Annalisa Laghi • Giorgio Luppi • Angelina Mazzocchetti

Territorial clustering and recent demographic trends

Abstract The demographic framework of Emilia-Romagna region shows, in the last ten years, fast and deep changes. After a long stationary period the total amount of population has recorded a positive variation. The demographic rebirth, like shown by many studies relative to the entire regional territory, is manly due to the immigration from South Italy and abroad, to the new life of the fertility process and to the decreasing trend of the ageing process. The proposed analysis concerns the identification of sub-regional areas, homogeneous with respect to the above mentioned population's process, focused on the last ten years where the change appeared with more importance; the scope is to describe and understand the demographic heterogeneity of the Emilia-Romagna region. Applying a cluster analysis with preliminary data treatment and posterior check, the study results in five homogeneous areas characterized by different demographic behaviour.

Keywords demographic indicators • territorial analysis • outliers • clustering

Annalisa Laghi
Servizio Controllo strategico e statistica, Regione Emilia-Romagna
V.le Aldo Moro 52, 40127 Bologna, Italy
E-mail: alaghi@regione.emilia-romagna.it

Giorgio Luppi (✉)
Via Rocca 15, 40014 Crevalcore, Italy
E-mail: giorgioluppi2@alice.it

Angelina Mazzocchetti
Servizio Controllo strategico e statistica, Regione Emilia-Romagna
V.le Aldo Moro 52, 40127 Bologna, Italy
E-mail: amazzocchetti@regione.emilia-romagna.it

1 Introduction

The strong development of population in the last ten years (about 260 thousand people, from 1996 to 2006) results to be one of the more consisting growth of the demographic history of the region; since 1861 only in the intercensus periods 1901-11 and 1911-21 a population increase of the same entity, in absolute terms, was observed.

The importance of this increase is underlined from the fact that in 2006 the regional population marked the double with respect to 1861 while in Italy had happened with the census of 1971.

The strong novelty and importance of the recent development can be appraised better to the light of some demographic reflections made in the course of the years' 90 and here synthetically summarized.

Golini, Citoni (1994): “The two phenomena that much more characterize the demographic evolution of the west society, namely prolongation of the middle life and fall of fertility, appeared , respect to Italy, with greater evidence and anticipation in Emilia-Romagna, that entered in a demographic phase of fall that can herald to a very implosion”.

Del Panta, Montanari (1994): The expansive minor spirit of the regional population respect to the national one is an almost systematic fact. If it was necessary little more of 100 years to assist the double of the Italian population, never Emilia-Romagna region has been able to reach this result. The turn in the upward trend is therefore happened in Emilia before reaching the double, anticipating of some year what will be, with every probability, the awaited inversion of tendency for the whole country".

Bonaguidi, Toigo (1994) "If it will persist the demographic regime observed at the beginning of the years' 90 the population of the Emilia-Romagna region would enter, after a long phase of growth, in a demographic phase of decline, limited in the next ten years, more emphasized in the successive twenty."

How does it note, attentions was pointed to the problems of the demographic decline; very different are the evaluations that can be found at the beginning of the years ‘2000.

Particularly, already in Nicolini (2000) it’s said that “based on our projections, in the next years it is permitted to await itself, also in case of constant fertility , a slight increase of the population (…) In conclusion, seems improbable that come true, to middle limit, the so feared demographic sudden fall”.

More recently, in the report that accompanies the demographic predictions of the Regione Emilia-Romagna (2004) it’s introduced the idea of " a demographic rebirth of the region" and says that "the new tendencies are leading the regional system towards new situations and new balances, demographic and territorial ones…. When a population returns to grow and the fertility is in renewal, births increases: as a result the trial of ageing undergoes a slowing, if not an inversion of tendency".

2 Emilia-Romagna demographic framework

In the last years Emilia-Romagna resident population has recorded some demographic changes which represent an important turning point respect to the past (Regione Emilia-Romagna 2005). In the decade 1996-2006 the population has increased of nearly 260 thousand of units (6,7%) to reach the size of 4.188 thousand of units while in the previous years it had been substantially stable and little above 3.900 thousand of people (since middle 70s the population ranges was between 3.900 and 3.960 thousand of units).

