Journal of Regional & Socio-Economic Issues

Volume 12, Issue 1, January 2022

ISSN 2049-1409

Guest-Editor

Prof. Dr. Maria Michailidis, University of Nicosia, Cyprus

Table of Contents

Indexed by Copernicus Index, DOAJ (Director of Open Access Journal), EBSCO, Cabell's Index
The journal is catalogued in the following catalogues: ROAD: Directory of Open Access Scholarly Resources, OCLC WorldCat, EconBiz - ECONIS, CITEFACTOR, OpenAccess

JOURNAL OF REGIONAL SOCIO-ECONOMIC ISSUES (JRSEI)

Volume 12, Issue 2, January 2022

Journal of Regional & Socio-Economic Issues (Print) ISSN 2049-1395

Journal of Regional & Socio-Economic Issues (Online) ISSN 2049-1409

Indexed by Copernicus Index, DOAJ (Director of Open Access Journal), EBSCO, Cabell's Index

The journal is catalogued in the following catalogues: ROAD: Directory of Open Access Scholarly Resources, OCLC WorldCat, EconBiz - ECONIS, CITEFACTOR, OpenAccess

JOURNAL OF REGIONAL SOCIO-ECONOMIC ISSUES (JRSEI)

ISSN No. 2049-1409

Aims of the Journal: Journal of Regional Socio-Economic Issues (JRSEI) is an international multidisciplinary refereed journal the purpose of which is to present papers manuscripts linked to all aspects of regional socio-economic and business and related issues. The views expressed in this journal are the personal views of the authors and do not necessarily reflect the views of JRSEI journal. The journal invites contributions from both academic and industry scholars. Electronic submissions are highly encouraged (mail to: **gkorres@geo.aegean.gr**).

Indexed by Copernicus Index, DOAJ (Director of Open Access Journal), EBSCO, Cabell's Index International Institute of Organized Research (I2OR) database

The journal is catalogued in the following catalogues: ROAD: Directory of Open Access Scholarly Resources, OCLC WorldCat, EconBiz - ECONIS, CITEFACTOR, OpenAccess

Table of Contents

Adopting Information Distance Measures for Geographical Data Analysis

Abstract:
In Geographical Data Analysis there are different measures describing the sense of distance, between two points in R^n. In this paper we adopt the distance measure between the two distributions describing the variable under consideration - the Information Distance Measures (idm) – between two regions. The idm does not obey the triangular inequality and there is, in principle, a problem in the symmetry, as well. We are working with those which can satisfy, eventually, the symmetry, while a compact discussion, under a general framework, of the most idm is developed. Eventually a real life statistical problem with a particular Data Analysis is performed for Geographical oriented Economical collected data in the Metropolitan Area of Athens (MAA), applying particular idm which are suggested for this particular kind of data sets.

Key Words: Distance measures, Divergence, Geographical Analysis, Multivariate Normal

Christos P. Kitsos[1] **and Polixeni Iliopoulou**[2]

[1] Corresponding-Address: Prof. Dr. Christos P. Kitsos, University of West Attika, Email: xkitsos@uniwa.gr
[2] Corresponding-Address: Dr. Polixeni Iliopoulou, University of West Attika. Email: piliop@uniwa.gr

1. Introduction

The quantified methods, even empirical, are essential in all areas of investigation. A typical example can be Hack's Law, Hack (1957), an empirical relation between the length λ of a stream and the area A of their basins as:

$$\lambda = c\mathrm{A}^{h} \qquad (1.1)$$

The exponent h plays an important role, known as Hack's exponent, and there is not a theoretical justification for (1.1). The values of h are around 0.52, for every particular case under investigation. But the empirical relation (1.1) provides food for thought for results based on probability theory, Rigon et. al. (1996). That is empirical laws, can be the basis for theoretical results, typical example can be the empirical distribution and normality. In the same manner theoretical results, can be adopted in applications, sometimes different for the ones that are proposed. This is the line of thought we present in this paper: we are referred to "distance measures", emphasis given to the information ones, usually developed and applied for the field of Electronics and provide the way that can be useful in real life data applications coming from Geographical Analysis.

The sense of distance, although fundamental in all fields, it is understood under a different line of thought to each one field, consider as an example Statistics and Geography. Traditionally, in Geography, Abler and Gould (1971), distance is considered either as absolute or relative. The former is the distance between two points, which is measured with conventional distance units (e.g. meters), while the latter represents the relative location, in terms of distances along dimensions. There are different results of distance measurements according to the path between two points. For example, along a transportation network, either absolute or relative distances will be different compared to measurements in Euclidean space, Lu et al. (2014), Comber et al. (2020). In the space of Riemann, the shortest distance between two points is the curved line. A typical example of Riemann's space, in practice, is when we route a highway around a mountain, as it is, in principle, difficult to cross it, through a straight line in Euclidean space. Manhattan distance is a variant of Euclidean space, in which the shortest distance between two points consists of line segments, which meet on right angles. On the other hand, relative distance can be measured in a two dimensional or multi-dimensional space, which is defined by two or more variables, as the coordinate system, instead of conventional distance units. In that case, two places have smaller or larger distance according to their similarities with respect to the variables defining space. Similar values of the variables result to smaller distances between observations.

Spherical distance it is not discussed in this paper. There is a number of Statistical considerations that it is not easy to be transferred in Spatial Statistics and Geographical Analysis, Iliopoulou and Kitsos (2019).

The measure of distance in R^n is adopted either from the principles in Analysis or from Linear Algebra. The latter provides the Euclidean distance in R^n, as well as other measures, while the former provides, in an early stage of development, the distance of a distribution from the empirical one, or the distance between two distributions, as a result of the distance of two functions. Mahalanobis' distance, compromise the above, as the Statistical way of approaching "distance", in the sense that tries to evaluate "how far" the data points are from the Normal distribution. We shall discuss briefly these lines of thought.

The Euclidean distance, as well as the Manhattan distance, is special cases of the Minkowski distance. Let the two points x = (x_i) and y = (y_i) with i=1,2…,n, i.e $x, y \in R^n$, then the Minkowski distance, of order m is, Johnson & Wichern (2007), among others:

$$d_m(x,y) = [\sum_{i=1}^{n} |x_{i-} y_i|^m]^{1/m} \quad (1.2)$$

For m=2 the distance $d_2(x, y)$ is the Euclidean one, while for m=1 we obtain the Manhattan one, which is strongly related to the city-blocks, under investigation:

$$d_1(x,y) = \sum_{i=1}^{n} | x_i - y_i |$$

The exponent m provides different weight on the "spread" of observations. Particularly when m tends to infinity (m→∞), the Chebyshev distance, equals to max {$|x_i - y_i|$, i=1,2,…,n}, is obtained. In principle for a given distance, d_{GIVEN} (x, y), between the points x, y, we can create a new distance, through a function g, d_{NEW} = g (d_{GIVEN} (x, y)), provided g is monotonically increased function. However, the distance axioms might not be satisfied for some forms of g. Moreover, a number of applications have been discussed adopting distance measures, Wilson and Martinez (1997). The Mahalanobis distance, used in Fisher's linear discriminant analysis, is a particular case of the Bhattacharyya distance, Bhattacharyya (1943). The Mahalanobis distance, d_M , the distance of the measurements $x \in R^n$ form the Normal distribution N(μ, Σ) with mean $\boldsymbol{\mu} \in R^n$ and covariance matrix $\boldsymbol{\Sigma} \in R^{nxn}$, Mahalanobis (1936), is defined as :

$$d_M = d_M(x, N(\mu, \Sigma)) = [(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})]^{1/2} \quad (1.3)$$

As it was noted d_M can be also applied even with correlated variables, Klecka (1980), Norušis (2011). Eventually whatever distance measure is adopted, practically the explanation is: "similar" observations provide small distance measure, while large distances correspond to "different" observations.

The sense of distance is also important in other aspects of Statistics. Among the pioneering work one could consider the work of Wolfowitz (1957), Blyth (1970), were it was pointed out that the Maximum Likelihood Estimators (MLE), the Least Square Estimators (LSE), the Chi-Square test the Kolmogorov-Smirnov test obey the minimum distance principle.

It has been defined, that the minimum distance estimators, are those which minimize the distance between the empirical distribution F_n and the actual distribution, F_θ , in terms that:

$$\inf d(F_\theta, F_n) = d(F_{\hat{\theta}}, F_n) \quad (1.4)$$

where $\theta, \hat{\theta} \in \Theta, \Theta$ being the parameter space, is the "minimum distance estimator", and $\hat{\theta}$ is the evaluated estimate. Notice that by the Law of "Large Numbers" the empirical distribution converges a.s to the distribution function F, while the Glivenko–Cantelli theorem insures that the convergence in fact happens uniformly in R, in terms that:

$$\| F_n - F \|_\infty = \sup \{ | F_n(x) - F(x) |, x \in R \} \quad (1.5)$$

Moreover by the Central Limit Theorem (CLT) the empirical distribution has asymptotically normal distribution, with √n rate of convergence, mean zero and variance F(x)(1-F(x)). In principle the discrepancy measure is reflected to $d(F_\theta, F_n)$, for the estimation case under consideration. Typical example can be the Kolmogorov – Smirnov (K-S) test, similar to the above line of thought, Kolmogorov (1933), Smirnov (1948), were:

$$d(F_{\hat{\theta}}, F_n) = \sup_{x \in \Re} |F_\theta(x) - F_n(x)|$$

So it is clear that the K-S statistic quantifies a distance measure between the empirical distribution function, of the collected sample, of size n and the cumulative distribution function of the reference distribution. In the same way of thinking, Least Squares have been considered as Euclidean distance. That is why, we believe, Maximum Likelihood estimation, offered by Fisher (1922) it is an impressed evolution: although no typical distance measure is used, at the end, it is a distance measure. Adopting the distance methods in Bioassays, Kitsos and Sotiropoulos (2009), a number of theoretical results can be proved and used, while the affine transformation preserves all the properties for a class of models, Kitsos et. al. (2014)

In applications not only the K-S test is adopted, but also the Shapiro–Wilk test as a test of normality in frequentist statistics, Shapiro and Wilk (1965). We shall apply both tests in section 3.

So starting from the Euclidean distance between two points, we moved to Minkowski's distance for given points, then the distance of a given point from a Normal distibution, see (1.3), then the distance between the empirical distribution and reference distribution, see (1.5). The next step is to consider the distance between any two distributions : Consider two probability measures *P* and *Q* say, on a sigma-algebra **A** of subsets of the sample space Ω. Then, the (total variation) distance between P and Q is defined :

$$d_{TV}(P, Q) = 2\sup[|P(A) - Q(A)|: A \in A] \quad (1.6)$$

It holds, that:

$$d_{TV}(P, Q) = \| P - Q \|_1 = \Sigma \, | P(\omega) - Q(\omega)|, \ \omega \in \Omega$$

Moreover from the Random-Nikoym theorem there exist functions f, g, respectively to P and Q integrable such that, with μ the appropriate measure:

$$f = \frac{dP}{d\mu}, \quad g = \frac{dQ}{d\mu}$$

So that $$d_{TV}(P,Q) = \int_A |f - g| d\mu \quad (1.7)$$

This is the appropriate background to proceed this paper and the line of thought we want to introduce. In this paper we are presenting another approach of the distance measure, we shall refer to it with capital D, The measures we are referred are more useful, we believe, in Geographical Analysis or Spatial Statistics, approaching the collected data, of the same framework, in different regions, say. We apply this distance-method, introduced in section 2 and the adopted information distance measures are applied in a real life data set, collected from municipal units in the Athens Metropolitan Area, see section 3.

2. The Information theory point of view of Distance

In principle, in Statistics and therefore in Information Theory, the term "distance" does not fulfill the "triangular inequality", as in Linear Algebra or in Analysis. Moreover, see section 3, the "symmetry" it is not also available in most cases. But we keep on using the term "distance" for these information distance measures! Moreover, there are "information distance" measures, which we discuss and we believe that are useful to be introduced in Geographical Data Analysis, as well in other Data Analysis fields.

The information distance measures (idm) succeed to qualify the information provided by the random variable (r.v) X, coming from the distribution P, say, in relation with the random variable Y, coming from the given distribution Q. In such case, the most well-known information measure, is that of Kullback-Leibler (1951), widely known as K-L divergence, also known as relative entropy, Kullback (1959). The K-L divergence it is neither symmetric nor triangular, but offers a measure to quantify how one probability distribution is different from a second, considering as reference probability distribution. This is why it is known as "divergence" and not as "distance" measure, and is not adopted in this paper, although it offers to us food for thought, as it is widely applied in Science. It varies from zero (when P and Q perfectly much) to infinity (the K-L divergence between Cauchy and Normal distribution, the perfect disagreement) and is considering as an information-based measure of disparity among probability distributions, defined for the continuous or discrete case as:

$$D_{KL}(X,Y) \equiv D_{KL}(P,Q) := \int p(x) log \frac{p(x)}{q(x)} dx \quad \text{Continuous case} \qquad (2.1)$$

$$D_{KL}(X,Y) \equiv D_{KL}(P,Q) := \sum_{i=1}^{n} p_i \, log \frac{p_i}{q_i} \qquad \text{Discrete case}$$

Usually p(x) is considered the "true" distribution and "q(x)" the approximate, so it is eventually the expectation of the log difference between the probability of data in the original distribution with the approximating distribution. Moreover K-L information needs no many regularity conditions, as Fisher's information needs, but still is related to it, Schervish (1995).

Example 1. Let the r.v X coming from the Normal distribution with mean (μ) zero and variance (σ^2) one, i.e X~N(0,1), while the r.v Y is coming from the Laplace, with same mean and variance, i.e Y~L(0,1). According to the introduced notation:

$$P = N(0,1), Q = L(0,1)$$

Recall that the pdf of the standard Normal and Laplace, let p(x) and q(x) respectively are:

$$p(x) = \frac{1}{\sqrt{2\pi}} exp\left[-\frac{1}{2}x^2\right], \qquad q(x) = \frac{1}{2} exp[-|x|], \qquad x \in \mathbb{R}$$

Then one can evaluate the Kullback and Leibler information distance, D_{KL}, information measure between P and Q, as well as Q and P, Schervish (1995):

$$D_{KL}(P,Q) = 0.07209 \ , \ D_{KL}(Q,P) = 0.22579$$

It is clear that there is no symmetry, $D_{K,L}(P,Q) \neq D_{K,L}(Q,P)$.
The K-L diverge usually works with the Normal distribution, while there is a broader class of "normal" distributions :The Generalized Normal Distribution (GND) was defined by Kitsos and Tavoularis (2009), through the Logarithm Sobolev Inequalities (LSI) for the r.v $X \sim N_\gamma^p(\mu, \Sigma)$ with three parameters: for position $\mu \in \mathbb{R}^p$, variance $\Sigma \in \mathbb{R}^{p\times p}$ and the shape parameter γ,with pdf :

$$f(x) = f(x;\mu,\Sigma,\gamma) := c_x exp\left\{-\frac{\gamma-1}{\gamma} : Q(x)_x^{\overline{\frac{\gamma}{2(\gamma-1)}}}\right\}, \gamma \in R - (0\ ,1) \qquad (2.2)$$

Where $C_x = \frac{\left(\frac{p}{2}\right)+1}{\pi^{\frac{p}{2}}\left(p\frac{\gamma-1}{\gamma}+1\right)\sqrt{|\Sigma|}} \left(\frac{\gamma-1}{\gamma}\right) p\,\frac{\gamma-1}{\gamma}$, $|\Sigma| =$ determinant of Σ

and

$Q(x) = (x-\mu)\Sigma^{-1}(x-\mu)^T, x \in \mathbb{R}^p, the\ parameter\ vector\ is\ \theta = (\mu, \Sigma, \gamma).$

In GND the introduced new parameter, the shape parameter γ, offers the possibility for creating a family of "normal distributions with fat tails". Moreover, special values of γ lead to a number of well-known distributions:

- When γ=2 the p- variable Normal is obtained, $N_2^p(\mu, \Sigma)$
- When γ=0 the Dirac distribution is obtained
- When γ=1 we obtain the Uniform distribution
- When $\gamma \to \pm\infty$ the Laplace distribution is obtained.

The appropriate choice of γ provides an approach to describe data not so close to typical Normal distribution, but still "normal with more probability in tails", otherwise to "fat tail distributions".

This can be proved useful to a number of applications mainly to Economic data sets. For the Generalized Normal distributions $KT_\gamma^p(\mu_1, \sigma_1^2 \mathrm{I}_p)$ and $KT_\gamma^p(\mu_0, \sigma_0^2 \mathrm{I}_p)$ the (information) distance KLI_γ^p of Kullback-Leibler, for given γ is equal to, Kitsos and Toulias (2010):

$$\mathrm{KLI}_\gamma^p = \frac{C(p,\gamma)}{\sigma_1^p}\left[\frac{p}{2}\left(\ln\frac{\sigma_0^2}{\sigma_1^2}\right)\int_{\mathbb{R}^p} e^{-q_1(x)}dx - \int_{\mathbb{R}^p} e^{-q_1(x)}q_1(x)dx + \int_{\mathbb{R}^p} e^{-q_1(x)}q_0(x)dx\right], \qquad (2.3)$$

where $q_i(x) = \frac{\gamma-1}{\gamma}\left(\sigma_i^{-1}\|x-\mu_i\|\right)^{\frac{\gamma}{\gamma-1}}, \ x \in \mathbb{R}^p, \ \ i = 0,1$.

Moreover, the evaluated KLI_2^p, between two classical Normal distributions, is a linear expression:

$$\mathrm{KLI}_2^p = \xi p + \frac{|\mu_1 - \mu_0|^2}{2\sigma_0^2}, \qquad \xi = \frac{1}{2}\left[\left(\ln\frac{\sigma_0^2}{\sigma_1^2}\right) - 1 + \frac{\sigma_1^2}{\sigma_0^2}\right] \geq 0 \qquad (2.4)$$

The “universal constant”, as it is known, the quantity:

$$(\frac{1}{\gamma-1})^{1/(\gamma-1)} \text{ or } (\frac{\gamma}{\gamma-1})^{\gamma/(\gamma-1)} \qquad (2.5)$$

and as it appears in Applied Analysis, Takashi and Takashi (2011), is associated with the γ order GND theory. We will work on this in another occasion, where the normality does not hold, while in this paper the normality of the observations is the main assumption.

A general form of the information distance measure (idm) was proposed by Kitsos and Toulias (2017). For given “smooth” functions g, h and f_X, f_Y the pdf pf given rv’s X and Y the idm is defined to be, D_{KT}, as:

$$D_{KT} = D_{KT}(X,Y) = g\left(\int_{\mathbb{R}^p} h(f_X, f_Y)\right) \qquad (2.6)$$

Under (2.6) the following cases can be considered for the function g:

Case 1: $\quad For \quad g(t) = \sqrt{t}$

Then with:

$$h(f_X, f_Y) = \frac{1}{2}\left(\sqrt{f_X} - \sqrt{f_Y}\right)^2 \qquad (2.7)$$

we come across to $D_H(X,Y)$ the Hellinger distance, see also Appendix A, as far probability
measures concerns.

Case 2: $\quad For \quad g(t) = id(t) = t$

With different values of $h(\cdot,\cdot)$ we come across to different distance measures:

2i. With $h(f_X, f_Y) = f_X log\frac{f_X}{f_Y}$ we evaluate the D_{KL}, the K-L distance

2ii. With $h(f_X, f_Y) = f_X\left[\log\frac{f_X}{f_Y}\right]^2$ we evaluate to the exponential distribution, D_e, distance

2iii. With $h(f_X, f_Y) = f_X\left[1 - \frac{f_Y}{f_X}\right]^\alpha, \alpha \geq 1$ we evaluate the Vajda, D_v, distance

2iv. With $h(f_X, f_Y) = \frac{1}{2} f_X \left[1 - \frac{f_Y}{f_X}\right]^2$ we evaluate the Kagan (or X^2), D_{X^2} , distance

2v. With $h(f_X, f_Y) = f_Y \phi\left(\frac{f_Y}{f_X}\right)$, ϕ convex, $\Psi(1) = 0$, we evaluate the Csiszar, D_c , distance

2vi. With $h(f_X, f_Y) = f_X{}^{\alpha} f_Y{}^{1-\alpha}$, $\alpha \in (0,1)$ we evaluate $D_{ch}^{(\alpha)}$, Chernoff''s distance

2vii. With $h(f_X, f_Y) = |f_X - f_Y|$ we obtain the, very popular, total variation distance, D_{TV}.

Case 3: For $g(t) = -\log(t)$

With h(f_X , f_Y) = $(f_X * f_Y)^{1/2}$ we come across the Bhattacharyya's , D_B, distance.

Notice that in most cases the ratio $\left(\frac{f_X}{f_Y}\right)$ is used and this can offer an easy explanation that there is, in principle, no symmetry for the idm. We shall take a special consideration for the cases where symmetry can be present. Moreover, under certain restriction these idm are related between them. One example is that Hellinger's D_H distance, which obeys to the triangle inequality, is related to Bhattacharyya's , D_B, which is symmetric but not under triangular inequality, as :

$$D_B = -\log(1 - D_H^2) \qquad (2.8)$$

The cases 2i and 2ii are related as well as the cases 2iii and 2iv. The 2v case offers a generalization for the ratio $\frac{f_Y}{f_X}$, with first term the function f_Y. The cases 2vi and 2vii are based on completely different line of thought. In this paper we adopt Hellinger, D_H , distance which obeys the triangular relation, and offers possibilities, discussed here, for symmetry and we work with it in section 3. The K-L divergence needs a special consideration, see section 2, as it is useful is a number of applications, as all the idm, Geoghegan (2008), that can be extended.

