Available online at http://scik.org

Commun. Math. Biol. Neurosci. 2021, 2021:67

https://doi.org/10.28919/cmbn/5926

ISSN: 2052-2541

# MODELING OF PREMATURE MORTALITY RATES FROM CHRONIC DISEASES IN EUROPE, INVESTIGATION OF CORRELATIONS, CLUSTERING AND GRANGER CAUSALITY

VASILEIOS PAPAGEORGIOU, GEORGE TSAKLIDIS*

Department of Mathematics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece



**Abstract:** In this study mortality rates due to chronic diseases in 31 European countries are examined, through certain time series modeling, based on the electronic database of Eurostat. The time series are grouped into 6 clusters using Fuzzy K-means method combined with Jensen-Shannon divergence enabling a more accurate investigation of the homogeneity within clusters. The intuitive interpretation of these results relies on studies concerning the differentiation of Europeans' habits due to geopolitical position, nutritional, sociopolitical and environmental factors. Spearman coefficient, distance correlation and cross-correlation are used to reveal the underlying correlations while causality is checked through a parametric test, aiming to disclose the existence of Granger causality by examining all the pairs of the respective time series of premature mortality rates. Correlation indices and causality tests, highlight many statistically significant relationships with respect to their dynamics, comprehending the relationships governing the living standards of European countries. Finally, Europe's average mortality is modelled while special emphasis is given to the emergence of the ideal representative of Europe's rates. A detailed investigation of the relationships and patterns that characterize premature mortality from chronic diseases across Europe is provided, examining each

*Corresponding author

E-mail address: tsaklidi@math.auth.gr

Received April 25, 2021

country either individually or in connection to the other countries, while each statistically generated outcome is supported by results concerning the economic status, environment, nutrition and especially status of public health.

**Keywords:** linear time series analysis; SETAR models; fuzzy K-means clustering; correlation indices; granger causality.

**2010 AMS Subject Classification:** 92C60, 62M10, 62P10, 62H30.

## 1. INTRODUCTION

The field of health and human mortality is an important research area, where the exported results and observations can play a decisive role in improving public health conditions aiming at a more targeted treatment of various diseases and illnesses encountered in nowadays while constantly increasing the average life expectancy. The usage of real and trustworthy data combined with proper modeling and effective highlighting of various types of relationships between observations may have a decisive role in providing important findings for utilization and further research.

In what follows we examine the dynamics and relations between time series describing the mortality rates caused by chronic diseases per 100,000 inhabitants under the age of 65 in 31 European countries, 27 of which belong to the European Union. Many studies have been carried out concerning time series of mortality rates using a variety of mathematical and statistical tools, while in cited references some examples of such research articles are mentioned. In Valea et al. [1] the cardiovascular mortality rates in southern Romania are studied; Yilan et al. [2] uses ARIMA methodology to model mortality rates from injuries in China, while Adeyinka and Muhajarine [3] and Bisong et al. [4], utilize ARIMA models to study the evolution of mortality rates along with neural networks or Kalman filters, respectively. In Megyesiova and Lieskovska [5], the differentiation in Europe's premature mortality between men and women is examined while highlighting the descending trend that characterizes premature mortality rates in the entire EU.

Our analysis can be divided in two parts: The first part focuses on modeling the rates of premature mortality in 31 European countries using linear and nonlinear time series methodologies.

It is noticed, that the majority of the time series are consisted of only 24 consecutive observations, constraining the utilization of complex time series models that would require the estimation of many unknown parameters and the invocation of many asymptotical properties. As a result, special attention is given to linear methodology, which is characterized by more "economic" models like autoregressive integrated moving average models (ARIMA), and random walk.

Moreover, in some time series where the fitting of any ARIMA model seems to be unsatisfactory, we turned our attention to nonlinear methodology and more specifically to low order self–exciting threshold autoregressive models (SETAR), which are adequate to describe more complicated relations between the observations of the time series. The short length of the examined time series does not allow the utilization of complex nonlinear statistical tests and tools, like correlation dimension, false nearest neighbors or Lyapunov exponents, aiming to uncover the hidden elements governing the dynamics of these time series.

In some time series where values are missing, we applied spline interpolation to supplement these values (Pratama et al. [6]). However, we have avoided extrapolation methods to increase the length of time series, as these methods involve larger errors especially when the length of the initial time series is relatively short.

The second part, is dedicated to the investigation of a variety of relations between the 31 aforementioned time series of premature mortality rates. These attempts of revealing relations among different mortality levels include clustering algorithms, correlation indices and causality tests. More specifically, we utilized the fuzzy K – means algorithm combined with the Jensen – Shannon divergence aiming to cluster the 31 European countries and measure the differentiation of mortality rates between countries of the same cluster.

Furthermore, we provide certain correlation indices such as Spearman coefficient, distance correlation and cross–correlation for different lags among all 465 pairs of time series, and we investigate causal relations between time series by examining the (Instantaneous) Granger Causality. Checking the existence of Granger causality between time series allows us to examine whether the dynamics of each time series determines the evolution of other time series included in

our research.

Of particular interest are the methods of investigating Granger causality between time series using conditional (multivariate) Granger causality (Papana et al. [7]), frequency domain-based Granger causality (Siggiridou et al. [8]) and non-linear Granger causality (Kugiumtzis [9]). In Siggiridou et al. [10], a comparative study of the suitability of the aforementioned methods in multivariate time series is presented. One of the main disadvantages of simple (bivariate) Granger causality over conditional causality, is that the simple version cannot separate the causality case (a) $x \rightarrow y$, $x \rightarrow z$ and $z \rightarrow y$, from case (b) $x \rightarrow z$ and $z \rightarrow y$. As a result, conditional Granger causality reveals only direct causal relationships between time series but not relationships that may emerge due to a third time series. In the present work, we do not examine causes using any of the 3 aforementioned controls because of the short length of the time series.

The mean mortality rates from chronic diseases in Europe are depicted in a new time series. This time series is created by considering the mean value of the mortality rates of all 31 examined countries at each time point without taking into account incompatibilities due to missing values in some of the time points. Doubtlessly, both the modeling and the exploration of causal and non-causal relations provide interesting results concerning the health conditions in this continent.

The paper is organized as follows: In section 2 we present the data sources we used. In section 3 we outline the utilized methods and mathematical tools in linear and non-linear analysis, clustering and correlation-causality investigation. In section 4 we provide the results concerning the model selection for the mortality rates, clustering and homogeneity examination within the clusters, correlations and Granger causality. Finally, in section 5 we discuss and summarize the derived results, paying special attention to the explanation of the general descending trend that characterizes the evolution of the mortality rates.

## 2. DATA

In this study, we utilized time series data of chronic disease mortality rates per 100,000 inhabitants under the age of 65 in 31 European countries, 27 of which belong to the European Union. The

datasets used, belong to the open access, electronic database of Eurostat, which can be visited by interested readers through its website. More specifically, the database used in our analysis includes observations contained in two databases with codes: hlth_cd_asdr and sdg_03_40. The observations of both of these databases are organized in tables, where each row corresponds to a different European country and each column to a different calendar year. The first database, coded as sdg_03_40 and entitled "Standardized death rate due to chronic diseases by sex", contains observations of mortality rates from chronic diseases per 100,000 people under the age of 65 for 17 consecutive years and more specifically rates from 2000 to 2017. In addition, these percentages include the cases that correspond to 6 chronic diseases. The 6 selected diseases are malignant neoplasms, diabetes mellitus, ischemic heart and cerebrovascular diseases, chronic diseases of the lower respiratory system and liver. On the other hand, the database with code hlth_cd_asdr and title "Causes of death - standardized death rate" includes mortality rates from 1994 to 2010 for the European countries. These rates are the cumulative rates of a much wider variety of chronic diseases. However, through Eurostat ‘s electronic system, it is possible to select only the 6 diseases that appear in database sdg_03_40 producing a valid combination of these datasets.

By merging these two databases, we obtained the mortality rates in European countries from 1994 to 2017, a total of 24 years. Although, this fact does not imply that all 31 time series consist of 24 observations. The overlap that exists to the time periods covered by these two databases (2000 - 2010), allowed us to check their compatibility avoiding controversial conclusions. In addition, Eurostat ‘s electronic system allows the examination of mortality rates of men and women separately, but we preferred to study the behavior of these two categories as a whole. Finally, in the merged databases, there were time series of mortality rates for countries such as Liechtenstein, Turkey, Albania, Serbia and Northern Macedonia, but their length was quite short and their processing was beyond the scopes of this article. A really important element for our analysis is the short length of the examined time series, which should be taken into account in order to avoid precarious results.

## 3. METHODS AND MATHEMATICAL TOOLS FOR THE ANALYSIS OF MORTALITY RATES

### 3.1 Linear Time Series Models

Throughout this paragraph we outline the mathematical tools utilized for the upcoming investigation of the linear dynamics determining the progress of mortality rates in the 31 European countries examined in our article. Initially, our approach will be based on ARIMA models; during the time series analysis we start by investigating the linear methodology before applying any nonlinear models. Apparently, the short length of the times series requires the fitting of low complexity statistical models, making ARIMA models an ideal choice. The selection of the best ARIMA model for each of the 31 time series was made by taking into account certain statistical criteria like autocorrelation and partial autocorrelation function, portmanteau test, Akaike information criterion (AIC) and the value of the NRMSE for various ratios of training–testing sets.

#### 3.1.1. Utilization of ARIMA models

After subtracting the trend of the time series, the dynamics of system $X_t$ with ARMA models (p,q) can be studied, where p denotes the order of the autoregressive (AR) part and q denotes the order of the moving average model (MA). The general form of an ARMA (p,q) model is given as

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t. \quad (1)$$

Special cases of ARMA models are the models AR(p) which consist only of p autoregressive terms and the models MA(q) which consist of q terms of moving average. Then, the ARIMA(p,d,q) model is defined, that describes the dynamics of a non-stationary time series. An ARIMA(p,1,q) model can be written as

$$Y_t = c + Y_{t-1} + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t. \quad (2)$$

The choice of the order of an ARIMA(p,d,q) model relies on a series of criteria which will be briefly presented below. There is a wide variety of statistical methods, aiming to model time series characterized by linear relationships; indicatively, we quote the articles Yilan et al. [2], Bisong et al. [4], Fattah et al. [11] and Atance et al. [12]. In case that the fitting of various ARIMA models is not satisfactory, the stationary time series $X_t$ can be deemed as white noise with variance $\sigma_\varepsilon^2$ while the corresponding non-stationary $Y_t$ can be deemed as random walk.

### 3.2. ARMA model selection criteria

In this paragraph, we will describe briefly ARMA models to examine the mortality rates of the 31 countries contained in the dataset and provide statistical criteria for the selection of the models. These criteria are useful for the case of linear analysis, while criteria like AIC and NRMSE can be used in parallel with nonlinear methodologies.

#### 3.2.1. Autocorrelation and partial autocorrelation function

Let $\{x_1, x_2, \dots, x_n\}$, be a realization of the stochastic process $X_t$. We can examine the dynamics of the linear relations of the time series through the autocorrelation function ρ(τ). Our goal is to calculate the autocorrelation of the time series for various lags $\tau$, using the function $r_\tau = \hat{\rho}(\tau)$ which is the quotient of the estimated autocovariance of lag τ ( *c(τ)* ) and the variance ($s_X^2 = c(0)$). Thus, we can estimate the autocorrelations of $X_t$ by means of the formula

$$r_\tau = \frac{c(\tau)}{s_X^2} = \frac{\frac{1}{n-\tau}\sum_{t=\tau+1}^{n}(x_t x_{t-\tau} - \bar{x}^2)}{\frac{1}{n}\sum_{t=1}^{n}(x_t^2 - \bar{x}^2)}, \quad (3)$$

where $\bar{x}$ denotes the unbiased estimator of theoretical mean value, $\mu$, of the time series. Usually,

the way that autocorrelations decline, indicates the existence of an MA(q) model, where q denotes the lag after which the autocorrelations become insignificant.

The autocorrelation function presented in the previous paragraph decreases exponentially in AR(p) procedures. The random variables $X_t$ and $X_{t-\tau}$ for $\tau > p$, are correlated through the intermediate random variables $X_{t-1}, \dots, X_{t-\tau-1}$. One measure that examines the correlation explicitly between $X_t$ and $X_{t-\tau}$ is the partial autocorrelation function, defined as $\varphi_{\tau,\tau} = Cor(X_t, X_{t-\tau} | X_{t-1}, \dots, X_{t-\tau-1})$. We can express the partial autocorrelation function through the autocovariance of lag $\tau$ and the Yule – Walker equations as

$$\varphi_{\tau,\tau} = \frac{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{\tau-2} & \rho_1 \\ \rho_1 & 1 & \dots & \rho_{\tau-3} & \rho_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{\tau-1} & \rho_{\tau-2} & \dots & \rho_1 & \rho_\tau \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{\tau-2} & \rho_{\tau-1} \\ \rho_1 & 1 & \dots & \rho_{\tau-3} & \rho_{\tau-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{\tau-1} & \rho_{\tau-2} & \dots & \rho_1 & 1 \end{vmatrix}}. \quad (4)$$

On the other hand, the theoretical partial autocorrelation function becomes statistically insignificant for $\tau > p$ in the AR(p) procedures, thus contributing to an efficient model selection. The partial autocorrelations of a time series can also be calculated using the Durbin – Levinson iterative algorithm.

### 3.2.2. Portmanteau test

The first step before fitting an ARMA model to a stationary time series is testing the independence between its values at different lags. Independence between $X_{t-\tau}$ for $\tau = 1, \dots, \frac{n}{2}$, leads to the conclusion that the examined time series is a white noise. The short length of the time series does not strongly recommend the use of the standard autocorrelation criterion of $r_\tau \notin [-\frac{2}{\sqrt{n}}, \frac{2}{\sqrt{n}}]$.

As a result, we culminate in testing the independence of time series according to Portmanteau

test which gives reliable results for small samples and allows us to examine $l$ autocorrelations at the same time. In our analysis we will utilize the correction proposed by Ljung and Box, where

$$Q = n(n+2)\sum_{\tau=1}^{l}\frac{r_\tau^2}{(n-\tau)}, \quad (5)$$

follows the $\chi^2$ distribution with $l$ degrees of freedom. The null hypothesis which states that the time series is white noise, is rejected when $Q > \chi^2_{l;1-a}$ where $\alpha$ represents the confidence level of the test.

### 3.2.3. Akaike information criterion (AIC)

There are criteria to determine the order of an ARMA(p,q), which are based on the likelihood of the data based on this model, where the variance of the residuals $\sigma_\varepsilon^2$ - considered as a likelihood indicator - is produced after fitting the model to the time series. A widely used criterion based on likelihood is the Akaike Information Criterion, defined as

$$AIC(p,q) = \ln(\sigma_\varepsilon^2) + \frac{2(p+q+1)}{n}, \quad (6)$$

which involves the variance of the residuals, the order of the model, as well as the length of the time series $n$. The value of AIC decreases as $\sigma_\varepsilon^2$ decreases, and increases as the order of ARMA(p,q) increases. As a result, we choose *p* and *q* to be the values that minimize AIC. In cases where the AIC-values for various *p* and *q* are close, we employ additional criteria for the optimal order selection.

