GSL Journal of Public Health and Epidemiology

Global Scientific Library

Research Article

Life Expectancy does not Depend on Classical Ecological Variables: Stochastic and Non-stochastic Analysis

U. Cornelli, M. Recchia, E. Grossi, G. Belcaro

[1]Loyola University School of Medicine Chicago-USA, Piazza Novelli 5, Milan-20129, Italy
[2]Dept BiostatisticsUniversity of Lugano-Swiss Via Salaino 7, Milano-20144, Italy
[3]Villa Santa Maria Institute, Tavernerio, Italy
[4]Irwin Labs, University of Chieti, Strada Statale 16 Bis 94 Spoltore-65010 (PE), Italy

Abstract

The Life Expectancy (LE) in 191 countries was compared on the base of 17 demographic and economic variables. Two approaches were followed: the stochastic and non-stochastic analysis. The first consisted of simple correlation coefficients followed by the, Main Component, Factorial and finally Segmentation Analysis. The second was the Artificial Neuronal Network Analysis using the Auto Contractive Map (Auto-CM). The results of the two approaches were very similar and all excluded at least 10 out the 17 variables. The surface covered by Forests, the Kmq^2 of forests, the ratio of the Domestic Gross Profit (GDP) with education, the number of hospital beds, the particulate matter, the population and the population density were not considered determinant. What were emerging as directly correlated with LE were internet, GDPs (GDP/inhab and GDP 2 and GDP 3 related respectively to the advance industry and economy), urban concentration, cars and cellphones. An inverse correlation was found with GDP 1 (related to agriculture, livestock, fishing) and unemployment rate. The last was detected by the ANN only. The conclusion is that LE in the world is far from the variables typically bound to the environment and more linked to the economic variables.

One sentence summary: The life expectancy measured with stochastic and non stochastic analysis excludes the correlation with the common ecological variables.

Introduction

Life Expectancy (LE) has been steadily rising in the past two centuries and projected to continue increasing. From the 40 year in 1990 [1,2] has doubled since the beginning of the 20th century, and in most developed countries now exceeds 80 years in both man and women [3] with an 80% probability that world population will increase to between 9.6 and 12.3 billion in 2100 [4]. LE is increasing in every area of the world and already in 2010 the International Future modeling system (IFs) was calculating values ranging from 73 to 85 years for males and 80 to 87 years for females [5]. Particular improvements were forecasted for Sub-Saharian Africa, whereas slower improvements were expected in high-income countries. Considering data deriving from The Global Burden Disease Study 2015 the total death increased by 4.1% from 2005 to 2015, but the age standardized death rate fell by 17% [6].

After 2015 many efforts were made by United Nation to achieve a world of prosperity, equity, freedom and peace-according to the Millennium Development Goals-to stimulate the implementation of primary health care. However, the gap between countries is still existing and occurs for different reasons that should by analyzed with details [7]. Indeed, a wide array of disease and injury sequelae affects the world's population. Globally, in 2013 only 4.3% of the population had no burden of disease or injury sequelae, slightly up from 4.2% in 1990 and the Years Lived with Disability (YLDs) was also increasing [6]. According to these considerations, globally LE at birth is projected to rise from 70 years in the period 2010-2015 to 77 years in 2045-2050 and achieve 83 years in 2095-2100.

Africa seems to gain about 19 years of LE by the end of the century. Such increases are contingent on further reductions in the spread of HIV, and combating successfully other infectious as well as non-communicable diseases. Both Asia and Latin America and the Caribbean are expected to gain 13-14 years of LE at birth by 2095-2100, while Europe, Northern America and Oceania are projected to gain 10-11 years. Globally, population aged 60 or over is the fastest growing. Furthermore, as fertility declines and LE rises, the proportion of the population above a certain age rises and this phenomenon-known as population ageing- is occurring throughout the world. In 2015, there were 901 million people aged 60 or over, comprising 12% of the global population. Europe has the largest percentage of its population at ages 60 or over (24%) and the number of older persons in the world is projected to reach 1.4 billion by 2030 and 2.1 billion by 2050.

Population ageing will have a profound effect on the number of workers per retiree in various countries, as measured by the Potential Support Ratio (PSR), defined as the number of people aged 20 to 64 divided by the number of people aged 65 or over.

Currently, African countries, on average, have 12.9 people aged 20 to 64 for every person aged 65 or over, while Asian countries have PSRs of 8.0, Latin America and the Caribbean 7.6, Oceania 4.8, Northern America 4.0 and Europe 3.5. Japan, at 2.1, has the lowest PSR in the world, although seven European countries also have PSRs below 3.

By 2050, seven Asian countries, 24 European countries and four countries of Latin America and the Caribbean are expected to have PSRs below 2, underscoring the fiscal and political pressures that many countries are likely to face in the not-too-distant future in relation to their public health care.

The increase of LE is undoubtedly evident, however what is not fully understood is why there are so many differences among countries. Life span and longevity are complex variables and belong to the interaction between environmental, genetic, epigenetic and stochastic factors that can be influenced also by foods and drugs which according to the common knowledge are considered valid factors to live longer.

The aim of our study was to consider the LE as a main variable to correlate with 17 ancillary variables that are recorded as routine in most of the countries in relation to the environment, the demography and the economy. The determinants of LE were measured through the stochastic and non stochastic analysis to discover the existence of common variables.