Fertility, after many years, has started increasing again; since the second half of the 1990s it has been raising at a rate definitely faster than all of other Italian region ones: the total fertility rate increases of 35,8% in Emilia-Romagna while the other regions recorded percentage increases in a range between 13,3% (Marche) and 27,3% (Friuli Venezia Giulia and Tuscany). The value reached by the Emilia-Romagna total fertility rate in 2004 is in line with Italian average (1,3) (Cantalini, 2005).

Particularly, a recent analysis of the National Statistical Institute (ISTAT) (August 2006) reminded that "between 1995 and 2004, in the center's and north's regions births increases was between 15% and 25% with a peak in Emilia-Romagna with nearly 37%. In the south's regions, on the contrary, in the same period the numbers of birth has a decline included in –9% of Campania region and –17% of Calabria region ".

The increase of birth is in part determined by immigrants: in the same ISTAT's study we find that in 2004 in Emilia-Romagna the average number of birth for woman was of 1,15 for Italian women and of 2,78 for the foreign one's (at national levels the two figures was 1,26 e 2,61 respectively).

The demographic transformation is due to the decisive contribution of immigration from the south of Italy and abroad. Emilia-Romagna as become the privileged destination of the south-north migration flows and the attractive capacity to foreign people is one of the highest among Italian regions. In the three-years period 1996-1998 the migration balance was of nearly 80 thousands of units while it is of nearly 180 thousands of units in the latest period (2003-2005). A sign of this attractively status of Emilia-Romagna, even if partial end not so updated, is the numbers of resident people born in southern or center Italian region: 360 thousands of units in 2001 census against 260 thousands ten years before (1991 census). Moreover, the ageing process has shown a decrease: the ageing index, ratio between people aged over 65 and people under 14 (x 100), which had been increasing from 1981 to 1998 (from little less than 100 to nearly 200), has started to decrease reaching the value of 182 in 2006. Considerable and systematic is the earnings in middle limits of life at birth: from 1990 it results in 4 years for male and 3 for females. In spite of a positive development they do not result however changes of speed or special accelerations: the earnings in limits of survival results moderate in the time.

In the national contest it's useful to remember the following points about Emilia-Romagna demographics: "a very dynamic demographic frame, decrease of the ageing level of the population, strong and determined effect of immigration , increase of fertility, population growth, privileged territory for the south-north

migrantion flows, unique case in the national contest" (Regione Emilia-Romagna 2005).

It can also be cited an intervention of Dalla Zuanna (2005) with the significant title "A new demographic springtime" how he asserts, between the other things, that "Emilia-Romagna is the most extraordinary case: started from 1990s the population replacement it's guaranteed almost in the same way to birth and immigration".

3 Demographic data sources and indicators

The above observations indicates the present change and its main components; in any case, it's reasonable considers not exhausted this phase of strong change and so it appears useful to continue the monitoring process to pick out the peculiarity of the regional demographics characteristics.

The aim of this sub-regional demographic analysis is to explore the territorial differences occurred in this period of changes detecting clusters of municipalities. The administrative aggregations of municipalities that can be used are several, i.e. provinces, scholastic or sanitary districts, etc., but clusters not linked to administrative bonds and based on the indicators characterizing the recent demographic development at municipalities level seems to be much more informative.

The changes under study occurred in the period from the middle of '90s to 2006, so referring the analysis to census data (surely interesting because of the availability of economic and social indicators) seems not to be completely right given that the more interesting changes occurred in the *intercensus* period and much more strong in the last five-six years; for this reasons we used the follow current demographic data :

- Vital statistics data on resident population collected by ISTAT.
- Age and sex distribution of the resident population collected since the beginning of the '80s by the Statistics Department of Regione Emilia-Romagna.

