briefly describe the research methodology used in this study.

### *3.1. Research Design*

In this paper, we set out a backtesting framework applicable to single-period ahead forecasts from time series methods and use it to evaluate the forecasting performance of three different univariate models applied to subnational (NUTS3) male and female monthly births and deaths data. The backtesting framework used in this paper involves the following steps:[1]

1. We begin by selecting the metric of interest, i.e., the forecasted variable that is the focus of the backtest (monthly births or deaths by sex and subpopulation);

2. We define and select the historical "lookback window" to be used to estimate the parameters of each time series model for any given year. We adopt a time series cross-validation approach, i.e., we consider a multi-step forecasting approach with re-estimation in which the training data or base period (the interval between the month of the earliest and the latest demographic data used to make a forecast) is extended before re-selecting and re-estimating the model at each iteration and computing forecasts. For instance, if we wish to estimate the parameters for year $t$ we estimate the parameters using observations from years $t_0$ to $t-1$, if we wish to estimate the parameters for year $t+1$ we estimate the parameters using observations from years $t_0$ to $t$, i.e., we adopt a expanding lookback window approach. The selection of the lookback window depends on several factors, including the past behaviour pattern of the time series, previous knowledge about the nature of the phenomenon being studied and the availability of statistical data.

3. We then select the forecasting horizon ("lookforward window") over which we will make our forecasts, based on the estimated parameters of the model. In the present study, we focus on relatively short-term horizon forecasts since our interest is on generating 1-year ahead of monthly births and deaths forecasts (12 observations) as an input for computing monthly estimates of resident population and a key input in producing the Labour Force Survey (LFS) in Portugal. The LFS is a quarterly sample survey of households living at private addresses in Portuguese territory, with the main objective of characterising the population in terms of the labour market. It is conducted by Statistics Portugal, in accordance with requirements under EU regulation, and makes quarterly and annual data available. Published data are calibrated by using resident population estimates by NUTS 3 regions, sex and five-year age-breakdown. The LFS quarterly results are published around forty days after the end of the survey period. This calendar is incompatible with the current production of resident population estimates since data on the three components – births, deaths and migration – are not yet available. To comply with the LFS calendar, Statistics Portugal produces advanced monthly estimates of resident population, i.e., at the beginning of each year $t$, monthly estimated values of resident population are computed for year $t$ by NUTS 3 regions, sex and age. As such, monthly forecasts of live births, deaths and migration must be used to produce advanced monthly estimates of resident population.

4. We select a rolling fixed-length horizon backtesting approach in which we consider the accuracy of forecasts over fixed-length horizons as the jump-off date moves sequentially forward through time. This

[1] For a similar approach used in evaluating the forecasting performance of stochastic mortality models and interest rate and credit risk models see, e.g., Dowd et al. (2010), Bravo & Silva (2006) and Chamboko & Bravo (2016, 2019a,b).

procedure involves comparing the births, and deaths mean forecast and prediction intervals for some fixed-length horizon (1-year) rolling forward over time with the corresponding observed outcomes.

5. Finally, we select the evaluation criteria which will be used to compare the forecasting performance of the different models. We computed several evaluation criteria but, given the large number of experiments conducted in this work, we opted to report a single error metric, the Mean Absolute Percent Error (MAPE). For a given lookback and lookforward window, the MAPE for model $j$ is defined as

$$MAPE_j = \frac{1}{n}\sum_{t=1}^{n}\frac{|\hat{y}_{t,j} - y_t|}{y_t} \times 100 \tag{7}$$

where $n$ is the number of forecasted values, $\hat{y}_t$ is the number of monthly births/deaths predicted by the model for time point $t$, and $y_t$ is the corresponding value observed at time point $t$.

Each of the different time series models constructed (using a different lookback window and jump-off year) implies a different set of prediction intervals for the forecast horizon. To better understand the performance of the models analysed in terms of predicting the uncertainty of future births and deaths we computed the number of birth and death counts falling outside the 95% prediction intervals associated with each set of forecasts. Parameter estimation and model forecasting assessment were carried out using a computer routine written in R-script (R Development Core Team 2019).

### *3.2. Data*

In this paper, we use demographic data for Portugal comprising monthly data on live births and deaths broken down by sex and 25 different NUTS 3 regions from January 2000 to December 2018 provided by Statistics Portugal. The demographic dataset consists of 228 monthly observations for each one of the 100 different subpopulations of different size, the smallest with 38,753 resident individuals in December 2017 (Beira Baixa, male), the largest with 1,505,435 individuals (Lisbon Metropolitan Area, female). Of the 100 subpopulations tested, four (Lisbon and Oporto metropolitan areas male and female populations) correspond to highly populated areas with, in the case of Lisbon, more than one million residents. In contrast, the dataset tested includes several small population areas with less than 50,000 residents (e.g., Beira Baixa, Alto Tâmega, Alentejo Litoral). This archive is a challenging dataset in which to assess the monthly forecasting performance of time series methods since the data exhibits significant trend and seasonal components and high volatility in some cases, particularly in small population areas. Figures 1 and 2 represent the time series plot of monthly births and deaths of two representative (small and large) NUTS3 subpopulations.