In principle the following Proposition 1 and the Corollary, for the multivariate and univariate case, Kitsos and Toulias (2017), among others, offer the way that these measures can be evaluated.

Proposition 1. Let us consider the p-dimensional normal distribution $N_p(\mu, \Sigma), \mu \in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p \times p}$. Then for the given rv's X and Y with

$$X \sim N_p(\mu_X, \Sigma_X), Y \sim N_p(\mu_Y, \Sigma_Y), \mu_X \neq \mu_Y, \Sigma_X \neq \Sigma_Y \qquad (2.9)$$

The corresponding information distance measures (idm) are :

(1) The square of idm of Hellinger's D_H is :

$$D_H^2(X,Y) = 1 - \frac{|\Sigma_X|^{1/4}|\Sigma_Y|^{1/4}}{|\Sigma|} exp[W] \qquad (2.10)$$

with $\Sigma = \frac{1}{2}(\Sigma_X + \Sigma_Y)$ and $W = -\frac{1}{8}[(\mu_X - \mu_Y)\Sigma^{-1}(\mu_X - \mu_Y)^T]$

(2) The α-idm of Chernoff is :

$$D_{ch}^{(\alpha)}(X,Y) = \frac{1}{2} log \frac{|\alpha\Sigma_X + (1-\alpha)\Sigma_Y|}{|\Sigma_X|^{\alpha}|\Sigma_Y|^{1-\alpha}}, \ \alpha \in (0,1) \qquad (2.11)$$

(3) The Bhattacharyya's idm is :

$$D_B = \frac{1}{2}\log\frac{|\Sigma|}{|\Sigma_X\Sigma_Y|^{1/2}} - W \text{ , with W and } \Sigma \text{ as in (2.10)} \qquad (2.12)$$

Restriction: In this study, and for the problem we discuss, it is essential to impose the restriction that:

$$\frac{|\alpha\Sigma_X + (1-\alpha)\Sigma_Y|}{|\Sigma_X|^{\alpha}|\Sigma_Y|^{1-\alpha}} > 1$$

So that to obtain positive values for the corresponding distance.
Hellinger's distance is related to the Bhattacharyya coefficient BC, see also Case 3 above, as $D_H = (1\text{-}BC)^{1/2}$ and as the Bhattacharyya's distance is originally defined to be $D_B = -\ln(BC)$ it is easy to see that (2.8), as above, holds. Moreover from (2.12) the existence of D_B depends on the value of det(Σ), it is necessary to be positive , as well as det($\Sigma_X\Sigma_Y$).
It is clear that Chernoff's idm in (2.11) it is not symmetric. But when α=1/2 it is symmetric, while Hellinger's distance in (2.10) can be considered symmetric, while obeys, in principle, to triangular inequality. Moreover in univariate case is clear, symmetry holds for idm of Hellinger, and also when α = ½ for the Chernoff idm, due to the following Corollary. Indeed:

Corollary 1. For the one dimension rv's $X \sim N(\mu_X, \sigma_X^2), Y \sim N(\mu_Y, \sigma_Y^2)$ holds:

(i) $$D_H^2(X,Y) = 1 - \left(\frac{2\sigma_X\sigma_Y}{\sigma_X^2 + \sigma_Y^2}\right)^{1/2} exp(w_0)$$
(2.10a)

with $w_0 = -\frac{(\mu_X - \mu_Y)^2}{4(\sigma_X^2 + \sigma_Y^2)}$

(ii) $$D_{ch}^{(\alpha)}(X,Y) = \frac{1}{2} log \frac{\alpha\sigma_X^2 + (1-\alpha)\sigma_Y^2}{\left[\sigma_X^{\alpha}\sigma_Y^{(1-\alpha)}\right]^2} \qquad \alpha \in (0,1)$$
(2.11a)

(iii) $D_B = -\log(1 - D_H^2)$ see also (2.8)
(2.12a)

Notice that $(\mu_{X} - \mu_{Y})^2$ is a Euclidean distance of the means of the Normal distributions that the data set follow. Therefore the involved parameter w_o is a weighted distance measure and so does any continuous function, therefore d* = exp(w_o) is also a distance of the means. Hence:

Corollary 2. A distance between the mean values of the Normal univariate variables X and Y under consideration is expressed by $d^* = \exp(w_o)$. Therefore, Hellinger's idm between two Normal distributions is a function of the distance d^* between their mean values, $D_H^2(X,Y) = D_H^2(X,Y;d^*)$.
Moreover, it easy to see that:

Corollary 3. For the ½-Chernoff information distance measure holds:

$$D_{Ch}^{1/2}(X,Y) = \frac{1}{2}\log[\frac{1}{2}\frac{\sigma_X^2 + \sigma_Y^2}{\sigma_X \sigma_Y}] = D_{Ch}^{1/2}(Y,X) \qquad (2.13)$$

Notice that as:

$$D_{Ch}^{1/2}(X,Y) = \frac{1}{2}\log[\frac{1}{2}\frac{\sigma_X^2 + \sigma_Y^2}{\sigma_X \sigma_Y}] = \log[\frac{\sigma_X^2 + \sigma_Y^2}{2\sigma_X \sigma_Y}]^{1/2}$$

The ½ - Chernoff distance is related to Hellinger's square distance, see (2.11) and (2.12). The distance, obeying to symmetry, but not to triangular inequality, Bhattacharyya's , D_B , is related to Hellinger's distance, see (2.8) is also adopting to this paper. Notice that Mahalanobis' distance is a special case of Bhattacharyya's distance, Comaniciu et. al. (2000). That explains why Hellinger's distance plays a dominant role – most of the other distances related to it. The Total Variation distance, D_{TV}, is approximated in this paper, see (3.1), through Hellinger's distance.

To evaluate the symmetric idm Hellinger's or ½- Chernoff's the involved calculations are not so hard, despite the theoretical insight, and their symmetry offers the possibility to be adopted in a Data Analysis different than the usual one in Electronics. We worked out symmetry which eventually provides this improvement, to apply the appropriate symmetric, idm in Geographical Data Analysis, which are strongly referred to real-life data. We adopt this idea in section 3, for data concerning the metropolitan area of Athens, Greece.

3. Data Analysis for Geographical Regions

In this section the introduced information distance measures (idm) are applied to real geographical data. The study region is mainly consisting of the Metropolitan Area of Athens and the Metropolitan Area of Piraeus. It includes 72 municipal units. Data concern houses for sale and they were derived from real estate websites. The sample comprises over 5000 houses. Several characteristics of the houses are available, among which several quantitative variables: price, size, age, floor, price per square meter, distance from the beach and distance from the metro station. We examined all these variables for normality and we selected two of them, "price/m^2" and "distance from metro station", which showed normality in a relatively large number of municipal units. In the present study it is not included those municipalities with less than 33 observations. The idea is to provide evidence that the methods we discussed can be easily applied, despite their widely use in Information Theory in Electronics, with some extra calculation, but this is getting easier and cheaper in our days.

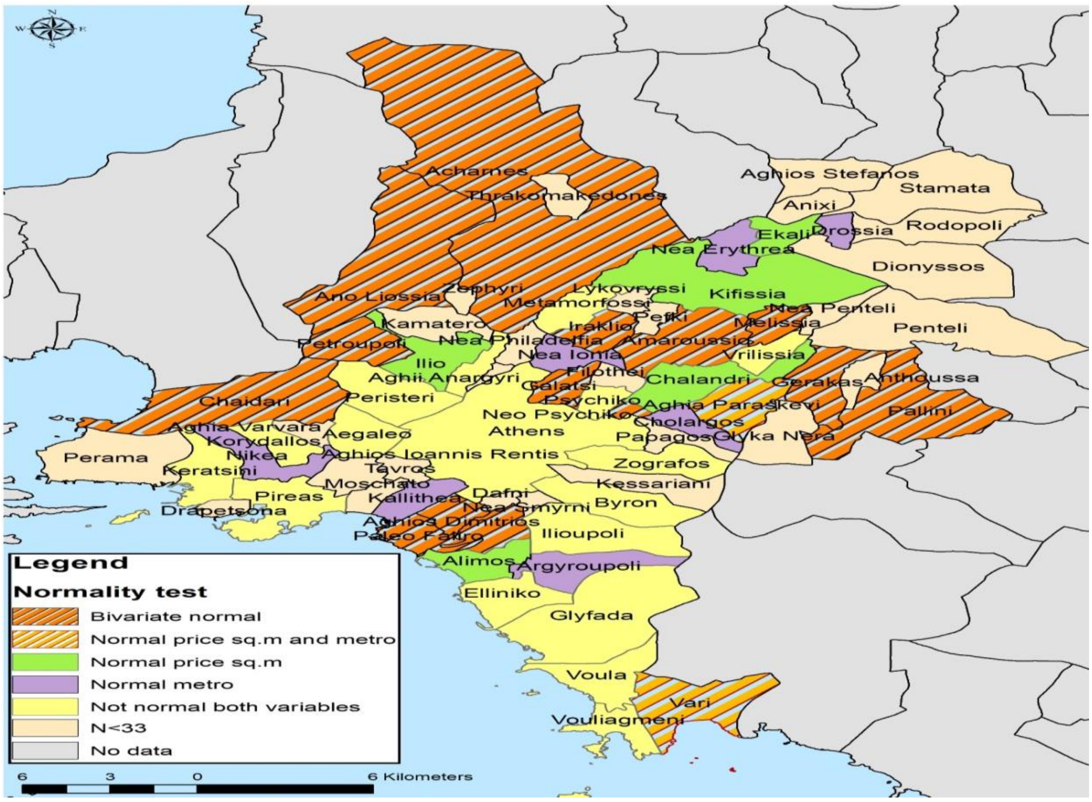
Thus for the 45 municipalities we examined which municipalities appear a normality on the random variables X =«price/m^2 » and Y= «distance from metro station». Normality was considered for the collected data set when both Kolmogorov-Smirnov (K-S) and Sharipo –Wilks (So-W) tests provided normality, under the light of no outliers. Adopting this definition, normality was insured for 15 municipalities for the random variable X, 7 are presented in Table 1a, while for the random variable Y there were 23, 7 from them are presented in Table 1b. Testing Bivariate Normality (BN) we came across cases, with both variables to be Normal, but not BN, namely the municipal units, see in Table 1c, those municipalities with an *.

Eventually was presented three Geographical regions, municipalities, with BN, for both variables X and Y, in Table 1c, with Pearson’s sample correlation r and Royston’s test statistic as well as the corresponding level of significance. In Table 1c the mean and the standard deviation of the random variables X and Y are presented, for the Bivariate Normal distribution.

To test the Bivariate Normality, the Royston’s H test was adopted, Royston (1983), employing the Statgraphics software. The Chi-Squared Q-Q plots to Assess Multivariate normality were also adopted and are presented in Figures 1-3.

The results of the normality tests are presented in map 1. It is worth noticing that the municipal unit with the most central location concerning the bivariate normal distribution is M12. This is the central feature in Euclidean space, as calculated by the ArcGIS software. This feature is associated with the smallest accumulated distance to all other features in the dataset and it is the most centrally located feature. Features can be points, lines or polygons. For polygon features, feature centroids are used in distance computations.

Map 1. Tests of normality



Recall Corollary 1 (i), were the evaluated that Hellinger's distance is symmetrical. That is the distance between the Normal distributions approaching the data, between different municipalities, remains symmetrical. There is certainly an underlying graph, the histogram of the collected data, the empirical distribution, approached from the Normal distribution, see Kitsos and Iliopoulou (2021) for details on these graphs, adopting SPSS26.

Tables 1 (1a and 1b) summarize most of the calculations performed in their columns, for the two variables (X and Y respectively), under the code system for municipalities, see Appendix, with common notation:

(1) Sample size
(2) The mean of the rv «cost/m^2 » in Table 1a , «distance from metro station» in Table 1b
(3) The variance of the rv
(4) The Kolmogorov – Smirnov test for the variable under consideration
(5) The corresponding significant level
(6) The Shapiro-Wilks test for the variable under consideration
(7) The corresponding significant level.

In Tables 2 (2a and 2b) for the two considered variables X and Y, the distance information measures are evaluated, as there are presented in Corollary 1. As the total variation measure, D_{TV}, lies between D^2 and $\sqrt{2}D$, $D_{TV} \in [D^2, \sqrt{2}D]$, we approximate D_{TV} as :

$$D_{TV} \cong [\sqrt{2}D - D^2]/2 \qquad (3.1)$$

Table 1a. Evaluated Statistical parameters for a number of municipalities in Athens, where the variable X "price/m^2" in Euro has been proved Normal

Municipal unit	n (1)	$\bar{X}$ (2)	S_y (3)	K-S (4)	α (5)	So-W (6)	α (7)
M1	115	2103.7	607.9	0.069	0.2	0.985	0.2
M4	112	2305.8	705.7	0.072	0.2	0.983	0.17
M5	126	1958.5	635.2	0.75	0.08	0.979	0.05
M2	104	1098.1	297.1	0.087	0.053	0.985	0.29
M6	86	2337.1	674.2	0.091	0.078	0.974	0.77
M7	71	2158.9	570.4	0.068	0.2	0.975	0.17
M8	145	2632	735.9	0.048	0.2	0.989	0.028

Table 1b. Evaluated Statistical parameters for a number of municipalities in Athens where the variable Y "distance from metro station" in m has been proved Normal.

Municipal unit	n (1)	$\bar{Y}$ (2)	S_y (3)	K-S (4)	α (5)	So-W (6)	α (7)
M1	115	3120.38	766.64	0.042	0.200	0.985	0.251
M5	126	1662.24	609.00	0.046	0.200	0.988	0.331
M2	104	9662.30	1700.02	0.077	0.142	0.978	0.083
M6	86	8704.05	633.45	0.066	0.200	0.973	0.064
M9	87	4654.92	674.79	0.072	0.200	0.982	0.259
M3	79	3472.86	543.31	0.048	0.200	0.991	0.845
M10	75	2452.28	660.81	0.048	0.200	0.987	0.614

Due to (3.1) the calculations through integration, see Appendix A, were avoided. Notice that the total variation it is typical example of measure distance between two "smooth" functions. Here the functions are the probability distributions. Distance measures between random variables may relate to the extent of dependence between them, rather than to their individual values. In Table 1c the appropriate parameters are presented, for those geographical regions, were a Bivariate Normality was under consideration.

Table 1c. Evaluated Statistical parameters for the Bivariate Normal distributions under investigation

Municipal unit	n (1)	$\bar{X}$ (2)	S_x (3)	$\bar{Y}$ (4)	S_y (5)	r (6)	Royston H (7)	p-value H (8)
M1	115	2103.7	607.9	3120.4	766.64	-0,263	2.902	0.234
M2	104	1098.1	297.1	9662.3	1700.0	0.180	4.127	0.127
M3	79	2614.0	545.3	3472.9	543.3	0.053	1.085	0.581
M11*	65	2083.9	567.7	1048.3	500.1	0.155	6.219	0.045
M6*	86	2337.1	674.2	8704.0	633.5	-0.146	6.536	0.038

*Not Bivariate Normal

The calculations on Tables 2a and 2b are based on the exact calculations; the values are truncated, while the total variation distance is an approximation, due to (3.1). Therefore the tedious analytic calculation was avoided, see Appendix A. The results are presented based on the rank of Hellinger's distance. Notice that in $D_{Ch}^{1/2}$ the exponent ½ declares the value of a=1/2 in a-Chernoff's distance and not the square root.

Table 2a. Calculating idm: Hellinger's, total variation, ½ Chernoff, Bhattacharyya's between Normal distributions , for the rv X "price/m^2" in the municipality units of Table 1a and exp(w_o).

No	Municipal units	e^{w_0}	D_H	$\sqrt{2D_H}$	D_{TV}	D_B	$D_{Ch}^{1/2}$
1	M4 - M6	0,9997#	0,0279*	0,0394	0,0193	0,0003*	0,00046
2	M1 - M7	0,9989	0,0459	0,0649	0,0014 *	0,0009	0,00043
3	M1 – M5	0,9932	0,0853	0,1206	0,0566	0,0031	0,00020
4	M5 - M7	0,9863	0,1286	0,1818	0,0826	0,0378	0,00125
5	M6 - M7	0,9899	0,1303	0,1843	0,0840	0,0074	0,00301
6	M1 - M4	0,9883	0,1310	0,1853	0,0840	0,0031	0,00240
7	M4 - M7	0,9935	0,1328	0,1878	0,0846	0,0077	0,00484
8	M1 – M6	0,9836	0,1379	0,1950	0,0846	0,0083	0,00232
9	M6 - M8	0,9784	0,1532	0,2166	0,0315	0,0103	0,00082
10	M4 - M8	0,9747	0,1603	0,2267	0,1005	1.0114#	0.00018*
11	M4 - M5	0,9671	0,1886	0,2667	0,1155	0,0157	0.00119
12	M5 – M6	0,9591	0,2043	0,2890	0,1236	0,0185	0.00038
13	M7 - M8	0,9375	0,2783	0,3935	0,1580	0,0349	0,00697
14	M1 - M8	0,9263	0,2865	0,4052	0,1615	0,0372	0,00393
15	M5 - M8	0,8869	0,3433	0,4855	0,1838	0,0544	0,00233
16	M5 - M2	0,6864	0,6314	0,8930	0,2471	0,2209	0,14991#
17	M1 - M2	0,5757	0,6990	0,9886	0,2471	0,2913	0,05145
18	M2 - M7	0,5065	0,7358	1,0406	0,2495#	0,3386	0,05145
19	M4 - M2	0,5369	0,7389	1,0449	0,2445	0,3428	0,07266
20	M2 – M6	0,4931	0,7592	1,0737	0,2481	0,3730	0,06596
21	M2 - M8	0,3930*	0,8201#	1,1598	0,2471	0,4947	0.07922

*= Minimum value #= Maximum value

As the evaluated idm are related mathematically, mainly as functions of the same quantities, a number of rank correlations, between the evaluated idm, were calculated. The rank correlation of d^* with the other evaluated idm, was not calculated, as d^* it is not an idm, but a generalized Euclidean distance measure of the means of the distributions describing the phenomenon under investigation. Kendall's tau was adopted for the rank correlations. For the estimations of idm in Table 2a, the rank correlation is declares either a close to 0,86 correlation, as for D_H and D_{TV}, due to the proposed approximation of D_{TV} we are using in this paper , while Hellinger's distance, D_H , provides a Kendall's tau to Bhattacharyya's D_B distance , as up to 0.79, due to (2.12a). The ½-Chernoff distance appears almost no correlation with the other idm, it is close 0.4 with D_B and 0.5 with D_H and D_{TV}.

Table 2b. Calculating idm : Hellinger's, total variation, ½ Chernoff, Bhattacharyya's between Normal distributions of the rv Y "distance from metro" in the units of Table 1b and e^{w_0} ,

No	Municipal units	e^{w_0}	D_H	$\sqrt{2}D_H$	D_{TV}	D_B	$D_{Ch}^{1/2}$
1	M2 – M6	0.9326	0.4955	0,7008	0.2276	0.1223	0.0920
2	M1 – M10	0.8966	0.3291	0.4654	0.1785	0.0497	0.0023
3	M1 – M5	0.5743	0.6126	0.8663	0.2166	0.2464	0.0057
4	M5 – M2	0.0073	0.9970	1.4100	0.2079	2.2291	0.0986
5	M3 – M10	0.7006	0.7136	1.0092	0.2500#	0.3090	0.0041
6	M6 - M10	$8*10^{-6}$	0.9999	1.4142	0.2071	4#	0.0001
7	M5 – M3	0.2921	0.8418	1.1905	0.2409	0.5355	0.0014
8	M1 – M6	0.0003	0.9998	1.4140	0.2071	3.5228	0.0039
9	M6 – M9	0.0083	0.9958	1.4083	0.2083	2.0809	0.0004
10	M2 – M9	0.1535	0.9342	1.3212	0.2242	0.8955	0.0819
11	M10 - M5	0.1986	0.9087	1.2851	0.2297	0.7587	0.0007
12	M5 – M6	$1.*10^{-7}$	0.9999#	1.4142	0.2071	4#	0.0006
13	M5 – M9	0.0665	0.9662	1.3665	0.2164	1.1778	0.0114
14	M1 – M3	0.9654	0.2492*	0.3525	0.1452*	0.0278*	0.0126
15	M9 – M10	0.2567	0.8621	1.2192	0.2379	0.5905	0.00005*
16	M3 - M2	0.0494	0.9810	1.3873	0.2125	1.4236	0.1183
17	M1 - M2	0.0461	0.9797	1.3856	0.2128	1.3979	0.19057#
18	M2 – M10	0.9316	0.4841	0.6846	0.2215	0.1159	0.0852
19	M9 – M3	0.6278	0.6160	0.8711	0.2458	0.2071	0.0050
20	M3 – M6	$5*10^{-5}$	0,9999#	1.4141	0.2071	4#	0.0025
21	M1 – M9	0.5687	0.6584	0.9312	0.2488	0.2468	0.0017

*= Minimum value #= Maximum value

The rank of the evaluated idm was applied, so that to order the 21 idm and decide for the minimum idm in Table 2b. Working for the ordinal association in Table 2b, as well as for Table 2a above, Kendall's tau was calculated, to calculate the rank correlation among the evaluated idm. The rank correlation was decided as in terms of mathematics the idm are related. Indeed, from a Mathematical point of view the adopted idm from the Information Theory, under the introduced in (2.6) general form are related but their interpretation in Electronics is different, Cover and Thomas (1991). That is why a particular attention on their rank correlation is needed, to provide evidence what idm are the appropriate for the researcher on a different field, and decide to what idm his decision will be developed. The Kendall's tau for is very small with all idm, actually near to -0.19 with DH and DB, while is -0.01 with DTV. Moreover the obtained rank correlation is significant at the 0.05 level (2-tailed) between DH and DTV ,DB and DTV , while rank correlation is significant at the 0.01 level (2-tailed between DH and DB Based on these results the researcher decides what idm will consider for his study, ie it is not need to use all at the same time for the decision

We emphasize that W in (2.10), as w_o , in (2.10a) , is a distance measure of the distance between the means, and so does, D^*, the extension of the measure d^*, defined above, $D^* = \exp(W)$.