These two source of data offers, in Emilia-Romagna, some important prerequisite, first of all, the well-timed availability of the data: six or seven months after the years of reference; the possibility to work with time series, the territorial level of aggregation- the municipalities- and the quality of the data assured by many check of them. They have, of course, also some disadvantages like some discordances linked to the census and the impossibility to measure the presence of immigrants with only the residence permit and not resident. For the migratory balance the data referred to foreign flow are affected from some legal measure in the years 1996, 1999-2000 and much more in the 2003 which involved in different way the Italian territory. Another problem is that data for big municipalities is the results of situations sometimes really different inside the municipalities so it would be better work with data at quarter level or aggregation of city-quarter.

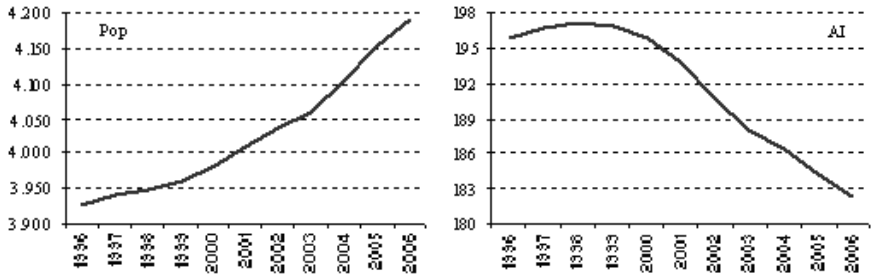
It goes underlined that these data were collected from the Italian National Statistics Office also for all the others Italian municipalities (balance data since many decades while the structure of the population by age since the beginning of the '90s), so the analysis we proposed could be repeated also in other Italian areas.

Homogeneous areas of municipalities have been detected by carrying out a cluster analysis. The indicators are chosen so that to measure the phenomena that have been characterizing the recent demographic development, ageing, fertility and migrations (Cantalini 2005). The variation of population hasn't been included because it's a 'final product' of the other indicators, particularly fertility and migration.

In details, the indicators considered are:

- ageing index (AI) which seems to be an indicator of the regional ageing process more informative than the only proportion of old people because its trend measures the relation between the growth of people aged 65 and over and that of people aged 14 years and less. In the decade 1996-2006 the growth of young people had been stronger than that of the elderly population: the number of elderly passed from 834 thousand units to 951 thousands of units (14,1%) while the people under 14 years increased of 96 thousands of units (22,6%), from 426 to 522 thousands of units. Furthermore, in the last five years, parallel of an equivalent development in absolute value (less more than 60 thousands of units), the recorded percentage increase was of 7,1% for elderly people and of 14% for young people.
- the proportion of children under 5 year for 1.000 women in fertility age, conventionally aged 15-49, (FR), proposed by Golini et al. (2000) to measure fertility because it's more stable at the municipality level than the gross birth rate.
- the total migration rate (MR) which permits to take into account in a complete way both the foreign migration flow and the important South-North one. It's worth noting that the internal migration has been composed by an increasing portion of foreign people. Emilia-Romagna population has recorded, throughout 2005, an increase of 4 thousands resident foreign people coming from other Italian municipalities (12,7% of the total increase of foreign population).

Fig. 1 shows the trend of the demographic indicators, together with the regional population one, during the last decade.



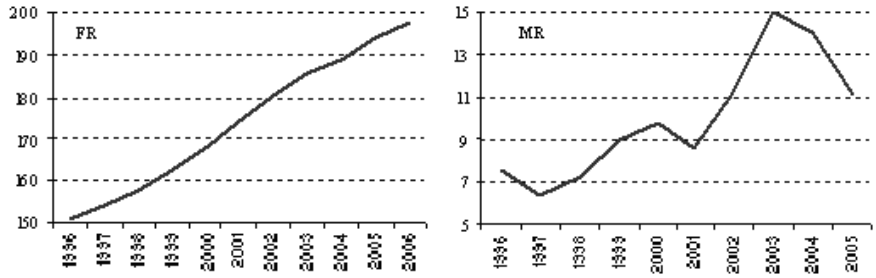


Fig. 1 Trends of Emilia-Romagna population and chosen indicators in the decades 1996-2006

4 Methods

Demographic indices built at the municipality level show high robustness and stableness problems because of the small population size of lots of these units: on the 1st of January 2006, 160 municipalities out of a total of 341 record less than 5 thousands of inhabitants. The indicators have been then built as weighted averages over a time period and then treated to reduce the impact of outliers.

