

Stochastic structures of world's death counts after World War II

Jae J. Lee^{1,a}

^aSchool of Business, State University of New York, USA

Abstract

This paper analyzes death counts after World War II of several countries to identify and to compare their stochastic structures. The stochastic structures that this paper entertains are three structural time series models, a local level with a random walk model, a fixed local linear trend model and a local linear trend model. The structural time series models assume that a time series can be formulated directly with the unobserved components such as trend, slope, seasonal, cycle and daily effect. Random effect of each unobserved component is characterized by its own stochastic structure and a distribution of its irregular component. The structural time series models use the Kalman filter to estimate unknown parameters of a stochastic model, to predict future data, and to do filtering data. This paper identifies the best-fitted stochastic model for three types of death counts (Female, Male and Total) of each country. Two diagnostic procedures are used to check the validity of fitted models. Three criteria, AIC, BIC and SSPE are used to select the best-fitted valid stochastic model for each type of death counts of each country.

Keywords: death counts, structural time series model, state space form, Kalman filter

1. Introduction

This paper analyzes yearly death counts after World War II of 8 countries in three regions, North America, Europe, and Asia-Pacific region to identify and to compare stochastic structures of death counts. The 8 countries are the United States and Canada in North America, United Kingdom, France, Italy, and Spain in Europe, and Taiwan and Australia in Asia-Pacific region. Death counts are from the year 1946 or 1970 (depending on availability) and are separated by gender and total counts to see whether gender influences the stochastic structures. The structural time series models (Harvey, 1981, 1989) assume that a time series can be formulated directly with the unobserved components such as trend, slope, seasonal, cycle, and daily effect. The random effect of each unobserved component is characterized by its stochastic structure and a distribution of its irregular component. Structural time series models that this paper entertained are a local level with a random walk model, a fixed local linear trend model, and a local linear trend model. These models are sensible choices based on the preliminary examination of death counts data. Structural time series models use the Kalman filter (Kalman, 1960) to estimate unknown parameters of the entertained model, forecast future values of time series, and estimate unobserved components in the stochastic model by filtering and smoothing. To apply the Kalman filter, the structural time series models need to be converted to a state space form (Durbin and Koopman, 2012) that is the standardized form for the Kalman filter. To check the validity, two diagnostic procedures are used for a fitted model: One is for checking the normality of

¹ School of Business, State University of New York, 1 Hawk Road, New Paltz, NY, USA. E-mail: leej@newpaltz.edu

residuals, for which the Shapiro-Wilk test is used and the normal QQ plot and density plot are used to confirm the results of the Shapiro-Wilk test. Second is for checking the independence of residuals, for which the Run test is used. To find the best-fitted model among valid models, Akaike information criterion (AIC), Bayesian information criterion (BIC) and sum of square of one-step-ahead prediction errors (SSPE) are used. The best-fitted valid models of death counts for each country by female, male and total are fully examined to see any differences or similarities among countries and regions. The organization of this paper is as follows. In Section 2, the structural time series model is presented. It also shows how to set up the state space form for each structural time series model entertained. In Section 3, the Kalman filter is introduced. This section also shows which R packages to use for the Kalman filter. In Section 4, results of analyzing death counts of 8 countries by female, male and total are presented. Finally, Section 5 concludes the paper.

2. Structural time series models and their state space forms

The structural time series models assume that a time series can be formulated directly with the unobserved components that are characterized by its own stochastic structure and an irregular term. By varying its own stochastic structure and the distribution of an irregular term of an unobserved component, structural time series models can fit a variety of time series in many fields. Outcomes of a structural time series model are estimates of unknown parameters, forecasts of future values of time series, and estimates of unobserved components in the model. Estimates of unobserved components of a model give an insight to fully understand a stochastic structure of a time series of interest. There are many papers regarding the structural time series model. For example, Harvey and Todd (1983) compared the structural time series model with Box and Jenkins' ARIMA model. Harvey and Peters (1990) showed the number of methods to compute the maximum likelihood estimator of unknown parameters of the structural time series model.

For death counts data, three structural time series models are entertained: local level with random walk model, fixed local linear trend model, and local linear trend model. Choices of these models are based on the examinations of plots of death counts of 8 countries by female, male and total. Plots of death counts of all 8 countries by female, male and total are presented in Section 4.

2.1. Local level with a random walk model

First structural time series model entertained in this paper is the local level with a random walk (LLRW) model:

$$\begin{cases} y_t = \mu_t + \epsilon_t, \\ \mu_t = \mu_{t-1} + \eta_t, \end{cases} \quad (2.1)$$

where y_t is an observed time series data, μ_t is an unobserved trend component that represents the long-term movement in a time series, η_t is an irregular component that shows the stochastic behavior of the trend of time series, and ϵ_t is an irregular component that shows the stochastic behavior other than the trend defined in the second equation in (2.1).

2.2. Local linear trend model

Second model entertained is the local linear trend (LT) model:

$$\begin{cases} y_t = \mu_t + \epsilon_t, \\ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \\ \beta_t = \beta_{t-1} + \zeta_t, \end{cases} \quad (2.2)$$

where y_t is an observed time series data, μ_t is an unobserved trend component, β_t is an unobserved slope component, ζ_t is an irregular component that shows the stochastic behavior of the slope of the trend of time series. η_t and ϵ_t are defined as (2.1). This model shows that not only a trend of series but also a slope of the trend is stochastic with its own irregular component.

2.3. Fixed local linear trend model

Third model is the fixed local linear trend (FT) model that is a variation of LT model:

$$\begin{cases} y_t = \mu_t + \epsilon_t, \\ \mu_t = \mu_{t-1} + \beta + \eta_t, \end{cases} \quad (2.3)$$

where y_t is an observed time series data, β is a deterministic slope of the trend, μ_t , η_t , and ϵ_t are defined as (2.1) and (2.2). This model assumes that a series has a deterministic slope, rather than a stochastic.

2.4. State space form

Structural time series model uses the Kalman filter to estimate unknown parameters, to forecast future values and to estimate unobserved components. To use the Kalman filter, a state space form (Durbin and Koopman, 2012) is required. The state space form is a standardized form of a stochastic model as an input for the Kalman filter. For the LLRW model (2.1), it is the state space form itself. Thus, the state space form (2.1) has 1×1 state vector, μ_t and two irregular components, η_t and ϵ_t . The state space form of the LT model (2.2) is:

$$\begin{cases} y_t = (1 \quad 0) \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} + \epsilon_t, \\ \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_t \\ \zeta_t \end{pmatrix}. \end{cases} \quad (2.4)$$

The state space form (2.4) has 2×1 state vector, $\begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}$ and three irregular components, η_t , ζ_t and ϵ_t . The state space form of the FT model (2.3) is:

$$\begin{cases} y_t = (1 \quad 0) \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} + \epsilon_t, \\ \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \eta_t. \end{cases} \quad (2.5)$$

The state space form (2.5) has 2×1 state vector, $\begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}$ and two irregular components, η_t and ϵ_t .

