SEEDS
Sustainability Environmental Economics and Dynamics Studies

Working Paper Series

Mapping energy-efficiency technological advances in home appliances.

by

Nicolò Barbieri and Alessandro Palma

2/2015

SEEDS is an interuniversity research centre. It develops research and higher education projects in the fields of ecological and environmental economics, with a special focus on the role of policy and innovation. Main fields of action are environmental policy, economics of innovation, energy economics and policy, economic evaluation by stated preference techniques, waste management and policy, climate change and development.

The SEEDS Working Paper Series are indexed in RePEc and Google Scholar.
Papers can be downloaded free of charge from the following websites:
http://www.sustainability-seeds.org/.

Enquiries:info@sustainability-seeds.org

SEEDS Working Paper 2/2015
January 2015
by Nicolò barbieri and Alessandro Palma.

Mapping energy-efficiency technological advances in home appliances.*

Nicolò Barbieri
Department of Economics
University of Bologna (Italy)
and SEEDS (Italy)
n.barbieri@unibo.it

Alessandro Palma†
Department of Economics
University of Roma Tre (Italy)
and SEEDS (Italy)
alessandro.palma@uniroma3.it

Abstract

The present study uses a patent dataset on four large energy-efficient appliances and provides a methodology for: i) mapping components related to energy efficiency improvements; ii) mapping their evolution over time; iii) testing the technological fungibility of these components. Our analysis model exploits an original patent selection process and the concept of technological relatedness using co-occurrence analysis of patent classes as input for Self-Organising Maps, an unsupervised artificial neural network able to represent high-dimensional data in visually attractive and low-dimensional distance-based maps. The results confirm the pervasive nature of energy efficiency to be nested in many technological components. In addition, we show that a dematerialisation process has affected the evolution of energy efficiency technologies over time, in a technological space characterised by a high level of complexity and variety.

Keywords: energy efficiency, Self-Organizing Maps, patent analysis, home appliances, ICTs.

J.E.L.: Q55, Q41, O33.

* We acknowledge financial support from the Italian Ministry of Education, University and Research (Scientific Research Program of National Relevance 2010 on "Climate change in the Mediterranean area: scenarios, economic impacts, mitigation policies and technological innovation"). We are also grateful to Davide Consoli, Valeria Costantini Francesco Crespi and Nicoletta Corrocher for their precious comments and suggestions. The usual disclaimer applies.

† Corresponding author, Department of Economics, Roma Tre University, Via Silvio D'Amico 77, 00144, Roma (Italy). E-mail: alessandro.palma@uniroma3.it

1. Introduction

The reduction of primary energy consumption through energy efficiency (EE) represents a cornerstone of the transition towards a resource-efficient green economy in Europe and an effective strategy to achieve energy independence (EC 2011). Nevertheless, energy saving and EE are not completely overlapping terms since EE is a sub-set of the energy saving (or energy conservation) domain. The latter is a broader concept since energy saving can be obtained through gains in EE or by simply reducing the level of economic activity. According to Patterson (1996), EE is the relationship between the produced output and the energy consumed to produce it, often called energy service. Thus, a general characteristic of EE is the use of less energy inputs for an equivalent level of economic activity or service (Gillingham *et al*. 2009) so that achieving higher EE performances intrinsically relies on the dynamics of technological innovation as a means for improving productivity of the energy input and reducing the associated energy consumption (Rennings and Rammer 2009; Florax *et al*, 2011).

While EE in the industrial sector has received much attention in economics, the residential sector has remained poorly explored even though the role of domestic consumption appears relevant and projected to be increasing over time (IEA 2012). At household level, the widespread presence of traditional large appliances (freezers, refrigerators, washing machines and dishwashers) is still responsible for 25% of households' electricity consumption as opposed to other appliances such as information and communication devices whose energy needs are negligible with respect to the so-called ''white goods" (Sadur *et al*. 2007)[1].

Even though traditional electrical appliances are characterised by mature technologies, their potential contribution to reducing aggregate energy consumption is still very large if we consider the combined effect of EE improvements in these technologies and the fact that, since traditional appliances are crucial to fulfilling primary needs, they are largely widespread among household dwellings (IEA 2009; 2010; EC-JRC 2012). Recent studies have confirmed the cost-effectiveness of EE gains deriving from electrical appliances with respect to those deriving from other sectors. In particular, cold appliances (freezers and refrigerators), washing machines and dishwashers seem to have had a large impact in terms of EE performances (McKinsey 2009).