***Corresponding author:** Umberto Cornelli, 1Loyola University School of Medicine Chicago-USA, Piazza Novelli 5, Milan-20129, Italy.

Received: January 24, 2018; **Accepted:** February 09, 2018; **Published:** February 13, 2018

Citation: Cornelli U, Recchia M, Grossi E, Belcaro G. Life Expectancy does not Depend on Classical Ecological Variables: Stochastic and Non-stochastic Analysis. GSL J Public Health Epidemiol 2018; 1:104.

Material and Methods

Source of data

All the 261 countries in the world were analyzed, considering the data of two years 2014 and 2015. Countries without some of the variables that were chosen for the analysis were excluded, and because of this the final sample was limited to 191 countries. The LE in each country was taken by the World Health Statistics Monitoring 2016 [8]. Those not reported such as for Montecarlo and San Marino were taken by the CIA World Fact book 2016 [9]. At the end all the data concerning the year 2015 were reported since those of the 2014 were superimposable.

List of the countries

The following 191 countries that have been analyzed are presented in terms of descending order for life expectancy: Monaco; Japan; San Marino; Singapore ;Andorra; Swiss; Australia; Sweden; Liechtenstein; Canada; France; Norway; Spain; Island; Nederland; New Zealand; Ireland; Germany; United Kingdom; Greece; Austria; Malta; Luxembourg; Belgium; Taiwan; South Korea; Finland; United States; Denmark; Portugal Bahrein; Chile; Qatar; Cyprus; Czech Republic; Panama; Costa Rica; Cuba; Albania; Slovenia; Dominican Republic; Kuwait; Argentina; Santa Lucia; Lebanon; United Emirates; Uruguay; Paraguay; Brunei; Slovakia; Poland; Morocco; Czech republic; Algeria; Ecuador; Sri Lanka; Bosnia and Herzegovina; Antigua Bermuda; Libya; Lithuania; Tonga; Macedonia; Georgia Brazil; Tunisia; Hungary; Mexico; Saint Kits and Nevis; Columbia; Mauritius; Maldives; China; Barbados; Oman; Salomon Islands; Saint Vincent Grenadine; Saudi Arabia; Romania; Malesia; Venezuela; Bulgaria; Seychelles; El Salvador; Thailand; Armenia; Jordan; Estonia; Grenada; Jamaica; Egypt; Latvia; Turkey; Uzbekistan; Peru; Samoa; Vietnam; Nicaragua; Vanuatu; Palau; Marshall Islands; Philippine; Micronesia; Indonesia; Belarus; Fiji Islands; Bahamas; Azerbaijan; Greenland; Guatemala; Suriname; Lebanon; Cape Verde; Iraq; Honduras; Iran; Bangladesh; Kazakhstan; Russia; Moldavia; Kirghizstan; North Korea; Turkmenistan; Ukraine; Bhutan; Mongolia; Bolivia; Belize; Syria; Guyana; India; Timor Est; Nepal; Tajikistan, Pakistan; Papua New Guinea; Burma; Tuvalu; Ghana; Kiribati; Madagascar; Yemen; Gambia; SãoTomè and Principe; Togo; Cambodia; Kenia; Eritrea; Laos; Equatorial Guinea; Comoros; Sudan ; Haiti; Djibouti; Mauritania; Tanzania; Benin; Senegal; Ethiopia; Malawi; Guinea; Burundi; Republic of Congo; Liberia; Ivory Coast; Sierra Leon; Cameroon; Democratic Republic of Congo; Zimbabwe; Angola; Mali; Burkina Faso; Niger; Uganda; Botswana; Lesotho; Nigeria; Mozambique; Gabon; Namibia; Zambia; Somalia; Central African Republic; Swaziland; Afghanistan; Guinea-Bissau; South Africa; Chad.

The variables

The following 17 variables were considered taken by CIA World Factbook 2016 [9]:

Population: as number of inhabitants.

Population density: In terms number of subjects/km^2.

Urbpop: Rate (%) of urban population in comparison to the total population.

GDP/inhab: Gross Domestic Product/inhabitants or the total values/inhabitants of goods and final services related to economical activities, capital investments.

Unempl: Unemployment rate (%) of people looking for a job in relation to the labor force.

GDP 1: GPD rate (%) of the Gross Domestic Product in relation to primary industry bound to agriculture, forests, livestock, fishing.

GDP 2: GDP rate (%) of the Gross Domestic Product in relation to industry, mining and construction industry.

GDP 3: GDP rate (%) of the Gross Domestic Product in relation to commerce, transportation, communication, tourism, insurance.

GDP 2+3: Sum of the rate (percentage) related to GDP 2+GDP 3.

Education: Rate (%) of the investments in public and private instruction in relation to GDP.

HB: Number of hospital beds/1000 inhabitants.

Forests: As rate (%) of surface covered by forests.

PM: Particulate matter ($PM_{2.5}$ and PM_{10}) in mcg/m^3 measured in cities with > 100,000 inhabitants.

Cars: Number of cars/1000 inhabitants.

Cell: Number of mobiles/1000 inhabitants.

Internet: Number of people connected to internet/1000 inhabitants.

Forests Kmq^2: Square kilometers of forest/1000 inhabitants.

