Longevity pattern in Emilia Romagna (Italy) in a dynamic perspective

G. Roli[a], R. Miglio[a], R. Rettaroli[a], A. Samoggia[a]

[a]*Department of Statistical Sciences "P.Fortunati", University of Bologna, Italy*

Abstract

In this paper, we investigate the pattern of longevity during the last 15 years in Emilia Romagna, a North-Eastern region of Italy, at a municipality level. We consider a specific index of extreme longevity based on people aged 95 and over in two different periods (1995-1999 and 2005-2009). Spatio-temporal modeling is used to tackle at both periods the random variations in the occurrence of people 95+, due to the increasingly rareness of such events, especially in small areas. This method exploits the spatial proximity and the consequent interaction of the geographical areas to smooth the observations, as well as to control for the effects of a set of regressors. As a result, clusters of areas characterized by high and low indexes of longevity are well identified and the temporal evolution of the phenomenon can be depicted. In a parallel analysis, we consider the past levels of mortality on the same cohort of individuals reaching 95 years and over in the second period and when they were aged 80-89 and 90-99. Within this longitudinal framework, the longevity outcome is modeled by a spatial regression. The area-specific structures of mortality are included as regressors, whose effects represent the causal link between the occurrence of people 95+ and the causes of death in the same cohort.

Keywords: centenarian rate, hierarchical Bayesian models, longevity, small areas, spatial and spatio-temporal models

1. Introduction

In the last decades, the study of human longevity and its development has drawn the attention of researchers belonging to different fields of analysis. Indeed, it is widely recognized that the length of life is influenced by a combination of genetic and environmental factors. Various studies performed in different Italian regions (Sardinia, Calabria, Sicily and Emilia Romagna) showed the presence of specific areas where the prevalence of oldest-old people is higher than elsewhere. For instance, a definite geographical area in Sardinia is characterized by an exceptional male longevity (Poulain et al., 2004), as well as a low female/male centenarians ratio; a significant negative correlation between surname abundance and index of longevity has been detected in Calabria (Southern Italy) where some isolated areas of male longevity present a high level of inbreeding (Montesanto et al., 2008); finally, in Emilia Romagna some longevity “clusters” have been identified and their persistence has been detected by comparing the results of different spatial scan statistic methods (Miglio et al., 2008).

The explanatory analysis of disparities in the frequencies of the oldest-old population reminds of environmental and genetics features differently spread at a geographical level. In order to deepen the several aspects of this complex phenomenon, one of the scientific approaches aims at mapping the geographical diffusion of extreme longevity using methods of spatial analysis techniques in order to identify areas, or clusters of areas, characterized by particularly high, or low, concentration of oldest-old population.

When the geographical analysis is performed on a fine territorial scale and the phenomenon under study is characterized by a scarce number of units, the territorial distribution of cases may be strongly influenced by random variation due to the sample data. This invalidates both the distribution of the cases in a given time point and its dynamic evolution in terms of comparison among successive time points. A recent solution adopted to deal with these problems consists in the use of spatial and spatio-temporal models which involve the geographical and time interactions resulting in a smoothing effect among the observations. These methods can be developed both within a frequentist framework (see for example Breslow and Clayton, 1993; Langford et al., 1999) and under a fully Bayesian perspective (e.g., Bernardinelli and Montomoli, 1992; Mollie, 1994; Banrjee, Carlin and Gelfand, 2004; Lawson, 2009). A great advance in the use of Bayesian methodology has been seen since the development of computational methods, such as Monte Carlo Markov Chain (MCMC) techniques together with Gibbs sampling or Metropolis-Hastings algorithm (Gelman et al., 2003; Carlin and Louis, 1998). This has led to a rapid increase in the use of Bayesian approaches also in the analysis of the geographical distribution of health data as feasible methods to fit almost any model involving multiple levels, random effects and complex dependence structures. Moreover, this kind of modeling allows to embed different levels of covariates aiming to explain the variability among the areas.

In this paper, we use a Bayesian spatio-temporal model to tackle both the geographical structure and the temporal dimension of longevity over the last 15 years in Emilia Romagna. In particular, we consider a relevant longevity indicator based on the numbers of people aged 95 years and over in two different periods (1995-1999 and 2005-2009), separately by municipality and gender. Some areal futures are further collected and included as covariates into the analysis. The model we employ allows to control for variations due to random occurrences, especially for small municipalities which have low numbers. The main purposes are to identify territorial groups of areas characterized by high or low levels of longevity and to investigate the development of these clusters by time.

Besides being influenced by spatial interactions and areal features, we suppose the variations in the measures of longevity across the areas can be partially explained by the different patterns of mortality. In order to quantify these contributions, we focus the attention on the most recent period (i.e., 2005-2009). A cohort perspective is adopted to depict the development of mortality for the same individuals who reach 95 years of age and over in the period under investigation. In detail, we calculate the Standardized Mortality Ratios (SMR) for each municipality according to the most considerable causes of death when the cohort was aged 80 to 89 and 90 to 99 years, separately. Within this longitudinal framework, the longevity outcome is modeled by a Bayesian spatial regression where the area-specific levels of mortality for each group of age in the cohort are included as regressors. The hierarchical structure of the model further allows to take into account for correlations among the mortality covariates, as well as spatial interactions. We further control for the effects of the areal features introduced above.