### 3.2.4. Normalized root mean squared error (NRMSE)

An important criterion for selecting a model is the assessment of its predictive capacity. For this purpose, we divide our time series into two parts, the training set and the testing set. The training set is used to adjust and estimate the parameters of the model, while the testing set is used to

estimate its predictive power through an error function. The NRMSE function is suitable for this approach, defined as

$$NRMSE(l) = \frac{\sqrt{\frac{1}{k-l+1}\sum_{j=n}^{n+k-l}\left(x_{j+l} - x_j(l)\right)^2}}{\sqrt{\frac{1}{k-l+1}\sum_{j=n}^{n+l-k}\left(x_{j+l} - \bar{x}\right)^2}}. \quad (7)$$

In (7), $l$ denotes a prediction of $l$ steps ahead, for a series of known observations at time points $n+1, n+2, \ldots, n+k$.

The closer the value of NRMSE is to 0, the better the predictive power of the fitted ARMA model. For $\text{NRMSE} > 1$ the forecasts given by the selected model are worse than those of the historic mean. During model selection for each of the 31 time series of mortality rates from chronic diseases, we derived predictions for $l = 1$, using as training set the first 60, 70 and 80%.

### 3.3. Modeling mortality from chronic diseases according to non-linear threshold models

As provided in the previous section, we fitted ARIMA models to 19 out of the 32 examined times series. For the remaining 13 time series, the adaptation of any ARIMA models was not suggested. Since, ARIMA models failed to describe the dynamics that rules the evolution of these time series, the next step was the employment of nonlinear methodologies. The field of nonlinear time series analysis contains a variety of methods that utilize different measures to capture model's dynamics, such as the mutual information function, correlation dimension, approximate entropy, Lyapunov exponents and others. However, all these measures require long time series in order to produce reliable results. Inevitably, we will limit our analysis to the usage of local linear models (Shi [13] and Bruin and Gooijer [14]). Typical instances of this category of nonlinear models are the self-exciting threshold autoregressive models, which are often referred to in literature as SETAR.

#### 3.3.1. Self - exciting threshold autoregressive models (SETAR)

The basic idea for SETAR models relies on the usage of different AR models for each of the *k*

segments of SETAR, while the selection of each segment is being accomplished according to the values of the time series itself at each time point $t$. Let AR($p$) be an autoregressive model with $k$ different sets of values for the estimated parameters $\varphi_{j,0}, \dots, \varphi_{j,p}$ corresponding to $k$ subsets of $\mathbb{R}^p$.

The $k$ subsets constitute a partition of $\mathbb{R}^p$, a partition obtained through the values of the component $Z_t$, while the determination of the values of $\varphi_{j,0}, \dots, \varphi_{j,p}$ depends on the subset of $\mathbb{R}$, to which $Z_t$ belongs each time point $t$. Parameter $d$ plays the role of the lag and in our case it takes the value $d = 1$. For a set of partition values of $\mathbb{R}$, let $\{x_0, x_1, \dots, x_k\}$, with $-\infty = x_0 < x_1 < \cdots < x_k = \infty$; this partition will be defined as $\mathbb{R} = R_1 \cup R_2 \cup \dots \cup R_k$ with $R_i = (x_{i-1}, x_i]$, $i = 1, \dots, k$. The general form that describes a SETAR model consisting of two AR parts can be given as

$$\begin{aligned} X_t = {} & \left(\varphi_{1,0} + \varphi_{1,1}X_{t-d} + \varphi_{1,2}X_{t-2d} + \cdots + \varphi_{1,L}X_{t-(L-1)d}\right) I(Z_t \le th) \\ & + \left(\varphi_{2,0} + \varphi_{2,1}X_{t-d} + \varphi_{2,2}X_{t-2d} + \cdots + \varphi_{2,H}X_{t-(H-1)d}\right) I(Z_t > th) \\ & + \varepsilon_t, \end{aligned} \tag{8}$$

where $th$ denotes the threshold's value chosen for the partition of $\mathbb{R}$, while $L$ is the order of the AR model corresponding to the part of SETAR when $Z_t \le th$, $H$ stands for the order of the AR model when $Z_t > th$ while $Z_t$ represents one of $X_{t-d}, X_{t-2d}, \dots$, chosen based on an information criterion. We denote by $I$ the index function taking the value 1 when the condition in parentheses is satisfied, otherwise it equals 0, while $m$ stands for the embedding dimension of the reconstructed time series, or the maximum order of the two parts (Firat [15] and Boero and Lampis [16]).

### 3.4. Clustering of European countries according to mortality levels

An interesting concept for this kind of analysis is the clustering of European countries in K defined groups according to their mortality levels. One of the most popular clustering algorithms is the K

– means algorithm that classifies a set of data points in K predefined and unequally distributed clusters based on some metric, e.g., Euclidean, Minkowski or Chebyshev distance, as it is described in KumarMalhotra et al. [17]. Usually, a preprocessing step of this procedure is the standardization of observations, targeting to minimize effects of heteroscedasticity (Mohamad and Usman [18] and Nasser et al. [19]).

However, in this article we prefer the production of more flexible results that can be used also for measuring the differentiation of mortality rates between countries of the same cluster. As a result, we utilized Fuzzy K – means algorithm to attain the desirable outcome.

### 3.4.1. Fuzzy K – means algorithm

The methodology of this iterative algorithm differs slightly from the methodology of K – means. The element that introduces fuzziness in this algorithm is the parameter $m$ (fuzzifier), according to which the amount of ambiguity included in the clustering results is determined. Each subgroup is characterized based on the elements that will be placed inside it, and by its centroid $c_k$. A weight $w_{ik}$ is assigned to each element of the clustering process. In our case, $i = 1,\dots,31$ represents the country while $k$ represents the cluster. For each $i$ we have

$$\sum_{k=1}^{K} w_{ik} = 1, \quad 0 < w_{ik} < 1. \tag{9}$$

while $c_k$ and $w_{ik}$ are calculated as

$$c_k = \frac{\sum_{i=1}^{N} (w_{ik})^m \; x_i}{\sum_{i=1}^{N} (w_{ik})^m}, \tag{10}$$

$$w_{ik} = \frac{1}{\sum_{l=1}^{K} \left( \frac{||x_i - c_k||^2}{||x_i - c_l||^2} \right)^{\frac{1}{m-1}}}. \tag{11}$$

produced by the minimization of the objective function (KumarMalhotra et al. [17] and Mohamad and Usman [18])

$$J(c,m)=\sum_{i=1}^{N}\sum_{k=1}^{K}(w_{ik})^{m}\left||x_i-c_k|\right|^2. \quad (12)$$

The parameters $m$ and $K$ are defined in advance, while the iterative procedure is being completed when $J(c,m)<\varepsilon$, where $\varepsilon$ is constant. As stated in Nasser et al. [19] when $m\to\infty$, the fuzziness of clustering results increases, while arbitrary definition of $m$ and $K$ can cause untrustworthy results. Hence, in many articles, like in Höppner et al. [20], the authors propose the usage of $m=2$, while in Schwämmle and Jensen [21] the authors propose the estimation of $m$ based on an analytical function that takes into account the number and the dimension of the data points. However, this function gives reliable results only for many data points. Moreover, the selection of $K$ can be determined by observing the graphs displaying the values of various functions in parallel with K. In our analysis, we relied on the diagrams of minimum centroid distance ($V_{MCD}$), partition coefficient ($V_{PC}$) and partition entropy ($V_{PE}$) against K (Schwämmle and Jensen [21]). The selection of K arises from the part of the graph where a plateau region emerges. On the other hand, there are cases where this criterion cannot be implemented due to large amount of noise.

### 3.4.2. Jensen – Shannon divergence

One of the most popular applications of Kullback – Leibler divergence ($D_{KL}$) is the measurement of the differentiation between two probability distributions. In the discrete case, letting $p$ and $q$ describing distributions, the Kullback – Leibler function is stated as

$$D_{KL}(p,q) = \sum_{x\in\mathcal{F}} p(x)\log\left(\frac{p(x)}{q(x)}\right), \tag{13}$$

where $\mathcal{F}$ denotes the σ – algebra containing all possible events of the sample space of the experiment Ω. More specifically, if we use 2 as the basis of the logarithm, Kullback – Leibler divergence measures the average number of extra bits needed to describe – encode samples of distribution *p* using distribution *q*.

In our analysis, we preferred a symmetric distribution measure, leading us to the alternative of Jensen – Shannon divergence that is usually referred as symmetric Kullback – Leibler. As stated (Belov [22], Nguyen and Vreeken [23], Menéndez et al. [24], Levene and Kononovicius [25]) the Jensen – Shannon divergence is defined as

$$D_{JS}(p,q) = \frac{1}{2}D_{KL}(p,\mu) + \frac{1}{2}D_{KL}(q,\mu) \tag{14}$$

where $\mu = \frac{p+q}{2}$. We can claim that the Jensen – Shannon divergence measures the average number of extra bits needed if we try to encode samples of the distributions *p* and *q* through their mean distribution. Endres and Schindelin [26], underline that the Jensen – Shannon function does not satisfy all distance‘s properties, however this is accomplished by taking the squared root of (14).

### 3.5. Correlations in Time Series Dynamics

After the presentation of clustering techniques, we focus on three correlation measures and more specifically on the Spearman coefficient, the distance correlation and the cross – correlation function. We underline that the following analysis concerns only the stationary equivalent of the initial time series, as the existence of the decreasing trend in all 31 time series would lead to extremely high values of the aforementioned correlation indices, producing unreliable results and failing to describe the dynamics of the examined time series.

### 3.5.1. Spearman coefficient

Initially, we chose to present the results of the time series correlation investigation through the Spearman coefficient. We notice that the implementation of the Spearman coefficient concerns observations that are independently distributed (i.i.d.), and not observations that are part of time series where linear or non-linear relationships are likely to occur; however, in many research articles this measure is used with caution in order to investigate time series correlations (Kodera et al. [27]). Another frequently used measure for investigating linear relationships between two random variables is the Pearson coefficient (Liang et al. [28]). Nevertheless, we focused on the option of the Spearman coefficient, firstly because it takes into account monotonous and not only linear relations and secondly it removes the hypothesis of normality that accompanies the calculation of Pearson‘s coefficient.

Spearman’s coefficient is given by equation

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}. \tag{15}$$

where $d_i = r_{x_i} - r_{y_i}$ while $r_{x_i}$, $r_{y_i}$ represent the order of random variables $x_i$ and $y_i$ respectively. Spearman’s coefficient takes values in $[-1, 1]$, where $r_s = 1$ displays perfect positive correlation, $r_s = -1$ perfect negative correlation and $r_s = 0$ linear independence. Coefficient’s values are statistically significant when the condition

$$\left| r_s \sqrt{\frac{n-2}{1-r_s^2}} \right| > t_{n-2;\frac{\alpha}{2}} \tag{16}$$

is satisfied for $n > 10$. Finally, we should be aware that the smaller the value of $n$, the more irresponsible the results become.

### 3.5.2. Distance correlation

In contrast to what we mentioned above about the Pearson and Spearman correlation coefficients, which provide an estimate for the linear correlation of two random variables or two time series, whether positive or negative, the distance correlation measures linear and non-linear correlations between the variables. However, by definition, this indicator cannot determine whether we have a positive or negative dependence.

We consider a pair of random values $(x_k, y_k)$, which are either coordinates of two random variables X and Y, or elements of two time series $\{x_t\}$ and $\{y_t\}$. First, we compute the $n \times n$ tables of distances $(a_{j,k})$ and $(\beta_{j,k})$ where the elements of the distances between observations are defined as

$$\alpha_{j,k} = \left|\left|x_j - x_k\right|\right|, \quad \beta_{j,k} = \left|\left|y_j - y_k\right|\right|, \; j,k = 1, \dots, n, \tag{17}$$

and ||.|| denotes the Euclidean metric. In addition, we calculate the values of

$$A_{j,k} = a_{j,k} - \bar{a}_{j.} - \bar{a}_{.k} + \bar{a}.. \tag{18}$$

and

$$B_{j,k} = \beta_{j,k} - \bar{\beta}_{j.} - \bar{\beta}_{.k} + \bar{\beta}.., \tag{19}$$

where $\bar{a}_{j.}$ stands for the mean value of the j-th column, $\bar{a}_{.k}$ denotes the mean value of the k-th row, while $\bar{a}..$ denotes the overall mean value of matrix $(a_{j,k})$. The interpretation of the quantities in equation (19) is similar to the interpretation of equation (18). As a result, the square of the sample distance covariance is the mean value of the sum of the two elements $A_{j,k}, B_{j,k}$ (Székely et al. [29]) and can be expressed as

$$dCov^2(X,Y) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} A_{j,k}\, B_{j,k}. \tag{20}$$

Furthermore, after displaying the sample covariance, we present the sample distance variance described by equation (21),

$$dVar^2(X) = dCov^2(X,X) = \frac{1}{n^2}\sum_{j,k} A_{j,k}^2. \tag{21}$$

Finally, according to equations (17) - (21) the sample distance correlation is defined as

$$R_n = dCor(X,Y) = \frac{dCov(X,Y)}{\sqrt{dVar(X)dVar(Y)}}. \tag{22}$$

taking values in the interval $[0,1]$. Greater values of $R_n$ are witnessing stronger dependence between the examined elements.

The statistical significance of the generated value of the distance correlation $R_n$ can be examined by conducting a hypothesis test, where the null hypothesis stating that the examined random variables or time series are independent is rejected when

$$\left|\sqrt{\nu-1}\frac{R_n}{\sqrt{1-R_n^2}}\right| > t_{\nu-1;\frac{\alpha}{2}} \tag{23}$$

as the quantity to the left of the inequality follows asymptotically the Student distribution with $\nu-1$ degrees of freedom where $\nu = \frac{n(n-3)}{2}$ (Székely and Rizzo [30]).

### 3.5.3. Cross – correlation of time series

The cross – correlation function is often characterized as a sliding inner product, while its calculation involves the process of continuous or discrete convolution depending on the problem to which it applies. The most common fields of application of this mathematical tool are pattern recognition, cryptanalysis, signal processing, neurophysiology and the investigation of correlations between stochastic processes.

The cross – correlation of two stochastic processes is the correlation between the values of the processes, while it is a function of two time points $t_1$ and $t_2$. This function is a generalization of the autocorrelation function which refers to the dependency relationships that are created within a time series for various lags $\tau$. Thus, a correlation function can also be expressed in terms of the lag $\tau = t_1 - t_2$. The sample correlation function of two time series $\{x_t\}$ and $\{y_t\}$ can be written (Sadiku et al. [31], Robbins and Fisher [32]) as

$$r_{XY}(\tau) = \frac{K_{XY}(\tau)}{s_X s_Y} = \frac{\frac{1}{n-\tau}\sum_{t=\tau+1}^{n}(x_t y_{t-\tau} - \bar{x}\bar{y})}{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(x_t^2 - \bar{x}^2)\frac{1}{n}\sum_{t=1}^{n}(y_t^2 - \bar{y}^2)}}, \tag{24}$$

The correlation function takes values in the interval $[-1, 1]$, where $r_{XY}(\tau) > 0$ denotes positive linear correlation between the examined time series, while $r_{XY}(\tau) < 0$ denotes negative linear correlation. Also, the symmetric property holds for this function, as $r_{XY}(\tau) = r_{XY}(-\tau)$. The statistical significance of this index, is examined by checking whether

$$r_{XY}(\tau) \notin \left[-\frac{Z_{\frac{a}{2}}}{\sqrt{n}}, \frac{Z_{\frac{a}{2}}}{\sqrt{n}}\right], \tag{25}$$

depending on the significance level $\alpha$, but also on the length of the time series. Because the most common choice is $a = 0.05$ the above control is simplified, considering statistically significant

values of $r_{XY}(\tau)$ that are outside the interval $[-\frac{2}{\sqrt{n}}, \frac{2}{\sqrt{n}}]$. As in the previous measures, the calculation of the cross – correlation presupposes the stationarity of the time series (Damos [33]). There are many articles using this measure to highlight the causal relationships of different stochastic processes due to the intuitive interpretation provided by this indicator and the ease of its application, although this approach seems to be insufficient.