At the same time, it is not difficult to notice that traditional electrical appliances have shown a growing level of technology integration, as more and more appliances incorporate information and communication technologies (ICTs) such as stand-by devices, digital displays, intelligent sensors, advanced firmware as well as many other improvements which allow for more sophisticated food conservation and washing processes while reducing the electricity required to operate. Moreover, the steadily technological advancement of electrical appliances is strongly linked to the concept of Smart House, in which a large use of ICTs is required to facilitate the interoperability of household products and services (Peine 2008). In a smart house, different products are mutually linked and controlled through a bus system, thus imposing the presence of ICTs in addition to traditional mechanics. In this respect, Antonelli states that 'no product or process can be manufactured without the substantial application of new information and communication technologies or without substantial effects of the application of new information and communication technologies' (Antonelli 2003, pp. 598; see also Antonelli 1992) and identifies fungibility as a further element characterising the ICTs. The concept of fungibility applies to technologies with applications that are useful to a great array of new products and processes and is strictly related to the concept of

[1] The portfolio of energy services available for households has massively increased over the last 20 years, with a strong penetration of new devices and appliances aimed at satisfying these services. See Burnwell (1990).

general-purpose technologies (GPTs) (Bresnahan and Trajtenberg 1995; Rosenberg and Trajtenberg 2004; Bresnahan 2010). In this regard, Panzar and Willig (1981) highlighted some important futures of GPTs, such as lower production costs due to the joint use of appliance components and gains in production efficiency through scope economies. According to Corrocher *et al*. (2007), ICTs are largely seen as one of the most important GPTs, able to spur the development of new technologies and applications spreading across different sectors.

The higher degree of technological variety as well as increasing level of ICT integration raises a series of questions and issues to be analysed which constitute the object of this paper. In particular, given the increasing technological complexity embodied in appliances and their strong contribution to decreasing household energy consumption, it is worth analysing the technological structure and evolution of EE in the most important electrical appliances. This implies a decomposition of the technological space in which EE has nested and evolved over time within a single appliance. In addition, we aim to analyse what role, if any, ICTs have played in transforming the technological space of energy-efficient domestic appliances.

Building on this debate, we break down the technological space of four groups of energy-efficient domestic appliances (freezers and refrigerators, dishwashers and washing machines) into different technological clusters that are able to affect EE performances, mapping and precisely identifying the position of each one, as well as to unveil their dynamics over time and across appliances. In doing so, we use patent maps obtained through the implementation of Self-Organizing Maps (SOMs), a class of unsupervised artificial neural network able to represent high-dimensional data in visually-attractive and low-dimensional distance-based maps.

Our analysis envisages three experiments. In the first one, we aim to identify niches of EE technologies by analysing the pattern of the technological space in each appliance. This allows us to easily identify those clusters in which EE technologies are nested and obtain a measure of technological variety. In the second experiment, we examine the mapping of EE clusters and explore the dynamics of clusters over time for patents belonging to cooling appliances (freezers and refrigerators) only. In the third and last experiment, we compare the technological space of two different appliances in search of technological fungibility, identified in patents with horizontal usefulness.

In order to achieve the objectives that we have just presented, we employ an original dataset of 688 unique triadic patent families belonging to the four energy-efficient appliances here analyzed and selected by developing an ad hoc methodology. In this respect, the role of patents has been largely exploited since they allow specific technology features to be analysed (Griliches 1990 and, more recently, Leydesdorff *et al*. 2014 among others). In particular, we consider the OECD Triadic Patent Family[2] database which assumes significant importance in recognizing high quality patents.

The rest of the paper is organised as follows. Section 2 describes the sample selection and the patent analysis model, along with the theoretical foundations of SOMs. Section 3 presents the experiment results and discusses them, and Section 4 concludes the paper.

2. Patent analysis model

There are a number of possibilities for measuring technological advances. As with most economic variables, the problem of measurement is directly related to the availability and the quality of specific data. In order to achieve the objectives introduced above, we employ patent data, which provide a public wealth of information on the nature of the invention and the applicant for

[2] While patent families are collection of patents filed at different patent offices and related to the same invention, triadic patent families are inventions protected in the main patent offices: USTPO, JPO and EPO.

rather long time series, indicating not only the countries where inventions are developed, but also where these new technologies are used and derive from. In addition, notably important for our paper, they provide the technological specification for the invention. However, the use of patents as a proxy of innovative activity was not exempt from criticism regarding, among others, heterogeneity in their technical and economic value (Griliches 1998; Hall et al. 2005) and differences in propensity of patenting across sectors and technologies (Arundel and Kabla 1998). A further limitation of patent data is represented by the partial and rough representation in the set of international patent classification when technological domains characterised by high complexity and rapid evolution are under scrutiny. Such an issue becomes crucial in eco-innovation studies since a growing number of green technologies are being developed and mapped by researchers and practitioners through *ad hoc* methodologies, as in the case of biofuels in Costantini *et al*. (2013).

The technological domain of residential EE constitutes a similar case, since only since 2013 the Cooperative Patent Classification includes classes related to EE technologies for domestic electrical appliances[3]. According to Costantini *et al.* (2014), the reason for such a lack is twofold. First, EE appears as a latent technological domain since the improvements in EE are not always explicitly mentioned by the main unstructured items of patent documents, namely title, abstract and claims. Consequently, the full document text, including long patent descriptions, must be analysed. The second reason lies with EE pervasiveness, the characteristic to be embodied in many components and devices. Indeed, an analysis of the relationship between EE and technological content shows that EE not only operates in the most advanced technologies but, comparatively, in the entire panorama of technologies using energy. In the field of patent analysis, this means that EE represents a complex cross-cutting technological space in which many CPC classes are involved in order to increase the efficiency of devices. Given the partial representation of the new CPC classification, a specific methodology for identifying EE technologies using patents is strongly required (Noailly and Batrakova 2010).

