Influence of Atmospheric Parameters on Human Mortality data at Different Geographical Levels

Joseph Sánchez-Balseca[a*], Agustí Pérez-Foguet[a]

[a]Research group on Engineering Sciences and Global Development (EScGD),

Civil and Environmental Engineering Department,

Universitat Politècnica de Catalunya – BarcelonaTech (UPC), Spain.

*Corresponding author: (+34) 690 132 602. Jordi Girona 31, UPC Campus Nord 08034, Barcelona, Spain.

E-mail address: joseph.sanchez@upc.edu, agusti.perez@upc.edu

HIGHLIGHTS

- The meteorological and air pollutant covariates allow identify extreme events.
- The proposed approach improves the fitting mortality data with great variability.
- Adding atmospheric parameters improves the basic demographic model.
- Human mortality data modelling should consider its over-dispersion.

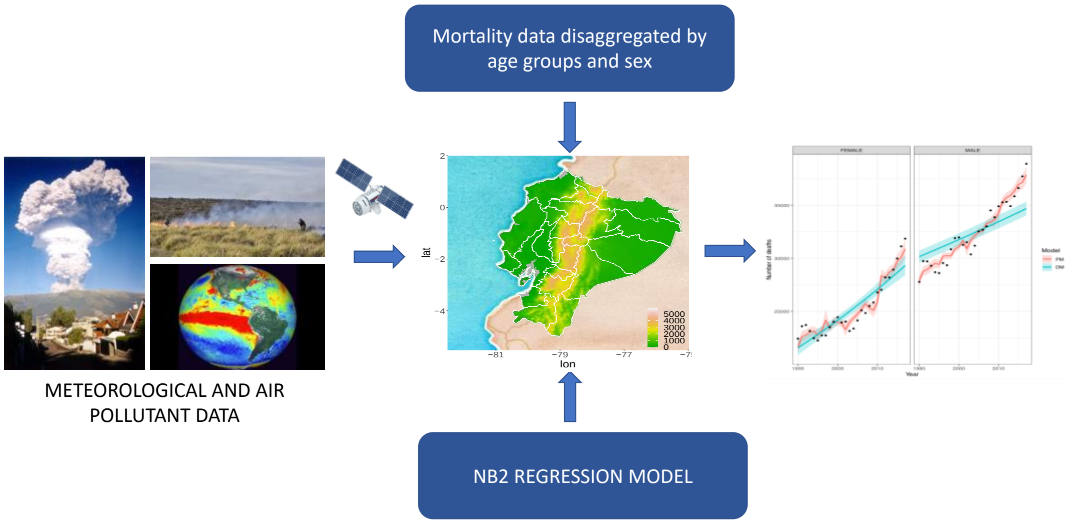
ABSTRACT

The human mortality data under a demographic approach often are modeled as a function of time. This approach does not present an adequate fit model for the number of deaths with great variability. For this reason, it is necessary for additional information (social, economic, and environmental) that complements and improves demographic modelling. This article evaluated the association between human mortality data (segregated by age groups and sex) with meteorological and air pollutant covariates at three geographical levels: country, macro-climate regions, and county. The modelling was based on a generalized linear modeling framework, and it takes into account the common

characteristic of overdispersion in human mortality data through the application of negative binomial distribution. The proposed approach improved the dynamic behavior of the Farrington-like model (basic demographic model), and it took into account the extreme meteorological and natural air pollution events. The proposed modelling worked well in cases where the amount of data was scarce.

Keywords: Air quality, Environmental statistics, ENSO, Volcanic, Negative Binomial, Human health

GRAPHICAL ABSTRACT



1. Introduction

In demographic studies is common to assess human mortality (count of deaths) across a period of time using non-linear regression models, where the time as covariate has a fundamental role (Holford, 1983; Koissi et al., 2006; Viner et al., 2014; Bozikas and Pitselis, 2018; Czaja et al., 2020). For instance, a widely used method is the Lee-Carter (LC) model, which has a high dependency to the time (Siu-Hang and Wai-Sum, 2013;

Neves et al., 2017). The Serfling regression model used to estimate the number of influenza-attributable deaths consists of applying a cyclic regression to the time series of mortality rates observed. This model uses two variables to adjust fluctuation (sine and cosine) and the linear and squared time term. Another example commonly used in the demographic field is the Farrington-like model, which is a basic regression model related to time; in this article, we referred to it as a common or basic demographic approach (Farrington et al., 1996; López-Cuadrado et al., 2012; Zhou et al., 2017). Furthermore, demographic models beyond the temporal characteristic may also have spatial characteristics. For instance, the Spatiotemporal Epidemiological Modeler proposed and used by Edlund et al. (2011) in Israel.

Socio-economic, biological and environmental factors could improve the mortality data modelling considering the different studies and territorial context. A better knowledge about the influence of environmental conditions over the human mortality is crucial to understand the resulting socio-economic impacts, especially if it is a region susceptible to climate change due to the drastic changes in the meteorological variables (Martínez et al, 2009; Rao et al., 2013; Nunes et al., 2016; Lucas et al., 2019; Alahmad et al., 2020, Tsekeri et al., 2020).

Globally, the air quality is being continuous deteriorated (Matus, 2012; Chen at al., 2017; Zhang and Zhou, 2020; Spezzano, 2021). Several statistical epidemiological studies have associated the air pollution with human mortality (Armstrong, 2006; Barnett et al., 2010, Tran et al., 2018; Saini and Sharma, 2020). Air pollution cause about seven million premature deaths each year (Neira and Prüss-Ustün, 2016).

It is common to use models with Poisson distribution in studies related to human mortality and atmospheric conditions (Liddle, 2011; Cutter, 2017; Liang et al., 2018; Guo et al., 2019; Wang et al., 2020; Hyun et al., 2020). However, it has a restrictive assumption that the variance is equal to the mean (Hilbe, 2007; Tsekeri, 2020). Count data in human mortality are often “overdispersed”, i.e, the mortality data exhibit more variation than given by the mean (Currie and Djeundje, 2010, Cruz et al., 2020). A way to deal with overdispersion for counts is to use a generalized linear model framework (GLM), where an approach few used and with satisfactory results in environmental studies is the negative binomial model (NB2, based on the Poisson-gamma mixture distribution) (Ver-Hoef and Boveng, 2007; Li et al., 2009; Cameron and Trivedi, 2013; Hilbe, 2014; Alahmad et al., 2020; Muche et al., 2020; Tsekeri, 2020).

This article proposes to evaluate at different geographical levels (country, macro-climates regions, and counties) the association between human mortality data disaggregated by sex and age, with meteorological and air pollution data, using negative binomial regression models (NB2). The remaining of this article provides the site description, datasets used, a brief background on statistical tools (Farrington-like model and GLM with negative binomial approach), and methodology (Section 2), the results (Section 3), the discussion (Section 4), and the principal conclusions (Sections 5).

2. Data and methodology

2.1. Site description

Ecuador is a tropical country located in South America (Figure 1a), it boasts an extraordinary array of geographical systems that range from high altitude glaciers (Andean Region) to tropical rainforest in the Amazon region (east) upper tributaries to dry tropical forest on the Pacific coast (northwest, Coast Region) (Figure 1b). Ecuador is crossed north to south by the Andes Highlands creating a highly variable topography and micro-climates. Given its geographical location and rugged topography, Ecuador is a highly vulnerable country to impacts of climate change. A clear example is El Niño Southern Oscillation, ENSO (Ministry of the Environment of Ecuador, 2000).

Ecuador is one of the countries in America with the highest volcanic activity. Its situation within the Pacific Ring of Fire determines that it is in an area of high seismic, volcanic and landslide risk, due to the presence of more than 75% of the 850 most active volcanoes in the world (Pan American Health Organization, 2005).

The administrative counties are distributed by macroclimate region, as is showed in the Figure 1c. The Coast Region is conformed by Esmeraldas (8), Manabí (14), Guayas (10), Los Ríos (13), El Oro (7), and Santa Elena (20); the Andean Region by Carchi (3), Imbabura (11), Pichincha (19), Cotopaxi (6), Tungurahua (23), Bolivar (2), Chimborazo (5), Cañar (4), Azuay (1), Loja, Santo Domingo de los Tsáchilas (21); and the Amazon Region by Sucumbíos (22), Napo (16), Pastaza (18), Orellana (17), Morona Santiago (15), and Zamora Chinchipe (24).

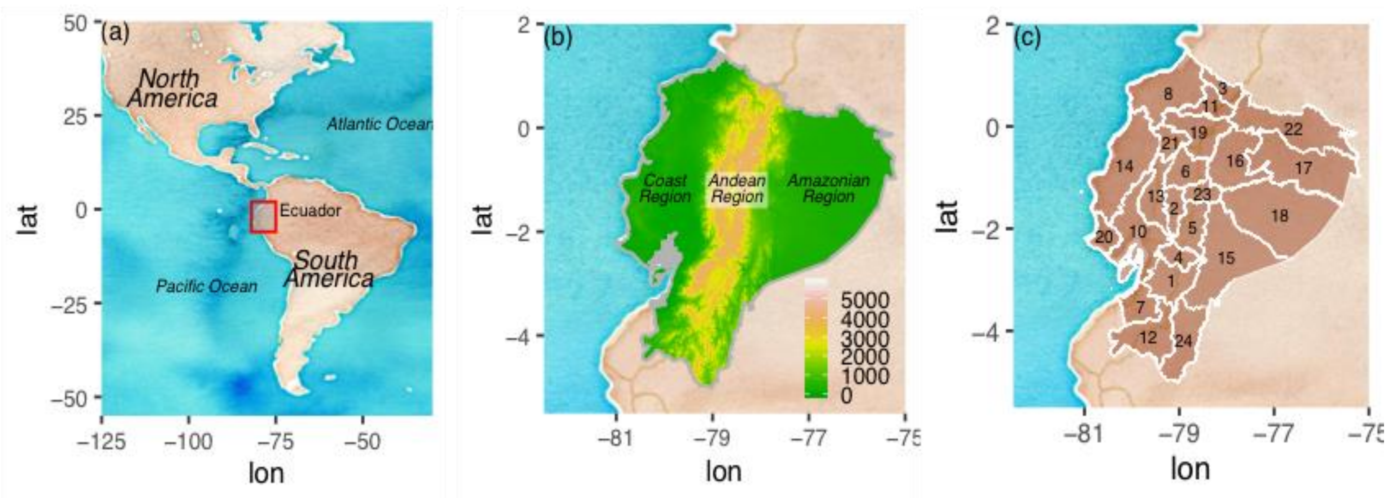


Figure 1. Study area. (a) Ecuador within South America, (b) three main macroclimate regions in Ecuador and altitude, and (c) Distribution of Ecuador by counties.

2.2. Data

2.2.1. Mortality data

The annually data of mortality by sex, age, and county level was extracted from the statistical register of general deaths, which are under the supervision of the National Institute of Statistics and Censures of Ecuador (INEC, acronym in Spanish). The data was collected during twenty-eight years (1990-2017).

2.2.2. Meteorological and air pollution data

The meteorological covariates used in this paper were air temperature (*K*), pressure (*Pa*), specific humidity ($kg \cdot kg^{-1}$), and precipitation (*mm*). The meteorological data was monthly compiled from the meteorological assimilation systems based on satellite data. The results presented in this article are derived from three data products: (1) air temperature (MATMNXSLV), it had a spatial resolution of 0.5° × 0.667° lat-lon; (2) pressure and specific humidity (FLDAS_NOAH01_C_GL_M), both had a spatial resolution of 0.1° × 0.1° lat-lon; (3) precipitation (GPCPMON), it had a spatial resolution of 0.5° × 0.5° lat-lon.

The air pollution covariates used in this paper were carbon monoxide (CO, *ppbv*), sulphur dioxide (SO_2, $kg \cdot m^{-3}$), and particulate matter with aerodynamic diameter less than 2.5 μm ($PM_{2.5}$, $kg \cdot m^{-3}$). The air pollution data was monthly compiled from the meteorological assimilation system based on satellite data. The Modern- Era Retrospective analysis for Research and Applications version 2 (MERRA-2). MERRA-2 published many analysis products used in air quality modelling (Kuo, 2017; Qin et al., 2018). The results presented in this article are derived from three data products: (1) carbon monoxide (M2TMNXCHM), (2) sulphur dioxide (M2TMNXAER), (3) particulate matter with aerodynamic diameter less than 2.5 μm (M2T1NXAER). The data products of air pollution had a spatial resolution of 0.5°×0.625° lat-lon.

2.3. Statistical model

Equation 1 represent the Negative Binomial distribution and it is used to model count data.

$$Y_t \sim NB(\mu_t) \qquad (1)$$

where Y_t denoted the count variable at the time *t* and μ_t its expectation. With the structure of a Farrington-like model, the expectation is defined as

$$\mu_t = \exp{(\beta_0 + \beta_1 t)} \qquad (2)$$

where β_0 is called the intercept, β_1 accounts for the linear time trend, and a time term that could be measured in days, weeks, months, years, etc. (Zhou et al., 2017). Our approach defined the expectation as

$$\mu_t = \exp{(\beta_0 + \beta_1 t + \beta_2 X_{1,t} + \beta_3 X_{2,t} + \ \dots + \beta_k X_{k-1,t})}. \qquad (3)$$

Where the regression coefficients $\beta_2, \beta_3, \dots,$ and β_k are unknown parameter that are calculated from a set of data. $X_{i,t}$ represents the vector of regressors that change temporally.