3. Kalman filter

Kalman filter has been applied to many fields since Kalman's first paper (Kalman, 1960) was published. Harrison and Stevens (1971) was the first paper to apply the Kalman filter to a time series analysis. Since then, the Kalman filter has been used to analyze several time series models such as ARIMA models (Box and Jenkins, 1976), structural time series models (Harvey, 1989), and ARMAX models (Hannan and Deistler, 1988). Some areas where it has been applied include disease control (Gove and Houston, 1996), actuary claim reserves forecasting (Chukhrova and Johannssen, 2017), rain fall forecasting (Asemota *et al.*, 2016; Zulfi *et al.*, 2018), and machine learning (Nobrega and Oliveira, 2019).

Kalman filter can be applied to either univariate and multivariate time series, and to either time variant structure or time invariant structure of time series. Time series data, Y_t , for $t = 1, 2, \dots, T$ denotes the observed values of a time series of interest. Y_t could be a univariate or multivariate. α_t denotes the unobserved component vector, called the state vector. For the Kalman filter, the observed data Y_t and the unobserved state vector α_t have the linear relationship as

$$Y_t = Z_t \alpha_t + \epsilon_t, \quad t = 1, 2, \dots, T, \quad (3.1)$$

where Y_t is $N \times 1$ ($N = 1$ for the univariate and $N > 1$ for the multivariate), Z_t is $N \times m$ matrix, α_t is $m \times 1$ vector, and ϵ_t is $N \times 1$ vector of irregular components. For equation (3.1), Z_t is assumed to be a known quantity that shows the relationship between Y_t and α_t , and ϵ_t is assumed to have a multivariate normal distribution with mean of $N \times 1$ vector of zero and covariance of $N \times N$ matrix, h_t . Equation (3.1) is called the observation equation. For a univariate Y_t , the observation equation (3.1) can be written as

$$Y_t = z_t' \alpha_t + \epsilon_t, \quad t = 1, 2, \dots, T, \quad (3.2)$$

where Y_t is a scalar, z_t' is $1 \times m$ vector, α_t is $m \times 1$ vector, ϵ_t is a scalar irregular component whose distribution is a normal with mean of zero and a scalar variance, h_t . The other equation required in the Kalman filter is called the transition equation, which shows the stochastic behavior of the state vector, α_t . The transition equation is also linear as

$$\alpha_t = T_t \alpha_{t-1} + R_t \gamma_t, \quad t = 1, 2, \dots, T, \quad (3.3)$$

where T_t is $m \times m$ matrix, R_t is $m \times g$ matrix, γ_t is $g \times 1$ vector of irregular terms whose distribution is assumed to be a multivariate normal with mean of $g \times 1$ vector of zero and covariance of $g \times g$ matrix, Q_t . The matrices Z_t , T_t and R_t , and covariance matrices, h_t and Q_t may or may not change over time. If the matrices do not change over time, the model is called the time-invariant Kalman filter. Otherwise, it is called the time-variant Kalman filter. Equation (3.1) and (3.3) together is called the multivariate state space form of the Kalman filter, and (3.2) and (3.3) together is called the univariate state space form of the Kalman filter. By comparing (3.2) and (3.3), the state space form of the LLRW model (2.1) has the following identities:

$$\begin{cases} \alpha_t = \mu_t, z_t' = 1, T_t = 1, \\ R_t = 1, \epsilon_t = \epsilon_t, \gamma_t = \eta_t, \\ h_t = \text{variance of } \epsilon_t := \text{VarE}, \\ Q_t = \text{variance of } \eta_t := \text{VarN}. \end{cases} \quad (3.4)$$

The state space form of the LT model (2.4) has the following identities:

$$\left\{ \begin{array}{l} \alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}, z'_t = (1 \ 0), T_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \\ R_t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \epsilon_t = \epsilon_t, \gamma_t = \begin{pmatrix} \eta_t \\ \zeta_t \end{pmatrix}, \\ h_t = \text{variance of } \epsilon_t := \text{VarE}, \\ Q_t = \text{variance of } \begin{pmatrix} \eta_t \\ \zeta_t \end{pmatrix} := \begin{pmatrix} \text{VarN} & 0 \\ 0 & \text{VarK} \end{pmatrix}. \end{array} \right. \quad (3.5)$$

The state space form of the FT model (2.5) has the following identities:

$$\left\{ \begin{array}{l} \alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}, z'_t = (1 \ 0), T_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \\ R_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \epsilon_t = \epsilon_t, \gamma_t = \eta_t, \\ h_t = \text{variance of } \epsilon_t := \text{VarE}, \\ Q_t = \text{variance of } \eta_t := \text{VarN}. \end{array} \right. \quad (3.6)$$

The Kalman filter has three assumptions: A1) the initial state vector, α_0 has a mean of \mathbf{a}_0 and Covariance matrix of \mathbf{P}_0 . A2) the disturbances η_t and γ_t are independent of each other. This assumption could be relaxed. A3) the disturbances η_t and γ_t are independent with the initial state, α_0 .

Given the information of initial state variable, α_0 , the Kalman filter starts off the recursive algorithm. The recursive algorithm provides two estimates of the state variable α_t . First is the filtered estimate. It is the estimate of α_t given $Y_t = (y_t, \dots, y_1)$. Second is the forecast estimate. It is the estimate of α_t for $t = T + 1, T + 2, \dots$ given $Y_T = (y_T, \dots, y_1)$. The filtered estimates of the state variable α_t provide estimates of unobserved components in the structural time series models. These estimates are minimum mean square estimators (MMSE).

The information required to start off the Kalman filter is mean and covariance matrix of the initial state vector, α_0 , which are \mathbf{a}_0 and \mathbf{P}_0 , respectively. If the state vector α_t is a nonstationary, then a distribution of α_0 is given as a diffuse prior whose mean is given as the first data in the series and covariance matrix is given as kI where k is a very large scalar and I is an $m \times m$ identity matrix where m is the dimension of the state vector, α_t . For the LLRW model, $\alpha_t = \mu_t$ and it is a nonstationary and $m = 1$. Therefore, \mathbf{a}_0 is set as y_1 and \mathbf{P}_0 is set as $1.2e + 10$. For the LT model, $\alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}$ is also a nonstationary and $m = 2$. Therefore, \mathbf{a}_0 is set as $\begin{pmatrix} y_1 \\ 0 \end{pmatrix}$ and \mathbf{P}_0 is set as $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ where $k = 1.2e + 10$. For the FT model, $\alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}$ is also a nonstationary and $m = 2$ and slope β_t is a deterministic. Therefore, \mathbf{a}_0 is set as $\begin{pmatrix} y_1 \\ \beta \end{pmatrix}$ and \mathbf{P}_0 is set as $\begin{pmatrix} k & 0 \\ 0 & 0 \end{pmatrix}$, where $k = 1.2e + 10$ and the value of β is the estimate of slope obtained by doing a simple linear regression of yearly data on time.

In the state space form in (3.1) and (3.3) for a multivariate time series or in (3.2) and (3.3) for a univariate time series, there are some unknown parameters in matrices, Z_t , T_t , h_t and Q_t . Before running the recursive algorithm, these unknown parameters should be estimated. From (3.4), (3.5) and (3.6), unknown parameters of LLRW, LT and FT models are $\psi_{LLRW} = (\text{VarN}, \text{VarE})$, $\psi_{LT} = (\text{VarN},$

VarK, VarE) and $\psi_{FT} = (\text{VarN}, \text{VarE})$, respectively. If the disturbances of ϵ_t and η_t are normally distributed, the likelihood function of the observations could be obtained from the Kalman filter via the prediction error decomposition (Harvey and Peters, 1990). Unknown parameters are estimated by maximizing the likelihood function with respect to the unknown parameters.