The paper develops as follows. We firstly describe the regional area under study and corresponding data and indicators. Section 3 defines the statistical analysis. The results are summarized and discussed in Section 4. The last section reports the main conclusions.

2. Area, data and indicators

Emilia Romagna is a North-Eastern Italian region (Figure 1) which shows one of the oldest age structures in Europe (with 22.5% persons aged 65+, and

6.9% persons aged 80+ in 2009 related to a total population of $4,337,966$). The region is characterized by a great geographical variability: mountains, hills and a wide flat land, with a consequent heterogeneity in environmental context, social conditions and economic resources. Emilia Romagna is split up into nine provinces (Piacenza, Parma, Reggio Emilia, Modena, Bologna, Ferrara, Ravenna, Forl-Cesena, Rimini) and 341 municipalities (Figure 1).

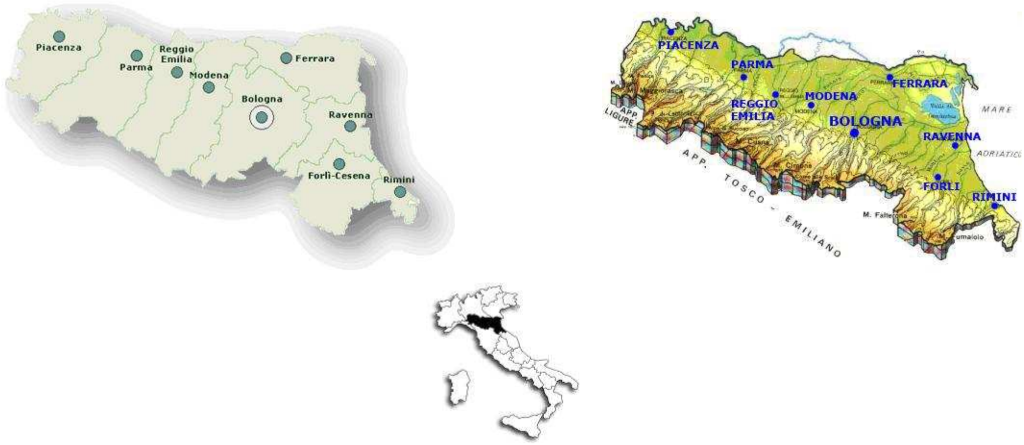


Figure 1: Political and physical map of Emilia Romagna.

In order to measure the different spread of longevity in the regional area, we use a modified version of the Centenarian Rate (CR), as firstly proposed by Robine and Caselli, 2005 and Robine et al., 2006[1]. In particular, we consider people aged 95 and over, denoted by P^{95+}, instead of centenarians,

[1]The CR has been showed to be an appropriate indicator of longevity as it takes into account for the effects of the work-related migration. Indeed, it is well known that in Italy, including Emilia Romagna region, migration was very common in the recent past, especially during working ages and among people who now belong to the oldest-old population. Under this perspective, the CR indicator measures and compares the dimensions of longevity in different areas as the ratio of the observed number of centenarians to the number of the cohort survivors 40 years before the current observation and not to the number of birth in the corresponding generation as usual. As a result, the CR removes the unknown influence of the migration process, which is assumed to be negligible only after the 60 years of age. Moreover, it is independent from the size of birth cohorts, infant mortality, past migrations, and policies of naturalization (Robine and Caselli, 2005; Robine and Paccaud, 2005).

separately by municipality area and sex, in order to avoid inconsistency or lack of data. Our indicator of longevity, now denoted by CR^{95+}, is thus obtained by dividing this count by the number of 55-64 years old persons living in the same area 40 years earlier (P^{55-64}), according to the following formula:

$$CR^{95+}_{kit} = \frac{P^{95+}_{kit}}{P^{55-64}_{ki(t-40)}} \tag{1}$$

where t denotes the current time under study, k is the gender and i is the municipality. More precisely, we draw the denominator of CR^{95+} from the nearest available Italian census, with analogous considerations to identify the cohort age group. In this framework, we assume that the individuals resident in an area 40 years before the time point t under study are the population exposed to the "risk" of becoming P^{95+} or, similarly, that P^{95+} observed at time t comes from the cohort of individuals resident 40 years earlier in the same area.

To depict the temporal dimension of the phenomenon, we calculate the CR^{95+} separately by sex and municipality area in two periods: 1995-1999 and 2005-2009[2]. The corresponding denominators refer to the same cohorts of individuals at the censuses 1961 and 1971, respectively.

Additional information on the features of the municipalities are also collected. In detail, we consider the classification of areas with respect to the altimetry zone (mountain, hill, coastal hill, plain) and population density (expressed as people per square kilometer). These are then combined to form the following groups[3]: mountain, hill and density lower than 78 people per km^2, hill and density between 78 and 198, hill and density greater than 198, coastal hill, plain and density lower than 78, plain and density between 78 and 198, plain and density greater than 198 (Figure 2).

For the cohort of P^{95+} in the period 2005-2009, we focus on the corresponding patterns of mortality measured when they were aged between 80 and 89 and between 90 and 99, separately. In particular, for each municipality we consider death counts by gender and cause of death according to the two age groups of the cohort. These data are averaged over 5 calendar years, corresponding to 1990-1994 for the age group 80-89 and 2000-2004 for 90-99.