2.4. Methodology

To propose a mortality model with atmospheric covariates at different geographical levels using a generalized linear modelling framework (GLM) with Negative Binomial distribution, our approach encompasses the following steps: (i) Data treatment (mortality data disaggregated, and meteorological and air pollution dataset with annual temporal resolution), (ii) applying the GLM model to data, and (iii) evaluating the models (between common or basic approach with the proposed approach).

Models were created using the packages *MASS* (*glm.nb* and *stepAIC* functions) and *lmtest* (*lrtest* fuction) in the R statistical Environment (Ripley et al., 2020; Zeileis, 2020). The R script is described in Sánchez-Balseca and Pérez-Foguet (2020).

Step 1. Data treatment

The monthly means value of each meteorological and air pollution covariate was obtained from all grid cells that were covered by the polygon of each geographical level. The polygon describes the different geographical level: county, macro-climate region, and country. Then, the meteorological and air pollution monthly dataset was calculated as the annual mean value over each geographic level. Wang et al. (2020) used a similar method to obtain the annual average concentrations of $PM_{2.5}$ in their epidemiological studies. The concentration of pollutants was transformed into $\mu g \cdot m^{-3}$. Finally, the trend analysis for each meteorological and air pollution covariate was performed.

The mortality data was disaggregated by sex and age group for each geographic level. Following the suggestion of Pan American Health Organization in 2017, we used the three “functional” age groups: A (young people under 24 years), B (middle-aged people between 25 and 54 years), and C (people over 55 years).

Step 2. Model Application

We applied GLM model to the dataset considering the Negative Binomial distribution, where Y_t denoted the observed annual deaths at year *t*, and $X_{i,t}$ denote the large-scale component including meteorological, geographical, and environmental (air pollution concentration) covariates.

First, we applied the GLM using the structure of a Farrington-like model to the disaggregated data by sex and age ranges for each geographic level, i.e., we only considered the time *t* (year) as the covariate.

To check the model selection between Poisson and BN2 in mortality data, we used a likelihood ratio test. NB2 assume the conditional means are not equal to the conditional variances. This inequality is captured by estimating a dispersion parameter that is keeping constant in a Poisson regression model. Taking account, the Poisson regression model is actually nested in the negative binomial model, we can then use a likelihood ratio test to compare these two and test the model selection (we used the function *lrtest* in R). For this purpose, we used the basic model explained in the last paragraph in a Poisson regression modelling framework at the country level for mortality data disaggregated by sex.

Then, we applied the model using the atmospheric covariates (air temperature, pressure, specific humidity, precipitation, CO, SO_2, and $PM_{2.5}$). We applied the GLM to the disaggregated data by sex and age ranges for each geographic level, i.e., in each geographic level we had six models. For covariates selection in GLM approach, the suggested Akaike Information Criterion (AIC) was used (Gelman et al., 2014). We used the function *stepAIC* in R to produce the optimal set of covariates (taking into account their significance), this function also removes the multicollinearity if it exists. The function *stepAIC* choose the model by AIC in a stepwise algorithm.

Step 3. Model Evaluation

The Nash-Sutcliffe Efficiency Index (NSE) and Pearson correlation were used in the model evaluation. We compared the common and basic approach (only time *t* as covariate) with the proposed approach (with significant covariates) for each geographic level. NSE (Eq. 5) is a widely used and potentially reliable statistic for assessing the goodness of fit of models. The NSE scale is from 0 to 1, whereby *NSE = 1* means the model fitted perfect. *NSE = 0* means that the model is equal to the average of the observed data, and negative values mean that the average is a better predictor (McCuen et al., 2006)

$$NSE = 1 - \frac{\sum(Yobs_i - Ysim_i)^2}{\sum(Yobs_i - \overline{Yobs})^2}. \qquad (3)$$

The NSE and Pearson correlation are independent of the scale of measurement of the variables. $Yobs_i$ denotes the observed annual count mortality. $Ysim_i$ denotes simulated annual count mortality. The quality metrics for both basic and proposed approaches for each age range and sex at different geographic level were obtained.

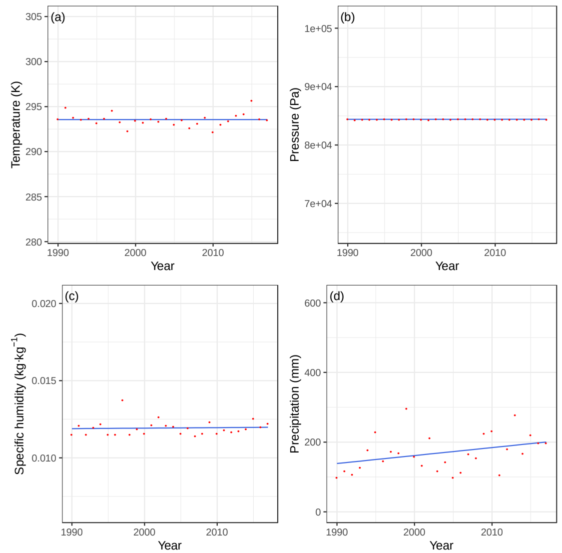
3. Results

Meteorological and air pollution data description

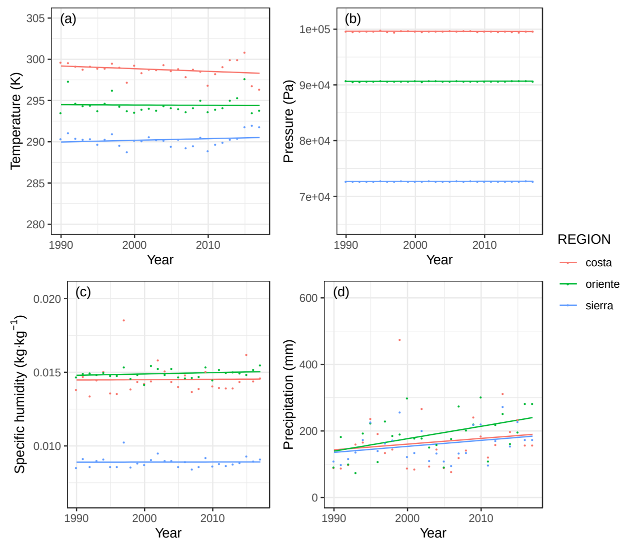
The meteorological variables at country level show some extreme events (see Figure 2), and they are related with the ENSO. For instance, in 1997-1999 were recorded pikes in the temperature, humidity, and precipitation. During these years Ecuador had the most extreme ENSO event in the last century (UCL, 2018). The temperature in the country level does not show a trend and it was constant along the time. The pressure and humidity show a slight growing trend, while, precipitation shows a strong trend increase.

The ENSO effects also are visible in the macro-climatic regions, but the most important pikes are in the Coast Region. In this region, all counties presented highest values in the humidity. Stand out with the highest values of precipitation, “Los Ríos” and “Santa Elena” in the Coast region; in the Andean region “Bolivar” and “Santo Domingo de los Tsáchilas”. While in the Amazon region, stand out with the highest values of temperature “Orellana”, “Pastaza”, and “Sucumbios”.

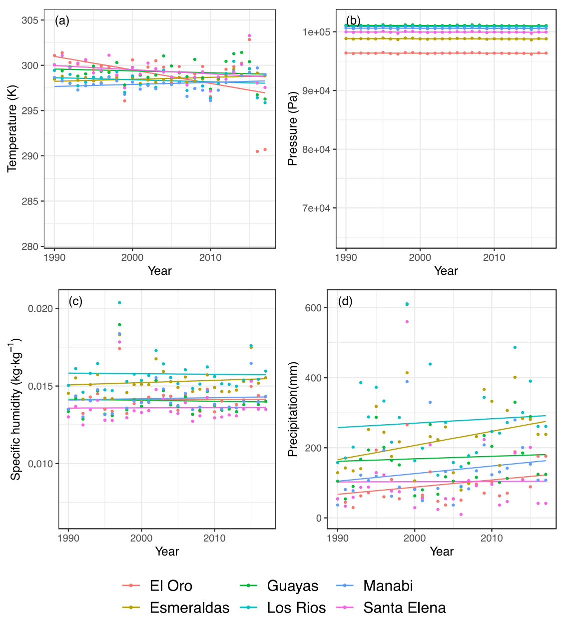
In general, the Coast region presented a decreasing trend of temperature, the Andean region presents a strong trend of temperature increase while the Amazon region does not show a trend. The three macro-climate regions present a trend of precipitation increase. Morán-Tejeda et al. (2016) and Eguiguren-Velepucha et al. (2020) presented similar behavior of temperature and precipitation in their study of climate trends in Ecuador.



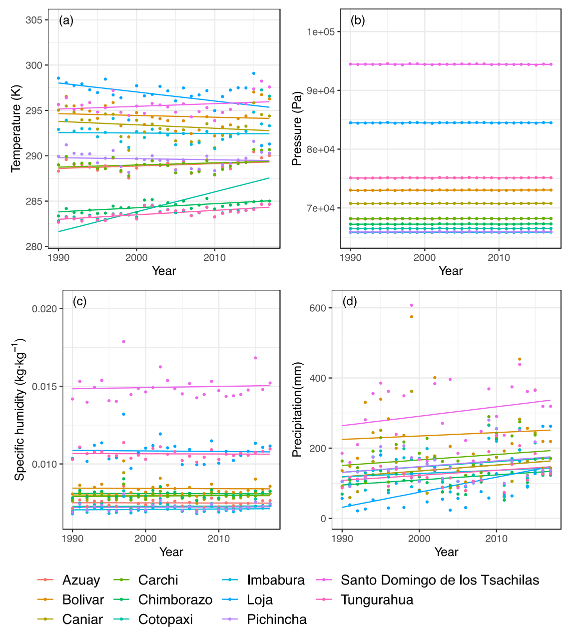
Country



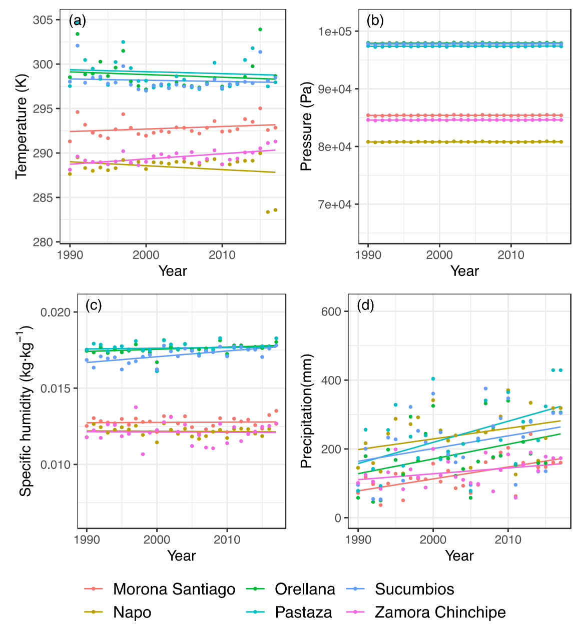
Macro-climate regions



Counties at Coast Region



Counties at Andean Region



Counties at Amazon region

Figure 2. Meteorological data trend. (a) Temperature (K), (b) Pressure (Pa), (c) Specific humidity ($kg \cdot kg^{-1}$), and (d) Precipitation (mm).

At the county geographical level, the variability of air pollutant was more evident. For instance, the annual mean concentration of SO_2 varied between 0.0746 to 3.54 $\mu g \cdot m^{-3}$. The threshold level suggested World Health Organization for the daily mean concentration for SO_2 was 125 $\mu g \cdot m^{-3}$, and it is statistically equivalent to 40 $\mu g \cdot m^{-3}$ for the yearly mean of hourly average (Yanagisawa, 1973). The annual mean concentration of $PM_{2.5}$ was between 0.0264 to 6.898 $\mu g \cdot m^{-3}$. The threshold level suggested by the World Health Organization for the annual mean concentration for $PM_{2.5}$ is 10 $\mu g \cdot m^{-3}$ (WHO, 2006). Finally, the annual mean concentration of CO was in the range from 67.75 to 176.68 $\mu g \cdot m^{-3}$. These values were lower than 20 000 $\mu g \cdot m^{-3}$ which is measured in urban traffic environments of large European cities (WHO, 2000).

The air pollutant data description in Figure 3 shows some pikes across the time, overview the most important pikes were identified in 1991-1993, 1996-1997, 2000-2002, 2005-2006, and 2010-2012. If we evaluated the period 1996-1997, the concentrations of CO and SO_2 stand out at the country level, in low proportion the concentration of $PM_{2.5}$. In the same period at macro-climate region, the Coast region had the most important variation in the concentrations of CO and SO_2.