### 3.6. Simple and Instantaneous Granger Causality

The goal in this section is to reveal causal relations between time series that can be proven as beneficial during the forecasting process of these series. Let $\{x_t\}$ and $\{y_t\}$ be two stochastic processes and let us investigate the cause – effect relationship between these processes. If such a relation is proven, where process $\{x_t\}$ functions as the cause for $\{y_t\}$ (effect), then the information provided by the past values of $\{x_t\}$ will be decisive in the predictive accuracy that is provided using only the past values of process $\{y_t\}$.

Before we move on to the rigorous mathematical representation of causality according to Granger, it should be noted that in many research papers exploring cause – effect relationships between stochastic processes, correlation, partial correlation or the mutual information function have been used as tools of investigating causality. Despite that these approaches are not incorrect, they fail to manage the concept of causality effectively, thus often yielding misleading results. The main difference between these techniques and Granger causality is that these measures can only give some indication of the correlation – dependence between processes and not whether the information provided by a stochastic process can contribute beneficially to the predictability of another. Once again, we base our analysis on stationary time series absolved from deterministic or stochastic trends (Papaioannou et al. [34]).

#### 3.6.1. Definition

**S**uppose that $I_t$ symbolizes the total information available until time point t. Hence, this set contains all the information provided by the time series $\{x_t\}$ and $\{y_t\}$. Let $\bar{x}_t$ =

$\{x_t, x_{t-1}, \dots, x_1\}$ be the set containing the current and past values of time series $\{x_t\}$ while $\bar{y}_t$ denotes the corresponding set for $\{y_t\}$. Finally, let $\sigma^2(.)$ representing the variance of the prediction errors. Then the following applies to the time series $\{x_t\}$ and $\{y_t\}$:

1. (Granger causality). The time series $\{x_t\}$ can be deemed as Granger causal to $\{y_t\}$ if and only if

$$\sigma^2(y_{t+1} \mid I_t) < \sigma^2(y_{t+1} \mid I_t - \bar{x}_t). \tag{26}$$

meaning that the future values of $\{y_t\}$ can be predicted more accurately (less forecasting error variance) when we take into account information provided by the current and past values of $\{x_t\}$ (Koutlis [35]),

2. (Instantaneous Granger causality). The process $\{x_t\}$ can be deemed as an instantaneous Granger causal to $\{y_t\}$ if and only if

$$\sigma^2(y_{t+1} \mid \{I_t, x_{t+1}\}) < \sigma^2(y_{t+1} \mid I_t), \tag{27}$$

considering that the future values of $\{y_t\}$ can be predicted more accurately by taking into account the information provided by $x_{t+1}$, the current and the past values of $\{x_t\}$ (Gianetto [36]),

3. (Feedback). There is a feedback between time series $\{x_t\}$ and $\{y_t\}$, when $\{x_t\}$ is Granger causal to $\{y_t\}$ and vice versa.

The existence of feedback makes sense to be considered only in the case of Granger causality, because the property of instantaneous Granger causality is symmetric.

### 3.6.2. Statistical test for Granger causality

The existence of Granger causality from $\{x_t\}$ to $\{y_t\}$, is examined by checking whether the

regression of past values of $\{x_t\}$ and $\{y_t\}$, to the present value $y_t$, produces less residual variance compared to the variance produced using only the past values of $\{y_t\}$. According to least squares method, the following equation combines the information provided by the past values of $\{x_t\}$ and $\{y_t\}$ for the prediction of $y_t$

$$y_t = a_0 + \sum_{k=1}^{k_1} a_{11}^k y_{t-k} + \sum_{k=k_0}^{k_2} a_{12}^k x_{t-k} + v_{1,t}, \tag{28}$$

where $k_0 = 1$. After the fitting of model (28), we utilize the least squares method once more, aiming to fit a classic AR model to the time series $\{y_t\}$, of the form

$$y_t = a_0 + \sum_{k=1}^{k_1} a_{11}^k y_{t-k} + v_{2,t}. \tag{29}$$

The model exhibited by equation (28) is called unrestricted, while model (29) is called restricted. The value of Granger causality is calculated as

$$GCI_{X\to Y} = ln\frac{Var(\hat{v}_{2,t})}{Var(\hat{v}_{1,t})}. \tag{30}$$

If $GCI_{X\to Y} > 0$ then we can claim that the time series $\{x_t\}$ is a Granger cause for $\{y_t\}$. During the calculation of $GCI_{X\to Y}$ between different time series, negative values may also occur. In this case, the past values of the time series $\{x_t\}$ do not show causality relations with the time series $\{y_t\}$ and act negatively in predicting future values.

To check the statistical significance of the Granger causality from $\{x_t\}$ to $\{y_t\}$, the use of the Snedecor – Fisher (F – test) hypothesis test is employed, where the null hypothesis states that $a_{12}^k = 0$ for all $k = 1, \dots, p$, which is equivalent to the fact that there is no causal relation from time series $\{x_t\}$ to $\{y_t\}$. The utilized F statistic, follows F distribution with $p$ and $n - 3p$

degrees of freedom

$$F = \frac{\frac{(SSE^R - SSE^U)}{p}}{\frac{SSE^U}{n-3p}} \sim F_{p,n-3p} \tag{31}$$

where $SSE^R$, $SSE^U$ denote the sum of the squared errors of the restricted and unrestricted model, $p$ is the order of the model of the relation (29), $2p$ is the order of the model of equation (28) (Kugiumtzis [9]).

The existence of a Granger causality from $\{y_t\}$ to $\{x_t\}$ can be examined in a similar way. The existence or absence of causality from $\{x_t\}$ to $\{y_t\}$ in no way implies the existence or absence of causality from $\{y_t\}$ to $\{x_t\}$. The instantaneous Granger causality can be controlled similarly, but considering in the model of equation (28) that $k_0 = 0$. In this way, the null hypothesis of the F – test is converted to $a_{12}^k = 0$ for all $k = 0, \dots, p$.

## 4. RESULTS

### 4.1. Trend Estimation

Before proceeding to the results of linear time series analysis, we dedicated the first part of this section to the trend estimation of the 31 time series. Observing Eurostat's dataset of premature mortality rates, all 31 time series were characterized by intense linear descending trend. Robertson and Ecob [37], provide a methodology on estimating the trend in time series of mortality rates that are not necessarily associated with chronic diseases, in the UK. Each of the non-stationary time series can be represented by equation $y_t = at + b + \varepsilon_t$, where $\varepsilon_t$ is a white noise following $N(0, \sigma_\varepsilon)$. The estimated slope and constant for each country, is presented in Table 1.

**Table 1**. Estimation of trend and constant coefficient for time series of mortality rates

| **Country** | $\hat{a}$ | $\hat{b}$ |
|---|---|---|
| Belgium | -3.1172 | 171.7156 |
| Bulgaria | -4.0330 | 279.9960 |
| Czech Republic | -5.8971 | 268.5678 |
| Denmark | -4.4761 | 205.4622 |
| Germany | -3.2601 | 180.9308 |
| Estonia | -9.4230 | 355.5714 |
| Ireland | -4.4379 | 192.3025 |
| Greece | -0.9641 | 141.4141 |
| Spain | -2.6096 | 152.8529 |
| France | -2.5519 | 141.0100 |
| Croatia | -3.4175 | 225.0550 |
| Italy | -3.0929 | 152.8575 |
| Cyprus | -1.3007 | 101.2978 |
| Latvia | -5.8450 | 345.7130 |
| Lithuania | -3.6026 | 315.5243 |
| Luxembourg | -3.8432 | 175.0819 |
| Hungary | -7.6580 | 421.8214 |
| Malta | -2.5630 | 156.3500 |
| Holland | -3.0266 | 163.6946 |
| Austria | -3.6019 | 184.8275 |
| Poland | -4.5247 | 261.4071 |
| Portugal | -2.7315 | 170.3482 |
| Romania | -5.2284 | 320.3263 |
| Slovenia | -4.8163 | 230.6246 |
| Slovakia | -4.2959 | 276.2481 |
| Finland | -2.8436 | 167.0960 |
| Sweden | -2.7256 | 139.1743 |
| UK | -1.3007 | 101.2978 |
| Iceland | -3.6751 | 154.5641 |
| Norway | -3.5577 | 158.2551 |
| Switzerland | -2.5487 | 131.0301 |

Performing hypothesis tests for the produced values of $\hat{a}, \hat{b}$, the presented estimated coefficients were deemed as statistically significant in all 31 cases of time series, at the level of significance $\alpha = 0.05$.

Looking at the results presented in Table 1, we observe that the most intense descending trend corresponds to the mortality rates of the inhabitants of Estonia with $\hat{\alpha} = -9.641$, while the

mortality rates of Hungary show the second most intense decreasing trend with $\hat{\alpha} = -7.658$. On the other hand, the smoothest descending tendency occurs in Greece 's occasion with $\hat{\alpha} = -0.9641$, followed by UK and Cyprus with $\hat{\alpha} = -1.3007$. The high value of $\hat{b}$ coefficient in the case of Hungary classifies it among the countries with the highest mortality rates in Europe despite its prevalent downward trend. At the same time, compared to the values of the estimated $\hat{b}$ coefficients of the 31 time series, value $\hat{b} = 141.414$ corresponding to Greece, is relatively small which states that Greece has relatively low mortality rates in Europe. However, the weak declining tendency shown by the 24 observations of the time series (1994 – 2017), leads to relatively high mortality rates in recent years compared to the rest of the European Countries.

### 4.2. Fitting ARIMA models on premature mortality rates

In the following analysis we utilized the aforementioned methods and models, keeping in mind the short length of time series that force us to choose models with limited number of parameters, while in some cases where the choice of unique and optimal model was not clear, priority was given to AR models as they offer a better understanding of system dynamics compared to MA models. We consider that the maximum number of estimated parameters for the proposed linear models should be 5, based on the length of 24 observations that characterizes the majority of the examined time series, avoiding precarious results. Furthermore, the most fundamental criterion in the selection of an ARIMA model was AIC, although in cases where AIC values were close for different combinations of $p$ and $q$, we compared the forecasting efficiency of these models by comparing their corresponding NRMSE-values. The prevalent descending trend of the time series was removed, before the application of the mentioned criteria, using 1st and 2nd order differences.

We next provide the methodology we followed for the selection of the two basic linear models, i.e. ARIMA and random walk models, for the 32 time series exhibiting the mortality rates. By eliminating the trend from the time series of premature mean mortality rates in Europe, we could study the dynamics of the stationary time series through the statistical tests and diagrams given below. Looking at the AIC graph, the smallest value is displayed for $p$ = 0 and $q$ = 2, with $AIC_{MA(2)} = 0.6721$. However, we choose model AR(2), with $AIC_{AR(2)} = 0.8375$. This choice is based on the difference in the forecasting capacity of the two models, as for 80% training set we

received $NRMSE_{AR(2)} = 0.799$, while $NRMSE_{MA(2)} = 0.9027$. Simultaneously, the difference between the AICs of these two models is quite small, allowing us to choose the AR(2), which is also supported by the partial autocorrelation diagram, according to which the autocorrelations after the 2nd lag become statistically insignificant. The equation describing the AR(2) model of mean mortality rates in Europe, is given as

$$X_t = \nabla Y_t = -5.1949 - 0.0816X_{t-1} - 0.4367X_{t-2} + \varepsilon_t, \quad (32)$$

with $\sigma_\varepsilon = 1.405$.

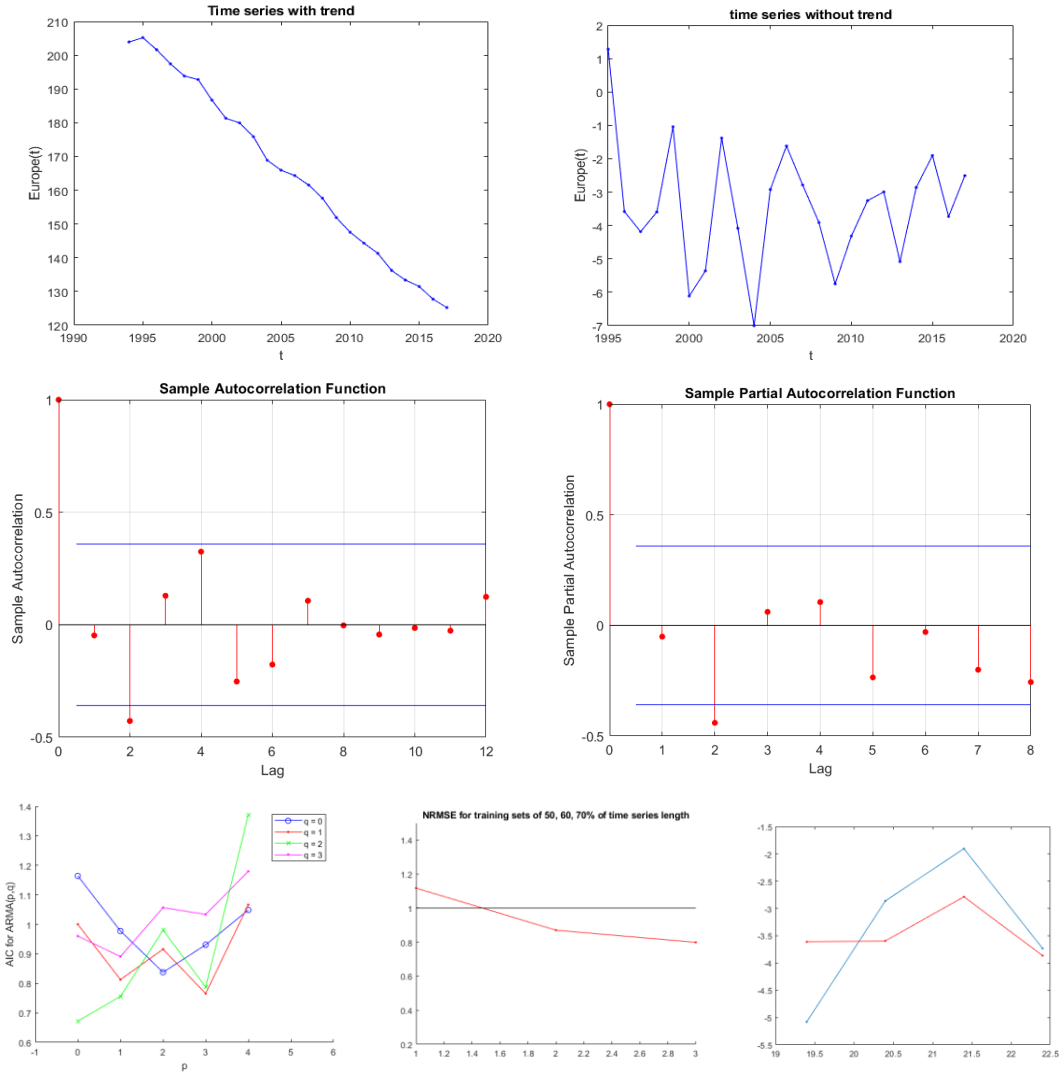


**Figure 1**. Europe's non-stationary and stationary time series of mean mortality rates and the corresponding autocorrelation diagrams, AIC and NRMSE.

