Sigma J Eng Nat Sci, Vol. 44, No. 3, pp. 1970–1992, June, 2026

Sigma Journal of Engineering and Natural Sciences
Web page info: https://sigma.yildiz.edu.tr
DOI: 10.14744/sigma.2026.2076

Research Article

A novel hybrid approach to forecast mortality using generalized estimation equations and autoregressive integrated moving average models

Öznur ÖZALTIN[1,*], Neslihan İYİT[2]

[1]*Department of Actuarial Science, Faculty of Science, Hacettepe University, Ankara, 06800, Türkiye*
[2]*Department of Statistics, Faculty of Science, Selcuk University, Konya, 42130, Türkiye*

ARTICLE INFO

Article history
Received: 08 February 2025
Revised: 05 May 2025
Accepted: 10 July 2025

Keywords:
Autoregressive Integrated Moving Average; Generalized Estimating Equations; Longitudinal Data; Mortality Forecasting; Quasi-Likelihood

ABSTRACT

Mortality forecasting is important for population policy and public health, particularly in the context of pandemics and aging populations. This study addresses this need by proposing a novel hybrid modeling framework that combines generalized estimating equations based on quasi-likelihood method and autoregressive integrated moving average models to predict mortality rates across Organization for Economic Cooperation and Development countries. The suggested method provides for parameter estimate in correlated panel data, notably longitudinal data, and precise temporal prediction of important predictors. The study, which examined the drivers of particular mortality, non-communicable diseases, accidents and other external causes of death, employed a large dataset covering 33 years from 1980 to 2013.
The prediction performance of the model was high with an average determination coefficient R^2 of 0.928 across countries, indicating an excellent model fit. The approach was shown to be reliable out of sample with forecasts produced for the years 2014 and 2023. The methodology offers a statistically interpretable alternative to black-box machine learning methods, and fills a gap in the literature where generalized estimating equations and autoregressive integrated moving average models have been utilized mostly in isolation.
In this research we combine the two methodologies to provide a clear and powerful method for predicting mortality rates. It may be a useful tool for governments, insurance companies, and healthcare planners to prepare for demographic shift. This can allow people to make informed decisions and prepare for what lies ahead.

Cite this article as: Özaltın Ö, İyit N. A novel hybrid approach to forecast mortality using generalized estimation equations and autoregressive integrated moving average models. Sigma J Eng Nat Sci 2026;44(3):1970–1992.

***Corresponding author.**
*E-mail address: oznurozaltin@hacettepe.edu.tr

This paper was recommended for publication in revised form by Editor-in-Chief Ahmet Selim Dalkilic

Published by Yıldız Technical University Press, İstanbul, Turkey

INTRODUCTION

Mortality is a critical indicator of both demographic and public health dynamics. In particular, the SARS-CoV-2 pandemic [1,2], which causes severe respiratory complications, continues to increase mortality worldwide [3-7]. Accurate modeling and forecasting of mortality trends are vitally essential for informing health policy, social planning, and economic stability worldwide.

Different modeling techniques have been used to investigate mortality in many studies [8-23]. In recent years, artificial intelligence (AI) techniques, such as machine learning and deep learning, have been popular for modeling in many fields [5, 24]. However, statistical approaches such as generalized linear models (GLMs), Generalized Estimating Equations (GEEs) [25] and time-series models such as Autoregressive Integrated Moving Average (ARIMA) [26] are still widely trusted and used for their interpretability and suitability for interdisciplinary data [2, 8, 27- 40].

GEEs, proposed by Liang and Zeger [25, 41, 42], are increasingly popular in diverse fields such as public health, social sciences, econometrics, etc. [43-45]. One key advantage of GEEs is their robustness to violations of classical statistical model assumptions, such as "normality" and "homogeneity of variance". The GEEs method combines the flexibility of the GLMs method [46] with the "quasi-likelihood" estimation theory [47], whereas the GLM approach is based on the "maximum likelihood" (ML) theory [48], enabling efficient modeling of within-cluster correlations through a "working correlation matrix" [25, 8, 43, 46, 49-58].

While GEEs and ARIMA models have each been widely applied to model and forecast mortality [59-74], studies integrating both within a single hybrid framework remain scarce, particularly for mortality forecasting across countries. Moreover, although machine learning methods are powerful, they often suffer from a lack of interpretability. This study addresses that gap by proposing an ARIMA integrated GEEs hybrid model that leverages the strengths of both methods for panel mortality data prediction in the Organization for Economic Co-operation and Development (OECD) countries.

In this study, we employ symmetric and asymmetric continuous distributions (e.g., normal, gamma, and inverse Gaussian) within the GEE framework, along with different link functions. Model selection is guided by the quasi-likelihood-independent model criterion (QIC) and its corrected version (QICC). For the aim of forecasting, ARIMA models are used to project significant major causes of mortality between 2014 and 2023. These projections are then integrated into the GEEs framework, establishing a novel hybrid strategy for the mortality estimation of the OECD countries.

Recent applications of GEEs and ARIMA models in a variety of domains, such as traffic safety [61], child obesity [64], energy forecasting [66], and clinical trials [63], demonstrate their versatility. However, few of these studies specifically address mortality or the use of GEEs and ARIMA methods for the aim of forecasting separately. This emphasizes the novelty and practical value of the proposed ARIMA integrated GEE hybrid approach in this study. Details of these studies and further studies in the literature are presented as follows;

Iqbal et al. [59] proposed a novel GLM approach using the Poisson regression model based on homogeneously weighted moving average (HWMA) and double HWMA charting schemes to monitor count manufacturing processes. Kim and Ha [60] suggested a control chart for highly related explanatory factors and a binary response by employing Bayesian variable selection, principal component analysis, a neural network regression model, and a deep learning regression model. They expressed that the deep learning control chart outperforms the GLM with probit and logit link functions, and also neural network control charts in keeping tabs on both simulated and real binary response data.

Jamal et al. [61] used three years of Saudi Arabian rural highway crash data to recommend statistical monitoring for real-time highway safety surveillance using Poisson, negative binomial, and Conway–Maxwell–Poisson (CMP) count regression models. They determined the CMP model as the most suitable crash model, showing that road type and surface conditions greatly affect collisions. Foster et al. [62] extended the Tweedie GLM for the summation of a Poisson number of gamma random variables, including both a composite link and a covariate-dependent mean-variance link, and also applied their proposed model to Southeast Australian fish trawl survey data.

Fernández et al. [63] modeled the arthritis clinical trial dataset using the ordered stereotype model and the GEE approach as a comparison to fill a gap in ordinal response variable modeling. He et al. [64] examined obesity in Chinese children before and after the COVID-19 epidemic to help combat childhood obesity via the GEE approach and demonstrated that ethnic minorities, less daily physical activity, lower sleep duration, and longer screen time were positively linked with obesity. Finally, they expressed that the COVID-19 epidemic has increased childhood obesity, which may be due to weight-related changes.

Tang et al. [65] utilized machine learning and finance to investigate optimal financial investment matching and reduce investor information asymmetry. They determined that investor personalities affected model-predicted investment returns and transaction volumes using the ARIMA model through the particle swarm algorithm, and so they evaluated investor trading behavior. Phyo et al. [66] forecasted energy and meteorological data of four Jeju Island locations by comparing five ML regressors, an ARIMA model, and a Voting Regressor (VR) approach. They concluded that the VR approach performed effectively in energy forecasting. Martino et al. [67] introduced fuzzy transform-based seasonal forecasting and applied this novel method to the daily weather data of Naples, Italy, from 2003 to 2015 to predict mean, max., and min. Temperatures. Furthermore, they

applied this method to the Liguria weather station "Chiavari Caperana" daily mean temperature data. They showed that their solution outperforms the average seasonal variation, ARIMA, and fuzzy transformations in both situations.

Ye and Kerr [68], Makinde et al. [69], Ghoshal [70], Barbiellini Amidei et al. [71], Kuo et al. [72], Allain and Collins [73] investigated the relationships between mortality from liver cirrhosis and alcohol consumption in the United States of America (USA); daily incidence and mortality belonging to the COVID-19 pandemic in African countries, surgery volumes and mortality before and after the COVID-19 in Massachusetts, specific causes of mortality and the effect of COVID-19 in Veneto Italy, the ambient environment and COVID-19 cases in Lima, ethnic groups and park equity in Miami USA by the GEEs approach and ARIMA models. Also, Beard et al. [74] provided a detailed overview for comparing the GLM and GEEs approaches, and also ARIMA models with other models.

While these studies demonstrate the flexibility of the hybrid GEE-ARIMA models individually, our research distinguishes itself by combining these methods within a single coherent structure tailored to the longitudinal mortality analysis. The following point headers that summarize the key contributions of this study are as follows:

1. To apply the GEE approach to the longitudinal mortality data, benefiting from its robustness to distributional assumptions.
2. To incorporate various disease categories as explanatory variables to examine mortality patterns across the OECD countries.
3. To identify the best GEE model based on the smallest QIC value.
4. To forecast statistically significant independent variables (as major causes of mortality) using the ARIMA model for each OECD country.
5. To observe a general decrease in mortality of the OECD countries, except for certain disease categories such as nervous system and psychiatric disorders.
6. To propose an original hybrid forecasting approach based on the integration of the GEE and ARIMA models.

The organization of this study is as follows: The materials and methods part of this study are presented in the next section. Empirical results from this study, the discussion, and also conclusion are given in this study, respectively.

MATERIALS AND METHODS

This section details the collected dataset, the proposed architecture for mortality prediction, and the evaluation metrics used to compare model performance. We also present the configuration details for the GEE and ARIMA models. A flowchart of the study is shown in Figure 1.

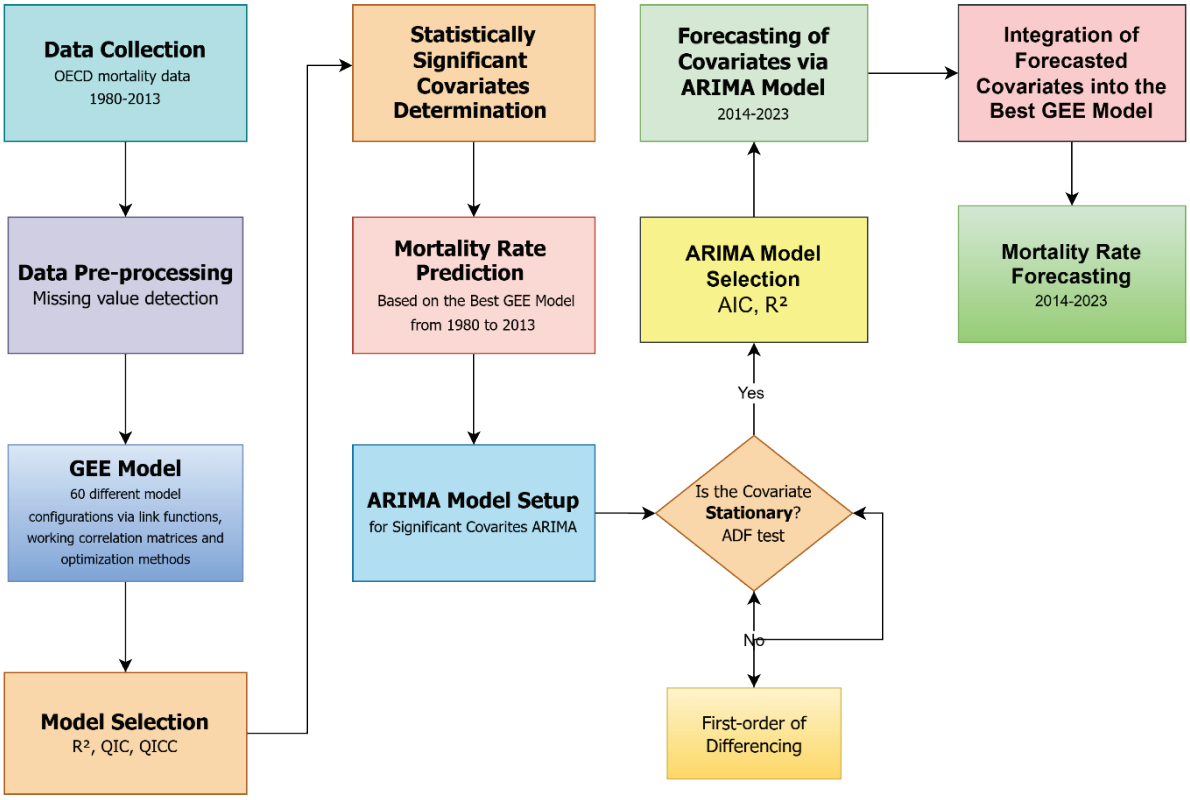

Figure 1. The hybrid GEE–ARIMA modeling framework.