Considering the period 2000-2002, the three air pollutants showed clear pikes in the 2002 at the country and macro-climate region, where the concentration of SO_2 stand out in the Andean Region, and the concentration of $PM_{2.5}$ stand out in the Coast Region. This event is mainly related with the volcanic activity, in 2001-2003 the volcanoes "Guagua Pichincha" and "Reventador" presented a high activity, with a Volcanic Explosivity Index (VEI) superior to three. Figure 3 at the county level shows important peaks for the period 2001-2003 in the counties of "Pichincha" and "Sucumbios for concentrations of CO and SO_2 where there are located the volcanoes "Guagua Pichincha" and "Reventador" respectively. Other important volcanic events are: In 2009-2012 the volcanoes

"Tungurahua" and "Sangay" presented volcanic activity; In 2015-2016 three volcanoes were actives ("Sangay", "Cotopaxi", and "Tungurahua").

The concentration of SO_2 and $PM_{2.5}$ at the country level increased, while the trend of CO decreased. In general, the three macro-climatic regions had positive trends for SO_2 and $PM_{2.5}$. Related with the concentration of CO, the Amazon region has the highest concentrations, while the Coast region had the lowest.

In the air pollution analysis at county level is necessary take account the human activities (chimneys, emissions from vehicles, etc). Guayas and Los Ríos had the highest concentrations of SO_2, Bolivar and Cañar in the Andean region, Orellana and Sucumbíos in the Amazon region. Zalakeviciute et al. (2020) showed through satellite images the spatial distribution of SO_2 concentration in Ecuador, and their results are similar to our description. For instance, Guayas is the principal port of Ecuador and it has the most important affectation due to the emissions of SO2 from ship chimneys (Quevedo, 2015). Related to CO at the county level, Guayas had the highest concentrations until 2005, after this year Los Ríos in the Coast Region. Pichincha had the highest concentrations of CO in the Andean Region. Orellana presented highest concentrations in the Amazon Region. These counties have the largest vehicle fleet and thus the most important CO emissions (Estrella et al., 2019). Santa Elena and Guayas had the highest concentrations of $PM_{2.5}$ in the Coast region, Cañar in the Andean region, and Zamora Chinchipe in the Amazon region.

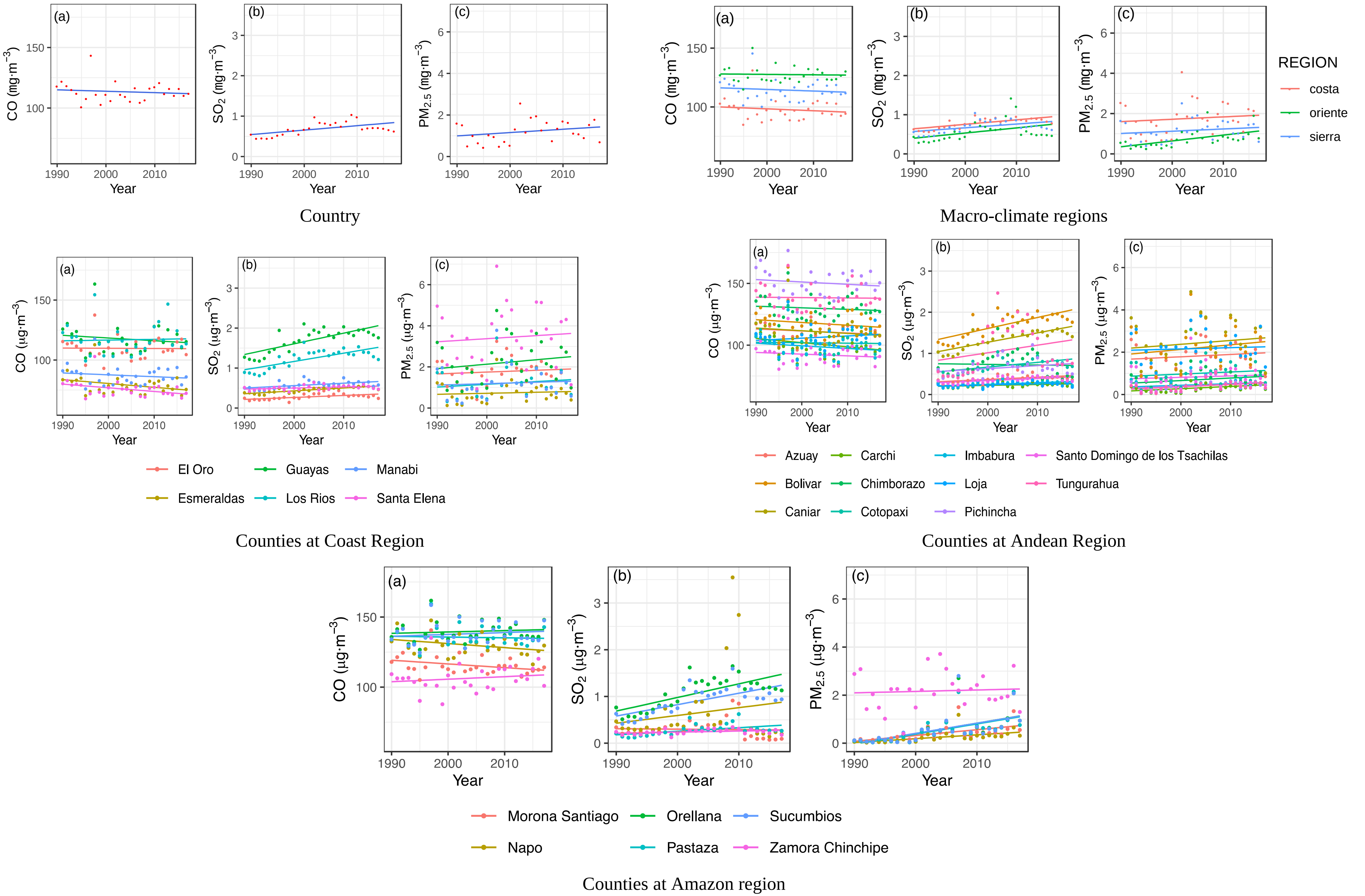


Figure 3. Air Pollution data trend. (a) CO ($\mu g \cdot m^{-3}$), (b) SO_2 ($\mu g \cdot m^{-3}$), and (c) $PM_{2.5}$ ($\mu g \cdot m^{-3}$)

Application and evaluation model

Geographical level: Country

If we aggregate the data between female and male people at the country level, skipping for a moment the age groups, we can see that the figure 4 shows that the number of deaths (denoted by dots) of male people was higher than female people. For instance, the numbers of death in 1990 was superior in male infants than female infants. In 2001, the number of male infant deaths was even higher in 1990 (INEC, 2001). Additionally, some pikes in the number of deaths are notorious in the years 1992, 2000, 2006, 2010 and 2011 for both female and male people.

In the assessment of model assumption using a likelihood ratio test, the p-value from the chi-squared distribution with one degree of freedom (Df) strongly suggested the BN2 model is more appropriate then the Poisson model for the mortality data used in this work.

Table 1. Results of the likelihood ratio test for Poisson and NB2 regression models by sex at the country level.

Country	Model	Df	LogLik	Df	Chisq	Pr (>Chisq)
Female	Poisson	2	-855.36			
	NB2	3	-235.93	1	1238.9	< 2.2e-16 ***
Male	Poisson	2	-446.81			
	NB2	3	-227.04	1	439.53	< 2.2e-16 ***

The mortality data modeling under a demographic approach (DM) (as a function of the time) presents good quality metrics (see figure 4). The demographic approach for female data mortality has an NSE equal to 0.8221, and a correlation coefficient equal to 0.9078. Under the same scenario, if we add meteorologically and air pollution covariates to the analysis, the quality metrics present better values. Table 2 shows the selected covariates for female mortality data modelling and these were time, CO, and SO_2. Using our approach (PM), we obtained an NSE equal to 0.9278 and a correlation coefficient equal to 0.9633 to female data mortality. The AIC criterion value was lower using our approach and it is equal to 456.16 while using the demographic approach the AIC is equal to 477.85.

The selected covariates for male mortality data modelling were time, pression, and SO_2. For these selected covariates the model presented an AIC criterion equal to 454.3, while under a demographic approach the AIC criteria was equal to 460.08 (see Table 2). For male data mortality, we obtained NSE values equal to 0.6756 and 0.9401 for both demographic and proposed approaches respectively. Additional, correlation coefficients equal to 0.9595 and 0.9696 for both demographic and proposed approaches respectively.

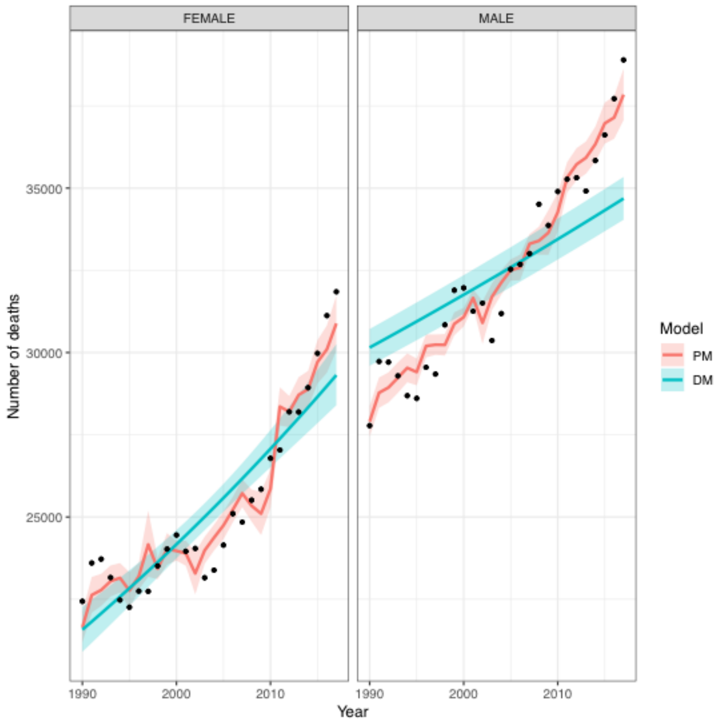


Figure 4. Mortality data modelling by sex at geographic level of country using both demographic (DM) and proposed (PM) approaches.

Table 2. Covariates selection using AIC for mortality data modelling by sex at the country level.

Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	AIC
Country: Female								
x	x	x	x	x	x	x	x	460.53
x	x		x	x	x	x	x	458.71
x	x		x	x	x	x		457.12
x	x		x		x	x		456.92
x			x		x	x		456.31
x					**x**	**x**		**456.16**
x								477.85
Country: Male								
x	x	x	x	x	x	x	x	462.67
x	x	x	x	x	x	x		460.67
x		x	x	x	x	x		459.01

x		x	x	x	x	457.6
x		x		x	x	456.09
x		x			x	455.71
x	**x**				**x**	**454.3**
x						460.08

If we disaggregated the mortality data at country level by age groups and sex (see Figure 5), the first age group had a pike in 2010 for both sexes. Mortality data in middle-aged people presented three notorious pikes in the number of deaths in male people in the years 2000, 2005, and 2009. The third age group had pikes in 1991 and 2000 for both sexes, and one additional pike in 2008 for male people. Additionally, the mortality in people over 55 years had a strong trend increase.

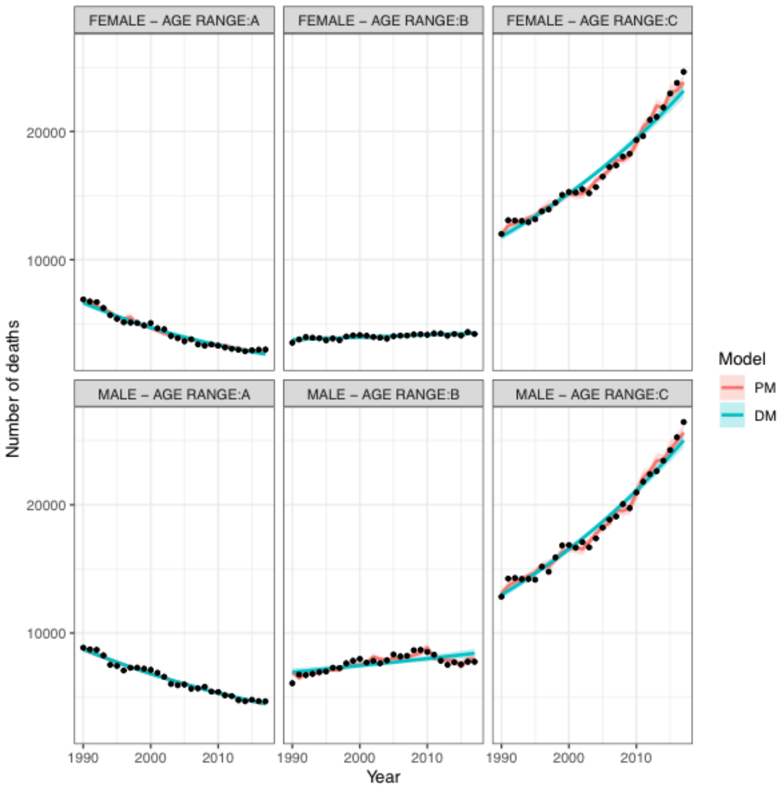


Figure 5. Mortality data segregated by sex and age range at geographic level of country using both demographic (DM) and proposed (PM) approaches.