As d^* it is not an idm, it is based on the classical line of thought of distance, as the measure D^* does.

Working from the results of Table 1c the covariance matrices were obtained and following Proposition 1, for the Bivariate Normal distributions, the following idm, were evaluated

Table 2c. Calculating idm : W, Hellinger's, total variation, ½ Chernoff, Bhattacharyya's between the Bivariate Normal distributions of Table 1c.

BN Municipal units	W	Exp(W)	D_H	D_{TV}	D_B	$D_{Ch}^{1/2}$
M1, M2	-8.5924	$1.85*10^{-4}$	0.9999	0.2071	0.1741	0.0218
M1, M3	- 0.1522	0.8587	0.9999	0.2071	8.1111	-*
M2, M3	33.2446	$2.74*10^{14}$	1	0.2071	-*	-*

* The imposed Restriction is not full field

Similar to a normal scores plot which assesses graphically the univariate normality of a given variable, the chi-squared Q-Q plot offers a very useful graphical way to assess whether a two variable distribution is distributed as a Bivariate Normal. In Figures 1-3 the Chi-square plots for the three municipal units with bivariate normal (BN) distribution are presented. In these figures the empirical cumulative distribution function (CDF) of the squared distances are the dot-points, while the solid line is the Chi-square distribution. The points are close to the line and within the Kolmogorov-Smirnov .95 confidence interval limits, which are plotted on either side of the solid line. Therefore, the Chi-square plots confirm, as well as the Royston H test, the conclusion that the data may well have come from a multivariate normal distribution. For the graphs of the normal distributions corresponding to M1, M2, M3 municipal regions see Figure 4.

Figure1. Bivariate Chi-Square test for the municipal unit of the M3 data set.

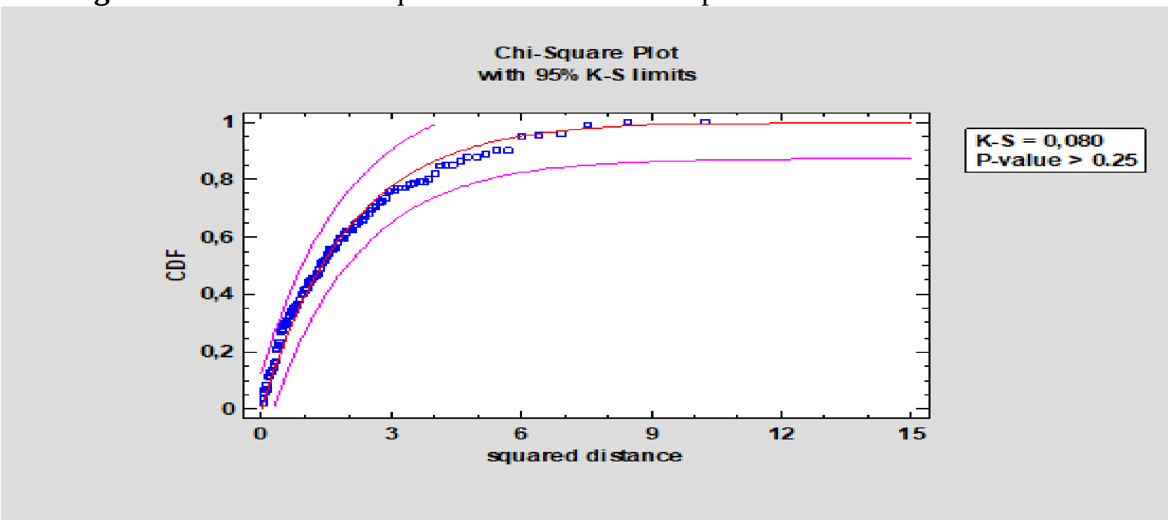


Figure2. Bivariate Chi-Square test for the municipal unit of the M2 data set.

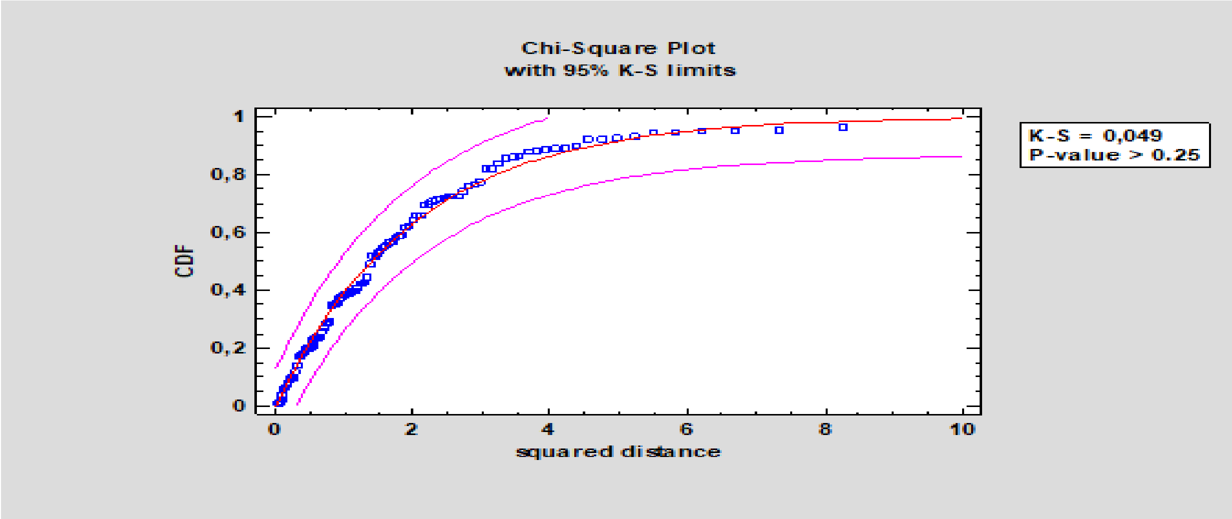


Figure3. Bivariate Chi-Square test for the municipal unit of the M1 data set.

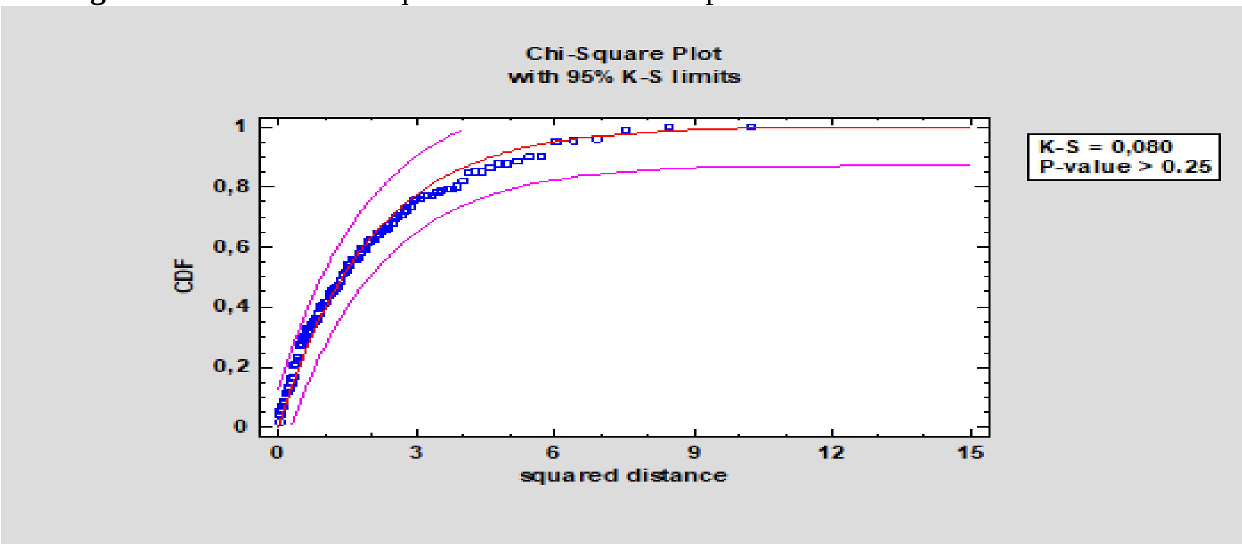
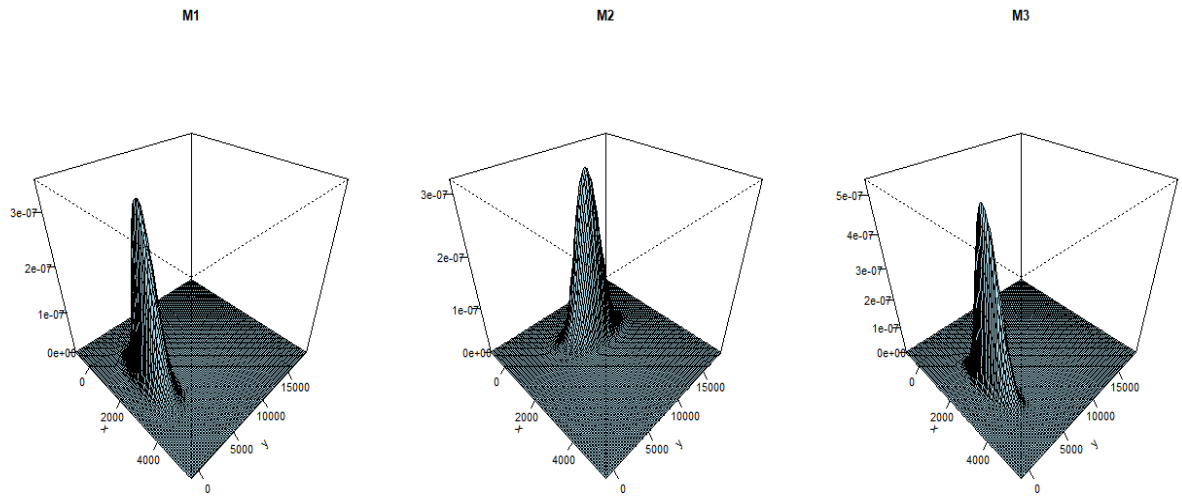


Figure 4 The Bivariate Normal distributions corresponding to M1, M2, M3



To standardize the evaluated information distance measure, we worked out with the proposed D^o:

$$D_j^o = \frac{D_j}{\max D_j} \in [0,1],\ j = H, TV, Ch, B \qquad (3.2)$$

This way of standardization provides an easy way to “group” the real – life data we analyze through the idm. Certainly under different idm, a different way to classify data is created. From Table 2a and for Hellinger’s distance can be easily obtained that $[0.034, 0.16] \cup [0.18, 0.41] \cup [0.76, 0.92] \subseteq [0, 1]$

three groups are created, with a clear information provided for them: G_1 with the first 8 distances (No 1-8), G2 with the next 7 distances (No 9-15), and G_3 with the last 6 distances (no 16-21). The G_1 group provides evidence that the distance of the Normal distributions describing the rv «price/m^2» between these municipalities/regions is, and therefore the prices/m ^{2}, are close. So the interval [0.034, 0.16] offers a clear result for the distances lying within it, while this is also the case for the “dual” group G_3 with idm within [0.76, 0.92] – the distances are large, so the prices between these municipalities/regions are not the same. It is not clear for the distances lying in between [0.18, 0.41], for corresponding case of group G_2. These results are in accordance with the social and economic structure of the municipality regions under consideration. Under Chernoff’s ½ distance two groups O_1, O_2 say, can be considered, as it is easily one can see that the [0, 1] interval can be decomposed to $[0.00, 0.02] \cup [0.15, 0.52] \subseteq [0, 1]$. The O_1 group of distances is including the No1-14 evaluated idm and O_2 the last seven. The first O_1 group declares an agreement in prices, while the O_2 a distance, a disagreement for the municipalities/regions within it.

Working under the light of (3.2) for Table 2b, let us consider the idm D_{TV} and D_B to provide food for thought, for the analysis proceeded already for Table 2a. As far as D_{TV} concerns one decision rule could be large idm to be within [0.88, 1] and this creates a group of distances K_1, say, with the No 1,5,7,10,11,18,19,21 corresponding idm in it, small value idm with value 0.58 for No15 and a middle value set of values within [0.71, 086] offering no information, for all the rest evaluated idm. This example makes clear there is not a unique rule helpful for the decision, see also Discussion. As far as D_B concerns that large values for idm lie within [0.88, 1] and correspond to idm for No 6,8,12,20 see Table 2b, the small distance for values within [0.29 , 0.55] corresponding to No 4, 13, 16 17. All other values correspond to not clear decision.

As far as the Bivariate Normal concerns, it is characteristic that there exists stability as far as the Total Variation distance, see Table 1c, due to the fact that the evaluated Hellinger’s distance for the three BN distributions describing the underlying existing analysis, for the three pairs of municipalities, are almost equal. Moreover equal to the largest possible value, and this can be interpreted due to the complete social-economical difference structure of M2 with M1 and M3. Another interesting point is that either for the first pair of municipalities (M1, M2) were the weighted distance of the mean value vectors is very small, or for the third pair of municipalities (M2, M3) the weighted distance of the mean value vectors is very large, still Hellinger’s distance remain very close to 1. The fact that there is not a metro station in M2 also influences the evaluated distances between the Bivariate Normal distributions.

4. Discussion

In this paper a number of symmetric information (or probability) distance measures were applied , mainly introduced fir the Information Theory development, Kitsos and Toulias (2010, 2017), in Geographical Data Analysis. Under this line of thought, the application was considered on the Metropolitan Area of Athens, as the reference region, which consists on a number of municipalities, the sub regions. It could be a country, as a reference region, divided in its provinces or the EU divided in the countries consist it. The sense of distance adopted concerns the "distance of the distributions" describing the phenomenon under investigation between the considered two (sub) regions. The method was applied not only for the one variable Normal distributions, but for the Bivariate Normal distributions, and it can be extended for more variables, for the particular symmetric measures, we worked with them. The information distance measures, although not distance measures, have been widely applied not only in Information Theory but also in various areas related to it, Comaniciu et al. (2000), Berger et al. (2021) among others. In all these application there is a reference to one particular idm, and a theoretical line of thought. This papers attempts a real-life applications, providing a strong theoretical framework. A number of attempts to bring more Statistical methods to Geographical analysis have been recorded, Lu et. al. (2014), Comber et. al. (2020), Wilson and Martinez (1997), Iliopoulou and Kitsos (2019, 2021). The introduction of these methods, under the framework developed above, offer qualifications to examine how close are two distributions. From the above analysis a Decision rule, can be proposed, for a uniform application offering the possibility to compare results:

For D_j^o within (0.7, 1] the idm is large, For D_j^o within [0, 0.3) the idm is small. No decision for values within [0.3 , 0.7].

To work with such a rule, emphasis is given that the calculations are not a problem nowadays, and the spirit of the discussed analysis remains within an adjusted theoretical background to quantify information under the spirit and line of thought of Applied Statistics.

AKNOWLEDGMENTS

We would like to thank Dr K-S. Nisiotis, Univrsity of West Attika, for his help with R language.

5. References

Abler, R., Adams, J.S., & Gould, P. (1971). *Spatial organization. The geographer's view of the world*. Englewood Cliffs, NJ: Prentice Hall.

Berger, K., Caicedo, J.P.R.,Martino, L., Wocher, M., Hank, T., Verrelst, J. (2021). A Survey of Active Learning for Quantifying Vegetation Traits from Terrestrial Earth Observation Data. *Remote Sens.,*13, 287

Bhatttcharryaya, A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of Calcutta Math. Soc.,35, 99-110.*

Blyth C. R. (1970). On the inference and decision models of Statistics. *Ann. Math Statistic, 41, 1034-1048.*

Comaniciu, D., Ramesh, V, Meer, P. (2000). Real-Time Tracking of Non-Rigid Objects using Mean Shift. *IEEE CVPR, 1-8.*

Comber A., Chi K., Huy Q. M., Nguyen Q., Lu B., Phe H. H. & Harris P. (2020) Distance metric choice can both reduce and induce collinearity in geographically weighted regression, *Environment and Planning B Urban Analytics and City Science* 47:3, 489-507.

Cover, T. M., Thomas, J. A. (1991). *Elements of Information Theory*. Wiley, New York.

Darling D. A. (1957). The Kolmogorov – Smirnov, Gramer – Von Miss Tests. *Ann Math Statistic, 28 283-838*

Fisher, R. A. (1922). On the Mathematical Foundation of Theoretical Statistics. *Phil. Trans. Roy. Soc., London, series A, Vol 22, 165-177.*

Geoghegan, D. B.(2008). The Histographic Conceptualization of Information: A Critical Survey. *IEEE Ann. Of History of Computing*, 30(1), 66-81.

Lu B., Charlton M., Harris P. & A. Fotheringham S. (2014) Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data, *International Journal of Geographical Information Science* 28:4, 660-681.

Iliopoulou, P., (2015). *Spatial Analysis* [e-book]. Hellenic Academic Libraries Link. Available Online at: http://hdl.handle.net/11419/2059 (in Greek).

Iliopoulou, P., Kitsos, C. (2019). Statistical applications in Geography: Spatial Analysis. In Korres G., Kourliouros E. & Kokkinou A. (eds.). *Contemporary essays in Social Sciences & Geography: Theory and Policies*, pp. 212-224.

Kitsos, C. Iliopoulou, P., (2021). Distance measures in Spatial Statistics, In Kalabokidis K., Korres G., Soulakellis N. & Feidas. H. (eds.). *Social Sciences & Geography: Theory, Methods and Spatial Analysis techniques* (in Greek), pp 96-108.

Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis*. Upper Saddle River, New Jersey: Pearson Education.

Kitsos, C. P., Toulias, L. T. (2010). New Information Measures for the Generalized Normal Distribution. *Information, 1(1) 13-27.*

Kitsos C.P.; Tavoularis N.K.(2009). New Entropy Type Information Measures. In : *Proceedings of*
*Information Technology Interface*s; Luzar-Stiffer V., Jarec I., Bekic Z.(Eds) pp. 255-259.

Kitsos, C. P., Toulias, T.L (2017): “Hellinger distance between generalized normal distributions”, *British Journal of Mathematics and Computer Science*, vol. 21, no. 2 (2017), pp. 1-16.

Kitsos, C. P., Tavoularis, K. N. (2009). Logarithmic Sobolev Inequalities for Information Measures. *IEEE Transactions on Information Theory*, Vol 55, 6, June 2009, 2554-2561.

Kitsos, P. C, Toulias, L. T. (2017). Hellinger Distance Between Generalized Normal Distributions. *British Journal of Mathematics & Computer Science* 21(2):1-16

Kitsos, C.P., Toulias, T.L. (2019) Inequalities in Statistics and Information Measures. In : D. Andrica, T. M. Rassias (eds.), *Differential and Integral Inequalities*, Springer Nature Switzerland AG 2019, Springer Optimization and Its Applications , 151, pp. 481 – 508.

Kitsos, C. P., Toulias, L. T., Papageorgiou, E. (2014). Statistical Invariance for the Logit and Probit Models. *Biometrie und Medizinisce Informatik Greifswalder Seminarberichte, Heft 23 :* Statistical and Biometrical Challenges, Theory and Applications, by T. Oliveira, K-E. Biebler, A. Oliveira, B. Jager (Eds), pg 203-216.

Kitsos, C. P., Sotiropoulos, M. (2009). Distance Methods for Bioassays. *Biometrie und Medizinische Informatik Greifswalder Seminarberichte*, Heft 15, pg 55-74

Kolmogorov A (1933). Sulla determinazione empirica di una legge di distribuzione. *G. Ist. Ital. Attuari.* ***4****: 83–91.*

Kullback, S. and Leibler, A. R. (1951). On Information and Sufficiency. *Ann. Math. Statist*., Vol 22 (1), 79-86.

Kullback, S. (1959). *Information Theory and Statistics*. Wiley, New York.

Mahalanobis, C. P. (1936). On the generalized distance in Statistics. *Proc. Nat. Inst. Of Sciences of India*, *Vol 2.*

Norušis, M. J. (2011). *IBM SPSS Statistics 19. Statistical procedures companion*. Upper Saddle River N.J.: Prentice Hall.

Royston, J. P. (1983). Some Techniques for Assessing Multivariate Normality Based on the Shapiro-Wilk W. Appl. Statist. (1983), 32, No. 2, pp. 121-133

Schervish, J. M. (1995). *Theory of Statistics*. Springer.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J. 27*, 379-423, 623-656.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 65, 591-611.

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions . *Annals of Math. Stat.* **19** (2): 279–281.

Toulias, L.T., Kitsos, C. P. (2012). Kullback-Leibler divergence of the γ-ordered Normal over t- distribution. *British Journal of Mathematics and Computer Science*, 2(4), 198-212.

