

Assessing Regional Wellbeing in Italy: An Application of Malmquist-DEA and Self-organizing Map Neural Clustering

Questa è la versione Post print del seguente articolo:

Original

Assessing Regional Wellbeing in Italy: An Application of Malmquist-DEA and Self-organizing Map Neural Clustering / Carboni, Oliviero Antonio; Russu, Paolo. - In: SOCIAL INDICATORS RESEARCH. - ISSN 0303-8300. - 122:3(2015), pp. 677-700. [10.1007/s11205-014-0722-7]

Availability:

This version is available at: 11388/59705 since: 2022-05-27T17:23:44Z

Publisher:

Published

DOI:10.1007/s11205-014-0722-7

Terms of use:

Chiunque può accedere liberamente al full text dei lavori resi disponibili come "Open Access".

Publisher copyright

note finali coverpage

(Article begins on next page)

Assessing Regional Wellbeing in Italy: An Application of Malmquist-DEA and Self-Organizing Map Neural Clustering

Oliviero A. Carboni, corresponding author

DiSEA and CRENoS, University of Sassari, Italy. Email: ocarboni@uniss.it

Via Muroni 25, 07100 Sassari, Italy. Tel.: +39 079 213016; fax: +39 079 213002.

Paolo Russu,

DiSEA, University of Sassari, Italy. Email: russu@uniss.it

Abstract

Interest in the measurement of wellbeing and quality of life has increased in recent decades and a wide range of statistical and econometric techniques have been used to investigate and measure individual quality of life. Following this line of research, this paper uses data envelopment analysis (DEA) to evaluate the wellbeing performance and ranking of the 20 Italian regions from 2005 to 2011. The analysis is based on 12 indicators which represent some of the different aspects of wellbeing. These include economic conditions, labour market conditions, neighbourhood relationships and the environment. The Malmquist indices (MPI) obtained from the DEA scores are then used to measure changes in wellbeing over time. The results reveal that northern regions have been performing with more efficiency than southern ones. This paper also uses the Self-Organizing Map (SOM) technique to cluster regions into homogeneous groups where the within-group-object dissimilarity is minimized and the between-group-object dissimilarity is maximized. The clustering analysis confirms a marked duality in regional wellbeing in Italy.

Keywords: Quality of life; Data envelopment analysis (DEA); Self-organizing map neural network; Panel data.

INTRODUCTION

Measurement of quality of life or welfare has become a matter of interest for researchers. Over the last few years, many empirical works have investigated the question of quality of life through theoretical and empirical research in various disciplines. One paramount idea is that the social and physical environment of an area can affect the wellbeing of the people living there. Another reason for measuring and comparing quality of life in different places is that it may provide useful indications when developing welfare policies. The concept of wellbeing is multifaceted, and many of these aspects are hard to quantify. Many works have assumed that individual wellbeing or satisfaction with life is a combination of satisfaction in various domains (Diener 1984; Diener et al. 1999; Sirgy 2002; van Praag et al. 2003; van Praag 2007, 2011). Moreover, many of the elements determining the level of wellbeing are subjective in nature and depend to a large extent on individual perceptions.

Individual wellbeing is mainly determined by the satisfaction that individuals assign to their personal life (Andrews and Withey 1976; Campbell et al. 1976; Diener 1984; Diener and Suh 1997; Diener et al. 1999, 2003; Sirgy et al. 2000, 2008, 2010; Sirgy and Cornwell 2001, 2002; Yiengprugsawan et al. 2010). Satisfaction in turn depends on several other aspects, such as government services, business, non-profit services (Sirgy et al. 2000; Sirgy and Cornwell 2001), neighbourhood characteristics (Sirgy and Cornwell 2002), and community services and conditions (Christakopoulou et al. 2001; Sirgy et al. 2008). Hence wellbeing at an individual level is the sense of satisfaction that the individual feels through consumption of market goods, leisure, public goods and the other physical and social characteristics of the environment in which individuals live (Gillingham and Reece 1979). Hence, improving the quality of life is not only a matter social equity, but also a crucial aspect of attracting people and investments in certain areas. In this sense it may play a crucial role in determining whether or not firms or individuals decide to locate themselves in a particular place (Rogerson 1999).

Though social welfare has always been a crucial issue in economic sciences, its measurement has traditionally been mainly linked to monetary factors, such as GDP, price levels and cost of life. Another strand of economic thinking claims that this is too reductive for reflecting the complexities of a society's economic situation and prefer more complex definitions. For instance Townsend (1979) and the so called Scandinavian Welfare Approach (Erikson et al., 1987; Erikson, 1993) claim that the multidimensionality of quality of life is an economic good. Slotje (1991) stresses the difficulties involved in this multidimensional approach, both in terms of its methodological and theoretical requirements.

Nordhaus (2002) highlights that national accounts are necessarily incomplete, because they do not contain many of the economic activities taking place within the marketplace, and thus may supply only a partial or biased estimate of wellbeing. Sen (2000) highlights the importance of considering elements that have an impact on the standard of living and of going beyond purely income based measurements. In addition the quality of life cannot be exclusively measured in terms of objective living conditions, because there are subjective aspects to people's perceptions of their quality of life, and qualitative ones must also be taken into consideration. Thus quality of life is linked to several aspects of life, some of which are hard to measure.

Murias et al. (2006), Jurado (2010), Gonzalez et al. (2011), Lawless and Lucas (2010) have all constructed multidimensional wellbeing indices at regional or administrative levels can be found in. Alternatively welfare can be measured directly through a subjective evaluation of an individual's quality of life (Diener 1984; Diener and Suh 1997; Diener et al. 1999, 2003; Kahneman and Krueger 2006; van Praag 2011; van Praag et al. 2003). The subjective wellbeing indicator method takes into consideration an individual's experience and perception of the multifaced social environment.

The use of utility functions, the analysis of national data, and the aggregation of a variety of social indicators which incorporate both social and economic aspects, have all helped our understanding and measurement of quality of life. The social indicators approach is one of the most important methods for measuring wellbeing. Social welfare is split into several components, each defining a certain social indicator. The aggregation of all the indicators provides the measurement of wellbeing. This method is a great improvement, because it does not require the indicators to be assigned monetary values for aggregation. The Genuine Progress Indicator (Cobb et al. 1995) and the Index of Economic Well-being (Osberg and Sharpe 1998, 1999) are two of the most important examples of applying the social indicator approach.

From the empirical point of view, there are two key problems in constructing a social welfare indicator. Firstly the measurement of a multidimensional concept is subjective in nature, since it is strongly linked to an individual's own perception of her or his life. This inevitably means that the indicators for all the relevant key dimensions of quality of life must be selected and quantified. Secondly an aggregation methodology has

to be defined and applied to the subset of indicators, in order to obtain a reliable measure of welfare (González et al. 2011). In a variety of cases in the empirical literature the particular sub set of indicators chosen usually depends on the available data. The empirical literature also offers a wide variety of choices in the aggregation methodologies used. Among the most relevant ones are: the synthetic indicator of multidimensional distance proposed by Pena (1977), the hedonic price methods proposed by Rosen (1979) and Roback (1982) and the data envelopment analysis approach suggested by Hashimoto and Ishikawa (1993).

Although DEA is a multidimensional technique originally developed in efficiency analysis, it is particularly appropriate for use in the social indicator context. This procedure has increasingly been used in several applications which are not necessarily production-oriented, thus highlighting the possibilities of DEA as a valid instrument in multidimensional analysis (Hashimoto 1999). One of the most relevant uses of such a procedure has been its application to analysing standards of living and social wellbeing as well as for the construction of social and economic wellbeing indicators (Cook and Kress 1990; Hashimoto and Ishikawa 1993; Hashimoto and Kodama 1997; Zhu 2001; Murias et al. 2006; Gonzalez et al. 2011).

The DEA procedure can be adapted to measure quality of life by considering the indicators that highlight the drawbacks of living in specific areas as inputs (costly aspects that should be minimised) and the indicators that show the advantages as outputs (positive factors that should be maximized). It is a reasonable way of aggregating indicators, because it can easily handle multiple dimensions (inputs/outputs) without imposing tight structures on the relationships between the variables.

Other methodologies, hedonic pricing for instance, require the functional forms of the relationship between the indicators to be specified. DEA generates a comparison frontier from the best units observed in the sample, based on a comparative assessment of the indicators. However, one must not forget that this procedure also has low discriminating power, especially when many dimensions are taken into account and the sample size is limited (Ali 1994). The DEA setting also gives information of which regions act as frontier benchmarks for each low performing region in the sample.

Regardless of the methodology and indicators used, there are several examples of it being used for city ranking analysis. For instance Liu (1976) employed a simple statistical method to compare the performance of 240 metropolitan areas in the U.S. and rank them on the basis of 50 indicators. These indicators were related to economic health, political performance, environmental conditions, health and education, and social concerns. In line with this, the Places Rated Almanac (Boyer and Savageu 1981), classified North American cities according to a previously selected and weighted group of indicators. Other works employ rankings of cities for different objectives. For instance, the yearly European Cities Monitor Report aims to rank leading European cities according to their attractiveness for business. In Italy, the newspaper Sole24ore produces an annual report on quality of life, with rankings of the 103 Italian provinces based on six groups of indicators: standard of living, employment and business, services, crime, population and leisure. In Spain, the Anuario Social de España supplies rankings of provinces and cities, based on eight social measurements: population characteristics, territorial distribution of population, demography, migration, housing, employment, children and third age.

Following this line of research, this paper intends to rank the 20 Italian regions according to their efficiency in terms of jointly maximizing the level of: a) leisure satisfaction; b) family relations; c) friendship and d) employment. The analysis is based on 12 indicators in several fields, such as economic, individual perception, social, neighbourhood and environment. The methodology is based on data envelopment analysis and the Malmquist productivity index (MPI). These procedures allow one to derive a virtual input and output computed within a multi-factor framework and to compare the regions.