Next, we present a case where the fitting of any ARIMA model seems to be unsatisfactory according to the statistical criteria presented above; in this case the selection of random walk model is required. Thus, considering the stationary time series of Bulgaria's mortality rates, we observe insignificant autocorrelations and partial autocorrelations just from the first lag. This phenomenon, in parallel with the p-values of the Portmanteau test underline the inappropriate fitting of a linear time series model in Bulgaria's premature mortality rates, leading us to deem the stationary time series as white noise and the nonstationary equivalent as random walk.

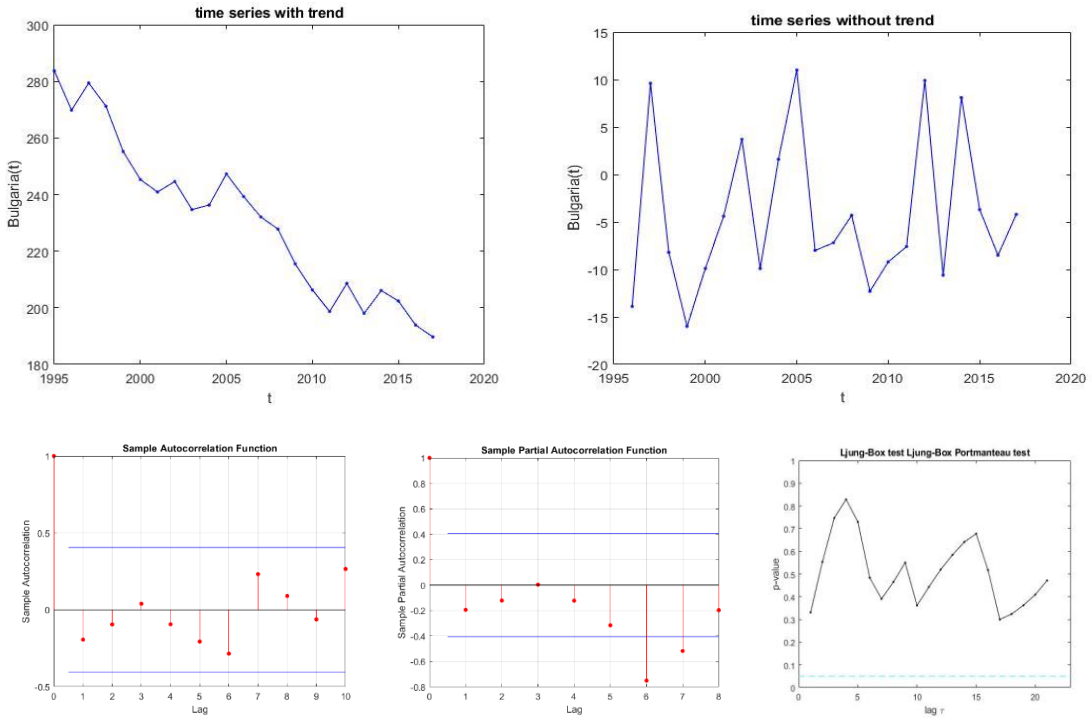


**Figure 2**. Bulgaria's non-stationary and stationary time series of mortality rates and the corresponding autocorrelation and Portmanteau diagrams.

The results of the ARIMA and random walk models to describe the mortality rates of the European countries of our analysis, are presented in Table 2, where the selection of the suitable models was made as presented in the aforementioned two examples.

**Table 2**. Best ARIMA models with estimated coefficients and values of selection criteria for each of the 31 countries plus the model of the mean mortality rates of Europe.

| Countries | ARIMA(p,d,q) | AIC | $\sigma_\varepsilon$ | $NRMSE$ | C | $\varphi_1$ | $\varphi_2$ | $\varphi_3$ | $\theta_1$ | $\theta_2$ | $\theta_3$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Belgium | ARIMA(0,1,2) | 1.476 | 1.930 | 0.676 | | - | - | - | 0.978 | 0.053 | - |
| Bulgaria | Random Walk | - | 8.069 | 1 | - | - | - | - | - | - | - |
| Czech Republic | ARIMA(1,1,2) | 1.830 | 2.271 | 0.826 | -3.895 | 0.371 | - | - | 1.338 | -0.874 | - |
| Denmark | ARIMA(1,1,0) | 2.338 | 3.151 | 0.947 | -7.489 | -0.631 | - | - | - | - | - |
| Germany | ARIMA(0,2,1) | 1.289 | 1.863 | 0.653 | - | - | - | - | 0.847 | - | - |
| Estonia | Random Walk | - | 11.240 | 1 | - | - | - | - | - | - | - |
| Ireland | Random Walk | - | 4.186 | 1 | - | - | - | - | - | - | - |
| Greece | ARIMA(1,1,2) | 1.296 | 1.739 | 0.679 | -1.646 | -0.721 | - | - | -0.01 | 0.991 | - |
| Spain | ARIMA(0,2,2) | 0.181 | 1.010 | 0.666 | - | - | - | - | 1.656 | -0.912 | - |
| France | Random Walk | - | 0.907 | 1 | - | - | - | - | - | - | - |
| Croatia | ARIMA(2,1,0) | 2.922 | 3.780 | 0.872 | -3.559 | -0.744 | -0.601 | - | - | - | - |
| Italy | ARIMA(0,2,1) | 1.137 | 1.691 | 0.698 | - | - | - | - | 0.717 | - | - |
| Cyprus | Random Walk | - | 5.695 | 1 | - | - | - | - | - | - | - |
| Latvia | Random Walk | - | 15.240 | 1 | - | - | - | - | - | - | - |
| Lithuania | ARIMA(0,1,1) | 4.852 | 11.077 | 0.960 | - | - | - | - | -0.753 | - | - |
| Luxembourg | ARIMA(1,1,3) | 4.461 | 8.227 | 0.790 | -4.832 | -0.318 | - | - | 0.471 | 0.172 | 0.357 |
| Hungary | ARIMA(2,1,0) | 3.267 | 4.736 | 0.809 | -11.245 | 0.030 | -0.453 | - | - | - | - |
| Malta | ARIMA(3,1,0) | 4.208 | 7.347 | 0.458 | -7.33 | -0.675 | -0.652 | -0.361 | - | - | - |
| Holland | ARIMA(2,1,0) | 1.013 | 1.534 | 0.817 | -4.808 | -0.116 | -0.471 | - | - | - | - |
| Austria | Random Walk | - | 2.200 | 1 | - | - | - | - | - | - | - |
| Poland | Random Walk | - | 3.630 | 1 | - | - | - | - | - | - | - |
| Portugal | Random Walk | - | 3.540 | 1 | - | - | - | - | - | - | - |
| Romania | Random Walk | - | 5.805 | 1 | - | - | - | - | - | - | - |
| Slovenia | ARIMA(1,1,0) | 3.511 | 5.666 | 0.704 | -7.99 | -0.632 | - | - | - | - | - |
| Slovakia | Random Walk | - | 6.209 | 1 | - | - | - | - | - | - | - |
| Finland | Random Walk | - | 2.593 | 1 | -3.265 | -0.113 | - | - | 0.109 | - | - |
| Sweden | Random Walk | - | 1.586 | 1 | - | - | - | - | -0.330 | 0.878 | - |
| UK | ARIMA(2,2,0) | 1.511 | 2.049 | 0.829 | 0.2425 | -0.635 | -0.499 | - | - | - | - |
| Iceland | ARIMA(1,1,0) | 5.091 | 12.233 | 0.799 | -4.858 | -0.451 | - | - | - | - | - |
| Norway | ARIMA(0,2,1) | 1.415 | 1.985 | 0.732 | - | - | - | - | 0.956 | - | - |
| Switzerland | ARIMA(2,1,1) | 1.147 | 1.614 | 0.337 | -7.499 | -1.277 | -0.686 | - | -0.964 | - | - |
| Europe | ARIMA(2,1,0) | 0.837 | 1.405 | 0.799 | -5.195 | -0.082 | -0.437 | - | - | - | - |

### 4.3. Fitting SETAR models to Mortality Rates

After the presentation of the results of linear analysis, we propose the fitting of SETAR models, examining the 13 stationary time series of countries that during the linear analysis were considered white noise. This part of the analysis was implemented because the standard ARIMA models do not have the ability to manage systems governed by nonlinear dependency relationships.

However, applying the criteria of non-linear analysis and fitting SETAR models as described by (8), for 5 out of 13 time series, we obtained notable results through the application of local linear models. The selection of the SETAR models of the following analysis is mainly based on the AIC values and the NRMSE values, as the (partial) autocorrelation function and the Portmanteau test can no longer provide helpful hints, because they are useful in highlighting the existence of linear correlations. The selection of only 2 linear segments for time series modeling, aims to limit the parameters that have to be estimated, due to the short length of the time series.

Often in non-linear time series analysis, the appropriate value for lag $d$ is selected by observing the mutual information diagram, where the lag that correspond to the first minimum is chosen as $d$. Since in our analysis the sampling rate is low, $d = 1$ could be considered as the ideal option. The appropriate value for $m$ in each case was determined after experimentation.

**Table 3**. Best SETAR models with estimated coefficients and values of selection criteria.

| Countries | AIC | $\sigma_\varepsilon$ | NRMSE | Thres | $\varphi_{1,0}$ | $\varphi_{2,0}$ | $\varphi_{1,1}$ | $\varphi_{2,1}$ | $\varphi_{1,2}$ | $\varphi_{2,2}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Austria | 1.43 | 1.79 | 0.551 | -4.1 | -0.379 | -2.425 | - | 0.573 | 0.385 | - |
| Bulgaria | 3.92 | 5.66 | 0.572 | -9.2 | 2.394 | -7.826 | 0.164 | -0.175 | - | 0.204 |
| Ireland | 2.80 | 3.54 | 0.800 | -5.6 | -10.980 | -4.035 | -0.738 | - | - | - |
| Romania | 2.41 | 2.69 | 0.747 | -3.6 | -3.067 | -4.719 | 0.106 | 0.903 | - | -0.285 |
| Slovakia | 3.78 | 5.74 | 0.858 | -1.9 | -4.11 | -13.1 | - | 0.774 | - | - |

Regarding the other 8 time series of premature mortality levels, SETAR models do not seem to give an adequate fit for the various values of $d$, $L$ and $H$. Thus, we suggest (maintain) the model of the random walk for them, taking as a means of forecasting for the next time steps the average value of the time series. However, in some cases such as the time series of France, Poland,

Portugal and Sweden where the standard deviation of white noise is quite small, this forecasting method is quite sufficient.

### 4.4. Clustering of European countries and measuring distances between participation distributions

In this paragraph we will present the experimental results of the application of Fuzzy K – means algorithm on the mortality rates of the 31 examined countries. For this purpose, we used the observations of non – stationary time series from 2003 to 2017, aiming to avoid the majority of missing values that concern the time interval between 1994 and 2002. Hence, we will cluster 31 data points denoted as $\boldsymbol{x}_i \in \mathbb{R}^{15}$, using $m = 2$ according to the suggestions of related work. In addition, we select to group $\boldsymbol{x}_i$ into 6 clusters, a choice that combines the results presented in Figure 3.

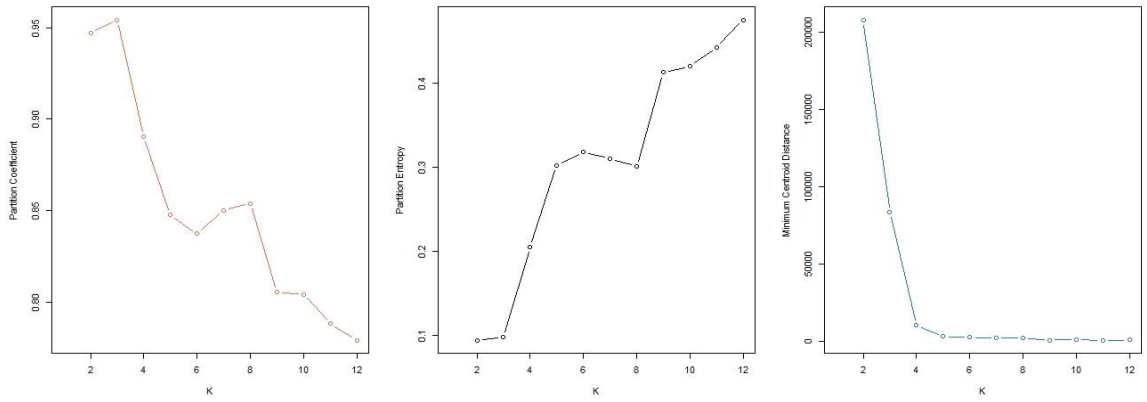


**Figure 3**. Diagrams of $\mathrm{V_{MCD}} - \mathrm{K}, \mathrm{V_{PC}} - \mathrm{K}$ and $\mathrm{V_{PE}} - \mathrm{K}$ for K = 2, …, 12.

After the implementation of Fuzzy K – means to the 31 data points, we culminate in the construction of matrix $\boldsymbol{W}$ that contains the estimated $w_{ik}$ for $i = 1, \dots, 31$ and $K = 1, \dots, 6$, that could be deemed as the participation probabilities of each time series into the 6 clusters. In Figure 4, we observe the estimated values of matrix $\boldsymbol{W}$**.** Based on the estimated $w_{ik}$ for each country, we have the ability to put the produced results on Europe's map, considering that $\max_{k=1,\dots,6} w_{ik}$ is the criterion used to get a unique classification of the countries into the 6 clusters.

The number of the cluster in which each of the 31 countries belongs, acts as an indicator of how low or high is the premature mortality from chronic diseases in this country. We observe that some countries, e.g., Iceland, Norway, Switzerland, et al, are characterized by the lowest mortality rates compared to the rest of European countries, while Latvia, Lithuania, Romania and Hungary seem to be the most impaired instances. The convergence of the algorithm was accomplished after the 40th iteration.

After the completion of the clustering process, we took advantage of the distributions of participation into clusters provided by the estimated $w_{ik}$, by measuring their divergence utilizing the Jensen – Shannon function. The presented results concern only pairs of countries belonging into the same cluster, as our goal is the examination of homogeneity within the determined groups. Deploying the symmetry of Jensen – Shannon divergence, we constructed matrix

$$\boldsymbol{JS} = \begin{pmatrix} \boldsymbol{C_1} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{C_6} \end{pmatrix}. \tag{33}$$

where $\boldsymbol{C}_i$ are upper triangular matrices, consisted of zeros in the main diagonal as there is no divergence in pairs consisting of the same country.