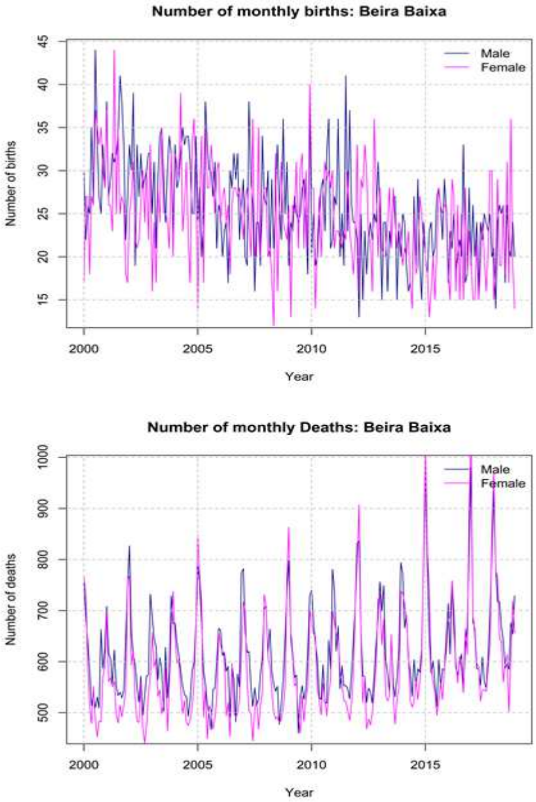


Figure 1 – Number of monthly births and deaths: Beira Baixa NUTS3 Region

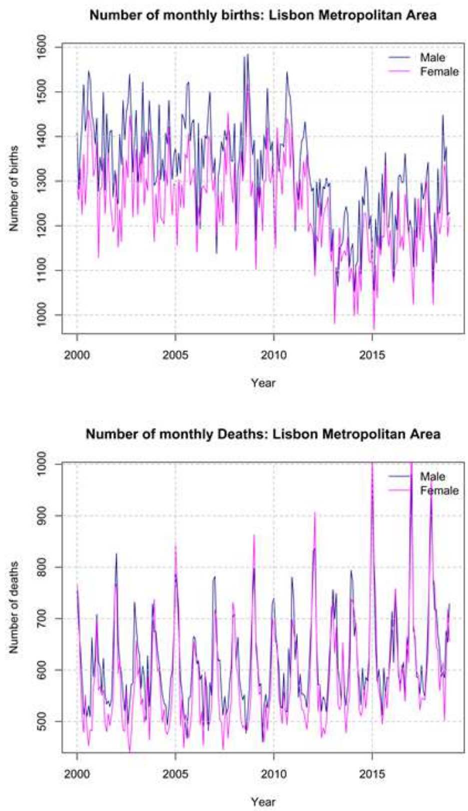


Figure 2 – Number of monthly births and deaths: Lisbon Metropolitan Area NUTS3 Region

Examination of the time plots revealed that there is a negative trend in the births series over the time period considered, although some recovery is observed in the Lisbon Metropolitan Area (LMA) in the years following the end of the Troika adjustment program; in the case of deaths time series we do not observe a significant trend over this period. Overall, a seasonal pattern is evident in the behaviour of live births and deaths, with the highest number of births in the spring and summer months while the highest number of deaths occurs during the winter months. Substantial changes are observed in the trend of fertility, with the number of live births showing a declining trend after 2000 in the majority of NUTS 3 regions. Since 2015, a relative stabilisation and even a small increase are being observed. Over time, albeit the slight increase in the total number of deaths in the last years, time mortality patterns are relatively stable, showing a strong seasonal pattern with a higher number of deaths in winter months.

## 4 EMPIRICAL RESULTS

The three univariate time series models are used as predictive models for making forecasts for future values of live births and deaths by sex and NUTS3 regions in Portugal. The MAPE results of 1-year ahead forecasts of monthly births and deaths by sex and NUTS3 regions for the period 2014-2018 averaged over all jump-off years with the different models are given in Tables 1 and 2, respectively. The results averaged (simple and weighted averages) over all 25 regions and five launch years are shown in the Tables. Additionally, Tables 1 and 2 include data on the population size of each NUTS3 region in December 2017 to ascertain whether the model's relative forecasting performance is a function of population size.

We first discuss the results related to monthly births forecasting. The all regions and launch years simple and weighted average forecasting performance for the three models tested are similar for both male and female subpopulations showing relatively small average MAPE results. The simple average results show that the precision of the SARIMA forecasts is better than that of Holt-Winters (HW) and State Space (SS) models for the female subpopulations but, for the male counterparts, SS models show slightly lower forecasting errors. Note, however, that when considering the weighted average results (with weights given by the proportion of the region's subpopulation in the total resident population) SS models exhibit higher forecasting accuracy due to their superior performance in highly populated areas. Using this later metric, the SS model advantages the SARIMA and HW models by 0,17 (0,18) and 0,16 (0,35) percentage points in the female (male) subpopulations, respectively.