To the best of our knowledge this is the first time this approach has been used to develop environmental efficiency measurements for the Italian regions (a recent exception is the work of Bernini et al. 2013 on the Emilia Romagna region). Italy is strongly characterised by consistent and persistent differences between the North and South for almost all economic and social indicators. One relevant objective of this work is to investigate if there is also such dualism for wellbeing. The results reveal that northern regions have been performing better in wellbeing than the southern ones, highlighting the strong geographical differences.

We also use the Self-Organizing Map (SOM) method to help the visual analysis of spatial data. SOM is a combination of clustering and dimension reduction. This makes SOM useful for illustrating low-dimension views of high-dimension data, akin to multidimensional scaling. This method was first described as an artificial neural network by Kohonen (1982) and it has since been applied in the geospatial data analysis domain, where its data aggregation and sorting properties are leveraged.

Clustering techniques can be employed for identifying structure in an unlabelled data set, by objectively organizing data into homogeneous groups, where the within-group-object dissimilarity is minimized and the between-group-object dissimilarity is maximized. As the SOM network itself represents an abstract map for data, colour-coding can be used to link the location of data elements in SOM space with their respective geospatial coordinates (Koua and Kraak 2008; Spielman and Thill 2008).

In this work we use the Self-Organizing Map procedure to cluster regions into homogeneous groups, where the within-group-object dissimilarity is minimized and the between-group-object dissimilarity is maximized. Cluster analysis confirms that there is a marked duality in regional wellbeing in Italy.

The paper is organised in the following way: section 2 describes the data set and the variables used, section 3 outlines the DEA and Malmquist methodology, section 4 contains the results, section 5 supplies a brief description of the Self-Organizing Map application. Section 6 concludes.

2 Data description and variables employed

In our analysis we use data collected from different databases. These were: Sole 24 Ore, National Institute of Statistics (ISTAT), Istituto Tagliacarne, and Legambiente. The indicators used in this work represent some of the dimensions of wellbeing, such as economic conditions, labour market conditions, neighbourhood relations and environmental aspects. More specifically, the analysis is based on four variables, which are maximised by the linear programming.

Three indicators express the level of satisfaction for leisure (LEIS), family relations (FAM) and friendship (FRIEN). The fulfilment of psychological needs and leisure activities affects individual wellbeing, because it supplies opportunities for the individual to establish social relationships, feel positive emotions, acquire additional skills and knowledge, and consequently improve their quality of life (Lloyd and Auld 2002; Leung and Lee 2005; Nimrod and Adoni 2006; Iwasaki 2007; Rodriguez et al. 2008; Brajša-Žganec et al. 2011; Dolnicar et al. 2011). The last output variable is the employment indicator (EMPL) used as a proxy for the level of economic conditions in the region.

The input variables are the regional pro-capita GDP (INCOM) and a measure of income inequality (INEQ). For the latter some works have investigated the correlation between income distribution and individual happiness (Morawetz et al. 1977; Alesina et al. 2001). An ecological indicator is also included among the inputs (ENVIR). This is based on Indice Legambiente Ecosistema urbano (2012) which employs 25 thematic indexes based on more than 120 parameters. Pioneering work by Dasgupta (2000) and Krutilla and Reuveny (2002) highlight the importance of the natural environment when measuring well-being.

The other inputs are the number of unemployed people per family (FAM_UNEM), the number of not studying and not working young people (NO_WORK_NO_STUDY), a measurement of cultural facilities (CULTU), an indicator of failures in the water supply system as a proxy for the level of supply of general public service (WATE) and an indicator of life style (SEDEN).

Given the nature of the DEA specifications, in order to maximize the process in a coherent way, desirable inputs (INCOM, CULTU, ENVIR, WATE) are taken as their inverse value (Scheel 2001; Seiford and Zhu 2002; Fare and Grosskopf 2004; Hua and Bin 2007). The characteristic of undesirability of some variables has been used in Färe et al. (1986); Tyceta (1996); Chung et al. (1997); Zofio and Prieto (2001). However, although this approach is widely accepted among the environmental economists it has also been criticised (Färe and Grosskopf 2003; Färe and Grosskopf 2009; Kuosmanen and Podinovski 2009). Several works based on the production process with different data sets have used distance functions to construct desirable and undesirable outputs and thus to measure the environmental performance of different units (Färe et al. 1989a,b; Chung et al. 1997; Zaim and Taskin 2000; Zaim 2004).

Table 1 shows the statistics for all the variables for the period 2005-2011. It clearly emerges that all the indicators are consistently worse in the southern part of the country. To be more precise, the indicators for the South as a percentage of those for the North are¹: leisure (68%), family relations (74%), friendship (74%), employment (74%), income (83%), the index of inequality (120%), unemployed people per family

¹ As fig. 1 shows, in this analysis the North consists of 12 regions (Piedmont, Valle d'Aosta, Lombardy, Liguria, Trentino, Friuli, Veneto, Emilia, Tuscany, Umbria, Marche, Lazio) and the South consists of 8 regions (Abruzzo, Molise, Campania, Apulia, Basilicata, Calabria, Sicily, Sardinia).

(320%) young people not studying and not working (203%), cultural facilities (0.82%), the environmental indicator (0.35%), water supply default (2.60%) and sedentary (1.60%).

3. Data Envelopment Analysis: general settings

Composite indicators have been increasingly used for analysing wellbeing. Such indicators are an aggregation of a set of sub-indicators. These quantify multi-dimension aspects cannot be represented by a single item. Their construction, however, requires that individual indicators have to be measured, and that a 'weighting and aggregation' technique needs be defined. This exercise involves a certain degree of subjectivity and this has direct effects on the quality of the indicator. An optimization procedure that identifies endogenous and entity-common weights and determines the sub-indicators weights may provide a solution to this problem of subjectivity. The DEA procedure is reasonable way of selecting the most advantageous factor weight' vector for individual decision-making units, without requiring prior knowledge of sub-indicator weights.

The attribution of weights ought be performed objectively, while respecting the subjectivity with which the distinct groups and individuals might interpret their preferences. DEA manages to meet these two requirements. On one hand, it is an objective tool, since it does not require the attribution of a priori weights. On the other, it includes an element of subjectivity, since it is more flexible when it comes to setting the weights for each of the units that are being compared (Murias et al. 2006). The main advantage of this technique is that once the objective is defined, it automatically determines the structure of the weights by selecting the factor weights that are the most advantageous for each decision making units (DMU) when they are calculating their wellbeing. The DEA maximises the composite indicator value of each entity, subject to the constraint that the obtained weights produce consistent results for all the DMUs (Bernini et al. 2013). However, DEA models are run separately for each DMU and, the set of weights is different for the various sub-indicators and DMUs. Hence the flexibility in selecting each single DMU's optimal weights does not allow a common base to be established which can be used for comparing DMUs.

Although DEA was initially developed for computing production efficiency, some applications of this procedure have been employed in empirical work that focuses on its properties as an aggregating tool. The aggregation is carried out by comparing the indicators for each unit with the best practices observed in the frontier of reference. The efficient production frontier allows one to calculate and compare a firm's efficiency with its own benchmark. The calculation of the efficiency of the DMUs is accomplished by using linear programming methods to encompass observed input-output vectors as tightly as possible. To be more precise, it may be an appropriate tool for quantifying the efficiency of a group of decision making units over time, as well as the performance of the relatively most productive decision units within the sample set. Employing a panel DEA also helps us to understand how well units process their inputs, compared to their past performances and their own particular benchmark. Furthermore, it is also possible to identify the source of inefficiency, thus providing useful information to economic agents, and helping them to improve their performance.

The variables employed in the DEA analysis are inputs (factors that ought to be kept to a minimum) and outputs (factors that have a positive value and ought to be increased to their maximum). DEA consistently weighs inputs and outputs, in order to compute an index of productive efficiency. Two indicators, pure technical efficiency and scale efficiency scores, relate the performance of each DMU to the estimated production frontier. This frontier represents all those points at which a relatively optimal capacity of transformation of inputs into outputs is achieved. Other methods for estimating the efficiency frontier are also available such as index numbers and the application of stochastic frontier production functions (Aigner et al. 1977; Meeusen and van der Broeck 1977). However non-parametric techniques are the most widely used, because they do not require an a priori specification of the functional form of the production function for the units under investigation and also have some other advantages, as discussed above.

The DEA model can be divided into the input-oriented model, which minimizes inputs for a given level of output in order to achieve full technical efficiency, and the output-oriented model, which maximizes outputs without requiring any of the observed input values to be expanded. DEA analysis allows multiple inputs and outputs to be employed at the same time without any assumption being made on data distribution. Efficiency is quantified in terms of the proportional change in inputs or outputs. The variables of interest can also be employed jointly, regardless of their scale of measurement (Köksal and Aksu 2007). Within a given sample of productive units, one subgroup is given a relative efficiency equal to 1 (or 100%) and the other DMUs are considered as inefficient if they have a score of less than 1. The output-oriented model can be

used for planning and strategic objectives (Cullinane et al. 2004). For example, it is useful in helping decision units to establish whether an expansion of their capacity is feasible, given the level of the inputs employed. Thus an output-oriented approach is generally more appropriate for estimating capacity and capacity utilization.

Although there are several types of DEA programs, in this paper we use variable returns to scale frontier (Banker et al. 1984). This is because the regions used in our analysis are of different sizes, and thus it is very likely that not all the regions are operating at an optimal scale. The less restrictive VRS frontier allows the best practice level of outputs to inputs to vary across the units in the sample. The VRS model has also been widely used in the last two decades.