Dataset

In this study, we obtained the dataset from the OECD. Stat [https://stats.oecd.org/] website for the OECD countries. This longitudinal dataset includes "mortality rates per 100,000 population" as the dependent variable, and also deaths from "non-communicable diseases", and also "external causes", and "accident falls" as independent variables for Australia, Austria, Belgium, Canada, Chile, Denmark, France, Hungary, Iceland, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Poland, Portugal, Spain, Sweden, Switzerland, United Kingdom (UK), and USA as the OECD countries included in the study between 1980 and 2013.

The selection of OECD countries was based on the consistency, availability, and reliability of longitudinal health and demographic records across these nations. The study period from 1980 to 2013 was chosen due to the completeness of time-series data for both mortality rates and cause-specific death indicators. The selected independent variables, namely, non-communicable diseases, external causes, and accidental falls, were determined based on their statistical significance in prior exploratory analyses and their policy relevance in public health. The forecast years of 2014 and 2023 were selected to evaluate the model's ability to make prospective mortality predictions beyond the observed range.

To provide a quick overview of the data distribution, Table 1 presents the descriptive statistics of mortality rates for several OECD countries included in the study. The table reports the mean, minimum, and maximum values calculated for neoplasm and the mean for diseases of the circulatory system, mental and behavioral diseases, diabetes mellitus, and diseases of the respiratory system over the 1980–2013 period.

Generalized Estimating Equations Approach

The GEEs approach proposed by Liang and Zeger [25] is considered as an extension of the generalized linear models (GLMs) approach introduced by Wedderburn [47] based on the quasi-likelihood estimation method as an appropriate alternative to the ML approach by avoiding distributional assumptions [46]. The most essential feature of the GEE approach is that a regression model can be obtained by using a link function to relate the independent variables to the mean structure of the dependent variable. For the longitudinal data structure, the GEEs also include relationships between within-cluster repeated measurements in the model [46]. The GEEs approach is analyzed in three parts for the longitudinal data as follows in Eqs (1)-(12):

1. Let y_{ij} be a continuous or discrete response variable for the i^{th} subject on the j^{th} independent variable ($i = 1, ..., K$; $j = 1, ..., n_i$), and x_{ij} denotes the $p \times 1$ dimensional covariates vector. Moreover, $y_i = (y_{i1}, y_{i2}, y_{i3}, ..., y_{in_i})'$ indicates the response vector with the mean vector noted by $\mu_i = (\mu_{i1}, \mu_{i2}, \mu_{i3}, ..., \mu_{in_i})$ where μ_{ij} is the corresponding j^{th} mean [75]. The conditional expected value (mean) of each dependent variable, $E(y_{ij} \mid x^{ij}) = \mu_{ij}$, is assumed to rely on the independent variables through a known g(.) link function:

$$g(\mu_{ij}) = \eta_{ij} = x_{ij}{}'\beta \tag{1}$$

where β is a $p \times 1$ dimensional regression parameters vector.

2. The conditional variance of each dependent variable y_{ij} is assumed to rely on the known variance function of the conditional mean $v(.)$:

$$Var(y_{ij}|x_{ij}) = \phi v(\mu_{ij}) \tag{2}$$

where ϕ is the scale parameter which might be required to be estimated [75]. It is worth noting that $v(.)$ and ϕ are

Table 1. Summary statistics of major causes of mortality by some OECD countries

Country	Neoplasm Mean	Neoplasm Min	Neoplasm Max	Diseases of the Circulatory System Mean	Mental and Behavioral Diseases Mean	Diabetes Mellitus Mean	Diseases of the Respiratory System Mean
Australia	259.8	174.5	229.1	400.3	140.2	55.1	123.4
Austria	234.4	210.1	259.8	390.2	138.4	52.3	118.5
Belgium	211.6	185.3	236.2	410.5	145.3	50.7	126.2
Canada	204.4	180.2	230.6	402.2	143.6	54.2	121.7
Chile	181.2	160.1	202.3	412.1	129.9	58.8	119.8
Denmark	199.3	180.2	215.6	420.6	130.0	51.1	130.1
France	217.3	196.1	237.4	408.8	148.2	49.9	134.0
Hungary	263.9	240.0	284.4	452.7	160.7	60.2	128.5
Iceland	172.9	155.0	190.1	384.5	126.8	47.6	117.3
Ireland	215.3	190.5	237.2	395.0	135.9	53.3	120.9

dependent on result distributions. If y_{ij} is a continuous variable, then $v(\mu_{ij}) = 1$, and ϕ is the error variance; if y_{ij} is a count variable, then $v(\mu_{ij})$ is equal to μ_{ij} and thus, $\phi = 1$.

In this study, normal, gamma, and inverse Gaussian distributions from symmetric and asymmetric continuous distributions are investigated using various link functions, parameter estimation methods, and working correlation matrices for the mortality of the OECD countries' longitudinal data by the GEEs approach. Detailed information about these distributions from the exponential family can be found in the following references [33, 61, 76-82].

3. The within-subject conditional relationships between the repeated measurements for the response variable y_{ij} observed at time t_{ij} are determined by the parameter α. Moreover, the variance-covariance matrix for y_i is noted by $V_i(\alpha)$ and estimation of the parameters vector β in the GEEs approach depends on it as follows:

$$V_i(\alpha) = \phi A_i^{\frac{1}{2}} R_i(\alpha) A_i^{\frac{1}{2}} \quad (3)$$

where $A_i = diag\{v(\mu_{ij})\}$ is a diagonal matrix consisting of diagonal elements of the variance function of the conditional mean $v(\mu_{ij})$, $R_i(\alpha)$ is a $n_i \times n_i$ dimensional correlation matrix known as the "working correlation matrix". Here, the term "working" is utilized to stress that V_i is only an approximate approach to the genuine mentioned covariance $\sum_i = Cov(y_i)$. The estimation of α is used an iterative fitting process based on the Pearson residuals $e_{ij} = \frac{(y_{ij}-\mu_{ij})}{\sqrt{v(\mu_{ij})}}$ computed via the current value of β [83]. Further, the scale parameter ϕ is estimated as follows:

$$\widehat{\phi} = \frac{1}{N-p} \sum_{i=1}^{K} \sum_{j=1}^{n_i} e_{ij}^2 \quad (4)$$

where $N = \sum_{i=1}^{K} n_i$ and p denote the total number of observations and the dimension of covariates, respectively.

According to Liang and Zeger [25], GEE possesses the desirable property of asymptotic consistency in estimating the parameter vector β, even in situations where the assumed "working" correlation structure $R_i(\alpha)$ is misspecified. This is achieved by resolving the following estimating equation:

$$U(\beta) = \sum_{i=1}^{n} D_i' [V_i(\widehat{\alpha})]^{-1} (y_i - \mu_i) = 0 \quad (5)$$

where D_i is $\frac{\partial \mu_i}{\partial \beta'}$. $\widehat{\beta}$ is asymptotically regularly distributed under mildly irregularity conditions, with a mean of β_0, and a covariance matrix computed using the sandwich estimator as follows:

$$\widehat{V}_{LZ} = \left(\sum_{i=1}^{K} D_i' [V_i(\widehat{\alpha})]^{-1} D_i\right)^{-1} M_{LZ} \left(\sum_{i=1}^{K} D_i' [V_i(\widehat{\alpha})]^{-1} D_i\right)^{-1} \quad (6)$$

with

$$M_{LZ} = \sum_{i=1}^{K} D_i' [V_i(\widehat{\alpha})]^{-1} Cov(y_i) [V_i(\widehat{\alpha})]^{-1} D_i \quad (7)$$

where $Cov(y_i) = \hat{r}_i \hat{r}_i'$ with $\hat{r}_i = y_i - \widehat{\mu}_i$ is an estimator of the variance-covariance matrix of y_i by replacing β, α, and ϕ via their consistent estimates [49, 75, 83, 84]. This "sandwich" estimator is robust since it remains consistent even when the correlation structure is incorrect.

In the GEEs approach, determining the working correlation matrix effectively will increase the reliability of the estimate of the parameter vector $\widehat{\beta}$, in order to best determine the structure of the actual relationships between the repeated measurements for the response variable [49]. For this aim, in this study, five different working correlation matrices are investigated to determine the relationships between the repeated measurements for the "mortality rates" of the OECD countries from 1980 to 2013 as follows:

Independent correlation matrix [49, 83, 85]:

$R = I$

$$Corr(y_{ij}, y_{ik}) = \begin{cases} 1, j = k \\ 0, j \neq k \end{cases} \quad (8)$$

Exchangeable correlation matrix [49, 83, 85]:

$R = \alpha$

$$Corr(y_{ij}, y_{ik}) = \begin{cases} 1, j = k \\ \alpha, j \neq k \end{cases} \quad (9)$$

First-order Autoregressive (AR (1)) correlation matrix [49, 83, 85]:

$R = a^k$

$$Corr(y_{ij}, y_{i,j+k}) = \alpha^k, \quad k = 0, 1, 2, \ldots, n_i - j \quad (10)$$

M-dependent correlation matrix [83, 86, 87]:

$$Corr(y_{ij}, y_{i,j+k}) = \begin{cases} 1, k = 0 \\ \alpha_k, k = 1, 2, \ldots, m \\ 0, k > m \end{cases} \quad (11)$$

Unstructured correlation matrix [49, 83, 85]:

$R = \alpha_{jk}$

$$Corr(y_{ij}, y_{ik}) = \begin{cases} 1, & j = k \\ \alpha_{jk}, & j \neq k \end{cases} \quad (12)$$

The hybrid GEE–ARIMA model proposed in this study was selected to address the dual challenges of modeling longitudinal panel data and forecasting future trends. While GEE is well-suited for capturing within-subject correlation and estimating parameters across countries, ARIMA provides the ability to model and forecast time-dependent predictors. By using ARIMA to generate future projections for significant independent variables and feeding these into the GEE framework, the proposed method leverages both temporal dynamics and intra-cluster correlation. This design also ensures statistical interpretability, which is often limited in more opaque machine learning models.

Model Selection Criteria

In order to evaluate and compare alternative model specifications within the hybrid GEE-ARIMA framework, appropriate model selection criteria must be utilized.

Akaike information criterion (AIC) [88] is the most widely used goodness-of-fit test statistic for the GLMs. However, GEEs cannot use the criteria directly, since it is not likelihood-based. Therefore, Pan [89] introduced the QIC based on the quasi-likelihood estimation method as a modification of the AIC criterion for the GEEs approach. In this study, this criterion is used to determine the best working correlation matrix in the GEEs approach. When $\boldsymbol{trace(\widehat{\Omega}_I^{-1}\widehat{V}_R) \approx trace(I) = p}$, there is a simplified version of QIC, called QICC [53] in Eq.(13) and Eq.(14):

$$\boldsymbol{QIC = -2Q(\hat{\mu}; I) + 2trace(\widehat{\Omega}_I^{-1}\widehat{V}_R)} \quad (13)$$

$$\boldsymbol{QICC = -2Q(\hat{\mu}; I) + 2p} \quad (14)$$

where I is the independent covariance structure, R is the working correlation matrix, $\boldsymbol{\widehat{V}_R}$ is the robust estimator of the working variance-covariance matrix, $\boldsymbol{\widehat{\Omega}_I}$ is the estimator of the other variance-covariance matrix obtained under the assumption of the independent correlation structure, and is the number of parameters in the model [53].

In this study, we decided to determine the most appropriate working correlation matrix through the smallest QIC and QICC. If QIC and QICC indicate different things in the study, we preferred the smallest QIC value, as suggested by Pan [90]. While QICC serves as a simplified form of QIC under the assumption that the trace term approximates the number of parameters (p), QIC provides a more general and accurate criterion for model selection in GEE, especially when model complexity increases. Therefore, QIC was prioritized for ensuring robust and reliable model selection.

After identifying the best-fitting GEE model based on these criteria, the next phase involved forecasting the significant explanatory variables using ARIMA models, selected via minimum AIC. The predicted values were subsequently incorporated into the GEE framework to generate mortality forecasts.

Autoregressive Integrated Moving Average Models

ARIMA models are a class of stochastic processes used in time series analysis [90]. ARIMA was originally developed by electrical engineers in 1930-1940 and emerged as an application of discrete-time filtration methods. Box and Jenkins [26] applied and developed this method on business and economic data. It is also called the Box-Jenkins method. ARIMA is an effective method used in many scientific studies. Generally, Box-Jenkins methods are expressed as ARIMA (p, d, q) where p, d, and q are the orders of autoregressive, the difference between the times of the series is stagnant, and the moving average, respectively. ARIMA (p, d, q) denoted in Eq. (15):

$$(\mathbf{1}-\boldsymbol{\phi_1}\boldsymbol{B}-\boldsymbol{\phi_2}\boldsymbol{B^2}-\ldots-\boldsymbol{\phi_p}\boldsymbol{B^p})(\mathbf{1}-\boldsymbol{B})^d\boldsymbol{z_t}=(\mathbf{1}-\boldsymbol{\theta_1}\boldsymbol{B}-\boldsymbol{\theta_2}\boldsymbol{B^2}-\ldots-\boldsymbol{\theta_q}\boldsymbol{B^q})\boldsymbol{\varepsilon_t} \quad (15)$$

where $(\mathbf{1} - \boldsymbol{B})^d$ can also be shown in ∇^d form in some sources that are *d.* order the difference process [91].