The analysis proposed in this paper is based on weighted averages over periods of three years, which is a time interval frequently considered in territorial demographic studies. To make stock indicators (AI and FR) more comparable with the flow one (MR), their weighted averages have been computed over values referred to the middle of each year.

The demographic development over the last decade have been appreciated by performing the cluster analysis on the indices computed over the starting and final three year periods: 1996-1998 and 2003-2005.

Outlier treatment has consisted in detecting the tails of each univariate indicator distribution and then trimming the range of variation by defining minimum and maximum thresholds. Values outlying those defined ranges have been replaced by the nearest threshold (Banett, Lewis 1984). This simple method allows to perform the analysis on complete data for each statistical unit.

The treatment revealed substantial mainly for AI leading to reduce the values of 25 municipalities of small size (1,000 people on average), located in the mountain area, to the threshold of 400. The range of variation of the fertility rate relative to 1996-98 period was limited between 90 and 215 determining the treatment for the values of 3 little inhabited municipalities (2 with less than 200 inhabitants, the third one with nearly 1,000). No treatment was necessary for the fertility rate relative to 2003-05 period. The migratory rates of both periods were treated for values outlying the [-5, 40] range (9 units in the first period, 7 in the latter, with a population average of nearly 2,000 units).

Cluster analysis has been performed on the whole set of the 6 treated indices (3 per each period) standardized so as to range in [0,1]. No preliminary transformations aiming at variable uncorrelation were applied to the data because they can hide the underlying cluster structure. The conjoint use of the 6 indicators, though correlated, has been evaluated appropriate and justified in this context, without loss of information about the pre-existing cluster structure, on the basis

that the differences among the municipality classifications detected by each single indicator aren't statistically significant (Faulkes et al. 1988, Soffritti 2003).

A two step clustering procedure has been adopted consisting in a hierarchical agglomerative algorithm followed by a reclassification of the statistical units via the standard k-means algorithm. In the hierarchical stage squared Euclidean distance and Ward algorithm have been considered as dissimilarity measure and linkage method respectively, as proposed by Lance and Williams (1967). The partitive phase was imposed to start the iteration process from the centroids of the k clusters detected in stage 1.

In this paper the results of two cluster analyses, carried out on the 341 Emilia-Romagna municipalities, are presented and compared. The first one considers all of the statistical units equally weighted 1, while the latter solution weights each municipality with a population size measure; here the amount of population on the middle of the decade (1st January 2001). The dendograms of the hierarchical algorithm both for the unweighted and weighted analysis point at 5 cluster partitions.

The described procedure, performed on indicators made steady by the time period average and then treated for outliers, detects clusters including municipalities often territorially contiguous in both the weighted and unweighted analysis. In a first stage of the research a geographic contiguity constraint clustering algorithm was also applied but it performed worse than the proposed standard procedure in detecting the peculiar demographics of the surroundings of the main cities, a phenomenon widely analysed by recent demographic studies.

The comparison between pairs of different partitions has been performed on the basis of the Rand index (Rand 1971), corrected for chance and normalized by Hubert and Arabie (1985). The adjusted Rand index is bounded above by 1, is equal to 0 when the two partitions are picked at random, subject to having a given number of classes and objects in each class, and can also take on negative values (which generally have no substantive use). It can be computed on the basis of a contingency table containing weighted or unweighted frequencies.

The indicators importance in detecting the cluster partition has been evaluated by applying random forests in the classification context, a non parametric method proposed by Breiman (2001). It aims at predicting, for of each statistical unit, the value of a particular feature (the cluster membership) on the basis of the values of other features (the demographic indicators). A random forest consists in generating a set of random classification trees to define an aggregate prevision function that allows also to study the importance of each indicator. In the weighted version each unit member of a cluster has been weighted by the percentage of population (out of the total regional population) resident in the municipalities belonging to that cluster, as suggested by Breiman. The general importance of each indicator in detecting a partition ranges between 0 and 1.