On average for all models and for 61,3% of the subpopulations the forecasting errors are smaller for the male subpopulations when compared to their female counterparts. As expected, the average MAPE results over the five launch years are larger, the smaller the region's population size. The largest average forecasting error (24,19%) is found in the Beira Baixa female subpopulation using the SS model whereas the highest accuracy (having 3,11% MAPE) is attained in the Lisbon metropolitan area ("Área Metropolitana de Lisboa") also using the SS model. The forecasting error is less than 10% in 40% of the subpopulations considered.

| Births | Females | | | | Males | | | |
|---|---|---|---|---|---|---|---|---|
| NUTS 3 | Pop. Size | ARIMA | HW | SS | Pop. Size | ARIMA | HW | SS |
| Alto Minho | 124583 | **13.20** | 13.38 | 14.13 | 107595 | **12.46** | 13.07 | 12.87 |
| Cávado | 211950 | 10.23 | 10.39 | **9.63** | 192003 | 11.61 | 10.44 | **9.62** |
| Ave | 215975 | 10.50 | 9.57 | **8.67** | 197879 | 9.52 | **6.90** | 8.13 |
| Área Metropolitana do Porto | 910200 | 6.14 | **5.35** | 5.45 | 809502 | **4.74** | 5.40 | 5.04 |
| Alto Tâmega | 46044 | **18.49** | 18.86 | 21.03 | 41113 | 23.66 | **21.11** | 21.26 |
| Tâmega e Sousa | 216999 | 8.96 | **8.39** | 9.18 | 201769 | 8.62 | **7.61** | 8.55 |
| Douro | 101142 | **13.49** | 13.76 | 13.97 | 90904 | 14.27 | 14.86 | **13.88** |
| Terras de Trás-os-Montes | 56870 | 15.33 | **15.02** | 16.95 | 51677 | **16.38** | 16.44 | 15.70 |
| Oeste | 186405 | 9.91 | 10.18 | **9.53** | 171301 | **7.84** | 8.82 | 8.07 |
| Região de Aveiro | 190926 | 8.84 | 8.75 | **8.22** | 172169 | 8.47 | 7.56 | **6.95** |
| Região de Coimbra | 231654 | **8.15** | 8.21 | 7.84 | 205294 | 7.86 | 7.70 | **7.34** |
| Região de Leiria | 149784 | 9.20 | 8.85 | **7.84** | 136525 | 9.86 | 10.79 | **9.75** |
| Viseu Dão Lafões | 134679 | 12.15 | **11.62** | 12.21 | 119952 | 12.64 | 12.53 | **12.12** |
| Beira Baixa | 43061 | 21.60 | **21.31** | 24.19 | 38753 | **16.40** | 17.92 | 17.23 |
| Médio Tejo | 123699 | **10.53** | 11.54 | 12.01 | 110956 | **9.99** | 10.63 | 10.21 |
| Beiras e Serra da Estrela | 114163 | 12.72 | **12.43** | 12.97 | 102025 | 11.11 | **10.75** | 12.37 |
| Área Metropolitana de Lisboa | 1505435 | 3.38 | 3.60 | **3.11** | 1328244 | 3.59 | 4.18 | **3.29** |
| Alentejo Litoral | 47551 | **16.99** | 17.56 | 17.78 | 46223 | **17.32** | 18.77 | 19.11 |
| Baixo Alentejo | 60669 | **11.59** | 12.45 | 12.57 | 57199 | 14.33 | **14.17** | 14.59 |
| Lezíria do Tejo | 124049 | **8.00** | 9.50 | 8.66 | 114666 | 11.56 | **10.94** | 11.48 |
| Alto Alentejo | 56092 | 18.80 | 19.01 | **18.54** | 50965 | **16.32** | 16.33 | 17.15 |
| Alentejo Central | 80677 | **12.81** | 13.87 | 13.01 | 73859 | **12.01** | 13.61 | 12.80 |
| Algarve | 229719 | 7.33 | 8.30 | **7.02** | 209898 | 7.40 | 7.46 | **7.24** |
| RA Açores | 125052 | 10.77 | **10.49** | 10.59 | 118810 | 9.78 | 9.78 | **9.52** |
| RA Madeira | 135957 | **12.00** | 12.30 | 14.02 | 118411 | **10.39** | 11.51 | 11.93 |
| All regions and launch years: | | | | | | | | |
| Simple Average | 216933 | **11.64** | 11.79 | 11.96 | 194708 | 11.53 | 11.57 | **11.45** |
| Weighted Average | | 8.00 | 7.99 | **7.83** | | 7.70 | 7.87 | **7.52** |
| Max | 1505435 | 21.60 | 21.31 | 24.19 | 1328244 | 23.66 | 21.11 | 21.26 |
| Min | 43061 | 3.38 | 3.60 | 3.11 | 38753 | 3.59 | 4.18 | 3.29 |