Following Coelli (1994), let assume we have data on K inputs and M outputs for each of N DMUs. For the i^{th} DMU these are represented by the vectors x_i and y_i , respectively. The $K \times N$ input matrix, X , and the $M \times N$ output matrix, Y , represent the data of all N DMU's. The objective of DEA is to obtain a non-parametric envelopment frontier such that all observed points lie on or below the production frontier. For the output-orientated VRS model, DEA can be expressed as follows:

$$\begin{aligned}
 & \max_{\phi, \lambda} \phi, \\
 \text{st} \quad & Y\lambda - \phi y_i \geq 0, \\
 & x_i - X\lambda \geq 0, \\
 & \mathbf{N}\mathbf{1}'\lambda = 1, \\
 & \lambda \geq 0,
 \end{aligned} \tag{1}$$

where λ is an $N \times 1$ vector of constants and $\mathbf{N}\mathbf{1}$ is an N -order column vector of ones, $1 \leq \phi < \infty$, and $\phi - 1$ is the proportional increase in outputs that could be achieved by DMU _{i} while input quantities are held constant. The term $(\mathbf{N}\mathbf{1}'\lambda = 1)$ represents the convexity constraint. It ensures that an inefficient production unit is only compared with production units of a similar size (for the constant returns to scale case, this constraint is excluded, the λ weights sum up to a value different from one and the benchmarking may occur against production units that are substantially different from the production unit being analysed (Färe et al, 1989). It is worth mentioning that $1/\phi$ defines the technical efficiency (TE) score, which varies between zero and one (and that this is the output-orientated TE score reported by DEAP). TE=1 implies that the performance of the i^{th} unit is (relatively) fully efficient and lies on the best-practice frontier.

3.1. The Malmquist Productivity Index

The Malmquist Productivity Index (MPI) was first proposed by Malmquist (1953) as a quantity for use in the analysis of consumption of inputs. Färe et al. (1994) developed a DEA-based Malmquist productivity index. This measures productivity change over time. Basically, MPI can be used for estimating the performance change between two points in time by calculating the ratio of the distance of each point from a common technology. The application of Malmquist DEA methods to panel data allows one to evaluate the dynamic performance of the DMUs over time. The choice of this method is based on the fact that regions may require more than one year to arrive at the output levels given by the input factors. As in the case of a moving average approach, regions in different years are treated as if they were different DMUs. This allows one to compare the efficiency of a DMU with its own efficiency in other years, as well as with the efficiency of the other DMUs' (Bosetti et al. 2003). Following Coelli et al. (2005), MPI between period, s , and the reference period, t , can be expressed as follows:

$$m_0^t(q_s, x_s, q_t, x_t) = \frac{d_0^t(q_t, x_t)}{d_0^t(q_s, x_s)} \tag{2}$$

This represents the productivity of the production point (q_t, x_t) relative to the production point (q_s, x_s) . $m_0 > 1$ indicates that there is a positive TFP growth from period, s to period, t . The opposite occurs for $m_0 < 1$.

To avoid either imposing this restriction or arbitrarily choose one of two performance changes between two periods, output-based Malmquist productivity is conveniently defined as the geometric mean of these

two output-based indices (Fare et al. 1994). One index uses period s technology and the other period t technology:

$$m_0(q_s, x_s, q_t, x_t) = \left[\frac{d_0^s(q_t, x_t)}{d_0^s(q_s, x_s)} \cdot \frac{d_0^t(q_t, x_t)}{d_0^t(q_s, x_s)} \right]^{1/2} \quad (3)$$

The application of Malmquist DEA methods to panel data measures productivity changes and their changes over time, and can be broken down into changes in efficiency and technology. The distance function in this productivity index can be broken down into two components, with one measuring the technical efficiency change index and the other measuring the index of technical change (Fare et al. 1994, Coelli 1996):

$$m_0(q_s, x_s, q_t, x_t) = \frac{d_0^t(q_t, x_t)}{d_0^s(q_s, x_s)} \left[\frac{d_0^s(q_t, x_t)}{d_0^s(q_s, x_s)} \cdot \frac{d_0^t(q_t, x_t)}{d_0^t(q_s, x_s)} \right]^{1/2} \quad (4)$$

The distance functions cannot be evaluated without knowing the frontier production set. Several methods have been used to estimate this frontier (among others index numbers, stochastic frontier production functions and non-parametric techniques, such as data envelopment analysis). The DEA approach which we use in this study, is one of the possible available methods.

4. Results of the Panel DEA and the Malmquist Index

The model employed in this work is based on 12 indicators and aims to rank the 20 Italian regions according to their scores for wellbeing, obtained from the DEA analysis. The model implicitly assumes that each region is one DMU, and that inputs and outputs are partial indicators, characterized by the maxims “the fewer the better” and “the more the better”, respectively (Murias et al. 2006).

The variables that we considered are three indicators which express the level of satisfaction for leisure, family relations and friendship, and an indicator of employment. These are the outputs which need to be maximised by the linear programming.

The input variables are: the regional pro-capita GDP, a measure of income inequality, the number of unemployed people per family, young people not working and not studying, cultural facilities, an ecological indicator, the water supply system as a proxy for the general level of public services and, sedentary as indicator of life style.

DEA model:

Outputs: LEIS, FAM, FRIEN, EMPL

Inputs: INCOM, INEQ, FAM_UNEM, NO_WORK_NO_STUDY, CULTU, ENVIR, WATE, SEDEN

As seven years' worth of data are used (2005-2011), there are 1680 observations (DMUs) (12 variables for 20 regions over 7 years). Over-fit problems are avoided, as the number of DMUs is more than twice the total number of inputs and outputs in the DEA (Min et al. 2008).

Using the DEA procedure, we begin the empirical analysis by identifying the scores for wellbeing and the best-practice frontier, This is defined as the set of the most efficient combinations of outputs and inputs. This is a necessary step in identifying the regions that are located on the frontier. Using the scores obtained from the DEA, a geometric mean of two alternatives Malmquist index is then calculated (Färe et al. 1994), so as to identify the score dynamics during the period under consideration.

The 2.1 Version of the DEAP software package is used for running the DEA panel analysis (Coelli 1996). This first calculates the distances (or technical efficiencies) necessary for the Malmquist calculations. Four distances are calculated for each firm in each year. These are relative to:

1. the previous period's CRS DEA frontier;
2. the current period's CRS DEA frontier;

3. the next period's CRS DEA frontier;
4. the current period's VRS frontier.

Table 2 summarizes the results of the output oriented DEA (VRS). The Scores are computed for a three successive year base period, moving progressively across the eight years considered. This allows for dynamic effects to be included when measuring efficiency in cross sectional and time varying data (Charnes and Cooper 1985). For 7 years period under analysis, the results show that there were marked inefficiencies in wellbeing in the 20 regions. Regions in the North of Italy had higher efficiency scores than did those in the South. To be more precise, Trentino and Veneto had the best performances and are both on the best practice frontier (TE=1). This can be interpreted as meaning that they could not make any (relative) improvements, given the data and the structure of the DEA program. Very closely behind these two regions are Aosta, Emilia and Friuli with efficiency score of 0.992, 0.989 and 0.983 respectively. From the same table it clearly emerges that Sicily, Calabria, Apulia and Campania consistently lag behind, as do the other Southern regions. As can be easily seen, regions in the south consistently lag behind during the whole period. The overall wellbeing efficiency score in the south is only 74% of that in the north.

The table also shows the results of the application of the Malmquist DEA model. This measure of wellbeing changes over time, and shows that, on average, wellbeing declined in both the northern and the southern regions, during the period, albeit by only a modest amount (0.998 and 0.996 respectively).

5. Self-organizing map neural network

In order to support the visual analysis of spatial data, this paper employs the Self-Organizing Map (SOM) method. SOM is as a combination of clustering and dimensionality reduction which allows for visualizing low-dimensional views of high-dimensional data, akin to multidimensional scaling. Clustering techniques are used to identify structure in an unlabelled data set by objectively organizing data into homogeneous groups. As the SOM network itself represents an abstract map for data, color-coding can be used to link the location of data elements in SOM space with their respective geospatial coordinates (Koua and Kraak 2008; Spielman and Thill 2008). This method was first employed as an artificial neural network by Kohonen (1982a,b) and it has found several successive applications in the geospatial data analysis field. Recent applications can be found in Hsu and Li (2010), Andrienko et al (2010), Chon (2011).

Kohonen nets are artificial neural networks which adapt themselves in responses to input signals and on the basis of the Kohonen algorithm. They consist of nodes, which are spread uniformly on a grid and are “functionally” connected to their neighbouring nodes. In most cases, two dimensional grids are used, however, Kohonen nets with one-dimensional or multi-dimensional grids are possible as well. In the following we will look at two-dimensional Kohonen nets which are also used in the Viscovery products (Viscovery SOMine lite 1998).

Each node has weight vector w^p . The components of this vector represent the strength of the synapse connections to the “input” neurons. The Kohonen algorithm enables these weights to adapt themselves in response to the input signals (self-organization).

The nodes are competing with each other. In particular if all of them are presented the same input signal (denoted by vector x^p), the node with the strongest response is the winner. The response of a node is defined as the distance $|w^p - x^p|$ between the vectors w^p and x^p . The closer the weight vector w^p of a node to the input vector x^p , the greater the response. The Kohonen algorithm can, thus defined as follows:

The winner node c changes its weight vector w_c^p to become more similar to the input vector x^p . All neighbours of c which lie within a predefined distance to the winner node also change their vectors to the direction x^p . This modification is proportional to the difference between the input vectors x^p , and the corresponding weight vector (the proportionality factor α is called the learning rate).

The weight vectors of the nodes are distributed throughout the entire data space if enough different data vectors are presented to the net. Thus, the weight vectors of neighbouring nodes become more and more similar as a result of their convergence with the winning node towards the input data vector. Consequently, the Kohonen net is ordered. After competition of the learning process, neighbouring nodes have similar values regarding the original data space.

The above process can mathematically be seen as a non-linear, non-parameterized regression. There is a corresponding Error Function $E(w^p)$ with an expectation value converging to a minimum during the training process (distortion measure):

$$E = \int \sum_i h_{ci} |w_i - x_i| g(x) d^n x, \quad (4)$$

Where h_{ci} is the neighbouring function of node i to the corresponding winner $c(x^\rho)$ (e.g. an exponential function) and $g(x^\rho)$ the density function of the vectors x^ρ in the n -dimensional data space. The Kohonen algorithm is obtained in a discrete data space by computing the optimal weight vectors (for minimizing $E(w^\rho)$) by gradient descent.