The mortality data modeling under a demographic approach presents good quality metrics, except to mortality data of male middle-aged people that it had an NSE equal to 0.4367 and a correlation coefficient equal to 0.6625 (see Table 1). For this same case, our approach presented an NSE equal to 0.7926 and a correlation coefficient equal to 0.8903.

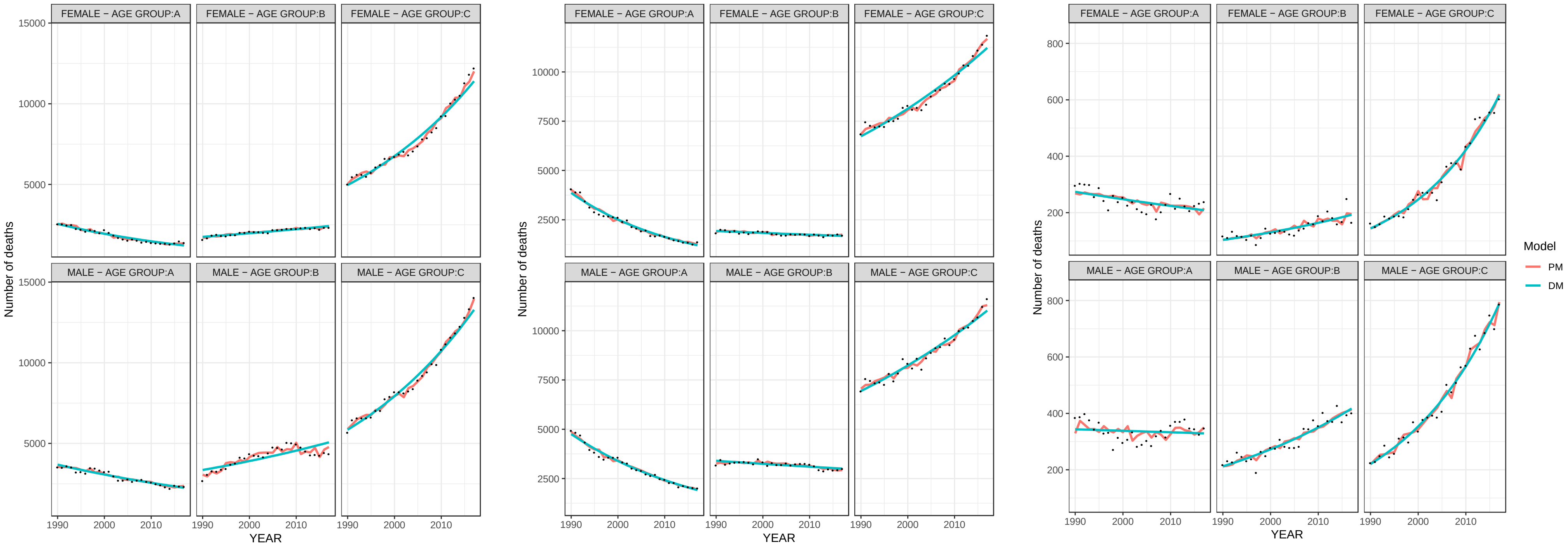
The time, temperature, and SO_2 were the variables selected. In general, the covariate pressure at the country level was not significant.

Table 3. NSE, and correlation coefficient for both basic and proposed approach for each sex and age range with its significant covariates at the country level.

Age Range	Sex	Proposed Model								NSE		Coef. Corr.	
		Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	DM	PM	DM	PM
A	F	x					x	x		0.969	0.975	0.985	0.987
	M	x								0.973	0.973	0.986	0.986
B	F	x	x						x	0.646	0.722	0.804	0.849
	M	x	x					x		0.436	0.792	0.662	0.890
C	F	x	x		x	x	x	x		0.968	0.989	0.984	0.994
	M	x			x		x	x		0.976	0.989	0.988	0.994

Geographical level: Macro-climate regions

If we evaluate the first age group (young people), the number of deaths had a decreased trend in all macro-climate regions, the Andean region had the strongest decreased trend (from around 5000 deaths in 1990 to around 1250 deaths in 2017). The Coast and Andean regions had similar pikes in 1991-1992, 1997 and 2000 (see Figures 6a, and 6b). These regions had higher deaths in Ecuador. The Amazon region also presented pikes in 1991-1992, 2000, and 2010. Additionally, the Amazon region presented over-dispersed data and low numbers of deaths in young people (see Figure 6c). The mortality data modeling for young people (age group A) under a demographic approach presented adequate quality metrics, except for the Amazon region. The mortality data of female young people in the Amazon region had an NSE value equal to 0.3896 and a coefficient of correlation equal to 0.6256. In the same scenario, our approach with time and $PM_{2.5}$ as selected covariates did not present adequate quality metrics, we obtained an NSE equal to 0.4434 and a coefficient of correlation equal to 0.6668. The quality metrics for mortality data of male young people were even lower than in the female young people (see Table



(a) (b) (c)

Figure 6. Mortality data segregated by sex and age group at geographic level of macro-climate region (a) Coast, (b) Andean, and (c) Amazon; using both demographic (DM) and proposed (PM) approaches.

In the second age group (middle-aged people), each macro-climate region had different behavior of mortality data. The Coast region presented a trend increase, where it is strongest in the male people. The Andean region did not show a notorious trend, keeping the mortality data apparently constant. The mortality data for Coast and Andean regions had similar pikes in 1991, 1998, 2000, 2005, and 2008-2010, being it more visible in the male people. The Amazon region present a similar trend to the Coast region, despite that the mortality data had overdispersion, was possible identify important pikes in 2000, and 2009-2010. For this age group, the mortality data modelling presented adequate quality metrics for all macro-climatic regions, our approach presented a slight improvement. The Andean region had the lowest quality metrics. The mortality data modelling of female middle-aged people had values of NSE equal to 0.633 and 0.7539, and correlation coefficients equal to 0.795 and 0.868 for both demographic and proposed approach respectively. Time, temperature, pressure, CO, and $PM_{2.5}$ were the selected covariates in this scenario (see Table 2).

Table 4. NSE, and correlation coefficient for both demographic and proposed approaches for each sex and age group with its selected covariates at the macro-climate regions.

Age group	Sex	Proposed Model								NSE		Coef. Corr.	
		Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	BM	PM	BM	PM
Climatic Region: Coast													
A	F	x			x			x		0.939	0.952	0.969	0.976
	M	x	x		x				x	0.904	0.944	0.951	0.972
B	F	x	x		x			x	x	0.866	0.909	0.931	0.954
	M	x	x				x	x	x	0.554	0.833	0.751	0.913
C	F	x					x	x		0.968	0.988	0.985	0.994
	M	x					x	x		0.981	0.993	0.991	0.997
Climatic Region: Andean													
A	F	x	x							0.975	0.979	0.988	0.989
	M	x	x		x					0.977	0.981	0.989	0.991
B	F	x	x	x			x		x	0.633	0.753	0.796	0.868
	M	x	x		x		x			0.550	0.743	0.742	0.862
C	F	x	x					x		0.963	0.985	0.982	0.993
	M	x						x		0.954	0.969	0.977	0.984
Climatic Region: Amazon													
A	F	x							x	0.389	0.443	0.626	0.667
	M	x		x				x		0.018	0.207	0.134	0.455
B	F	x	x							0.660	0.696	0.813	0.835
	M	x					x			0.871	0.875	0.933	0.936
C	F	x		x				x		0.971	0.979	0.986	0.989
	M	x				x	x		x	0.977	0.985	0.989	0.992

The third age group presented a strong increasing trend in all macro-climate regions. The Coast and Andean regions had a higher number of deaths. The mortality data in the Coast region showed some pikes and great data variability in 1992-1993, 1998-2000, and 2009-2010. The mortality data in the Andean region presented notorious pikes in 1991, 1992, and 1999-2000. In the case of Amazon region, the pikes of female deaths were notorious in 2000, 2006, and 2010-2012. The mortality data modelling for this age group presented good quality metrics for the demographic and proposed approach with NSE > 0.9543 and correlation coefficient values over 0.9682.

Geographical level: County

At this stage, we will describe a significant county as an example for each macro-climate region, and then, we will present the general results of the modelling process. For the Coast, Andean, and Amazon regions we choose the counties of “Los Ríos”, “Pichincha”, and “Morona Santiago” respectively. Each county fulfills three conditions: (i) founded/created before 1990, (ii) having difference in the quality metrics between demographic and proposed approach, and (iii) having the majority of meteorological and air pollutant data as selected covariates.

“Los Ríos” county belongs to the Coast region. The number of deaths in “Los Ríos” county for young male people is superior to the young female people (see Figure 7). The mortality data for female young people had a decreased trend, while the mortality data for male people had an increasing trend. The mortality data in young female people had notable pikes in 1999-2001. The data for young male people showed overdispersion, but it had some pikes in 1993, 1998-2001, 2008-2009, and 2016. The number of deaths for female middle-aged people was superior to male people. The mortality data for middle-aged people had an increasing trend, but it is stronger in the mortality data of female people with some pikes in 2000, and 2008-2010. Finally, the mortality data in people over 55 years old had an increasing trend and were notorious two pikes in 2000 and 2010 for both sexes.

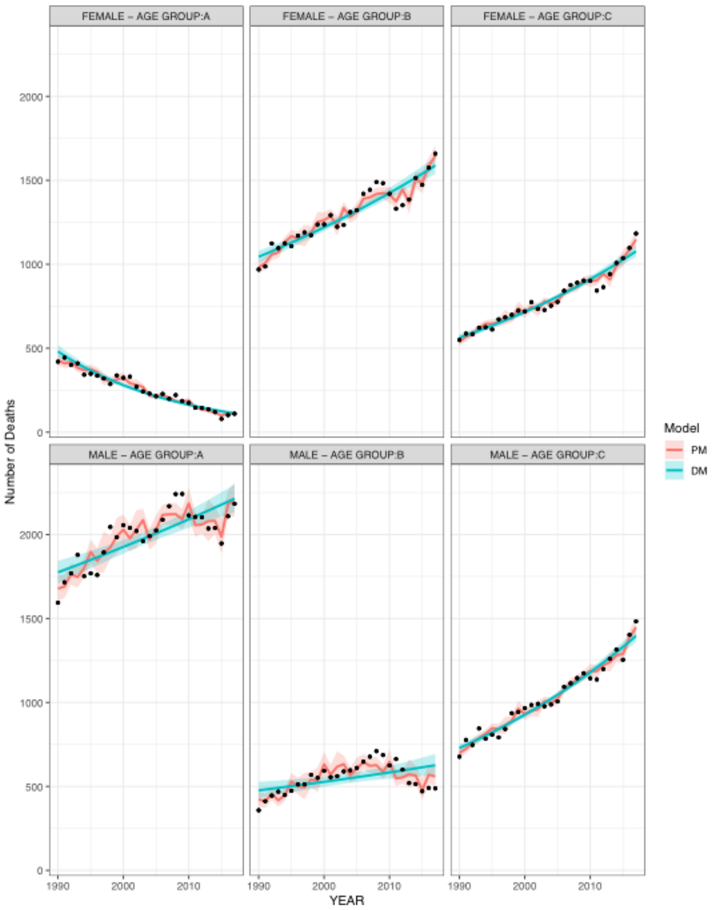


Figure 7. Mortality data segregated by sex and age group at geographic level of county "Los Ríos" using both demographic (DM) and proposed (PM) approaches.

The counties located in the Coast region presented adequate quality metrics in the mortality data modeling of young people (age group A) for both demographic and proposed approaches, except for Santa Elena using a demographic approach (see Table 3). Santa Elena was created in 2007, thus the data collected was scarce in comparison with the other counties. Overview, the NSE values were over 0.712 and 0.809, and the correlation coefficient values over 0.844 and 0.899 for demographic and proposed approaches respectively. These threshold values corresponding to Esmeraldas county. Santa Elena county presented important improvements in the quality metrics using the proposed approach in the modeling of mortality data. For example, the mortality data modelling of female young people had an NSE value equal to 0.950, and a correlation coefficient equal to 0.975; while using the demographic approach we obtained an NSE equal to 0.323 and a correlation coefficient equal to 0.569. The selected covariates in this example were time, temperature, pressure, humidity, precipitation, and SO_2.

The modelling under a demographic approach did not showed adequate quality metrics for mortality data in middle-aged people for both male and female sex at the county level (see Table 3). Only two counties presented a well-fitted modelling under a demographic approach. “Guayas” county presented good quality metrics for mortality data of female and male people, also, “El Oro” county presented good quality metrics for mortality data of male people (NSE equal to 0.528 and correlation coefficient equal to 0.727). The proposed modelling approach for mortality in middle-aged people presented important improvements in the quality metrics. However, the modelling of female mortality data in “Esmeraldas” did not have adequate quality metrics (NSE equal to 0.437 and correlation coefficient equal to 0.661). In this case, the time, temperature, humidity, and CO were the selected covariates.

The both demographic and proposed approach in the mortality data modelling for people over 55 years had good quality metrics, and respectively, with NSE values over 0.732 and 0.885; and correlation coefficients over 0.856 and 0.941. These threshold values corresponding to Esmeraldas county, similar to mortality data modelling in young people.