Table 1 – Births Forecasting - Average MAPE by Model, Sex and NUTS3

**Source:** Authors preparation; **Notes:** Average Mean Absolute Percent Error (MAPE) by model (ARIMA; Holt-Winters (HW); State Space (SS)) Sex and NUTS3 Region for the period 2014-2018. Weighted Average computed using the proportion of region's male or female population in the corresponding (sex) total population. The best (smaller MAPE) values are highlighted in bold.

Moving now to the results related to 1-year ahead monthly deaths forecasting, Table 2 shows once again that the all regions and launch years simple and weighted average forecasting performance for the three models was relatively similar for both the male and female subpopulations, although the differences between the worst and the best performing model is higher in the male subset. Compared to births results, the average (weighted) forecasting accuracy of the alternative univariate time series methods is lower in the female subpopulations but higher in the male group. The weighted average results show that the precision of SARIMA forecasts is consistently better than that of the Holt-Winters (HW) and State Space (SS) models.

The SARIMA model advantages the HW and SS models by 0,58 (0,31) and 0,19 (0,08) percentage points in the female (male) subpopulations, respectively. On average for all models and for 76% of the subpopulations the forecasting errors are notably smaller for the male subpopulations when compared to their female counterparts.

| Deaths | Females | | | | Males | | | |
|---|---|---|---|---|---|---|---|---|
| NUTS 3 | Pop. Size | ARIMA | HW | SS | Pop. Size | ARIMA | HW | SS |
| Alto Minho | 124583 | **10.36** | 11.27 | 10.64 | 107595 | **9.20** | 9.29 | 9.44 |
| Cávado | 211950 | 11.45 | 11.21 | **11.07** | 192003 | **8.88** | 9.14 | 8.99 |
| Ave | 215975 | **7.73** | 9.27 | 8.55 | 197879 | **8.89** | 9.31 | 9.05 |
| Área Metropolitana do Porto | 910200 | **7.24** | 8.07 | 7.85 | 809502 | **5.76** | 6.33 | 6.12 |
| Alto Tâmega | 46044 | 11.71 | 12.00 | **11.35** | 41113 | **12.69** | 15.07 | 14.94 |
| Tâmega e Sousa | 216999 | **9.02** | 11.21 | 10.33 | 201769 | **8.94** | 9.61 | 8.99 |
| Douro | 101142 | **11.25** | 13.74 | 12.21 | 90904 | 9.88 | 10.27 | **9.73** |
| Terras de Trás-os-Montes | 56870 | 11.46 | 12.35 | **11.35** | 51677 | **10.53** | 11.07 | 10.71 |
| Oeste | 186405 | **7.30** | 8.11 | 7.65 | 171301 | 8.09 | 7.99 | **7.75** |
| Região de Aveiro | 190926 | 10.29 | 10.47 | **10.04** | 172169 | **7.94** | 9.20 | 8.20 |
| Região de Coimbra | 231654 | **7.54** | 7.56 | 7.57 | 205294 | 7.29 | **7.16** | 7.38 |
| Região de Leiria | 149784 | **9.57** | 9.98 | 9.63 | 136525 | 9.74 | 9.90 | **9.62** |
| Viseu Dão Lafões | 134679 | 9.91 | 10.35 | **9.80** | 119952 | 8.28 | 8.79 | **7.80** |
| Beira Baixa | 43061 | 14.26 | **14.13** | 14.96 | 38753 | 12.75 | **11.73** | 13.95 |
| Médio Tejo | 123699 | 8.10 | 7.95 | **7.74** | 110956 | **8.87** | 9.12 | 9.11 |
| Beiras e Serra da Estrela | 114163 | **10.29** | 11.48 | 10.34 | 102025 | 8.46 | 8.10 | **8.07** |
| Área Metropolitana de Lisboa | 1505435 | 6.01 | 6.07 | **5.89** | 1328244 | 5.07 | 5.01 | **4.99** |
| Alentejo Litoral | 47551 | 11.97 | 13.04 | **11.46** | 46223 | **13.24** | 16.11 | 15.24 |
| Baixo Alentejo | 60669 | **11.80** | 13.06 | 12.48 | 57199 | 10.09 | 10.29 | **10.00** |
| Lezíria do Tejo | 124049 | **9.48** | 10.33 | 9.97 | 114666 | 9.07 | 9.85 | **8.87** |
| Alto Alentejo | 56092 | 10.65 | 11.56 | **10.57** | 50965 | **11.29** | 11.71 | 11.48 |
| Alentejo Central | 80677 | **9.74** | 10.49 | 10.93 | 73859 | 9.22 | 9.52 | **8.98** |
| Algarve | 229719 | 9.26 | 9.65 | **8.94** | 209898 | 7.57 | 7.50 | **7.33** |
| RA Açores | 125052 | **10.90** | 11.72 | 11.33 | 118810 | **9.67** | 11.31 | 10.52 |
| RA Madeira | 135957 | **9.78** | 10.86 | 9.82 | 118411 | 9.53 | 10.05 | **9.50** |
| All regions and launch years: | | | | | | | | |
| Simple Average | 216933 | **9.88** | 10.64 | 10.10 | 194708 | **9.24** | 9.74 | 9.47 |
| Weighted Average | | **8.25** | 8.83 | 8.44 | | **7.35** | 7.66 | 7.43 |
| Max | 1505435 | 14.26 | 14.13 | 14.96 | 1328244 | 13.24 | 16.11 | 15.24 |
| Min | 43061 | 6.01 | 6.07 | 5.89 | 38753 | 5.07 | 5.01 | 4.99 |