Quite a few improvements have been made to the Kohonen algorithm. They are used in state-of-art software based on the above mentioned organization principle and are known under the general term of Self-organizing Maps.

Concerning the present work, the SOM procedure supplies maps that are obtained employing the 12 indicators used in the DEA analysis. The data vectors in table 1 were spread over 35×7 or, equivalently, in 241 nodes (neurons) that are most similar were grouped together, while vectors that are most different appear further away from each other on the map. The purpose is not to detect non-linear dependencies between the indicators (inputs), but, to identify patterns of similarity or dissimilarity among the process variables in the 20 Italian regions under analysis. The first map shows the clustering into six zones (Fig 2). It can be noted that the region of Trentino represents a single zone meaning that it is different (better position in the map) from all the rest of the country.

An analysis of the main differences among these six zone is reported in table 2. It is straightforward to observe that the difference between the best and the worst performing cluster identified by the SOM analysis (cluster 1 and cluster 6 respectively) in all the 12 indicators employed. The indicator of leisure for instance is more than twice in cluster 1, the family relation and friendship satisfaction, the employment and the income indicators are about 50% bigger in cluster 1 when compared to cluster 6.

This comparison follows a similar pattern with respect to inequality (1.5 times in cluster 6) FAM_UNEM is 6 times bigger in cluster 6, the culture indicator is 57%, the environment and the water service indicators are about 11 times smaller, the sedentary indicator is 3.5 bigger.

A part from the differences among the six zones, there also emerges a clear geographic divide between the north and the south of Italy. Regions in the north are located in the left side of the Viscosity SOM map, the central regions are mostly depicted in middle of the map while the southern regions are all in the right side of the map. Hence the zones identified by the SOM analysis on the base of the process variables correspond quite well to the geographical locations of the regions and are in line with the DEA results. An exception is represented by the Veneto region which though being in the DEA efficiency frontier together with Trentino it has been grouped in the immediately following cluster 2. The other exception is Lazio and Sardinia which both belong to cluster 4 while their DEA efficiency position is 7 and 14 respectively.

The component windows (fig. 3) display the distribution of the values of the respective component over the map, thereby representing a cross-section through the map. The colour scale at the bottom of the a Component window match the colours of the nodes with the value of the respective component at that node. Blue is used for small values, green for mid range values and red for high values. The colour changes gradually among the clusters, similarity of indicators types is reflected in the similarity of colours. For instance, the component planes show the influence of the inverse of income indicator, blue is high income and red low income. From the component windows it can be seen that 12 process variables are all dominant variables since they appear to be distributed uniformly. Hence they all play an important role in the overall distribution.

From the single component analysis (fig. 3) it can be seen that the desirable inputs: LEIS, FAM, FRIEN, EMPL are concentrated in the northern regions and are indicated as red zones (high values). Conversely, the same inputs are indicated as blue zones (low values) for the southern regions. The same applies for the other desirable inputs: INCOM, CULTU, ENVIR, which as described above, are taken as their inverse value in order to make the DEA process of maximization coherent (colours are obviously reversed). In the same way, the “undesirable” inputs (INEQ, FAM_UNEM, NO_WORK_NO_STUDY, WATE, SEDEN) are represented by blue zones in the northern regions while they are represented by red zones in southern regions. This strongly highlights a clear geographic divide between the two macro-areas of the country.

6. CONCLUSIONS

This work uses data envelopment analysis (DEA) to ascertain the wellbeing performance and ranking of the 20 Italian regions according to their efficiency in terms of: a) leisure satisfaction; b) family relations; c) friendship and d) employment. Malmquist indices (MPI) obtained from the DEA scores are then used to measure changes in wellbeing over time. These procedures allow one to derive a virtual input and output result, computed within a multi-factor framework, and to compare the regions. Italy is strongly characterised by consistent and persistent differences between the North and South for almost all economic and social indicators. One of the objectives of this work is to discover whether or not such dualism is also true for wellbeing.

The analysis is based on 12 indicators. These represent some of the dimensions that are linked to wellbeing, such as economic conditions, labour market conditions, neighbourhood relations and environmental aspects. The period under consideration runs from 2005 to 2011. The results show that, considering leisure, family relations, friendship, and employment jointly as variables to be maximized in order to create a proxy for wellbeing, northern regions have been performing better than the southern ones, indicating the clear geographical division in the country. The Malmquist procedure also shows that Italy has not succeeded in improving wellbeing.

We also use the Self-Organizing Map method in our visual analysis of the spatial data. SOM is a combination of clustering and dimension reduction. This makes it useful for visualizing low-dimensional views of high-dimensional data, akin to multidimensional scaling. We apply the Self-Organizing Map procedure to cluster regions into homogeneous groups, where the within-group-object dissimilarity is minimized and the between-group-object dissimilarity is maximized. Cluster analysis confirms that there are marked regional differences in wellbeing in Italy.

References

- Aigner, D.J., Lovell, C.A.K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production models. *Journal of Econometrics*, 6, 21-37.
- Alesina, A., Di Tella, R., & McCulloch, R. (2001). Inequality and happiness: Are European and American different?’, NBER Working Paper, No 8198, Cambridge.
- Ali, A.I., (1994). Computational aspects of Data Envelopment Analysis. In DEA: Theory, methodology and applications, edited by A. Charnes, W.W. Cooper, A.Y. Lewin, and L.M. Seiford, 63-88. Boston: Kluwer Academic Publishers.
- Andrews, F.M., & Withey, S.B. (1976). Social indicators of well-being: America’s perception of quality of life. New York: Plenum Press.
- Anuario Social de España, Fundación La Caixa, Barcelona: various years.
- Banker, P.C., Charnes, A., & Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30, 1078-1092.
- Bosetti, V., Cassinelli, M., & Lanza, A. (2003). Using data envelopment analysis to evaluate environmentally conscious tourism management. In: International Conference on Tourism and Sustainable Development, CRENoS Cagliari, Sassari University and World Bank, September 19-20, Chia, Sardinia.
- Boyer, R., & Savageu, D. (1981). Places Rated Almanac: Your guide to finding the best places to live in America (Rand McNally, New York).
- Brajša-Žganec, A., Merkaš, M., & Šverko, I. (2011). Quality of life and leisure activities: How do leisure activities contribute to subjective well-being? *Social Indicators Research*, 102(1), 81-91.
- Campbell, A., Converse, P.F., & Rodgers, W.R. (1976). The quality of American life. New York: Sage.
- Charnes, A., & Cooper, W.W. (1985). Preface to topics in data envelopment analysis. *Annals of Operation Research*, 2, 59-94.

- Chon ,Tae-Soo (2011). Self-Organizing Maps applied to ecological sciences. *Ecological Informatics*, 6, 50-61.
- Christakopoulou, S., Dawson, J., & Gari, A. (2001). The community well-being questionnaire: Theoretical context and initial assessment of its reliability and validity. *Social Indicators Research*, 56(3), 321-345.
- Chung, Y.H., Färe, R., & Grosskopf, S. (1997). Productivity and undesirable outputs: A directional distance function approach. *Journal of Environmental Management*, 51, 229-240.
- Cobb, C., Halstead T. & Rowe J. (1995), The genuine progress indicator: summary of data and methodology (Redefining Progress, San Francisco).
- Coelli, T, Rao D.S.P, O'Donnel, C.J, Battese G.E. (2005). An introduction to efficiency and productivity analysis, 2nd edn. Springer, Berlin.
- Coelli, T.J. (1996). A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program”, University of New England, Department of Econometrics, CEPA Working Paper No. 8/96, Armidale, NSW 2351, Australia. (Available at: <http://www.une.edu.au/econometrics/cepa2.htm#software>).
- Cook, W.D., & Kress, M. (1990). A data envelopment model for aggregating preference rankings. *Management Science*, 36(11), 1302-1310.
- Bernini, C., Guizzardi, A. & Angelini, G. (2013), DEA-Like Model and Common Weights Approach for the Construction of a Subjective Community Well-Being Indicator. *Social Indicators Research*, 114, 405-424.
- Cullinane, K., Song, D.W., & Wang, T. F., 2004. An application of DEA windows analysis to container port production efficiency. *Review of Network Economy* 32, 184-206.
- Dasgupta, P. (2000). Population resources and Welfare: An exploration into reproductive and environmental externalities, Working paper.
- Diener, E. (1984). Subjective well being. *Psychological Bulletin*, 95(3), 542-575.
- Diener, E., & Suh, E. (1997). Measuring quality of life: economic, social and subjective indicators. *Social Indicators Research*, 40, 189-216.
- Diener, E., Oishi, S., & Lucas, R. (2003). Personality, culture and subjective well-being: Emotional and cognitive evaluations of life. *Annual Review of Psychology*, 54, 403-425.
- Diener, E., Suh, E.M., Lucas, R.E., & Smith, H.L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, 125(2), 276-302.
- Dolnicar, S., Yanamandram, V., & Cliff, C. (2011). The contribution of vacations to quality of life. *Annals of Tourism Research*, 39(1), 59-83.
- González, E, Cárcaba, A., & Ventura, J. (2013).The Importance of the Geographic Level of Analysis in the Assessment of the Quality of Life: The Case of Spain, *Social Indicators Research*, 102, 209-228.
- Erikson, R., Hansen, E.J., Ringen S. & Uusitalo, H. (1987), The Scandinavian Model: Welfare States and Welfare Research (M.E.Sharpe, New York).
- Erikson, R. (1993). Description of inequality: The Swedish approach to welfare research, in M. Nussbaum and A. Sen (eds.), *The Quality of Life* (Clarendon Press, Oxford).
- Färe, R., & Grosskopf, S. (2003). Non-parametric productivity analysis with undesirable outputs: comment. *American Journal of Agriculture Economics*, 85, 1070-1074.
- Fare, R., & Grosskopf, S. (2004). Modelling undesirable factors in efficiency evaluation: Comment. *European Journal of Operational Research*, 157, 242-245.
- Färe, R., & Grosskopf, S. (2009). A comment on weak disposability in nonparametric production analysis. *American Journal of Agriculture Economics*, 91, 535-538.