LE was considered as the main variable to compare with all the others using two approaches: stochastic and non stochastic analysis. The first was based on the simple correlations and on the Main Component Analysis (MCA with rotation Verimax), to explain the correlation among the variables [10,11]. This evaluation was followed by the Factorial Analysis that used "factor scores" to transform the cluster of observations in more simple structures or Factors [12,13]. The Segmentation analysis was also used [13] on the base of CHAID (Chi square Automatic Interaction Detection).

For the non stochastic analysis The Artificial Neural Networks Analysis (ANN) was applied to establishing the hierarchy of variables within a specific set. The Auto Contractive Map (Auto-CM) was used with this aim [14-17].

Data analysis

The methods employed for the analysis were the following:

Correlation matrix among all the variables, considering as limit of significance the value of $r \geq 0.6$. The Main Component Analysis (MCA) was applied to explain the correlation among the variables as cause of "non observable" or hidden factors [10,11].

This was followed by the Factorial Analysis that used the "factor scores" to transform the cluster of observations in more simple structures or Factors [12,13]. Followed the analysis to measure if the LE was correlated to the mutation of any single factor. After these three steps, only the factor scores capable to explain a consistent percentage of variability (> 80%) will be considered, with the aim of isolating two orthogonal factors, each consisting of a combination of variables, and project the need to define a target variable 191 countries on these two new axis.

The segmentation analysis is based on the construction of a decisional tree based on Chi Square Automatic Interaction Detection (CHAIDs). The method needs at first to identify the target variable that in the current case was the LE that ranges between 45 to 85 years. LE was divided into 4 groups of similar dimension (Between 45 to 50 countries) and CHAIDs allows the identification of "optimal splits" that consent to maximize the differences among the groups [13]. In a first step the sample was divided into two clusters of countries according to the most solid "predictor". The best predictor is chosen among the variables that contain 100% of the variance and were isolated with the MCA (In these case 7 variables). The cluster isolated by the first predictor can be further divided on the base of other predictors. The analysis gives back the probability of a given LE based upon the dimension of the predictors.

All the stochastic calculations were carried out with JMP 12 SAS (Sas Institute inc. e XLSTAT). The Artificial Neural Networks Analysis

(ANN) was also applied because complex mathematical networks can help in establishing the hierarchy of variables within a specific set. The Auto Contractive Map (Auto-CM) was used with this aim.

Auto-CM system is a fourth generation unsupervised ANN able to highlight the natural links among variables with a graph based on minimum spanning tree theory, where distances among variables reflect the weights of the ANN after successful training phase [14-16]. The Auto-CM system finds a square matrix of "similarities" (Weights mathematically speaking) among the variables of a given dataset. Once the Auto-CM weights matrix is obtained it is then filtered by a Minimum Spanning Tree algorithm (MST) [17]. MST shows the best way to connect the variables in a tree and the shortest possible combination allows to present data in a simplified graph. The assumption is that since all biological systems tend naturally to the minimal energetic states, this graph express the fundamental biological information of the system. The ultimate goal of this model is to discover hidden trends and associations among variables in non linear associations in connection with Hubs. Hubs can be defined as variables with the maximum amount of connections in the map.

In simple words Auto-CM "spatializes" the correlation among different variables (Closeness) and converts it into a compelling graph that identifies only the relevant associations and organizes them into a coherent picture, building a complex global picture of the whole pattern of variation. In the Auto-CM the discrimination between "high and low values" allows a more consistent representation of the Hubs and MST. This is important because in non linear systems, the position of high and low values of a given variable is not necessarily symmetric.

Because of this the 18 continuous input variables were split in 36 input variables: high and low. To do this a particular complement transformation, by scaling original variables values from 0 to 1 and creating a complement variable by subtracting the scaled value from 1 [17]. In this way the projection of the original variables will tended to show high values while the complement transformation tended to show low values of the original variables. In the map these two different forms were named simply "high and low".

Results

Here are summarized the data of 2015 because those of 2104 were practically superimposable to 2015. The correlation matrix is reported in Table 1 and is limited to variables with r value ≥ 0.6. According to the r ≥ 0.6 cut-off, five variables only were found correlated to LE, namely: GDP 1; GDP 3; Cars; GDP 2+3; Internet. This last showed the highest r value (r=0.7697) Vs LE. In the light of this simple analysis, two other aspects were considered. The first was to isolate those variables that were considered "inhert" because they were not correlated with any other variable. These were 10: Population, population density, unemployment rate, GDP 2; Instr/GDP, HB, Forests, PM, Cell, Forests Kmq^2. The second was to determine how the remaining variables were mutually crossing, as reported in Table 1.

What can be drown is that GDP in all the aspects is either positively (GDP2 and GDP 3) or negatively (GDP 1) correlated with LE and the other highly interconnected variables are Internet, Cars and GDP 3 or GDP 2+3. GDPs are all correlates among them either in positive or negative way.

The proposed exploratory Factor Analysis (With rotation Verimax) that was following allowed to explain the correlation between different variables in terms of a reduced number of factors unobservable or latent. This analysis eliminate 10 of the 17 variables because their information was already contained in the remaining 8. The following variables were excluded: Population and population density, unemployment, GPD 2 and GPD 3, HB, Instr/GDP, forests, forest Kmq^2 and PM.

The seven remaining variables instead were capturing 100% of the variance, namely: urb pop; GPD/inhab; GDP 1; Cars; Cell; GDP2+GDP 3; Internet (Figure 1).