To analyze our data by the Kalman filter, a R function, `StructTS` in the R package, `stats` is used. The function `StructTS`, first, estimates unknown parameters in matrices, Z_t , T_t , h_t and Q_t , and then given the estimates of unknown parameters, provides predictions and filtered estimations for univariate time series based on the state space form of (3.2) and (3.3). To estimate unknown parameters in matrices, Z_t , T_t , h_t and Q_t , `StructTS` calls a R function, `optim` in the R package, `stats`. For the method for optimization in `optim`, L-BFGS-B is used. L-BFGS-B is described by Byrd *et al.* (1995) which allows box constraints, which means that each variable can be given a lower and/or upper bound. For our models, since unknown parameters are variances of irregular terms, the lower bound is set as 0 and the upper bound is set as infinity. The initial values to start the function `optim` are any nonnegative numbers. This method uses a limited-memory modification of the BFGS, which is a quasi-Newton method. Outputs of `StructTS` are estimates of unknown parameters, loglikelihood, standardized residuals, and filtered estimates of state vector α_t . The standardized residuals are used for the diagnostics of fitted models. From the filtered estimates of α_t , estimates of unobserved components of LLRW, LT and FT are obtained.

4. Analysis of data

4.1. Death counts data

Yearly death counts data of 8 countries analyzed in this paper are extracted from the Human Mortality Database (HMD). HMD provides several data such as death counts, census counts, birth counts, and population estimates for calculations of death rates and life tables. The main goal of the HMD is to document the longevity revolution of the modern era and to facilitate research in its causes and consequences. HMD includes relatively wealthy and highly industrialized countries since it is based on design to populations where death registration and census data are virtually complete. In this paper, death counts data of 8 countries are analyzed: Two countries in North America (U.S. and Canada), four countries in Europe (U.K., France, Italy, and Spain), two countries in Asia-Pacific (Taiwan and Australia). From death counts from HMD, three death counts data are generated for each country: Female death counts, Male death counts, and Total death counts. The annual periods of data for countries are for U.S. (1946–2017), Canada (1946–2016), U.K. (1946–2016), France (1946–2017), Spain (1946–2018), Italy (1946–2017), Australia (1946–2018), and Taiwan (1970–2014). HMD keeps death counts data for Taiwan only from the year 1970. This paper uses data after World War II (1939–1945) since there are serious outliers during World War II, especially for data of countries in Europe.

4.2. Diagnostics and Best fitted model selection

All three models, LLRW, FT, and LT are fitted to three types of death counts (female, male and total) of all 8 countries. For each model fitted, two diagnostic procedures based on residuals are applied to see the validity of the fitted model. The residuals in the Kalman filter are obtained from one-step-ahead-prediction errors (also called innovations). One-step-ahead-prediction errors are obtained by the Kalman filter with the estimated values of unknown parameters of the fitted model. Standardized residuals are obtained from one-step-ahead-prediction errors divided by the standard deviation of one-step-ahead-prediction errors. If a fitted model is valid, then the standardized residuals are independent and identically distributed by a normal distribution (Harvey, 1989). Thus, the first diagnostic

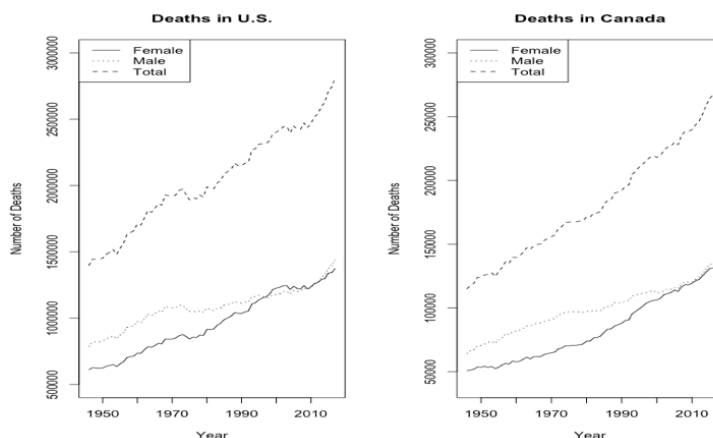


Figure 1: Death counts for female, male and total of U.S. and Canada.

Table 1: p -values of Shapiro-Wilk tests for death counts of U.S. and Canada

| | U.S. | | | Canada | | |
|--------|-------|-------|-------|--------|-------|-------|
| | LLRW | FT | LT | LLRW | FT | LT |
| Female | 0.936 | 0.666 | 0.717 | 0.724 | 0.610 | 0.723 |
| Male | 0.400 | 0.585 | 0.659 | 0.578 | 0.406 | 0.198 |
| Total | 0.930 | 0.956 | 0.993 | 0.867 | 0.916 | 0.603 |

procedure is to check the normality of the standardized residuals and the second diagnostic procedure is to check the independence of the standardized residuals. Shapiro-Wilk test is used to check the normality, and normal QQ plots and density plots are used to confirm the conclusions of Shapiro-Wilk test. Run test is used to check the independence of the standardized residuals. Models that pass two diagnostics procedures are treated as the valid ones. Once the valid models are identified for each type of data after the diagnostic procedures, a best-fitted valid model is selected based on AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and SSPE (Sum of Square of one-step-ahead-prediction errors). If these three criteria do not recommend a model unanimously, then the most recommended one is selected as the best-fitted valid model.

4.3. North American countries, U.S. and Canada

Figure 1 shows plots of death counts of female, male, and total for U.S. and Canada. Both U.S. and Canadian data show a similar linear trend for death counts of female, male and total. Also, both countries' data show that death counts for males are larger than those of females up to the year 2000. Since then, death counts of females and males are close to each other. Since both countries' data show a clear linear trend, it is sensible to fit the LT and FT model. The LT model assumes that the slope of linear trend is stochastic, and the FT model assumes that it is deterministic. Outcomes of the LLRW model are also provided to compare with them. Table 1 shows p -values of Shapiro-Wilk tests for testing normality of the standardized residuals of three models fitted to three types of death counts in U.S. and Canada. The p -values of two countries show that none of the three models failed for normality on the level of 0.01. Figure 2 and 3 show normal QQ plots for U.S. and Canada, respectively. These plots confirm the results of Shapiro-Wilk tests. Table 2 shows the p -values of Run tests for independence of standardized residuals. The p -values of two countries show that none of the

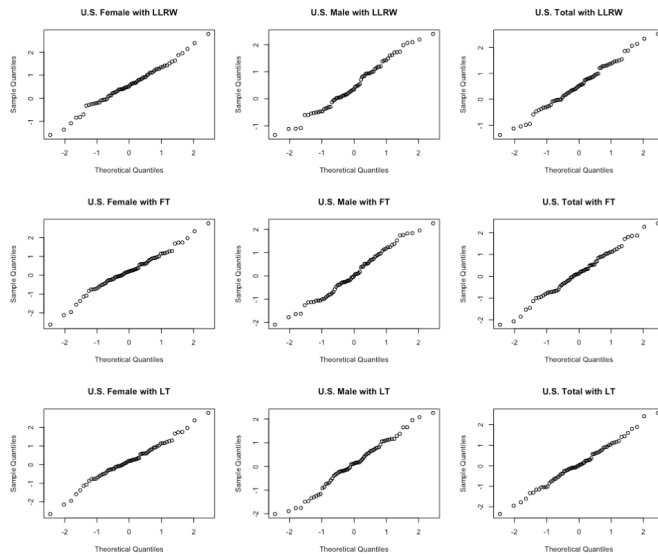


Figure 2: QQ plots of U.S.