When patents are used as a means for investigation, the outcome usually takes the name of patent analysis. The latter relies on a broad set of methods and techniques aimed at identifying coherent information for different purposes. Basically, patent analysis techniques automatically reduce the large amount of information provided by patent documents to useful low-dimensional information. Lee *et al*. (2009) classifies patent data analysis as techniques based on structured and unstructured items. Structured patent items (SI) represent standardised text elements of patent documents such as IPC classes, priority year or citation count, which are analysed mostly by exploiting bibliometric techniques such as citation analysis (Yoon and Park 2004, among others). On the other hand, unstructured patent items (UI) consist of free text strings which are often very long as in the case of patent descriptions or claims and analysed using text-mining (TM) techniques as a means of knowledge extraction (Kim *et al*. 2008).

The visualization process (patent mapping) also constitutes an important part in explorative patent data analysis, especially when there is high information complexity and data dimensionality (Vesanto 1999). In this respect, the representation of patent analysis using maps is widespread in the literature of patent analysis and there are numerous techniques devoted to this aim. Broadly speaking, a patent map[4] is able to show complex and invisible relationships between different patent documents as well as their peculiar features by exploiting a simpler low-dimensional visualization. Abbas *et al*. (2014) classify maps as patent networks and cluster-based maps. In a

[3] The Cooperative Patent Classification (hereafter referred to as CPC) was established in 2010 as a joint partnership between the United States Trademark and Patent Office (USPTO) and the European Patent Office (EPO) to provide harmonization between the two classification systems developed by each office, European Classification and United States Patent Classification, respectively.

[4] Map is here used as a generic term, being synonymous with diagram, chart or graph.

patent network, the relationships between objects are investigated by analysing the relationships between ties and arches and exploiting the framework of the graph theory. Although network analysis was initially employed in sociological studies, such methodology now represents a widespread technique in innovation economics with a number of tools for visualizing and interpreting both SIs and UIs patent data (Narin 2000; Huang *et al*. 2004; Yoon 2004; Verspagen 2007; Sternitzke *et al*. 2008; Lee *et al*. 2009, among others). A patent map can also derive from a clustering process in which observations are divided into groups 'internally homogeneous (internal cohesion) and heterogeneous from group to group (external separation) [...] reducing the space dimensionality' (Giudici 2003, pp.76; see also Kim *et al*. 2008).

Patent maps may also differ in the outputs they produce. Van Eck and Waltman (2010), distinguish between graph-based maps and distance-based maps. While the focus of the former is on the presence of links between items, the latter captures the strength of these relationships and projects them in a spatially ordered space that captures the relatedness between objects. In this class of maps, the similarity between the input data is measured and represented in a low dimensional space where, usually, the lower the distance between items, the greater their similarity.

Despite the increasing number of studies that use citation data and connectivity analysis (Verspagen 2007, among others), we rely on distance-based maps to unveil the domains of EE technologies. The reason for our choice is twofold. First, the characteristics of EE technologies may permeate many technological domains decreasing the usefulness of using citation data to detect these domains. In addition, the use of citations would underestimate the technological content in the most recent patents, considering that these show lower probability to be cited than older patents (Jaffe and Trajtenberg 2004). In this respect, the choice of using Triadic patents instead of citations allow us to exploit a long time-series including recent patents while reducing the drawbacks of citations.

A promising approach to building patent maps relies on the use of artificial neural networks (ANN) which have shown a high level of efficiency in managing high-dimensional observed data, incomplete information, errors or inaccuracies. Although there are many types of ANNs[5], a first important distinction can be made between supervised and unsupervised ANNs. Differing from supervised[6] ANNs, which are more suitable for prediction analysis, a SOM is based on unsupervised learning processes able to map every dimensional observation in an output spatial grid. Such a feature makes the SOM particularly effective for classification and clustering analysis. Indeed, the nodes are placed in such a way that adjacent ones will be more similar than distant output nodes, thus introducing a topological dependence between clusters and preserving spatial correlation between the input vectors and the clusters.

The use of SOMs has been increasingly adopted in several applications and for mapping different types of data, such as, for instance, scientific journal networks (Campanario 1995), author co-citation data (White *et al.* 1998), conferences (da Silva Almendra *et al*. 2013) or industrial districts (Carlei and Nuccio 2014). In addition, several papers have employed patent data to visualise technological landscapes (Park *et al*. 2013; Polanco *et al*. 2001; Lee *et al*. 2009). For instance, Yoon *et al*. (2002) applied SOMs in order to show complex relationships and dynamic patterns in different technologies. They built technology vacuum, claim point and technology portfolio maps to identify technology-missing areas, potential infringements and technology classifications, respectively. A further recent contribution employing TM techniques as input for SOMs is provided