Royston J. P. (1983). Some Techniques for Assessing Multivariate Normality Based on the Shapiro- Wilk W, *Journal of the Royal Statistical Society*. *Series C* (Applied Statistics), Vol. 32, No. 2, pp. 121-133.

Shapiro, S. S., and Wilk, M. B. (1965), An analysis of variance test for normality (complete samples). *Biometrika*, 52, 591–611.

Takashi, S., Takash, S. (2011). *Applied Analysis: Mathematical Methods in Natural Science*. 2d Edition. Imperial College Press.

Toulias, T. L, Vassiliadis, V., Kitsos, C. P. (2014). MLE for the γ-order Generalized Normal Distribution. Discussiones Mathematicae Probability and Statistics 34 , pp 143-158.

Wilson, R., Martinez, T. (1997). Improved Heterogeneous Distance Functions. J. *of Artificial Intelligence Research*, 6, pp 1-37.

Wodny, M. Jager, B., Biebler, K. E. (2003). Cluster and discriminant analysis with both metric and categorial data. In : *Between data science and applied data analysis*, M. Schader, W. Gaul, M. Vichi (Eds), pp127-130. Springer, Berlin.

Wolfowitz, J (1957). The minimum Distance Method. *Ann. Math. Stat*, 28, 75-87.

APPENDIX Coding Municipal units

Code	**Municipal units**
M1	Amaroussio
M2	Acharnes
M3	Paleo Faliro
M4	Kifissia
M5	Nea Smyrni
M6	Vari
M7	Chalandri
M8	Alimos
M9	Nea Ionia
M10	Kallithea
M11	Ag. Paraskevi
M12	Galatsi

A Decision Support Tool for more Effective Legal and Financial Management of Civil Medical Liability Cases

Abstract:

Mistakes during medical care can occur anywhere in the healthcare system. In Greece, the assessment of the nature and total financial burden of medical errors is difficult and cannot be accurately approached, due to the lack of data from an organized information system. In addition, the medical liability landscape is extremely foggy. The scope of this paper is to perform a Risk Analysis on a sample of 173 cases of compensation awarded by civil courts for the years 2000 to 2019 for medical errors of Obstetricians-Gynecologists. To perform the risk analysis and represent the problem studied, a model structure was created using TreeAge Pro software. The scope, methodology and material of the research are presented, and the results are analyzed and discussed. Our research pointed out that the amount and level of compensation awarded by civil courts for medical errors is worryingly high. In addition, the frequency and amount of mean compensation have dramatically increased in the late years. Our research findings are consistent and agree with the findings of other surveys. Taking the multiple uncertainties and the complexity of the relevant cases into account, medical liability litigation constitutes an ideal ground for a litigation risk analysis tool to effectively manage a dispute and, hence, cope with uncertainties in a reasoned way. Decision tree analysis will have a positive social and financial impact and can be proved beneficial to all the parties involved in medical injury litigation.

Keywords: medical liability; decision trees; litigation risk analysis

Athanassios Vozikis[1], Athanasios Panagiotou LL.M.,[2] and Stefanos Karakolias, PhD[3]

[1] Corresponding address: 80, M. Karaoli & A. Dimitriou St., 18534 Piraeus. Athanassios Vozikis Associate Professor, Laboratory of Health Economics and Management, University of Piraeus, Economics Dept., Piraeus, Greece. Email: avozik@unipi.gr

[2] Corresponding address: 80, M. Karaoli & A. Dimitriou St., 18534 Piraeus. Athanasios Panagiotou LL.M., Researcher University of Piraeus. Email: panagiotou.law@gmail.com

[3] Corresponding address: 80, M. Karaoli & A. Dimitriou St., 18534 Piraeus. Stefanos Karakolias, PhD, Researcher University of Piraeus. Email: s.karakolias@gmail.com

1. Introduction

Mistakes during medical care can occur anywhere in the healthcare system (European Commission, 2006). In Greece, the assessment of the nature and total financial burden of medical errors is difficult and cannot be accurately approached, due to the lack of data from an organized information system (Riga et al., 2014). According to the European Commission Eurobarometer (2014), 78% of the Greek respondents think it is likely patients could be harmed by hospital care and 71% of the respondents think it is likely patients could be harmed by non-hospital healthcare. Moreover, 20% of the Greek respondents have -personally or through a member of their family-experienced an adverse event while receiving healthcare (European Commission, 2014). The Greek redress system is a traditional tort system. Even though according to the Greek Civil Law a claim for medical negligence can be based on either contract or tort law (or –cumulatively- on both), tort is the prevailing legal basis. The medical liability landscape in Greece is extremely foggy due to the doctrinal uncertainties which lie at the heart of the current medical liability system (Panagiotou, 2016), the inherent uncertainty/inexactness of medical science and the human body's complexity, the need to attribute liability and specify vague legal concepts in particular cases and the ineffectiveness of litigation. Thus, there is a need to develop decision support tools for a more effective legal and financial management of civil medical liability cases.

2. Scope, Methodology and Material of the Research

The scope of this paper is to perform a Risk (economic) Analysis on a sample of 173 cases of compensation awarded by civil courts for the years 2000 to 2019 for medical errors of Obstetricians-Gynecologists in Greece. We focused on the specialty of Obstetricians-Gynecologists, as it is considered one of the most injurious (either in terms of frequency or of the mean compensation amount awarded) (The Risk Authority Stanford, 2015) (Pinnacle Actuarial Resources, Inc., 2004) (Irish Health analysis, 2016) (Vozikis, Riga, Pollalis., 2016) (Riga, Vozikis, Pollalis, 2014). With respect to every case, we recorded various characteristics (i.e. year of publication of the case, the legal status of the health care organization, the type of medical error, the severity of the adverse event and the amount awarded).

To perform our Risk Analysis and represent the problem studied we create a model structure using TreeAge Pro software. The model structure will include decision points and all events that could occur and have an impact on the outcome. Different node types within the structure reflect whether branches are alterative options or possible events.

3. Analysis- Results

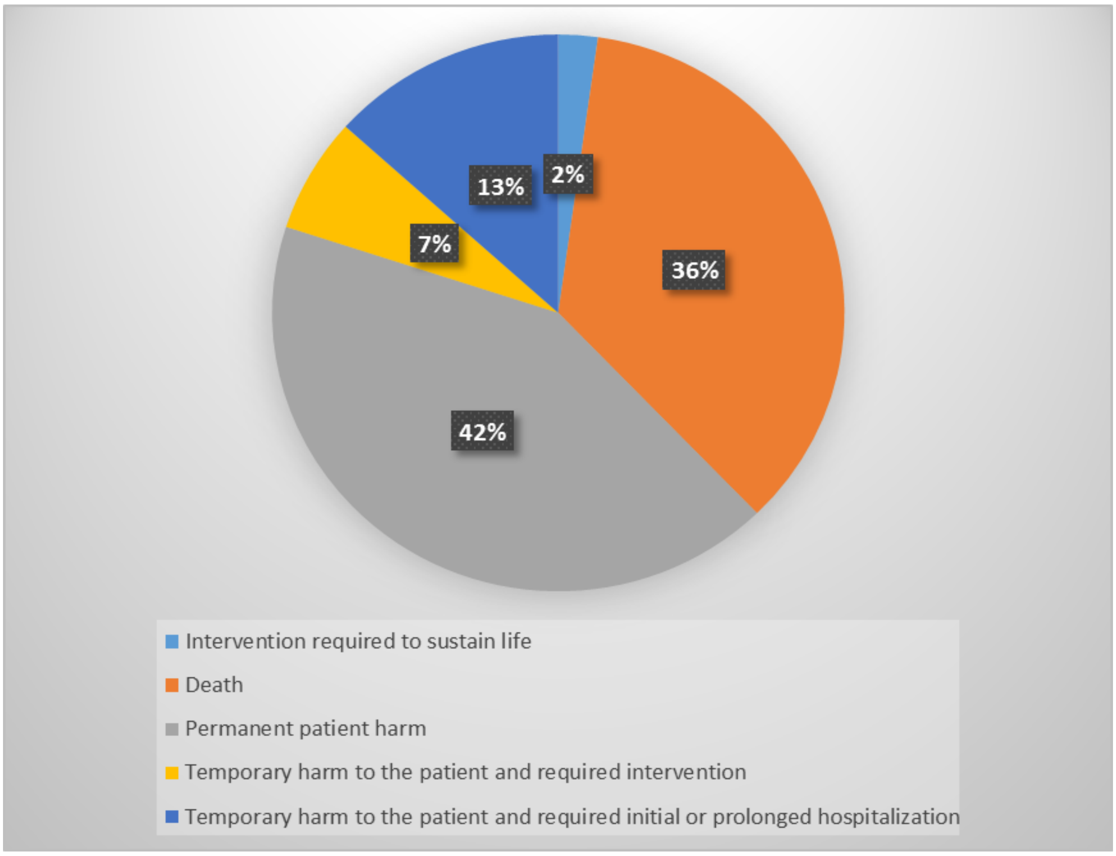
In the model we construct, there is a single decision, whether to litigate or accept a settlement offer. In addition, there are points where the outcome is unknown (win/lose case and damage amounts), represented by various severity types:

Category E: Temporary harm to the patient and required intervention Category F: Temporary harm to the patient and required initial or prolonged hospitalization Category G: Permanent patient harm Category H: Intervention required to sustain life Category I: Patient death

Source: Vozikis A (2009) Information management of medical errors in Greece: The MERIS Proposal. Int J Inf Manage 29: 15–26.

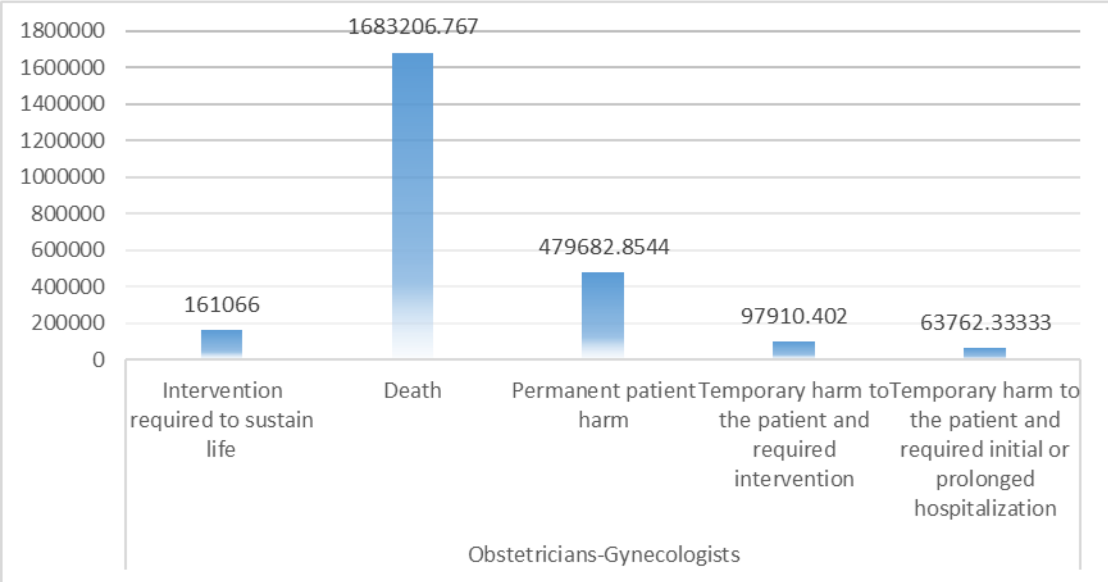
Our data analysis showed that the most frequent adverse events by severity type fall under *Category G: Permanent patient harm* (42% of the total cases) followed by *Category I: Patient death* (36% of the total cases) (See Graph 1).

Graph 1 Frequency of medical errors by severity



One of the unexpected findings of our analysis is that the mean compensation awarded for the *Category H: Intervention required to sustain life* is much lower than the compensation awarded for adverse events with minor severity (*Category G: Permanent patient harm*) (See Graph 2)

Graph 2 Average amount of awarded compensation by severity

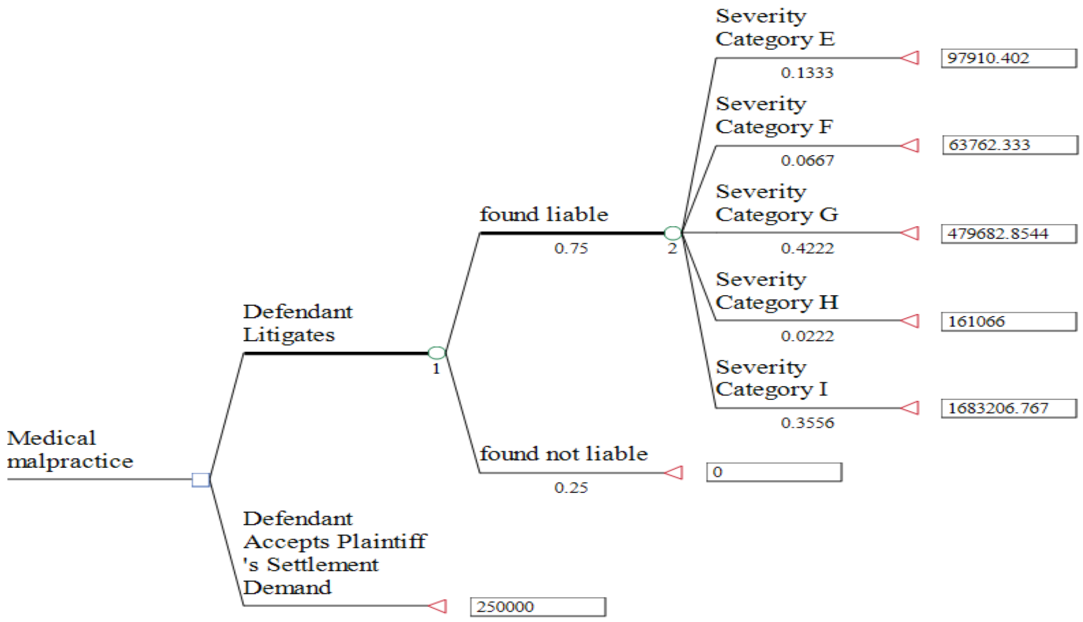


The scenario constructed based on the research data, is presented below in Graph 3. Our scenario is a combination of branches read from left to right, with possible litigation scenarios. Probabilities represent the quantitative best guesses of the relative likelihood of the possible outcomes at each branch, based on the findings of the analysis. They are shown under their respective branches. Probabilities at a chance node always sum to 1.00 (100%). This is logical since the branches must be mutually exclusive and collectively exhaustive. There are no probabilities under the branches following a decision node because they represent the strategy that is chosen and followed. Finally, there are terminal nodes which reflect the overall value of each possible expected outcome, represented by triangles.

3.1. Interpreting the Decision Tree

The decision node shows that the Defendant is debating whether to litigate or pay a 250,000 € settlement demand. If he/she rejects the demand and litigates, he/she faces possible consequences ranging from 0 € (a defense verdict) to 1.683€ million. Reading from the left, we see that the ultimate issue is whether the Court will conclude that the Defendant's malpractice caused the Plaintiff's injury. If the Defendant succeeds in this issue, he/she will be found "*not liable*". If the Plaintiff prevails, he/she will be able to recover some damages.

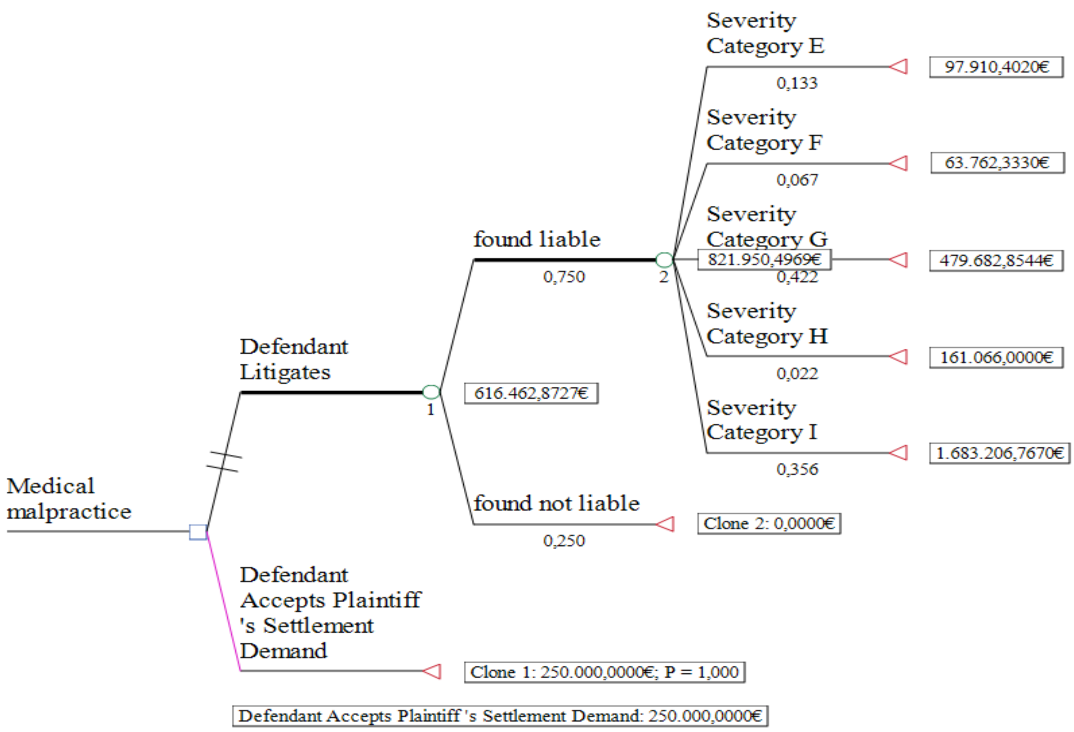
Graph 3 The Decision Tree



3.2. Understanding the results

Once we document that the decision tree is a reasonable representation of the major components of litigation, and that the probabilities and verdict ranges best reflect the related evidence and arguments, then it is time to calculate and interpret the results (See Graph 4)

Graph 4 The Decision Tree with Results



There are two principal ways of calculating the values and evaluating the Decision Tree. The first is to determine the COMPOUND PROBABILITY of each scenario and then to plot the various litigation awards and their respective probabilities in a graph. To determine the compound probability of our scenario we multiply together the probabilities that lie under the branches comprising that scenario. The logic behind multiplying probabilities together is that this ensures and is the only way of ensuring that each issue is given just the right amount of weight and that the conclusion is the most suitable for carefully arriving at conclusions/decisions regarding each of the many underlying issues.

The path probabilities are shown in the Table 1 below:

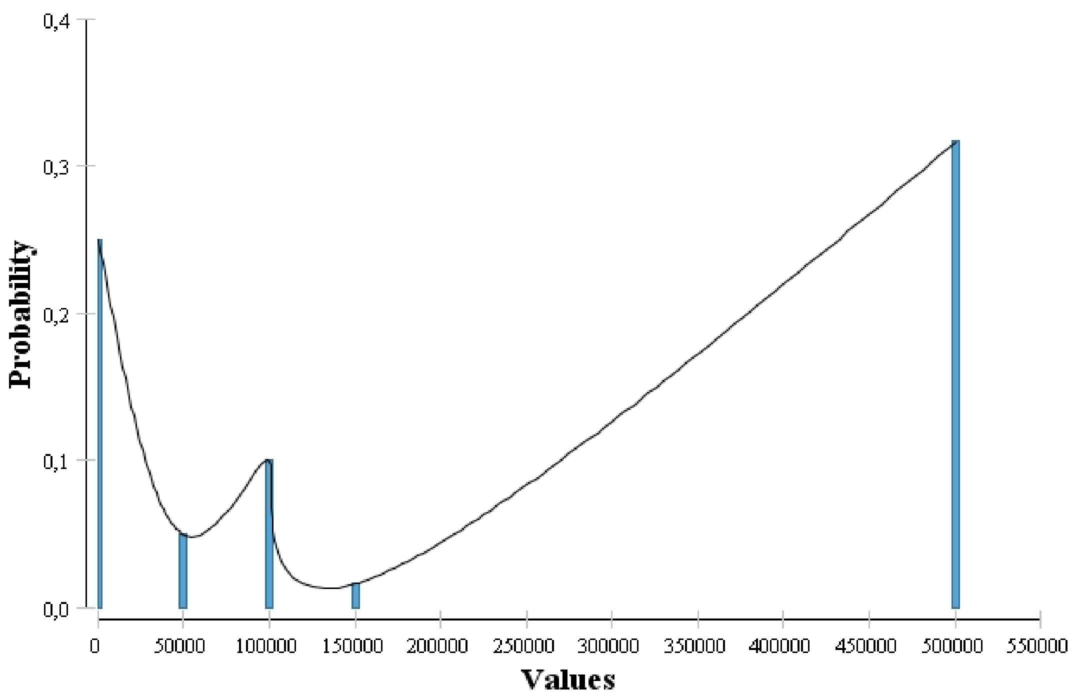
Table 1 The Decision's Tree Path Probabilities

value	prob	label
0.0	0.25	not liable
50000.000000000044	0.050025	Severity Category F
100000.00000000009	0.099975	Severity Category E
150000.00000000012	0.01665	Severity Category H
500000.0000000005	0.31665	Severity Category G

The second principal way of evaluating a tree is to calculate its EXPECTED VALUE. This is a probability-weighted average value (or mean value). It results by weighting each of the possible outcomes by its probability of occurring. If a party can afford to play the averages, the expected value is a fair settlement. In our analysis the Expected Value in the case the Defendant litigates is 616,463 €. That means that the settlement amount of 250,000 € (which Plaintiff demanded) should be accepted by the Defendant.