In the 1st cluster, the greatest homogeneity is presented by the pair Cyprus – Norway with $D_{JS} = 0.0004$, while extremely close to them is Switzerland where the $D_{JS}$ of Cyprus – Switzerland is 0.0014. In the 2nd subgroup, the highest homogeneity is given by the pair France – Netherlands with $D_{JS} = 0.0031$, while Finland shows the greatest inhomogeneity, as the minimum $D_{JS}$ we encounter is of the order of 0.0102. The most homogeneous pair of the 3rd subgroup, seems to be the pair of Germany – UK, having the lowest value we have encountered in all 6 subgroups with $D_{JS} = 0.0001$. On the other hand, the distribution of Slovenia seems to deviate considerably from the distribution of any other of the 6 countries in the 3rd subgroup, as all values of the Jensen - Shannon deviation are greater than $0.1$. Furthermore, the Croatia - Estonia pair of the 4th subgroup is presented as quite homogeneous with $D_{JS} = 0.0084$, while

the distribution of Romania in the 6th subgroup seems to differ slightly compared to that of Lithuania ($D_{JS} = 0.0023$), and that of Hungary ($D_{JS} = 0.0093$).

Finally, in this paragraph investigation is carried out to find out the country whose mortality rates indicate the greatest homogeneity with the mean mortality rates of Europe. For this approach, we took advantage of two statistical criteria that are the value of RMSE and the combination of Fuzzy K – means and Jensen – Shannon divergence. Calculating the RMSE between Europe's mortality rates and the other 31 countries, the minimum value was given by Slovenia ($RMSE = 4.7686$) which was approximately 4 times lower than the next lowest value given by Czech Republic. Simultaneously, the minimum $D_{JS} = 0.00011$ was given by Slovenia, although this time the next lowest value was 360 times greater, corresponding again to the Czech Republic.

The aforementioned criteria advocate that Slovenia shows the greatest homogeneity with Europe's mean mortality rates, making Slovenia the best indicator of the evolution of the average behavior of the examined phenomenon in Europe. The Slovenia's selection is maybe not surprising, because Slovenia is classified in the intermediate clusters, being in some sense a middle point between the technologically and medically evolved countries of northern and western Europe in contrast with other EU members.

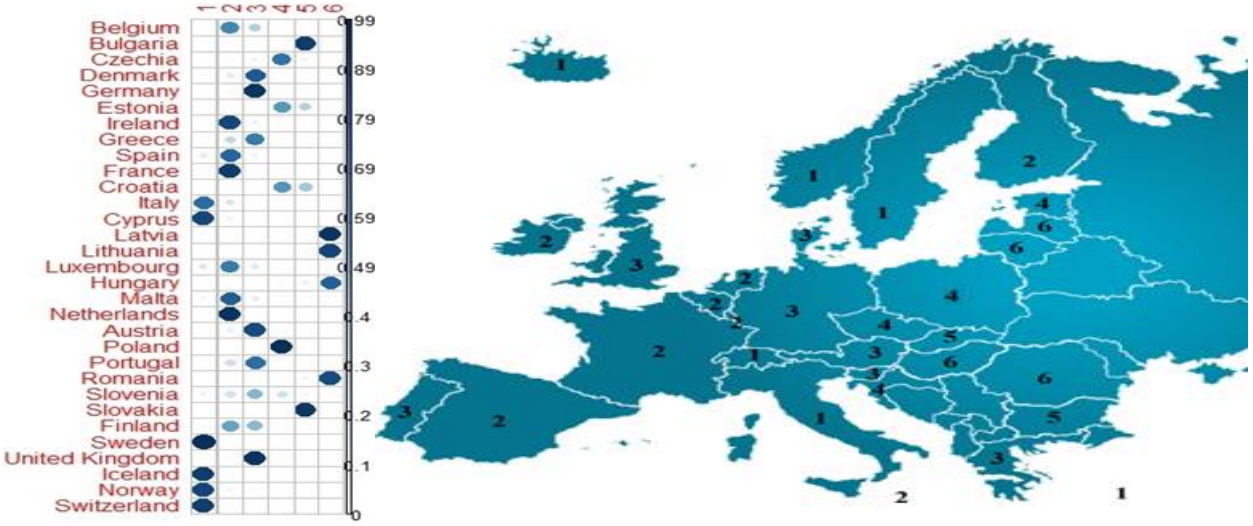


**Figure 4**. Representation of $w_{ik}$ for each country corresponding to the participation probability in each cluster and the distribution of clusters on Europe's map

### 4.5. Application of correlation indices on mortality rates

We begin with the representation of estimated Spearman indices that display the monotonous correlation between the 31 examined time series. This presentation is depicted in Figure 5, where all $\binom{31}{2}$ indices rounded in the first decimal place are given.

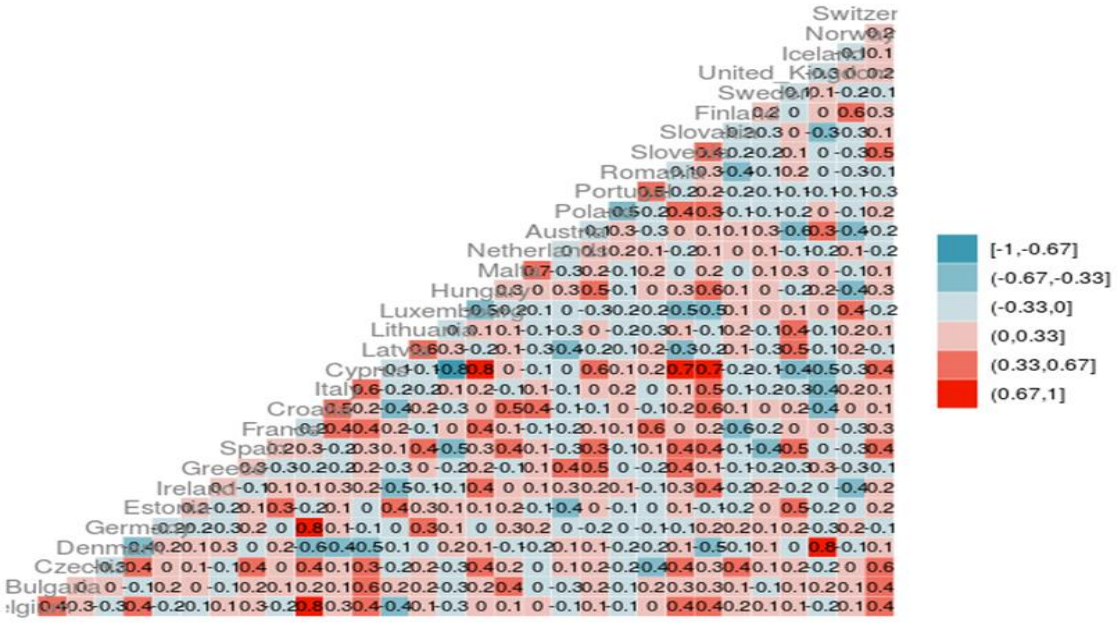


**Figure 5**. Triangle matrix displaying all 465 Spearman correlations

According to the data presented in Figure 5, the strongest monotonous positive correlation corresponds to three pairs of countries with a value of 0.77. The time series pairs that show this high level of correlation are Denmark – Iceland, Germany – Croatia and Cyprus – Hungary, while the pair Belgium – Croatia also shows a high positive linear correlation with $r_s$ = 0.758. At the same time, a strong positive correlation accompanies the pairs of the time series Cyprus – Slovenia and Cyprus – Slovakia with values $r_s$ = 0.69 and $r_s = 0.68$. On the other hand, the strongest monotonous negative correlation appears in the case of Cyprus and Luxembourg with $r_s = -0.75$. The negative correlation indicates the opposite course in the dynamics of the 2 time series.

At this point, we present the most important dependence relationships that occurred between the 31 stationary time series, analyzed through the distance correlation coefficient. Table 5 of the

Appendix shows the time series pairs with a statistically significant distance correlation of $R_n \geq 0.55$. Many of the above time series pairs also appear to have strong monotonous correlations as shown in Figure 5, which is rather expected considering that the distance correlation takes into account linear and non-linear relationships. Quite interesting are the cases of Italy – Hungary and Italy – Slovakia with distance correlations of 0.73, while these pairs do not reveal a significant Spearman correlation.

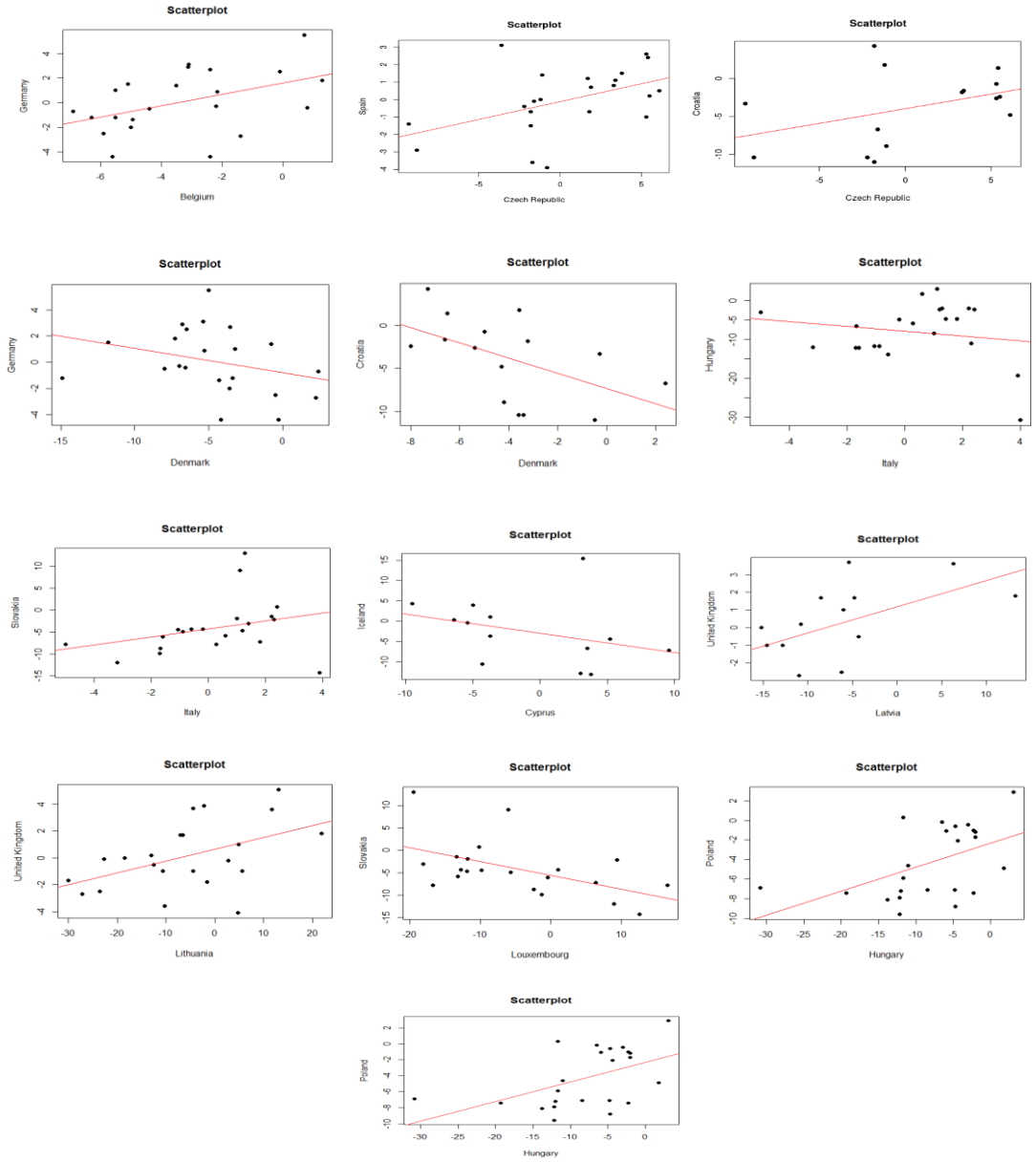


**Figure 6**. Scatterplots of cases with statistically significant distance correlation and low Spearman coefficient

Based on the above 13 diagrams, the examined time series pairs are not necessarily characterized by linear relations, as the regression line of the diagrams does not seem to fit satisfactory the data in most cases. So, we outline the importance of this measure as it reveals in 13 cases, relationships between time series which could not be revealed through the Pearson or Spearman coefficients.

Finally, we will present all the important correlations between the 31 stationary time series concerning chronic mortality rates in Europe. In most cases, values outside the interval $[-0.45, 0.45]$ are considered statistically significant, according to (25). Table 6 of the Appendix includes only the cases where $r_{XY} \notin [-0.5, 0.5]$ and for lags up to $\tau = \pm 4$, aiming to give an adequate intuitive interpretation of the produced results. A few $r_{XY}$ values within the interval $[-0.5, 0.5]$ are displayed exceptionally in time series pairs that had already resulted in some stronger correlation in another lag.

Strong positive correlations appear in the dynamics of the time series Belgium – Croatia, Denmark – Iceland, Germany – Croatia and Greece – Finland with values of 0.748, 0.771, 0.734 and 0.734 respectively. Quite interesting is the fact that these values correspond to a lag of $\tau = 0$, which means that the associated time series pairs evolve in parallel and simultaneously. Still, strong positive correlations are found in the pairs Estonia – Spain and Greece – Hungary with $r_{XY} = 0.644$ for lag $\tau = -1$. The interpretation of these 2 cases is that for every $t$, this year's mortality rates in Estonia and Greece are moving parallel to last year's rates of mortality in Spain and Hungary respectively. Finally, for each year, $t$, the values of the Croatian mortality rates move in parallel with the respective rates of Malta for three years ahead ($r_{XY}$ = 0.672). Recall that the reported correlation values examine the existence of linear relationships between time series.

Regarding negative correlations, strong cross – correlations were revealed by the time series pairs of Estonia – Malta, Italy – Norway, Cyprus – Luxembourg and UK – Switzerland, with correlations $-0.62$, $-0.594$, $-0.697$ and $-0.697$ for lags $\tau = -3, -3, 0$ and 2 respectively. Negative correlation indicates the opposite evolution in time series dynamics, while the lag element is interpreted similarly as in the instances of positive correlations.

### 4.6. Causality investigation between European countries' premature mortality

The selection of $k_1 = k_2$ is often suggested for model (28) as Kirchgässner and Wolters [38] state. In this case, we consider $2k = k_1 + k_2$. At the same time, the results of the aforementioned test are sensitive to the number of past values of the time series $\{x_t\}$ used in the model, since the more observations we include in the regression, the smaller the variance of the residuals. However, high order models can lead to overfitting, resulting in additional parameters describing the noise rather than the system dynamics. For this reason, the use of information criteria (Kirchgässner and Wolters [38]) is proposed, such as the Akaike (AIC), the Schwarz Bayesian Information Criterion (BIC), or the FPE that are defined as

$$BIC(k) = \ln\left(\sigma_{v_1}^2\right) + \frac{\ln(n)(1+2k)}{n}, \tag{34}$$

$$FPE(k) = \sigma_{v_1}^2 \frac{T + (1+2k)}{T - 1 - 2k}. \tag{35}$$

Aiming to obtain the best possible balance between the reduction of residuals variance and the avoidance of cases of overfitting, we selected the value $k$ that minimizes the above criteria. In fact, these criteria can be used in parallel to select the order of the model (28), as in some cases the 3 criteria do not propose the usage of the same optimal order. For example, the Schwarz criterion gives more conservative model choices than the one proposed through the AIC criterion. In cases where the proposed order of AIC and BIC criteria do not agree, the results of the F – test was examined for both different propositions for the class k.