Table 5. NSE, and correlation coefficient for both basic and proposed approach for each sex and age range with its significant covariates at county level of Coast region.

Age group	Sex	Proposed Model								NSE		Coef. Corr.	
		Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	BM	PM	BM	PM
El Oro													
A	F	x	x							0.931	0.968	0.965	0.984
	M	x	x							0.841	0.879	0.917	0.937
B	F	x					x	x		0.343	0.526	0.585	0.726
	M	x					x	x		0.528	0.755	0.727	0.869
C	F	x				x				0.975	0.991	0.987	0.995
	M	x						x		0.967	0.979	0.983	0.989
Esmeraldas													
A	F	x							x	0.896	0.927	0.949	0.963
	M	x	x							0.707	0.827	0.842	0.909

	B	F	x	x		x		x			0.035	0.437	0.188	0.660
		M	x	x					x		0.173	0.537	0.424	0.733
	C	F	x		x	x	x	x			0.794	0.899	0.892	0.948
		M	x	x	x	x		x			0.732	0.886	0.856	0.941
Guayas														
	A	F	x							x	0.712	0.809	0.844	0.899
		M	x							x	0.659	0.795	0.812	0.892
	B	F	x							x	0.864	0.885	0.930	0.941
		M	x					x	x	x	0.597	0.845	0.778	0.919
	C	F	x					x			0.956	0.975	0.979	0.988
		M	x	x	x	x		x		x	0.964	0.980	0.982	0.990
Los Ríos														
	A	F	x	x		x			x	x	0.929	0.967	0.967	0.983
		M	x	x					x		0.849	0.923	0.927	0.961
	B	F	x					x	x	x	0.024	0.546	0.154	0.739
		M	x	x					x	x	0.185	0.655	0.438	0.809
	C	F	x	x		x		x			0.942	0.971	0.970	0.985
		M	x	x						x	0.960	0.975	0.979	0.987
Manabí														
	A	F	x								0.913	0.913	0.955	0.955
		M	x		x						0.907	0.926	0.952	0.962
	B	F	x			x		x			0.103	0.505	0.322	0.711
		M	x	x	x						0.353	0.734	0.599	0.857
	C	F	x						x		0.962	0.978	0.981	0.988
		M	x								0.976	0.976	0.9877	0.987
Santa Elena														
	A	F	x	x	x	x	x		x		0.323	0.950	0.569	0.974
		M	x	x	x	x	x		x		0.198	0.979	0.445	0.989
	B	F	x	x	x	x	x		x	x	0.238	0.918	0.488	0.957
		M	x	x	x	x	x		x		0.021	0.861	0.144	0.928
	C	F	x	x	x	x	x	x	x	x	0.849	0.992	0.921	0.996
		M	x	x	x	x	x	x	x	x	0.896	0.994	0.947	0.997

“Pichincha” county presented a similar behavior to “Los Ríos” county in the mortality data in young people. The number of deaths in young male people was superior to the number of deaths in female people, and it followed an increasing trend (see Figure 8). The pikes in 2000, 2007, and 2015 were more notorious in the mortality data of young male people. The number of deaths in female middle-aged people is superior to the number of deaths in male people, and it showed an increased trend with important pikes in 2000, 2006, and 2012. Finally, the mortality data for people over 55 years presented an increasing trend with pikes in 2000, 2007, and 2012.

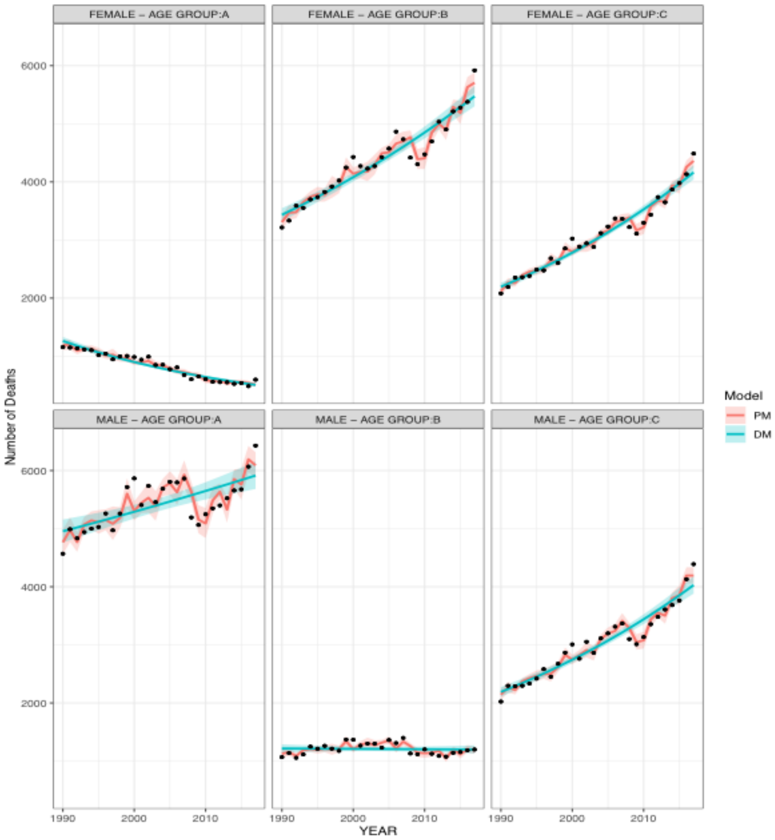


Figure 8. Mortality data segregated by sex and age group at geographic level of county "Pichincha" using both demographic (DM) and proposed (PM) approaches.

The counties located in the Andean region presented adequate quality metrics in the mortality data modeling of young people (age group A) for both demographic and proposed approaches, except for "Santo Domingo de los Tsáchilas" using a demographic approach (see Table 4). Santo Domingo was created in 2007, thus the data collected was scarce in comparison with the other counties in its macro-climate region. However, the quality metrics for "Santo Domingo de los Tsáchilas" showed improvements using the time, temperature, humidity, SO_2, and $PM_{2.5}$ as selected covariates. We obtained an NSE equal to 0.649 and a correlation coefficient equal to 0.820.

The modelling of mortality data for female middle-aged people using a demographic approach did not present adequate quality metrics for Azuay, Bolivar, Cañar, Carchi, Imbabura, and Pichincha. The low threshold values corresponding to Pichincha county with an NSE value of 0.045 and correlation coefficient value of 0.211. In the same scenario, the addition of meteorological and air pollutant covariates improvement the quality metrics, except to Imbabura and Pichincha, with NSE values of 0.389 and 0.402, and correlation coefficient values of 0.629 and 0.634 respectively. The modeling of mortality data for male middle-aged people using a demographic approach did not present adequate quality metrics to Bolivar, Carchi, and Pichincha. For Carchi, adding meteorological and air pollutant covariates did not improve the modelling of male middle-aged people with an NSE equal to 0.499 and a correlation coefficient equal to 0.707. For Bolivar and Pichincha, the improvement was notorious.

The modelling of mortality data for female people over 55 years using a demographic approach presented unacceptable quality metrics for Bolivar, and Cotopaxi; with NSE values of 0.259 and 0.239, and correlation coefficient values of 0.510 and 0.489 respectively. Adding meteorological and air pollutant covariates did not register significantly improves in the modelling of mortality data of female people over 55 years in Cotopaxi, while the modelling in Bolivar with time and temperature as selected covariates showed improvements. The modelling of mortality data for male people over 55 years using a demographic approach presented bad quality metrics for Carchi, Chimborazo, and Cotopaxi. In the same scenario, adding meteorological and air pollutant covariates only improvements the modelling of mortality data of Carchi county, where the selected covariates were time and pressure.

Table 6. NSE, and correlation coefficient for both basic and proposed approach for each sex and age range with its significant covariates at county level of Andean region.

Age group	Sex	Proposed Model								NSE		Coef. Corr.	
		Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	BM	PM	BM	PM
Azuay													
A	F	X	X		X		X	X		0.816	0.945	0.909	0.972
	M	X			X		X	X		0.759	0.909	0.877	0.954
B	F	X	X				X	X		0.383	0.519	0.619	0.721
	M	X			X					0.646	0.778	0.804	0.882
C	F	X	X		X		X	X		0.889	0.964	0.944	0.982
	M	X			X		X	X		0.877	0.938	0.937	0.969
Bolivar													
A	F	X					X			0.908	0.980	0.955	0.990
	M	X	X	X				X		0.854	0.915	0.924	0.956
B	F	X	X							0.229	0.506	0.479	0.712
	M	X	X							0.372	0.744	0.612	0.863
C	F	X	X							0.259	0.643	0.510	0.802

		M	x								0.521	0.521	0.722	0.722
Cañar														
	A	F	x			x		x	x		0.813	0.925	0.907	0.962
		M	x		x	x		x	x		0.607	0.914	0.792	0.956
	B	F	x	x							0.425	0.559	0.652	0.748
		M	x								0.699	0.699	0.836	0.836
	C	F	x	x							0.842	0.896	0.918	0.947
		M	x			x			x	x	0.675	0.807	0.822	0.898
Carchi														
	A	F	x								0.899	0.899	0.948	0.948
		M	x			x					0.928	0.945	0.963	0.972
	B	F	x							x	0.389	0.503	0.624	0.709
		M	x	x							0.282	0.499	0.531	0.707
	C	F	x								0.604	0.604	0.777	0.777
		M	x		x						0.310	0.526	0.557	0.726
Chimborazo														
	A	F	x			x		x			0.939	0.958	0.969	0.979
		M	x			x			x		0.928	0.959	0.964	0.979
	B	F	x		x						0.874	0.911	0.935	0.954
		M	x						x		0.896	0.929	0.946	0.964
	C	F	x								0.728	0.728	0.853	0.853
		M	x								0.493	0.493	0.702	0.702
Cotopaxi														
	A	F	x			x					0.963	0.976	0.981	0.988
		M	x	x							0.968	0.978	0.984	0.989
	B	F	x	x							0.759	0.817	0.872	0.904
		M	x			x					0.691	0.809	0.832	0.899
	C	F	x								0.239	0.239	0.489	0.489
		M	x								0.256	0.256	0.506	0.506
Imbabura														
	A	F	x								0.926	0.926	0.964	0.964
		M	x	x							0.964	0.967	0.982	0.983
	B	F	x	x				x			0.078	0.389	0.279	0.624
		M	x								0.659	0.659	0.812	0.812
	C	F	x		x						0.656	0.783	0.810	0.885
		M	x		x				x		0.582	0.767	0.763	0.876
Loja														
	A	F	x								0.939	0.939	0.969	0.969
		M	x								0.970	0.970	0.985	0.985
	B	F	x				x		x		0.561	0.765	0.749	0.875
		M	x				x		x		0.096	0.250	0.309	0.5001
	C	F	x						x		0.923	0.949	0.961	0.975
		M	x			x	x	x	x		0.905	0.947	0.952	0.973
Pichincha														
	A	F	x			x		x			0.927	0.955	0.963	0.978
		M	x			x	x	x	x		0.839	0.937	0.919	0.968
	B	F	x				x	x			0.045	0.402	0.211	0.634
		M	x			x	x	x		x	0.003	0.619	0.050	0.787
	C	F	x			x	x	x	x		0.949	0.981	0.974	0.990
		M	x				x	x	x		0.917	0.969	0.957	0.985
Santo Domingo de los Tsáchilas														
	A	F	x	x		x			x	x	0.499	0.649	0.732	0.820
		M	x	x	x	x	x	x	x	x	0.505	0.635	0.734	0.812
	B	F	x	x		x			x	x	0.662	0.771	0.832	0.888
		M	x	x	x	x			x	x	0.598	0.693	0.794	0.846
	C	F	x	x	x			x	x		0.786	0.853	0.901	0.928
		M	x	x	x	x	x	x	x	x	0.776	0.836	0.895	0.920
Tungurahua														
	A	F	x				x	x			0.956	0.971	0.978	0.986
		M	x						x		0.947	0.957	0.973	0.978
	B	F	x	x							0.530	0.578	0.728	0.760
		M	x								0.543	0.543	0.737	0.737
	C	F	x								0.880	0.880	0.938	0.938
		M	x		x						0.771	0.822	0.878	0.907

"Morona Santiago" had low number of deaths in comparison with the counties located in the Coast and Andean macro-climate regions (Figure 9). The number of deaths in young male people was superior to female people, even it has an increasing trend. Despite the overdispersion, it was possible to identify pikes in 1991, 1998-2000, 2009-2010, and 2014-2015. The mortality data in middle aged people had an increased trend, being it more notorious in the mortality data of female people. Additionally, the number of deaths of female middle-aged people is superior to male people and were notorious pikes in 2000, 2006-2009, 2013, and 2016. The mortality data for people over 55 years had increasing trend, with pikes in 1999, 2006, 2013-2015.