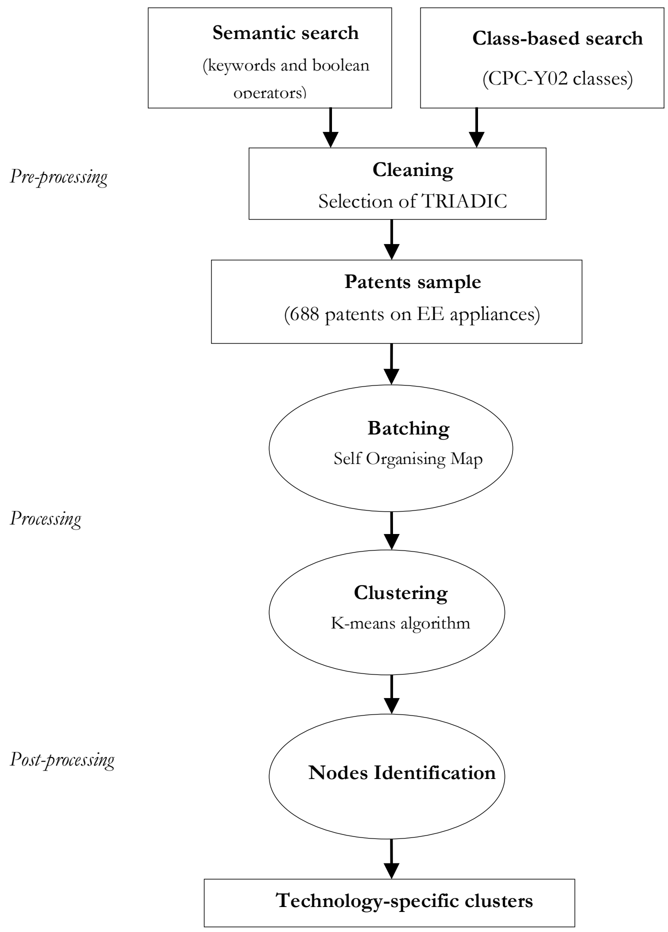
[5] For a comprehensive classification, see Giudici (2003).

[6] In a supervised ANN, training data need both input and output results while an unsupervised ANN only requires input data.

by Segev and Kantola (2012) who used a Term Frequency-Inverse Document Frequency (TF-IDF) algorithm to extract patent knowledge, then represented by SOMs. Moreover, they compared SOM performances with K-means (MacQueen 1967) and with Density-Based Spatial Clustering of Applications with Noise (DBSCAN, Ester *et al*. 1996) cluster classifications and found that SOM achieves better clustering performances, followed by K-means and DBSAN.

In light of these contributions, our Patent Analysis Model (Figure 2) is characterised by three phases. The first phase, i.e. pre-processing, is devoted to obtaining a set of patents belonging to energy-efficient electrical appliances while the last two, i.e. processing and post-processing, aim to identify the technological clusters in which EE is nested, as well as analysing their technological evolution over time.

Figure 2 – Patent Analysis Model



2.1 Sample selection

In order to collect a coherent set of patent documents using the 'Thomson Innovation database'. In doing this, we follow a combination of a top-down and bottom-up approach for catching patents filed at several patent offices. The top-down search method employs the recent CPC-Y02 classes, defined as 'Climate Change Mitigation' technologies. These include specific sub-classes referring to EE technologies in electrical appliances such as thermal insulation for freezers or heat recovery devices for dishwashers (see Table A1). Nonetheless, CPC fails to capture different technical expedients able to increase appliance efficiency, such as for instance, lighter motion devices deriving from the use of new materials or more efficient refrigerant liquid in cooling appliances. We filled this gap in the bottom-up search approach by employing two levels of searching with selected keywords together with Boolean operators (AND, OR, NOT) as in Costantini *et al*. (2014).

Considering the latent nature of EE technologies, two search levels were performed on the full patent UIs (title, abstract, claims and description). The first level defines the EE macro-domain with respect to the universe of patent applications in the considered period, while the second level reduces the macro-domain to an end-use level on those patents classifiable as inherent to domestically employed EE technologies using appliance-specific words such as 'refrigerator', 'washing machine' and so on[7]. Therefore, we reduced the set of applications to four energy-intensive electrical appliances: refrigerators and freezers (cooling appliances), washing machines and dishwashers (washing appliances). The patents obtained by the combination of a top-down and bottom-up approach are then used to collect the patent families[8]. Furthermore, in order to increase the quality of our dataset with higher value patents, we only rely on triadic patent families[9] (TPF) (Martinez 2010). Unfortunately, the use of TPFs further reduces our sample since the cost to file for patent protection in many countries is associated with higher application fees that are likely not to be incurred for low-value inventions. As a final step, we tested a sample of 15% as a further manual validation process, obtaining a unique dataset of 688 patents over the time span 1990-2014 and divided into three groups of domestic appliances as shown in Table 1.

Table 1- Patent sample, by appliance.

	Patents	Share
Dishwashers	66	9.60%
Refrigerators and Freezers	489	71.08%
Washing machines	133	19.33%
	688	100.00%

Source: own elaboration.