However, if a party is unable to bear the risk posed by litigating (usually when its net worth is small in relation to the magnitude of the possible gains or losses), he/she will need to go beyond the expected value and see the full bar chart in order to decide on its litigate-versus-settle strategy. This can be estimated by the Probability Distribution of the awarded amount (see Graph 5).

Graph 5 Probability Distribution of the awarded amounts



4. Results and Proposals

Although Litigation Risk Analysis strictly relies on the professional judgment of attorneys and claims managers, it enhances their intuition, expertise, and traditional thought processes, and offers a sound, proven and reasoned methodology for placing a monetary value on actual and potential legal problems.

This approach has wide application — from "pre-constructed" software models to quickly evaluate small claims, to full risk analyses of the largest and most complex commercial disputes. The result is always a comprehensive understanding of the risks of a lawsuit expressed in the business language of probabilities and potential outcomes. That is exactly the information a researcher needs.

Our research pointed out that the amount and the level of compensation awarded by civil courts for medical errors in Greece is worryingly high. In addition, the frequency and the amount of mean compensation have dramatically increased in the late years (Riga, Vozikis, Pollalis, Souliotis, 2015). The highest mean compensation is awarded to the severity of adverse events *Category I: Patient death*, *Category G: Permanent patient harm* and *Category H: Intervention required to sustain life*. Based on our research findings, we note that the mean compensation for the Obstetrics-Gynecology specialty scaled up to 616,463 €.

Our research findings are consistent and agree with the findings of other surveys, according to the literature (Pollalis, Vozikis, Riga, 2012). In a USA survey the specialties of General Surgery and Obstetrics and Gynecology occurred in the first two positions as responsible for causing harm due to medical malpractice (U.S.D.H.H., 2002). The same findings regarding Greece are presented in other similar research studies (Riga, Vozikis, Pollalis, 2014) (Riga, Vozikis, Pollalis, Souliotis, 2015). In the recent 2016 Medical Malpractice Annual Report, General Surgery and Obstetrics and Gynecology are the physician specialties that had the largest number of claims with paid indemnity, while Pediatrics and Obstetrics and gynecology had the higher Median paid indemnity (Office of the Insurance Commissioner, 2016). This is consistent with the findings of a study which mentions that both the Emergency Department and the specialty of Obstetrics-Gynecology cover 28% of all financial demands or $2 million (Larcher and Dulberger, 2008). Other recent research from Utah shows that 58% of adverse events are surgical, of which 25% refer to foreign objects retained after surgery, especially in surgeries of Obstetrics-Gynecology (Utah Department of Health, 2010).

When deciding and reviewing the probabilities of winning or losing a medical liability case, it is essential to go behind the numbers based on different important considerations. For example: 1) what supports your judgment on both sides of the specific issues like (the existence or non-existence of fault/malpractice, negligence (breach or non-breach of the required standard of care), and causal link between the physician's conduct and the damage caused to the patient etc.) 2) What evidence (unfavorable as well as favorable such as the existence of relevant clinical guidelines and/or clinical protocols, the content of relevant medical records, relevant medical literature, written medical expert opinions) do you have? 3) What witnesses (harmful as well as helpful) are we going to confront? 4) Are the witnesses physicians of the same or different specialty? 5) What experience does each expert witness have? 6) What specific/special expertise (if any) does each expert witness have? 7) Are there any previous decisions (bad as well as good) on the topic? 8) How have the courts handled/resolved similar cases? Based on which arguments and jurisprudence? 9) What general factors (such as a. the physician's specialty, expertise and/or specialized training in the specific medical subfield, b. the severity of the patient's injury/

damage, c. the fact that the patient was the only source of income for the family, d. the family's financial status and/or lack of social insurance, e. his/her children's age, f. the inefficiencies of the health and social care provided by the state) may come into play?

Taking the data as well as the multiple uncertainties and the complexity of the relevant cases into account, medical liability litigation constitutes an ideal ground for a litigation risk analysis structure to effectively manage a dispute and, hence, cope with uncertainties in a reasoned way. The key aim of such a structure described above is to apply a litigation risk estimation system to medical liability cases in Greece as a tool to: a) ameliorate decision- making, b) reduce medical liability litigation uncertainty, c) evaluate litigation alternatives in a reasoned and organized way and, thus d) facilitate settlements and enhance the use of mediation in civil cases and, hence, e) decongest Greek courts f) contribute to the sustainability of the National Healthcare System.

Decision tree analysis will have a positive social and financial impact and can be proved beneficial to all the parties involved in medical injury litigation. Particularly it will benefit a) *patients and/or their families (claimants)*: Good settlements will be facilitated and the disadvantages of malpractice litigation (i.e. it is slow, expensive, and emotionally draining) could be avoided. b) *Physicians (defendants)*: Facilitating good settlements and avoiding the negative consequences of litigation such as the financial ramifications, the repercussions on their professional reputation and the psychological impact on physicians, which in turn lead to the practice of defensive medicine. c) *Hospitals*: they will be able to make effective litigation risk analysis and, thus, make the most advantageous (from a financial perspective) decision. d) *Attorneys:* they can express their opinion on the possible outcome of the relevant cases and their alternative options in quantitative terms away from the inaccuracy and highly speculative nature of subjective –based on experience- judgments and estimations. e) *Mediators*: They may convince the parties reach a commonly accepted solution to their dispute. In Greece, based on the current legal framework (Law 4512/2018), mediation is a mandatory stage of disputes concerning claims for compensation of patients or their relatives against physicians.

Therefore, the use of a litigation risk analysis tool to come to agreement in the context of mediation in medical liability cases presents prospect of wider practical application. f) *Insurance companies:* They can conduct more effective malpractice litigation risk analysis, and this can contribute to their long-term financial sustainability. *g) Health Care and Social Insurance System:* Given that significant financial resources of public hospitals are diverted to the payment of damages for medical liability, decision analysis may contribute to their financial sustainability and stability. Decision trees will assist patients harmed by medical errors (and their families) make informed decisions whether to start a claim or not and many of them may abandon the litigation option and prefer either the settlement or the use of alternative dispute resolution methods (like mediation). Taking into account that the use of mediation in the field of administrative law (and specifically for the resolution of disputes between individuals and the state) may be discussed in the future, decision tree analysis could also be useful for effective decision-making and facilitation of settlement with respect to disputes between patients and public hospitals. The resolution of these disputes through mediation could result in more efficient controlling of administrative fees, savings from the courts' operational costs, quicker administration of justice and decongestion of administrative courts. At the same time, it could contribute to combatting bureaucracy and maladministration by enabling public services to improve substantially. g) *The judicial/civil justice system:* The early settlements and the use of ADR methods will lead to the decongestion of courts (since fewer cases will finally reach the courts), will make the administration of justice quicker (the cases which finally reach the courts will be resolved more quickly), and will secure cost savings.

5. Limitations to our research findings

The criticism is often that the “inputs” to a decision tree — the probabilities of various outcomes — are imprecise. The answer is that it’s not a question of probabilities being “precise” or “imprecise” — the idea is for them to be “realistic.” In our case, the assignment of a probability is based on “a frequency of similar events to count.” In other cases, a probability is simply a reflection of someone’s opinion of the likelihood of success in a particular situation.

In addition, other influencing factors whose inclusion would considerably aid (and modify) the assessment of probabilities regarding an existing issue, have been omitted in our model, due to lack of data and because this would result in a very extended and complicated decision tree. If an influencing factor affects only one issue in the tree and the tree is getting unwieldy, a SUBTREE should be created. To illustrate, assume that the Defendant’s medical malpractice insurance coverage or the existence of an expert testimony affects the amount awarded. A subtree containing just those two uncertainties can be drawn and solved to represent the probability that would be used in the main tree.

Finally, our research findings (which are based on a limited number of cases, covering a specific period of time and with limited data on the available cases) should not be generalized with respect to other medical specialties and other countries with different legal systems and insurance coverage for medical malpractice.

Acknowledgements - Funding

This paper is one of the deliverables of the Proposal entitled "Support for researchers with emphasis on young researchers - cycle B" (Code: EDBM103) which is part of the Operational Program "Human Resources Development, Education and Lifelong Learning" which is Co-financed by Greece and the European Union (European Social Fund) and specifically of the Operation entitled “*Decision-making support system for the effective legal and financial management of civil medical malpractice liability cases*” (MIS 5050634)

6. References

European Commission. (2006): Special Eurobarometer 241, *Medical Errors*. Available at: https://data.europa.eu/data/datasets/s403_64_3_241?locale=en , retrieved 29 December 2021.

European Commission. (2014): Special Eurobarometer 411, *Patient Safety and Quality of Care*. Available at: https://data.europa.eu/data/datasets/s1100_80_2_411?locale=en , retrieved 20 December 2021.

Irish Health analysis (2016) Malpractice crisis hits obstetrics. Dublin, Ireland.

Larcher, G. & Dulberger, T. (2008) Trends in Medical Malpractice. Patients versus Profits. CAS Spring Meeting, Quebec, 15-18 June 2008.

Office of the Insurance Commissioner (2016) Medical Malpractice Annual Report. WA, OIC.

Panagiotou (2016). Clinical Guidelines as the fundamental criterion for the establishment of “Fault” in Medical Liability: Advantages, issues and proposals. *Rostrum of Asclepius/Vima tou Asklipiou*, *15*(1).

Pinnacle Actuarial Resources, Inc. (2004) Oregon professional panel for analysis of medical professional liability insurance: a report on factors impacting medical malpractice insurance availability and affordability. Pinnacle Actuarial Resources, Inc., Bloomington.

Pollalis Y, Vozikis A, Riga M (2012) Qualitative patterns of medical errors: research findings from Greece. Rostrum of Asclepius 11: 577-592.

Riga M., Vozikis A., Pollalis Y. (2014) Medical Errors in Greece: An Economic Analysis of Compensations Awarded by Civil Courts (2000-2009), Open Journal of Applied Sciences, 4(4 Special Issue on Applied Economics):168-175

Riga M., Vozikis A., Pollalis Y., Souliotis K. (2015) MERIS (Medical Error Reporting Information System) as an innovative patient safety intervention: A health policy perspective, Health Policy 119(4): 539–548

The Risk Authority Stanford (2015) Medical malpractice trend review. Stanford University Medical Network, CA.

U.S. Department of Health and Human Services (2002) Confronting the new health care crisis: improving health care quality and lowering costs by fixing our medical liability system. Washington, DC.

Utah Department of Health (2010) Utah Sentinel Events Data Report, Identifying Opportunities for Improvement. Health Insight and Utah Hospitals & Health Systems Association. http://health.utah.gov/psi/

Vozikis A (2009) Information management of medical errors in Greece: The MERIS Proposal. Int J Inf Manage 29: 15–26.

Vozikis A, Riga M, Pollalis Y. (2016) Medical Malpractice Risk Factors: An Economic Perspective of Closed Claims Experience, Journal of Health & Medical Economics, 2(3): 12

Students Constructing Their Own "Deep Ecology": An Application of the Richard E. Gross Problem-Solving Model

Abstract:
In problem-based learning students work together in order to find a solution to a "real world" problem. Such an approach is different from traditional teaching practices, mainly lectures, where knowledge is transmitted through the teacher. In the process of finding a solution to the problem, students cultivate skills in collecting, evaluating, and synthesizing resources. Most importantly, problem-based learning helps students become self-directed learners. This paper discusses the classroom use of a problem-based learning model known as the Richard E. Gross Problem-Solving Model. Following presentation of the steps of the model, the paper attempts to show how students could apply the model by constructing their own environmental philosophy, in particular, their own "Deep Ecology".

Keywords: Problem-based learning, Cooperation, Environmental philosophy

Evangelos Manolas[1]

[1] Corresponding author: Prof. Dr. Evangelos Manolas, Professor, Department of Forestry and Management of the Environment and Natural Resources, School of Agricultural and Forestry Sciences, Democritus University of Thrace, Orestiada, Greece. E-mail: emanolas@fmenr.duth,gr

1. Introduction

Traditional instruction, such as lecturing, which involves the delivering of information without the participation of the audience, encourages students to be passive in the classroom. Passivity promotes memorization (Unangst 2021).

Engaging students in problem-based learning means helping them build higher order thinking skills such as collecting, evaluating, and synthesizing information rather than simply recalling it.

Problem-based learning is a pedagogy in which problems are simultaneously both the context and the stimulus for learning. In problem-based learning classes, students work in teams to complete a specific task, to solve "real world" problems. Such problems enable students to make connections between past experiences and the problem they are trying to solve. The instructor becomes a facilitator, helping students become independent learners and responsible for their own learning (Cotton 2014, Nair et al. 2020)).

This paper discusses how a problem-based learning model, in particular the one designed by Richard E. Gross, Stanford University (Gross 1958) can be used in the classroom. Following a presentation of the steps to be followed, the paper attempts to apply Gross' model to the issue of how students can construct their own philosophical scheme with regard to protecting the environment, in particular their own "Deep Ecology".

2. The Richard E. Gross Problem-Solving Model

Gross' Problem-Solving Model consists of six steps (Gross, 1958; Chilcoat & Ligon, 2004). It should be noted that the steps are connected and that the lesson usually flows naturally from one step into another.

Steps in the Richard E. Gross Problem-Solving Model

Step 1. Defines the problem, taking account of societal values
Step 2. Lists the various feasible courses of action
Step 3. Collects and interprets pertinent data
Step 4. Reaches a tentative decision based on the data
Step 5. Acts in accordance with the decision
Step 6. Evaluates the results and modifies future action accordingly

Step 1: Define the problem, taking account of societal values. In deciding the problem to be presented to the class, the teacher needs to determine whether the problem will be suitable. An important question to ask is whether the problem is of concern to the students and if it is important enough for the class to spend time studying it. After the presentation of the problem to the class, the students are encouraged to give answers to questions such as how solving the problem will be useful to them or how it might affect their everyday life. The answers of the students are written on the board.

Step 2: List the various feasible courses of action. The second step of the model is to discuss possible courses of action that might be useful in solving the problem. Using questions may be helpful in guiding the discussion.

Step 3: Collect and interpret pertinent data. The third step of Gross' model is the collection of data. The teacher and students decide what kinds of information might be needed in order to solve the problem at hand. Necessary information for solving the problem may be obtained through methods such as reading, experimentation, interviews and surveys, role play, or student discussion. After the collection of the information is completed, the students must decide what kind of information is most useful for helping them select the best course of action.

Step 4: Reach a tentative decision based on the data. In the fourth step the class reaches a tentative decision based on the information. Here, the students answer questions such as if they are biased in their decisions and what the implications of various courses of action might be. They also decide which actions are most feasible to follow and which are less important. At this point, the class is left with only a few courses of action. They must now understand the reasons for choosing among the options left. The questions which need to be answered now is if each option will provide a good solution or if it will complicate the situation. The class then arrives at a group decision which may involve a compromise.

Step 5: Act in accordance with the decision. The fifth step is the action part of the model. This step might involve using community surveys or interviews, doing a play, writing letters to political figures, or designing a video.

Step 6: Evaluate the results and modify future action accordingly. During this final step of the model the students evaluate what they have accomplished and learned and how they might use what they learned in their life.

3. Using the Richard E. Gross Problem-Solving Model in the Classroom

The following is an example of how the Gross Problem-Solving Model can be used in the classroom.

Step 1: Define the problem, taking account of societal values. The problem chosen and written on the board is the following: Environmental philosophies such as Deep Ecology which attempt to fundamentally change human attitudes and behavior towards the environment have proven much more influential than other philosophies. Drawing upon the principles and implications of Deep Ecology, construct your own "Deep Ecology". You do not need to develop a complete philosophy, but you should select ideas that engage you intensely (Manolas 2006).

Following announcement of the problem the following question is put to the students: Why should we engage in such an activity? The discussion of such a question should establish the rationale for such an activity and settle the audience. After a few answers the class moves to the next stage.

Step 2: Lists the various feasible courses of action. The eight principles of Deep Ecology may be considered such courses of action. These are:

1) The well-being and flourishing of human and nonhuman Life on Earth have value in themselves (synonyms: intrinsic value, inherent value). These values are independent of the usefulness of the nonhuman world for human purposes.
2) Richness and diversity of life forms contribute to the realization of these values and are also values in themselves.
3) Humans have no right to reduce this richness and diversity except to satisfy vital needs.

4) The flourishing of human life and cultures is compatible with a substantial decrease of the human population. The flourishing of nonhuman life requires such a decrease.
5) Present human interference with the nonhuman world is excessive, and the situation is rapidly worsening.
6) Policies must therefore be changed. These policies affect basic economic, technological, and ideological structures. The resulting state of affairs will be deeply different from the present.
7) The ideological change is mainly that of appreciating life quality (dwelling in situations of inherent value) rather than adhering to an increasingly higher standard of living. There will be a profound awareness of the difference between big and great.
8) Those who subscribe to the foregoing points have an obligation directly or indirectly to try to implement the necessary changes (Naess 1989).

Step 3: Collect and interpret pertinent data. In this step, students determine what additional information they need in order to understand the deep ecology principles better. The teacher places the students in small groups with each group covering each of the eight principles described in step 2. When the formation of the eight groups is completed, the teacher assigns each group various activities with the aim of gathering additional information. The activities that students in each group could do include interviews with other teachers; specialists; and articles they can read about their assigned task. An effective way for the students to keep track of the information they collect is to keep a research log or journal in which they write down what they find as well as theirs or other peoples' comments on the information gathered.

Step 4: Reach a tentative decision based on the data. In this step the groups are reformed so that each group has an expert in each of the eight principles described in step 2. Each of these new groups should now produce its vision of what their own "Deep Ecology" would look like in a text of 300-400 words in length. The text produced by these groups should be presented to the class. The class then discusses the different visions, eliminating ideas that might not work. The basic elements of the new "Deep Ecology" as they emerge from the discussion are written on the board.

Step 5: Act in accordance with the decision. In this step the students should determine how to carry out the decision they made in step 4. One idea is for the students to produce, as a class now, a 300-400 words text to be posted in notice boards or published in student magazines, the internet or / and sent to other schools.

Step 6: Evaluate the results and modify future action accordingly. The evaluation stage is the key moment in time when deeper learning can happen (Lenz 2014, Rothman 2018). The type of questions which need be answered during this stage could be:

- In reviewing the new "Deep Ecology" you constructed, what would you drop, modify, or add? Why?
- What were your most positive learning experiences of the process followed in the activity? What made them so positive? List as many things as possible and be specific.

- What were your most negative learning experiences of the process followed in the activity? What made them so negative? List as many things as possible and be specific (Apps 1991 adapted in University of Waterloo n.d.)
- What did you learn from the process? Have your attitudes changed? How will you act in the future on the basis of the new knowledge you acquired? (Wilson et al. n.d.)
- What questions does your own "Deep Ecology" raise?
- If you had to summarize in a few lines the new philosophical scheme what would these be?

4. Conclusion

Problem-based learning may promote deeper learning much more than conventional teaching methods such as lectures. It emphasizes higher order thinking skills, cooperation and self-directed learning. This paper discussed the classroom use of a problem-based instructional model known as the Richard E. Gross Problem-Solving Model. After presenting the steps of the model, the paper attempted to apply the model to the problem of designing a new "Deep Ecology". The example based on the Gross model as analyzed in this paper may be used for a variety of subjects, grade levels or course structures.

5. References

Chilcoat, G.W., Ligon, J.A. (2004). "Issues-Centered Instruction in the Social Studies Classroom: The Richard E. Gross Problem-Solving Approach Model". Social Studies Review, Vol. 32, No. 3, pp. 374-381.

Cotton, C. (2014). "Real-world and Active – The Benefits of Problem-Based Learning". Teacher, 15 May. https://www.teachermagazine.com/au_en/articles/real-world-and-active-the-benefits-of-problem-based-learning

Gross, R.E. (1958). "The Problems Approach". In: Gross, R.E., Zeleny, L.D. (Editors), Educating Citizens for Democracy: Curriculum and Instruction in Secondary Social Studies. New York: Oxford University Press, pp. 341-367.

Lenz, B. (2014). "How Assessment Can Lead to Deeper Learning". Edutopia. 21 March. https://www.edutopia.org/blog/how-assessment-can-lead-to-deeper-learning-bob-lenz

Manolas, E.I. (2006). "A Critical Introduction to Deep Ecology in Higher Education: An Application of Kolb's Model of Experiential Learning". Economic and Environmental Studies, No. 8, 2006, pp. 245-252.

Naess, A. (1989). Ecology, Community and Life Style: Outline of an Ecosophy. Cambridge: Cambridge University Press.

Nair, S.S., Smritika, S.P., Thomas, K.A. (2020). "Revitalizing Education through Problem-Based Learning". Shanlax International Journal of Education, Vol. 9, No. 1, pp. 109-117.

Rothman, R. (2018). Measuring Deeper Learning: New Directions in Formative Assessment. Students at the Center: Deeper Learning Research Series. Boston, MA: Jobs for the Future.