[2]We introduce multiple-years aggregation of data to avoid the drawback of scarcity or lack of occurrences in small areas.

[3]Tercile subsets on the distribution of the population density are considered.

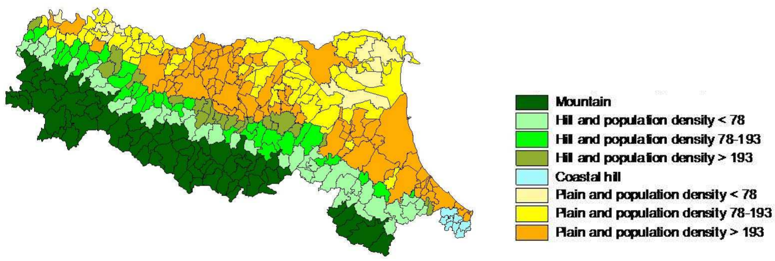


Figure 2: Altimetry and population density (people per km^2) distribution in Emilia Romagna.

Moreover, data are grouped into 2 categories of causes of death according to the International Classification of Diseases (ICD), Revision IX (ISTAT, 1997): malignant neoplasms and diseases of the circulatory system. These causes jointly represent the most common causes of death for people aged 80 and over (approximately 70% of the cases excluding injuries and poisonings). As a result, their different contribution to longevity is certainly an interesting topic to be investigated.

The structure of mortality depicted by the death counts can be appropriately summarized by the relative risk of death. An estimate of this measure is represented by the Standardized Mortality Ratio (SMR), defined as the ratio between the observed O and the expected E death counts with respect to any specific territorial unit. The SMR allows us to control for the influence of the age structure within the age groups, making possible the comparison of different mortality levels among the municipalities. In our case, the SMR indexes with respect to the cohort P^{95+} in the period 2005-2009 are computed for each gender k, municipality i, age group a=x-$x+9$ and cause of death d (with $i = 1, \ldots, 341$) as follows

$$SMR_{kiad} = \frac{O_{kiad}}{E_{kiad}}. \tag{2}$$

In particular, the expected death counts E_{kiad} are calculated, separately by sex and cause of death, through the average mortality rate m_{kad} for the overall region in the age group a and multiplied by the population in area i with ages in a (*internal standardization*).

Then, the SMRs are grouped into 3 levels of mortality based on their quintile distribution, separately by cause of death and age group. The first quintile subset identifies the municipalities with low mortality levels according to the regional average. Conversely, areas characterized by high mortality corresponds to the last quintile subset. Finally, remaining areas are included into the last group with values of SMR around 1, i.e., close to the overall regional rate.

The sources of the data we referred above are all official statistics published by the Italian institutional agency for statistical data collection (ISTAT) and by the regional authorized agency of Emilia Romagna.

3. Methods

3.1. Spatio-temporal model

We study the space-time pattern of “risk” of becoming P^{95+} among the 341 municipalities of Emilia Romagna by adopting a hierarchical spatio-temporal approach. This method allows to deal with the random variation due to the increasingly rareness of such events, by exploiting the spatial proximity and the consequent interaction of the geographical areas. We further associate a temporal dimension to the phenomenon, by considering the evolution of P^{95+} in the two periods 1995-1999 and 2005-2009, separately by gender. As a final result, spatio-temporal models can provide more consistent estimates on quantities of interest and inferences. Indeed, the consequent two-fold smoothing effect in both spatial and temporal sense allows to control for the variation in the population size across the geographical areas (Clayton and Kaldor, 1987) and for extra-variation among the units (Breslow, 1984).

In detail, we employ a hierarchical Bayesian model for areal data by gender, separately. With respect to the i-th geographical area (i.e., the 341 municipalities of Emilia Romagna), we assume the observable P^{95+} at each time period t (1995-1999 and 2005-2009), denoted by y_{it}, are Poisson distributed with means $p_{it}\theta_{it}$. In this formulation, p_{it} represents the potential P^{95+} and θ_{it} is the estimate of CR^{95+} in the required location and period. Then, we follow the conventional log-linear formulation on the rate θ_{it} and we allow for the possibility of different components that additively contribute to explain the spatio-temporal distribution of these rates. Several models for the log rate θ_{it} have been compared and assessed in terms of model complexity by evaluating the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2003)[4]. The resulting model we firstly assume is the following

$$log\,(\theta_{it}) = \alpha_t + u_i + v_i \tag{3}$$

where α_t represents the time-varying intercepts and u_i and v_i are the correlated and uncorrelated spatial heterogeneity, respectively, which are both assumed to be constant in time.

Under a fully Bayesian framework, we specify a prior distribution for each parameter involved in (3). For the correlated spatial component u_i, we assume a Gaussian conditionally autoregressive (CAR) model (Best et al., 1999, Banrjee, Carlin and Gelfand, 2004) of the form

$$p(u_1, \ldots, u_{341}|\tau_u) \propto exp\left\{-\frac{\tau_u}{2}\sum_{i \neq j} w_{ij}(u_i - u_j)^2\right\} \tag{4}$$

where τ_u represents the precision parameter and w_{ij} are the adjacency weights across the areas. Although improper, the CAR prior leads to a posterior distribution which is proper, allowing the Bayesian inferences still proceed.