The model's fitting is considered acceptable when the resulting residuals follow normal distribution and are uncorrelated to each other in terms of the various lags. We tested these two conditions, using the non – parametric Kolmogorov – Smirnov normality test, while to evaluate the correlation of residuals, we relied on the use of the Portmanteau test, and the Lagrange test of independence. We note that in many cases, the hypothesis of residuals' independence is tested through the Durbin – Watson test. However, we chose the Portmanteau and Lagrange tests mainly

because they have the advantage of examining the correlation of residuals for multiple lags simultaneously, in contrast to the Durbin – Watson test that must be performed for each lag separately (Kirchgässner and Wolters [38]). Barnett and Seth [39] present the use of a set of functions for calculating various measures related to Granger causality for Matlab software. With regard to instantaneous causality, this kind of causality may occur either because the sampling period is longer than the causal effect period, or because the causally active variables of the system remain hidden during the calculations (Hyvärinen et al. [40]). The phenomenon of long sampling periods is often found in time series that resulted from annual recordings.

The networks shown in what follows display the existence of Granger and instantaneous Granger causality relationships between the 31 time series of mortality rates. In the Appendix we can see the table that contains all the Granger causality relations, the p – values of the corresponding F – tests as well as the order $k$ of the fitted models.

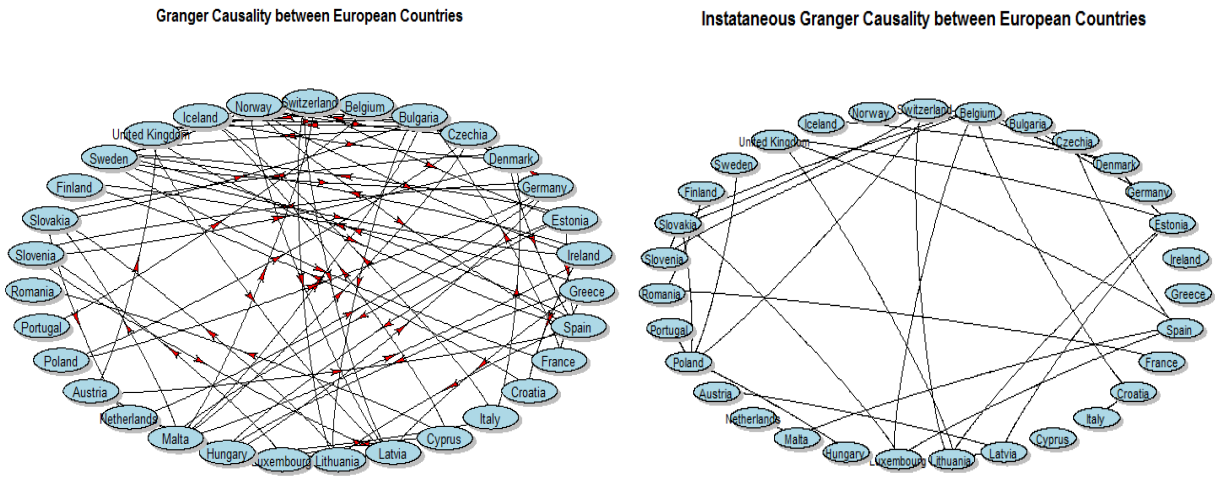


**Figure 7**. (Instantaneous) Granger causality between the premature mortality rates of European countries

In this part of analysis, we conducted 930 tests aiming to reveal all causal relationships existing between the $\binom{31}{2}$ pairs of time series belonging to the dataset. The short length of them led us to the selection of $k \leq 3$ because $k = 3$ implies that for the fitting of model (28), 7 unknown parameters should be estimated based on the time series' observations. In some cases

where the length of the time series was less than 18 observations, we tested the existence of causality using only $k = 1$. The residuals of the fitted models satisfied all normality and independence conditions.

Combining the results of the above diagrams, we can identify all types of Granger causality that occur between the 465 time series pairs. Initially, it becomes clear that there is no symmetry to the property of Granger causality between two time series. Moreover, with regard to the feedback property, only one case, that of Sweden – Ireland, satisfied this property's conditions using a model of order $k = 2$. Also of particular interest is the fact that in three cases, those of Iceland – Denmark, Spain – Luxembourg and the UK – Spain, we observed simple and instantaneous Granger causality simultaneously. None of the 465 time series pairs displayed feedback along with instantaneous Granger causality.

### 4.7. (Instantaneous) Granger causality of Europe's average mortality and the rest of European countries

Similar procedure was followed for the examination of causal relations between the average mortality rates in Europe and the other 31 European countries' rates. The analysis revealed 12 Granger causality and 5 instantaneous causality relations regarding the average European mortality rates.

Statistically significant causality occurred when the time series of Ireland ($p = 0.0482$, $k = 2$), Latvia ($p = 0.0352$, $k = 1$) and Slovakia ($p = 0.0459$, $k = 2$) were used as an effect. In these occasions, the average rates of Europe could be utilized to support the forecasting process of these countries.

On the other hand, the results of statistically significant Granger causality are more interesting when we use Europe's time series this time as an effect. The time series of the countries that showed statistically significant Granger causality having the role of cause, were Cyprus ($p = 0.0447$, $k = 1$), Latvia ($p = 0.0006$, $k = 1$), Austria ($p = 0.0173$, $k = 1$), Slovenia ($p = 0.0122$, $k = 3$), the UK ($p = 0.0056$, $k = 2$) and Norway ($p = 0.03116$, $k = 1$).

The case of Latvia is of particular interest, as there is an obvious feedback between this country's rates and the average rates of Europe. The important feature that accompanies the case where the European time series was used as a result, is that the countries mentioned above, offer important information in predicting the behavior that characterizes the mean mortality rates of Europe in the coming years. We further emphasize that the interpretation of the results concerns only the system dynamics and not the non-stationary time series.

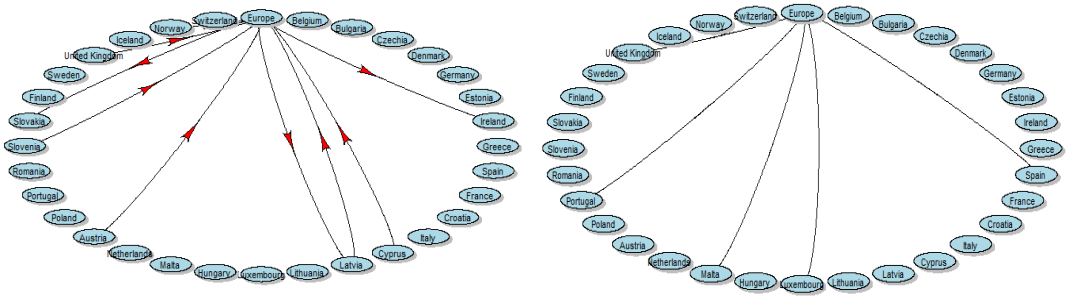


**Figure 8**. (Instantaneous) Granger causality between Europe's average mortality and mortality rates of European countries

## 5. DISCUSSION

In our analysis we presented the clustering results on a map, aiming to examine how the 6 clusters are distributed geographically. We conclude from Figure 4, that as we move from western to eastern Europe, mortality rates from chronic diseases increase, as Eastern European countries have been included in clusters 4 to 6 corresponding to mortality levels that are above average. At the same time, as we go either north or south on this map there is a decrease in mortality rates as Sweden, Norway, Iceland, Italy and Cyprus belong to the 1st subgroup, that is associated with the lowest mortality rates. Finland is ranked in the 2nd group showing slightly higher levels, differentiating itself from the rest of the Nordic countries. Quite interesting is the case of Greece, that differs significantly from the other Balkan and Eastern European countries and belongs to the 3rd cluster; however, one would probably expect Greece to belong to groups 1 or 2 due to its dietary habits which do not differentiate from those of the Mediterranean ones, in many areas;

apparently, other factors do also affect the classification in certain clusters.

We will try to interpret the results of this grouping by investigating social, economic and environmental factors that have probably played an important role in shaping these mortality rates. In order to define the possible causes that result in the deterioration of patients' health suffering from the aforementioned diseases, we should look for the reasons that cause this deterioration, but also search for the factors that have been proven to cause the occurrence of such diseases. It is obvious that mortality rates are increased in countries where the incidence of chronic diseases is also increased.

In terms of eating habits, the beneficial effect of the Mediterranean diet based on olive oil and low – saturated fats on human health is well known. In Romagnolo and Selmin [41] and Georgoulis et al. [42] a systematic study has been carried out according to which, the adoption of the Mediterranean diet, seems to significantly reduce the incidence of malignancies and diabetes, while in the case of patients already suffering from these diseases, mortality rates are also reduced provided that the patients follow these dietary patterns. We observe, therefore, that the Mediterranean countries (Italy, Cyprus, Malta, France, Spain) belong to clusters 1 and 2 that have the lowest mortality rates confirming the above hypothesis.

Cluster 1 includes 2 Mediterranean countries, the Scandinavian countries Sweden and Norway, Iceland, while Finland belongs to the next group of mortality rates. The significantly low mortality levels in these countries are probably associated with the increased omega-3 fatty diet of the inhabitants of these countries, due to the abundance of fishery products that exist in those areas. More specifically, Freitas and Campos [43] show that there is a statistically significant drop in cancer incidence levels when consuming sufficient amounts of Ω3 fats. Their contribution seems to be beneficial in the case of chronic liver disease (Shi et al. [44]) and in the reduction of various strokes (Bu et al. [45]), because they are claimed to enhance the regeneration of brain cells, as well as cardiac ischemia (Ajith and Jayakumar [46]). On the other hand, according to Simopoulos [47], the high omega-6 fat content of the diet of Western European countries' inhabitants, in contrast to that of the Nordic countries, leads to an increase in the incidence of various cancers, while it also

worsens the health of those already suffering. Equally aggravating is the tendency in those countries where people base the necessary fat intake on animal rather than vegetable fats. In Bobak et al. [48], Britton and McKee [49] and Pomerleau et al. [50], authors emphasize that smoking, rich in saturated fats and low in antioxidants nutrition, as well as alcohol, play a key role for the high levels of cardiovascular disease found in eastern countries compared to those in western Europe. In addition, the relatively high level of health care provided in western and northern European countries can be seen to keep premature mortality rates from chronic diseases low, which is confirmed by the fact that these countries belong to the top 3 subgroups in contrast to the eastern European countries that make up subgroups 4,5 and 6.

Returning now to the case of Greece, we have to consider smoking and the consumption of alcoholic beverages as two of the main causes of the deterioration of a patient's health. Both of these factors have been shown to cause cancer, while there is close association between smoking and deteriorating of patients' health with chronic respiratory problems (U.S. Department of Health and Human Services [51]), and also association between alcohol and the health of patients with liver or brain diseases, as reported in numerous studies (Osna et al. [52] and Klatsky et al. [53]). According to published results of Eurostat surveys for 2017 (Eurostat [54]), Greece is in the 1st place in terms of rates of smokers aged 15 and over, a position that has been maintained since 2012. Equally high are the rates of countries such as Croatia, Bulgaria, France, Latvia, Poland, etc., while the rest of the Mediterranean countries are characterized by much lower rates. The rates are low also for the Nordic countries, while Sweden has the lowest rates of smokers.

According to a study conducted in 9 European countries, an increase of one liter in per capita alcohol consumption, can cause up to 3 in 4 additional liver deaths per 100,000 men and 1 additional death per 100,000 women (Ramstedt [55]). Excessive alcohol consumption in southern European countries appeared to be statistically lower than in Western European countries, a fact that verifies the inclusion of southern countries in the first two subgroups, with the exception of Greece. In some cases, such as those in Ireland and Finland, alcohol overdoses are up to three times higher than in Italy, in contrast to Sweden's rates that seem to be close to those of Italy,

reinforcing the inclusion of Finland in the 2nd cluster of the analysis and not in the 1st like its neighboring country. At the same time, the rates of alcohol abuse in the Baltic countries were presented as similar to those of Finland, as reported in Helasoja et al. [56].

In the field of medical care, the superiority of the health systems of the countries of western and northern Europe enables the extension of the life expectancy of patients with chronic diseases. An interesting statistic presented by Eurostat in the 2018 edition, is constituted by the rates of self-reported cases of unfulfilled medical care. In the corresponding table for 2016, Greece is in the 2nd place behind Estonia, while in 2016 Greece was much further behind Bulgaria, Poland, Romania and Latvia, which as we saw above belong to the two most burdened groups 5 and 6. We consider that this observation probably affects the differentiation of Greece from the rest of the Mediterranean countries compared to the rates of premature mortality.

Undoubtedly, the income of the inhabitants of European countries is a vital factor for the formation of the rates of premature mortality. The countries of western and northern Europe are known for their increased earnings, compared to countries of central and eastern Europe, such as those of the Balkans, the countries of the former Soviet Union and countries such as Poland, Czech Republic and Hungary. It is no coincidence that the income factor affects the proposed clustering, as it may not be responsible for the reduced incidence of chronic diseases, but plays a key role in maintaining patients' living standards and extending their life expectancy, especially when we have to deal with diseases whose treatment requires the continuous purchase of drugs and monitoring by specialists, thus burdening patients' financial situation (Marmot [57]).

In addition, it becomes necessary to highlight some environmental conditions that, in parallel with the above factors, justify the grouping that emerged through the Fuzzy K – means algorithm. Many studies have attributed the increased incidence of chronic diseases, mainly cancerous tumors (Cardis, Krewski et al. [58] and Cardis, Howe et al. [59]) among the inhabitants of countries around Ukraine, to the radiation released into the atmosphere in 1986 after the explosion of the Chernobyl nuclear reactor. Looking at Figure 4, we can see that these neighboring countries for which we have data from Eurostat statistics seem to be burdened, while as we move away the mortality rates

are much more moderate. Looking at the Eurostat's table of greenhouse gas emissions in the 2018 edition (Eurostat [54]), which affects the health of people with respiratory problems, we see increased emissions mainly in countries of the 2nd and 3rd subgroups, such as Luxembourg, Belgium, Ireland, the Netherlands and Finland. Limiting these emissions would most likely lead to their inclusion in subgroups 1 and 2 respectively.

Regarding the wider descending trend observed in all EU countries, one of the most probable causes is the significant improvement of medical equipment and the strengthening of preventive medicine as mentioned in Gavurova et al. [60]. Special emphasis has been placed on the deteriorated mortality levels from cardiovascular diseases through the beneficial use of medical technology – methodology and in general the improvement of the conditions of the health systems in countries such as Lithuania, Sweden and Norway (Lisauskiene et al. [61]). In parallel, Meslé [62] underlines that the downward trend in deaths from cardiovascular diseases is due to the multifaceted development of medicine in areas such as improving systemic prognosis, the adoption of screening, but also the spread of new treatment techniques and surgeries.

The most common cancers in Europe during 2006 were breast cancer, accounting for 14% of all cases, prostate cancer with 13.2%, colorectal cancer with 13% and lung cancer accounting for 11.6% of all cases of tumor incidents (Ferlay [63]). Based on Boyle [64] and Sant et al. [65] it is considered that a vital role in the treatment of breast cancer, in countries such as the UK, had the briefing of the citizens but also the establishment of screening in combination with the more efficient methods of treatment. Even in countries such as Spain and Slovakia, where screening has not been adopted at the same scale, there is an improvement in the medical practices. Tumors can now be diagnosed at earlier stages than in the past, due to the constant improvement in the accuracy of diagnostic tools such as MRI scans, while surgery is used in advanced cases. For example, during the existence of a prostate cancer the treatment can be relayed on the usage of radiation. For the above reasons, mortality rates from cancerous tumors, such as breast cancer, are declining even in eastern EU countries (Tyczynski et al. [66]). As mentioned above, the incidence of lung cancer, as well as other chronic respiratory diseases, seems to be related to tobacco smoking.