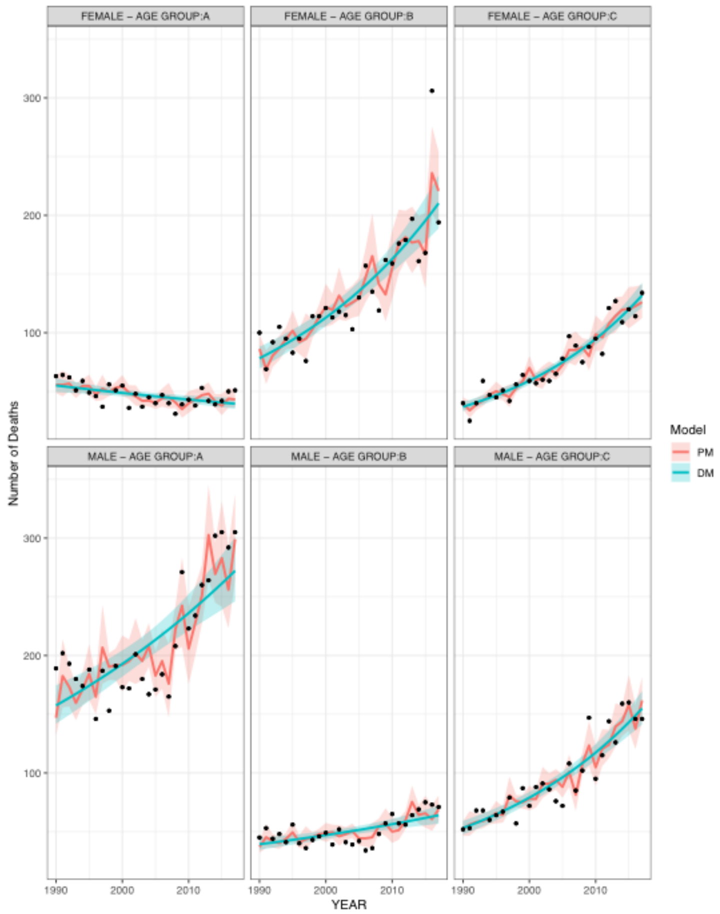


Figure 9. Mortality data segregated by sex and age group at geographic level of county "Morona Santiago" using both demographic (DM) and proposed (PM) approaches.

The modelling of mortality data for young female people using a demographic approach presented bad quality metrics for "Morona Santiago", "Orellana", "Pastaza", and "Sucumbíos"; with NSE values under 0.315 and correlation coefficients under 0.492. The selected covariates time, temperature, CO, SO_2, and $PM_{2.5}$ improvement the quality metrics in the modelling of mortality data for young female people only in “Orellana” (see Table 5). The modelling of mortality data for young male people using a demographic approach presented inadequate quality metrics for “Morona Santiago”, “Orellana”, “Pastaza”, and “Sucumbíos”. Only in “Orellana” our approach did not present adequate quality (NSE equal to 0.429, and correlation coefficient equal to 0.655), while in the remain of counties had well quality metrics.

The modelling of mortality data for female middle-aged people using a demographic approach presented problems in “Morona Santiago”, “Napo”, “Orellana”, and “Zamora Chinchipe”. For this scenario, adding extra information worked well for “Orellana”, where the selected covariates were time, temperature, CO, SO_2, and $PM_{2.5}$. The modelling of mortality data for male middle-aged people using meteorological and air pollutant data improvement the demographic approach, and it had adequate quality metrics, except to “Napo” and “Zamora Chinchipe”.

In general, the modelling of mortality data for people over 55 years using our approach presented better quality metrics than using the demographic approach. However, “Napo” did not present adequate quality metrics, with an NSE equal to 0.480 and a correlation coefficient equal to 0.693. For this case, the time and SO_2 were the selected covariates.

Table 7. NSE, and correlation coefficient for both basic and proposed approach for each sex and age range with its significant covariates at county level in the Amazon macro-climate region.

Age group	Sex	Proposed Model								NSE		Coef. Corr.	
		Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	BM	PM	BM	PM
Morona Santiago													
A	F	X				X		X		0.315	0.492	0.562	0.702
	M	X				X	X			0.008	0.358	0.089	0.598
B	F	X	X							0.212	0.483	0.463	0.708
	M	X					X			0.461	0.606	0.681	0.778
C	F	X			X					0.902	0.924	0.949	0.961
	M	X			X					0.843	0.894	0.918	0.946
Napo													
A	F	X	X							0.683	0.794	0.840	0.897
	M	X						X		0.582	0.700	0.779	0.849
B	F	X								0.285	0.285	0.536	0.536
	M	X						X		0.171	0.349	0.415	0.591
C	F	X		X				X		0.550	0.758	0.743	0.872
	M	X						X		0.392	0.480	0.627	0.693
Orellana													
A	F	X	X				X	X	X	0.081	0.715	0.284	0.845
	M	X	X				X	X	X	0.067	0.429	0.258	0.655
B	F	X	X				X	X	X	0.361	0.624	0.601	0.789

	M	x	x	x		x	x	x		0.599	0.796	0.776	0.892
C	F	x	x				x	x	x	0.860	0.953	0.929	0.976
	M	x	x					x	x	0.781	0.921	0.885	0.959
Pastaza													
A	F	x							x	0.128	0.403	0.359	0.635
	M	x			x				X	0.098	0.568	0.314	0.753
B	F	x							X	0.502	0.585	0.709	0.765
	M	x							X	0.467	0.641	0.687	0.802
C	F	x							x	0.846	0.870	0.919	0.933
	M	x							x	0.876	0.950	0.936	0.975
Sucumbíos													
A	F	x							x	0.001	0.422	0.033	0.649
	M	x							x	0.412	0.643	0.642	0.802
B	F	x						x		0.612	0.699	0.782	0.836
	M	x						x		0.736	0.889	0.862	0.943
C	F	x						x		0.919	0.941	0.959	0.970
	M	x	x							0.948	0.961	0.973	0.980
Zamora Chinchipe													
A	F	x						x		0.656	0.704	0.809	0.839
	M	x	x							0.714	0.801	0.845	0.895
B	F	x					x			0.033	0.146	0.182	0.383
	M	x						x		0.019	0.179	0.141	0.425
C	F	x					x			0.911	0.939	0.955	0.969
	M	x	x							0.902	0.941	0.950	0.970

4. Discussion

This article proposed modelling the association between mortality data segregated by age group and sex with atmospheric parameters using GLM with negative binomial distribution at different geographic levels. We used the three "fundamental" age groups for practicality and simplicity in data management. However, our approach presented difficulties in modeling mortality data in middle-aged people, this limitation could be caused by using wide age ranges. However, for further works, each age group proposed could be divided to identify the association between mortality data and atmospheric covariates in more detail.

The negative binomial regression model is often used for over-dispersed data (Currie and Djeundje, 2010), however, additional information through both meteorological and pollution covariates improvements the dynamic behavior of modelling with a demographic approach. This characteristic allows modeling better the peaks of numbers of deaths. In this work, we analyzed the relationship between the mortality data and atmospheric factors through the selected covariates in our approach. Additionally, we have sought to answer the high variability of atmospheric factors through extreme events related to ENSO and volcanoes. The meteorological and air pollutant covariates influence directly mortality in the with or without ENSO/volcano-related-extremes.

The peaks of number of deaths in this study were related with some meteorological and air pollution extreme events. The ENSO in the period of 1991-1992 with a moderate intensity generated high values of temperature at the country level. It was notable a high number of deaths in people over 55 years and we used the time, temperature, humidity, precipitation and CO as covariates to modelling. At the macro-climate region level, stand out the high values of temperature in all regions and the high values in the precipitation for the Amazon region. In this context the high number of deaths in young people for all regions was notable, where the selected covariates for modelling were time, temperature, and humidity. Also, a high number of deaths in middle-aged people and people over 55 years was evident mainly in the Coast and Andean regions, where the selected covariates were temperature, humidity, and CO. Additionally, the concentration of carbon monoxide was affected for ENSO events, for this reason, the concentration presented pikes. This behavior has been widely studied and it is related to the affectation of the carbon cycle in the biospheric uptake of CO_2 (e.g. due to drying of tropical land regions) (Chatterjee et al., 2017). Rowlinson et al. (2019) explained that the ENSO events influence fire occurrences, wetland emission, and atmospheric circulation and thus the CO concentration increase; for instance, they concluded that during the ENSO event in 1997 and 1998, there were increased the emission from biomass burning globally, causing global CO concentrations to increase by more than 40%. Ecuador had the most extreme ENSO event in the last century in 1997-1998 (Espinoza et al., 2009; Morán-Tejada et al. 2016), where the meteorological covariates and concentration of CO in this work showed important association with the modelling of mortality data. For ENSO in 1997 the number of deaths showed a slight increase in relation to the previous years. At the country level, there were observed pikes in the number of deaths of male people for all age groups, while in the female data only in middle-aged people and people over 55 years. At macro-climate region, the high number of deaths for this event were identified in all age groups in the coast region, in the young male people of the Andean region, and in the female people in the Amazon region. The main selected covariates were time, temperature, humidity, and CO. The effects also were visible at the county level, “Los Ríos” and “Pichincha” presented a high number of deaths of male people for all age groups. The selected covariates for modelling were time, temperature, humidity, precipitation, and CO.

In the proposed work, SO_2 was a selected covariate in most cases. High emissions of SO_2 allow us to determine volcanic activities in our work, despite previous studies did not present consistent conclusions on the effect of SO_2 in the mortality data (Park et al., 2011; Analitis et al., 2014, Chen et al., 2017). For instance, the number of deaths for young male people and male middle-aged people in 2009 at country level were high. At macro-climate region, the Amazon region had highest number of young people, the three regions showed pikes in the middle-aged people, and Coast and Amazon regions showed pikes in the number of deaths of people over 55 years. In this year “Tungurahua” and “Sangay” volcanoes had activity, they are located in the counties of Tungurahua and Morona Santiago. Additional, 2010 had the highest number of reported disasters related to the climatic change (meteorological, hydrological, and climatological), Ecuador had between 31 to 60 disaster events (Leonard, 2018).

In the proposed work we have three counties, which were created recently and thus the mortality data was less than the other counties. These counties presented better quality metrics with our approach. Although the Negative Binomial approach is not recommended for small data samples (Hilbe, 2007), the additional information through meteorological and air pollutant covariates helped to obtain improvements in the modelling. For further works under our approach for count data with excess zeros of unknown sources such as the number of deaths, the Hurdle negative binomial model and the Zero-inflated model are the suggested options (Muche et al., 2020).

Hilbe (2007, 2041) explained that the confidence intervals for the negative binomial regression model are narrower as compared to those from a Poisson regression model, thus the uncertainties in fitted values could be lower. For this reason, the likelihood ratio test showed that is more appropriate use BN2 for count data with overdispersion.

For further works, the proposed model could be applied to make short-term mortality predictions using different environmental scenarios. Also, it could be used with environmental predictions and their uncertainty on a specific country, macroclimate region, or county to obtain the prediction of the number of deaths by sex and age.

5. Conclusions

The proposed approach with meteorological and air pollution data at different geographical levels improved the Farrington-like model, which is a classical demographic approach to modeling mortality data as a function of time. The additional information through the meteorological and air pollutant covariates adequately described the great variability present in the mortality data. Thus, it improved the dynamic behavior of the classical demographic modelling. The meteorological and air pollutant data allowed us to identify extreme environmental events, like ENSO and volcanic activity, and their relationship with the most important pikes in the human mortality data of Ecuador. The modeling presents better quality metrics in the aggregated data related to the geographic level. For instance, at the country level, the proposed model presents better quality metrics than on at the county level. The macro-climate regions with a high number of human deaths present better-fitted values using our approach. Additionally, in cases with limited mortality data, the proposed model presented adequate behavior as well. The proposed model could support making short-term projections in the number of human deaths by sex and age in different environmental scenarios.

Acknowledgements

Joseph Sánchez Balseca is the recipient of a full scholarship from the Secretaria de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT), Ecuador. The authors want to thank the research group on Engineering Sciences and Global Development (EScGD) and the Agència de Gestió d'Ajuts Universitaris i de Recerca de la Generalitat de Catalunya (Ref. 2017 SGR 1496).

Appendices

Table A.1. Covariates selection using AIC for mortality data modelling by sex and age range at the country level.

Covariates								AIC
Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	
Country: Female								
x	x	x	x	x	x	x	x	460.53
x	x		x	x	x	x	x	458.71
x	x		x	x	x	x		457.12
x	x		x		x	x		456.92
x			x		x	x		456.31
x					x	x		456.16
Country: Male								
x	x	x	x	x	x	x	x	462.67
x	x	x	x	x	x	x		460.67
x		x	x	x	x	x		459.01
x			x	x	x	x		457.6
x			x		x	x		456.09
x			x			x		455.71
x		x				x		454.3
Country: Female-Age range: A								
x	x	x	x	x	x	x	x	391.06
x	x	x		x	x	x	x	389.13
x	x	x			x	x	x	387.23
x				x	x	x	x	385.32
x				x	x	x		383.48
x					x	x		381.8
Country: Female-Age range: B								
x	x	x	x	x	x	x	x	353.01
x	x	x		x	x	x	x	351.02
x	x	x		x	x		x	349.04
x	x	x		x			x	347.24
x	x			x			x	345.5

x	x						x	345.35
Country: Female-Age range: C								
x	x	x	x	x	x	x	x	421.77
x	x	x	x	x	x	x		420.08
x	x		x	x	x	x		418.72
Country: Male-Age range: A								
x	x	x	x	x	x	x	x	391.54
x	x	x	x	x	x	x		389.55
x	x	x		x	x	x		387.63
x	x	x			x	x		385.67
x	x				x	x		383.9
x	x					x		382.65
x						x		380.87
x								379.27
Country: Male-Age range: B								
x	x	x	x	x	x	x	x	409.87
x	x	x	x		x	x	x	408.75
x	x	x			x	x	x	407.57
x	x	x				x	x	406.21
x	x					x	x	405.09
x	x					x		405.08
Country: Male-Age range: C								
x	x	x	x	x	x	x	x	430.23
x		x	x	x	x	x	x	428.27
x		x	x	x	x	x		426.55
x			x	x	x	x		425.38
x			x		x	x		424.47

Table A.2. Covariates selection using AIC for mortality data modelling by sex and age range at macro-climate region: Coast.