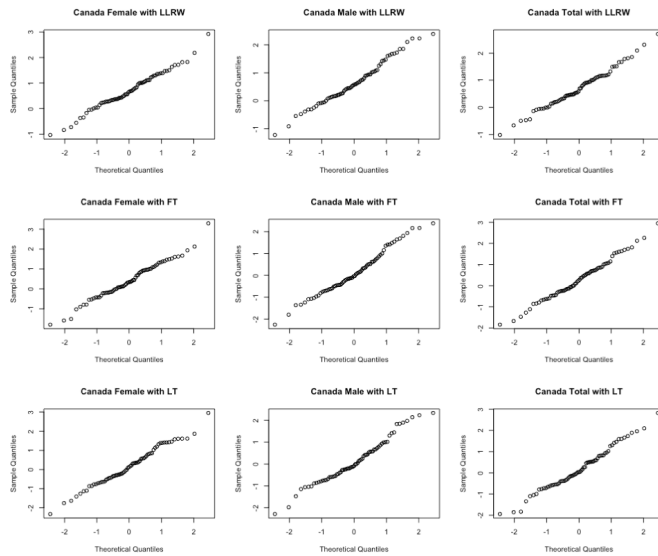


Figure 3: QQ plots of Canada.

three models failed for independence on the level of 0.01. Thus, for two countries, none of the three models failed to be valid. Table 3 presents the best model for each type of data based on AIC, BIC and SSPE for the valid models of U.S. and Canada. If a model is not selected unanimously by all three criteria, then the most commonly recommended one is selected as the best one. For example, for total deaths of U.S., the FT is the selected as the best one. Thus, for the U.S., the FT models are the best for both female and total deaths and the LT model is the best for male deaths. For Canada, the LT models are the best for both female and total deaths and the FT model is the best for male deaths. Table 4

Table 2: p -values of Run tests for death counts of U.S. and Canada

| | U.S. | | | Canada | | |
|--------|-------|-------|-------|--------|-------|-------|
| | LLRW | FT | LT | LLRW | FT | LT |
| Female | 0.042 | 0.549 | 0.631 | 0.809 | 0.336 | 0.543 |
| Male | 0.281 | 0.281 | 0.998 | 0.471 | 0.471 | 0.395 |
| Total | 0.281 | 0.904 | 0.998 | 0.809 | 0.809 | 0.543 |

Table 3: Best models for death counts of U.S. and Canada

| | U.S. | | | Canada | | |
|--------|------|-----|------|--------|-----|------|
| | AIS | BIS | SSPE | AIS | BIS | SSPE |
| Female | FT | FT | FT | LT | LT | LT |
| Male | LT | LT | LT | FT | FT | LT |
| Total | FT | FT | LT | LT | FT | LT |

Table 4: Estimates of parameters for death counts of U.S. and Canada

| | U.S. | | | Canada | | | |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | VarN | VarK | VarE | VarN | VarK | VarE | |
| Female by FT | 1.922503E+08 | NA | 3.115608E+07 | Female by LT | 5.618963E+05 | 1.730798E+04 | 3.000166E+05 |
| Male by LT | 5.625876E+07 | 1.459582E+07 | 7.079134E+07 | Male by FT | 1.437916E+06 | NA | 4.166670E+04 |
| Total by FT | 7.584381E+08 | NA | 1.080131E+08 | Total by LT | 3.156514E+06 | 3.567977E+04 | 7.245420E+05 |

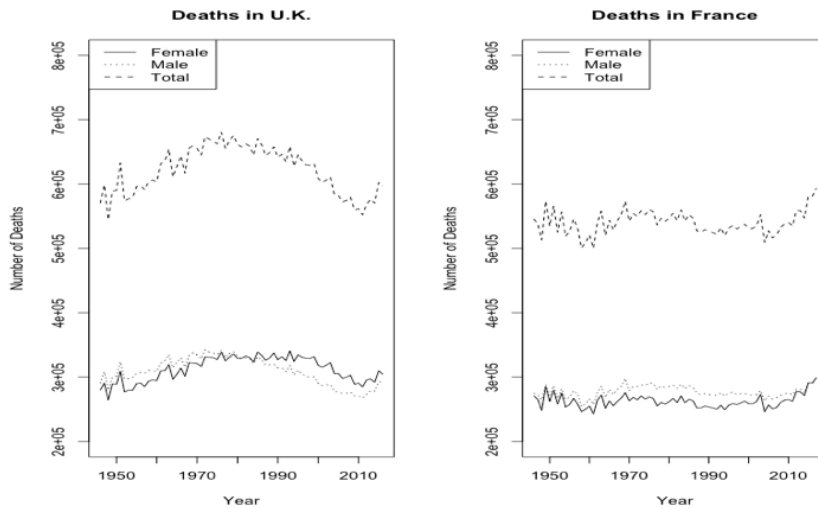


Figure 4: Death counts for female, male and total of U.K. and France

presents the estimated values of unknown parameters of the best model for each type of deaths in U.S. and Canada. It is noted from Section 3 that unknown parameters of LLRW, LT and FT model are, respectively, $\psi_{LLRW} = (\text{VarN}, \text{VarE})$, $\psi_{LT} = (\text{VarN}, \text{VarK}, \text{VarE})$ and $\psi_{FT} = (\text{VarN}, \text{VarE})$. Thus, for FT and LLRW models, the parameter estimate for VarK is NA.

4.4. European countries, U.K. and France

Figure 4 and 7 show plots of death counts of female, male and total for U.K. and France, and Italy and Spain, respectively. Deaths of European countries do not show a clear linear trend that deaths of North American countries show. Table 5 shows p -values of the Shapiro-Wilk tests for testing normality of

Table 5: p -values of Shapiro-Wilk tests for death counts of U.K. and France

| | U.K. | | | France | | |
|--------|-------|-------|-------|--------|-------|-------|
| | LLRW | FT | LT | LLRW | FT | LT |
| Female | 0.999 | 0.995 | 0.869 | 0.819 | 0.553 | 0.489 |
| Male | 0.829 | 0.902 | 0.345 | 0.377 | 0.251 | 0.306 |
| Total | 0.996 | 0.969 | 0.807 | 0.802 | 0.448 | 0.314 |

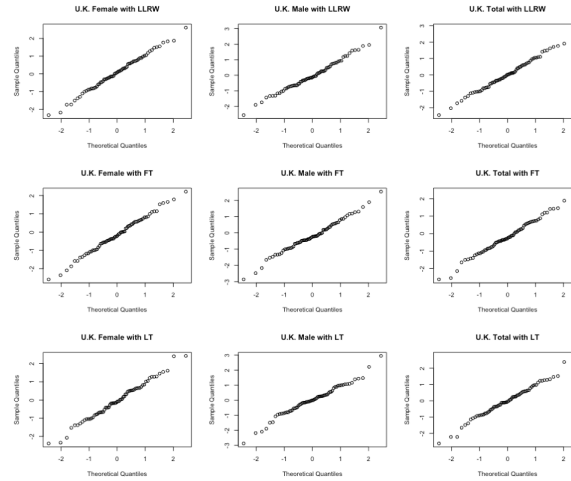


Figure 5: QQ plots of U.K.