The advantages of using both CPC-search and string-search approaches are manifold. First, as shown in Table 2, the results obtained from a combination of the two methods are complementary since the use of only one would imply underestimation of EE technologies in the considered appliances. In particular, when the patent search is conducted using the top-down approach, a large portion of patents is missed. The latter is, on the other hand, captured by the bottom-up approach. Moreover, if we consider the number of CPC codes per patent as a proxy of technological diversification[10], the top-down approach shows a lower level of variety as measured by the standard deviation (SD) of CPC codes per patent, which is higher for those patents collected through a semantic search (keywords). This confirms the latent nature of the EE technological domain and provides evidence of its pervasiveness, given that the general objective of EE spreads across several classes as already highlighted in Noailly and Batrakova (2010) and Costantini *et al*. (2014).

[7] Search strings are provided in Table A2

[8] This procedure allows the double counting of patents that refer to the same technology whose protection had been extended to many patent offices to be dropped.

[9] While patent families are collections of patents filed at different patent offices and related to the same invention, triadic patent families are inventions protected in the main patent offices: USTPO, JPO and EPO.

[10] Breschi *et al*. (2003) have used the number of technological classes (that patents belong to) as a proxy for firm technological effort diversification. In our paper we adopt a similar approach using patents as the observation unit (instead of firms). We assume that the higher the number of CPC codes assigned to each patent, the greater their technological diversification.

Table 2 – Number of patents and their CPC code standard deviation (SD) for each sampling approach.

	Top-down approach	Bottom-up approach	Combined Approach
Number of patents	42	659	688
Full CPC code SD	4,490	9,748	9,567
8-digit CPC code SD	2,282	3,152	3,122
4-digit CPC code SD	0,975	1,187	1,164

A further validation method for capturing the pervasiveness of EE in different technological areas is based on a graphical analysis of the technological space of refrigerators obtained with 4-digit and 8-digit CPC co-occurrence maps[11] for the collected sample (Figure 3). Each node, which represents a single CPC class, is located according to the similarity level of other nodes, as the map captures the relative importance of the CPC codes with respect to the others. Hence, the more central the node, the greater its co-occurrence with other classes. As shown, the portion of technological space captured by the bottom-up approach (red line) is larger than that obtained using only a top-down approach (blue line), the latter only representing a subset of the former. However, despite the fact that the top-down approach fails to catch EE peripheral technological areas, it captures the most important CPC classes, located in the central part of the maps, since it is characterised by a high degree of co-occurrence.

2.2 SOM implementation

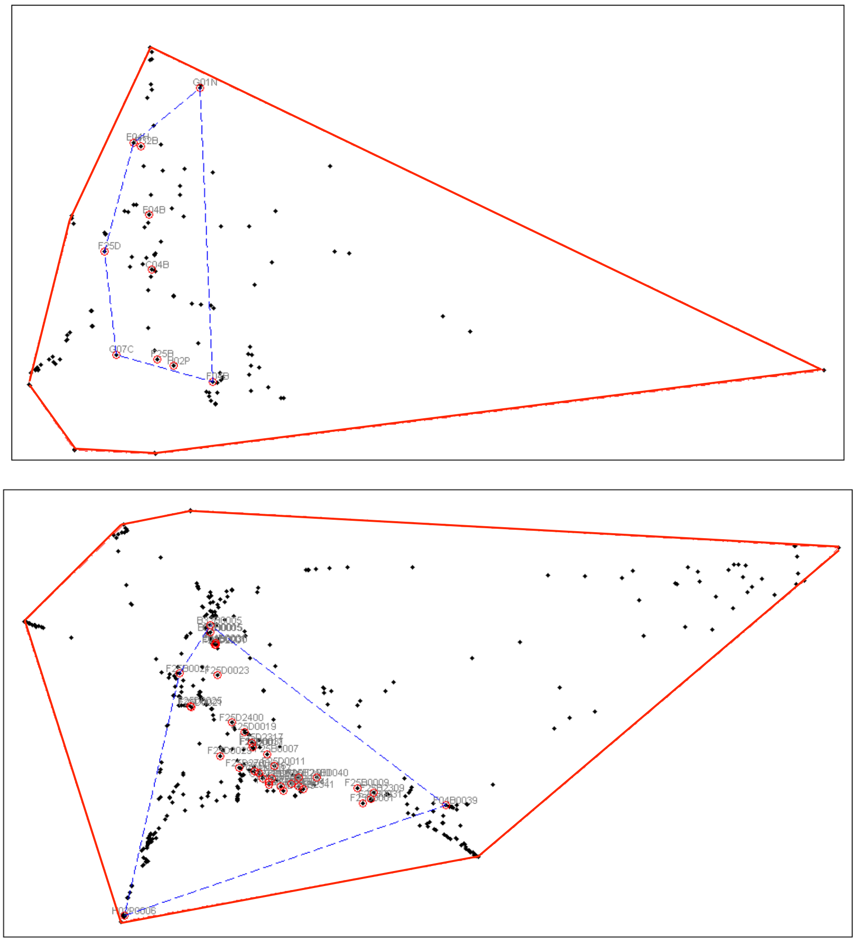
In order to create the PM, we apply the Self-Organising Map (SOM), a topological ordered mapping technique firstly introduced by Kohonen (1990; 1988). This constitutes a two-layer unsupervised competitive Artificial Neural Network (NN) able to represent multidimensional data on a two dimensional topological grid (Kohonen 2001). This technique reduces complex non-linear statistical relationships into a more simple, easy to understand and graphically attractive low-dimensional display in which the topological relationship between input data, that tend to be clustered, is preserved (Kohonen *et al*. 1996).