Considering this factor, we realize that the reduction of active or passive smoking will continuously lead to a reduction in premature deaths from respiratory diseases. Thus, the anti-smoking law in all closed and some open spaces in EU, first discussed in 2002 and widely implemented since 2009, is likely to lead to lower rates of premature mortality from respiratory diseases over the next years.

In addition, especially in the last two decades, artificial intelligence reinforced the fields of preventive and diagnostic medicine. In 2008, following the standards of the Association for the Advancement of Medical Instruments and the British and Irish Hypertension Society, an innovative blood pressure tool was developed based on a combination of convolutional and recurrent neural networks to solve the problem of the extraction of characteristic pulse waveforms of low strength. Using the generated results of this algorithm, the diagnosis of the condition of the patient's heart becomes more efficient (Zhang et al. [67]).

Furthermore, there has been a significant improvement in the quality of MRI scans and diagnostic measures in general using image or video. More specifically, the case of cervical cancer, where its existence in the early stages is mainly not accompanied by symptoms, makes the need for early diagnosis urgent, something that can be achieved more commonly with the use of convolutional neural networks (Bielak et al. [68]). Also, it is possible to enhance the video-based diagnostic process, as in the case of Barrett's esophagus, a precancerous condition that begins in the mucosa of the lower esophagus and is due to chronic gastroesophageal reflux (Hashimoto et al. [69]). Equally important is the contribution to the prediction of ischemic strokes and thromboembolism, as reported in Li et al. [70]. A general summary of the contribution of artificial intelligence to medicine is given in Rong et al. [71] and Becker [72].

## 6. CONCLUSIONS

Public health is a field of research with many benefits for the inhabitants and the community of the examined countries or continents. The capability of forecasting, interpretation and

investigation of causality relationship between factors of a variety of scientific fields may be proven beneficial for the living standards and life expectancy of people nowadays.

The proposed in this study ARIMA and SETAR models that are presented for each of the 31 time series of European countries should be utilized to make predictions for the upcoming years, acquiring a practical tool for the understanding of premature mortality in these countries. A decrease in the predicted mortality rates of chronic diseases can be deemed as an encouraging hint for public health and medicine, although, an increase should raise the awareness of health systems. The proposed AR(2) model for Europe's average mortality levels, shows that the prediction of a year's mortality can be based on the average mortality of the two previous years.

After the implementation of the clustering procedure, we gained the ability to examine the magnitudes of mortality rates in European countries and distinguish the differentiation that exists between western and eastern Europe. The usage of Jensen – Shannon divergence reinforces this attempt, providing an indicator of the differentiation within clusters. Furthermore, research in the fields of nutrition, economy, environment, public health and medicine facilitate the interpretation of the proposed division validating the credibility – reliability of the produced results. Quite interesting is the presentation of an EU country -i.e., Slovenia- as the best indicator of the average premature mortality in Europe, a fact that is supported unanimously by two statistical criteria.

The probe of the existence of correlation and causality between all the time series of this research, enhances the general understanding of the relationships that accompany EU countries' mortality rates proposing more robust predictions concerning the future mortality of each country. Finally, the case of Latvia, displays the ability to use the rates of an EU country to efficiently predict EU's average mortality and vice versa.

# APPENDIX

**Table 4**. Spearman coefficients with $r_s \geq 0.55$

| $\{x_t\}$ | $\{y_t\}$ | $r_s$ | p-value |
|---|---|---|---|
| **Belgium** | Croatia | 0.76 | 0.001 |
| **Bulgaria** | Cyprus | 0.62 | 0.029 |
| **Czech Republic** | Switzerland | 0.59 | 0.005 |
| **Denmark** | Croatia | -0.55 | 0.033 |
| | Iceland | 0.77 | < 0.001 |
| **Germany** | Croatia | 0.77 | < 0.001 |
| **France** | Romania | 0.59 | 0.021 |
| | Finland | -0.61 | 0.015 |
| **Croatia** | Slovakia | 0.62 | 0.013 |
| **Italy** | Cyprus | 0.64 | 0.022 |
| **Cyprus** | Luxembourg | -0.75 | 0.004 |
| | Hungary | 0.77 | 0.002 |
| | Poland | 0.64 | 0.018 |
| | Slovenia | 0.69 | 0.011 |
| | Slovakia | 0.68 | 0.013 |
| **Latvia** | Lithuania | 0.56 | 0.008 |
| **Hungary** | Slovakia | 0.55 | 0.009 |
| **Malta** | Holland | 0.65 | < 0.001 |
| **Austria** | UK | 0.60 | 0.004 |
| **Finland** | Norway | 0.59 | 0.004 |

**Table 5**. Statistically significant distance correlations with $R_n \geq 0.55$

| $\{x_t\}$ | $\{y_t\}$ | $R_n$ | p-value |
|---|---|---|---|
| **Belgium** | Croatia | 0.69 | < 0.001 |
| | Germany | 0.58 | 0.025 |
| **Czech Republic** | Spain | 0.55 | 0.002 |
| | Croatia | 0.56 | 0.022 |
| **Denmark** | Germany | 0.61 | 0.042 |
| | Croatia | 0.61 | 0.022 |
| | Iceland | 0.80 | < 0.001 |
| **Germany** | Croatia | 0.76 | < 0.001 |
| **France** | Romania | 0.66 | < 0.001 |
| **Spain** | Luxembourg | 0.56 | < 0.001 |
| **Italy** | Cyprus | 0.60 | 0.011 |
| | Hungary | 0.73 | 0.012 |
| | Slovakia | 0.73 | < 0.001 |
| **Cyprus** | Luxembourg | 0.79 | < 0.001 |
| | Slovenia | 0.63 | < 0.001 |
| | Slovakia | 0.66 | < 0.001 |
| | Iceland | 0.56 | 0.043 |
| **Latvia** | Lithuania | 0.73 | < 0.001 |
| | Austria | 0.71 | < 0.001 |
| | UK | 0.69 | < 0.001 |
| **Lithuania** | UK | 0.58 | 0.009 |
| **Luxembourg** | Slovakia | 0.67 | 0.006 |
| **Hungary** | Poland | 0.55 | < 0.001 |
| | Slovakia | 0.59 | 0.018 |
| **Malta** | Holland | 0.64 | < 0.001 |
| **Austria** | UK | 0.69 | < 0.001 |
| **Finland** | Norway | 0.57 | < 0.001 |

**Table 6**. Cross – correlations between the examined mortality rates

| $\{x_t\}$ | $\{y_t\}$ | $r_{XY}$ | Lag (τ) |
|---|---|---|---|
| **Belgium** | Bulgaria | 0.541 | 3 |
| | Croatia | 0.748 | 0 |
| | Cyprus | 0.481, - 0.573 | 0, 2 |
| **Bulgaria** | France | -0.508 | 1 |
| | Croatia | 0.548 | -3 |
| | Italy | 0.511 | -3 |
| | Malta | -0.563 | 1 |
| | Switzerland | -0.540 | 1 |
| **Czech Republic** | Cyprus | 0.551, -0.614 | -2, -3 |
| **Denmark** | Italy | 0.539 | -3 |
| | Iceland | 0.771 | 0 |
| **Germany** | Croatia | 0.544, 0.734 | -1, 0 |
| | Malta | 0.565 | 3 |
| | Slovenia | 0.500 | -2 |
| **Estonia** | Malta | -0.620 | -3 |
| | Slovakia | -0.511 | -2 |
| | Sweden | 0.506 | -2 |
| **Ireland** | Malta | 0.518 | 1 |
| | Slovakia | -0.590 | 3 |
| **Greece** | Hungary | 0.644 | -1 |
| | Finland | 0.734 | 0 |
| | Sweden | 0.565 | 3 |
| | Norway | 0.508 | 1 |
| **Spain** | Luxembourg | 0.584 | -1 |
| | UK | 0.532 | -1 |
| | Norway | 0.509, 0.514 | -2, 1 |
| **France** | Holland | -0.532 | -4 |
| | Romania | 0.562 | 0 |
| | Finland | -0.562 | 0 |

| | | | |
|---|---|---|---|
| **Croatia** | Malta | 0.672 | 3 |
| | Holland | 0.554 | 3 |
| | Portugal | -0.503 | 2 |
| | Finland | 0.505 | -3 |
| **Italy** | Cyprus | 0.505 | 0 |
| | Iceland | -0.535, 0.511, 0.484 | 4, 3, 1 |
| | Norway | -0.594, 0.692 | -3, -4 |
| **Cyprus** | Luxembourg | -0.697 | 0 |
| | Poland | 0.560 | 0 |
| | Slovenia | 0.675 | 0 |
| | Slovakia | 0.602 | 0 |
| | Sweden | 0.525 | -3 |
| **Latvia** | Lithuania | 0.636 | 0 |
| | UK | 0.638, -0.494 | 0, 2 |
| **Luxembourg** | Slovakia | -0.516 | 0 |
| **Malta** | Holland | 0.642 | 0 |
| | Slovakia | -0.512 | 2 |
| **Holland** | Slovakia | 0.532 | 1 |
| **Austria** | UK | -0.560 | 0 |
| **Poland** | Portugal | -0.504 | 0 |
| | Slovakia | 0.510 | 0 |
| | Sweden | -0.615 | -3 |
| **Portugal** | Romania | 0.523 | 0 |
| **Romania** | Finland | -0.554 | 0 |
| **Slovakia** | Sweden | -0.511 | -3 |
| | UK | 0.580 | -3 |
| **UK** | Switzerland | -0.631 | 2 |
| **Sweden** | Norway | 0.507 | -2 |

**Table 7**. Pairs of mortality time series with statistically significant Granger causality.

| $\{x_t\}$ (cause) | $\{y_t\}$ (effect) | p-value | Order ($k$) |
|---|---|---|---|
| **Bulgaria** | Denmark | 0.0238 | 2 |
| | Lithuania | 0.0353 | 1 |
| | UK | 0.0357 | 1 |
| **Czech Republic** | Belgium | 0.0351 | 2 |
| | Sweden | 0.0439 | 1 |
| | UK | 0.0353 | 3 |
| | Iceland | 0.0106 | 1 |
| **Denmark** | Spain | 0.0157 | 2 |
| | Slovakia | 0.0035 | 1 |
| | Switzerland | 0.0469 | 1 |
| **Germany** | Denmark | 0.0034 | 3 |
| | Slovenia | 0.0277 | 2 |
| **Estonia** | Spain | < 0.0001 | 1 |
| | Hungary | 0.0352 | 1 |
| | Malta | 0.0086 | 3 |
| | Sweden | 0.0009 | 2 |
| **Ireland** | Croatia | 0.0465 | 1 |
| | Sweden | 0.0356 | 2 |
| **Greece** | France | 0.0363 | 1 |
| | Hungary | 0.0013 | 2 |
| | Finland | < 0.0001 | 2 |
| | Latvia | 0.0111 | 1 |
| **Spain** | Lithuania | 0.0012 | 1 |
| | Sweden | 0.0206 | 1 |
| **France** | Bulgaria | 0.0300 | 1 |
| **Croatia** | Finland | 0.0482 | 1 |
| **Italy** | Denmark | 0.0444 | 3 |
| | Iceland | 0.0214 | 1 |
| **Cyprus** | Luxembourg | 0.0421 | 1 |
| | Romania | 0.0242 | 1 |
| | Iceland | 0.0335 | 1 |
| **Lithuania** | Slovakia | 0.0371 | 2 |
| **Luxembourg** | Hungary | 0.0286 | 2 |
| | Slovenia | 0.0223 | 2 |
| **Hungary** | Germany | 0.0447 | 2 |
| | Italy | 0.0377 | 1 |

| | | | |
|---|---|---|---|
| | Latvia | 0.0158 | 1 |
| | Malta | 0.0461 | 2 |
| | Holland | 0.0286 | 2 |
| **Malta** | Bulgaria | 0.0010 | 1 |
| | Czech Republic | 0.0317 | 1 |
| | Ireland | 0.0089 | 1 |
| | Austria | 0.0354 | 1 |
| | Slovakia | 0.0105 | 1 |
| | Switzerland | 0.0224 | 3 |
| **Holland** | Germany | 0.0213 | 2 |
| **Austria** | Spain | 0.0127 | 1 |
| | UK | 0.0443 | 1 |
| **Poland** | Czech Republic | 0.0406 | 3 |
| | Germany | 0.0032 | 3 |
| **Portugal** | Belgium | 0.0126 | 3 |
| **Slovenia** | Denmark | 0.0437 | 3 |
| | Latvia | 0.0172 | 1 |
| **Slovakia** | Bulgaria | 0.0340 | 2 |
| | Holland | 0.0066 | 1 |
| **Sweden** | Germany | 0.0382 | 2 |
| | Ireland | 0.0199 | 2 |
| | France | 0.0467 | 1 |
| | Norway | 0.0418 | 1 |
| **UK** | Spain | 0.0366 | 2 |
| | Latvia | 0.0006 | 2 |
| | Switzerland | 0.0388 | 2 |
| **Iceland** | Bulgaria | 0.0150 | 1 |
| | Denmark | 0.0311 | 1 |
| **Norway** | Czech Republic | 0.0351 | 2 |
| | Ireland | 0.0087 | 2 |
| | Spain | 0.0036 | 1 |
| | Latvia | 0.0056 | 3 |
| **Switzerland** | Estonia | 0.0036 | 2 |
| | Latvia | 0.0438 | 2 |
| | Lithuania | 0.0029 | 2 |

**Table 8**. Statistically significant instantaneous Granger causalities between mortality rates.

| $\{x_t\}$ | $\{y_t\}$ | p-value | Order ($k$) |
|---|---|---|---|
| **Belgium** | Denmark | 0.0422 | 2 |
| | Croatia | 0.0360 | 1 |
| | Luxembourg | 0.0249 | 3 |
| | Slovenia | 0.0242 | 2 |
| | Slovakia | 0.0142 | 2 |
| **Bulgaria** | Germany | 0.0429 | 3 |
| **Czech Republic** | Spain | 0.0134 | 1 |
| | Switzerland | 0.0396 | 1 |
| **Denmark** | Estonia | 0.0220 | 1 |
| | Iceland | 0.0077 | 1 |
| **Estonia** | Latvia | 0.0293 | 2 |
| | Lithuania | 0.0295 | 1 |
| | UK | 0.0272 | 1 |
| **Spain** | Luxembourg | 0.0249 | 2 |
| | Malta | 0.0346 | 1 |
| | UK | 0.0268 | 2 |
| **France** | Romania | 0.0301 | 1 |
| **Croatia** | Italy | 0.0458 | 1 |
| **Latvia** | Lithuania | 0.0105 | 2 |
| | Austria | 0.0233 | 2 |
| **Lithuania** | UK | 0.0132 | 1 |
| **Luxembourg** | Slovakia | 0.0424 | 1 |
| **Hungary** | Poland | 0.0236 | 2 |
| **Malta** | Holland | 0.0060 | 2 |
| **Poland** | Portugal | 0.0404 | 1 |
| | Slovakia | 0.0450 | 1 |
| | Sweden | 0.0410 | 2 |
| | Switzerland | 0.0270 | 2 |
| **Romania** | Finland | 0.0487 | 1 |
| **Slovenia** | Slovakia | 0.0445 | 1 |
| **Slovakia** | Switzerland | 0.0459 | 2 |

## CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

## REFERENCES

[1] L. Velea, M.T. Udristioiu, R. Bojariu, S.C. Sararu, L. Prunariu, The influence of climate conditions on the mortality related to cardiovascular diseases in Dolj county (Southern Romania), in: Heidelberg, Germany, 2017: p. 040003.