Table 2 – Deaths Forecasting - MAPE by Model, Sex and NUTS3

**Source:** Authors preparation; **Notes:** Average Mean Absolute Percent Error (MAPE) by model (ARIMA; Holt-Winters (HW); State Space (SS)) Sex and NUTS3 Region for the period 2014-2018. Weighted Average computed using the proportion of region's male or female population in the corresponding (sex) total population. The best (smaller MAPE) values are highlighted in bold.

Similar to the births results, the average MAPE results over the five launch years are smaller, the more populated the region is. The largest average forecasting error (16,11%) is found in the Alentejo Litoral male subpopulation using the HW model whereas the highest accuracy (4,99%) is attained in the Lisbon metropolitan area ("Área Metropolitana de Lisboa") male subpopulation using the SS model. The forecasting error is less than 10% in 57% of the subpopulations considered. Table 3 reports the percentage of monthly birth/death counts falling outside the 95% prediction interval estimated for each model, sex and NUTS3 Region.

| NUTS 3 | Births | | | | | | Deaths | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Females | | | Males | | | Females | | | Males | | |
| | AR | HW | SS | AR | HW | SS | AR | HW | SS | AR | HW | SS |
| Alto Minho | **1.7** | 7.0 | 2.3 | **1.4** | 4.2 | 1.8 | **1.3** | 1.7 | 1.7 | 1.7 | 9.3 | **1.0** |
| Cávado | 5.0 | 9.0 | **2.0** | 3.6 | 4.2 | **0.8** | 2.0 | 5.0 | **1.7** | **1.3** | 1.0 | 1.7 |
| Ave | 4.0 | 7.0 | **1.0** | 2.8 | **0.8** | **0.8** | 1.0 | **0.3** | 0.7 | 2.3 | 4.0 | **2.0** |
| Área Metropolitana do Porto | 3.3 | 3.7 | **3.0** | **1.0** | **1.0** | 0.2 | **2.0** | 12.7 | **2.0** | **1.0** | 10.3 | 2.3 |
| Alto Tâmega | **2.3** | 10.7 | 2.7 | 1.8 | 2.6 | **1.2** | 0.3 | 9.7 | **0.0** | **1.3** | 11.7 | 2.0 |
| Tâmega e Sousa | 2.3 | 2.3 | **1.3** | 2.0 | **1.4** | **1.4** | 1.7 | 6.0 | **1.0** | **1.0** | 1.7 | 1.3 |
| Douro | 3.7 | 6.3 | **0.3** | 1.8 | 4.8 | **1.4** | **1.0** | 1.3 | 1.3 | 1.7 | 6.3 | **1.0** |
| Terras de Trás-os-Montes | 3.0 | 8.7 | **2.0** | 1.2 | 5.4 | **0.2** | **1.0** | 1.0 | 1.3 | 1.3 | 1.0 | **0.7** |
| Oeste | 1.3 | 6.0 | **0.0** | **0.8** | 3.4 | 0.6 | 0.7 | 2.3 | **0.3** | 1.7 | 0.7 | **1.0** |
| Região de Aveiro | 2.0 | 2.0 | **0.7** | 1.4 | 3.2 | **0.6** | **1.3** | 4.7 | **1.3** | **1.0** | 9.0 | **1.0** |
| Região de Coimbra | 2.0 | 3.3 | **0.7** | 1.2 | 3.2 | **0.8** | **1.3** | 9.0 | **1.3** | 1.3 | 10.3 | **1.0** |
| Região de Leiria | 1.7 | 7.0 | **1.3** | 1.4 | 3.8 | **0.6** | **1.7** | 6.3 | **1.7** | 3.3 | 1.7 | 3.3 |
| Viseu Dão Lafões | 2.3 | 8.3 | **1.3** | 1.4 | 5.8 | **0.0** | **1.3** | 7.7 | **1.3** | **1.0** | 7.3 | **1.0** |
| Beira Baixa | **0.7** | 8.7 | **0.7** | 0.2 | 0.2 | **0.0** | 1.3 | 13.3 | **0.3** | 1.7 | 10.3 | **0.7** |
| Médio Tejo | **2.0** | 9.0 | **2.0** | 1.4 | 3.4 | **0.4** | 0.3 | 0.3 | **0.0** | 1.0 | 0.7 | **0.7** |
| Beiras e Serra da Estrela | 2.0 | 8.3 | **1.0** | 1.4 | 3.8 | **0.8** | 1.3 | 2.0 | **0.7** | **1.0** | 0.3 | **1.0** |