Covariates								AIC
Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	
Coast region: Female-Age range: A								
x	x	x	x	x	x	x	x	347.98
x	x	x	x	x		x	x	345.98
x	x	x	x	x		x		344.12
x	x		x	x		x		342.46
x			x	x		x		341.54
x			x			x		340.57
Coast region: Female-Age range: B								
x	x	x	x	x	x	x	x	330.71
x	x	x	x		x	x	x	328.75
x	x	x	x			x	x	327.31
x	x		x			x	x	325.68
Coast region: Female-Age range: C								
x	x	x	x	x	x	x	x	396.25
x	x	x	x	x	x	x		394.25
x		x	x	x	x	x		392.31
x			x	x	x	x		390.86
x				x	x	x		389.67
x					x	x		388.41
Coast region: Male-Age range: A								
x	x	x	x	x	x	x	x	361.44
x	x	x	x	x		x	x	359.44
x	x	x	x			x	x	357.48
x	x		x			x	x	355.55
x	x		x				x	353.76
Coast region: Male-Age range: B								
x	x	x	x	x	x	x	x	404.63
x	x	x	x		x	x	x	400.66
x	x	x			x	x	x	398.9
x	x				x	x	x	398.35
Coast region: Male-Age range: C								

x	x	x	x	x	x	x	x	393.04
x	x	x	x		x	x	x	391.28
x	x	x	x		x	x		389.83
x		x	x		x	x		388.11
x			x		x	x		386.28
x					x	x		386.25

Table A.3. Covariates selection using AIC for mortality data modelling by sex and age range at macro-climate region: Andean.

Covariates								AIC
Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	
Andean region: Female-Age range: A								
x	x	x	x	x	x	x	x	350.56
x	x	x	x	x	x	x		348.57
x	x		x	x	x	x		346.87
x	x		x	x	x			345.22
x	x		x		x			343.63
x	x		x					343.57
x	x							342.07
Andean region: Female-Age range: B								
x	x	x	x	x	x	x	x	310.27
x	x	x	x	x	x		x	308.64
x	x	x	x		x		x	307.62
x	x	x			x		x	306.27
Andean region: Female-Age range: C								
x	x	x	x	x	x	x	x	387.37
x	x		x	x	x	x	x	385.41
x	x		x		x	x	x	383.55
x	x		x		x	x		381.75
x	x		x			x		380.16
x	x					x		378.75
Andean region: Male-Age range: A								
x	x	x	x	x	x	x	x	354.47
x	x		x	x		x	x	350.74
x	x		x	x			x	348.97
x	x		x	x				347.64
x	x		x					346.8
Andean region: Male-Age range: B								
x	x	x	x	x	x	x	x	340.68
x	x	x	x	x	x	x		338.75
x	x		x	x	x	x		337.08
x	x		x	x	x			335.37
x	x		x		x			334.51
Andean region: Male-Age range: C								
x	x	x	x	x	x	x	x	400.39
x	x	x	x		x	x	x	398.5
x	x		x		x	x	x	396.59
x	x				x	x	x	394.86
x					x	x	x	393.49
x						x	x	392.31
x						x		391.87

Table A.4. Covariates selection using AIC for mortality data modelling by sex and age range at macro-climate region: Amazon.

Covariates								AIC
Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	
Amazon region: Female-Age range: A								
x	x	x	x	x	x	x	x	275.58
x	x	x	x	x		x	x	273.58
x		x	x	x		x	x	271.6
x		x		x		x	x	269.71
x				x		x	x	268.08
x				x			x	266.76
x							x	265.73

Amazon region: Female-Age range: B								
x	x	x	x	x	x	x	x	254.42
x	x	x	x		x	x	x	252.42
x	x	x	x			x	x	250.45
x	x	x	x				x	248.57
x	x	x					x	246.94
x	x	x						245.6
x	x							245.5
Amazon region: Female-Age range: C								
x	x	x	x	x	x	x	x	258.08
x	x	x	x	x		x	x	256.1
x	x	x	x			x	x	254.21
x	x		x			x	x	253.01
x	x		x			x		252.47
x			x			x		251.56
Amazon region: Male-Age range: A								
x	x	x	x	x	x	x	x	288.71
x	x	x	x	x		x	x	286.71
x	x	x		x		x	x	284.79
x	x	x		x		x		282.89
x	x	x				x		281.07
x		x				x		278.18
Amazon region: Male-Age range: B								
x	x	x	x	x	x	x	x	269.51
x	x	x		x	x	x	x	267.51
x	x	x		x	x	x		265.53
x		x		x	x	x		263.79
x		x			x	x		262.83
x					x	x		262
x					x			261.53
Amazon region: Male-Age range: C								
x	x	x	x	x	x	x	x	266.46
x	x	x	x	x	x		x	262.76
x		x			x		x	261.15
x				x	x		x	259.65

Table A.5. Covariates selection using AIC for mortality data modelling by sex and age range at the county level: Los Ríos.

Covariates								AIC
Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	
Los Ríos County: Female-Age range: A								
x	x	x	x	x	x	x	x	258.08
x	x	x	x		x	x	x	256.15
x	x	x	x			x	x	254.28
x	x		x			x	x	253.26
Los Ríos County: Female-Age range: B								
x	x	x	x	x	x	x	x	251.49
x	x		x	x	x	x	x	249.59
x	x			x	x	x	x	247.69
x					x	x	x	246.99
Los Ríos County: Female-Age range: C								
x	x	x	x	x	x	x	x	281.06
x	x		x	x	x	x	x	279.13
x	x		x		x	x	x	277.25
x	x		x		x			276.29
Los Ríos County: Male-Age range: A								
x	x	x	x	x	x	x	x	291.2
x	x	x	x		x	x	x	289.21
x	x	x			x	x	x	287.24
x	x				x	x	x	285.39
x	x					x		283.97
Los Ríos County: Male-Age range: B								
x	x	x	x	x	x	x	x	315.91

Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	AIC
x	x		x	x	x	x	x	314.08
x	x		x		x	x	x	312.39
x	x					x	x	311.51
Los Ríos County: Male-Age range: C								
x	x	x	x	x	x	x	x	293.81
x	x	x	x	x	x		x	292.13
x	x	x	x	x			x	290.6
x	x	x		x			x	289.09
x	x			x			x	287.51
x	x						x	286.41

Table A.6. Covariates selection using AIC for mortality data modelling by sex and age range at the county level: Pichincha.

Covariates								AIC
Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	
Pichincha county: Female-Age range: A								
x	x	x	x	x	x	x	x	316.05
x	x	x	x	x	x	x		314.51
x		x	x	x	x	x		312.82
x		x	x	x	x			311.61
x		x	x		x			310.17
x			x		x			309.96
Pichincha county: Female-Age range: B								
x	x	x	x	x	x	x	x	296.43
x	x	x	x	x	x		x	294.46
x		x	x	x	x		x	292.71
x		x	x	x	x			291.41
x	x		x			x	x	288.92
Pichincha county: Female-Age range: C								
x	x	x	x	x	x	x	x	343.07
x		x	x	x	x	x	x	341.09
x			x	x	x	x		339.68
Pichincha county: Male-Age range: A								
x	x	x	x	x	x	x	x	333.33
x	x		x	x	x	x	x	331.42
x			x	x	x	x	x	329.84
x			x	x	x	x		328.89
Pichincha county: Male-Age range: B								
x	x	x	x	x	x	x	x	326.39
x	x		x	x	x	x	x	324.44
x			x	x	x		x	323.03
Pichincha county: Male-Age range: C								
x	x	x	x	x	x	x	x	354.49
x		x	x	x	x	x	x	353.54
x				x	x	x		352.62

Table A.7. Covariates selection using AIC for mortality data modelling by sex and age range at the county level: Morona Santiago.

Covariates								AIC
Year	T	P	H	Prec.	CO	SO_2	$PM_{2.5}$	
Morona Santiago county: Female-Age range: A								
x	x	x	x	x	x	x	x	200.06
x	x	x	x	x	x	x		198.06
x	x	x		x	x	x		196.07
x	x			x	x	x		194.37
x	x			x		x		193.04
x				x		x		191.9
Morona Santiago county: Female-Age range: B								
x	x	x	x	x	x	x	x	208.95
x	x		x	x	x	x	x	206.96
x	x		x		x	x	x	205.22
x	x		x		x		x	203.46
x	x							202.26

Morona Santiago county: Female-Age range: C								
x	x	x	x	x	x	x	x	213.71
x	x	x	x	x		x	x	211.77
x	x		x	x		x	x	209.91
x	x		x	x		x		208.28
x	x		x			x		207.63
x			x					207.04
Morona Santiago county: Male-Age range: A								
x	x	x	x	x	x	x	x	233.54
x	x		x	x	x	x	x	231.54
x	x		x	x		x	x	229.74
x	x		x	x		x		228.68
x				x	x			228.34
Morona Santiago county: Male-Age range: B								
x	x	x	x	x	x	x	x	209.03
x	x		x	x	x	x	x	207.15
x			x	x	x	x	x	205.81
x			x	x		x	x	204.5
x				x		x	x	203.42
x					x			203.37
Morona Santiago county: Male-Age range: C								
x	x	x	x	x	x	x	x	228.34
x	x	x	x	x	x		x	226.34
x	x	x	x		x		x	224.39
x	x	x	x				x	222.63
x	x	x					x	222.14
x			x					221.85

Bibliography

Alahmad, B., Shakarchi, A., Khraishah, H., Alseaidan, M., Gasana, J., Al-Hemoud, A., Fox, M. (2020). Extreme temperatures and mortality in Kuwait: Who is vulnerable? Science of the Total Environment, 732(139289). https://doi.org/10.1016/j.scitotenv.2020.139289

Analitis, A., Michelozzi, P., D'Ippoliti, D., De'Donato, F., Menne, B., Matthies, F., Katsouyanni, K. (2014). Effects of heat waves on mortality: effect modification and confounding by air pollutants. *Epidemiology*, *25*(1), 15-22. DOI: 10.1097/EDE.0b013e31828ac01b

Armstrong, B. (2006). Models for the Relationship Between Ambient Temperature and Daily Mortality. *Epidemiology, 17*(6), 624-631. DOI: 10.1097/01.ede.0000239732.50999.8f

Barnett, A., Tong, S., Clements, A.C.A. (2010). What measure of temperature is the best predictor of mortality? *Environmental Research, 110*(6), 604-611. https://doi.org/10.1016/j.envres.2010.05.006

Bozikas, A., & Pitselis, G. (2018). An Empirical Study on Stochastic Mortality Modelling under the Age-Period-Cohort Framework: The Case of Greece with Applications to Insurance Pricing. Risks, 6(2), 44. https://doi.org/10.3390/risks6020044

Cameron, C., & Trivedi, P. (2013). *Regression Analysis of Count Data, 2nd edition*. Cambridge University Press.

Chatterjee, A., Gierach, M., Sutton, A., Feely, R., Crisp, D., Eldering, A., Schimel, D. (2017). Influence of El Niño on atmospheric CO_2 over the tropical Pacific Ocean: findings from NASA's OCO-2 mission. *Science, 358*(6360), eaam5776. DOI: 10.1126/science.aam5776

Chen, F., Deng, Z., Deng, Y., Qiao, Z., Lan, L., Meng, Q., Li, X. (2017). Attributable risk of ambient PM10 on daily mortality and years of life lost in Chengdu, China. *Science of the Total Environment, 581-582*, 426-433. https://doi.org/10.1016/j.scitotenv.2016.12.151

Cruz, M., Awruch, C., Gilardoni, C., Demetrio, M., & Palacios, M.-G. (2020). Immunity and health of two wild marine fishes naturally exposed to anthropogenic pollution. *Science of the Total Environment, 726*(138303).