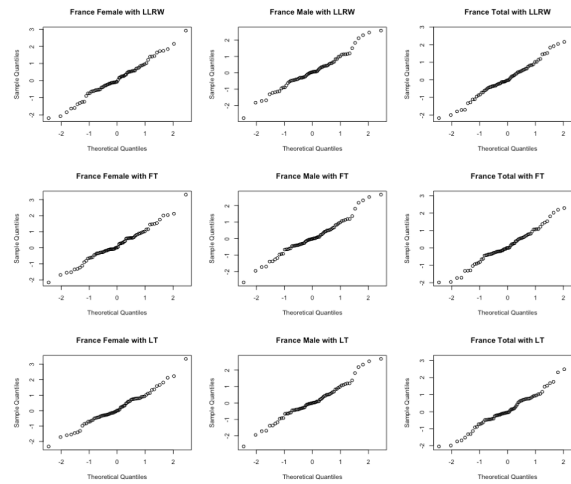


Figure 6: QQ plots of France.

the standardized residuals of models fitted to three types of death counts in U.K. and France. The p -values of two countries show that none of the three models failed for normality on the level of 0.01. Figure 5 and 6 show normal QQ plots of U.K. and France, respectively. These plots confirm the results of Shapiro-Wilk tests. Table 6 shows the p -values of Run tests for independence of standardized residuals. The p -values of two countries show that none of the three models failed for independence on the level of 0.01. Thus, for two countries, none of the three models failed to be

Table 6: p -values of Run tests for death counts of U.K. and France

| | U.K. | | | France | | |
|--------|-------|-------|-------|--------|-------|-------|
| | LLRW | FT | LT | LLRW | FT | LT |
| Female | 0.471 | 0.471 | 0.902 | 0.719 | 0.719 | 0.631 |
| Male | 0.092 | 0.809 | 0.902 | 0.073 | 0.073 | 0.054 |
| Total | 0.809 | 0.809 | 0.543 | 0.401 | 0.188 | 0.631 |

Table 7: Best models for death counts of U.K. and France

| | U.K. | | | France | | |
|--------|------|-----|------|--------|------|------|
| | AIS | BIS | SSPE | AIS | BIS | SSPE |
| Female | LT | LT | LLRW | LLRW | LLRW | LLRW |
| Male | LT | LT | LLRW | LLRW | LLRW | LLRW |
| Total | LT | LT | LLRW | LLRW | LLRW | LLRW |

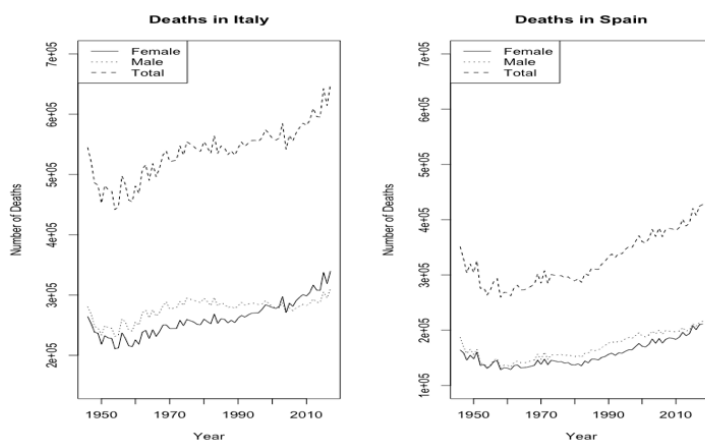


Figure 7: Death counts for female, male and total of Italy and Spain.

Table 8: Estimates of parameters for death counts of U.K. and France

| | U.K. | | | France | | | |
|--------------|--------------|--------------|--------------|----------------|--------------|------|--------------|
| | VarN | VarK | VarE | VarN | VarK | VarE | |
| Female by LT | 6.732207E+06 | 2.012691E+05 | 4.868096E+07 | Female by LLRW | 1.462518E+07 | NA | 4.612672E+07 |
| Male by LT | 0.000000E+00 | 5.816742E+05 | 3.787161E+07 | Male by LLRW | 9.614802E+06 | NA | 2.995601E+07 |
| Total by LT | 0.000000E+00 | 2.140323E+06 | 1.745933E+08 | Total by LLRW | 4.473024E+07 | NA | 1.475258E+08 |

valid. Table 7 presents the best model for each type of data based on AIC, BIC and SSPE for the valid models for each type for the U.K. and France. As North American countries, if a model is not selected by unanimously, then the most commonly recommended one is selected. For the U.K., the LT models are the best for all three types and for France, the LLRW models are the best for all three types. Table 8 shows estimates of unknown parameters of the best model for each type of deaths in U.K. and France. As Table 4, for the LLRW models, parameter estimates for VarK are NAs.

4.5. European countries, Italy and Spain

Table 9 shows p -values of Shapiro -Wilk tests for testing normality of the standardized residuals of models fitted to three types of death counts in Italy and Spain. The p -values show that for Italy, none of models failed for normality on the level of 0.01, and for Spain, however, the LLRW models for

Table 9: p -values of Shapiro-Wilk tests for death counts of Italy and Spain

| | Italy | | | Spain | | |
|--------|-------|-------|-------|--------|-------|-------|
| | LLRW | FT | LT | LLRW | FT | LT |
| Female | 0.689 | 0.303 | 0.168 | 0.019 | 0.124 | 0.503 |
| Male | 0.453 | 0.089 | 0.129 | 0.000* | 0.021 | 0.119 |
| Total | 0.519 | 0.215 | 0.098 | 0.002* | 0.066 | 0.278 |

*Significance with 0.01

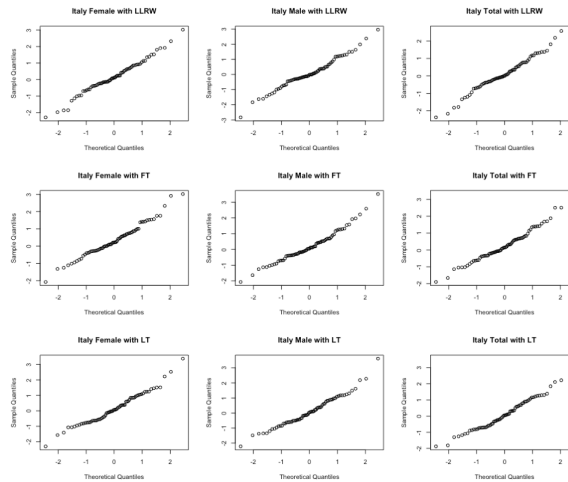


Figure 8: QQ plots of Italy.