In the Box-Jenkins method, the determination of the appropriate model for the series consists of four steps. These are the stages of model identification, parameter estimation, residual analysis, and forecasting for the future [27]. First of all, it should be decided whether the series is stagnant before these operations are performed.

The stationarity of each time series was checked prior to model fitting to ensure that the ARIMA models were specified appropriately. The Augmented Dickey Fuller (ADF) test was conducted to test for the presence of unit roots in the predictor variables. If any series was found to be non-stationary, first order differencing was applied. The ADF test was then performed again on the differenced series to test for stationarity. The value of differencing (d) used in the ARIMA model structure was then determined based on the above, to comply with the stationarity assumptions needed for successful time series forecasting.

RESULTS AND DISCUSSION

This study uses IBM SPSS Statistics version 23 to apply the GEE technique to the mortality data of OECD countries from 1980 to 2013. Many combinations of continuous distributions (normal, gamma, inverse Gaussian), link functions (identity, log, inverse, inverse square), parameter estimation techniques (hybrid, Newton-Raphson, Fisher-Scoring, fixed value), and working correlation structures (independent, exchangeable, AR(1), M-dependent, unstructured) were investigated in order to identify the best GEE model. The QIC and its corrected variant, QICC, were used to evaluate the model fit.

Table 2. Goodness-of-Fit test statistics to select the best combination

Distribution	Link function	Working correlation matrix	Parameter estimation method	Scale parameter estimation method	QIC	QICC
Normal	Identity	Independent	Hybrid	MLE	20103.707	19981.009
Gamma	Identity	Exchangeable	Hybrid	Deviance	121.205*	76.015
Inverse Gaussian	Identity	Independent	Hybrid	Deviance	136.331	76.000*

*The smallest QIC and QICC values.

Table 3. Parameter estimates are obtained by using the best model structure belonging to OECD countries' mortality data

Covariates	$\widehat{\boldsymbol{\beta}}$	s.e($\widehat{\boldsymbol{\beta}}$)	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi -Square	df	Sig.
Certain infectious and parasitic diseases	0.770	0.1421	0.491	1.048	29.325	1	0.000
Neoplasms	0.989	0.1234	0.748	1.231	64.286	1	0.000
Leukemia	0.909	0.2442	0.431	1.388	13.865	1	0.000
Endocrine, nutritional, and metabolic diseases	0.733	0.0915	0.554	0.912	64.187	1	0.000
Diabetes mellitus	0.316	0.1166	0.087	0.544	7.331	1	0.007
Mental and behavioral diseases	1.022	0.0292	0.965	1.079	1222.553	1	0.000
Diseases of the nervous system	0.998	0.0342	0.931	1.065	849.757	1	0.000
Diseases of the circulatory system	1.001	0.0043	0.993	1.010	54043.045	1	0.000
Diseases of the respiratory system	1.014	0.0547	0.907	1.122	344.039	1	0.000
Diseases of the digestive system	0.664	0.1791	0.313	1.015	13.760	1	0.000
Diseases of the skin	2.141	0.5119	1.138	3.145	17.497	1	0.000
Diseases of the musculoskeletal system	1.409	0.1351	1.145	1.674	108.792	1	0.000
Diseases of the genitourinary system	1.123	0.0639	0.998	1.248	308.427	1	0.000
Certain conditions originating in the perinatal period	1.835	0.4558	0.941	2.728	16.201	1	0.000
Abnormalities	0.962	0.2550	0.462	1.462	14.237	1	0.000
Symptoms, signs, and ill-defined causes	1.036	0.0163	1.004	1.068	4021.097	1	0.000
External causes	0.923	0.1818	0.566	1.279	25.748	1	0.000
Accidental falls	-0.152	0.0764	-0.301	-0.002	3.935	1	0.047

The following configurations were found to be the best: (1) distribution: gamma; (2) link function: identity; (3) working correlation structure: exchangeable; (4) parameter estimation method: hybrid (Fisher Scoring followed by Newton-Raphson); and (5) scale estimation: deviance. The structure's fit for the data was confirmed by the lowest QIC score (121.205). Pan [90] stressed that QIC should be given priority when determining the optimal model, even if the inverse Gaussian distribution obtained the lowest QICC value (76). Gamma-based models were therefore chosen for additional examination.

Table 2 summarizes the model selection process. Although 60 different model configurations were evaluated, only the most representative are reported for brevity

Table 3 presents the parameter estimates, standard errors, Wald confidence intervals, chi-square statistics, degrees of freedom (df), and significance levels for statistically meaningful covariates in the selected GEE model.

Table 3 presents the regression coefficients ($\hat{\beta}$), their standard errors, 95% Wald confidence intervals, chi-square statistics, degrees of freedom, and p-values for the explanatory variables identified as statistically significant within the GEE framework. While most covariates show strong statistical significance, Accidental Falls was found to be marginally significant (p = 0.047). Nevertheless, it was retained in the final model due to its relevance and contribution to model completeness. From a methodological point of view, it did not introduce multicollinearity or instability, and from a public health perspective, falls continue to be a major cause of mortality for the elderly in many OECD countries. As a result, its inclusion provides both practical application and statistical purity.

Eq. (16) below is the final GEE model with an exchangeable correlation structure and a gamma distribution with an identity link function based on these estimates:

$$\begin{aligned}(\textit{Mortality})_{ij} = {} & 0.770(\textit{Infectious})_{ij} + 0.989(\textit{Neoplasms})_{ij} \\ & + 0.909(\textit{Leukemia})_{ij} + 0.733(\textit{Endocrine})_{ij} \\ & + 0.316(\textit{Diabetes})_{ij} + 1.022(\textit{Mental})_{ij} \\ & + 0.998(\textit{Nervous})_{ij} + 1.001(\textit{Circulatory})_{ij} \\ & + 1.014(\textit{Ischaemic})_{ij} + 0.664(\textit{Digestive})_{ij} \\ & + 2.141(\textit{Skin})_{ij} + 1.409(\textit{Musculoskeletal})_{ij} \\ & + 1.123(\textit{Genitourinary})_{ij} + 1.835(\textit{Perinatal})_{ij} \\ & + 0.962(\textit{Abnormal})_{ij} + 1.036(\textit{IlldefinedCauses})_{ij} \\ & + 0.923(\textit{ExternalCauses})_{ij} - 0.152(\textit{AccidentalFalls})_{ij}; \\ & i=\textit{Australia, Austria, Belgium, Canada, Chile, Denmark, France, Hungary, Iceland, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Poland, Portugal, Spain, Sweden, Switzerland, UK, USA,} \\ & j = 1980, 1981, 1982, ..., 2011, 2012, 2013\end{aligned} \tag{16}$$

This equation integrates both cross-sectional and longitudinal effects on mortality, capturing the fixed-effects structure obtained from significant covariates.

ARIMA models were then used to forecast these important independent variables. For every covariate between 2014 and 2023, country-specific ARIMA models were created using R software. We provide a representative case of circulatory system-related mortality in Hungary to demonstrate the forecasting methods used in this work. The ADF test revealed non-stationarity in the initial time series, with a p-value of 0.963 and a value of 0.055. Following first-order differencing, stationarity at the 5% significance level was indicated by the ADF test statistic of -3.28 with a p-value of 0.0237.

Plots of autocorrelation and partial autocorrelation as well as the R output of the arima() function indicated that an ARIMA(1,1,1) model would be suitable. After the model was calculated, the z-statistics showed that every parameter was statistically significant. The Ljung–Box test was used for residual diagnostics, and the results showed a p-value higher than 0.05, indicating that the residuals are a white noise process.

The circulatory-related mortality rates in Hungary from 2014 to 2023 were then forecast with the ARIMA(1,1,1) model. To forecast future all-cause mortality, these projected values were incorporated into the hybrid GEE–ARIMA framework.

All OECD country-specific models, including the ARIMA specifications for each statistically significant covariate, were constructed following the same systematic procedure. The forecasted covariate values were incorporated into Eq. (16) to generate mortality projections.

The GEE model's predicted and actual mortality values for certain OECD countries are shown in Figure 2. The model's predictive accuracy is demonstrated by the satisfactory agreement between observed and anticipated values.

Using both historical and forecasted data, Figure 3 shows the average mortality trend for all countries from 1980 to 2023. There is a clear pattern that reflects improvements in public health policy, illness prevention, and healthcare delivery.

Table 4 summarizes a portion of these forecasts, while detailed country-specific information is provided in Tables 5-10 in the Appendix. Figures 4 and 5, which contrast average mortality rates by country from 1980 to 2013 and 2014 to 2023, provide a visual representation of these projections.

Figure 4. The highest mortality is in Hungary, Poland and Ireland, while the lowest is in Japan, Switzerland and France. Historical validation of the model performance. As can be seen in Figure 5, these rankings are generally maintained in the forecasted period, although a convergence in mortality rates is visible (probably driven by improvements in health systems).

As shown in Table 5, mortality in Hungary dropped steadily from 1134 to 1093 per 100,000. However, there are some concerns as indicated by slight increases in deaths related to the mental and nervous system.

Table 6 forecasts a modest decline in Japan's overall mortality. However, deaths from nervous and mental disorders are projected to rise, consistent with Japan's aging population structure.

Table 7 presents an exception: the Netherlands is expected to see an increase in mortality, from 772 to 818 per 100,000. The rise in deaths related to mental and behavioral disorders is particularly remarkable, with an expected more than doubling by 2023.

Table 8 shows that Norway's mortality is projected to decline from 738 to 691, with stable or slightly decreasing rates for nervous and mental health causes, marking a contrast to other countries.

Table 9 reports that Spain will experience one of the steepest declines in mortality, falling from 638 to 485 per 100,000. Mortality due to nervous system and mental disorders also trends downward.

Table 10 shows a gradual decline in overall mortality in the USA. However, deaths from neoplasms, diabetes, and especially psychiatric and neurological disorders are expected to increase.

More details on disease-specific trends can be found in Figures 6 and 7. Figure 6 shows the average mortality by cause across six countries from 2014 to 2023. The most frequent causes of death are neoplasms and circulatory diseases, with Hungary having the highest number of deaths related to circulatory diseases and the USA leading in mortality related to neoplasms. In most categories, Japan and Spain have lower mortality.

Figure 7 focuses on mental and behavioral disorders in Hungary, Japan, and the USA. Hungary's rates are stable; Japan is slightly up, but the USA shows a significant upward trend, indicating an increasing mental health burden.

Figures 2 to 7 collectively show the potential usefulness of the hybrid GEE–ARIMA model in capturing both the overall level mortality trends and the granular disease-specific patterns in various settings in the OECD.

This study investigates the mortality trends in OECD countries over the period 1980-2013 using a hybrid forecasting framework based on GEE and ARIMA models. The choice of OECD countries was determined by the availability of complete and consistent mortality data over the period of interest.

Different combinations of distributional assumptions (normal, gamma, inverse Gaussian), link functions, working correlation structures, parameter estimation techniques and scale estimation methods were used to first model mortality rates using the GEE approach (Eq. (1)–(14) and Eq. (16)). Table 2 presents the optimal model according to the minimum QIC, which used a gamma distribution, an identity link function, an exchangeable working correlation structure, a hybrid parameter estimation method and a deviance-based scale estimation. From this configuration, the statistically significant predictors were identified and forecasted for each country for the period of 2014–2023 using ARIMA models defined in Eq. (15): These predicted covariates were then reinserted into the GEE model to estimate the future mortality rates and allow decade ahead

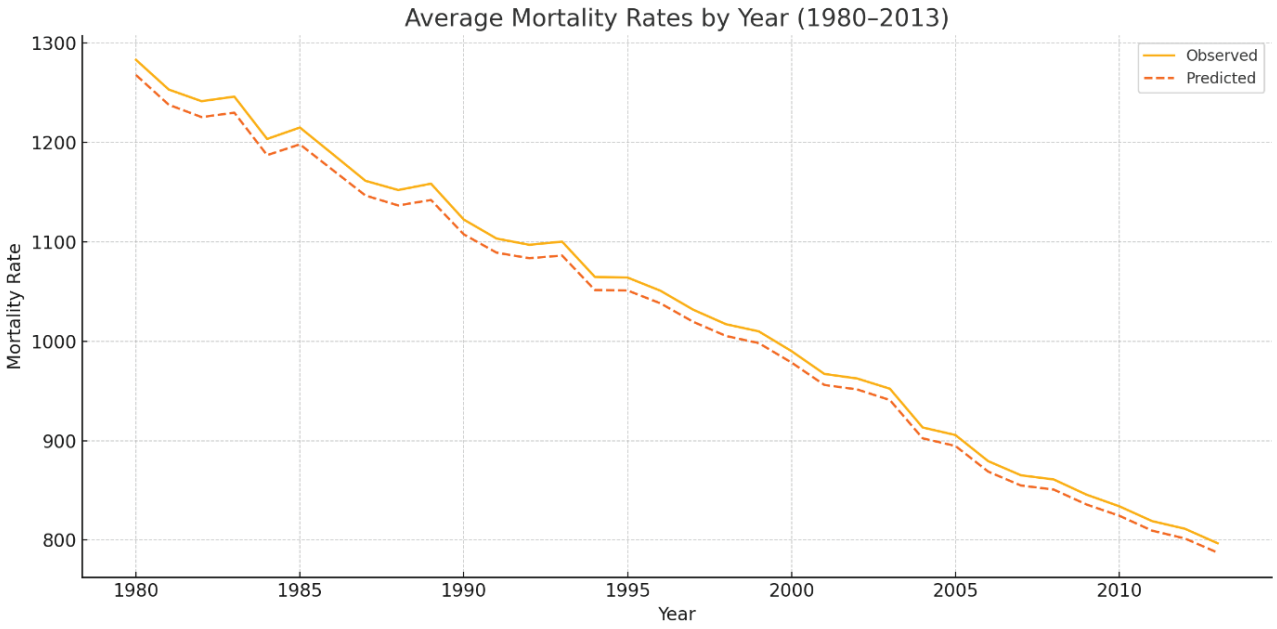

Figure 2. Actual and predicted average OECD countries' mortality rates using the GEE approach from 1980 to 2013.