The 7 Numbers were explaining 100 of the variance and two of them (see Numbers 1 and 2 in Figure 1) capture as much as 80% of the variance. The Number 1 contains 66.83% of the variance and represents the combination of the following four variables: urban pop, GDP 2+GDP 3, Cell, and GDP1.

The Number 2 contains 12.83% of the variance and is represented by the combination of the following three variables: GDP/inhab, Cars and Cell. These two Numbers were called Factor 1 and Factor 2 respectively. The correlations with LE of each variable contained in the Factors was calculated and it was possible to determine the importance of every single variable by the "greatness" of the relative coefficient (Figure 2).

By their "greatness" one may notice that: Factor 1 decreases with

Eigenvalues

Number	Eigenvalue	Percent	20 40 60 80	Cum Percent
1	4.6781	66.830		66.830
2	0.8982	12.831		79.661
3	0.5857	8.367		88.028
4	0.4151	5.930		93.958
5	0.2499	3.570		97.528
6	0.1708	2.440		99.968
7	0.0022	0.032		100.000

1=urb pop; 2=GDP/inhab; 3=GDP1; 4=Cars; 5=Cell; 6=GDP 2+GDP 3; 7=Internet

Figure 1: Weight of the 7 Factors that have been isolated from the initial 17 variables. Eigen values, percentages and cumulative percentages

Rotated Factor Loading

	Factor 1	Factor 2
% urb pop	0.682752	0.423625
GDP/inhab	0.162872	0.912412
GDP 1	-0.831122	-0.409403
Cars/1000	0.334113	0.859529
Cell/1000	0.803700	0.062570
Inter/1000	0.603184	0.697391
GDP 2+3	0.833336	0.410305

Figure 2: Correlations coefficients of Factor 1 and the Factor 2 Vs the original variables composing the factors. Analysis with rotation Verimax.

Correlations	LE	GDP 1	GDP 3	Cars/1000	Inter/1000	GDP 2+3
LE	1.0000	-0.6734	0.667	0.6521	o.7697	0.6751
GDP 1	-0.6734	1.0000	-0.9258	-0.6029	-0.7554	-0.99n
GDP 3	0.667	-0.9258	1.0000	0.5959	0.7274	0.9319
CARS/1000	0.6521	-0.6029	0.5959	1.0000	0.7839	0.6004
INTER/1000	0.7697	-0.7554	0.7274	0.7839	1.0000	0.7582
GDP 2+3	0.6751	-0.9977	0.9319	0.6004	0.7582	1.0000

Table 1: Variables that is significantly correlate with r ≥ 0.6.

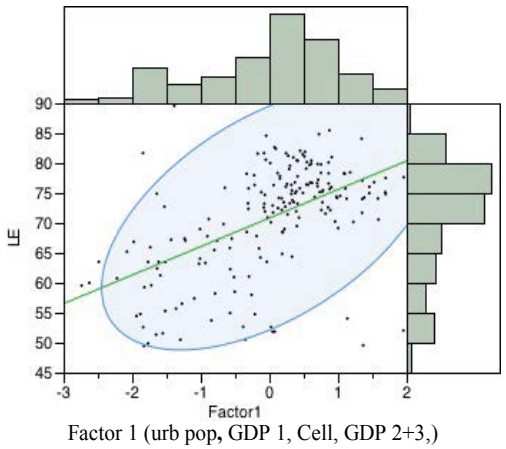
GDP 1 (negative coefficient) and increases with increasing GDP 2+3, urban pop and Cell, whereas GDP/inhab and Cars were almost inconsistent; Factor 2 values increase mainly with the growth of GDP/inhab, Cars, Internet. In this last case pop urb, GDP 1, Cell, and GDP 2+GDP 3 were shown less consistent. Internet correlation is present in both Factors but its weight prevail in Factor 2. Only values > 0.6 were considered consistent.

The "greatness" allowed identifying within the two Factors those variables that were better defining the Factor's characteristics. For Factor 1 they were represented by urban pop, GDP 1, Cell and GDP 2+GDP 3, whereas for Factor 2 they were represented by GDP/inhab, Cars and Internet.

A regressive models comparing LE respectively Vs Factor 1 and Factor 2 was calculated and the results are represented in the following Figure 3 and 4.

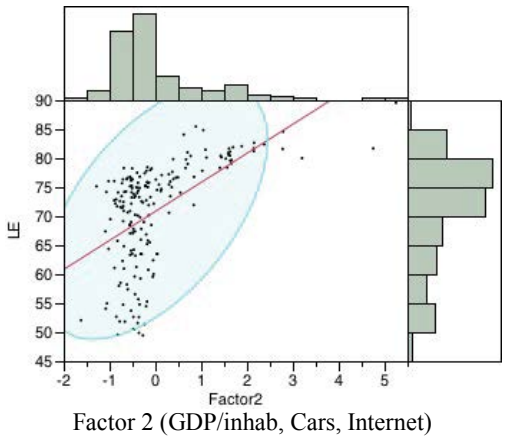
The regression line marked in the Figure 3 represents the growth of LE as a function of Factor 1. For what concern Factor 2, represented by GDP/inhab, Cars and Internet, the regression line and the points disposition are reported in Figure 4.