- Färe, R., Grosskopf, S., & Pasurka, C. (1986). Effects on relative efficiency in electric power generation due to environmental controls. *Resources and Energy*, 8, 167-184.
- Färe, R., Grosskopf, S., Pasurka, C. (1989). The effect of environmental regulations on the efficiency of electric utilities: 1969 versus 1975. *Applied Economics*, 21, 225-235.
- Färe, R.S., Grosskopf, S., & Lovell, C.A.K. (1994). *Production Frontiers*. Cambridge University Press.
- Andrienko G., Andrienko, N., Bremm, S., Schreck, T., von Landesberger, T., Bak, P.& Keim, D. (2010), Space-in-Time and Time-in-Space Self-Organizing Maps for Exploring Spatiotemporal Patterns, Eurographics/ IEEE-VGTC Symposium on Visualization, Volume 29(3).
- Gillingham, R. & Reece W.S. (1979), A new approach to quality of life measurement. *Urban Studies* 16(3), 329-332.
- Gonzalez, E., Carcaba, A., & Ventura, J. (2011). The importance of the geographic level of analysis in the assessment of the quality of life: The case of Spain. *Social Indicators Research*, 102(2), 209-228.
- Hashimoto, A., & Ishikawa, H. (1993). Using DEA to evaluate the state of society as measured by multiple social indicators. *Socio-Economic Planning Sciences*, 27, 257-268.
- Hashimoto, A., & Kodama, M. (1997). Has liveability of Japan gotten better for 1956-1990? A DEA approach. *Social Indicators Research*, 40, 359-373.
- Hashimoto, A. (1999), Proposing non-uniform evaluation in social systems analysis, Discussion Paper Series, 827 (Institute of Policy and Planning Science, University of Tsukuba).
- Hsu, K.C., & Li, S.T. (2010). Clustering spatial-temporal precipitation data using wavelet transform and self-organizing map neural network. *Advances in Water Resources*, 33, 190-200.
- Hua, Z., & Bin, Y. (2007). DEA with undesirable factors. In: Zhu, J., Cook, W.D. (Eds.), *Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis*. Springer Science Series (Chapter 6).
- Istituto Tagliacarne, various years, <http://www.tagliacarne.it/>
- Iwasaki, Y. (2007). Leisure and quality of life in an international and multicultural context: What are major pathways linking leisure to quality of life? *Social Indicators Research*, 82, 233-264.
- Jurado, A., & Perez-Mayo, J. (2012). Construction and evolution of a multidimensional well-being index for Spanish Regions. *Social Indicators Research*, 107(2), 259-279.
- Kahneman, D., & Krueger, A. (2006). Developments in the measurement of subjective well being. *Journal of Economic Perspective*, 20, 3-24.
- Kohonen, T. (1982,a). Self-organized formation of topologically correct feature maps. *Biol. Cybern.* 43, 59-69.
- Kohonen, T. (1982,b). Analysis of a simple self-organizing process. *Biol. Cybern.* 44, 135-140.
- Köksal, C.D., & Aksu, A.A. (2007). Efficiency evaluation of a group travel agencies with data envelopment analysis DEA: a case study in the Antalya region, Turkey. *Tourism Management*, 28, 830-834.
- Koua, E.L., & Kraak, M.J. (2008). An integrated exploratory geovisualization environment based on self-organizing map. In *Self-Organising Maps: Applications in Geographic Information Science*, Agarwal P., Skupin A., (Eds.). Wiley, 45-66.
- Krutilla, K. & R. Reuveny: 2002, The quality of life in the dynamics of economic development. *Environment and Development Economics*, 7, 23-45.
- Kuo-Chin, Hsu, & Sheng-Tun, Li (2010). Clustering spatial-temporal precipitation data using wavelet transform and self-organizing map neural network. *Advances in Water Resources*, 33, 190-200.
- Kuosmanen, T., & Podinovski, V. (2009). Weak disposability in nonparametric production analysis: reply to Färe and Grosskopf. *American Journal of Agricultural Economics*, 91, 539-545.

- Lawless, N., & Lucas, R. (2010). Predictors of regional well-being: A county level analysis. *Social Indicators Research*, 101, 341-357.
- Legambiente Ecosistema Urbano (2012), XIX Rapporto sulla qualità ambientale dei comuni capoluogo di provincia.
- Leung, L., & Lee, P.S.N. (2005). Multiple determinants of life quality: The roles of Internet activities, use of new media, social support, and leisure activities. *Telematics and Informatics*, 22(3), 161-180.
- Liu, B.C. (1976). *Quality of Life Indicators in U.S. Metropolitan Areas: A Statistical Analysis* (Praeger Publishers, New York).
- Lloyd, K.M., & Auld, C.J. (2002). The role of leisure in determining quality of life: Issues of content and measurement. *Social Indicators Research*, 57, 43-71.
- Malmquist, S. (1953). Index numbers and indifference surfaces. *Trabajos de Estadística*, 4, 209-242.
- Meeusen, W., & van den Broek, J. (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review* 18(2), 435-444.
- Min, H., Min, H., & Joo, S.J. (2008). A data envelopment analysis-based balanced scorecard for measuring the comparative efficiency of Korean luxury hotels. *International Journal of Quality and Reliability Management*, 25, 349-365.
- Morawetz, D. (1977). Income distribution and self-rated happiness: Some empirical evidence. *The Economic Journal*, 87, 511-522.
- Murias, P., Martínez, F., & Miguel, C. (2006). An economic well-being index for the Spanish Provinces. A data envelopment analysis approach. *Social Indicators Research*, 77(3), 395-417.
- Nimrod, G., & Adoni, H. (2006). Leisure-styles and life satisfaction among recent retirees in Israel. *Ageing and Society*, 26, 607-630.
- Nordhaus, W.D. (2002), The health of nations: the contribution of improved health to living standards, in K.M. Murphy and R.H. Topel (eds.), *Exceptional Returns* (University of Chicago Press, Chicago).
- Osberg, L. and Sharpe, A. (1998). An index of economic well-being for Canada. Applied Research Branch, Research Paper R-99-3E, Human Resources Development Canada, Ottawa, Ontario.
- Osberg, L. & Sharpe, A. (1999), An index of economic well-being for Canada and the United States, paper presented at the annual meeting of the American Economic Association, New York, January.
- Pena, J.B. (1977). Problemas de la medición del bienestar y conceptos afines. Una aplicación al Caso Español. (I.N.E.: Madrid).
- Roback, J. (1982). Wages, rents, and quality of life, *Journal of Political Economy*, 90, 1257-1278.
- Rodriguez, A., Làtková, P., & Sun, Y.Y. (2008). The relationship between leisure and life satisfaction: Application of activity and need theory. *Social Indicators Research*, 86, 163-175.
- Rogerson, R.J. (1999). Quality of life and city competitiveness. *Urban Studies*, 36(5-6), 969-985.
- Rosen, S. (1979). Wage-based indexes of urban quality of life, in P. Mieszkowski and M. Straszheim (eds.), *Current Issues in Urban Economics* (John Hopkins Press, Baltimore), 74-104.
- Scheel, H. (2001). Undesirable outputs in efficiency valuations. *European Journal of Operational Research*, 132, 400-410.
- Seiford, L., & Zhu, J. (2002). Modelling undesirable factors in efficiency evaluation. *European Journal of Operational Research*, 142, 16-20.
- Sen, A. (2000). *Desarrollo y libertad*, (Editorial Planeta, Barcelona).
- Sirgy, M.J. (2002). *The psychology of quality of life*. Dordrecht, The Netherlands: Kluwer.
- Sirgy, M.J., & Cornwell, T. (2001). Further validation of the Sirgy et al.'s measure of community quality of life. *Social Indicators Research*, 56(12), 5-143.

- Sirgy, M.J., & Cornwell, T. (2002). How neighborhood features affect quality of life. *Social Indicators Research*, 59(1), 79-102.
- Sirgy, M.J., Gao, T., & Young, R. F. (2008). How residents' satisfaction with community services influence quality of life (QOL) outcomes? *Applied Research in Quality of Life*, 3(2), 81-106.
- Sirgy, M.J., Rahtz, D., Cicic, M., & Underwood, R. (2000). A method for assessing residents' satisfaction with community-based services: A quality-of-life perspective. *Social Indicators Research*, 49, 279-316.
- Sirgy, M.J., Widgery, R.N., Lee, D.J., & Yu, G.B. (2010). Developing a measure of community wellbeing based on perceptions of impact in various life domains. *Social Indicators Research*, 96(2), 295-311.
- Slottje, D. (1991), Measuring the quality of life across countries, *The Review of Economics and Statistics* 73(4), 684-693.
- Sole24ore, various years, Dossier sull'Italia: Qualità di vita, www.ilsole24ore.com.
- Spielman S. E., & Thill J.C. (2008). Social area analysis, data mining, and gis. *Computers, Environment and Urban Systems*, 32, (2), 110-122.
- Townsend, P. (1979), Poverty on the United Kingdom (Penguin, London).
- Tyceta, D. (1996). On the measurement of the environmental performance of firms: a literature review and a productive efficiency perspective. *Journal of Environmental Management*, 46, 281-308.
- Van Praag, B. (2007). Perspectives from the happiness literature and the role of new instruments for policy analysis. CESifo. *Economic Studies*, 53, 42-68.
- Van Praag, B. (2011). Well-being inequality and reference groups: An agenda for new research. *Journal of Economic Inequality*, 9, 111-127.
- Van Praag, B.M.S., Frijters, P., & Ferrer-i-Carbonell, A. (2003). The anatomy of subjective well-being. *Journal of Economic Behaviour & Organization*, 51, 29-49.
- Viscovery SOMine lite Version 2.1 (1998). User's Manual Eudaptics software gmbh, Austria.
- Zaim, O. (2004). Measuring environmental performance of state manufacturing through changes in pollution intensities: a DEA framework. *Ecological Economics*, 48, 37-47.
- Zaim, O., & Taskin, F. (2000). A Kuznets curve in environmental efficiency: an application on OECD countries. *Environmental and Resource Economics*, 17, 21-36.
- Zhu, J. (2001). Multidimensional quality of life measure with an application to fortune's best cities. *Socio-Economic Planning Sciences*, 35, 263-284.
- Zofio, J.L., & Prieto, A.M. (2001), Environmental efficiency and regulatory standards: the case of CO2 emissions from OECD industries. *Resource and Energy Economics*, 23, 63-83.