5 Results

The municipalities belonging to each of the 5 detected groups for the unweighted and the 2001 population weighted analyses are shown in Fig. 2 and Fig. 3 respectively.

The unweighted analysis represents substantially an updating of the territorial clustering proposed in Laghi et al. (2006) made possible by the availability of the 2005 migration data and the resident population by sex and age on the 1st January 2006. The resulting partition is equivalent to the one presented in the previous study, with groups that reveal interesting and informative because of the differences among the demographic indicator levels, as shown by the cluster centroids (Table 1), and for including municipalities often territorially contiguous. This analysis detects a wide and potentially not homogeneous area, cluster 4, including 133 municipalities and more than the half of the total regional population.

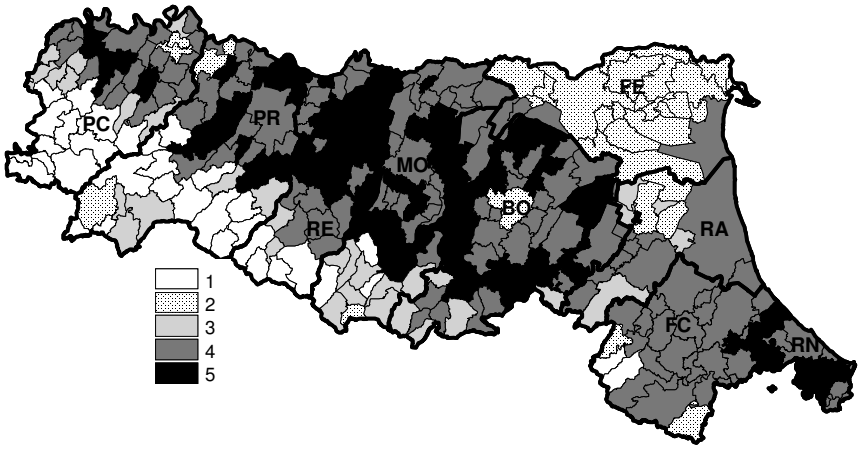


Fig. 2 Clusters of Emilia-Romagna municipalities – unweighted analysis

Table 1 Cluster centroids and sizes, 2001 population – unweighted analysis

Cluster	AI		FR		MR		N	2001 Pop
	1996/98	2003/05	1996/98	2003/05	1996/98	2003/05		
1	386.8	389.3	137.3	145.9	4.3	6.8	34	52,862
2	259.6	278.9	124.0	150.3	1.1	7.1	34	744,710
3	298.6	278.1	163.6	181.1	8.4	14.8	33	111,348
4	186.0	179.0	161.5	194.9	7.2	12.8	133	2,211,617
5	153.8	136.7	176.5	216.3	17.4	22.6	107	888,304
Total	196.7	185.2	156.2	191.7	7.1	13.5	341	4,008,841

The main characteristics of the partition detected by the weighted analysis are:
Cluster 1: mountain area of Piacenza, Parma e Reggio Emilia where the aging index is very high and stable over the analysed period and all the indicators differs negatively from the regional average; this area pertains to little more of 60 thousand inhabitants distributed in small municipalities (with average population inferior to 2000 inhabitants).
Cluster 2: Ferrara district (inclusive of Ferrara city) and the Bologna municipality with indicators negatively divergent from the regional average but with aging index inferior to the value recorded for the cluster 1 and in slight decreasing. The two province-capital municipality cover about the 2/3 of the population of the cluster.
Cluster 3: municipalities that present, on average, values not very distant from the average and regional tendency; at territorial level it is not presented with a solid distribution but can be individualized two areas: to) Parma and some municipalities mainly towards Piacenza province; b) part of the Romagna in an area that come from the costs to the Apennine and includes also the municipalities of Ravenna, Faenza and Forlì. Municipalities beyond 50 thousand livings (Parma, Ravenna, Faenza and Forlì) cover almost 50% of the population of the cluster.
Cluster 4: municipalities with a *restrained* aging index in slight increase and an immigration lower than the regional average; such communes are adjacent to areas in stronger development particularly in the provinces of Reggio Emilia, Modena