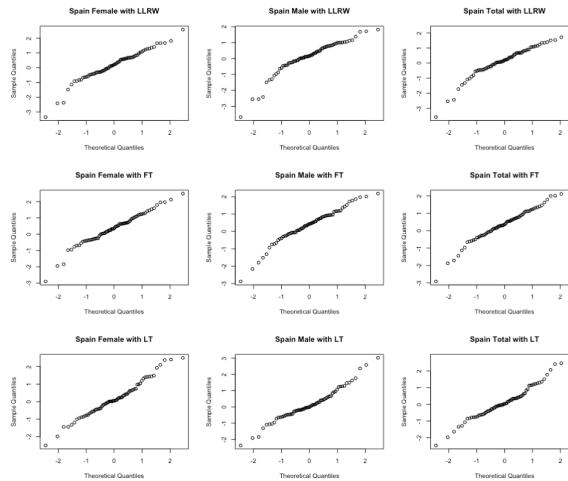


Figure 9: QQ plots of Spain.

both male and total deaths failed for normality on the level of 0.01. Figure 8 and 9 show the normal QQ plots for Italy and Spain, respectively. For Spain, normal QQ plots of both male and total deaths of the LLRW model show clearly the non-normality (left skewed shape) of data which confirms the results of Shapiro-Wilk tests. Figure 10 shows the normal QQ plots of both male LLRW and total LLRW model of Spain again and corresponding density plots for standardized residuals. residuals.

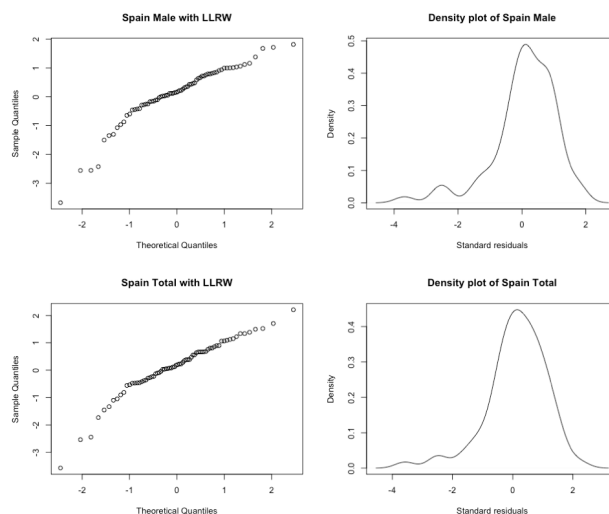


Figure 10: QQ plots and Density plots for male and total of Spain.

Table 10: *p*-values of Run tests for death counts of Italy and Spain

| | Italy | | | Spain | | |
|--------|-------|-------|-------|-------|-------|-------|
| | LLRW | FT | LT | LLRW | FT | LT |
| Female | 0.402 | 0.188 | 0.336 | 0.812 | 0.476 | 0.719 |
| Male | 0.719 | 0.719 | 0.631 | 0.476 | 0.476 | 0.719 |
| Total | 0.402 | 0.188 | 0.336 | 0.476 | 0.235 | 0.719 |

*Significance with 0.01

Table 11: Best models for death counts of Italy and Spain

| | Italy | | | Spain | | |
|--------|-------|------|------|-------|-----|------|
| | AIS | BIS | SSPE | AIS | BIS | SSPE |
| Female | LT | LT | LLRW | LT | LT | LLRW |
| Male | LLRW | LLRW | LLRW | LT | LT | LLRW |
| Total | LT | LLRW | LLRW | LT | LT | LLRW |

Table 12: Estimates of parameters for death counts of Italy and Spain

| | Italy | | | Spain | | | |
|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | VarN | VarK | VarE | VarN | VarK | VarE | |
| Female by LT | 0.000000E+00 | 7.834581E+05 | 4.880840E+07 | Female by LT | 0.000000E+00 | 2.607934E+05 | 1.998147E+07 |
| Male by LLRW | 3.068192E+07 | NA | 2.589466E+07 | Male by LT | 0.000000E+00 | 5.494487E+05 | 2.062101E+07 |
| Total by LLRW | 1.113686E+08 | NA | 1.315861E+08 | Total by LT | 0.000000E+00 | 1.511941E+06 | 7.905366E+07 |

Density plots show clearly the left skewed shape of densities. Table 10 shows the *p*-values of Run tests for Italy and Spain. The *p*-values of two countries show that none of the three models failed for independence on the level of 0.01. Thus, for countries, all models except male LLRW of Spain and total LLRW of Spain, pass both diagnostic procedures and thus are valid ones. Table 11 presents the best model for each type of data based on AIC, BIC and SSPE for the valid models for Italy and Spain. For Italy, the LT model is the best for female and the LLRW models are the best for both male and total deaths, and for Spain, the LT models are the best for all three types. Table 12 shows estimates of unknown parameters of the best-fitted model for each type of deaths in Italy and Spain.

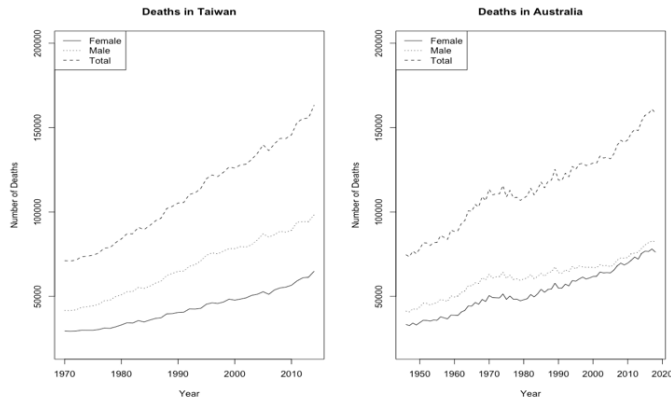


Figure 11: Death counts for female, male and total of Taiwan and Australia.

Table 13: p -values of Shapiro-Wilk tests for death counts of Taiwan and Australia

| | Taiwan | | | Australia | | |
|--------|--------|-------|-------|-----------|-------|-------|
| | LLRW | FT | LT | LLRW | FT | LT |
| Female | 0.623 | 0.570 | 0.939 | 0.243 | 0.243 | 0.358 |
| Male | 0.623 | 0.862 | 0.916 | 0.265 | 0.187 | 0.483 |
| Total | 0.450 | 0.655 | 0.938 | 0.167 | 0.189 | 0.198 |

*Significance with 0.01

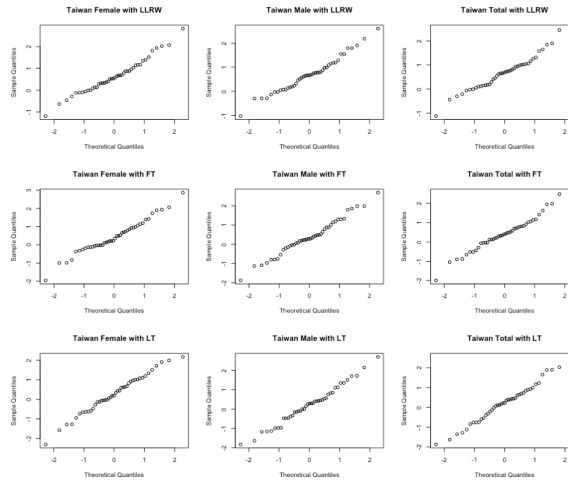


Figure 12: QQ plots of Taiwan.