TABLE 1: descriptive statistics: 2005-2011 mean values (*)

REG	LEIS	FAM	FRIEN	EMPL	INCOM	INEQ	FAM- UNEM	NO- WORK NO- STUDY	CULTU	ENVIR	WATE	SEDEN	EFFIC (DEA)
<i>NORTH</i>													
Piedmont	15.47	39.13	27.94	68.30	0.251	4.61	3.33	14.23	0.010	0.050	7.51	32.67	0.94
Aosta	16.61	34.70	24.89	71.26	0.289	4.23	3.00	12.66	0.010	0.056	6.80	32.56	0.99
Liguria	15.17	37.11	26.16	66.79	0.240	4.57	4.36	14.16	0.018	0.052	4.89	39.06	0.96
Lombardy	16.13	40.47	27.74	69.97	0.259	4.73	2.54	13.01	0.011	0.056	6.74	32.03	0.96
Trentino	23.03	45.93	33.96	72.61	0.298	3.93	2.19	10.06	0.018	0.065	2.84	16.17	1.00
Veneto	15.16	38.20	26.53	69.24	0.264	3.97	2.29	12.49	0.017	0.053	7.53	25.94	1.00
Friuli	17.81	43.10	30.81	68.01	0.266	3.99	2.87	12.61	0.011	0.056	3.91	29.03	0.98
Emilia	17.04	42.57	29.83	72.83	0.269	4.59	2.50	11.84	0.014	0.057	5.64	30.71	0.99
Marche	14.57	33.11	25.29	68.14	0.236	4.34	2.84	13.80	0.013	0.053	6.50	38.04	0.93
Tuscany	16.79	40.09	29.03	68.39	0.248	4.21	3.06	13.80	0.010	0.054	11.94	34.83	0.96
Umbria	16.10	37.87	27.94	67.29	0.219	4.36	3.17	13.84	0.011	0.058	10.37	40.36	0.92
Lazio	12.93	30.84	22.43	63.59	0.223	5.44	4.34	17.41	0.020	0.043	12.70	44.00	0.96
<i>SOUTH</i>													
Abruzzo	11.80	30.86	21.09	61.39	0.199	4.33	4.07	16.49	0.005	0.049	16.97	43.23	0.86
Molise	13.01	29.73	21.19	56.39	0.181	4.79	6.07	20.24	0.002	0.046	15.70	50.77	0.77
Campania	9.40	24.84	18.11	46.06	0.157	6.14	13.81	32.79	0.007	0.046	15.26	57.21	0.63
Apulia	10.24	25.83	18.86	49.26	0.148	5.06	8.80	28.70	0.004	0.046	15.50	53.86	0.67
Basilicata	12.10	29.24	21.96	53.17	0.170	4.91	7.67	24.91	0.002	0.050	11.87	49.66	0.73
Calabria	10.69	29.17	20.70	47.99	0.148	5.71	12.83	29.80	0.002	0.039	32.94	53.81	0.68
Sicily	10.66	31.11	20.13	47.77	0.154	6.31	13.13	33.24	0.007	0.035	29.69	60.83	0.69
Sardinia	13.90	32.94	24.91	55.53	0.177	4.79	7.04	24.97	0.007	0.045	17.41	43.34	0.78

(*) s.d. omitted for space reasons. They are available upon request

TABLE 2: Regional wellbeing scores: Panel DEA and Malmquist results^(*)

Region	2005	2006	2007	2008	2009	2010	2011	Period Average score	Rank	Malmquist Period Average Index ^(*)
<i>NORTH</i>										
Trentino	1	1	1	1	1	1	1	1	1	1
Veneto	1	1	1	1	1	1	1	1	1	1
Emilia	1	1	1	1	0,990	0,975	0,981	0,992	3	0,997
Aosta	0,982	1	1	0,985	0,988	1	0,969	0,989	4	0,998
Friuli	1	0,951	1	0,943	1	0,984	1	0,983	5	1,000
Liguria	0,902	1	0,929	0,916	1	1	1	0,964	6	1,017
Lazio	1	1	1	0,88	0,897	1	0,938	0,959	7	0,989
Lombardy	0,965	0,968	0,953	0,965	0,955	0,947	0,939	0,956	8	0,995
Tuscany	1	0,952	0,935	0,942	0,946	0,992	0,92	0,955	9	0,986
Piedmont	0,939	0,948	0,933	0,941	0,934	0,921	0,931	0,935	10	0,999
Marche	0,943	0,945	0,933	0,936	0,93	0,928	0,913	0,933	11	0,995
Umbria	0,911	0,924	0,932	0,94	0,933	0,915	0,906	0,923	12	0,999
<i>SOUTH</i>										
Abruzzo	0,858	0,858	0,836	0,854	1	0,814	0,831	0,864	13	0,995
Sardinia	0,772	0,767	0,763	0,765	0,747	0,797	0,832	0,778	14	1,013
Molise	0,777	0,776	0,787	0,787	0,772	0,753	0,744	0,771	15	0,993
Basilicata	0,751	0,758	0,728	0,727	0,737	0,704	0,704	0,730	16	0,989
Sicily	0,674	0,672	0,731	0,653	0,721	0,706	0,643	0,686	17	0,992
Calabria	0,681	0,683	0,663	0,656	0,644	0,704	0,737	0,681	18	1,013
Apuglia	0,669	0,679	0,683	0,682	0,666	0,658	0,661	0,671	19	0,998
Campania	0,67	0,661	0,646	0,624	0,615	0,596	0,586	0,628	20	0,978
Italy mean	0,875	0,877	0,873	0,86	0,874	0,87	0,862	0,870		0,997
North mean	0,970	0,974	0,968	0,954	0,964	0,972	0,958	0,966		0,998
South mean	0,732	0,732	0,730	0,719	0,738	0,717	0,717	0,726		0,996

^(*) Output oriented DEA (VRS), scores are computed on a three successive years base period. TE scores relative to t_1 in year 2005 and t_{+1} in year 2011 are not defined

^(*) Malmquist index averages are geometric means

TABLE 3: Clusters statistics: 2005-2011 group mean values ^(*)

LEISU	FAMI	FRIEN	EMPL	INCO	INEQU	FAM- UNEM	NO- WORK NO- STUD Y	CULT	ENVIR	WATE	SEDEN	EFFIC (DEA)
<i>CLUSTER 1: Trentino</i>												
23.03 (0.00)	45.93 (0.00)	33.96 (0.00)	72.61 (0.00)	0.25 (0.00)	3.93 (0.00)	2.19 (0.00)	10.06 (0.00)	0.07 (0.00)	0.35 (0.00)	2.84 (0.00)	16.17 (0.00)	1.00 (0.00)
<i>CLUSTER 2: Piedmont, Aosta, Liguria, Lombardy, Veneto, Friuli, Emilia, Umbria, Marche, Tuscany</i>												
16.09 (0.951)	38.64 (3.003)	27.62 (1.832)	69.02 (1.765)	0.23 (0.013)	4.36 (0.250)	3.00 (0.547)	13.24 (0.782)	0.05 (0.002)	0.15 (0.048)	7.18 (2.285)	33.52 (4.346)	0.96 (0.026)
<i>CLUSTER 3: Abruzzo</i>												
11.80 (0.00)	30.86 (0.00)	21.09 (0.00)	61.39 (0.00)	0.23 (0.00)	4.33 (0.00)	4.07 (0.00)	16.49 (0.00)	0.05 (0.00)	0.06 (0.00)	16.97 (0.00)	43.23 (0.00)	0.86 (0.00)
<i>CLUSTER 4: Lazio, Sardinia</i>												
13.41 (0.486)	31.89 (1.050)	23.67 (1.243)	59.56 (4.029)	0.20 (0.013)	5.11 (0.329)	5.69 (1.350)	21.19 (3.779)	0.04 (0.001)	0.07 (0.011)	15.06 (2.357)	43.67 (0.329)	0.87 (0.091)
<i>CLUSTER 5: Molise, Campania, Apulia, Basilicata</i>												
11.19 (1.437)	27.41 (2.111)	20.03 (1.589)	51.22 (3.905)	0.19 (0.018)	5.23 (0.539)	9.09 (2.895)	26.66 (4.634)	0.05 (0.002)	0.07 (0.009)	14.58 (1.573)	52.88 (2.940)	0.70 (0.055)
<i>CLUSTER 6: Calabria, Sicily</i>												
10.67 (0.014)	30.14 (0.971)	20.41 (0.286)	47.88 (0.107)	0.17 (0.008)	6.01 (0.300)	12.98 (0.150)	31.52 (1.721)	0.04 (0.002)	0.03 (0.002)	31.31 (1.629)	57.32 (3.507)	0.68 (0.002)
<i>NORTH</i>												
16,40 (2,346)	38,59 (4,124)	27,71 (2,895)	68,87 (2,471)	0,23 (0,019)	4,41 (0,403)	3,04 (0,673)	13,33 (1,673)	0,05 (0,005)	0,16 (0,075)	7,28 (2,907)	32,95 (7,046)	0,97 (0,026)
<i>SOUTH</i>												
11,24 (1,479)	28,73 (2,750)	20,56 (2,018)	51,51 (5,068)	0,19 (0,024)	5,35 (0,689)	9,69 (3,531)	27,10 (5,590)	0,04 (0,005)	0,06 (0,016)	18,96 (6,808)	52,21 (5,750)	0,72 (0,073)
<i>ITALY</i>												
14,43 (3,149)	34,84 (5,819)	24,97 (4,215)	62,20 (8,960)	0,21 (0,027)	4,75 (0,668)	5,50 (3,738)	18,55 (7,410)	0,05 (0,007)	0,12 (0,078)	12,14 (7,779)	40,41 (11,254)	0,87 (0,127)