| Área Metropolitana de Lisboa | **0.0** | 1.3 | **0.0** | 0.6 | 2.2 | **0.2** | **1.3** | 8.7 | 1.7 | **2.0** | 10.0 | **2.0** |
| Alentejo Litoral | **1.3** | 6.0 | **1.3** | 0.4 | 0.0 | **0.2** | 1.3 | 0.7 | **1.0** | 2.0 | 1.7 | **1.7** |
| Baixo Alentejo | 1.0 | 3.0 | **0.7** | 1.2 | 6.4 | **0.8** | 2.0 | 0.7 | **1.0** | 1.7 | 5.7 | **0.7** |
| Lezíria do Tejo | 0.3 | 2.7 | **0.0** | **1.6** | 5.6 | 1.8 | **1.3** | 10.3 | 2.3 | 2.3 | 2.0 | **1.3** |
| Alto Alentejo | 1.7 | 4.7 | **1.3** | 0.8 | 7.0 | **0.0** | 0.7 | 12.0 | **0.3** | **1.7** | 7.0 | **1.7** |
| Alentejo Central | 0.7 | **0.0** | 0.3 | 0.8 | 3.2 | **0.4** | **0.3** | 11.0 | 1.0 | 1.0 | 0.7 | **0.3** |
| Algarve | **0.3** | 3.7 | 0.7 | 0.6 | 0.6 | **0.2** | **1.7** | 9.0 | 2.0 | **2.7** | 12.7 | **2.7** |
| RA Açores | 1.7 | 8.7 | **0.7** | 1.2 | 3.0 | **0.6** | **1.3** | 0.3 | **1.3** | 1.3 | **0.7** | **0.7** |
| RA Madeira | 2.7 | 10.0 | **1.7** | 1.2 | 3.0 | **0.4** | **0.7** | 0.3 | **0.7** | 1.0 | **0.7** | 1.7 |
| All regions and launch years: | | | | | | | | | | | | |
| Simple Average | 2.0 | 5.9 | **1.2** | 1.3 | 3.3 | **0.6** | 1.2 | 5.5 | **1.1** | 1.6 | 5.1 | **1.4** |
| Weighted Average | 1.7 | 4.3 | **1.1** | 1.2 | 2.6 | **0.5** | **1.4** | 7.2 | **1.4** | **1.6** | 7.1 | 1.7 |
| Max | 5.0 | 10.7 | 3.0 | 3.6 | 7.0 | 1.8 | 2.0 | 13.3 | 2.3 | 3.3 | 12.7 | 3.3 |
| Min | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.3 | 0.3 | 0.0 | 1.0 | 0.3 | 0.3 |

Table 3 – Percentage of monthly birth/death counts falling outside the 95% prediction interval by model, sex & NUTS3

**Source:** Authors preparation; **Notes:** Average Mean Absolute Percent Error (MAPE) by model (AR=SARIMA; Holt-Winters (HW); State Space (SS)) Sex and NUTS3 Region for the period 2014-2018. Weighted Average computed using the proportion of region's male or female population in the corresponding (sex) total population. The best (smaller percentage error) values are highlighted in bold.

The goal is to measure how well the models analysed in this paper perform in terms of predicting the uncertainty of future monthly birth/death counts over 1-year forecasting horizons. Each cell in the table is based on 60 forecasts (five years and 12 monthly observations per year). Considering the 95% prediction intervals a valid measure of uncertainty means they will encompass 57 of the 60 out-of-sample observed monthly birth/death counts or, conversely, only 3 of the 60 observations will fall outside the 95% prediction interval boundaries. According to this criterion, the prediction intervals for the SARIMA and SS models consistently provide appropriate measures of uncertainty for short-term forecasting horizons. The SARIMA and SS models perform equally well in terms of predicting the uncertainty of future monthly death counts, with SS models slightly overperforming in births forecasting. On the contrary, the HW model consistently fails in predicting the uncertainty of future monthly birth and deaths with up to 13,3% of observed death counts falling out of the 95% prediction interval.