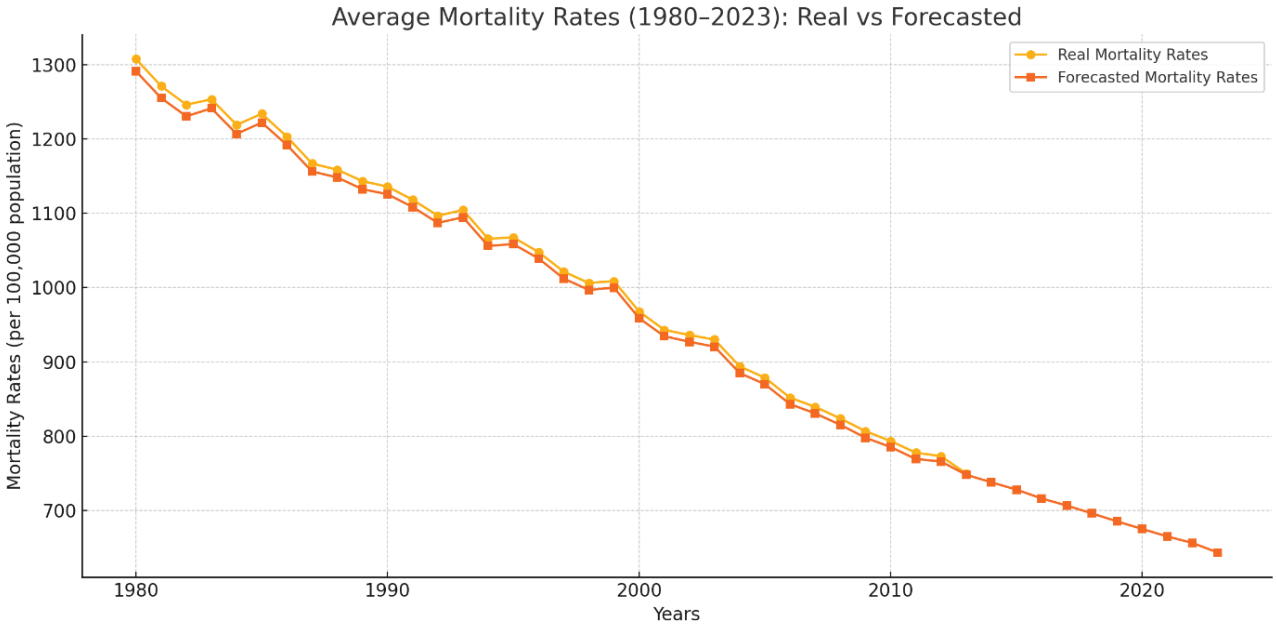

Figure 3. Actual and forecasted average OECD countries' mortality rates using the GEE approach from 1980 to 2023.

predictions. These methodological steps and their outcomes are shown in Figures 2-5. The variable "accidental falls" was kept in the model despite marginal significance ($p = 0.047$) due to its public health relevance, especially for aging populations. Statistical validity is not compromised by the inclusion of the latter, which makes the model more comprehensive.

The hybrid GEE-ARIMA modeling framework has several advantages:

(i) GEE models provide a flexible framework for modeling longitudinal data without making solid distributional assumptions.

(ii) The model is further enriched by the inclusion of a number of cause-specific death categories as independent variables.

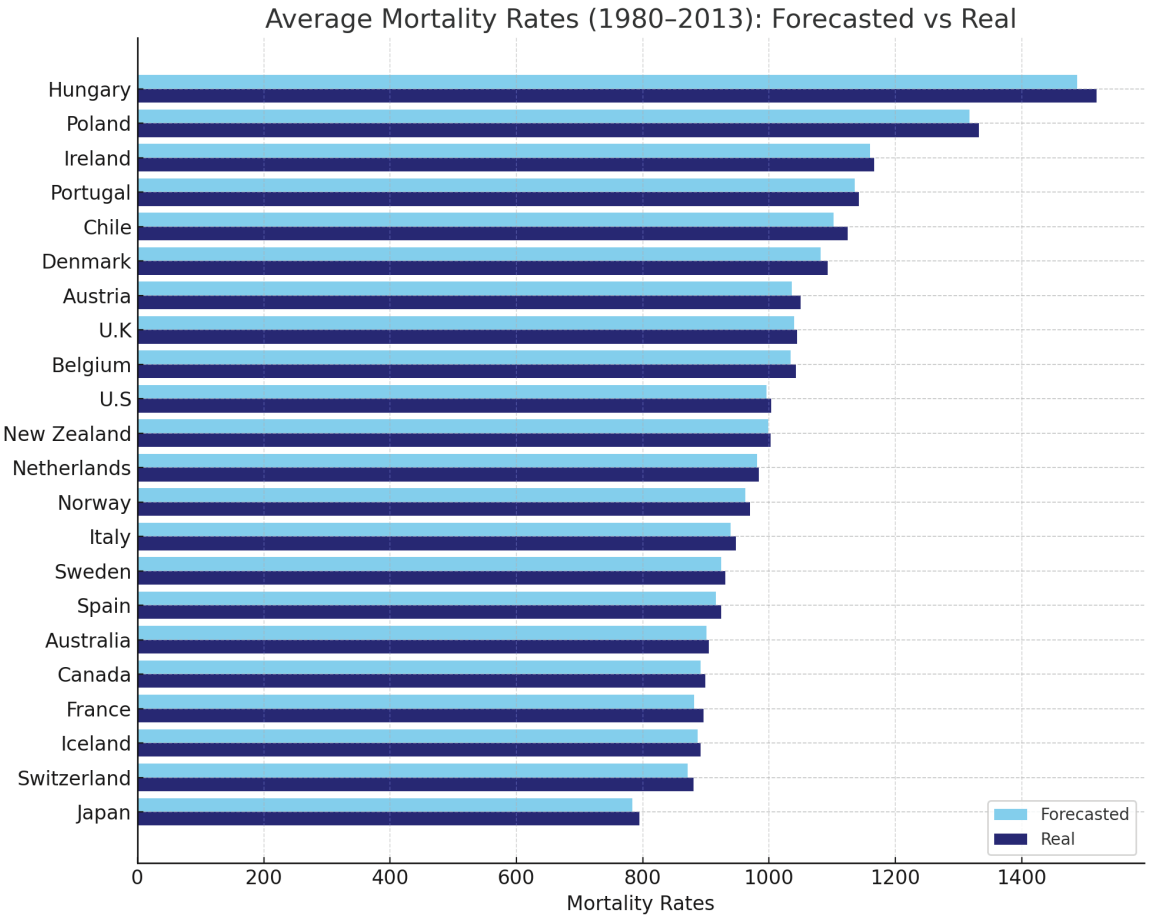

Figure 4. Predicted average mortality rates for OECD countries from 1980 to 2013.

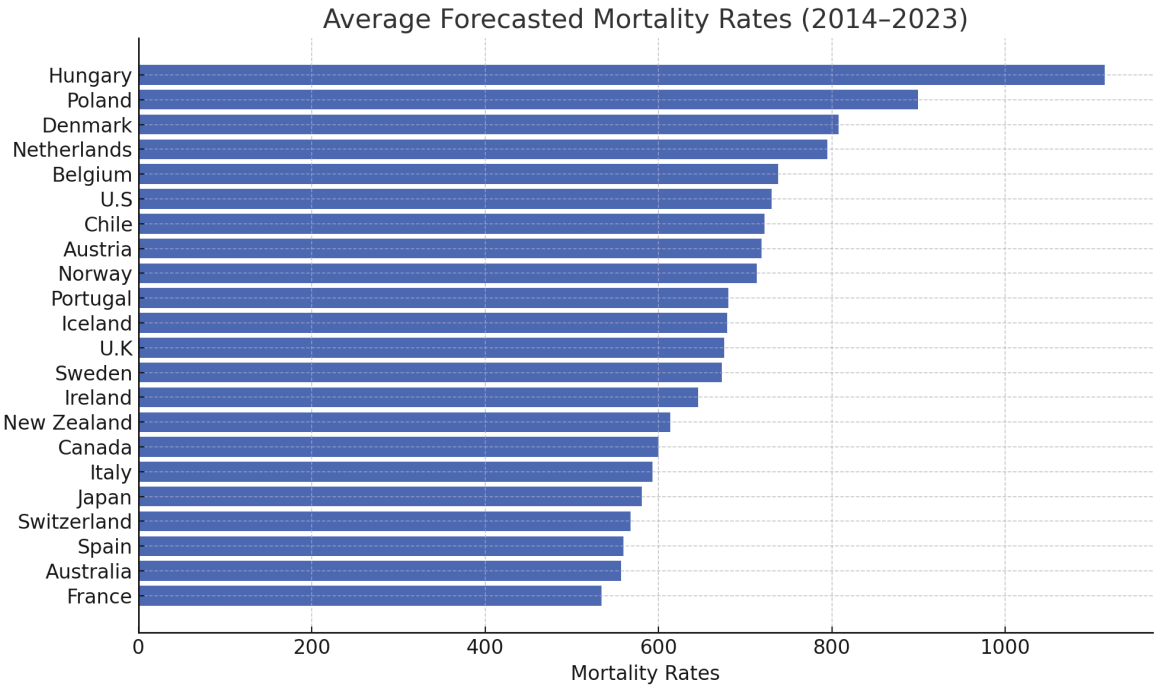

Figure 5. Comparison of average forecasted mortality rates for OECD countries from 2014 to 2023 .