The random effects v_i which capture the region-wide heterogeneity are supposed to follow an ordinary exchangeable normal prior

$$v_i \overset{iid}{\sim} N\left(0, \frac{1}{\tau_v}\right) \tag{5}$$

where τ_v is the precision term.

Then, we need to specify (hyper-)prior distributions for the precision parameters τ_u and τ_v. In particular, we consider vague and proper distributions which are also "fair", i.e., yield the proportion of the variability due to the spatial homogeneity to be $\approx \frac{1}{2}$ *a priori*. As suggested by Best et al., 1999, we use the following Gamma priors

$$\tau_u \sim Gamma\,(0.1, 0.1) \tag{6}$$

$$\tau_v \sim Gamma\,(0.001, 0.001)\,. \tag{7}$$

In order to explain the differences in the presence of P^{95+} across the areas, some covariates are included into equation (3). In particular, we control for

[4]As an example of these different model specifications, we considered the presence of a spatio-temporal interaction term or time-varying correlated and uncorrelated spatial effects.

some areal features and consider the 8 categories of altimetry and population density as introduced in section 2 .

As a result, the model with covariates for the logarithm of the rate θ_{it} can be analytically expressed as follows:

$$log\,(\theta_{it}) = \alpha_t + \beta x_i + u_i + v_i \tag{8}$$

where x_i is the altimetry and population density group of the i-th municipality and β is the corresponding effect on the log rate. For this covariate effect we assign a vague proper prior.

3.2. Cohort analysis

A parallel analysis is carried out following a longitudinal perspective. We attempt to investigate the effect of the past structure of mortality for people aged 80-89 and 90-99 on the same cohort of individuals who reach 95 years of age and over in the more recent period 2005-2009.

With this aim, a spatial model is used separately by gender. The outcome Y now represents the count of P^{95+} in years 2005 to 2009, lonely. For each municipality i, the y_i are again assumed to follow a Poisson distribution of parameter $p_i\theta_i$, with analogous meanings to those described in the previous section. The log-linear model on the rate θ_i involves the spatial structured (u_i) and unstructured (v_i) components, besides controlling for the effects altimetry and population density (x_i). In addition, we include information on the structure of mortality at ages 80-89 and 90-99 for malignant neoplasms and diseases of circulatory system, separately, based on the 3 levels of SMR as defined in section 2. These quantities are denoted by z_{iad}, where a is the age group of the cohort and d is the specific cause of death. Thus, the model can be formalized as follows:

$$log\,(\theta_i) = \alpha_0 + \beta x_i + \sum_{a,d} \gamma_{ad} z_{iad} + u_i + v_i \tag{9}$$

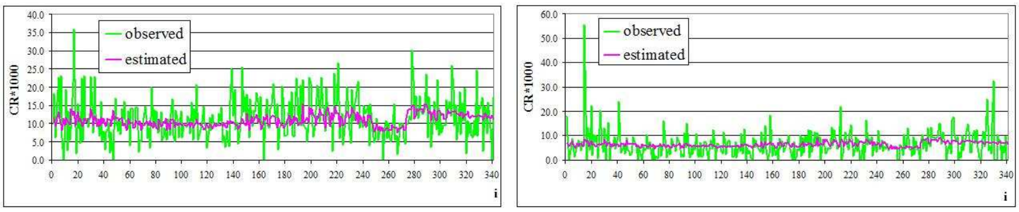
where the estimation of parameters γ_{ad} now represents our main objective.

An additional correlation structure can be further introduced among different causes of death within the observations collected in the same time period. An additional correlation can further be supposed among different causes of death within the same time of observation. Both these correlations are controlled by modeling the priors on the effects γ_{ad} through a noninformative multivariate normal distribution with unknown population mean vector and a variance covariance matrix Σ. The inverse of Σ corresponds to a precision matrix Ω, which is then assumed to follow a vague Wishart distribution (Spiegelhalter *et al.*, 2002). For the other parameters, we assign prior distributions analogous to those specified in subsection 3.1.

4. Results and discussion

4.1. Space-time development

The application of the hierarchical spatio-temporal model introduced above offers the opportunity to investigate different aspects of the longevity pattern, as well as control for variations due to random occurrences. The last feature is pursued through the so-called 'smoothing effect', which can be in practice highlighted by comparing the observed with the estimated values of the CRs^{95+} for each areas. Figure 3 plots these numbers in the first and second period for males in the 341 municipalities of Emilia Rimagna, but the same findings can be drawn out for the female group. The observed CRs^{95+} show a great variability with peaks of extremely high and extremely low numbers, while the CRs^{95+} estimated by our model are closer to each other (i.e., to the regional average). All the results we report for the CRs^{95+} are referred to number of 1000 individuals exposed to "risk".

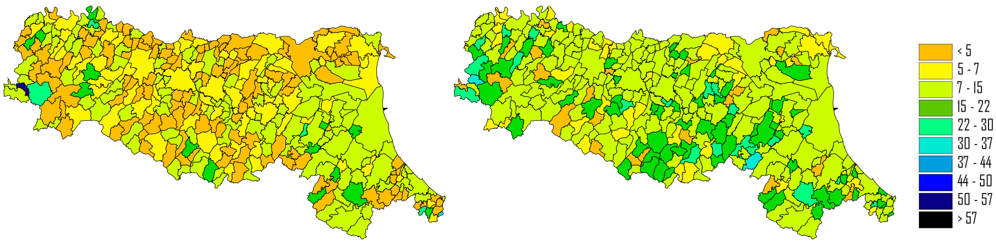


(a) 1995-1999.