The last step of the analysis was the connection of the two Factors (Figure 5). This is possible because they explain 80% of the variance and may form a "complex" variable to be used for the final correlation. The results were represented in ascending order in Figure 5, consisting of the new regression LE Vs Factor 1+Factor 2. The correlation is statistically significant: $r=0.77$ $p < 0.001$).



Factor 1 (urb pop, GDP 1, Cell, GDP 2+3,)

Figure 3: LE of 191 countries in relation to Factor 1.



Factor 2 (GDP/inhab, Cars, Internet)

Figure 4: LE of 190 countries in relation to Factor 2.

For the Segmentation analysis four classes of LE were considered: 49-64 years, 65-73 years, 73-77 year and finally > 77 years composed respectively by 47, 44, 53 and 47 countries (Figure 6).

The CHAID (Chi square Automatic Interaction Detection) was used to determine the best predictors among the seven variables isolated by the Factor analysis with rotation Verimax (see Figure 2).

The first predictor hierarchically most important was Internet presenting two cut-off: > 436 and < 436 (/1000 people). The first collect 90% of countries with LE > 73 years, whereas the second only the 13%. The second predictor was GDP/inhab > 21633 and < 21633 that were concentrating respectively 95% and 86% of the countries with LE > 73. The third predictor was Car > 360 and < 360. In this last case the first allowed a concentration of 82% of the cases with LE > 73 and the second the 100%. In other terms, a country characterized by internet > 436, GDP > 21633 and Cars > 360 was offering a LE > 77 years with 98% of probability. Without taking into consideration the three predictors, the probability of LE > 77 years was 52% only (52/191 countries).

For the non stochastic analysis a type of ANN was chosen where the variables were split into "low" and "high values" generating a matrix composed by 36 variables. In this case the LE was considered as one of the variables. This method allows a more precise definition of the MSP (Minimum Spanning Tree). The semantic connectivity map obtained with the analysis was forming the MST as reported in Figure 7.

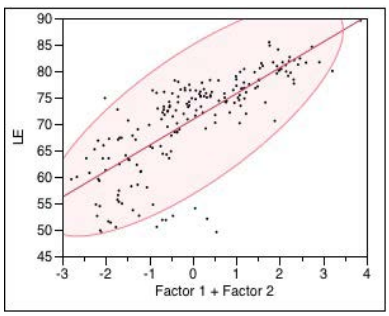
The first element to note was that the MST LE was set into two different poles where "high" and "low" are very well separated. The dotted line of the figure indicates the main cut of the MST to isolate respectively LE low and high. A numerical value (In red on the Figure 7) was applied to each edge of the graph and was proportional to the strength of the connection between the two variables. This strength ranges from 0=No connection and 1=Maximally connected. Most of the variables "high" are in the same cluster together with high LE (Upper part of the Figure 7) with the exception of GDP 1 which is low, whereas all the variables "low" are in the opposite pole together with LE low (Lower part of the Figure 7), with the exception of unemployment and GDP 1 that are high.

Variables such as forests, forest Kmq^2, PM, cell, pop density, education low and HB low were very far from the LE poles (High and low) in a position underlining their limited importance. It is possible to note how high LE values node at first was directly connected with high GDP 3 node, and with high internet node. Low LE values node was bound to the low internet and to high unemployment nodes.

Internet high coordinates car high, HB high, GDP/inhab high and population density high, while internet low coordinates cars low, urban population low, GDP 3 and GDP 2+3 low and GDP 1 high.

Discussion

This is the first time that LE is compared with many variables that are usually employed to define the condition/performance of a country.



Factor 1 (urb pop, GDP 1, Cell, GDP 2+3,)+Factor 2 (GDP/inhab, Cars, Internet)

Figure 5: LE in relation to the sum of Factor 1 and Factor 2.

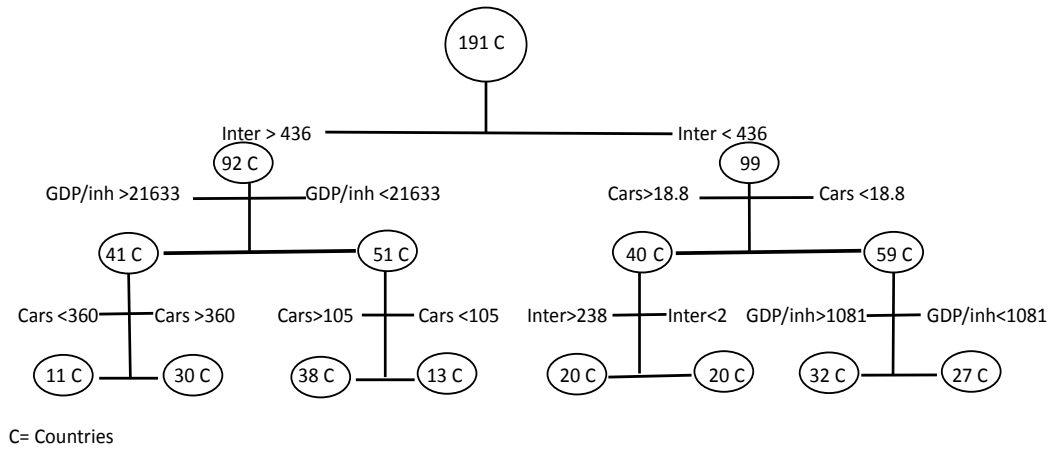


Figure 6: Segmentation analysis with CHAID on the base of the 7 Factors isolated by the MCA.