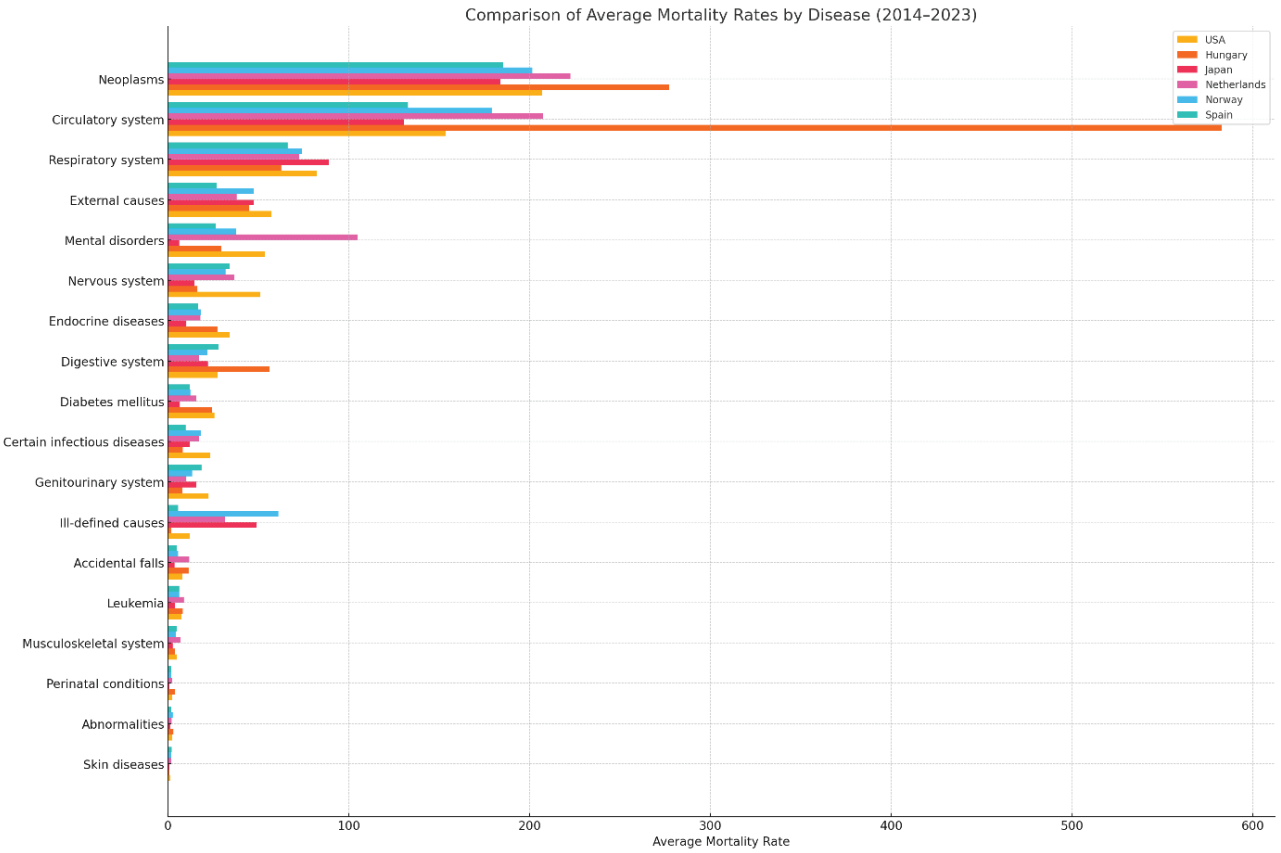

Figure 6. Comparison of average forecasting mortality rates by disease for some OECD countries from 2014 to 2023.

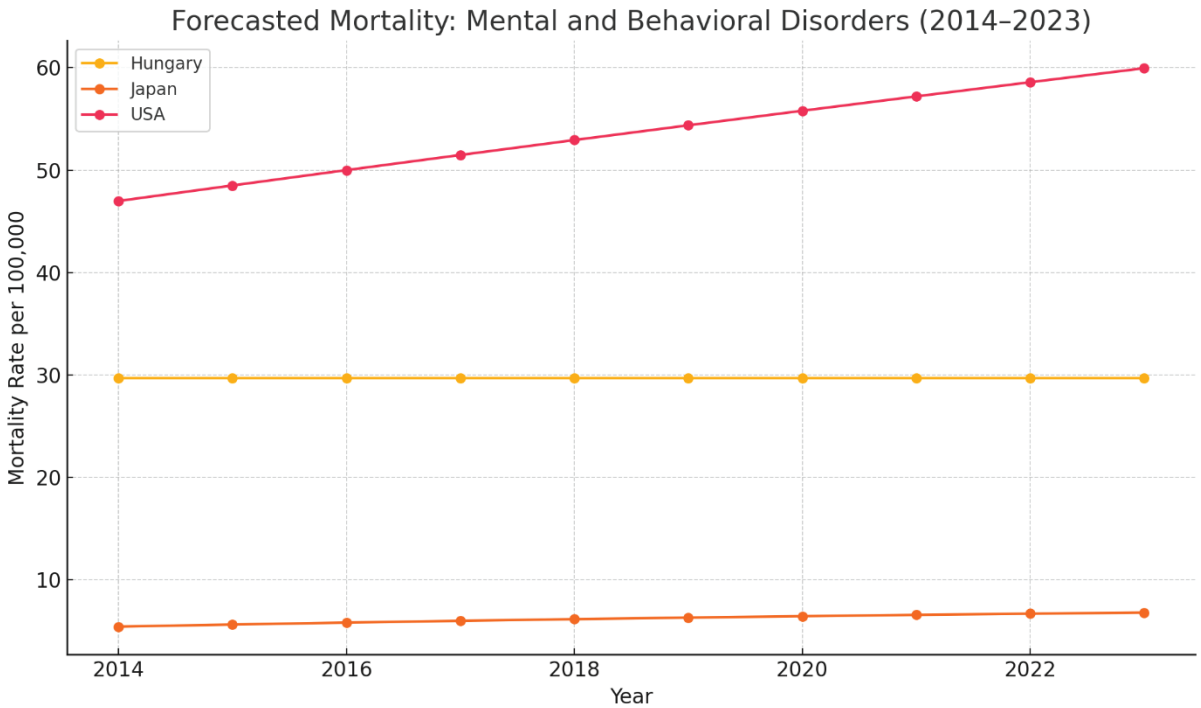

Figure 7. Comparison of average forecasting mortality rates by mental and behavioral disorders for Hungary, Japan, and the USA from 2014 to 2023.

Table 4. Some of the OECD member countries' actual, estimated, and forecasted values between 1980 and 2023 by using the hybrid GEE-ARIMA approach

Country	Year	Actual value	Predicted value	Forecasted value
Hungary	1980	1771.6	1747.9502	-
	1990	1690.4	1655.2769	-
	2000	1509.3	1474.7588	-
	2010	1217.2	1195.5728	-
	2015	-	-	1131.861117
	2020	-	-	1107.620769
	2023	-	-	1093.814827
Japan	1980	1093.5	1079.9863	-
	1990	886.6	874.4023	-
	2000	708.1	696.5455	-
	2010	622.4	613.2708	-
	2015	-	-	591.7269195
	2020	-	-	575.9700875
	2023	-	-	566.2642829
Netherland	1980	1133.1	1130.6958	-
	1990	1060.2	1057.9063	-
	2000	995.2	993.0784	-
	2010	789.5	785.4648	-
	2015	-	-	776.1473402
	2020	-	-	803.5354028
	2023	-	-	818.8284283
Norway	1980	1151.5	1143.366	-
	1990	1096.6	1090.3125	-
	2000	939.9	932.2111	-
	2010	774.9	767.9937	-
	2015	-	-	731.0043563
	2020	-	-	704.7410678
	2023	-	-	691.6829106
Spain	1980	1156.8	1142.1631	-
	1990	1042.5	1033.965	-
	2000	871.3	863.6636	-
	2010	704.2	698.0725	-
	2015	-	-	622.657563
	2020	-	-	531.375311
	2023	-	-	485.122471
USA	1980	1184.2	1175.9625	-
	1990	1075.4	1071.3668	-
	2000	994.6	985.7404	-
	2010	835.3	825.1929	-
	2015	-	-	764.3866416
	2020	-	-	715.8845212
	2023	-	-	687.4465377

(iii) The QIC criterion gives the statistical reliability of model selection.
(iv) The ARIMA component enables accurate time series forecasting of key predictors at the national level.
(v) The country-level mortality estimates for 2014-2023 presented in Tables 5-10 and Figures 5-7 show consistent declines in overall mortality.
(vi) Most importantly, the study provides an original integrated forecasting framework that coherently combines GEE and ARIMA methods.

There are some limitations too.

(i) While we have examined a wide range of disease categories, the inclusion of other independent variables such as environmental, socio-economic or policy-related factors could further enhance predictive accuracy. We investigated dimensionality reduction techniques, but did not proceed with them because of interpretability issues in ARIMA forecasting.
(ii) Some OECD countries were excluded because of large amounts of missing data.
(iii) The time window was restricted to 1980–2013 because of data consistency; extending the time window may pose a risk to the reliability of data.

The proposed hybrid framework performed better than the previous studies that used GEE or ARIMA independently in mortality studies. For instance, Ye and Kerr [68] applied time series models to study alcohol-related mortality trends, while Makinde et al. [69] employed ARIMA to study COVID-19 mortality patterns in African countries. In Peru, Kuo et al. [72] achieved an R^2 of 0.788 with a traditional ARIMA model to forecast mortality patterns in response to environmental conditions, which confirms the utility of time-series forecasting and also illustrates the advantage of our integrated approach. In contrast, the model proposed in this study shows high predictive accuracy for OECD countries with a mean R^2 value of 0.928. The proposed hybrid GEE-ARIMA, by taking into account both intra-subject correlation and the temporal evolution simultaneously, provides a more comprehensive and reliable forecasting strategy for complex multinational datasets.

CONCLUSION

Mortality is one of the main indicators for the evaluation of the demographic and health structures of the countries. With the continued progress of technology, medicine, and socioeconomic development, overall mortality rates are expected to decline, and the population is aging in many countries. Accurate forecasting of mortality trends is thus important for effective long-term planning in health care, insurance, and economic policy.

In this paper, we propose a new hybrid modeling framework that combines the GEE and ARIMA models to predict mortality rates across OECD countries. We used the GEE-based approach on the longitudinal data on mortality. The method of ARIMA was used to predict key explanatory variables over time. The best model was chosen using QIC, and the best structure was estimated using an exchangeable working correlation. In addition, the parameter estimation methods, namely the Hybrid, Newton-Raphson, Fisher Scoring, and Fixed Value algorithms, were taken into consideration.

The findings indicate that mortality is likely to decline in most OECD countries, which is consistent with the expectation of increasing life expectancy. This finding is consistent with the previous studies, such as Ye and Kerr [68], Ghoshal [70], and Lazoğlu and Büyükyazıcı [92], which also reported declining trends in overall mortality among developed countries.

The main finding, however, is that mortality for nervous system impairment and psychiatric or behavioral disorders is expected to increase in several countries. This is consistent with the findings of Barbiellini Amidei et al. [71], Kim et al. [93], Dreviskaite et al. [94], and Kostev et al. [95], who reported an increase in cause-specific mortality for neurological and mental health conditions in the aging population. Notably, there will be a decline in these categories for exceptions such as Norway and Spain. Detailed forecasts at the country level are presented in Tables 5-10.

The hybrid modeling framework proposed in this study has both methodological contributions and practical implications for public health and policymaking. The model provides accurate long-term mortality forecasts in multiple countries and thus helps prepare health systems, plan demographics, and allocate resources. It can help government agencies, insurance companies, and international health organizations to forecast future mortality trends, identify high-risk areas, and develop targeted intervention strategies. Such forecasting tools are becoming increasingly important in the context of post-pandemic recovery and aging populations.

Contrary to the literature, this study proposes a hybrid approach that is statistically interpretable and that accounts for the within-panel correlation and temporal dynamics in mortality forecasting, which has been paid little attention.

In future work, we hope to extend this work to include further demographic and economic factors, such as healthcare expenditures, retirement policies, and insurance dynamics. Furthermore, the comparative studies that will be conducted on the interaction of fertility and mortality by employing the hybrid GEE–ARIMA structures will be useful in guiding the demographic policy at the country level. Further, the study would also look into the use of advanced machine learning models such as LSTM and gradient boosting techniques to assess their potential in enhancing the forecasting accuracy and flexibility of the proposed framework.

ACKNOWLEDGMENTS

This study is based on Oznur Ozaltin's master's thesis, supervised by Neslihan Iyit, and its title is "An Application

of Generalized Estimating Equations (GEE) Approach on Organization for Economic Cooperation and Development (OECD) Countries Mortality Data".

AUTHORSHIP CONTRIBUTIONS

Ozaltin O.; methodology, software, validation, analysis, writing—original draft preparation. Iyit N.; methodology, validation, writing-- review and editing, supervision.

DATA AVAILABILITY STATEMENT

https://stats.oecd.org/.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

ETHICS

There are no ethical issues with the publication of this manuscript.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence was not used in the preparation of the article.

APPENDIX A

The appendix includes forecasted variables with different ARIMA models per OECD country.