(b) 2005-2009.

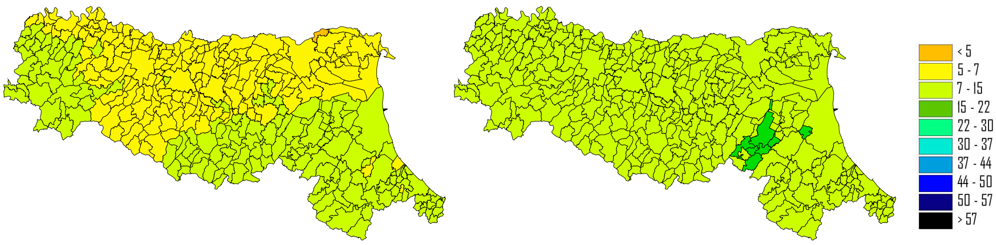
Figure 3: CRs^{95+} ($*1000$) observed and estimated by the model. Men.

The smoothing effect, which finally reflects the efficiency of such a model, is also evident when comparing observed and estimated CRs^{95+} maps (Figures 4-7). For both sexes, the model yields more homogeneous estimates according to the geographical proximity of areas. The observed CRs^{95+} are



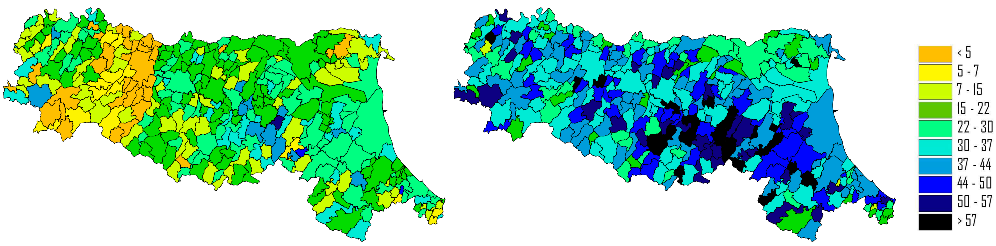
(a) 1995-1999. (b) 2005-2009. (c)

Figure 4: Observed CRs^{95+} ($*1000$). Men.



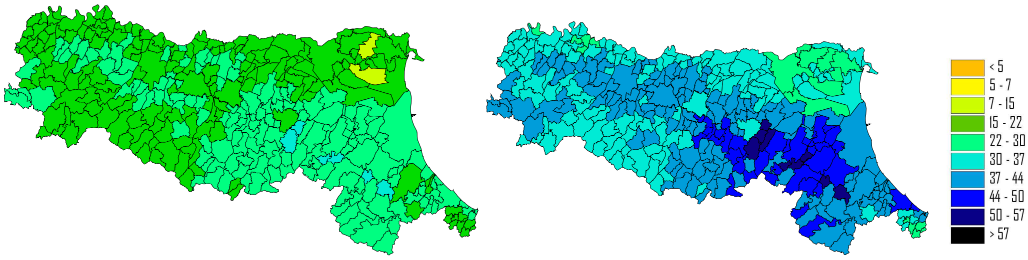
(a) 1995-1999. (b) 2005-2009. (c)

Figure 5: CRs^{95+} ($*1000$) estimated by the Bayesian spatio-temporal model. Men.



(a) 1995-1999. (b) 2005-2009. (c)

Figure 6: Observed CRs^{95+}. Women.

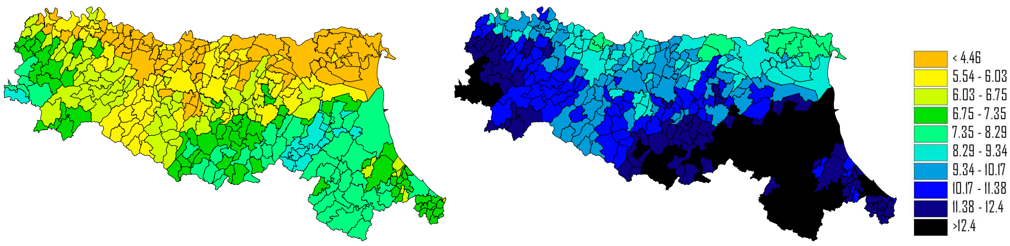


(a) 1995-1999. (b) 2005-2009. (c)

Figure 7: CRs^{95+} ($*1000$) estimated by the Bayesian spatio-temporal model. Women.

instead denoted by a set of spots, identifying locations with strange different values from the other nearby areas.

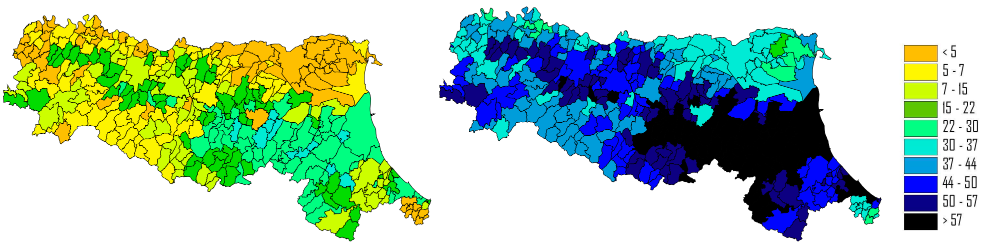
The comparison of the estimated CRs^{95+} across time shows a great increase in the numbers of P^{95+}, especially for women. The males are characterized by lower values with a small range of variability among the areas. Therefore, in order to evaluate the different territorial contributions to the CRs^{95+}, separately by sex, we further consider the ranking of the municipalities according to their decile distribution in both periods (Figures 8 and 9).