and Bologna and in part of the Romagna area. This cluster includes six municipalities beyond 50 thousand livings (Piacenza, Modena, Carpi, Imola, Cesena, Rimini) that covers little more of 50% of the population of the cluster.
Cluster 5: area with indicators exceeding the regional average. It includes Reggio Emilia municipality and a large area of its surroundings, several municipalities near the cities of Piacenza, Parma, Modena and Rimini as well as Bologna city second hinterland. In many points from Piacenza to Bologna this cluster shown territorial contiguity. Only one municipality beyond 50 thousand livings (Reggio Emilia) pertains to this cluster and it covers approximately the 15% of the cluster's population.

The 5 clusters presents population variations clearly differentiated in the decade 1996-2006 ranging from cluster 1 which record a substantial population decrease (nearly 7%) to cluster 5 with a population growth of about 19%. Cluster 3 population increase is on the regional average.

The distribution of population amount and of municipalities number in the 5 cluster results to be not very dissimilar; the very exception is cluster 1, with a *contained* demographic dimensions, but the solution is justified because specify a wide mountain area clearly identified by the value of the used indicators.

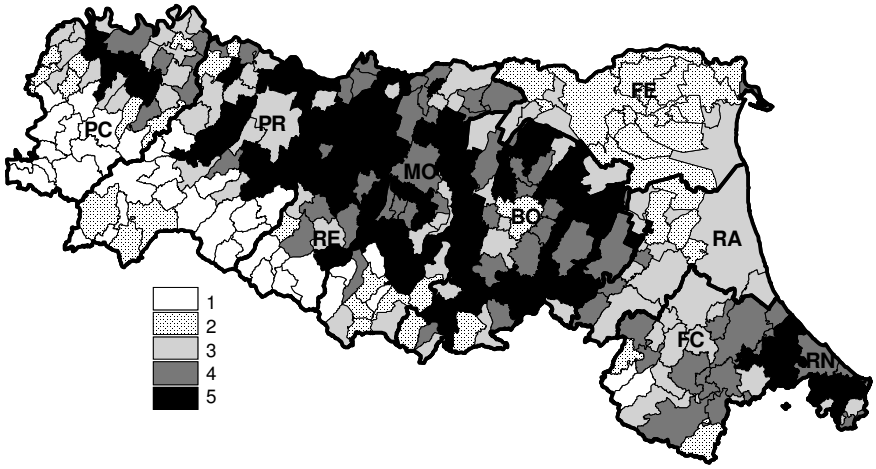


Fig. 3 Clusters of Emilia-Romagna municipalities – 2001 population weighted analysis

Table 2 Cluster centroids and sizes, 2001 population – 2001 population weighted analysis

Cluster	AI		FR		MR		N	2001 Pop
	1996/98	2003/05	1996/98	2003/05	1996/98	2003/05		
1	376.5	377.4	142.1	152.1	5.5	7.5	38	61,304
2	289.3	275.0	133.0	161.6	2.7	7.1	46	774,153
3	211.1	198.0	148.6	186.5	6.2	16.5	75	1,021,991
4	170.8	172.5	161.3	192.1	4.4	9.0	63	1,169,710
5	153.8	136.6	176.9	217.5	15.3	19.8	119	981,683
Total	196.7	185.2	156.2	191.7	7.1	13.5	341	4,008,841

The clusters can also be characterized by other indicators on the last tree-years period that permit a deeper investigation of the demographic behaviour of the clusters (Table 3).