REFERENCES

[1] İyit N, Sevim F, Kahraman ÜM. Investigating the impact of CO2 emissions on the COVID-19 pandemic by generalized linear mixed model approach with inverse Gaussian and gamma distributions. Open Chem 2023;21:20220301. [CrossRef]

[2] Rangasamy M, Chesneau C, Martin-Barreiro C, Leiva V. On a novel dynamics of SEIR epidemic models with a potential application to COVID-19. Symmetry 2022;14:1436. [CrossRef]

[3] Almars AM, Atlam ES, Noor TH, ELmarhomy G, Alagamy R, Gad I. Users opinion and emotion understanding in social media regarding COVID-19 vaccine. Computing 2022;104:1481–1496. [CrossRef]

[4] Noor TH, Almars A, Gad I, Atlam ES, Elmezain M. Spatial impressions monitoring during COVID-19 pandemic using machine learning techniques. Computers 2022;11:52. [CrossRef]

[5] Malki Z, Atlam ES, Ewis A, Dagnew G, Ghoneim OA, Mohamed AA, et al. The COVID-19 pandemic: Prediction study based on machine learning models. Environ Sci Pollut Res 2021;28:40496–40506. [CrossRef]

[6] Atlam ES, Ewis A, Abd El-Raouf M, Ghoneim O, Gad I. A new approach in identifying the psychological impact of COVID-19 on university student's academic performance. Alexandria Eng J 2022;61:5223–5233. [CrossRef]

[7] Malki A, Atlam ES, Hassanien AE, Ewis A, Dagnew G, Gad I. SARIMA model-based forecasting required number of COVID-19 vaccines globally and empirical analysis of peoples' view towards the vaccines. Alexandria Eng J 2022;61:12091–12110. [CrossRef]

[8] Ozaltin O, Iyit N. Modelling the US diabetes mortality rates via generalized linear model with the Tweedie distribution. Int J Sci Res 2018;7:1326–1334. [CrossRef]

[9] Grant WB, Boucher BJ. An exploration of how solar radiation affects the seasonal variation of human mortality rates and the seasonal variation in some other common disorders. Nutrients 2022;14:2519. [CrossRef]

[10] Kodera S, Rashed EA, Hirata A. Correlation between COVID-19 morbidity and mortality rates in Japan and local population density, temperature, and absolute humidity. Int J Environ Res Public Health 2020;17:5477. [CrossRef]

[11] Acharya D, Lee K, Lee DS, Lee YS, Moon SS. Mortality rate and predictors of mortality in hospitalized COVID-19 patients with diabetes. Healthcare 2020;8:341. [CrossRef]

[12] Ioacara S, Popescu AC, Tenenbaum J, Dimulescu DR, Popescu MR, Sirbu A, et al. Acute myocardial infarction mortality rates and trends in Romania between 1994 and 2017. Int J Environ Res Public Health 2020;17:285. [CrossRef]

[13] Almetwally EM, Alharbi R, Alnagar D, Hafez EH. A new inverted topp-leone distribution: Applications to the COVID-19 mortality rate in two different countries. Axioms 2021;10:25. [CrossRef]

[14] Khan IU, Aslam N, Aljabri M, Aljameel SS, Kamaleldin MMA, Alshamrani FM, et al. Computational intelligence-based model for mortality rate prediction in COVID-19 patients. Int J Environ Res Public Health 2021;18:6429. [CrossRef]

[15] Kato R, Okada M. Can financial support reduce suicide mortality rates? Int J Environ Res Public Health 2019;16:4797. [CrossRef]

[16] Zahid MN, Perna S. Continent-wide analysis of COVID 19: Total cases, deaths, tests, socio-economic, and morbidity factors associated to the mortality rate, and forecasting analysis in 2020-2021. Int J Environ Res Public Health 2021;18:5350. [CrossRef]

[17] La Vignera S, Cannarella R, Condorelli RA, Torre F, Aversa A, Calogero AE. Sex-specific SARS-CoV-2 mortality: Among hormone-modulated ACE2 expression, risk of venous thromboembolism and hypovitaminosis D. Int J Mol Sci 2020;21:2948. [CrossRef]

[18] Alhazzani AA, Mahfouz AA, Abolyazid AY, Awadalla NJ, Katramiz K, Faraheen A, et al. In hospital stroke mortality: Rates and determinants in southwestern Saudi Arabia. Int J Environ Res Public Health 2018;15:927. [CrossRef]

[19] Sugawara N, Yasui-Furukori N, Ishii N, Iwata N, Terao T. Lithium in tap water and suicide mortality in Japan. Int J Environ Res Public Health 2013;10:6044–6048. [CrossRef]

[20] Esquível ML, Krasii NP, Guerreiro GR, Patrício P. The multi-compartment SI (RD) model with regime switching: An application to COVID-19 pandemic. Symmetry 2021;13:2427. [CrossRef]

[21] Al Noufaey KS. Stability analysis of a diffusive three-species ecological system with time delays. Symmetry 2021;13:2217. [CrossRef]

[22] Song G, Liang G, Tian T, Zhang X. Mathematical modeling and analysis of tumor chemotherapy. Symmetry 2022;14:704. [CrossRef]

[23] Aslam M, Albassam M. Application of neutrosophic logic to evaluate correlation between prostate cancer mortality and dietary fat assumption. Symmetry 2019;11:330. [CrossRef]

[24] Ozaltin O, Yeniay O, Subasi A. OzNet: A new deep learning approach for automated classification of COVID-19 computed tomography scans. Big Data 2023;11:245–258. [CrossRef]

[25] Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13–22. [CrossRef]

[26] Box GE, Jenkins GM. Time series analysis, control, and forecasting. San Francisco: Holden-Day; 1976. p. 10.

[27] Chesneau C, Bakouch HS, Akdoğan Y. The binomial-discrete Poisson-Lindley model: Modelling and applications to count regression. Commun Math Res 2022;38:28–51. [CrossRef]

[28] Bakouch H, Chesneau C, Karakaya K, Coşkun K. The Cos-Poisson model with a novel count regression analysis. Hacettepe J Math Statis 2021;50:559–578. [CrossRef]

[29] Nasiru S, Chesneau C, Abubakari AG, Angbing ID. Generalized unit half-logistic geometric distribution: Properties and regression with applications to insurance. Analytics 2023;2:438–462. [CrossRef]

[30] Anju RP, Prince M, Raju R, Meriat J, Anitta T, Chesneau C, et al. Impact of smartphone usage on social interactions and attitudes among Keralites (India) in the post-pandemic period: A descriptive statistics approach. J Stat 2023;27:42–58. [CrossRef]

[31] Korkmaz MÇ, Chesneau C, Korkmaz ZS. A new alternative quantile regression model for the bounded response with educational measurements applications of OECD countries. J Appl Stat 2023;50:131–154. [CrossRef]

[32] Bakouch HS, Karakaya K, Chesneau C, Akdoğan Y. A non-negative integer-valued model. Applic Anal Discr Math 2022;16:467–484. [CrossRef]

[33] Irshad MR, Chesneau C, Shibu DS, Monisha M, Maya R. A novel generalization of zero-truncated binomial distribution by Lagrangian approach with applications for the COVID-19 pandemic. Stats 2022;5:1004–1028. [CrossRef]

[34] Korkmaz MÇ, Chesneau C. On the unit Burr-XII distribution with the quantile regression modeling and applications. Computational and Applied Mathematics 2021;40:29. [CrossRef]

[35] Merupula J, Vaidyanathan V, Chesneau C. Prediction interval for compound Conway-Maxwell-Poisson regression model with application to vehicle insurance claim data. Math Comput Appl 2023;28:39. [CrossRef]

[36] SK KB, Chesneau C, Anthonysamy V. Rice crop growth analysis using auto regressive models. JProbab Statis Sci 2023;21:1–15. [CrossRef]

[37] Nasır JA, Chesneau C, Imran M, Jamal F. A study on mortality rate for infant in Pakistan: An application of biostatistics. Asian J Adv Res 2021;4:38–47.

[38] Yonar H, Neslihan İ. Modeling the causality relationships between Gdp/Gni and electricity consumption according to income levels of countries by Generalized Estimating Equations. Selçuk Üniv Sosyal Bil Enst Derg 2018;39:191–200.

[39] Iyit N. Modelling world energy security data from multinomial distribution by generalized linear model under different cumulative link functions. Open Chem 2018;16:377–385. [CrossRef]

[40] Yonar H, İyit N. Some Generalized Estimating Equations models based on causality tests for investigation of the economic growth of the country groups. Found Comput Decis Sci 2021;46:297–315. [CrossRef]

[41] Zeger SL, Liang KY, Self SG. The analysis of binary longitudinal data with time independent covariates. Biometrika 1985;72:31–38. [CrossRef]

[42] Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 1986;42:121–130. [CrossRef]

[43] Zorn CJ. Generalized estimating equation models for correlated data: A review with applications. AmJ Politic Sci 2001;45:470–490. [CrossRef]

[44] Ballinger GA. Using generalized estimating equations for longitudinal data analysis. Organiz Res Methods 2004;7:127–150. [CrossRef]

[45] Ziegler A, Kastner C, Blettner M. The generalised estimating equations: An annotated bibliography. Biometric J 1998;40:115–139. [CrossRef]

[46] Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G. Longitudinal data analysis. Boca Raton: CRC Press; 2008. [CrossRef]

[47] Wedderburn RW. Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. Biometrika 1974;61:439–447. [CrossRef]

[48] McCullagh P, Nelder JA. Generalized linear models. 2nd ed. London: Chapman & Hall; 1989. [CrossRef]
[49] Davis CS. Statistical methods for the analysis of repeated measurements. New York: Springer Science & Business Media; 2002. [CrossRef]
[50] Firth D. Generalized linear models. In: Hinkley DV, Reid N, Snell EJ, editors. Statistical theory and modelling. London: Chapman & Hall; 1991. p. 55–82.
[51] Diggle P. Analysis of longitudinal data. 2nd ed. Oxford: Oxford University Press; 2002. [CrossRef]
[52] Pan W, Connett JE. Selecting the working correlation structure in generalized estimating equations with application to the lung health study. Statistica Sinica 2002;12:475–490.
[53] Cui J. QIC program and model selection in GEE analyses. Stata Journal 2007;7:209–220. [CrossRef]
[54] Hardin JW, Hilbe JM. Generalized estimating equations. 2nd ed. Boca Raton: CRC Press; 2013. [CrossRef]
[55] Dobson AJ, Barnett AG. An introduction to generalized linear models. 3rd ed. Boca Raton: CRC Press; 2008. [CrossRef]
[56] Gosho M. Criteria to select a working correlation structure for the generalized estimating equations method in SAS. J Stat Softw 2014;57:1–10. [CrossRef]
[57] İyit N, Yonar H, Genç A. Generalized linear models for European Union countries energy data. Acta Phys Polon A 2016;130:111–114. [CrossRef]
[58] Zeger SL, Liang KY, Albert PS. Models for longitudinal data: A generalized estimating equation approach. Biometrics 1988;44:1049–1060. [CrossRef]
[59] Iqbal A, Mahmood T, Ali Z, Riaz M. On enhanced GLM-based monitoring: An application to additive manufacturing process. Symmetry 2022;14:122. [CrossRef]
[60] Kim JM, Ha ID. Deep learning-based residual control chart for binary response. Symmetry 2021;13:1389. [CrossRef]
[61] Jamal A, Mahmood T, Riaz M, Al-Ahmadi HM. GLM-based flexible monitoring methods: An application to real-time highway safety surveillance. Symmetry 2021;13:362. [CrossRef]
[62] Foster SD, Bravington MV. A Poisson-Gamma model for analysis of ecological non-negative continuous data. Environmental and Ecological Statistics 2013;20:533–552. [CrossRef]
[63] Fernández D, McMillan L, Arnold R, Spiess M, Liu I. Goodness-of-fit and generalized estimating equation methods for ordinal responses based on the stereotype model. Stats 2022;5:507–520. [CrossRef]
[64] He Y, Luo B, Zhao L, Liao S. Influences of the COVID-19 pandemic on obesity and weight-related behaviors among Chinese children: A multi-center longitudinal study. Nutrients 2022;14:3744. [CrossRef]
[65] Tang X, Xu S, Ye H. The way to invest: Trading strategies based on ARIMA and investor personality. Symmetry 2022;14:2292. [CrossRef]
[66] Phyo PP, Byun YC, Park N. Short-term energy forecasting using machine-learning-based ensemble voting regression. Symmetry 2022;14:160. [CrossRef]
[67] Di Martino F, Sessa S. Time series seasonal analysis based on fuzzy transforms. Symmetry 2017;9:281. [CrossRef]
[68] Ye Y, Kerr WC. Alcohol and liver cirrhosis mortality in the United States: Comparison of methods for the analyses of time-series panel data models. Alcoholism 2011;35:108–115. [CrossRef]
[69] Makinde OS, Oseni BM, Adepetun AO, Olusola-Makinde OO, Abiodun GJ. The significance of daily incidence and mortality cases due to COVID-19 in some African countries. In: Data science for COVID-19. Amsterdam: Elsevier; 2022. p. 667–680. [CrossRef]
[70] Ghoshal S, Rigney G, Cheng D, Brumit R, Gee MS, Hodin RA, et al. Institutional surgical response and associated volume trends throughout the COVID-19 pandemic and postvaccination recovery period. JAMA Network Open 2022;5:e2227443. [CrossRef]
[71] Barbiellini Amidei C, Fedeli U, Gennaro N, Cestari L, Schievano E, Zorzi M, et al. Estimating overall and cause-specific excess mortality during the COVID-19 pandemic: Methodological approaches compared. Int J Environ Rese Public Health 2023;20:5941. [CrossRef]
[72] Kuo TC, Pacheco AM, Iswara AP, Dermawan D, Andhikaputra G, Chiang Hsieh LH. Sustainable ambient environment to prevent future outbreaks: How ambient environment relates to COVID-19 local transmission in Lima, Peru. Sustainability 2020;12:9277. [CrossRef]
[73] Allain ML, Collins TW. Differential access to park space based on country of origin within Miami's Hispanic/Latino population: A novel analysis of park equity. Int J Environ Rese Public Health 2021;18:8364. [CrossRef]
[74] Beard E, Marsden J, Brown J, Tombor I, Stapleton J, Michie S, et al. Understanding and using time series analyses in addiction research. Addiction 2019;114:1866–1884. [CrossRef]
[75] Wang M, Kong L, Li Z, Zhang L. Covariance estimators for generalized estimating equations (GEE) in longitudinal analysis with small samples. Statistics in Medicine 2016;35:1706–1721. [CrossRef]
[76] Simmachan T, Phaphan W. Generalization of two-sided length biased inverse Gaussian distributions and applications. Symmetry 2022;14:1965. [CrossRef]
[77] Chen R, Zhang C, Wang S, Hong L. Bivariate-dependent reliability estimation model based on inverse Gaussian processes and copulas fusing multisource information. Aerospace 2022;9:392. [CrossRef]
[78] Reyes J, Gómez-Déniz E, Gómez HW, Calderín-Ojeda E. A bimodal extension of the exponential distribution with applications in risk theory. Symmetry 2021;13:679. [CrossRef]