Unangst, G. (2021). "Passive Learning vs Active Learning". Arizona State University. 15 April. https://www.asuprepdigital.org/student_blog/passive-learning-vs-active-learning/

University of Waterloo (n.d.). "Exploring Your Teaching Philosophy: Sample Exercises". Center for Teaching Excellence. Available: https://uwaterloo.ca/ce ntre-for-teaching-excellence/teaching-resources/teaching-tips/professional-devel opment/enhancing-your-teaching/exploring-your-teaching-philosophy

Wilson, J.A., Fernandez, M.L., Hadaway, N. (n.d.). "Mathematical Problem Solving". https://faculty.tarleton.edu/brawner/coursefiles/507/Problem%20solving%20article% 20by%20Wilson.pdf

Development Assistance for Health and the Role of NGOs in the Africa Region: The Case of the Central African Republic

Abstract:
In today's globalized society, financial aid for health is not only an obligation, but also a necessity to safeguard public health. Although efforts began long before the organization of the United Nations and later the targeted actions of the World Health Organization, a strong increase in DAH has been identified since our entry into the last millennium. Clearly, the development of new treatments and the use of financial tools have influenced this increase, but the rapid spread of communicable diseases that have been a direct threat to global health cannot be overlooked. Mainly, from 1990 onwards we have very detailed data on disease trends, reference indicators in the general population, as well as on the amount of financial assistance with detailed reports as to the scope of health operations (development of infrastructures, provision of drugs, educational programs, prophylaxis, etc.). In this study the Central African Republic is examined, one of the poorest countries in the world with an extremely low health budget. Without taking for granted the extent to which health indicators are influenced by health operations financing, the correlation between health indicators and financing amounts is highlighted.

Key Words: International Health Aid, Africa, NGOs, Development Assistance, Donors

Symeon Sidiropoulos[1], Stavroula Valachea[2], Maria-Eirini Kanakaki[3], Alkinoos Emmanouil-Kalos[4], Grigorios Tsimogiannis[5], Athanassios Vozikis[6]

---

[1] Symeon Sidiropoulos, Department of Economics, University of Piraeus; President of the Hellenic Association of Political Scientists, email: sec.president@hapsc.org
[2] Stavroula Valachea, Department of International and European Economic Studies, Athens University of Economics and Business; Hellenic Association of Political Scientists.
[3] Maria-Eirini Kanakaki, Department of International and European Studies, University of Piraeus; Hellenic Association of Political Scientists.
[4] Alkinoos Emmanouil-Kalos, Department of Economics, National and Kapodistrian University of Athens; Hellenic Association of Political Scientists, HAPSc Committee on Political Economy.
[5] Grigorios Tsimogiannis, Department of Informatics, University of Piraeus; Laboratory of Health Economics and Management, University of Piraeus, Greece.
[6] Athanassios Vozikis, Associate Professor, Department of Economics, University of Piraeus; Director of the Laboratory of Health Economics and Management, University of Piraeus, Greece.

1. Introduction

Global Health is defined as the field of study, research, and practice that focuses on promoting health and establishing health equity for all people around the world. However, global health should not be limited to health-related challenges that traverse international borders. Rather, in this context, global refers to any health issue that affects a large number of countries or is influenced by transnational causes such as climate change or urbanization (Koplan et al, 2009). Global health is significantly more complex than most people realize, and it takes more than just development assistance to attain. It is essentially an integrated approach in a global environment that includes action on social factors, social protection, local and national support for large health-care programs, and global agreements and commitments to handle the obligations and the perspectives of the different players involved (Kickbusch, n.d.).

Since its foundation, the United Nations has been actively concerned in promoting and preserving global health. The ambitious Sustainable Development Goals (SDGs), which contained 17 global goals aimed at economic and social development, were endorsed by UN member states in 2015. The third goal aims to achieve universal health coverage, better access to safe, effective, and inexpensive medicine, and the abolition of the HIV, malaria, and tuberculosis epidemics, by 2030 (United Nations, n.d. a). As noted by the WHO (2018), SDGs acknowledge good health as something bigger than identifying and managing specific diseases. To monitor health and well-being, the WHO focuses on "*...1. The level and distribution of healthy life that individuals and communities have, 2. The level and distribution of conditions that affect health and well-being, 3. The level and distribution of risk factors whose presence would affect health and well-being"* (WHO, 2018: 13).

The World Health Organization (WHO), whose constitution went into effect in April 1948 – a date we now commemorate as World Health Day – is in the forefront of this effort within the UN system. Malaria, women's and children's health, tuberculosis, venereal disease, nutrition, and environmental degradation were chosen as WHO's top objectives from the start (United Nations, n.d. b). Many of these, as well as relatively recent diseases such as HIV/AIDS, diabetes, cancer, and emerging diseases like SARS (Severe Acute Respiratory Syndrome), Ebola, and the Zika virus, remain on WHO's agenda today. The WHO is leading the international response to the Coronavirus Disease Pandemic (COVID-19). The International Classification of Diseases, which has become the international standard for classifying and reporting diseases and health problems, was given to WHO in 1948. WHO has been a part of several important achievements in global public health since its inception (United Nations, n.d. b). For example, the discovery of antibiotics began in 1950, and WHO was instrumental in counseling countries on how to use them. Furthermore, the involvement of WHO during the world's largest Ebola virus disease outbreak in 2014 in West Africa cannot be overlooked, as the activation of WHO was immediate. Medical doctors, public health specialists, scientists, epidemiologists, and other experts from the WHO are working on the ground in 150 countries across the world. They give technical advice to ministries of health and support with prevention, treatment, and care services within the health sector. WHO interventions touch on every aspect of global healthcare (United Nations, n.d. b). WHO, for example, intervenes in crises and responds to humanitarian disasters. It also strives to establish International Health Regulations, which countries must adhere to in order to detect and prevent disease outbreaks. In addition, WHO's work aids in the prevention of chronic diseases and the achievement of the Sustainable Development Goals connected to health (United Nations, n.d. b). According to the World Health Statistics

2020, WHO's annual snapshot of the world's health, Major infectious disease prevention and treatment coverage, as well as maternity, neonatal, and child health care, have all improved significantly in the last two decades, resulting in a continuous drop in disease incidence and mortality. The current rate of change, however, is insufficient to meet the 2030 SDG targets, and COVID-19 risks putting the world even more off track to meet the SDGs (WHO, 2020).

The global health governance (GHG) system has been embedded since the early 1990s on the international policy agenda. Clear evidence is found in the increased Development Assistance for Health (DAH) funding. Most of the non-state actors are not connected to issues related to DAH (Ng et al., 2011). As a result, states are the dominant actors within international society. GHG uses formal and informal institutions, rules and mechanisms, operated by states, intergovernmental organizations, and nonstate actors in order to collectively act and address cross-border health issues, such as access to vaccines (Fidler, 2010). When all actors act together, they mobilize resources and provide services. The institutionalized system tends to embody substantive normative commitment, increased active actors and funds. GHG presents some liabilities. It is underfunded, it has distorted funding structures, engaged in debates over the legitimacy of the actors, displays poor responses to some critical health plights.

In 1905 an International Congress on Tuberculosis was based on diverse multilateralism. It was observed that the involvement of NGOs seems to rise when governments need them and to fall when governments and international bureaucracies gain self-confidence (Charnovitz, 1997). In 1945, NGOs participated as consultants to the US delegation in San Francisco and assisted drafting the Article 71 of the UN Charter concerning NGO involvement in certain activities of the UN reaffirming the value of NGO participation.

Article 71 of the United Nations Charter, which established the Economic and Social Council (ECOSOC), states the following:

"The Economic and Social Council may make suitable arrangements for consultation with non-governmental organizations which are concerned with matters within its competence. Such arrangements may be made with international organizations and, where appropriate, with national organizations after consultation with the Member of the United Nations concerned."

— United Nations Charter, Chapter X, Article 71

ECOSOC in its guidelines in 1950 urged national organizations to present their views through international non-governmental organizations. NGOs' role is mostly common as part of CS in UN conferences, which were constituted as the new form of global governance (Tabbush, 2005). As they participated extensively, they could assist the international decision-making process, and not only take part in operational national or global affairs. The emergence of worldwide media helped NGOs gain visibility and thus, pursue transparency and public participation (Charnovitz, 1997). CS was seen as an international organizations' multiplier of organizations' legitimacy and constituency (Tabbush, 2005). The role of CS in forums and world conferences ranges from expert information and part of consultation to stakeholders to advocate. United Nations Non-Governmental Liaison Service (NGLS) was established in 1975 to promote the relation between the UN system and the civil society organizations (CSOs) facilitating the multistakeholder dialogue and the alliance-building (United Nations, n.d. c). The 1990s consultation arrangements with ECOSOC allowed NGOs to seek accreditation. The European Convention on the Recognition of the Legal Personality of International Non-Governmental Organizations set the legal basis for the existence and work of international NGOs in Europe (CoE, 1986). The 1992 Conference on Environment and Development, as well as the Commission for Sustainable Development (CDS) fostered NGOs' participation (Tabbush, 2005). NGOs facilitate the implementation of participatory democracy (UNDSD, 1992). The expansion of international society increased its ability to respond to crises (Tabbush, 2005). The model of CS as a pressure group inspired the creation of Social Watch. Much of the success of CSOs depends on the ability to create consensus.

The relationship between government and CS changed. Since the institutional arrangement treated CS as one of its stakeholders, the CS needed to construct a harmonized position (Ghaus-Pasha, 2005). Some donors and UN agencies increased their funds for CS's participation in UN conferences. The NGOs' participation compounds from three phases; the outsiders, the informal insiders, and the official and active participants. Since the 2000s NGOs have been more open to partnerships with businesses. The partnerships between UN organizations, civil society and corporations consist of modus operandi in the UN system. CS is one of the prime driving forces on HIV/AIDS, and among those who launch universal policy debates. The ODA for the support of NGOs erupted after 1999. Private foundations and NGOs have come to provide an increasing percentage of global health aid (at least 20% of the world total since 2000) (Youde, 2017). Three channels of influence of NGOs in global social policy making can be identified (Youde, 2017). The most common way that NGOs affect the global policy agenda is lobbying. But due to the diversity, NGOs were holding talks among themselves. There is also a surveillance-monitoring channel, and policy and decision-making channel (Lee, 2006; Youde, 2017). The relationship between NGOs and the UN varies from euphemistic to equivalent in health areas where the participation as executing agencies is pursued by IGOs. NGOs were aiming to raise awareness rather than organizing and having a political agenda, and as a result they did not build coalitions or broad long-term links with other NGOs. Though, they did create new linkages with NGOs and define the problem area. The UNCED conference process forced NGOs to develop their credibility and establish relationships with the most vocal governments, UN agencies and transnational corporations (Youde, 2017).

Primary Health Care (PHC) is the official WHO policy, which is supported by numerous international NGOs (Litsios, 2004). NGOs can be effective at any stage of the development of primary health care initiatives and promote improved understanding and positive attitudes about primary health care through various ways. To begin with, NGOs encourage and maintain discussion within and among NGOs, as well as with government officials. In the domains of health care and integrated human development, NGOs can help shape national policy. They can communicate health-care needs to relevant donor agencies based on their relationships with communities, and they can also interpret primary health-care plans for them. NGOs are also involved in all stages of the global health research cycle, helping to ensure that the study is relevant and effective, as well as identifying priorities and translating knowledge into action. They play an important role in stewardship (promoting and advocating for relevant global health research), research resource mobilization, knowledge generation, utilization, and management, and capacity building (Anbazhagan and Anbazhagan, 2016).

It is undisputed that NGOs play a vital role in the provision of Development Assistance for Health (DAH) and health system strengthening. DAH is defined by the

Institute for Health Metrics and Evaluation (IHME) as *"financial and in-kind contributions made by [...] Institutions whose principal mission is to provide development aid to developing countries in order to enhance their health"* (IHME, n.d.). Financial aid can also come from governments. Specifically, the US (33.8%) and UK (8.4%) governments, the German (4.2%) and Japanese (3.1%) governments, as well as the Bill & Melinda Gates Foundation (8.3%) were the primary sources of DAH in 2018 (Dieleman et al., 2019). Governments and NGOs can strengthen together the existing health systems. This can be achieved through strong partnerships between NGOs and Ministries of Health (MOH). In a true relationship, NGOs and the MOH work together to achieve the same aims. If a country lacks a vision or a plan, NGOs should provide their assistance and share their skills to aid in its development (Connelly, n.d.).

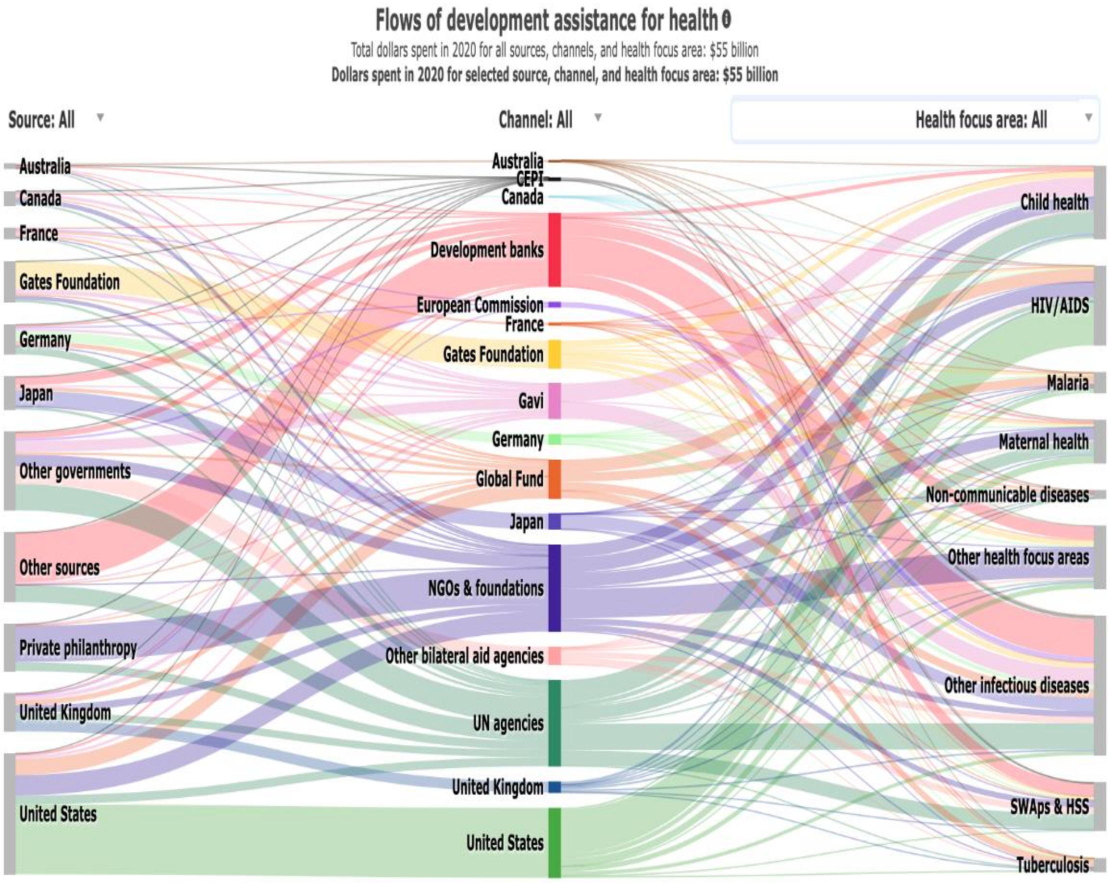
2. DAH, DALYs, YLLs

As mentioned above, the Development Assistance for Health (DAH) concerns financial and in-kind resources that are delivered via major international development agencies to low and middle-income countries, with the aim to maintain and improve health status (IHME, n.d.). Such agencies could be UNICEF, the United Kingdom's Department for International Development (DFID), or the Gates Foundation. Official Development Assistance (ODA) to the health sector is the largest constituent of DAH. DAH contains non-concessional loans, which may be provided by the World Bank, regional development banks in emergent states, and capital from private associations, as well as Non-Governmental Organizations, and which are directed to developing countries. The traditional actors in global health, national health ministries (MOH), and the World Health Organization (WHO), are now joined by the community, NGOs, private corporations, individual philanthropists, and public–private partnerships (PPPs) shaping the multilateral system actors (Szlezak et al., 2010). Programs, such as WHO's Roll Back Malaria Partnership (RBM), Stop TB, the Global Alliance for Vaccines and Immunization (GAVI), and the Global Fund to Fight AIDS, Tuberculosis, and Malaria (GFATM) were established. Governments provide ODA through two major routes: through direct bilateral aid, or multilateral organizations (WHO, 2002). The redefinition of the existing DAH system has been further stropped by the economic crisis and the following stagnation from the traditional donor countries (Moon and Omole, 2013). Since the 2000s, international NGOs and institutions were already providing a significant proportion of assistance in health.

Disability-adjusted life years (DALYs), as regards for a disease or health condition, are the total of years lost due to early death (YLLs) and years lived with disability (YLDs) concerning a population living in states of low good health levels due to a particular cause. DALYs is also defined as years of healthy life lost (WHO, n.d. a). The DALYs indicator is a more inclusive measure (WHO, 2004). DALYs is mostly used in international public health policy (Ferrucci et al., 2007). According to the World Bank (2013), DALY measures both the global burden of disease and the effectiveness of health interventions, as revealed by reductions in the disease burden. This measure was introduced in 1993's World Development Report. The DALY information faction is comprised of age, sex, disability status, and period, excluding any personal socioeconomic circumstances (Abbasi, 1999).

Years of life lost (YLLs) due to premature mortality are estimated as the figure of cause-specific deaths multiplied by a loss capacity determining the years lost with reference to the global standard life expectancy at the age at which death occurs. The loss capacity is grounded on the frontier national life expectancy forecasted for the year 2050 by the World Population Prospects 2012, with a life expectancy of 92 years. To estimate YLLs, age, sex, and the cause are counted (WHO, n.d. b). Such an indicator can be used in public health strategy in order to correlate the relative significance of distinct premature death causes in a certain population, to set priorities for prevention, and compare the premature mortality backgrounds between populations. YLLs has been used to describe non-communicable diseases, drug misuse, and suicides, as well as to characterize deaths due to coronavirus disease 2019 (COVID-19), both in multinational and single-country analyses (Quast et al, 2021).

Image 1. Flows of development assistance for health. Total dollars spent in 2020 for all sources, channels, and health focus area: $55 billion. Dollars spent in 2020 for selected source, channel, and health focus area: $55 billion.



Source: Institute for Health Metrics and Evaluation (2021a)

3. Africa is Under-Resourced

Africa is the region with countries of the worst standards in diseases leading to morbidity and mortality. This makes it the region with the lowest healthy life expectancy (at 52,3 years), 16,4 years less than Western Pacific Region (WHO, 2018). The Global Action Plan 2013-2020 identifies four conditions as a priority (chronic respiratory disease, cardiovascular disease, cancer and diabetes), by addressing the risk factors that are mainly associated with them (alcohol abuse, insufficient physical activity, unhealthy diets and tobacco use) (WHO, 2013). The primary health care strategy was limited implemented in the 1980s due to economic recession, the rise of neoliberalism, market ideologies, structural adjustment policies and new epidemics. For these reasons, it was replaced by selective primary health care, which focused on efficient and effective interventions. The Declaration of Alma Ata (1978) established the notion of international responsibility for health, that wealthy countries should assist poorer states to achieve health goals. The Bamako Initiative (BI) in 1987 aimed to deal with the problem of financing health services experienced in the 1980s in many countries especially in sub-Saharan Africa by increasing resource availability, providing start-up funds for community financing schemes, such as the revolving drug funds (Anaemene, 2017).

The development in Africa is affected by the triple burden (communicable, non-communicable, injuries). Sub-Sahara achieved the Millenium Development Goal (MDG) in regards to HIV in 2015 as the access to HIV treatment increased (Prendergast, 2015). Infections were reduced. Malaria in children declined from 26% in 2000 to 14% in 2013 and, in general, by 42% between 2000 and 2015. Tuberculosis treatment success noted a figure of 86% in 2013. Still, the high rates of hunger and malnutrition reveal the socio-economic barriers, chronic poverty, inadequate water and sanitation, and access to safe food. In 2015 the majority of member states did not sufficiently finance the health sector (Anyangwe, et al., 2007). External aid is high and social protection system as well as health insurance inversely to the population. Key investment strategies are not implemented. Also, investment is directed towards urban secondary or tertiary health infrastructure instead of primary care, and towards curative care instead of prevention. In addition, the medical supplies are limited. The non-transparency, unaccountability and insufficient engagement of stakeholders contribute to the health sector inefficiencies. Also, the participation of the innovative private sector to co-finance interventions is not encouraged. In the absence of a third payer system, the financial burden of care falls on patients and families, which in a background limited in resources equals a challenge for a large segment of the population (Kaze et al., 2012).

Another common reason for Africa's underfinancing is the natural as well as man-made disasters. Some states are constant battlefields and as a result, governments invest their budget in the security sector. Governments find difficulties in maintaining financing levels in the health sector. NGOs assist their maintenance especially since long-term health strategies may not be yet achieved. However, in some cases, state governments do not trust NGOs' motives and they complicate the financing procedure (Abdallah and Khondlo, 2015). Also, specific population groups, such as Internally Displaced Population (IDP), are not an attractive market for private donors (Abdallah and Khondlo, 2015).