By resembling the Vector Quantisation (VQ) process[12], the SOM is a suitable tool for multidimensional reduction which provides a spatial and global order within the output map (Kohonen 2012) since similar input data are placed closer in the map, with different input data gradually further away (Kohonen 2012). Such a feature is provided by the adaptive units that form the map and are able to modify their response in such a way that the position of nodes in the map becomes representative of the patterns found in the inputs (Yoon *et al.* 2002). In addition, SOMs are able to learn from input data for a more effective representation and dimensionality reduction.

[11] The maps are built using VOS mapping technique (Van Eck and Waltman 2007) implemented in the freely available software used to map bibliometric items called VOSviewer (http://www.vosviewer.com).

[12] A standard methodological tool in modern digital signal processing in which n-dimensional input vectors are assigned to contiguous regions, each of them represented, in an optimal way, by codebook vectors

Figure 3 – Technological space obtained through 4-digit and 8-digit CPC code co-occurrence.



The maps refer to refrigerator patents. Red circles indicate the CPC classes retrieved using the top-down approach and black dots using the bottom-up approach. Red lines show the technological space captured by the bottom-up approach, while blue lines indicate the one obtained through a top-down approach. The figures have been produced for patents belonging to freezers and refrigerator only; maps for dishwashers and washing machines provide similar results and are available on request. Source: own elaboration.

To build the SOM we employed the SOM Toolbox[13] using the Batch Algorithm (BA) that provides higher accuracy as well as lower computational efforts (Kohonen 2012). The BA is calculated as follows:

(a) definition of map dimension[14] and assignment of the map node weight vectors (initialization phase);

(b) selection of a single input vector from the dataset;

(c) finding the most similar map node to each input data (the one that minimises the Euclidean distance between the input and map vectors) and tracking it as Best Matching Unit (BMU);

---

[13] A free Matlab© function package developed by the SOM Toolbox Team at the Helsinki University of Technology (Vesanto et al. 1999). The SOM Toolbox is downloadable under GNU General Public License at: http://www.cis.hut.fi/projects/somtoolbox/
SOM Toolbox is Copyright (C) 2000-2005 by Esa Alhoniemi, Johan Himberg, Juha Parhankangas and Juha Vesanto.

[14] We employed a heuristic formula proposed in the SOM Toolbox: $n = 5\sqrt{dlen}$, where n is the number of units that make up the final map and $dlen$ the number of observations that are mapped. As stated above, the shape of the lattice is defined by the two largest eigen vectors of the training data, during the initialization phase.

(d) when all input data are assigned to their BMU, updating the weight of each map neuron by computing the mean of the input data placed in the kernel defined by the neighbourhood function.

The BA differs from the classical sequential algorithm (SA) in how input data are presented to the grid of neurons since the whole set of input data is presented to the map at each epoch and only in a second step, the map weights are adjusted to reproduce the similarity between them. In this way, the order in which the input data are presented to the map does not influence the final output[15].

The SOM implementation provides a PM that returns the technological clusters in which inventive efforts in EE can be easily identified. The PM is developed using 8-digit CPCs assigned to each patent. These classes label patents according to their technological content through a hierarchical, language-independent classification system, thus characterising the technological cluster that the invention refers to. Therefore, the similarity between two patents, in terms of CPC classes, can be used as a proxy for the strength of their technological relatedness (Sherer 1982; Jaffe 1986; Verspagen 1997, among others). In the related literature, many efforts have been made to measure the technological relatedness between patents. For instance, Leydesdorff *et al.* (2014) built a matrix for measuring the number of times an IPC class is cited by other classes, using the cosine index as a measure of similarity. Breschi *et al.* (2003) and Nesta and Saviotti (2005), after describing each patent by a set of technological classes, used a matrix of co-occurrences to measure the strength between these technological fields. Accordingly, we use co-occurrences among CPC classes to define patent similarity. Therefore, the input data of the SOM is constituted by a matrix having, in each column, the frequency of 8-digit CPC classes assigned to each patent, with the univocal patent identification (Application Number) in the rows , as shown in Table 2:

Table 2 - SOM input matrix

	CPC 1	CPC 2	...	CPC m
Patent 1	...	...	...	...
Patent 2	...	...	...	...
...	...	...	...	...
Patent n	...	...	...	...

As a result, the SOM provides a PM where patents providing similar (different) technological improvements are placed closer (distant). Figure 5 illustrates the mechanisms, showing that Patent 1 and Patent 2 share only one CPC class (CPC 2), while Patent 3 and patent 4 show the entire set of CPC codes. Therefore, due to their technological similarity, Patent 3 and 4 are placed in the same position far away from Patent 1 and 2 that refer to different technical developments. Finally, Patent 1 and 2 are closely located, but not in the same position (since they share just one CPC class).