(*) s.d. in parenthesis



Fig. 1: Italy, regions

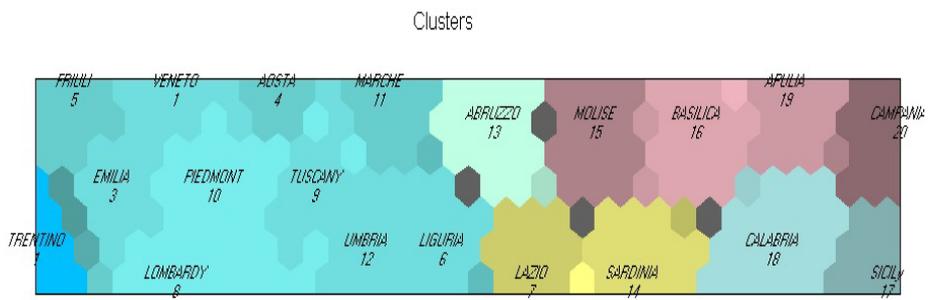
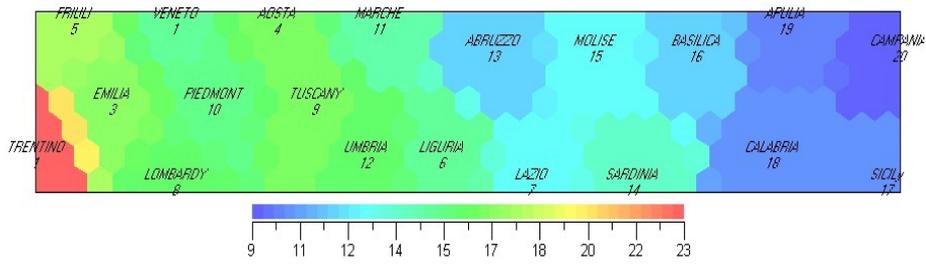
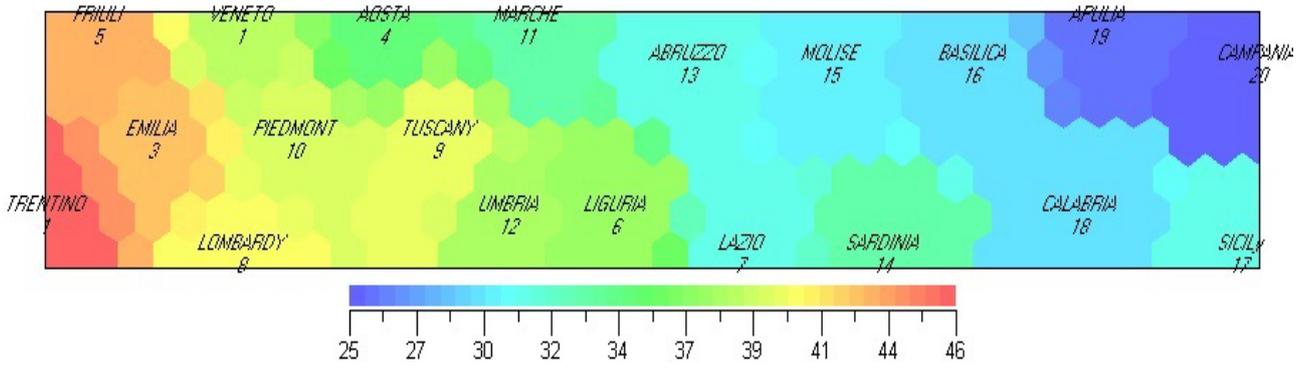


Fig. 2: Clusters regions (DEA efficiency rank)

[LEIS]



[FAM]



[FRIEN]

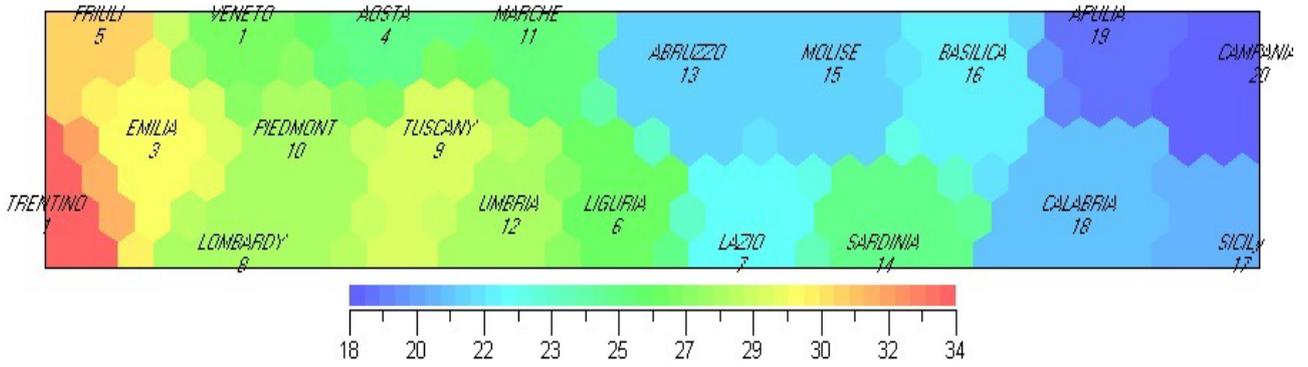
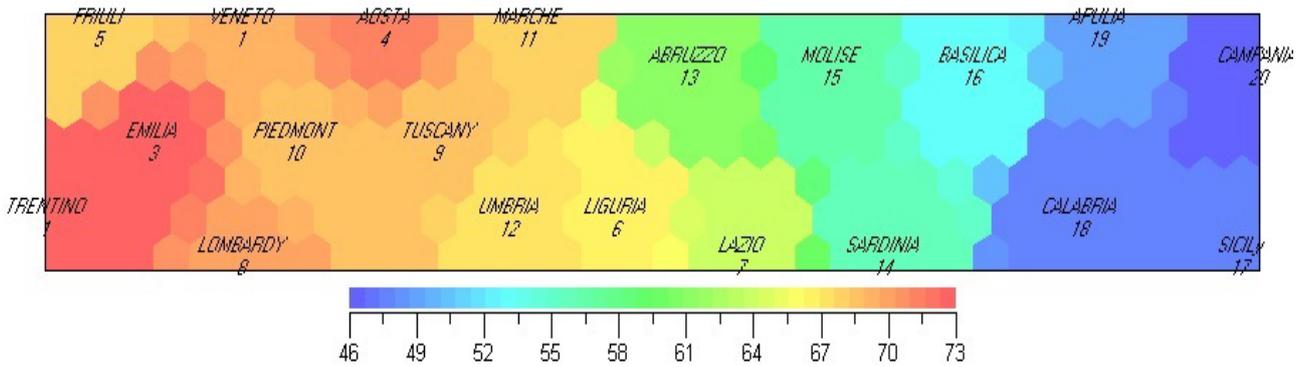
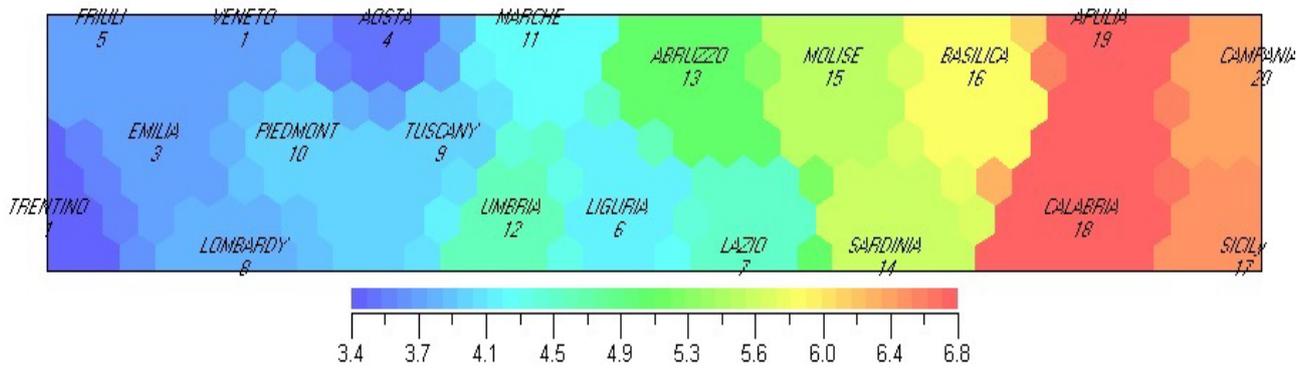


Fig. 3: single component window (DEA efficiency rank)

[EMPL]



[INV_INC]



[INEQ]