## 5 CONCLUSIONS AND POLICY IMPLICATIONS

Ageing populations and internal and international migration flows are key demographic trends facing developed and developing countries and its regions in the coming years. There is a territorial dimension of demographic change, with clear territorial differences in the ageing pattern around the world. This means that the responses proposed and implemented in multiple policy areas (e.g., economical, social, infrastructure, spatial planning) have to act on different spatial levels and in different ways. For end users of population projections, it is critical to improve the accuracy of projection series, particularly at the regional and local level, and to fully understand their reliability and limitations. Being conscious about how projections are computed and the potential sources of uncertainty in the population numbers is expected to assist policymakers in appropriately incorporating projections in their planning and decision-making process. The population of a given territorial area and its age distribution changes over time through the interaction of three possibly correlated factors: fertility, mortality, and migration. To project population size at a future date, economists and demographers use stochastic time series methods to project the dynamics of three components and eventually incorporate expert-based assumptions on long-term demographic trends.

Monthly time series of live births and deaths exhibit significant and persistent seasonality patterns, requiring the adoption of appropriate forecasting methods to increase the accuracy of population forecasts. In this paper we empirically evaluated the forecasting performance of seasonal ARIMA, Holt-Winters and State Space models applied to birth and death monthly forecasting by sex and NUTS 3 regions for Portugal, and investigate how well these models perform in terms of predicting the uncertainty of future monthly birth and death counts using a backtesting framework and monthly data for the period 2000-2018. The all regions and launch years simple and weighted average forecasting performance for the three models was relatively similar for both male and female subpopulations births and deaths; however, SS models showed slightly better performance for births and seasonal ARIMA for deaths. As expected, the weighted average precision is higher, the more populated the region is. The prediction intervals for the SARIMA and SS models consistently provide appropriate measures of uncertainty for short-term forecasting horizons. Further research should check for the robustness of these results against alternative forecasting horizons and fixed lookback windows using rolling fixed-length horizon backtests. Future research will also investigate the robustness of these results against alternative primary, extended, composite, and hybrid performance metrics used in machine learning regression, forecasting and prognostics, considering for competing distance measures and normalization and aggregation procedures.

Our study contributes to improve projections of future regional and local populations which are essential for public and private sector planning. They are used to determine the budget allocation from central to local government departments and agencies, are vital in the design and implementation of spatial (e.g. housing, education, infrastructure, land use, environmental, service networks) policies, assist in the design and reform of public and private pension schemes and public finance planning. Better sub-national population projections are needed to help in the design and implementation of migration and immigration policies since significant geographical differences in income, living standards and long-term development are likely to cause sizeable migration flows. Additionally, the expected decline of the total population and of the working age population in many regions in Portugal and in Europe may increase the number of labour migrants. Because of the compounding effects of current and past fertility levels on future population numbers and age structure, it is crucial to accurately forecast the number of births at the regional level. Fertility levels below replacement rate have a compounding effect on the dynamics of future population which can only partially offset by positive net immigration.

## REFERENCES


Abel, Guy and Bijak, Jakub and Forster, Jonathan J. and Raymer, James and Smith, Peter W.F. and Wong, Jackie S.T. (2013) Integrating uncertainty in time series population forecasts: An illustration using a simple projection model. *Demographic Research*, 29 (43). pp. 1187-1226.

Ahlburg, D. (1992). Error measures and the choice of a forecast method. *International Journal of Forecasting*, 8, 99–100.

Alho, J. M. and Spencer, B.D. (1985). Uncertain Population Forecasting. *Journal of the American Statistical Association* 80(390): 306–314.

Alho, J., Bravo, J. M. & Palmer, E. (2012). *Annuities and Life Expectancy in NDC*. In Holzmann, R. E. Palmer and D. Robalino (Eds.), Nonfinancial defined contribution Pension Schemes in a Changing Pension World, Vol. 2 Gender, Politics, and Financial Stability, 395 - 436.

Ayuso, M., Bravo, J. M., & Holzmann, R. (2017a). Addressing Longevity' Heterogeneity in Pension Scheme Design. *Journal of Finance and Economics*, 6(1), 1–21.

Ayuso, M., Bravo, J. M., & Holzmann, R. (2017b). On the Heterogeneity in Longevity among Socioeconomic Groups: Scope, Trends, and Implications for Earnings-Related Pension Schemes. *Global Journal of Human Social Sciences - Economics*, 17(1), 31–57.

Ayuso, M., Bravo, J. M. & Holzmann, R. (2019). Getting Life Expectancy Estimates Right for Pension Policy: Period versus Cohort Approach. *Journal of Pension Economics and Finance*. Forthcoming.

Box, G. E., & Jenkins, G. M. (1976). *Time series analysis, control, and forecasting*. San Francisco, CA: Holden Day, 3226(3228), 10.

Box, G.E.P., & G.M. Jenkins (1970). *Time Series Analysis: Forecasting and Control*, San Francisco: Holden Day (revised ed. 1976).