Currie, I., & Djeundje, V. (2010). Smoothing dispersed counts with applications to mortality data. *Annals of Actuarial Science, 5*(1), 33-52. DOI:10.1017/S1748499510000047

Cutter, S. (2017). The forgotten casualties redux: Women, children, and disaster risk. *Global Environmental Change*, *42*, 117-121. https://doi.org/10.1016/j.gloenvcha.2016.12.010

Czaja, C., Miller, L., & Colbor, K. (2020). State-level estimates of excess hospitalizations and deaths associated with influenza. *Influenza and other respiratory viruses, 14*(2), 111-121. https://doi.org/10.1111/irv.12700

Edlund, S., Kaufman, J., Lessler, J., Douglas, J., Bromberg, M., Kaufman, Z., Leventhal, A. (2011). Comparing three basic models for seasonal influenza. *Epidemics, 3*(3-4), 135-142. https://doi.org/10.1016/j.epidem.2011.04.002

Eguiguren-Velepucha, P. A., Maita Chamba, J. A., Aguirre Mendoza, N. A., Ojeda-Luna, T. L., Samaniego-Rojas, N. S., Carol Howe, M. J., & Aguirre Mendoza, Z. H. (2016). Tropical ecosystems vulnerability to climate change in southern Ecuador. *Tropical Conservation Science, 9*(4), 1-17. https://doi.org/10.1177/1940082916668007

Espinoza, J.-C., Ronchail, J., Guyot, J.-L., Cochonneau, G., Naziano, F., Lavado, W., Philippe, V. (2009). Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador). *International Journal of Clymatology*, 1574-1594. https://doi.org/10.1002/joc.1791

Estrella, B., Sempértegui, F., Franco, O., Cepeda, M., & Naumova, E. (2019). Air pollution control and the occurrence of acute respiratory illness in school children of Quito, Ecuador. *Journal of Public Health Policy*, *40*, 17-34. https://doi.org/10.1057/s41271-018-0148-6

Farrington, C., Andrews, N., Beale, A., & Catchpole, M. (1996). A Statistical Algorithm for the Early Detection of Outbreaks of Infectious Disease. *Journal of the Royal Statistical Society. Series A, 159*(3), 547-563. DOI: 10.2307/2983331

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 997–1016. https://doi.org/10.1007/s11222-013-9416-2

Guo, B., Sun, C., Fan, C., Ma, W., Zhang, H., Qiao, X., Zhao, X. (2019). Using spatio-temporal lagged association pattern to unravel the acute effect of air pollution on mortality. *Science of the Total Environment, 664,* 99-106. https://doi.org/10.1016/j.scitotenv.2019.02.018

Hilbe, J. (2007). *Negative Binomial Regression.* Arizona: Cambridge University Press.

Hilbe, J. (2014). *Modeling Count Data.* Arizona: Cambridge University Press.

Holford, T. (1983). The Estimation of Age, Period and Cohort Effects for Vital Rates. *Biometrics, 39,* 311-324. DOI: 10.2307/2531004

Hyun, H., Prasad, R., Maquiling, A., & Smith-Doiron, M. (2020). Temporal trends in associations between ozone and circulatory mortality in age and sex in Canada during 1984–2012. *Science of the Total Environment, 724*(137944). https://doi.org/10.1016/j.scitotenv.2020.137944

Instituto Nacional de Estadística y Censos. (2001). *Estudios Demográficos en Profundidad: La Mortalidad en el Ecuador en el período 1990-2001*. Quito: INEC.

Koissi, M.-C., Shapiro, A., & Högnäs, G. (2006). Evaluating and extending the Lee–Carter model for mortality forecasting: Bootstrap confidence interval. *Insurance:* Mathematics and Economics, 38(1), 1-20. https://doi.org/10.1016/j.insmatheco.2005.06.008

Kuo, C.-L. (2017). Assessments of Ali, Dome A, and Summit Camp for mm-wave Observations Using MERRA-2 Reanalysis. *The Astrophysical Journal, 848*(1), 1-11. DOI: 10.3847/1538-4357/aa8b74

Leonard, P. (2018). *Situación de las estadísticas e indicadores de eventos extremos y desastres.* Buenos Aires: Comisión Económica para América Latina y el Caribe (CEPAL).

Li, J. S.-H., Hardy, M., & Tan, K. (2009). Uncertainty in Mortality Forecasting: An Extension to the Classical Lee-Carter Approach. *ASTIN Bulletin, 39*(1), 137-164. https://doi.org/10.2143/AST.39.1.2038060

Liang, H., Qiu, H., & Tian, L. (2018). Short-term effects of fine particulate matter on acute myocardial infraction mortality and years of life lost: A time series study in Hong Kong. *Science of the Total Environment, 615*, 558–563. https://doi.org/10.1016/j.scitotenv.2017.09.266

Liddle, B. (2011). Consumption-driven environmental impact and age structure change in OECD countries: A cointegration-STIRPAT analysis. *Demographic Research, 24*(30), 749-770. DOI: 10.4054/DemRes.2011.24.30

López-Cuadrado, T., de Mateo, S., Jiménez-Jorge, S., Savulescu, C., & Larrauri, A. (2012). Influenza-related mortality in Spain, 1999-2005. *Gaceta Sanitaria, 26*(4), 325-329. http://dx.doi.org/10.1016/j.gaceta.2011.09.033

Lucas, P., Hilderink, H., Janssen, P., KC, S., van Vuuren, D., Niessen, L. (2019). Future impacts of environmental factors on achieving the SDG target on child mortality- A synergistic assessment. Global Environmental Change, 57, 101925. https://doi.org/10.1016/j.gloenvcha.2019.05.009

Matus, K., Nam, K., Selin, N., Lamsal, L., Reilly, J., Paltsev, S. (2012). Health damages from air pollution in China. *Global Environmental Change*, 22(1), 55-66. https://doi.org/10.1016/j.gloenvcha.2011.08.006

Martínez, J., Vega-Garcia, C., & Chuvieco, E. (2009). Human-caused wildfire risk rating for prevention planning in Spain. *Journal of Environmental Management, 90*(2), 1241-1252. https://doi.org/10.1016/j.jenvman.2008.07.005

Ministry of the Environment of Ecuador (2000). First National Communication about Climate Change. Quito: MAE.

McCuen, R. H., Knight, Z., & Cutter, A. G. (2006). Evaluation of the Nash–Sutcliffe Efficiency Index. *Journal of Hydrologic Engineering, 11*(6). https://doi.org/10.1061/(ASCE)1084-0699(2006)11:6(597)

Morán-Tejada, E., Bazo, J., López-Moreno, J., Aguilar, E., Azorín-Molina, C., Sánchez-Lorenzo, A., Vicente-Serrano, S. (2016). Climate trends and variability in Ecuador (1966–2011). *International Journal of Climatology, 36*(11), 3839-3855. https://doi.org/10.1002/joc.4597

Muche, S., Mekonnen, H., & Meseret, G. (2020). The best statistical model to estimate predictors of under-five mortality in Ethiopia. *Journal of Big Data, 7*(63). https://doi.org/10.1186/s40537-020-00339-0

Neira, M., & Prüss-Ustün, A. (2016). Preventing disease through healthy environments: A global assessment of the environmental burden of disease. *Toxicology Letters, 259*(10), S1. https://doi.org/10.1016/j.toxlet.2016.07.028

Neves, C., Fernandes, C., & Hoeltgebaum, H. (2017). Five different distributions for the Lee–Carter model of mortality forecasting: A comparison using GAS models. *Insurance: Mathematics and Economics, 75*, 48-57. https://doi.org/10.1016/j.insmatheco.2017.04.004

Nunes, A., Lourenço, L., & Castro-Meira, A. (2016). Exploring spatial patterns and drivers of forest fires in Portugal (1980–2014). *Science of the Total Environment, 573*, 1190-1202. https://doi.org/10.1016/j.scitotenv.2016.03.121

Pan American Health Organization. (2017). *Lineamientos básicos para el análisis de la mortalidad.* Washington D.C.: OPS.

Pan American Health Organization. (2005). *Guía de preparativos de salud frente a erupciones volcánicas*. Quito: PAHO.

Park, A., Hong, Y., & Kim, H. (2011). Effect of changes in season and temperature on mortality associated with air pollution in Seoul, Korea. *J Epidemiol Community Health*, *65*(4), 368-375. http://dx.doi.org/10.1136/jech.2009.089896

Quevedo, O. (2015). Reference affectation of SO_2 in the mangrove ecosystem of the port of Guayaquil. *Alternativas, 16*(3), 62-66.

Qin, W., Zhang, Y., Chen, J., Yu, Q., Cheng, S., Li, W., Tian, H. (2018). Variation, sources and historical trend of black carbon in Beijing, China based on ground observation and MERRA-2 reanalysis data. *Environmental Pollution*, 245, 853-863. https://doi.org/10.1016/j.envpol.2018.11.063

Rao, S., Pachauri, S., Dentener, F., Kinney, P., Klimont, Z., Riahi, K., Schoepp, W. (2013). Better air for better health: Forging synergies in policies for energy access, climate change and air pollution. *Global Environmental Change*, 23(5), 1122-1130. https://doi.org/10.1016/j.gloenvcha.2013.05.003

Ripley, B., Venables, B., Bates, D., Hornik, K., Gebhardt, A., Firth D. (2020). *Support Functions and Datasets for Venables and Riple's MASS*. Retrieved October 2020, from cran: http://www.stats.ox.ac.uk/pub/MASS4/

Rowlinson, M. J., Rap, A., Arnold, S. R., Pope, R. J., Chipperfield, M. P., McNorton, J., Forster, P., Gordon, H., Pringle, K. J., Feng, W., Kerridge, B. J., Latter, B. L., and Siddans, R. (2019). Impact of El Niño–Southern Oscillation on the interannual variability of methane and tropospheric ozone. *Atmos. Chem. Phys.*, 19, 8669–8686, https://doi.org/10.5194/acp-19-8669-2019.

Saini, P., & Sharma, M. (2020). Cause and Age-specific premature mortality attributable to $PM_{2.5}$ Exposure: An analysis for Million-Plus Indian cities. *Science of The Total Environment, 710*, 135230. https://doi.org/10.1016/j.scitotenv.2019.135230

Spezzano, P. (2021). Mapping the susceptibility of UNESCO World Cultural Heritage sites in Europe to ambient (outdoor) air pollution. *Science of The Total Environment, 754*(142345). https://doi.org/10.1016/j.scitotenv.2020.142345

Siu-Hang, L., & Wai-Sum, C. (2013). The Lee-Carter Model for Forecasting Mortality, Revisited. *North American Actuarial Journal, 11*(1), 68-69. https://doi.org/10.1080/10920277.2007.10597438

Tran, H., Kim, J., Kim, D., Choi, M., & Choi, M. (2018). Impact of air pollution on cause-specific mortality in Korea: Results from Bayesian Model Averaging and Principle Component Regression approaches. *Science of The Total Environment, 636*, 1020-1031. https://doi.org/10.1016/j.scitotenv.2018.04.273

Tsekeri, E., Kolokotsa, D., & Santamouris, M. (2020). On the association of ambient temperature and elderly mortality in a Mediterranean island - Crete. *Science of the Total Environment, 738*(139843). https://doi.org/10.1016/j.scitotenv.2020.139843

Universite catholique de Louvain (UCL). (2018). *The Emergency Events Database*. Retrieved September 2020, from EMDAT: http://www.emdat.be

Ver-Hoef, J., & Boveng, P. (2007). Quasi-poisson vs. Negative binomial regression: how should we model overdispersed count data? *Ecology, 88*(11), 27662772.

Viner, R., Hargreaves, D., Coffey, C., Patton, G., & Wolfe, I. (2014). Deaths in young people aged 0–24 years in the UK compared with the EU15+ countries, 1970–2008: analysis of the WHO Mortality Database. The Lancet, 384(9946), 880-892. https://doi.org/10.1016/S0140-6736(14)60485-2

Wang, N., Mengersen, K., Tong, S., Kimlin, M., Zhou, M., Liu, Y., & Hu, W. (2020). County-level variation in the long-term association between $PM_{2.5}$ and lung cancer mortality in China. *Science of the Total Environment, 738*(140195). https://doi.org/10.1016/j.scitotenv.2020.140195

World Health Organization. (2000). *Air Quality Guidelines*. Copenhagen: WHO Press.

World Health Organization. (2006). *WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide*. Switzerland: WHO Press.

Yanagisawa, S. (1973). Air Quality Standards National And International. *Journal of the Air Pollution Control Association, 23*(11), 945-948. https://doi.org/10.1080/00022470.1973.10469863

Zalakeviciute, R., Vasquez, R., Bayas, D., Buenano, A., Mejia, D., Zegarra, R., Diaz, A. and Lamb, B. (2020). Drastic Improvements in Air Quality in Ecuador during the COVID-19 Outbreak. *Aerosol Air Qual. Res.* 20: 1783–1792. https://doi.org/10.4209/aaqr.2020.05.0254

Zeileis, A. (2020). *Testing Linear Regression Models*. Retrieved October 2020, from cran: https://www.rdocumentation.org/packages/lmtest/versions/0.9-38

Zhang, P., & Zhou, X. (2020). Health and economic impacts of particulate matter pollution on hospital admissions for mental disorders in Chengdu, Southwestern China. *Science of the Total Environment, 733*(139114). https://doi.org/10.1016/j.scitotenv.2020.139114

Zhou, H., Burkom, H., Strine, T., Katz, S., Jajosky, R., Anderson, W., & Ajani, U. (2017). Comparing the historical limits method with regression models for weekly monitoring of national notifiable diseases reports. Journal of Biomedical Informatics, 34-40. https://doi.org/10.1016/j.jbi.2017.10.010