[79] Wang S, Gui W. Corrected maximum likelihood estimations of the lognormal distribution parameters. Symmetry 2020;12:968. [CrossRef]
[80] Liu G, Guan Q, Tang Y, Tzeng Y. Interval modeling for gamma process degradation model. Symmetry 2022;14:954. [CrossRef]
[81] Pang J, Dai J, Zhang C, Zhou H, Li Y. A new dynamic fault tree analysis method of electromagnetic brakes based on Bayesian network accompanying Wiener process. Symmetry 2022;14:968. [CrossRef]
[82] Muhammad I, Wang X, Li C, Yan M, Mukhtar M, Muhammad M. Reliability analysis with Wiener-transmuted truncated normal degradation model for linear and non-negative degradation data. Symmetry 2022;14:353. [CrossRef]
[83] Wang M. Generalized estimating equations in longitudinal data analysis: A review and recent developments. Adv Stat 2014;2014:1–11. [CrossRef]
[84] Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13–22. [CrossRef]
[85] Iyit N, Genc A. Constitution of random intercept and slope model (RISM) as a special case of linear mixed models (LMMs) for repeated measurements data. Appl Math Comput 2011;218:827–831. [CrossRef]
[86] Hilbe JM, Hardin JW. Generalized estimating equations. Boca Raton: CRC Press; 2003. [CrossRef]
[87] Hardin JW, Hilbe JM. Generalized estimating equations. 2nd ed. Boca Raton: CRC Press; 2013. [CrossRef]
[88] Akaike H. A new look at the statistical model identification. IEEE Transactions on Automat Control 1974;19:716–723. [CrossRef]
[89] Pan W. Akaike's information criterion in generalized estimating equations. Biometrics 2001;57:120–125. [CrossRef]
[90] Contreras J, Espinola R, Nogales FJ, Conejo AJ. ARIMA models to predict next-day electricity prices. IEEE Trans Power Syst 2003;18:1014–1020. [CrossRef]
[91] Box GE, Jenkins GM, Reinsel GC. Time series analysis: Forecasting and control. 3rd ed. Englewood Cliffs: Prentice Hall; 1994.
[92] Lazoğlu Ç, Büyükyazıcı M. Pricing for longevity risk in long-term care insurance. Sigma J Eng Nat Sci 2024;42:1115–1127. [CrossRef]
[93] Kim S, Lee H, Woo S, Lee H, Park J, Kim T, et al. Global, regional, and national trends in drug use disorder mortality rates across 73 countries from 1990 to 2021, with projections up to 2040: A global time-series analysis and modelling study. EClinicalMedicine 2025;79:102980. [CrossRef]
[94] Drevinskaite M, Kaceniene A, Germanavicius A, Smailyte G. Increased mortality risk in people with schizophrenia in Lithuania 2001-2020. Schizophrenia 2025;11:7. [CrossRef]
[95] Kostev K, Landré B, Yon DK, Haro JM, Gyasi RM, Hajek A, et al. Psychiatric comorbidity and in-hospital mortality in patients hospitalized for physical conditions in Germany. J Psychiatr Res 2025;181:142–149. [CrossRef]

APPENDIX A

Table 5. Forecasted values of the response variable and covariates for *Hungary* mortality data between 2014 and 2023

Response value /covariates	Model	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Mortality Rate	GEE	1134.237224	1131.861117	1127.692567	1121.301932	1116.197594	1112.469037	1107.620769	1102.771261	1097.91087	1093.814827
Certain infectious and parasitic diseases	ARIMA(2,1,2)	8.18167	8.18167	8.18167	8.18167	8.18167	8.18167	8.18167	8.18167	8.18167	8.18167
Neoplasms	ARIMA(6,3,6)	287.7983	287.6162	285.2636	281.0755	278.0613	276.1238	273.5541	270.3544	267.4938	265.0007
Leukemia	ARIMA(0,1,1)	8.211101	8.197164	8.198913	8.198693	8.198721	8.198717	8.198718	8.198718	8.198718	8.198718
Endocrine, nutritional, and metabolic diseases	ARIMA(0,1,1)	27.18967	27.43143	27.43143	27.43143	27.43143	27.43143	27.43143	27.43143	27.43143	27.43143
Diabetes mellitus	ARIMA(2,3,2)	24.09637	24.44598	24.56865	24.6117	24.6268	24.6321	24.63396	24.63461	24.63484	24.63492
Mental and behavioral diseases	ARIMA(2,2,3)	29.66334	29.66334	29.66334	29.66334	29.66334	29.66334	29.66334	29.66334	29.66334	29.66334
Diseases of the nervous system	ARIMA(0,1,2)	16.44344	16.36044	16.36044	16.36044	16.36044	16.36044	16.36044	16.36044	16.36044	16.36044
Diseases of the circulatory system	ARIMA(1,1,1)	582.9771	582.9771	582.9771	582.9771	582.9771	582.9771	582.9771	582.9771	582.9771	582.9771
Diseases of the respiratory system	ARIMA(1,1,2)	62.9001	62.54986	62.54986	62.54986	62.54986	62.54986	62.54986	62.54986	62.54986	62.54986
Diseases of the digestive system	ARIMA(0,1,2)	55.91554	56.08071	56.04987	56.05563	56.05455	56.05475	56.05472	56.05472	56.05472	56.05472
Diseases of the skin	ARIMA(0,1,1)	0.8391805	0.788182	0.8391534	0.8448604	0.8806364	0.8993464	0.9275341	0.9494243	0.9741344	0.9963186
Diseases of the musculoskeletal system	ARIMA(1,2,3)	3.884309	3.825018	3.844091	3.837955	3.839929	3.839294	3.839499	3.839433	3.839454	3.839447
Diseases of the genitourinary system	ARIMA(1,1,1)	8.02746	8.177434	8.024136	7.853654	7.775081	7.785886	7.596901	7.516759	7.454422	7.370989
Certain conditions originating in the perinatal period	ARIMA(0,1,1)	4.003069	4.003069	4.003069	4.003069	4.003069	4.003069	4.003069	4.003069	4.003069	4.003069
Abnormalities	ARIMA(1,1,1)	3.096544	3.037492	3.067402	3.052252	3.059925	3.056039	3.058008	3.05701	3.057515	3.05726
Symptoms, signs, ill-defined causes	ARIMA(0,1,1)	1.66774	1.710745	1.738045	1.755377	1.76638	1.773365	1.777799	1.780614	1.782401	1.783536
External causes of mortality	ARIMA(1,1,2)	54.89527	52.47831	50.14605	47.89553	45.72386	43.6283	41.60618	39.65492	37.77203	35.95513
Accidental falls	ARIMA(2,1,3)	12.52546	11.705115	9.928364	9.955383	10.829352	10.400187	12.346784	11.313417	13.146898	12.542399

Table 6. Forecasted values of the response variable and covariates for *Japan* mortality data between 2014 and 2023

Response value /covariates	Model	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Mortality Rate	GEE	598.1090244	591.7269195	586.2644384	586.51027	582.5780727	578.5410242	575.9700875	573.670827	568.9321866	566.2642829
Certain infectious and parasitic diseases	ARIMA(0,1,1)	12.09728	12.09728	12.09728	12.09728	12.09728	12.09728	12.09728	12.09728	12.09728	12.09728
Neoplasms	ARIMA(0,1,1)	183.763	183.763	183.763	183.763	183.763	183.763	183.763	183.763	183.763	183.763
Leukemia	ARIMA(1,1,2)	4.01639	3.987458	3.9835	3.982959	3.982884	3.982874	3.982873	3.982873	3.982873	3.982873
Endocrine, nutritional, and metabolic diseases	ARIMA(0,1,1)	10.18505	10.18505	10.18505	10.18505	10.18505	10.18505	10.18505	10.18505	10.18505	10.18505
Diabetes mellitus	ARIMA(0,1,1)	6.60856	6.60856	6.60856	6.60856	6.60856	6.60856	6.60856	6.60856	6.60856	6.60856
Mental and behavioral diseases	ARIMA(1,1,1)	5.421231	5.625852	5.815109	5.990156	6.152059	6.301806	6.44031	6.568414	6.6869	6.796489
Diseases of the nervous system	ARIMA(2,1,1)	13.23144	13.76666	14.14424	14.50498	14.77088	15.0154	15.2015	15.36796	15.49764	15.61131
Diseases of the circulatory system	ARIMA(2,2,3)	153.6278	148.2028	142.9929	138.4608	132.6203	128.4013	122.5603	118.083	112.6731	107.6904
Diseases of the respiratory system	ARIMA(1,1,1)	89.47145	88.99877	89.28839	89.11093	89.21967	89.15304	89.19386	89.16885	89.18418	89.17479
Diseases of the digestive system	ARIMA(5,1,4)	24.72248	24.84109	23.63623	22.76675	22.15488	21.24112	21.50572	20.80842	20.01579	19.8482
Diseases of the skin	ARIMA(1,1,4)	0.7262154	0.7456651	0.7436923	0.7439343	0.7439505	0.7439516	0.7439516	0.7439516	0.7439516	0.7439516
Diseases of the musculoskeletal system	ARIMA(0,2,1)	2.783871	2.767743	2.751614	2.735485	2.719357	2.703228	2.687099	2.670971	2.654842	2.638713
Diseases of the genitourinary system	ARIMA(2,2,3)	16.10558	15.83325	15.95852	15.79933	15.79108	15.77212	15.63046	15.72595	15.49848	15.6447
Certain conditions originating in the perinatal period	ARIMA(4,3,1)	0.7037849	0.6936113	0.7005203	0.6870167	0.7211691	0.7594637	0.775698	0.8191778	0.8695761	0.9302319
Abnormalities	ARIMA(2,1,1)	1.823158	1.769412	1.708746	1.650174	1.59099	1.532006	1.472978	1.413979	1.354988	1.296011
Symptoms, signs, ill-defined causes	ARIMA(2,1,2)	39.6575	41.93718	44.01135	46.10904	48.11104	50.08272	51.99153	53.85679	55.67026	57.43826
External causes of mortality	ARIMA(2,1,1)	50.97363	47.42804	44.69288	48.06256	47.77171	46.09374	47.13688	47.49862	46.74088	46.94396
Accidental falls	ARIMA(0,1,1)	3.697598	3.697598	3.697598	3.697598	3.697598	3.697598	3.697598	3.697598	3.697598	3.697598

Table 7. Forecasted values of the response variable and covariates for *Netherlands* mortality data between 2014 and 2023