- The percentage distribution of the population by big age classes (less than 14, 15-64, over 65) shown a big diversification on the weight of the working-age class starting with nearly 56% in cluster 1 to the nearly 67% in cluster 5. The other two extreme age class shown an opposite trend: while for the young class the value in cluster 5 in nearly double than that in cluster 1 for the oldest age class the trend is opposite;
- The active population replacement index (IR, ratio between the amount of population aged 15-19 years and the population aged 60-64) vary from nearly 130 in cluster 5 to nearly 220 in cluster 1;
- The natural balance (NB) pass from negative value (-15 x 1000) in cluster 1 to positive, even if with low value, in cluster 5;
- The migratory balance point out homogeneous value respect to the foreign part (MR_*A*), 8-9 x 1,000, but very different for the internal one; the internal migratory balance (MR_*IT*) is very concentrated in cluster 3 and 4.

Significant are the changes in the age structure in the consider decade reflected by the 5 identified cluster: a slight decrease of the working-age population, more than 1 percentage point increase of the young population (except for cluster 1) and of the oldest one (except for cluster 5). Young people percentage growth is more evident in cluster 5 while the increase in old population is more evident in cluster 3.

Table 3 Demographic indicators (2003/05) and population increase (PI - 2003/05 with respect to 1996/98) – 5 clusters partition, weighted analysis

Cluster	IR	%0-14	%15-64	%65	MR_*IT*	MR_*A*	NB	PI
1	222.1	7.9	55.9	36.1	-0.1	7.5	-15.1	-7.0
2	212.2	9.8	63.3	26.9	-0.8	7.9	-6.2	-2.4
3	166.4	11.8	64.9	23.3	7.4	9.2	-2.8	4.7
4	153.6	12.6	65.7	21.6	0.1	9.0	-1.5	3.3
5	133.0	14.2	66.5	19.3	11.7	8.6	1.3	13.0
Total	160.9	12.2	65.1	22.7	4.7	8.7	-2.2	4.6

The results of this analysis agrees with those of Golini *et al.* (2000) in the identification of two areas with many similarities and demographics problems: the Apennines from Piacenza to Reggio Emilia (cluster 1) and the Ferrara area (part of cluster 2). The novelty results in the comparison of the migratory process that is, now, stronger than in the period analysed by Golini *et al.*. In any case it's useful to remember that our analysis is different from that proposed by Golini et al. in the methodology, in the indicators and, obviously, in the reference period.

Compared with the previous version of the work (Laghi et al. 2006) the development and improvement of the methodological procedure permits to directly detect the complexity and diversity of some sub-regional areas. Particularly, a big cluster of over 2,200 thousand of inhabitants now results in two separate areas (cluster 3 and 4) with less more than 1,000 thousand of inhabitants.

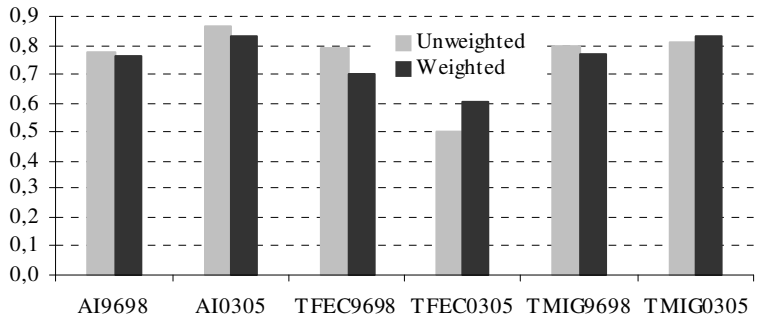


Fig. 4 Indicators importance in detecting 5 cluster partition for the weighted and unweighted solutions

6 Conclusions and perspectives

The obtained aggregations results particularly interesting and informative about the difference on the indicators levels and population dynamics and also because they includes in many areas municipalities territorially contiguous. The cluster procedure, as a matter of fact, recognize, without any territorial limit, contiguous areas; this is particularly evident, on one side, in two demographic frailty areas (cluster 1 and 2) and, on the other side, in a strong demographic development area (cluster 5). This result seems to be considerable for the evident demographic connotation of the individualised area and even because the procedure is completely renewable each year with the availability of data updates.

In a period of demographic mutation like underlined in this analysis seems necessary to continue the monitoring process of the population at a regional and sub-regional level to better understand the undergoing change and to have helpful material for the local programming.