In 2007, the African Union developed the first Africa Health Strategy (AHS) 2007 – 2015, where the main issues of the health sector are described as health systems' weakness, services too under-resourcing, and not universal access. Another issue was that health interventions do not always comply with the scale of the problem. AHS 2016-2030 describes the strategic direction in Africa's Member States to create better performing health sectors. In AHS 2016-2030, the strengthening of multi-sectoral collaboration, private-public partnerships, the capacity of MOH to facilitate and monitor activities of non-state suppliers, and the ensuring of effective accountability for its results are crucial goals. Collaboration between the member states, regional committees, multilateral agencies, bilateral development agencies and other partners in Africa is essential. The AU has called for international resources and cooperation between the WHO, UNAIDS, other UN and regional organizations. Another issue that needs to be addressed is the limited availability of data in the African Region, which undermines the accurate monitoring of health and well-being (WHO, 2018).

4. Donors

As Radelet (2006) notes, most help has generally been provided as respective help starting with one government then onto the next. Benefactors additionally give help in a roundabout way through multilateral help, which consolidates subsidizing from various sources. The World Bank, the International Monetary Fund, the African, Asian, and Inter-American Development Banks, and several United Nations agencies such as the United Nations Development Program are among the major multinational institutions.

There are various ways in which financial aid can be measured. The first way to estimate aid is in complete dollars. The second one is as a portion of GDP and finally per capita. Different things are revealed by each form of action. Absolute dollar sums are definitely a significant indicator; however, they don't recount the whole story. Help estimated as a portion of GDP demonstrates its size comparative with the whole economy and is maybe the most widely recognized measure (Radelet, 2006).

One of the World Bank pillars and one of its main priorities is the universal health coverage. The focus areas of the WB are the following (updated as to 2022): a. Pandemic Preparedness and COVID-19 (coronavirus) *"Pandemics, which are large disease outbreaks that affect several countries, pose major health, social, and economic risks."*, b. Universal Health Coverage *"Quality, affordable health care is the foundation for individuals to lead productive and fulfilling lives and for countries to have strong economies."*, c. Infectious Diseases and Vaccines *"Globally the incidence of infectious diseases has declined since 2000, but they continue to have major health and economic costs. Infectious diseases can lock people into poverty, undermine the resilience of communities, and have devastating consequences for a country's economy."*, d. Nutrition. *"Malnutrition is one of the world's most serious but least-addressed development challenges. Its human and economic costs are enormous, falling hardest on the poor, women, and children."* (World Bank, 2022).

In 2020, the US government had donated more than $7.4 billion in humanitarian relief around the world. Germany came in second with roughly 2.1 billion dollars, followed by the United Kingdom with 1.3 billion dollars (Szmigiera, 2021). Apart from governments, NGOS contribute to distributing and executing assistance dynamically. When the recipient countries are not able to conduct the essential transformations to meet the expectations of their donors, then NGOS are becoming partners of the countries in need and get involved in the process of macroeconomic structural transformation (Nair, 2013).

Foundations also play a crucial role in global health. The first philanthropic foundation with a strong impact on international aid was The Rockefeller Foundation's International Health Division, which was founded in 1913. The Rockefeller Foundation provided the largest amounts of financial aid until the 1940s. Nowadays, there are numerous foundations engaging with global health issues such as the Bill & Melinda Gates, the Clinton, and Bloomberg Foundations (Rushton & Williams, 2011). The Bill & Melinda Gates has created three major initiatives. The first is a US program aiming at secondary and post-secondary education. The second program is an international development program aimed at eradicating hunger and poverty, while the third is a worldwide health initiative (Bill & Melinda Gates Foundation, 2007).

The Institute for Health Metrics and Evaluation (n.d.) divides funding changes in distinct phases. From 1990-2001, global health funding underwent a 'moderate growth' phase. From 2001-2010, DAH increased at an average annual rate of 11.2%.

Since 2010 donor states' financing has been characterized as a 'no growth' period where DAH funding was relatively stagnant. When decline in funding was observed, as with the Global Fund to Fight AIDS, Tuberculosis and Malaria which reported a decline in funding from $8.7 billion in 2009 and $7.6 billion in 2010, the fund was restructured, refurbishing its grant administration process and engaging new leadership. In 2011, WHO recorded a $300 million shortage in budget.

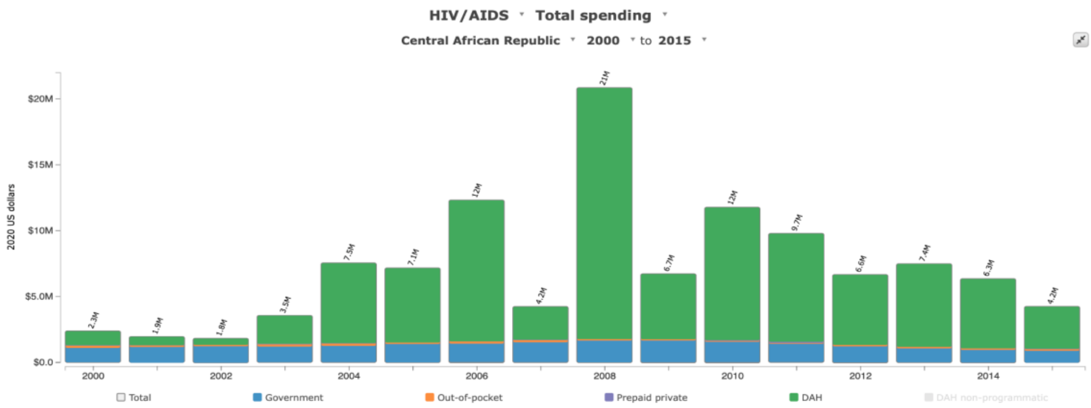
In 1990, 90% of all DAH came through bilateral channels, UN agencies, and regional development banks. Private foundations and NGOs played an insignificant role. In the following years, as the amount of money dedicated to DAH increased rapidly, more and more types of actors got involved in providing such aid, particularly a growth in private foundations and NGOs and public-private partnerships. A significant portion of that increase comes from the involvement of the BMGF, which is the world's prosperous philanthropic organization. It provides grants through its Global Health Program and with additional global health-related funding for issues through its Global Development Program. IHME's accounting shows that NGOs and other US-based private foundations provide significant funding for global health, too. Between 1990-2013 NGOs provided more than $57 billion for global health.

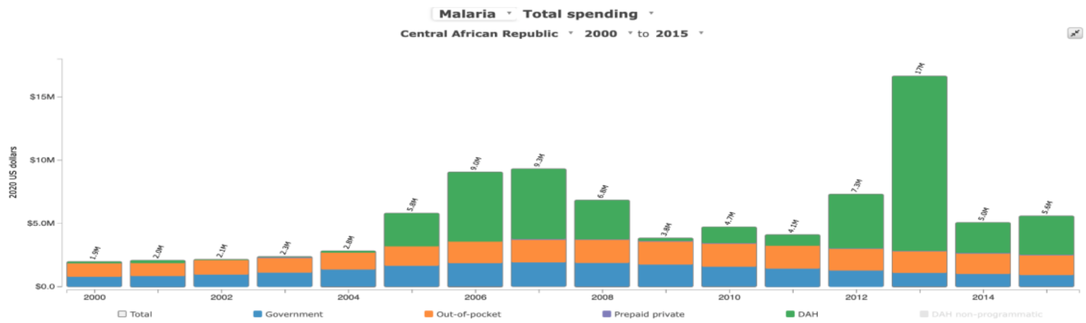
5. The Case of the Central African Republic

The Central African Republic (CAR) is one of the poorest countries in the African Region (WHO, 2018), with 75% of its population living below the poverty line and 50% of its population suffering from starvation (OCHA, 2022). Access to healthcare is extremely limited and the health statistics are mainly unreliable or non-existent (OCHA, 2022). On top of that, the HIV/AIDS epidemic and conflicts in the area have worsened mortality rates even further (WHO, 2016). In order to achieve the SDGs, the issues of poor governance and the heavily underfunded, uncoordinated and ineffective health system have to be addressed. At the same time, the high maternal and children mortality, as well as the high prevalence of both communicable and non-communicable diseases, must be prioritized. Concerning the reduction of communicable diseases, and in particular HIV, tuberculosis and malaria, the development of national policies and programs is supported by the WHO for both the prevention and the treatment of the diseases, through expertise, campaigns and the use of strategic information (WHO, 2016).

As shown in Images 2 and 3, especially in the cases of HIV and Malaria, the vast majority of the financing comes through DAH, rendering the efforts to combat them highly dependent on external aid, threatening its sustainability.

Images 2,3. Total Health Spending in the Central African Republic from 2000 to 2015 for HIV/AIDS and Malaria respectively.





Source: Institute for Health Metrics and Evaluation (2021a).

Central African Republic's DAH and DALYs
1995 - 2015

	HIV/AIDS		Malaria	
Year	DAH HIV	DALYs HIV	DAH Mal	DALYs Mal
1995	2.2M	570.0K	1.0K	240.0K
1996	60.0K	630.0K	3.0K	270.0K
1997	140.0K	680.0K	11.0K	290.0K
1998	130.0K	710.0K	3.0K	310.0K
1999	100.0K	730.0K	12.0K	330.0K
2000	1.0M	740.0K	92.0K	350.0K
2001	600.0K	730.0K	140.0K	370.0K
2002	410.0K	730.0K	7.0K	400.0K
2003	2.2M	720.0K	22.0K	430.0K
2004	6.2M	700.0K	48.0K	460.0K
2005	5.7M	680.0K	2.6M	470.0K
2006	11.0M	650.0K	5.4M	500.0K
2007	2.5M	620.0K	5.5M	500.0K
2008	19.0M	570.0K	3.0M	490.0K
2009	4.9M	530.0K	190.0K	460.0K
2010	10.0M	520.0K	1.2M	440.0K
2011	8.4M	520.0K	810.0K	420.0K
2012	5.3M	510.0K	4.2M	400.0K
2013	6.2M	500.0K	14.0M	380.0K
2014	5.3M	460.0K	2.4M	360.0K
2015	3.2M	410.0K	3.0M	350.0K

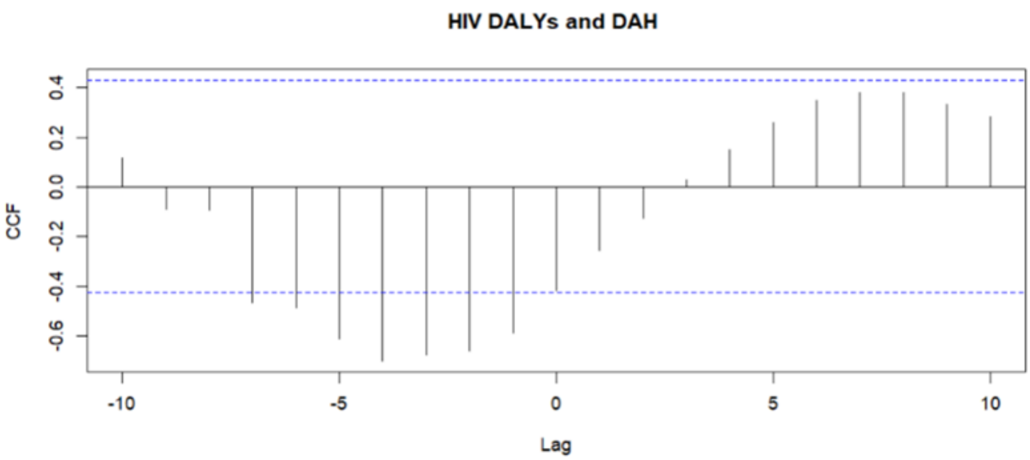
It is essential to analyze the correlation between all the DAH allocated to the Central African Republic and the corresponding DALYs indicator. In an effort to do so, an analysis on the time-series data is required.

The sample cross-correlation helps identify lags of the x-variable (DAH HIV), which might be useful predictors of y-variable (DALYs HIV).

$$r_k = \frac{\sum_{i=1}^{n-k} (X_i - \bar{X})(Y_{i+k} - \bar{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2 \sum_{i=1}^{n} (Y_i - \bar{Y})^2}}$$

CCF plot suggests that the highest correlation value is located at a lag of four (4) years. If the lag is negative then it is said that x leads y, in other words, DAH leads DALYs in HIV cases. Furthermore, a negative correlation between two variables indicates that these two variables tend to move in different directions. For instance, in case of an increase in DAH, the DALYs will decrease with a delay.

Sources: Institute for Health Metrics and Evaluation (2016, 2018, 2020, 2021b)



It is difficult to read the exact values of correlations and the lags from the plot, so a lag-correlation table is needed in order to interpret the results.

HIV/AIDS DAH and DALYs year-lagged correlation 1995 - 2015

Correlation Coefficient	Lag
-0.48542292	-6
-0.61160525	-5
-0.70231664	**-4**
-0.67541928	-3
-0.65979438	-2
-0.58846258	-1
-0.41637331	0
-0.25627651	1
-0.12450751	2
0.02699418	3
0.15112264	4
0.25903457	5
0.34944330	6

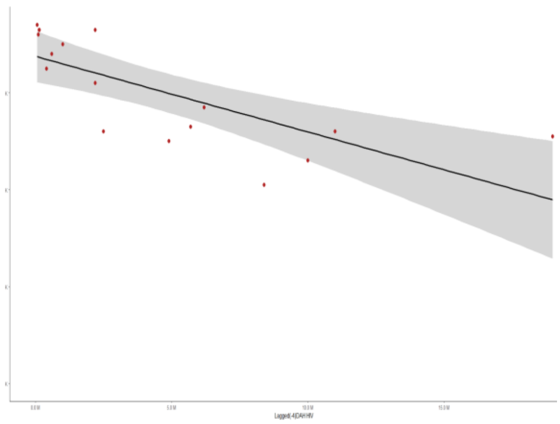
Correlation coefficient with a value of -0.7 is considered a strong negative linear relationship between the two variables. As a means to analyze the data more, a simple linear model using ordinary least squares (OLS) method will be created between dependent y-variable (DALYs HIV) and the lagged (-4 years) independent x-variable (DAHs HIV).

Linear Model's Summary:

```
MODEL INFO:
Observations: 17 (4 missing obs. deleted)
Dependent Variable: DALYs HIV
Type: OLS linear regression

MODEL FIT:
F(1,15) = 17.96, p = 0.00
R² = 0.54
Adj. R² = 0.51

Standard errors: OLS
------------------------------------------------------
                     Est.       S.E.   t val.      p
---------------- ---------- --------- -------- ------
(Intercept)      675437.00  24684.00    27.36   0.00
lag_m_4_DAH          -0.02      0.00    -4.24   0.00
------------------------------------------------------
```



Linear Regression Plot with 95% C.I.

6. References

Abbasi, K. (1999). The World Bank on world health: under fire. *British Medical Journal*, *318*(7189), 1003–1006. https://doi.org/10.1136/bmj.318.7189.1003

Abdallah, I. A. Y. and Khondlo, M. (2015). The role of non-governmental organizations in providing curative health services in North Darfur State, Sudan. *African Health Sciences*, 15(3): 1049–1055.

African Union (2007). Africa Health Strategy 2007 – 2015. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.469.1991&rep=rep1&type=pdf

African Union. (2016). Africa Health Strategy 2016 – 2030. https://au.int/sites/default/files/documents/24098-au_ahs_strategy_clean.pdf (Accessed: 13 July 2021)

Anaemene, B. U. (2017). Health, Agenda 2030 and the Future of Africa's Development. In: Anaemene, B. U. and Bolarinwa, O. J. (eds). *Agenda 2030 and Africa's Development in the 21st Century.* United Nations University, International Institute for Global Health.

Anbazhagan, S. and Anbazhagan, S. (2016). Role of non-governmental organizations in global health. *International Journal of Community Medicine and Public Health, 3*(1), 17-22.

Anyangwe, S. C.E., Chipayeni, M. (2007). Inequities in the Global Health Workforce: The Greatest Impediment to Health in Sub-Saharan Africa. *International Journal of Environmental Research and Public Health, 4*(2), 93-100. https://doi.org/10.3390/ijerph2007040002

Bill & Melinda Gates Foundation (2007). Annual report 2007. https://docs.gatesfoundation.org/Documents/2007Gates%20Foundation%20Annual%20Report.pdf

Charnovitz, S. (1997). Two Centuries of Participation: NGOs and International Governance. *Michigan Journal of International Law*, *18*(2), 183-286. https://repository.law.umich.edu/mjil/vol18/iss2/1

Connelly, M. (n.d.). *The Role of NGOs in Strengthening Health Systems*. Global Health Delivery online.

Council of Europe (CoE) (1986). *European Convention on the Recognition of the Legal Personality of International Non-Governmental Organisations.* ETS 124. Available at: https://www.coe.int/en/web/youth/-/european-convention-on-the-recognition-of-the-legal-personality-of-international-non-governmental-organisations (Accessed: 27 July 2021)

Dieleman, J. L., Micah, A. E., & Murray, C. J. (2019). Global health spending and development assistance for health. *Journal of the American Medical Association, 321*(21), 2073-2074.

Ferrucci, L., Koh, C., Bandinelli, S., Guralnik, J.M., Birren, J. E. (eds). (2007). Disability, Functional Status, and Activities of Daily Living. *Encyclopedia of Gerontology* (Second Edition). Elsevier. 427-436. https://doi.org/10.1016/B0-12-370870-2/00075-5 (Accessed 7 July 2021).

Fidler, D. P. (2010). The Challenges of Global Health Governance. *Council on Foreign Relations,* Working Paper. http://www.sadil.ws/bitstream/handle/123456789/160/Challenges-of-Global-Health-Governance.pdf?sequence=1&isAllowed=y

Ghaus-Pasha, A. (2005). Role of Civil Society Organizations in Governance. *6th Global Forum on Reinventing Government Towards Participatory and Transparent Governance.* 24 – 27 May 2005, Seoul, Republic of Korea. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.625.1861&rep=rep1&type=pdf

Institute for Health Metrics and Evaluation (2016). *Financing Global Health 2015: Development assistance steady on the path to new Global Goals*. Seattle, WA

Institute for Health Metrics and Evaluation (2018). *Global Health Spending 1995-2015*. Global Burden of Disease Health Financing Collaborator Network. Seattle, United States

Institute for Health Metrics and Evaluation (2020). *Global Burden of Disease Study 2019 (GBD 2019) Results.* Global Burden of Disease Collaborative Network. Seattle, United States. http://ghdx.healthdata.org/gbd-results-tool

Institute for Health Metrics and Evaluation (2021a). *Financing Global Health Visualization*. Seattle, WA: IHME, University of Washington. http://vizhub.healthdata.org/fgh/ (Accessed: 20 January 2022).

Institute for Health Metrics and Evaluation (2021b). *Development Assistance for Health Database 1990-2020*. Seattle, United States of America.

Institute for Health Metrics and Evaluation (n.d.). *Terms Defined*. http://www.healthdata.org/terms-defined

Kaze, F. F., Ashuntantang, G., Kengne, A. P., Hassan, A., Halle, M. P., Muna, W. (2012). Acute hemodialysis complications in end-stage renal disease patients: The burden and implications for the under-resourced Sub-Saharan African health systems. *Hemodialysis international*, 16(4), 526-531. https://doi.org/10.1111/j.1542-4758.2012.00692.x

Kickbusch, I. (n.a.). *Advancing the Global Health Agenda*. United Nations. https://www.un.org/en/chronicle/article/advancing-global-health-agenda

Koplan, J. P., Bond, T. C., Merson, M. H., Reddy, K. S., Rodriguez, M. H., Sewankambo, N. K., & Wasserheit, J. N. (2009). Towards a common definition of global health. *The Lancet, 373*(9679), 1993-1995.

Lee, J. (2006). *Comparing NGO Influence in the EU and the U.S*. Centre for Applied Studies in International Negotiations, Brief. https://www.files.ethz.ch/isn/25102/ngoinfluenceinuandusa.pdf

Litsios, S. (2004). Primary Health Care, WHO and the NGO Community. *Society for International Development*. 47 (2), 57- 63. http://dx.doi.org/10.1057/palgrave.development.1100030

Moon, S. and Omole, O. (2013). Development Assistance for Health: Critiques and Proposals for Change. *Chatham House (The Royal Institute of International Affairs).* https://www.chathamhouse.org/sites/default/files/public/Research/Global%20 Health/0413_devtassistancehealth.pdf

Nair, S. (2013). Governance, representation and international aid. *Third World Quarterly, 34*(4), 630-652.

Ng, N. Y., & Ruger, J. P. (2011). Global Health Governance at a Crossroads. *Global Health Governance, 3*(2), 1–37. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3983705/

Prendergast, A., Essajee, S. and Penazzato, M. (2015). HIV and the Millennium Development Goals. *Archives of Disease in Childhood*, 100, 48-52. https://adc.bmj.com/content/100/Suppl_1/S48.short

Quast, T., Andel, R., Gregory, S. Storch, E. A. (2021). Years of life lost associated with COVID-19 deaths in the USA during the first year of the pandemic, *Journal of Public Health*, fdab123. https://doi.org/10.1093/pubmed/fdab123

Radelet, S. (2006). A primer on foreign aid. *Center for Global Development*, working paper 92. https://www.cgdev.org/publication/primer-foreign-aid-working-paper-92

Rushton, S., & Williams, O. (Eds) (2011). *Partnerships and foundations in global health governance*. Springer.