Bravo, J. M. (2007). *Tábuas de Mortalidade Contemporâneas e Prospectivas: Modelos Estocásticos, Aplicações Actuariais e Cobertura do Risco de Longevidade*. Dissertação de Doutoramento em Economia pela Universidade de Évora, Maio. Retrieved from: http://www.rdpc.uevora.pt/handle /10174/11148?locale=pt

Bravo, J. M. (2016). Taxation of Pensions in Portugal: A Semi-Dual Income Tax System. *CESifo DICE Report - Journal for Institutional Comparisons*. 14 (1), 14-23, 1/2016 (Spring), Ifo Institute, Munich.

Bravo, J. M. & Silva, C. M. P. (2006). Immunization Using a Stochastic Process Independent Multifactor Model: The Portuguese Experience. *Journal of Banking and Finance*, 30 (1), 133-156.

Bravo, J. M., & El Mekkaoui de Freitas, N. (2018). Valuation of longevity-linked life annuities. *Insurance: Mathematics and Economics*, 78, 212–229.

Bravo, J. M., Coelho, E., & Magalhães, M. G. (2010). *Mortality projections in Portugal*. In EUROSTAT - European Commission (eds.), Work session on demographic projections, EUROSTAT-EC Collection: Methodologies and working papers, Theme: Population and Social Conditions, 241–252.

Bravo, J. M., Rodrigues, T., Ribeiro, S. & Inácio, A. (2018). *Portugal. Projeções de População Residente 2011-2040*. In Teresa Rodrigues & Marco Painho (Coord.). Modelos Preditivos e Segurança Pública. Fronteira do Caos Editores, 168-214.

Canova, F. & Hansen, B.E. (1995). Are Seasonal Patterns Constant over Time? A Test for Seasonal Stability. *Journal of Business and Economic Statistics* 13, 237–252.

Chamboko, R., & Bravo, J. M. (2016). On the modelling of prognosis from delinquency to normal performance on retail consumer loans. *Risk Management*, 18(4), 264–287.

Chamboko, R. & Bravo, J. M. (2019a). Modelling and forecasting recurrent recovery events on consumer loans. *International Journal of Applied Decision Sciences*, 12(3), 271-287.

Chamboko, R. & Bravo, J. M. (2019b). Frailty correlated default on retail consumer loans in developing markets. *International Journal of Applied Decision Sciences*, 12(3), 257–270.

Holt, C.C. (1957). Forecasting seasonals and trends by exponentially weighted averages. O.N.R. Memorandum 52/1957, Carnegie Institute of Technology. Reprinted with discussion in 2004, *International Journal of Forecasting*, 20, 5–13.

Hyndman, R. & Athanasopoulos, G. (2013). *Forecasting: principles and practice*. OTexts (OTexts.org/fpp2/).

Hyndman, R.J., Koehler, A. B., Snyder, R.D., & Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18, 439–454.

Hyndman, R.J., Koehler, A.B., Ord, J.K., & Snyder, R.D. (2008). *Forecasting with Exponential Smoothing: The State Space Approach*. Springer.

Keilman, N., Pham, D., & Hetland, A. (2002). Why population forecasts should be probabilistic – illustrated by the case of Norway. *Demographic Research*, 6, 409–453.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. & Shin, Y. (1992). Testing Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure are we that Economic Time Series have a Unit Root?. *Journal of Econometrics* 54, 159–178

Land, K. & Cantor, D. (1983). ARIMA models of seasonal variation in U. S. birth and death rates. *Demography*, 20(4), 541-568.

Lee, R. (1974). Forecasting births in post-transition populations: Stochastic renewal with serially correlated fertility. *Journal of the American Statistical Association*, 69, 607–617.

Lee, R. (1992). Stochastic demographic forecasting. *International Journal of Forecasting*, 8, 315–327.

Lee, R., & Tuljapurkar, S. (1994). Stochastic population forecasts for the United States: Beyond high, medium, and low. *Journal of the American Statistical Association*, 89, 1175–1189

McNown, R., & Rogers, A. (1989). Forecasting mortality: A parameterized time series approach. *Demography*, 26, 645–660.

Ord, J. K., Koehler, A. B., & Snyder, R. D. (1997). Estimation and prediction for a class of dynamic nonlinear statistical models. *Journal of American Statistical Association*, 92, 1621–1629.

Pflaumer, P. (1992). Forecasting U.S. population totals with the Box–Jenkins approach. *International Journal of Forecasting*, 8, 329–338.

R Development Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL www.R-project.org.

Saboia, J. L. M. (1974). Modeling and Forecasting Populations by Time Series: The Swedish Case. *Demography* 11(3): 483–492.

Smith, S., Tayman, J., & Swanson, D. (2001). *State and local population projections: Methodology and analysis*. New York, NY: Kluwer Academic/Plenum Publishers.

Swanson, D., & Beck, D. (1994). A new short-term county population projection method. *Journal of Economic and Social Measurement*, 20, 25–50.

Tayman, J., Smith, S. K., and Lin, J. (2007). Precision, bias, and uncertainty for state population forecasts: an exploratory analysis of time series models. *Population Research and Policy Review*, 26(3), 347–369.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6, 324–342