In this step is useful to take particularly attention at those areas with extreme characteristics (namely cluster 1,2 and 5). It will be important the characterization of the individuated area by a socio-economics points using the census indicators at 2001; a positive contribution for this characterization will be arrive considering the updated data of ASIA (Statistical Archive of the Active Firms) at municipality level.

Besides the application of the proposed methodology with updated data it seems to stands out other possible deepening and development:

- individuation of other temporal periods on the construction of the indicators average;
- evaluation of the contribution of the decreasing mortality to the increasing population trend in the different individualized areas; in this case it's needed the comprehension of an appropriate indicator at municipality level;
- use of smoothing procedure to better read the indicators on the territorial levels;
- use of data on the neighbouring municipalities with the scope of verify the presence of continuity in the demographics behaviour over the regional bounds;

- in deep evaluation of the contribution of the foreign migration flow.

Acknowledgments We thank Giuliano Galimberti and Gabriele Soffritti, researchers at the Department of Statistics of the University of Bologna, for their suggestions and methodological support. The helpful comments of the referee and Alessio Pollice (University of Bari), our discussant at the Turin 2006 SIS Conference, are grateful acknowledged.

References

Banett V, Lewis T (1984) Outliers in Statistical Data. 2nd ed. Wiley, New York

Bonaguidi A, Toigo M (1995) Le proiezioni demografiche per aree sub-regionali in Emilia-Romagna: l'applicazione del modello multiregionale. In Regione Emilia-Romagna, Atti della conferenza "La popolazione dell'Emilia-Romagna alle soglie del 2000", Ferrara 24-25 novembre 1994

Breiman L (2001) Random forests. Machine Learning 45:5–32

Cantalini B (2005) Recenti sviluppi dell'evoluzione demografica: un confronto tra le regioni. Invited speaking at the Bologna Cisis Conference, October 2005: Le previsioni demografiche. http://www.cisis.it/archivi/seminari/index.html.

Citoni F, Golini A (1995) L'Emilia-Romagna nel contesto nazionale ed europeo. In Regione Emilia-Romagna, Atti della conferenza "La popolazione dell'Emilia-Romagna alle soglie del 2000", Ferrara 24-25 novembre 1994

Dalla Zuanna G (2005) Una nuova primavera demografica. Il Mulino, 6:1061-1071

Del Panta L, Montanari A (1995) La storia recente della popolazione dell'Emilia-Romagna. In Regione Emilia-Romagna, Atti della conferenza "La popolazione dell'Emilia-Romagna alle soglie del 2000", Ferrara 24-25 novembre 1994

Faulkes EB, Gnanadesikan R, Kettenring JR (1988) Variable selection in clustering. Journal of classification 5:205-228

Golini A, Mussino A, Savioli M (2000) Il malessere demografico in Italia. Il Mulino, Bologna

Hubert L, Arabie P (1985) Comparing partitions. Journal of Classification2:193-218

ISTAT (2006) Natalità e fecondità della popolazione residente: caratteristiche e tendenze recenti – Anno 2004. Nota informativa 1 agosto 2006

Laghi A, Luppi G, Mazzocchetti A (2006) Detecting homogeneus areas in a sub-regional demographic analisys. In: Società Italiana di Statistica, Atti della XLIII Riunione Scientifica, Università di Torino, 14-16 giugno 2006, Sessioni spontanee

Lance GN, Williams WT (1967) A general theory of classification sorting strategies. Computer Journal 9:373-380

Nicolini L (2000) Le proiezioni di popolazione e famiglie. In Regione Emilia-Romagna, Lo sviluppo demografico in Emilia-Romagna, Tendenze della popolazione regionale. Franco Angeli

Rand WM (1971) Objective criteria for the evaluation of clustering methods. Journal of the American Statistical Association 66:846-850

Regione Emilia-Romagna (2004) Le previsioni demografiche. Bologna

Regione Emilia-Romagna (2005) Lo sviluppo demografico in Emilia-Romagna. Bologna

Soffritti G (2003) Identifying multiple cluster structures in a data matrix. Communications in Statistics, Simulation and Computation 32 (4):1151-1177