Szlezák, N. A., Bloom, B. R., Jamison, D. T., Keusch, G. T., Michaud, C. M., Moon, S., Clark, W. C. (2010) The Global Health System: Actors, Norms, and Expectations in Transition. *PLOS Medicine*, 7(1), e1000183. https://doi.org/10.1371/journal.pmed.1000183

Szmigiera, M. (2021). *Largest donors of humanitarian aid worldwide 2020, by country*. Statista. https://www.statista.com/statistics/275597/largers-donor-countries-of-aid-worldwide/

Tabbush, C. (2005). Civil Society in United Nations Conferences. A Literature Review. *United Nations Research Institute for Social Development (UNRISD)*. 17. https://unrisd.org/unrisd/website/document.nsf/ab82a6805797760f80256b4f005da1ab/2698fce2fdac8826c125708800269fe8/$FILE/tabbush.pdf (Accessed: 15 July 2021)

UNDSD (1992). *United Nations Conference on Environment & Development, Rio de Janerio, Brazil, 3 to 14 June 1992. AGENDA 21.* United Nations Division for Sustainable Development (UNDSD). https://sdgs.un.org/sites/default/files/publications/Agenda21.pdf (Accessed: 19 July 2021)

United Nations (n.d. c). *Civil Society. Our story*. https://www.un.org/en/civilsociety/ngls/ourstory (Accessed: 16 July 2021)

United Nations (UN) (n.d. a). *Sustainable Development Goals*. https://www.un.org/sustainabledevelopment/.

United Nations (UN) (n.d. b). *Global Issues: Health*. https://www.un.org/en/global-issues/health

WHO (1978). *Declaration of Alma-Ata*. Available at: https://www.who.int/publications/almaata_declaration_en.pdf

WHO (2002). *International Development Assistance and Health*. Commission on Macroeconomics and Health, Working Group 6. http://apps.who.int/iris/bitstream/handle/10665/42548/9241590140.pdf?sequence=1

WHO (2004). *Global burden of disease: DALYs by age, gender, region and cause for the year 2004*. UNESCO. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/SC/temp/wwap_pdf/Disability_adjusted_life_year.pdf (Accessed: 7 July 2021)

WHO (2013). *Global Action Plan for The Prevention and Control of Noncommunicable Diseases 2013-2020.* http://apps.who.int/iris/bitstream/handle/10665/94384/9789241506236_eng.pdf?sequence=1

WHO (2016). Country Cooperation Strategy at a Glance: Central African Republic. WHO/CCU/16.03/ Central African Republic.

WHO (2018). *The state of health in the WHO African Region: an analysis of the status of health, health services and health systems in the context of the Sustainable Development Goals*. Brazzaville: WHO Regional Office for Africa

WHO (2020). World Health Statistics 2020: Monitoring Health for the Sustainable Development Goals (SDGs). *World Health Organization (WHO).* https://apps.who.int/iris/bitstream/handle/10665/332070/9789240005105-eng.pdf (Accessed: 22 July 2021).

WHO (n.d. a). Disability-adjusted life years (DALYs). https://www.who.int/data/gho/indicator-metadata-registry/imr-details/158 (Accessed: 7 July 2021).

WHO (n.d. b). *Years of life lost (YLL) (per 100.000 population)*. https://www.who.int/data/gho/indicator-metadata-registry/imr-details/4427

World Bank (1993). *World Development Report 1993*. https://doi.org/10.1596/0-1952-0890-0

World Bank (2022). https://www.worldbank.org/en/topic/health (Accessed: 22 January 2022)

Youde J. (2017). Global health Governance in International Society. *Global Governance: A review of Multilateralism and International Organizations*, *23*(4), 583-600. https://doi.org/10.1163/19426720-02304005

COVID-19 and Spatial Interaction: Evidence from the Regions of Greece[1]

Abstract:
This article attempts to provide an examination of the COVID-19 cases across the regions of Greece. Moreover, it attempts to examine the extent of spatial interaction across the NUTS-3 regions is assessed by means of simple descriptive statistics.

Keywords: Covid19, Greece, Spatial interaction, descriptive statistics

Alexiadis, Stilianos [2]

[1] The findings, interpretations and conclusions are those entirely of the author and do not necessarily represent the official position, policies or views of the Ministry of Rural Development & Food and/or the Greek Government.

[2] Corresponding-Address: Dr. Alexiadis, Stilianos Ministry of Rural Development & Food, Department of Strategic Planning of Rural Development, Evaluation & Documentation, Division of Documentation & Agricultural Statistics, Room 519, 2 Acharnon St., 104 32, Athens, Greece

1. COVID-19 and Spatial Interaction:

This brief note provides an examination of the COVID-19 cases across the regions of Greece. As a first step, the extent of spatial interaction across the NUTS-3 regions is assessed by means of simple descriptive statistics. Several statistics have been put forward to describe interaction across space and one measure of spatial dependence that is used extensively in empirical studies is the Moran's *I* statistic, calculated as follows:

$$I_t = \frac{n}{s} \frac{\sum_i \sum_j w_{ij}(x_i - \mu)(x_j - \mu)}{\sum_i (x_i - \mu)^2}, \; i \neq j \qquad (1)$$

where n is the number of observations for a variable x and w_{ij} represents the spatial weight for each pair of regions i and j. The term μ denotes the mean of the data set for the variable x while s is a scaling constant, calculated as the sum of all spatial weights. In this task a principal issue is the construction of the spatial weights. A common practice is to allow these weights to take the value of 1 if a region is contiguous to another and 0 otherwise. In this case spatial interaction is presumed to occur only between regions that share a common border. Alternatively, the spatial weights may be continuous variables, constructed so as to produce declining weights as distance between regions increases. Thus,

$$w_{ij} = \frac{1/d_{ij}}{\sum_j 1/d_{ij}} \qquad (2)$$

Here, d_{ij} denotes the distance between two regions i and j, typically represented by the distance between the regions' main cities where the majority of economic activities are located (the centroid of a region). The denominator is the sum of the (inverse) distances from all regions surrounding region i, within a selected boundary. Equation (2) implies that interactions between regions, such as spillover effects, decay as the distance from one area to another increases (hence weights decline as distance increases) and that such effects are dominated by the leading area[3].

In constructing the spatial weights matrix to examine spatial dependence amongst the prefectures (NUTS-3 regions) of Greece a similar approach is adopted, where the weights are calculated using distances between the main cities of regions, given that economic activities are typically concentrated in the main city of each region. The numerator, thus, consists of the distance of the nearest main city in a region from the main city of the contiguous region.

This choice is based on the assumption that spillover effects are dominated by an area in which COVID-19 incidents were reported in the initial time of the analysis, and that such effects are greater in the nearby locations. Moreover, equation (2) allows accounting for the existence of island regions in the data set. To be more precise, in a simple contiguity matrix the weights for island-regions would be zero. That would change the sample size and the interpretation of the results. A spatial matrix based on the geographical distance provides a better approach to the issue of spatial dependence.

[3] Equation (2) is used extensively in spatial analysis (e.g. Richardson, 1974; Cliff and Ord, 1981; Anselin, 1988). Nevertheless, the spatial matrix can be constructed also using the inverse distances to the square as denominator. Results using this kind of spatial matrix were very similar.

Depending upon the spatial weights matrix, if $I_t > 0$ then this is indicative of high spatial autocorrelation, suggesting positive perfect spatial dependence. On the other hand, if $I_t < 0$, then this implies perfect negative spatial autocorrelation while if $I_t = 0$, then this indicates the absence of spatial dependence, that is no significant spatial links among the observational units.

Throughout the examined period (January 2021 to March 2021), the Moran's *I* statistic values were negative and statistically significant indicating some degree of spatial interaction across the NUTS-2 regions of the EU-27. Moreover, the correlation between the coefficient of variation and Moran's *I* statistic is positive, indicating that changes in the degree of spatial dependence follow a similar trend to that of changes in geographical distribution of COVID-19 incidents.

Spatial interaction can be depicted by Moran's scatterplot, which plots a regional characteristic against its spatial-lag (Anselin, 1988). In the present context, a region's spatial-lag is the weighted average of C19 incidents of its surrounding regions. On the vertical axis in Moran's scatterplot the spatial-lag of COVID-19 incidents in a region is measured while the horizontal axis measures COVID-19 incidents in a region. Thus, each Moran's scatterplot contains four quadrants that identify four different kinds of spatial interaction between regions (Figure 1).

Quadrant I includes regions with a high (above average) number of COVID-19 incidents that are surrounded also by regions with a high number of COVID-19 incidents, while quadrant III refers to those regions with a relatively low number of COVID-19 incidents which are surrounded also by regions in which a similar number of incidents is reported.

These two quadrants indicate positive spatial association, or correlation. On the other hand, quadrant II represents regions with low number of COVID-19 incidents, which are surrounded by regions with a relatively high number of COVID-19 incidents, while the reverse case is found in quadrant IV. Thus, quadrants II and IV exhibit negative spatial association.

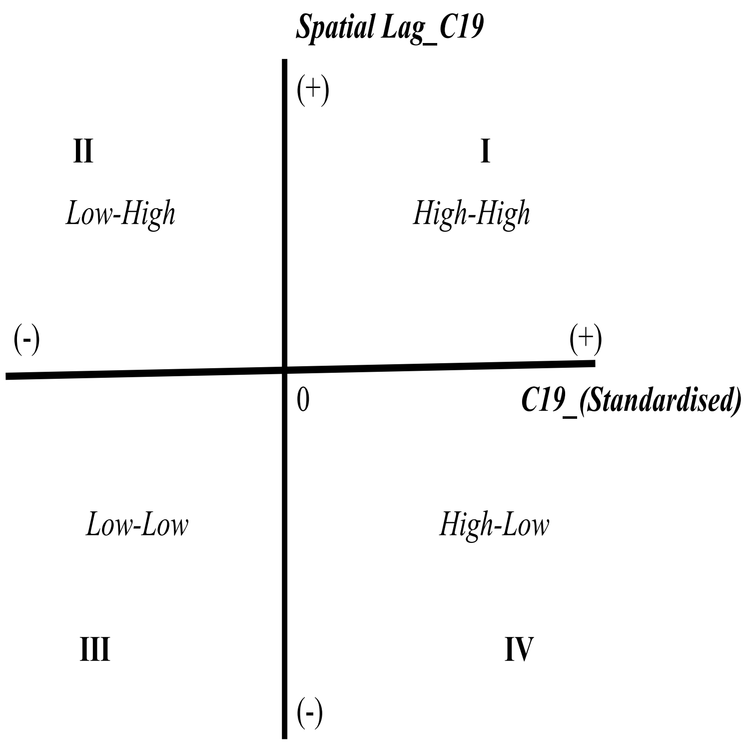


Fig. 1. Spatial autocorrelation in Moran's scatterplot

Significant positive spatial association implies that regions would spatially cluster into two distinctive groups, i.e. regions with a relative high number of COVID-19 and regions with relatively low number of COVID-19, in quadrants I and III. On the other hand, negative spatial association suggests regional grouping in quadrants II and IV, described by Rey and Montouri (1999) as ‘doughnut’ and ‘diamond in the rough’. It may be argued, therefore, that if observations scatter in both quadrants I and III, then a pattern of spatial clustering in two separate groups is a possibility. If, on the other hand, regions are scattered in quadrant IV, then this is compatible with a ‘core and periphery’ pattern, in the sense that regions with high presence of COVID-19 (centre or core) are surrounded by regions with low presence of COVID-19 (periphery) implying a pattern of regional divergence.

Figures 2 and 3 present the Moran’s scatterplots for the initial and terminal time of the analysis.

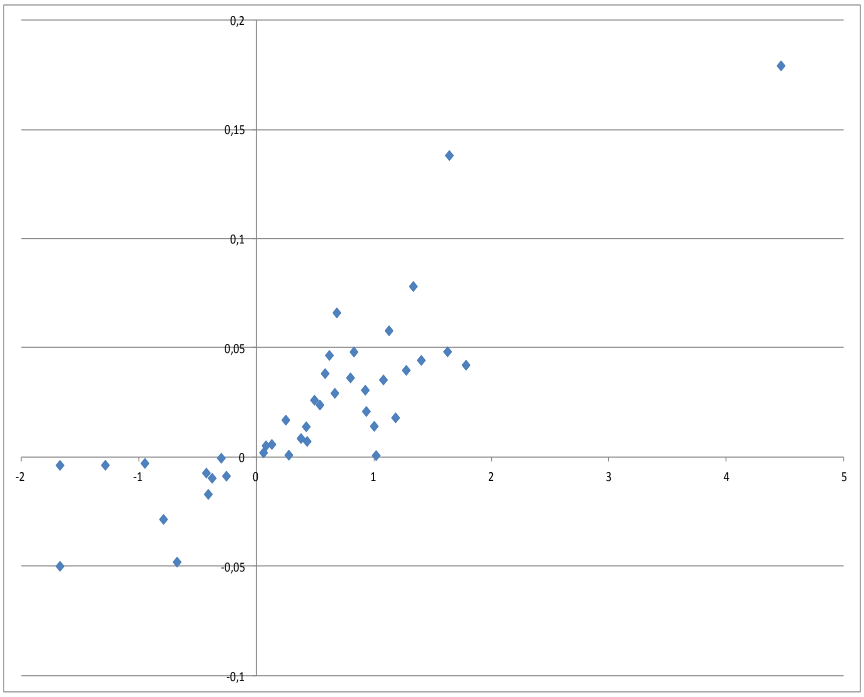


Fig. 2. Moran scatterplot, Initial Time, COVID-19 incidents, 51 Greek Prefectures

Figures 2 and 3 clearly suggest a pattern of clustering in the first and the third quadrants of both figures such that regions with relatively high levels of COVID-19 cases are likely to be neighbours of other high-cases regions. On the other hand, regions in which relatively low levels of COVID-19 cases are grouped with regions in which similar levels are located.

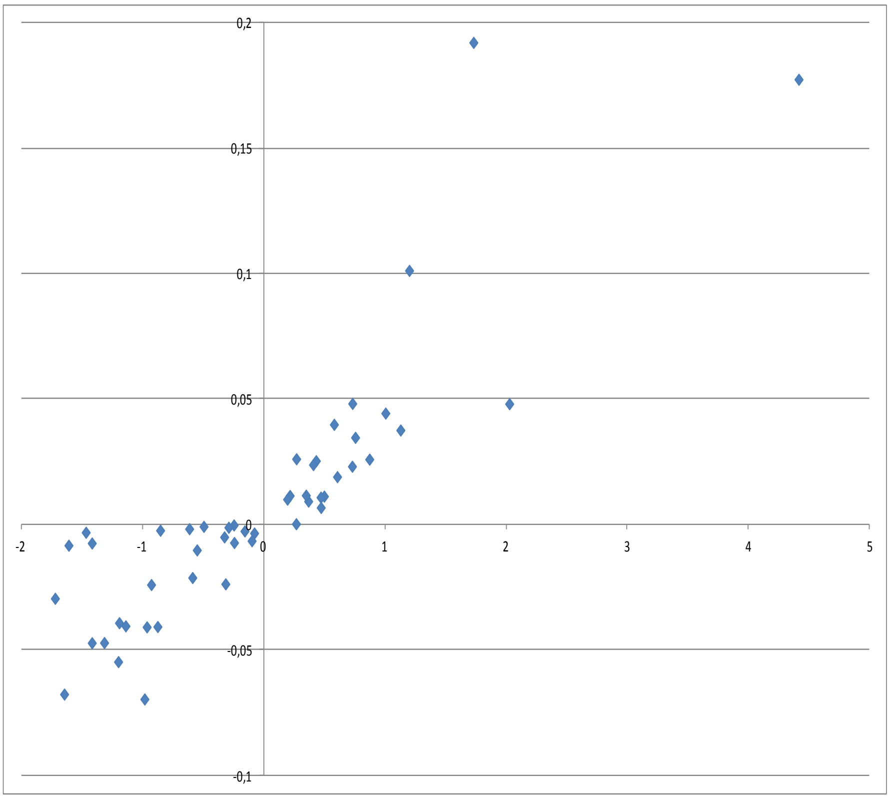


Fig. 3. Moran scatterplot, Terminal Time, COVID-19 incidents, 51 Greek Prefectures

This pattern is observed both at the beginning and at the end of the time period, suggesting that the data are not randomly distributed, but indicate a systematic spatial pattern, which is compatible with the 'regional-grouping' hypothesis, where there are two spatially connected groups. Identifying spatial dependence across this set of observational units suggests that any analysis or forecast of the COVID-19 pandemic should be modified to include an explicit spatial dimension.

2. References

Anselin, L. 1988. *Spatial Econometrics*: *Methods and Models*, Kluwer.

Cliff, A. and Ord, J. 1981. *Spatial Processes*: *Models and Applications*, Pion.

Richardson, H. 1974. Agglomeration potential: A generalisation of the income potential concept, *Journal of Regional Science*, 14 (3): 325-336.

Journal of Regional & Socio-Economic Issues

Call for Papers

Journal of Regional & Socio -Economic Issues (Print) ISSN 2049 -1395

Journal of Regional & Socio -Economic Issues (Online) ISSN 2049 -1409

The Journal of Regional Socio -Economics Issues (JRSEI, *indexed by Copernicus Index, DOAJ (Director of Opesn Access Journal), EBSCO & Cambell Index*) is scheduled to be published three times a year. Articles are now welcome for the forthcoming issue of this journal (JRSEI). The benefits of publishing in the Journal of Regional Socio -Economics Issues (JRSE I) include:

1. Fast publication times: your paper will appear online as soon as it is ready, in advance of print version
2. Excellent editorial standards
3. Free color electronic version
4. Free on -line access to every issu e of the journal
5. Rigorous, fast and constructive peer review process
6. The journal will be indexed in scientific databases.
7. All abstracts and full text are available free on -line to all main universities/institutions worldwide, ensurin g promotion to the widest possible audience.

For full paper submission guidelines, please visit the webpage:
www.jrsei.yolasite.com/
For further inquiry, please contact:
Professor Dr. George M. Korres, JRSEI Managing and Chief Editor
Professor, University of the Aegean, Department of Geography, Email: gkorres@geo.aegean.gr

Journal of Regional & Socio-Economic Issues (JRSEI)

Instructions to Authors

Journal of Regional & Socio-Economic Issues (Print) ISSN 2049-1395

Journal of Regional & Socio-Economic Issues (Online) ISSN 2049-1409

Aims of the Journal:
Journal of Regional Socio-Economic Issues (JRSEI) is an international multidisciplinary refereed journal the purpose of which is to present manuscripts that are linked to all aspects of regional socio-economic and all related issues. The journal indexed by Copernicus Index, DOAJ (Director of Open Access Journal), EBSCO & Cabell's Index and welcomes all points of view and perspectives and encourages original research or applied study in any of the areas listed above. The views expressed in this journal are the personal views of the authors and do not necessarily reflect the views of JRSEI journal. The journal invites contributions from both academic and industry scholars. If you have any questions about the journal, please contact the chief editor. Electronic submissions are highly encouraged (mail to: **gkorres@geo.aegean.gr**).

Review Process:

Each suitable article is blind-reviewed by two members of the editorial review board. A recommendation is then made by the Editor-in-Chief. The final decision is made by the Editor-in-Chief. If a revision is recommended, the revised paper is sent for a final approval to the Chief-Editor.

Instructions to Authors:

In order for a paper to be submitted to the Journal for publication, the following should be taken into consideration:

1. All papers must be in English.
2. Papers for publication should be sent both in electronic format (MS Word and MS Excel for charts) to the Chief Editor (mail to: **gkorres@geo.aegean.gr**).
3. The Editor takes for granted that:

- the submitted paper contains original, unpublished work that is not under consideration for publication elsewhere;
- authors have secured any kind of permission necessary for the publication from all potential co-authors, along with having agreed the order of names for publication;
- authors hold the copyright, have secured permission for the potential reproduction of original or derived material and are ready to transfer copyright of the submitted paper to the publisher, upon acceptance for publication.

4. The cover page should include the name of the author and coauthors, their affiliations, and the JEL category under which the paper primarily belongs. The cover page is the only page of the manuscript on which the names and affiliations of the authors and coauthors should be listed.
5. Submission of manuscripts in electronic form: Authors must submit electronic manuscripts. The submission should only contain the file(s) of the papers submitted for publication, in MS Word and MS Excel for charts. If more than one file, a compressed file (.zip) should be submitted instead.
6. Formatting requirements: Everything should be double-spaced (main text, footnotes, bibliography, etc.)
7. Footnotes should be as few and as short as possible (preferably devoid of tables or formulae), marked in the manuscript by superscripts in Arabic figures.

8. Formulae should be numbered by consecutive, Arabic figures (such as (1), (2), etc.), placed on the right-hand side of the page.
9. Tables and Figures should be numbered consecutively in Arabic figures and have a heading and a title.
10. References are citations of literature referred to in the text and should not appear as footnotes. Abbreviations are only accepted in the authors' first names. Place all references, alphabetized by author's last name (with last name first), on **separate pages** in a section titled "References" at the end of the paper. Indent the second and subsequent lines of each reference.

    Journals

    Include all authors, article title, full title of journal, volume number, issue number, month, year, and full page numbers. Example:

    Michael Mahmood. "A Multilevel Government Model of Deficits and Inflation," Economic Journal**,** 24, 2, June 2010, pp. 18-30.

    Books

    Include name of author, full title of book, edition, city and state (or country) of publisher, name of publisher, and year of publication. Example:

    Shapiro, John. Macroeconomics, 4th ed., New York, NY: Harcourt Brace Jovanovich, 2009.

    Use the following style when an author's work appears in a publication edited by another:

    George Summers, "Public Policy Implications of Declining Old-Age Mortality," in Gary ed., Health and Income, Washington, DC: The Brookings Institution, 1987, pp. 19-58.

    Public Documents

    Include the department or agency responsible for the document, title, any further description such as number in a series, city and state (or country) of publication, publisher, and date of publication. Example:

    World Bank. Educational Attainment of Workers, Special Labor Force Report 186, Washington, 2010.