[2] L. Yilan, C. Min, C. Guowei, W. Xiaoqing, L. Tianquan, Application of an autoregressive integrated moving average model for predicting injury mortality in Xiamen, China, BMJ Open, 5(12) (2015), e008491.

[3] D.A. Adeyinka, N. Muhajarine, Time series prediction of under-five mortality rates for Nigeria: comparative analysis of artificial neural networks, Holt-Winters exponential smoothing and autoregressive integrated moving average models, BMC Med. Res. Methodol. 20 (2020), 292.

[4] B. Hu, W. Qiu, C. Xu, J. Wang, Integration of a Kalman filter in the geographically weighted regression for modeling the transmission of hand, foot and mouth disease, BMC Public Health. 20 (2020), 479.

[5] S. Megyesiova, V. Lieskovska, Premature Mortality for Chronic Diseases in the EU Member States, International J. Environ. Res. Public Health, 16(20) (2019), 4021.

[6] I. Pratama, A.E. Permanasari, I. Ardiyanto, R. Indrayani, A review of missing values handling methods on time-series data, in: 2016 International Conference on Information Technology Systems and Innovation (ICITSI), IEEE, Bandung - Bali, Indonesia, 2016: pp. 1–6.

[7] A. Papana, C. Kyrtsou, D. Kugioumtzis, C.G.H. Diks, Financial networks based on Granger causality: a case study, Physica A: Stat. Mech. Appl. 482 (2017), 65–73.

[8] E. Siggiridou, V. Kimiskidis, D. Kugiumtzis, Dimension reduction of frequency-based direct Granger causality measures on short time series, J. Neurosci. Meth. 289 (2017), 64–74.

[9] D. Kugiumtzis, Time series and complex networks, in: 26th Summer School and Conference on Dynamical Systems and Complexity, (2019).

[10] E. Siggiridou, C. Koutlis, A. Tsimpiris, D. Kugiumtzis, Evaluation of Granger causality measures for constructing networks from multivariate time series, Entropy, 21(11) (2019), 1–24.

[11] J. Fattah, L. Ezzine, Z. Aman, H.E. Moussami, A. Lachhab, Forecasting of demand using ARIMA model, Int. J. Eng. Bus. Manage. 10(2) (2018), 184797901880867.

[12] D. Atance, A. Debón, E. Navarro, A comparison of forecasting mortality models using resampling methods,

Mathematics, 8(9) (2020), 1550.

[13] Y. Shi, Forecasting mortality rates with the adaptive spatial temporal autoregressive model, J. Forecast. 40(3) (2020), 528–546.

[14] P.T. De Bruin, J.G. De Gooijer, A comparison of ARMA and SETAR Forecasts, (1998). http://www.jandegooijer.nl/uploads/8/3/2/7/83270090/report-ae-6-98.pdf

[15] E.H. Firat, SETAR (Self-exciting Threshold Autoregressive) non-linear currency modelling in EUR/USD, EUR/TRY and USD/TRY parities, Math. Stat. 5(1) (2017), 33–55.

[16] G. Boero, F. Lampis, The forecasting performance of SETAR models: an empirical application, Bull. Econ. Res. 69(3) (2016), 216–228.

[17] V.K. Malhotra, H. Kaur, A. Alam, An analysis of fuzzy clustering methods, Int. J. Computer Appl. 94(19) (2014), 9–12.

[18] I. Mohamad, D. Usman, Standardization and its effects on K-means clustering algorithm, Research Journal of Applied Sciences, Engineering and Technology 6(17), (2013), 3299–3303.

[19] S. Nasser, R. Alkhaldi, G. Vert, A modified fuzzy K-means clustering using expectation maximization, in: 2006 IEEE International Conference on Fuzzy Systems, IEEE, Vancouver, BC, Canada, 2006: pp. 231–235.

[20] F. Höppner, F. Klawonn, R. Kruse, T. Runkler, Fuzzy Cluster Analysis, John Wiley & Sons, Inc., New York, (1999).

[21] V. Schwämmle, O.N. Jensen, A simple and fast method to determine the parameters for fuzzy c–means cluster validation, Bioinformatics, 26(22) (2010), 2841–2848.

[22] D.I. Belov, R.D. Armstrong, Distributions of the Kullback–Leibler divergence with applications, Br. J. Math. Stat. Psychol. 64(2) (2011), 291–309.

[23] H.-V. Nguyen, J. Vreeken, Non-parametric Jensen-Shannon Divergence, in: A. Appice, P.P. Rodrigues, V. Santos Costa, J. Gama, A. Jorge, C. Soares (Eds.), Machine Learning and Knowledge Discovery in Databases, Springer International Publishing, Cham, 2015: pp. 173–189.

[24] M.L. Menéndez, J.A. Pardo, L. Pardo, M.C. Pardo, The Jensen – Shannon divergence, J. Franklin Inst. 334(2) (1997), 307–318.

[25] M. Levene, A. Kononovicius, Empirical survival Jensen-Shannon divergence as a goodness-of-Fit measure for

maximum likelihood estimation and curve fitting, Commun. Stat. – Simul. Comput. (2019), https://doi.org/10.1080/03610918.2019.1630435.

[26] D.M. Endres, J.E. Schindelin, A new metric for probability distributions, IEEE Trans. Inform. Theory 49(7) (2003), 1858–1860.

[27] S. Kodera, E.A. Rashed, A. Hirata, Correlation between COVID-19 morbidity and mortality rates in Japan and local population density, temperature, and absolute humidity, Int. J. Environ. Res. Public Health, 17(15) (2015), 5477.

[28] L.-L. Liang, C.-H. Tseng, H.J. Ho, C.-Y. Wu, Covid-19 mortality is negatively associated with test number and government effectiveness, Sci Rep. 10 (2020), 12567.

[29] G.J. Székely, M.L. Rizzo, N. Bakirov, Measuring and testing dependence by correlation of distances, Ann. Stat. 35(6) (2007), 2769–2794.

[30] G.J. Székely, M.L. Rizzo, The distance correlation t-test of independence in high dimension, J. Multivar. Anal. 117 (2013), 193–213.

[31] M.N.O. Sadiku, S.M. Musa, S. Nelatury, R.G. Perry, Correlation: a brief introduction, Int. J. Electric. Eng. Educ. 51(2) (2014), 93–99.

[32] M.W. Robbins, T.J. Fisher, Cross-correlation matrices for tests of independence and causality between two multivariate time series, J. Bus. Econ. Stat. 33(4) (2015), 459–473.

[33] P. Damos, Using multivariate cross correlations, Granger causality and graphical models to quantify spatiotemporal synchronization and causality between pest populations, BMC Ecol. 16 (2016), 33.

[34] G.P. Papaioannou, C. Dikaiakos, C. Kaskouras, G. Evangelidis, F. Georgakis, Granger causality network methods for analyzing cross-border electricity trading between Greece, Italy, and Bulgaria, Energies, 13(4) (2020), 900.

[35] C. Koutlis, Evaluation of Granger causality measures for constructing networks from multivariate time series with applications to neurophysiology, National Archive of Doctoral Dissertations, (2017).

[36] Q.G. Gianetto, Testing instantaneous causality in presence of nonconstant unconditional variance, J. Bus. Econ. Stat. 33(1) (2012), 46-53.

[37] C. Robertson, R. Ecob, Simultaneous modelling of time trends and regional variation in mortality rates, Int. J. Epidemiol. 28(5) (1999), 955–963.

[38] G. Kirchgässner, J. Wolters, Introduction to modern time series analysis, Springer, New York, (2013).

[39] L. Barnett, A.K. Seth, The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference, J. Neurosci. Meth. 223 (2014), 50–68.

[40] A. Hyvärinen, K. Zhang, S. Shimizu, P.O. Hoyer, Estimation of a structural vector autoregression model using non-gaussianity, J. Mach. Learn. Res.11(56) (2010), 1709–1731.

[41] D.F. Romagnolo, O.I. Selmin, Mediterranean diet and prevention of chronic diseases, Nutrition Today, 52(5) (2017), 208–222.

[42] M. Georgoulis, M.D. Kontogianni, N. Yiannakouris, Mediterranean diet and diabetes: prevention and treatment, Nutrients, 6(4), (2014) 1406–1423.

[43] R.D.S. Freitas, M.M. Campos, Protective effects of omega-3 fatty acids in cancer-related complications, Nutrients, 11(5) (2019), 945.

[44] X.Y. Shi, S.M. Fan, G.M. Shi, J. Yao, Y. Gao, Y. Xia, Q. Chen, Efficacy and safety of omega-3 fatty acids on liver-related outcomes in patients with nonalcoholic fatty liver disease: A protocol for a systematic review and meta-analysis, Medicine, 99(24) (2020), e20624.

[45] J. Bu, Y. Dou, X. Tian, Z. Wang, G. Chen, The role of omega-3 polyunsaturated fatty acids in stroke, Oxidat. Med. Cellular Longevity. 2016 (2016), 6906712.

[46] T.A. Ajith, T.G. Jayakumar, Omega-3 fatty acids in coronary heart disease: recent updates and future perspectives, Clinic. Experiment. Pharmacol. Physiol. 46(1) (2018), 11–18.

[47] A. Simopoulos, Omega6/omega3 essential fatty acid ratio and chronic diseases, Food Rev. Int. 20(1) (2004), 77–90.

[48] M. Bobak, Z. Skodova, Z. Pisa, R. Poledne, M. Marmot, Political changes and trends in cardiovascular risk factor in Czech Republic, 1985-92, J. Epidemiol. Commun. Health, 51(3) (1997), 272–277.

[49] A. Britton, M. McKee, The relationship between alcohol and cardiovascular disease in Eastern Europe: Explaining the paradox, J. Epidemiol. Commun. Health, 54(5) (2000), 328–332.

[50] J. Pomerleau, M. McKee, A. Robertson, K. Kadziauskiene, A. Abaravicius, S. Vaask, I. Pudule, D. Grinberga, Macronutrient and food intake in the Baltic republics, Eur. J. Clinic. Nutrition, 55(3) (2001), 200–207.

[51] U.S. Department of Health and Human Services, The health consequences of smoking: 50 years of progress, A

report of the surgeon general. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, (2014).

[52] N.A. Osna, T.M.Jr Donohue, K.K. Kharbanda, Alcoholic liver disease: pathogenesis and current management, Alcohol Res. 38(2) (2017), 147–161.

[53] A.L. Klatsky, M.A. Armstrong, G.D. Friedman, Alcohol use and subsequent cerebrovascular disease hospitalizations, Stroke, 20(6) (1989), 741–746.

[54] Eurostat, Sustainable development in the European union: monitoring report on progress towards the SDGS in an eu context, 2018 edition, (2018).

[55] M. Ramstedt, Population drinking and liver cirrhosis mortality Is there a link in eastern Europe? Addiction 102(8) (2007), 1212–1223.

[56] V. Helasoja, E. Lahelma, R. Prättälä, J. Klumbiene, I. Pudule, M. Tekkel, Trends in the magnitude of educational inequalities in health in Estonia, Latvia, Lithaunia and Finland during 1994 – 2004, Public Health, 120(9) (2006), 841-853.

[57] M.G. Marmot, The influence of income on health: views of an epidemiologist, Health Affairs, 21(2) (2002), 31–46.

[58] E. Cardis, D. Krewski, M. Boniol, V. Drozdovitch, S.C. Dardy, E.S. Gilbert, et al., Estimates of the cancer burden in Europe from radioactive fallout from the Chernobyl accident, Int. J. Cancer 119(6) (2006), 1224–1235.

[59] E. Cardis, G. Howe, E. Ron, V. Bebeshko, T. Bogdanova, A. Bouville, Z. Carr, et al., Cancer consequences of the Chernobyl accident: 20 years on, J. Radiol. Protect. 26(2) (2006), 127–140.

[60] B. Gavurova, D. Tucek, V. Kovac, Investigation of Relationship Between Spatial Distribution of Medical Equipment and Preventable Mortality, Int. J. Environ. Res. Public Health, 16(16) (2019), 2913.

[61] I. Lisauskiene, K. Garuoliene, J. Gulbinovic, Utilization of Cardiovascular Medicines and Cardiovascular in Lithuania, Sweden and Norway in 2003 – 2012, Medicina 53(4) (2017), 259–267.

[62] F. Meslé, Mortality in eastern Europe and the former Soviet Union: Long-term trends and recent Upturns, IUSSP/MPIDR Workshop “Determinants of diverging trends in mortality”, (2002), 19–21.

[63] S. Ferlay, et al., Recrudescent tuberculosis in Russian Federation, The global impact of drugs-resistant

tuberculosis, Harvard Medical School/Open Society Institute, (2003), 39–83.

[64] P. Boyle, Breast cancer control: signs of progress, but more work required, Breast, 14(6) (2003), 429-438.

[65] M. Sant, S. Francisci, R. Capocaccia, A. Verdecchia, C. Allemani, F. Berrino, Time trends of breast cancer survival in Europe in relation to incidence and Mortality, Int. J. Cancer, 119(10) (2007), 2417–2422.

[66] J.E. Tyczynski, I. Plesko, T. Aareleid, M. Primic-Zakelj, M. Dalmas, J. Kurtinaitis, A. Stengrevics, D.M. Parkin, Breast cancer mortality patterns and time trends in 10 new EU Member States: Mortality declining in young women but stull increasing in the elderly, Int. J. Cancer, 112(6) (2004), 1056–1064.

[67] J.S. Zhang, L.Y. Gu, S.Y. Jiang, Blood pressure measurement model based on convolutional recurrent neural network, Beijing Biomed. Eng. 37(5) (2018), 494–501.

[68] L. Bielak, N. Wiedenmann, A. Berlin, N. Nicolay, D.D. Gunashekar, L. Hägele, T. Lottner, A.L. Grosu, M. Bock, Convolutional neural networks for head and neck tumor segmentation on 7-channel multiparametric MRI: a leave-one-out analysis, Radiation Ontol. 15(1) (2020), 181.

[69] R. Hashimoto, J. Requa, T. Dao, A. Ninh, E. Tran, D. Mai, M. Lugo, N.E. Chehade, K.J. Chang, W.E. Karnes, Artificial intelligence using convolutional neural networks for real-time detection of early esophageal neoplasia in Barrett's esophagus (with video), Gastrointestinal Endoscopy, 91(6) (2020), 1264–1271.

[70] X. Li, H. Liu, X. Du, P. Zhang, G. Hu, G. Xie, S. Guo, X. Meilin, X. Xie, Integrated machine learning approaches for predicting ischemic stroke and thromboembolism in atrial fibrillation, AMIA Ann. Symp. Proc. 2016 (2016), 799–807.

[71] L. Rong, R. Yan, Z. Peng, A review of medical artificial intelligence, Glob. Health J. 4(2) (2020), 42–45.

[72] A. Becker, Artificial intelligence in medicine: What is it doing for us today? Health Policy Technol. 8(2) (2019), 198–205.