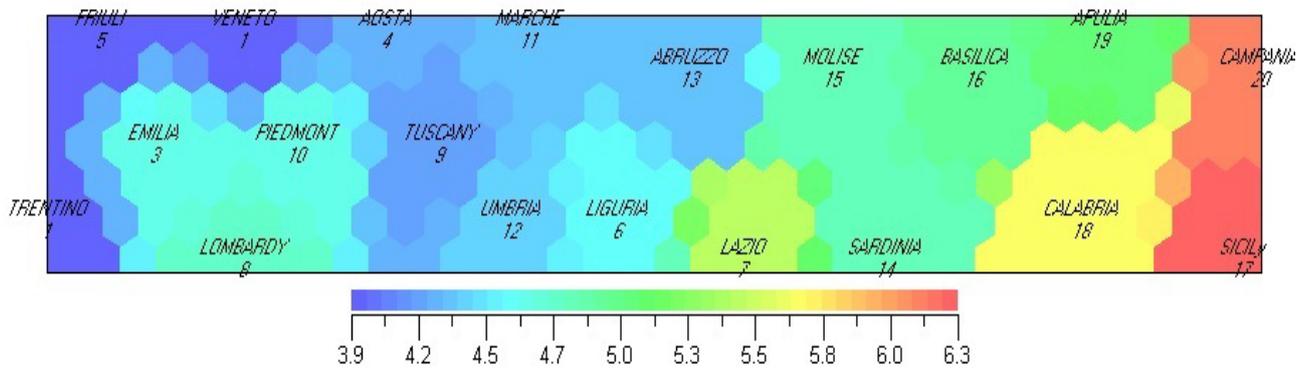
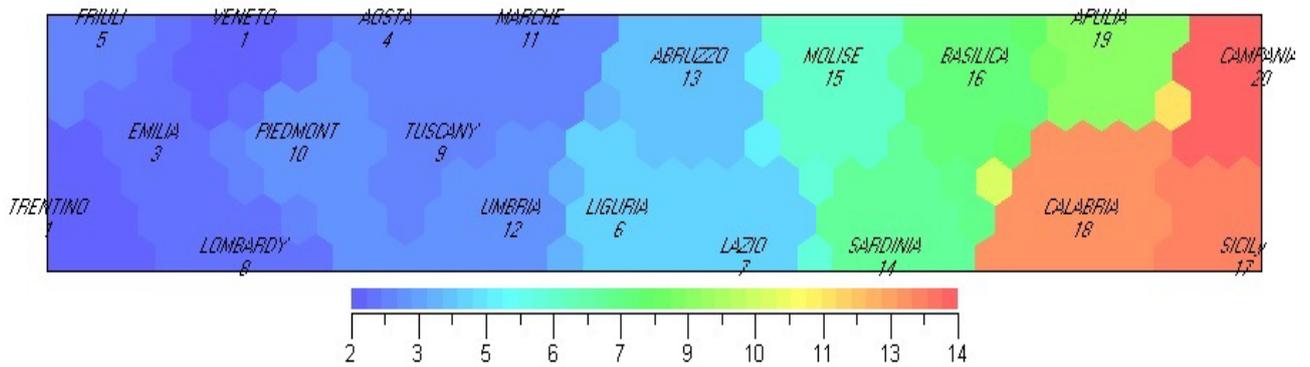
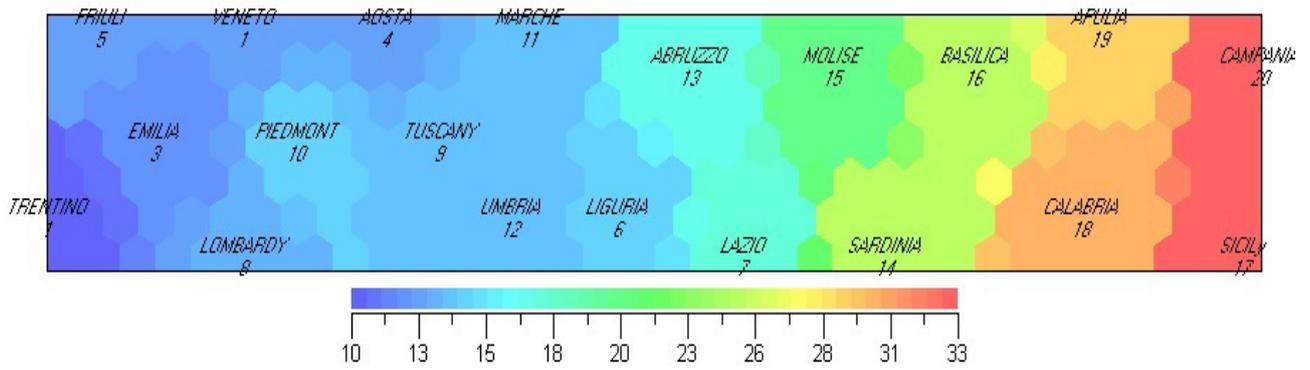


Fig. 3 (continued)

[FAM_UNEM]



[NO_WORK_STUDY]



[INVCULTU]

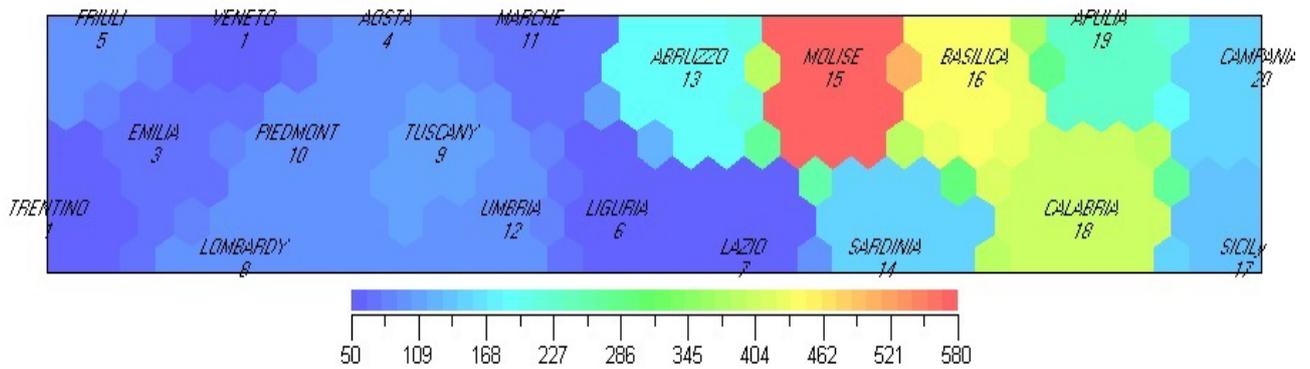
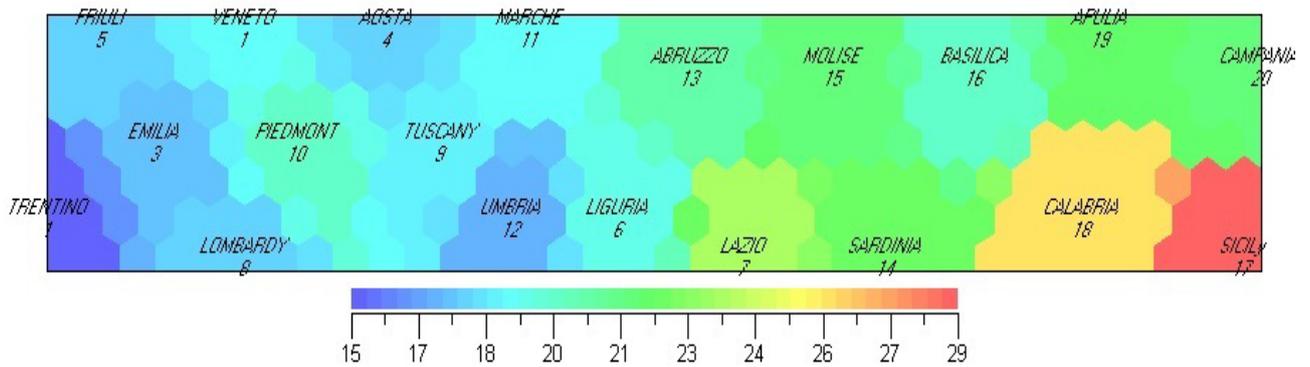
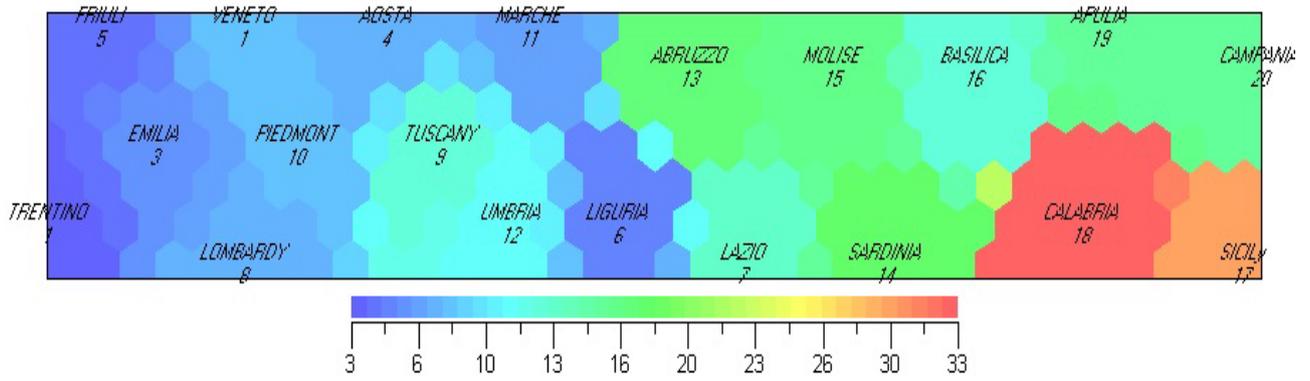


Fig. 3 (continued)

[INV_ENVIR]



[WATE]



[SEDEN]

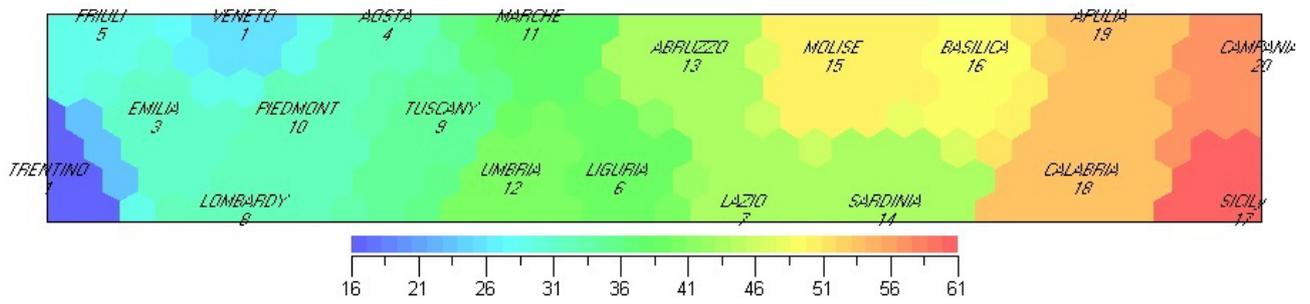


Fig 3 (continued)