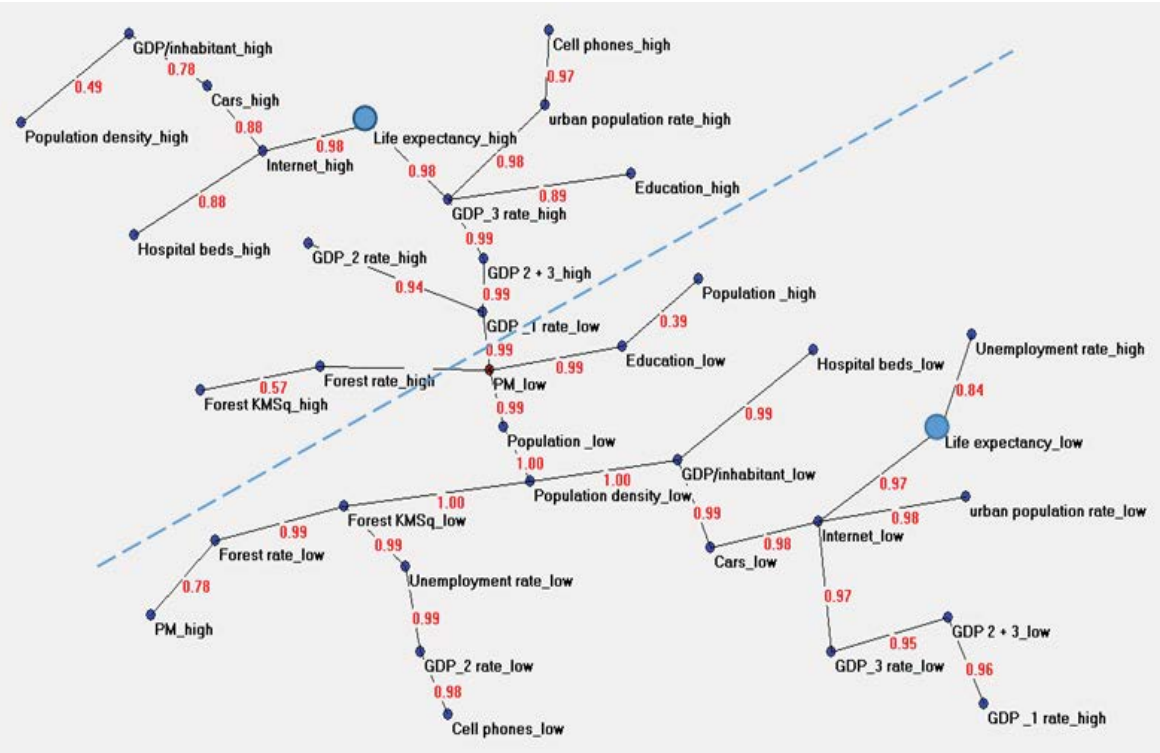


Figure 7: Auto Contractive MAP of the variables divided into high and low values for each variable.

The weakness of this study belongs to some points that have to be underlined.

They are the precision/quality of some variables, such as the PM, unemployment, and urban population. PM in fact was not taken in a standardized way in terms of the period of the year, number of measure and cities of a given country. Furthermore, they belong to different origin, ranging from inert matter to very toxic substances. In other terms the PM of cities close to the desert cannot be compared to PM of the industrial areas.

The unemployment can fluctuate in some country during the year, and the urban population may be modified quite rapidly by the massive immigrations. The urban population can change due to the emigration, since the war exploding in some countries can suddenly and dramatically modify the characteristics of the closest areas.

However, in general most of the variables can be defined as solid, and among the 191 countries the vast majority was stable and free from dramatic events. The first observation emerging from the analysis concerns those variables that were excluded as determinant for the LE and confirmed in all the type of analysis.

The classical and common concerns like PM, the Kmq^2 of forests or the forests dimension in a given country, the Instr/GDP, were excluded as direct factors that may have impact on LE. The number of HB was found important only in the ANN limited to the high LE only, as it was for the unemployment involved only in case of low LE.

For the extent of the forests, the HB and the Instr/GDP one may only take the data as they are, and the reason why they are not emerging as determinants in above the scope of this research. The PM instead can be matter of discussion due to same particular aspect. PM is a complex

mixture of chemical components that should be considered together with many gas such as Methane (CH_4), Ozone (O_3), Carbon monoxide (CO), Sulfate (SO_3), Nitrogen dioxide (NO_2) aerosols and all the possible widespread air pollutants, present wherever people live.

These particles are able to penetrate deeply into the respiratory tract and therefore constitute a risk for health by increasing mortality from respiratory infections and diseases, lung cancer, and selected cardiovascular diseases. The WHO estimated in 2000 that the exposure to PM caused 800,000 deaths and 6.4 million YLDs and that the developing countries accounted for two third of this burden [18].

In general the WHO stated that there is no evidence of a safe level of exposure to PM or a threshold below which no adverse health effects occur, and globally > 30% of the population lived in areas exceeding the WHO level Target of 35 mcg/m^3.