4.6. Asia-Pacific countries, Taiwan and Australia

Figure 11 shows plots of death counts of female, male and total deaths for Taiwan and Australia. Deaths of these two countries show a clear linear trend similar to deaths of the two North American countries mentioned above. Table 13 shows p -values of Shapiro tests for testing normality of the standardized residuals of models fitted to three types of death counts in Taiwan and Australia. The p -values show that none of models for all three types of deaths in Taiwan and Australia failed for normality on the level of 0.01. Figure 12 and 13 show the normal QQ plots for Taiwan and Australia,

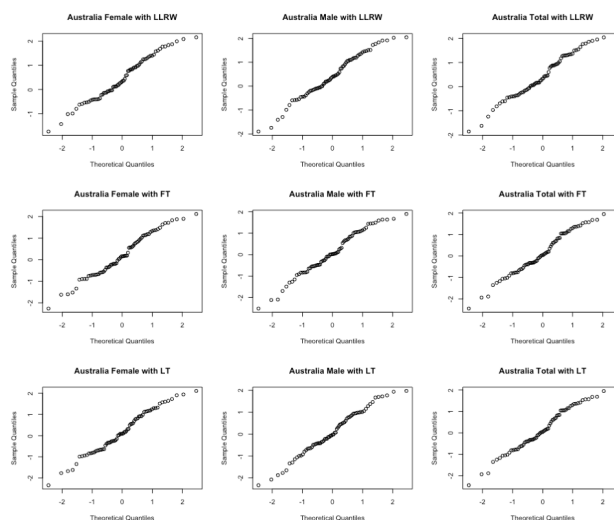


Figure 13: QQ plots of Australia.

Table 14: *p*-values of Run tests for death counts of Taiwan and Australia

| | Taiwan | | | Australia | | |
|--------|--------|-------|-------|-----------|-------|-------|
| | LLRW | FT | LT | LLRW | FT | LT |
| Female | 0.361 | 0.127 | 0.278 | 0.018 | 0.018 | 0.402 |
| Male | 0.127 | 0.761 | 0.442 | 0.342 | 0.998 | 0.719 |
| Total | 0.761 | 0.361 | 0.641 | 0.154 | 0.342 | 0.402 |

*Significance with 0.01

Table 15: Best models for death counts of Taiwan and Australia

| | Taiwan | | | Australia | | |
|--------|--------|-----|------|-----------|-----|------|
| | AIS | BIS | SSPE | AIS | BIS | SSPE |
| Female | LT | LT | FT | FT | FT | LT |
| Male | LT | FT | FT | FT | FT | LLRW |
| Total | LT | LT | LT | FT | FT | LLRW |

Table 16: Estimates of parameters for death counts of Taiwan and Australia

| | Taiwan | | | Australia | | | |
|--------------|--------------|--------------|--------------|--------------|--------------|------|--------------|
| | VarN | VarK | VarE | VarN | VarK | VarE | |
| Female by LT | 0.000000E+00 | 2.669466E+04 | 3.679238E+05 | Female by FT | 7.231256E+05 | NA | 7.907837E+05 |
| Male by FT | 1.282648E+06 | NA | 3.192780E+05 | Male by FT | 9.721123E+05 | NA | 7.088836E+05 |
| Total by LT | 1.765875E+05 | 1.316371E+05 | 2.106885E+06 | Total by FT | 3.093034E+06 | NA | 2.932334E+06 |

respectively. These plots confirm the results of Shapiro-Wilk tests. Table 14 shows the *p*-values of Run tests for Taiwan and Australia. The *p*-values of two countries show that none of the three models failed for independence on the level of 0.01. Thus, for two countries, none of the three models failed to be valid. Table 15 presents the best model for each type of data based on AIC, BIC and SSPE for the valid models for Taiwan and Australia. For Taiwan, the LT models are the best for both female and total deaths and the FT model is the best for male deaths, and for Australia, the FT models are the best for all three types of deaths. Table 16 shows estimates of unknown parameters of the best-fitted model for each type of deaths in Taiwan and Australia.

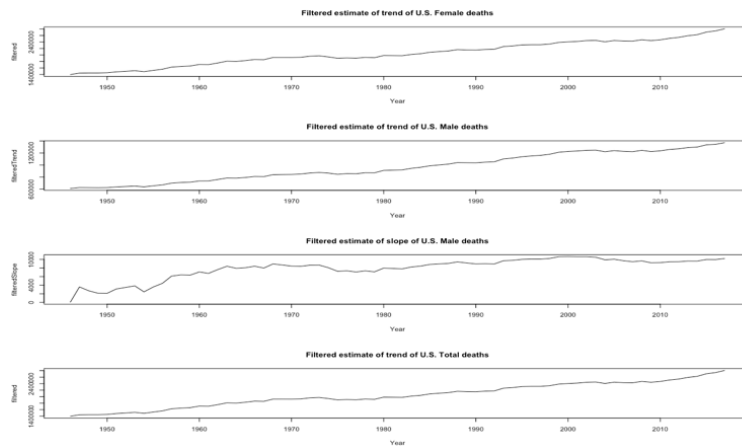


Figure 14: Filtered estimates of unobserved components for U.S. female, male and total Deaths.

Table 17: Signal to Noise Ratio for U.S. and Canada

| | U.S. | | Canada | | |
|--------------|-----------|-----------|--------------|-----------|-------|
| | VarN/VarE | VarK/VarE | VarN/VarE | VarK/VarE | |
| Female by FT | 6.171 | NA | Female by LT | 1.873 | 0.058 |
| Male by LT | 0.795 | 0.206 | Male by FT | 34.510 | NA |
| Total by FT | 7.022 | NA | Total by LT | 4.357 | 0.049 |

4.7. Signal extraction for unobserved components

One advantage to fit a structural time series model using the Kalman filter is that unobserved components of the time series can be estimated. Both the LLRW and the FT model have an unobserved trend component, μ_t . The LT model has two unobserved components, trend and slope, μ_t and β_t , respectively. The Kalman filter provides filtered estimates for unobserved components. For U.S. deaths, it is noted from Table 3 that the best-fitted model for female, male and total deaths is the FT, the LT and the FT model, respectively. Figure 14 shows filtered estimates of μ_t of U.S. female deaths, filtered estimates of μ_t and β_t of U.S. male deaths, and filtered estimates of μ_t of U.S. total deaths.