(a) 1995-1999. (b) 2005-2009. (c)

Figure 8: Estimated CRs^{95+} ($*1000$) (deciles). Men.

There is a notable persistence of the areas of lower and higher occurrences of P^{95+}, in spite of the rise in the CRs^{95+} during time. These results show



(a) 1995-1999. (b) 2005-2009. (c)

Figure 9: Estimated CRs^{95+} ($*1000$) (deciles). Women.

mean and median values higher than the regional ones in the municipalities belonging to the provinces of Ravenna, Bologna and Forli-Cesena, spreading out in the Adriatic coast, at one side, and in the Apennines municipalities of Modena, on the other. Conversely, the municipalities of the province of Ferrara are characterized by a lower propensity to longevity. Especially for men, some areas of the province of Piacenza still stand out for high values of CR^{95+}.

The spatio-temporal models here adopted further allows us to split the spatial variability into two parts. The first concerns the geographical structure modeled by the component $\mathbf{u}$. It detects the extra-Poisson variability in the log rate "that varies locally, so that nearby regions will have more similar rates" (Banerjee, Carlin, and Gelfand, 2004), and is revealed by yielding homogeneous clusters of adjacent areas. The second part of variability is due to the peculiarities of individual areas and results from municipalities with observed values significantly different from the other nearby areas. This is reflected by the component $\mathbf{v}$ in the model specification and captures the overdispersion over the whole region.

The contribution of the clustering and heterogeneity effects in terms of total territorial variability can be firstly evaluated through the estimates of the posterior variance of the two components $\mathbf{u}$ and $\mathbf{v}$, given respectively by $\widehat{Var}(U|Y)$ and $\widehat{Var}(V|Y)$ (Table 1) and of the posterior proportion of the variability in the random effects that is due to clustering, i.e.

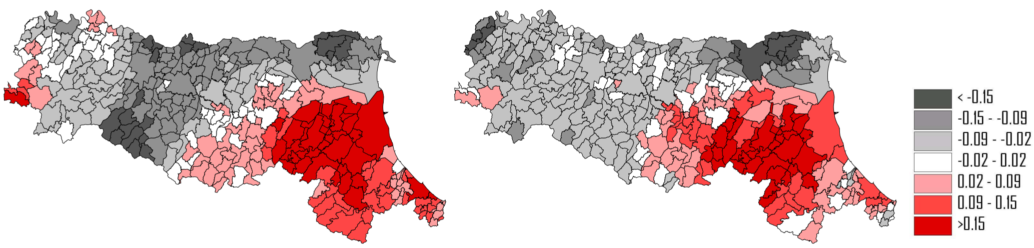
$$\delta = \frac{sd(\mathbf{u})}{sd(\mathbf{u}) + sd(\mathbf{v})} \tag{10}$$

where $sd(\cdot)$ is the empirical marginal standard deviation function. For both sexes, the structured spatial variability seems to prevail. As a consequence, the global model representations for the CRs^{95+} are more influenced by the territorial clusters with relatively similar longevity risks, rather then by the heterogeneity effect (80% and 78% for males and females, respectively). Once the covariates are included into the model specification, they mainly contribute to explain the homogeneity across the municipalities, as the estimate of the corresponding posterior variance decreases.

	Empty model			With covariates		
	$\widehat{Var}(U\|Y)$	$\widehat{Var}(V\|Y)$	$\widehat{\delta}$	$\widehat{Var}(U\|Y)$	$\widehat{Var}(V\|Y)$	$\widehat{\delta}$
Men	0.062	0.003	0.799	0.054	0.003	0.783
Women	0.053	0.003	0.766	0.037	0.005	0.670

Table 1: Posterior estimates of variances for clustering and heterogeneity components and δ.

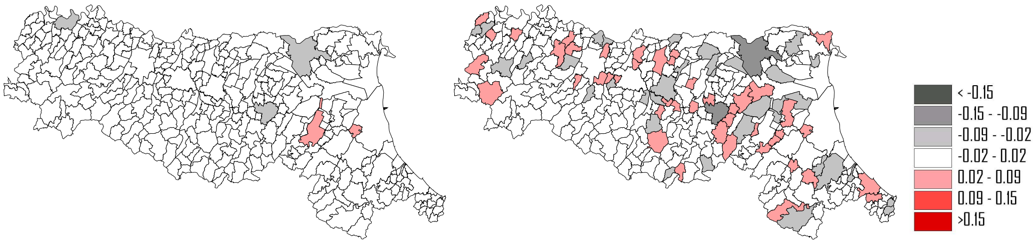
In order to compare such different contributions, distinct maps for each component, **u** and **v**, can be further provided (Figures 10 and 11). These results refer to the spatio-temporal model with covariates (8).



(a) Men. (b) Women. (c)

Figure 10: Clustering maps.