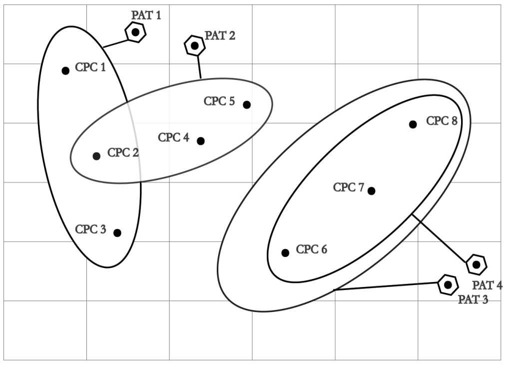
Once the SOM has been trained, a visualisation process is required in order to show the resulting output. More precisely, the process consists of a method for locating the BMUs in an effective and visually-attractive map (projection). Among the different techniques (Vesanto 1999), the Unified Distance Matrix (U-Matrix) is proposed here, choosing a grid with a hexagonal lattice[16] (Ultsch and Siemon 1990). By assigning different colour hues according to the distance between each map node, distance matrices show the similarity level between SOM nodes.

---

[15] For further details on SOM algorithms, see Appendix B.

[16] The choice of the lattice only reflects aesthetical reasons since it does not produce bias in the data representation.

Figure 5- Example of patent map created through the SOM.



Source: own elaboration.

Based on the spatial order of input data, obtained using the SOM, map nodes are clustered through the non-hierarchical K-means algorithm (MacQueen 1967). By applying the K-means method, the nodes are partitioned into k groups and clustered according to the method of centroids, i.e. points with a low distance between them and the other elements in the cluster. The number of clusters is defined by choosing k based on the sum of squared errors. Finally, the Davis-Bouldin (DB) index (Davis and Bouldin 1979) is calculated for each k, a clustering performance index that measures compactness and separation between nodes and clusters. The synergy that arises from the use of this two-stage method, i.e. SOM and clustering, is found to produce more powerful results than using them singularly (Chi and Yang 2008; Kuo *et al*. 2002).

3. Results and discussion

3.1 Experiment I – Identification of EE niches.

In this first experiment, we use the entire patent sample to discover which technical components are impacted by EE improvements. These components are presented on the maps as clusters that include nodes with similar technological content, previously identified by the SOM. We repeat this experiment for each of the three appliances, providing a clear picture of the pervasive nesting of EE in different appliance components such as mechanical, electro-mechanical, digital and chemical ones as well as in operational processes. The resulting framework is thus characterised by high technological complexity, generated by numerous clusters referring to a wide array of scientific and industrial contributions.

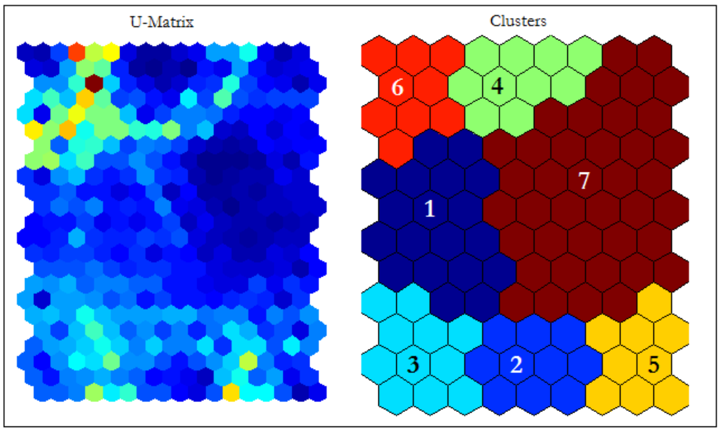
In the case of refrigerators and freezers, we identified seven clusters (Figure 6). A first, generic cluster (#1), includes patents on new refrigerators and freezers as a whole; excluding this, the others constitute a set of specific technology clusters which form the technological space of the appliance under scrutiny. As shown in Table 3, clusters 7, 3, 5, 4, 2 and 6, belong to various technological fields, ranging from electrical components (i.e. energy management systems, which include sensors, microprocessors, displays) to chemical components (i.e. refrigerant compositions, insulating foams and lubricant oils).

The pervasiveness of EE appears, as in the previous case, also when we repeat the experiment in the set of washing appliances, i.e. washing machines and dishwashers, whose results are represented in Fig. 7 and 8 respectively. It is worth noting that, although the number of patents belonging to these two appliances is lower, EE still affects a variety of different clusters. In particular, we identified 5 clusters (133 patents in total) for washing machines and 4 clusters (66 patents in total) for dishwashers shown in Table 4 and 5, respectively. The identification of these clusters confirms the hypothesis on the presence of the resource pooling effect also in these two electrical appliances. Indeed, the niches aimed at improving the level of EE pool a variety of different industries that produce technologies deriving from the application of several scientific fields. Moreover, since each cluster includes a variable set of patents, we can also derive a measure of innovation effort in each specific technological cluster. In other words, it is possible to identify where most of the efforts for EE gains have been addressed within each single appliance over the entire period of analysis, specifying that such a rank only assumes a qualitative nature since an assessment of the technological patent value does not constitute the main objective of this work.

Table 3 - Cluster identification for refrigerators and freezers.