5. Conclusions

This paper analyzes female, male and total death counts of 8 countries in several regions in the world using three structural time series models with the Kalman filter. Three structural models are a local level with a random walk (LLRW) model, a fixed local linear trend (FT) model and a local linear trend (LT) model. LLRW model implies that the level of data moves stochastically based on a random walk model. Thus, this model is good to fit data without a clear linear pattern. FT model implies that the level of data moves stochastically with a deterministic slope. Thus, this model is good to fit data with a clear linear pattern. LT model implies that the level of data moves stochastically with a stochastic slope. That is, both level and slope move stochastically and thus this model is the most flexible. Thus, this model is good to fit data both with and without a clear linear pattern. Death counts of all three types of deaths in two North American countries and two Asia-Pacific countries show similar linear trends. Best fitted stochastic models for both North American countries and Asia-Pacific countries are either FT or LT model. Table 17 shows the signal to noise ratios for trend (VarN/VarE) and those of slope (VarK/VarE) of the best-fitted stochastic models for two North American countries. For the

Table 18: Signal to Noise Ratio for Taiwan and Australia

| | Taiwan | | Australia | | |
|--------------|-----------|-----------|---------------|-----------|----|
| | VarN/VarE | VarK/VarE | VarN/VarE | VarK/VarE | |
| Female by LT | 0.00 | 0.07 | Female by FT. | 0.91 | NA |
| Male by FT | 4.02 | 0.04 | Male by FT | 1.37 | NA |
| Total by LT | 0.08 | 0.06 | Total by FT | 1.05 | NA |

Table 19: Signal to Noise Ratio for U.K. and France

| | U.K. | | France | | |
|--------------|-----------|-----------|----------------|-----------|----|
| | VarN/VarE | VarK/VarE | VarN/VarE | VarK/VarE | |
| Female by LT | 0.14 | 0.00 | Female by LLRW | 0.32 | NA |
| Male by LT | 0.00 | 0.02 | Male by LLRW | 0.32 | NA |
| Total by LT | 0.00 | 0.01 | Total by LLRW | 0.30 | NA |

Table 20: Signal to Noise Ratio for Italy and Spain

| | Italy | | Spain | | |
|---------------|-----------|-----------|--------------|-----------|------|
| | VarN/VarE | VarK/VarE | VarN/VarE | VarK/VarE | |
| Female by LT | 0.00 | 0.02 | Female by LT | 0.00 | 0.01 |
| Male by LLRW | 1.18 | NA | Male by LT | 0.00 | 0.03 |
| Total by LLRW | 0.85 | NA | Total by LT | 0.00 | 0.02 |

U.S., female deaths show higher signal to noise ratio for trend than that of male deaths. For Canada, however, female deaths show lower signal to noise ratio for trend than that of male deaths. Signal to noise ratios of slope are small for both countries. Note that NA in the table is for the FT model where VarK is not presented. Table 18 shows the signal to noise ratios for two Asia-Pacific countries. It shows that two countries do not have high signal to noise ratios for trend. Male deaths of both countries have higher signal to noise ratios than those of female death. Taiwan has low signal to noise ratios of slope as well. European countries do not show any clear linear trend. Best fitted stochastic models for European countries are either LLRW or LT. Table 19 and 20 show that European countries have low signal to noise ratios for both trend and slope.

Table 4, 8, 12 and 16 show the estimates of values of variances of irregular components, VarN, VarK and VarE. These variances show how much each component in the structural model moves stochastically up and down. For example, value of VarN shows how much the level of the trend, μ_t moves stochastically up and down over time, value of VarK shows how much the slope, β_t moves stochastically up and down over time and value of VarE shows how much the observation, y_t moves stochastically up and down around the unobserved trend, μ_t over time. Thus, the large variances do not imply that data has the large magnitudes of the trend and the slope, but imply that data has the large stochastic movements of the trend and the slope. Thus, large values of VarN and VarK of our data imply that our data has the trend and the slope which move a lot stochastically. To see the point, it is noted that for Australia, the best models for both female deaths and total deaths are the FT and from Table 18, ratios of VarN/VarE are similar each other, but from Table 16, VarN of total deaths is a lot larger than that of Female. Thus it implies that the level of the trend for total deaths moves stochastically more than that of female deaths. Figure 15 shows the estimates of the trend for both female deaths and total deaths from the Kalman filter. The plots of two estimated trends show that the trend of total deaths shows more volatility than that of female. Standard deviation of the trend of total deaths is 23064.75 and that of female deaths is 12681.7.

In Table 18 through 20, value of signal to noise ratio for trend (VarN/VarE) of LT model is zero for female model of Taiwan, male and total model of U.K., female and total model of Italy, and all

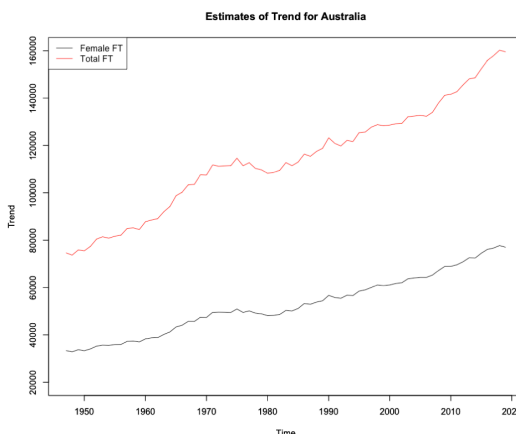


Figure 15: Estimates of trend component for female and total deaths of Australia.

models of Spain. For these cases, LT model (2.2) is adjusted as

$$\begin{cases} y_t = \mu_t + \epsilon_t, \\ \mu_t = \mu_{t-1} + \beta_{t-1}, \\ \beta_t = \beta_{t-1} + \zeta_t. \end{cases} \quad (5.1)$$

This adjusted model implies that the level of the trend at time t moves stochastically based on only the level of trend and slope at time $t - 1$ without an irregular component of the trend. In this case, the stochastic movement of the level of the trend is purely based on the stochastic movement of slope, VarK.

References

- Asemota OJ, Bamanga MA, and Alaribe OJ (2016). Modelling seasonal behavior of rainfall in north-east Nigeria. A State Space Approach, *International Journal of Statistics and Applications*, **6**, 203–222.
- Box GEP and Jenkins GM (1976). *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.
- Byrd RH, Lu P, Nocedal J, and Zhu C (1995). A limited memory algorithm for bound constrained optimization, *SIAM Journal on Scientific Computing*, **16**, 1190—1208.
- Chukhrova N and Johannssen A (2017). State space models and the Kalman-Filter in stochastic claims reserving: Forecasting, Filtering and Smoothing, *Risks*, **5**, 30.
- Durbin J and Koopman SJ (2012). *Time Series Analysis by State Space Methods* (2nd Ed), Oxford University Press, Oxford.
- Gove JH and Houston DR (1996). Monitoring the growth of American beech affected by beech bark disease in Maine using Kalman filter, *Environmental and Ecological Statistics* **3**, 167–187.
- Harvey AC (1981). *Time Series Models*, Deddington, Oxford University Press, Oxford.
- Harvey AC and Todd PHJ (1983). Forecasting economic time series with structural and Box-Jenkins models, (with discussion), *Journal of Business and Economic Statistics*, **1**, 299–315.
- Harvey AC (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.

- Harvey AC and Peters S (1990). Estimation procedures for structural time series models, *Journal of Forecasting*, **9**, 89–108.
- Hannan EJ and Deistler M (1988). *The Statistical Theory of Linear System*, Wiley, New York.
- Harrison PJ and Stevens CF (1971). A Bayesian approach to short-Term forecasting, *Operations Research Quarterly*, **22**, 341–362.
- Kalman RE (1960). A new approach to linear filtering and prediction problems, *Journal of Basic Engineering*, **82**, 34–45.
- Nobrega JP and Oliveira ALI (2019). A sequential learning with Kalman filter and extreme learning machine for regression and time series forecasting, *Neurocomputing*, **337**, 235–250.
- Zulfi M, Hasan M, and Purnomo KD (2018). The development rainfall forecasting using Kalman filter, *Journal of Physics: Conf. Series*, **1008**, 012006.

Received November 4, 2021; Revised December 18, 2021; Accepted December 22, 2021