The data recorded for this study represent an average of the cities where the monitoring stations were available. In order to present air quality largely representative for human exposure, measurements of residential areas, commercial and mixed areas were used. Stations characterized as particular "hot spots" or exclusively industrial areas were not included.

Furthermore, in some of countries particles < PM_{10} was largely based on estimates [19]. There are several studies conducted in different part of the world showing the negative effect of the PM on health [20-22] but still there is need for further research to define the long term toxicity [23] and whether some components and sources of PM may be more toxic than others [24]. The indoor air pollution also is something that should be considered since it is causing apparently 3.7 million of death [25].

In the Update report of 2106 [26] a comparison was done to determine the trend of PM in the world between 2008 and 2013, ending up with an estimation of 5% increase, despite some fluctuation within the macro-regions that were analyzed. In the same period the LE was increased also, practically in every of the 191 countries that have been considered in this study, no matter if the PM was increasing or not.

This indicate that more precise measures should be taken for PM, because in the present scenario it seems that they have a positive effect on the LE. Instead of the "classical threats", other variables have been isolated as determinants in the present analysis.

Factor 1: Which was bound to urban population, Cell, GDP2+GDP 3, GDP1.

Factor 2: Which was bound to GDP/inhab, Cars, Internet.

These factors should to be considered as "colors" that are made by the combination of fundamental colors which may end up with different shade/aspect according to the quantity of each base color. A more precise connection between variables, without any overlapping was emerging from the ANN. This analysis was confirming the importance of all the same factors in terms of nodes connected to the LE nodes (Low and high) with the exception of cell.

Considering in detail some of the variables, the following observation can be made. The urban concentration (urb pop) has a positive value for LE and seems far from the threat of the biodiversity [27].

This finding may justify why people tend to concentrate and the urban development is expected to increase nearly to 5 billion by 2030 for the 55% in Asia, particularly China and India [27]. The urban concentration is one aspect of the more complex and dramatic problem of emigration from the people living in the "South of the World" toward the "North of the World". Most of the humans seems to be attracted by the concentration, which means living together. This allows LE becoming longer even though is bringing as collateral effect the increase of YLDs. Aging is associated with a phenomenon of decline and YLD represents this aspect which is rapidly increasing, since in 1990 it was 537.6 million years and jumped up to 764.8 million in 2013 [6]. Going to the leading diseases that characterize the YLDs one may discover that low back pain and depressive disorders are in the top 10 diseases in every country. Should the people afford all this in the hope to live longer? Good point, that would need a discussion much above the scope of this research.

A very short answer can be found in the Psalm 89 (or 90 in some case) "*Our days may come to seventy years or eighty if our strength endures; yet the best of them are but troubles and sorrow for they quickly pass and fly away*". Like to say humans have to pay a bill if they want to live longer. Another variable to be discussed are Cell that was emerging as important in the MCA. Cellular phones are known to spread more than other modern technology since they are perceived to improve the livelihood of the poorest people in developing countries. In 2013 the UN Deputy Secretary-General Jan Eliason declared that "6 billion of people have mobile phones and only 4.4 billion have access to toilets or latrines [28].

The total number of cell phones is now exceeding the world population [29] since they are regarded not only as instruments to communicate but also to make business such as for microenterprise, in particular where internet is not available [30], offering work to people with little education and few resources. In conclusion, it seems that the direct correlation between Cell and LE is consistent.

The variables related to the GDPs were found important determinants in all the analysis. They have a dichotomic behavior, since the GDP 1 in terms of LE behaves as a negative variable whereas GDP 2 and 3 together with the urban concentration have positive effects. The increase of GDP 2+GDP 3 corresponds to the decrease of GDP1, which is bound to the "primary" way to collect resources such as with agriculture, fishing, live stock. More resources and in more rapid way will be available depending upon the GDP 2 and GDP 3.

The variables GDP/inhab, Cars and Internet should be also focused. GDP/inhab and Cars mirror respectively the prosperity and the possibility of the independent capacity to move within the territory.

Internet instead can be considered one of the easiest methods to communicate and collect information, carrying an extensive range of opportunities and services enabling or accelerating new forms of personal interactions. Its use grew rapidly in all world, such that as the average penetration in 2016 reached about 3.68 billion of people (50% of the total world population). The penetration rate goes from 28.7 of Africa, to 45.6 in Asia, 73.9 in Europe and 89.0% in North America, almost parallel to the LE growth.

The possibility of online shopping through internet give access to small businesses and entrepreneurs. In short, internet is a mixed variable containing both aspects of communication and business, which are directly bound to GPD. The importance of the variables that composed the Factor 2 (GDPs, Cars and Internet) was confirmed by the Segmentation Analysis indicating that, people living in countries characterized by Internet > 436 and GRD/inhab > 21633 have a probability of 89% to reach a LE > 77 years. The probability increases up to 98% in those countries with Cars > 360. The ANN was detecting more information than the stochastic analysis and found more complex relations among variables. The analysis was confirming that Internet is the most important determinant since it is the node more directly connected with LE nodes (High or low). When LE is low, the other node indirectly connected is unemployment, whereas when LE is high the closest node is GDP 3. The other variables that has been isolated with the previous analysis (GDP/inhab, Cars, urban population, GDP 1, GDP 2, GDP 2+ GDP3) seems to have less impact, but still are "indirectly" connected nodes having some influence.