Especially for men, two big clusters with a positive contribution on the CRs^{95+} can be clearly identified. These substantially corresponds to the same areas described before as characterized by high and low values of CR^{95+}. On the other hand, the province of Ferrara is confirmed again the worst area in terms of homogeneity contribution to the values of the CR^{95+}.



(a) Men. (b) Women. (c)

Figure 11: Heterogeneity maps.

Conversely, the heterogeneity maps identify areas with a (positive-red or negative-black) notable component not due to the territorial adjacency. A larger number of areas for the females are highlighted, showing the higher residual variability than in the male group and suggesting the presence of additional observable or latent variables which fully explain these territorial differences.

Tables 2 and 3 includes the estimates of the time-dependent intercepts and the effects of the covariates included into the analysis. Corresponding 90% confidence intervals (CI) are also reported.

	Men		Women	
	Fitted rate	90% CI	Fitted rate	90% CI
Intercept :				
1995-1999	0.007	0.006 - 0.008	0.022	0.020 - 0.023
2005-2009	0.012	0.011 - 0.013	0.038	0.035 - 0.040

Table 2: Intercepts.

The fitted rates for the intercepts in the two periods, $exp(\alpha_1)$ and $exp(\alpha_2)$, represents the level of probability of observing P^{95+} in the mountain areas of Emilia Romagna, that is our reference group. For males, the other municipalities all experienced rate ratios which are lower than 1, with significant values for hill zones with high density and plain areas with low or medium density. Conversely, in the female group, people living in hill areas with medium density have a significant higher rate of reaching 95 years of age and over than mountain residents. Significant values are also provided for

	Men		Women	
	Fitted rate ratios	90% CI	Fitted rate ratios	90% CI
Mountain	1	-	1	-
Hill and density < 78	0.961	0.827 - 1.114	0.956	0.872 - 1.047
Hill and density $78 - 193$	0.978	0.846 - 1.127	1.132*	1.037 - 1.233
Hill and densiti > 193	0.852	0.741 - 0.973	1.027	0.936 - 1.126
Coastal hill	0.895	0.620 - 1.295	0.751	0.590 - 0.952
Plain and densiti < 78	0.772	0.611 - 0.972	0.743*	0.640 - 0.862
Plain and density $78 - 193$	0.794*	0.690 - 0.912	0.942	0.859 - 1.033
Plain and densiti > 193	0.906	0.801 - 1.021	1.015	0.935 - 1.103

* Results significant at a 5% level.

Table 3: Covariates.

coastal hill and plain with low density municipalities. They are both areal characteristics which appear to be unfavorable to the presence of P^{95+}.

4.2. The regression effect of the mortality patterns

Through the cohort spatial analysis we can further investigate the effects of the recent (age group 90-99) and past (80-89) structure of mortality on the corresponding cohort of P^{95+}. We consider the period 2005 to 2009, lonely, where, moreover, the larger territorial differences have been observed for both sexes. The results on the fitted rate ratios and corresponding 90% CI are reported in Table 4.

Generally, high mortality levels in the SMRs yield rate ratios lower than 1 when compared with areas with medium values for both periods, cause of death and sex; conversely, areas characterized by lower intensity of mortality are more likely to have higher values of CR^{95+}. For both sexes, the mortality levels for the circulatory diseases between ages 90 and 99 seems to mostly affect the presence of P^{95+}. Moreover, a high mortality caused by malignant neoplasms at age 90 to 99 significantly prevent the longevity.

For men the chance of becoming P^{95+} considerably depends on the mortality levels between 80 and 89 years of age and, especially, on the deaths for malignant neoplasms. Conversely, women longevity appears weakly affected by the level of mortality in the first considered period (i.e. 80-89).

5. Conclusions

Statistical theory, several simulation studies and a large number of applications all support the use of hierarchical Bayesian modeling for spatial

	Men		Women	
	Fitted rate ratios	90% CI	Fitted rate ratios	90% CI
Circulatory diseases (80-89):				
Medium mortality	1	-	1	-
Low mortality	1.091	0.996 - 1.195	1.005	0.947 - 1.067
High mortality	0.980	0.878 - 1.094	0.979	0.920 - 1.042
Circulatory diseases (90-99):				
Medium mortality	1	-	1	-
Low mortality	1.197*	1.081 - 1.326	1.068	1.009 - 1.130
High mortality	0.797*	0.714 - 0.914	0.848	0.795 - 0.906
Malignant neoplasms (80-89):				
Medium mortality	1	-	1	-
Low mortality	1.124	1.012 - 1.248	1.000	0.937 - 1.066
High mortality	0.828*	0.752 - 0.909	0.997	0.943 - 1.054
Malignant neoplasms (90-99):				
Medium mortality	1	-	1	-
Low mortality	1.070	0.963 - 1.190	1.036	0.974 - 1.100
High mortality	0.896	0.808 - 0.991	0.926	0.871 - 0.983

* Results also significant at a 5% level.

Table 4: Effect of the past mortality on the cohort of P^{95+} in 2005-2009.

and spatio-temporal data as a powerful method which allows to yield strong gains in the accuracy of predictions and effect estimates. The improvement is mainly due to the smoothing effect which control for random occurrences, especially when the areas under study are small, and for overdispersion. Moreover, the opportunity of including different levels of covariates allows to investigate some crucial effects for the study at hand and explain some parts of variability.