Response value /covariates	Model	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Mortality Rate	GEE	772.3442577	776.1473402	779.745686	785.2032783	792.9899944	798.1636296	803.5354028	808.2742824	813.7379163	818.8284283
Certain infectious and parasitic diseases	ARIMA(1,1,2)	17.64273	16.9376	17.38356	17.10151	17.2799	17.16708	17.23843	17.1933	17.22184	17.20379
Neoplasms	ARIMA(1,1,1)	231.9516	229.8281	227.7292	225.6546	223.6041	221.5772	219.5739	217.5937	215.5937	213.7019
Leukemia	ARIMA(1,1,1)	6.958206	6.95847	6.95173	6.954149	9.954151	9.954151	9.954151	9.954151	9.954151	9.954151
Endocrine, nutritional, and metabolic diseases	ARIMA(1,2,1)	20.32549	19.90898	19.40565	18.87117	18.32551	17.77585	17.22474	16.67312	16.12131	15.56944
Diabetes mellitus	ARIMA(0,1,1)	15.64643	15.64643	15.64643	15.64643	15.64643	15.64643	15.64643	15.64643	15.64643	15.64643
Mental and behavioral diseases	ARIMA(1,3,1)	64.99008	73.48194	81.87719	90.63776	99.56939	108.75378	118.15657	127.79221	137.65461	147.74635
Diseases of the nervous system	ARIMA(0,1,1)	36.83629	36.83629	36.83629	36.83629	36.83629	36.83629	36.83629	36.83629	36.83629	36.83629
Diseases of the circulatory system	ARIMA(0,1,1)	207.5178	207.5178	207.5178	207.5178	207.5178	207.5178	207.5178	207.5178	207.5178	207.5178
Diseases of the respiratory system	ARIMA(2,1,0)	74.13148	73.19441	71.74643	72.57098	72.46802	72.62341	72.55384	72.55384	72.55545	72.57324
Diseases of the digestive system	ARIMA(2,2,2)	23.6786	22.39587	20.57511	18.93132	17.7313	16.67244	15.4118	13.94587	12.49097	11.17309
Diseases of the skin	ARIMA(1,1,3)	1.7949	1.69766	1.903796	1.896296	1.896569	1.896559	1.896559	1.896559	1.896559	1.896559
Diseases of the musculoskeletal system	ARIMA(4,3,4)	5.652348	6.055727	5.956917	6.374092	6.61583	6.898678	7.463178	7.706152	8.448382	8.82375
Diseases of the genitourinary system	ARIMA(2,3,2)	16.701026	15.341209	14.608605	13.323503	11.984974	10.346098	8.517123	6.445432	4.155066	1.634724
Certain conditions originating in the perinatal period	ARIMA(1,1,1)	2.542847	2.524539	2.532362	2.529019	2.530447	2.529837	2.530098	2.529986	2.530034	2.530014
Abnormalities	ARIMA(2,1,2)	2.439768	2.372716	2.306265	2.240409	2.175142	2.11046	2.046358	1.982829	1.919869	1.857473
Symptoms, signs, ill-defined causes	ARIMA(0,1,2)	30.98058	31.70591	31.70591	31.70591	31.70591	31.70591	31.70591	31.70591	31.70591	31.70591
External causes of mortality	ARIMA(1,1,2)	36.32994	36.75035	37.15479	37.54387	37.91819	38.27829	38.62473	38.95801	39.27864	39.5871
Accidental falls	ARIMA(0,1,1)	11.92262	11.92262	11.92262	11.92262	11.92262	11.92262	11.92262	11.92262	11.92262	11.92262

Table 8. Forecasted values of the response variable and covariates for *Norway* mortality data between 2014 and 2023

Response value /covariates	Model	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Mortality Rate	GEE	738.2035136	731.0043563	726.6927904	719.3790803	715.32947	709.9112773	704.7410678	700.8503869	696.0619818	691.6829106
Certain infectious and parasitic diseases	ARIMA(1,1,2)	17.96959	18.08191	18.18281	18.27343	18.35483	18.42795	18.49363	18.55262	18.60561	18.6532
Neoplasms	ARIMA(0,1,2)	202.003	201.4331	201.4331	201.4331	201.4331	201.4331	201.4331	201.4331	201.4331	201.4331
Leukemia	ARIMA(2,1,1)	6.229325	6.229325	6.229325	6.229325	6.229325	6.229325	6.229325	6.229325	6.229325	6.229325
Endocrine, nutritional, and metabolic diseases	ARIMA(3,1,2)	18.15323	17.91546	18.87881	18.1114	18.34784	18.03459	18.49784	18.21473	18.367	18.16121
Diabetes mellitus	ARIMA(3,1,0)	12.31501	12.42031	12.83848	12.46218	12.44489	12.39903	12.56184	12.52117	12.50807	12.46314
Mental and behavioral diseases	ARIMA(0,1,1)	37.76095	37.76095	37.76095	37.76095	37.76095	37.76095	37.76095	37.76095	37.76095	37.76095
Diseases of the nervous system	ARIMA(1,1,2)	32.19028	31.64012	31.91349	31.77765	31.84515	31.81161	31.82827	31.81999	31.82411	31.82206
Diseases of the circulatory system	ARIMA(1,1,2)	209.4741	201.8981	194.6513	187.7196	181.0892	174.7471	168.6806	162.8779	157.3274	152.0182
Diseases of the respiratory system	ARIMA(1,1,3)	74.3872	73.75988	74.20734	74.11859	74.13619	74.1327	74.13339	74.13326	74.13328	74.13328
Diseases of the digestive system	ARIMA(1,1,1)	22.20327	21.8269	22.10837	21.89788	22.05529	21.93757	22.02561	21.95977	22.00901	21.97218
Diseases of the skin	ARIMA(2,2,1)	1.838686	1.877894	1.898826	1.924789	1.951737	1.977753	2.0039	2.030133	2.056324	2.082516
Diseases of the musculoskeletal system	ARIMA(0,1,1)	4.469828	4.469828	4.469828	4.469828	4.469828	4.469828	4.469828	4.469828	4.469828	4.469828
Diseases of the genitourinary system	ARIMA(2,1,2)	12.8594	13.63117	13.46832	12.89368	13.72169	13.72169	12.96538	13.77598	13.19541	13.06711
Certain conditions originating in the perinatal period	ARIMA(1,1,2)	1.593643	1.7527	1.785451	1.792195	1.793584	1.79387	1.793929	1.793941	1.793944	1.793944
Abnormalities	ARIMA(1,1,1)	2.680158	2.678351	2.678186	2.678171	2.67817	2.67817	2.67817	2.67817	2.67817	2.67817
Symptoms, signs, ill-defined causes	ARIMA(1,2,1)	56.2911	57.34902	58.40575	59.46246	60.51917	61.57588	62.63259	63.6893	64.74601	65.80273
External causes of mortality	ARIMA(0,1,1)	47.57908	47.57908	47.57908	47.57908	47.57908	47.57908	47.57908	47.57908	47.57908	47.57908
Accidental falls	ARIMA(3,2,1)	8.27993	7.616328	6.981984	6.425969	5.884434	5.3433006	4.794065	4.243791	3.693202	3.143343

Table 9. Forecasted values of the response variable and covariates for *Spain* mortality data between 2014 and 2023

Response value/covariates	Model	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Mortality Rate	GEE	638.6493938	622.657563	601.9661482	586.2931227	566.3348631	551.8485077	531.375311	516.8450135	497.6533035	485.122471
Certain infectious and parasitic diseases	ARIMA(2,2,1)	10.311942	10.127921	10.020496	9.933205	9.848931	9.764251	9.679058	9.593635	9.5088141	9.422633
Neoplasms	ARIMA(3,2,0)	196.0516	194.0738	191.262	188.6272	186.5909	184.2394	181.7304	179.3231	177.0442	174.6504
Leukemia	ARIMA(3,2,1)	6.290295	6.207335	6.270056	6.200342	6.313735	6.245593	6.300835	6.292409	6.342321	6.308841
Endocrine, nutritional, and metabolic diseases	ARIMA(2,2,1)	19.11787	18.31603	17.71245	17.19611	16.75688	16.36489	16.00747	15.67316	15.35502	15.04792
Diabetes mellitus	ARIMA(1,1,2)	15.027587	14.238337	13.504495	12.822171	12.187748	11.597865	11.049393	10.539426	10.06526	9.626383
Mental and behavioral diseases	ARIMA(1,2,1)	26.39477	26.37322	26.36138	26.34872	26.33613	26.32354	26.31094	26.29835	26.28576	26.27316
Diseases of the nervous system	ARIMA(3,2,2)	36.25628	35.3553	35.29499	34.28266	34.15837	33.81788	33.67263	33.16858	32.75181	32.34055
Diseases of the circulatory system	ARIMA(3,1,2)	180.76225	170.425777	158.91489	149.29143	137.47192	128.00812	116.14285	106.71385	94.86202	85.43977
Diseases of the respiratory system	ARIMA(2,1,3)	69.93665	70.79244	67.91649	68.77265	65.89807	66.75459	63.88139	64.73828	61.86645	62.7237
Diseases of the digestive system	ARIMA(1,1,1)	31.90935	31.01876	30.12822	29.23774	28.34731	27.45694	26.56662	25.67636	24.78615	23.896
Diseases of the skin	ARIMA(0,1,1)	2.001312	2.001312	2.001312	2.001312	2.001312	2.001312	2.001312	2.001312	2.001312	2.001312
Diseases of the musculoskeletal system	ARIMA(0,1,1)	4.896652	4.896652	4.896652	4.896652	4.896652	4.896652	4.896652	4.896652	4.896652	4.896652
Diseases of the genitourinary system	ARIMA(1,1,0)	18.70061	18.65589	18.66586	18.66364	18.66413	18.66402	18.66405	18.66404	18.66404	18.66404
Certain conditions originating in the perinatal period	ARIMA(2,2,1)	2.054246	2.069544	1.991298	1.928151	1.898652	1.847743	1.791806	1.747327	1.699942	1.648903
Abnormalities	ARIMA(1,1,0)	1.791302	1.790545	1.790479	1.790473	1.790473	1.790473	1.790473	1.790473	1.790473	1.790473
Symptoms, signs, ill-defined causes	ARIMA(2,2,1)	13.5983839	11.702819	9.8033625	8.0410654	6.1975613	4.3652741	2.5484344	0.7192763	0.106717	0.09312447
External causes of mortality	ARIMA(0,1,1)	27.16536	27.16536	27.16536	27.16536	27.16536	27.16536	27.16536	27.16536	27.16536	27.16536
Accidental falls	ARIMA(2,2,1)	4.677224	4.698629	4.751375	4.876108	5.023152	5.150917	5.260655	5.37001	5.486586	5.606643

Table 10. Forecasted values of the response variable and covariates for *USA* mortality data between 2014 and 2023

Response value/covariates	Model	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Mortality Rate	GEE	775.7973354	764.3866416	754.2418892	744.3455617	734.5577981	725.1245957	715.8845212	706.2853494	696.8382873	687.4465377
Certain infectious and parasitic diseases	ARIMA(1,1,2)	23.44106	23.53848	23.44237	23.53719	23.44364	23.53593	23.44488	23.5347	23.44609	23.53351
Neoplasms	ARIMA(2,1,3)	202.2696	202.4181	203.0351	204.0499	205.3837	206.954	208.6774	210.4735	212.2669	213.9903
Leukemia	ARIMA(0,1,1)	7.70504	7.70504	7.70504	7.70504	7.70504	7.70504	7.70504	7.70504	7.70504	7.70504
Endocrine, nutritional, and metabolic diseases	ARIMA(1,1,2)	34.27066	34.27968	34.28336	34.28485	34.28546	34.28571	34.28581	34.28586	34.28587	34.28588
Diabetes mellitus	ARIMA(5,2,4)	24.04607	24.11972	24.72169	25.29889	25.77109	26.221	26.43792	26.62869	26.93324	27.25523
Mental and behavioral diseases	ARIMA(1,1,1)	46.98619	48.50255	49.99933	51.47678	52.93515	54.37469	55.79565	57.19825	58.58275	59.94937
Diseases of the nervous system	ARIMA(0,1,1)	51.14834	51.14834	51.14834	51.14834	51.14834	51.14834	51.14834	51.14834	51.14834	51.14834
Diseases of the circulatory system	ARIMA(1,1,1)	208.51236	196.2852	184.05852	171.83231	159.60659	147.38134	135.15656	122.93227	110.70845	98.48511
Diseases of the respiratory system	ARIMA(1,1,3)	82.55506	82.09725	82.38406	82.39847	82.39923	82.39923	82.39923	82.39923	82.39923	82.39923
Diseases of the digestive system	ARIMA(1,1,1)	29.10215	28.73568	28.37056	28.0068	27.6444	27.28335	26.92363	26.56526	26.20822	25.85251
Diseases of the skin	ARIMA(2,2,1)	1.450314	1.451351	1.446895	1.442056	1.435623	1.42891	1.421716	1.414391	1.406917	1.399389
Diseases of the musculoskeletal system	ARIMA(3,2,1)	4.81154	4.855233	4.90075	4.946764	4.993081	5.039499	5.08597	5.132461	5.17896	5.225464
Diseases of the genitourinary system	ARIMA(1,2,2)	22.15151	22.21051	22.28587	22.37097	22.46187	22.55622	22.65262	22.75025	22.8486	22.94739
Certain conditions originating in the perinatal period	ARIMA(2,2,1)	3.059989	2.911454	2.762838	2.614166	2.465489	2.316812	2.316836	2.019459	1.870783	1.722106
Abnormalities	ARIMA(1,1,1)	2.64287	2.565754	2.488652	2.411564	2.33449	2.257429	2.180383	2.10335	2.026331	1.949327
Symptoms, signs, ill-defined causes	ARIMA(1,1,1)	12.01776	12.00926	12.00702	12.00643	12.00628	12.00624	12.00623	12.00622	12.00622	12.00622
External causes of mortality	ARIMA(0,1,2)	57.41402	57.37231	57.37231	57.37231	57.37231	57.37231	57.37231	57.37231	57.37231	57.37231
Accidental falls	ARIMA(1,2,1)	7.914455	7.922515	7.942832	7.96595	7.98971	8.013616	8.037555	8.061502	8.085451	8.109401