In conclusion the two different analysis, stochastic and non

stochastic, are indicating similar variables that are determinant for LE: GDPs, Internet and Cars are emerging in every analysis. There is no simple explanation of why the "classical threats" seems not involved.

Apparently, what can be depicted as the heaven on earth, represented by fresh air, forest, lucky solitude, health, knowledge, resources coming from the environment, and finally "unemployment" are not bound to LE. Much more important are resources deriving from a modern way of leaving, no matter if they are carrying also some troubles.

At the end a more holistic point of view should be considered: Probably Adam and Eve were kicked out from the Paradise not because of an apple, but to give them the chance to live longer.

Acknowledgments

Funding: No funding was requested for this study.

Author Contributions: UC and GB conceived the study and collected all the data; UC and MR did the calculations for the stochastic analysis; EG performed the non-stochastic neural network analysis (ANN); UC and GB wrote the paper.

Competing Interests: there are no competing interests

References

1. Juvin H. The coming of the body. London UK: Verso, 2010.
2. Lutz W, Sanderson W, Scherbov S. The coming acceleration of global population ageing. Nature. 2008; 7: 716-719.
3. Gerland P, Raftery AE, Ševciková H, et al. World population stabilization unlikely this century. Science. 2014; 364: 234-237.
4. United Nation World Population Prospect the 2015 revision.
5. Hugues BB, Kuhn R, Peterson CM, et al. Projection of global health outcomes from 2005 to 2060 using the International Future integrated forecasting model. Bull World Heath Organ. 2011; 89: 478-486.
6. GDB 2015 Mortality and Causes of Death Collaborators. Global, regional, and national LE, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016; 388: 1549-1544.
7. Laaser U, Brandt H. Global Health in the 21st century. Global Health Action. 2014; 7: 23694.
8. WHO-Global Health Observatory (GHO). 2016.
9. The CIA World Factbook, Central Intelligence Agency. 2016.
10. Dillon W, Goldstein M. Multivariate Analysis-Methods and Applications. John Wiley, NY. 1984.
11. Cureton EE, D'Agostino RB. Factor Analysis-An Applied Approach, Lawrence Erlbaum Associates, Hillsdale, New Jersey. 1983.
12. Jae-On K, Mueller CW. Introduction to factor Analysis. Beverly Hills: Sage. 1978.
13. Fabbris L. Statistica Multivariata McGraw-Hill Libri Italia. 1997.
14. Buscema M, GrossiE.. The semantic connectivity map: an adapting self-organizing knowledge discovery method in data bases. Experience in gastro-oesophageal reflux disease. Int J Data Min Bioinform. 2008; 2: 362-404.
15. Buscema MBM, Sacco PL. Auto-contractive Maps, the H Function, and the Maximally Regular Graph (MRG): A New Methodology for Data Mining Applications of Mathematics in Models, Artificial Neural Networks and Arts, eds. Dordrecht (Springer), 2010; 227-275.
16. Buscema M, Grossi E, Snowdon D, Antuono P. Auto-Contractive Maps: An Artificial Adaptive System for Data Mining. An Application to Alzheimer Disease. Curr Alzheimer Res. 2008; 5: 481-498.
17. Gironi M, Saresella M, Rovaris M, et al. A novel data mining system points out hidden relationships between immunological markers in multiple sclerosis. Immun Ageing. 2010; 10: 1.
18. Cohen AJ, Anderson HR, Ostro B, et al. Comparative of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. 1st Edi., World Health Organization Geneva. 2004; 2: 1353-1453.
19. Brauer M, Amann M, Burnett RT, et al. Exposure assessment for estimation of the global burden of disease attributable to outdoor pollution. Environ Sci Technol. 2012; 46: 652-660.
20. Lippmann M, Chen LC, Gordon T, et al. National Particle Component Toxicity (NPACT) initiative: integrated epidemiologic and toxicologic studies of the health effects of particulate matter component. Res Rep Health Eff Inst. 2013; 177: 5-13.
21. Crouse DL, Peters PA, Brook JR, et al. Ambient $PM_{2.5}$, O_3, and NO_2 exposures and associations with Mortality over 16 years of follow-up in the Canadian Census Health and Environment cohort (CanChec). Environ Health Perspect. 2015; 123: 1180-1186.
22. Baccarelli AA, Hales N, Burnett RT, et al. Particulate air pollution, exceptional aging, and rate of centenarians: a nationwide analysis of the United States, 1980- 2010. Environ Health Perspect. 2016; 124: 1744-1750.
23. Henschel S, Atkinson R, Zeka A, et al. Air pollution and their impact on public health. Int J Public Health. 2012; 57: 757-768.
24. Adams K, Greenbaum DS, Shaikh R, et al. Particulate matter components, sources, and health: systematic approaches to testing effects. J Air Waste Assoc. 2015; 65: 544-558.
25. WHO. World Health Assembly closes, passing resolution on air pollution and epilepsy, Environ. 2016.
26. WHO-WHO's Urban Ambient Air Pollution database-Update. 2016.
27. Seto K, Güneralp B, Hutyra LR. Global forecast of urban expansion to 2030 and direct impacts on biodiversity. PNAS. 2012; 109: 16083-16088.
28. UN News Centre. 2013.
29. DH Deccan Herald. 2017.
30. Lynn, Jonathan. Mobile phones help lift poor out of the poverty: UN study News Centre March October 14 2010.