A critical aspect of these methods is certainly represented by the sensitivity to the choice of the hyperpriors (Bernardinelli, Clayton, and Montomoli, 1995), which can be sometimes avoided by the use of an empirical-Bayes approach (Carlin and Louis, 1998), rather than previous knowledge or by making assumptions which are reasonable with the problem at hand (Best et al., 1999; Banrjee, Carlin and Gelfand, 2004).

The use of these methods to analyze the longevity pattern in small areas, such as the municipalities of our north-eastern Italian region, represents a quite innovative application, which aims to confirm the usefulness of hierarchical modeling also in socio-demographic researches.

In the spatial and spatio-temporal application, the hierarchical Bayesian model has allowed us to investigate the effect of some crucial time dependent causes of mortality, as well as a set of areal features and correlations. As a result, some light on the factors that affect longevity in the population have been shed. The estimated values of the adopted longevity index have allowed us to identify some groups of areas where the people reaching 95+ years of age are likely to occur with a substantial persistence with time. In addition, this approach has offered the opportunity of adopting a cohort perspective, in order to study the causal link between the occurrence of P^{95+} and the level of mortality related to people belonging to the same cohort.

Since mortality and longevity may be differently affected by the same risk factors and the level of a specific-cause mortality might encourage/inhibit the presence of longevity over a region, a further development in both the applicative and methodological sense would be a multivariate spatial model to properly analyze this kind of data in a cohort framework (Jin, Carlin and Banerjee, 2005). In such case, additional broad categories of causes of death should need to be considered aiming to outline and control for all the mortality patterns.

References

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2004). Hierarchical Modeling and Analysis for Spatial Data. London: Chapman and Hall.

Bernardinelli, L., Clayton, D. G., and Montomoli, C. (1995). Bayesian estimates of disease maps: how important are priors? Statist. Med. 14: 2411-2431.

Bernardinelli, L. and Montomoli, C. (1992). Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. Statist. Med. 11: 983-1007.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems (with discussion). J. Roy. Statist. Soc., Ser. B, 36, 192-236.

Best, N.G., Waller, L.A., Thomas, A., Conlon, E.M. and Arnold, R.A. (1999). Bayesian models for spatially correlated diseases and exposure data. In Bayesian Statistics 6, eds. J.M. Bernardo et al. Oxford: Oxford University Press.

Breslow, N. E. (1984). Extra-Poisson variation in log-linear models. Appl. Statist. 33(1): 33-44.

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. J. Am. Statist. Ass. 88(421): 925.

Carlin B., Louis T. (1998). Bayes and empirical Bayes methods for data analysis. Chapman and Hall/CRC.

Clayton, D. G. and Kaldor, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. Biometrics 43: 671-681.

Gelman A., Carlin J.B., Stern H.S., Rubin D.B. (2003). Bayesian Data Analysis. 2nd edn. New York: Chapman and Hall/CRC.

ISTAT (1997). Classificazione delle malattie, traumatismi e cause di morte, IX Revisione 1975. Roma: Istituto Nazionale di Statistica, Metodi e Norme, Serie C, No. 10, Voll. 1 e 2, V ristampa.

Jin X., Carlin B.P. and Banerjee S.(2005). Generalized Hierarchical Multivariate CAR Models for Areal Data. Biometrics. 61: 950-961.

Langford, I. H., Leyland, A. H., Rasbash, J., and Goldstein, H. (1999). Multilevel modelling of the geographical distributions of diseases. J. Roy. Stat. Soc. Ser. C(48): 253-268.

Lawson, A.B., (2009), Bayesian disease mapping, CRC press.

Mollie, A. (1994). Bayesian Mapping of Disease. In: Gilks, W., Richardson, S., and Spiegelhalter, D. J. (eds.). Markov Chain Monte Carlo in Practice. London: Chapman and Hall.

Montesanto A., Passarino G., Senatore A., Carotenuto L., De Benedictis G, (2008), Spatial Analysis and Surname Analysis: complementary Tools for Shedding Light on Human Longevity Patterns, Ann. Hum. Genet. 72: 253-260.

Miglio R., Marino M., Rettaroli R., Samoggia A., Spatial analysis of longevity in a Northern Italian region, paper presented at the PAA Annual Conference, Detroit, MI, 28/4-2/5 2009

Poulain M., Pes G., Grasland C., Carru C., Ferrucci L., Baggio G., Franceschi C., Deiana L. (2004), Identification of a geographic area characterized by estreme longevity in the Sardinia Island: the AKEA study, Exp. Gereontol. (39): 1423-1429.

Robine, J. M., Caselli, G.,(2005) An unprecented increase in the number of centenarians, Genus, LXI: 57-82.

Robine, J. M., Caselli, G., Rasulo, D. Cournil, A. (2006) Differentials in the femininity ratio among centenarians: variations between northern and southern Italy from 1870. Pop. Stud., 60(1): 99-113.

Robine, J. M. , Paccaud, F. (2005) Nonagenarians and centenarians in Switzerland, 1860-2001: a demographic analysis. J. Epidemiol. Comm. H. 59:31-37.

Spiegelhalter D., Thomas A., Best N., Lunn D. (2003). WinBUGS User Manual, Version 1.4.