Refrigerators and Freezers				
Cluster #	Technology description	% patent share per cluster	Within distance (nodes)	Between distance (clusters)
#7	Mechanical and electrical components (compressors, pumps etc.)	43.15 (1)	0.32(1)	0.53 (1)
#3	Refrigerant circulation systems	12.27 (2)	0.51 (6)	0.82 (6)
#5	Components for power and control management	12.07 (3)	0.42 (3)	0.77 (5)
#1	New refrigerators and freezers (as a whole)	11.45 (4)	0.43 (4)	0.55 (2)
#4	Heat transfer and refrigerant compositions	8.18 (5)	0.35 (2)	0.71 (4)
#2	Insulation panels and foams	6.54 (6)	0.44 (5)	0.647 (3)
#6	Lubricant oils	6.34 (7)	0.74 (7)	1.26 (7)
	Total	100	Mean 0.46	Mean 0.76

Figure 6 - U-matrix and K-Means clustering for refrigerators and freezers.

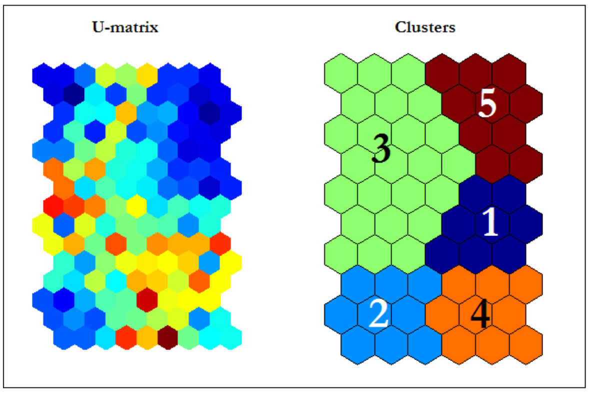


Source: own elaboration on MatLab.

Table 4- Cluster identification for washing machines.

Washing Machines				
Cluster #	Technology description	% patent share per cluster	Average node distance within the cluster	Average distance between CLs
#3	Mechanical and electromechanical components.	47.37 (1)	0.38 (2)	0.67 (2)
#5	Washing process/methods and washing machine as a whole	21.80 (2)	0.32 (1)	0.86(4)
#4	Digital components for energy management	12.03 (3)	0.45 (3)	0.95(5)
#2	Sensors	11.28 (4)	0.52 (5)	0.85(3)
#1	Motion and heating electrical controllers.	7.52 (5)	0.47 (4)	0.61 (1)
	Total	100	Mean 0.43	Mean 0.79

Figure 7 - U-matrix and K-Means clustering for washing machines.

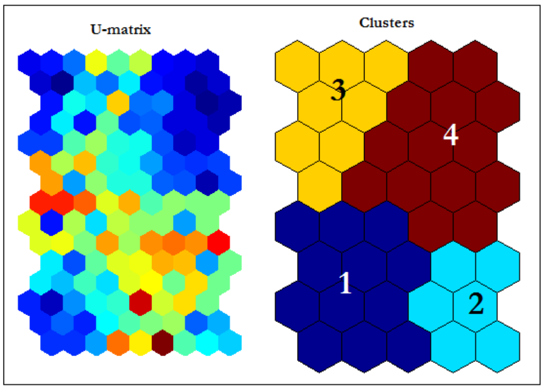


Source: own elaboration on MatLab.

Table 5- Cluster identification for dishwashers.

Dishwashers				
Cluster #	Technology description	% patent share per cluster	Average node distance within the cluster	Average distance between CLs
#4	Other components (mechanical, electromechanical and chemical)	40.91 (1)	0.35(2)	0.59 (1)
#3	New washing methods and dishwashers as a whole	24.24 (2)	0.35 (1)	0.76(3)
#1	Components for energy management	22.73 (3)	0.53(4)	0.71(2)
#2	Controllers and sensors	12.12 (4)	0.44(3)	0.79(4)
	Total	100	Mean 0.42	Mean 0.71

Figure 8 - U-matrix and K-Means clustering for dishwashers.



Source: own elaboration on MatLab.

Further investigation is based on a spatial analysis of clusters. By exploiting the topology-preserving feature of SOMs and their geometrical properties (preservation of the initial degree of CPC relatedness projected in an Euclidean space) we calculated two additional measures within the cluster maps: the average distance between each cluster centroid (*between* distance) and the average distance between the nodes included in each cluster (*within* distance)[17]. The *between* distance is here used as a proxy for measuring the degree of spatial agglomeration not captured by

[17] Whereas the within distance is calculated using the average Euclidean distance between each node within each cluster, the 'between' distance is the average Euclidean distance between cluster centroids. Here, the centroid is measured using the geometric centre of the cluster nodes.

the cluster map. Accordingly, lower average distances imply a denser technological space since the cluster centroid agglomerates closer to others. On the contrary, higher averages *between* distances indicate that the nodes remain sparse in the clustering space and, within each cluster, do not tend to be 'attracted' by other clusters. In the case of refrigerators and freezers, the core cluster is graphically represented by #7, which includes mechanical and electromechanical components which include the highest number of patents (43 % of the sample). Such centrality is confirmed by comparing the *between* distances, which is the lowest among the clusters (0.53). The level of centrality approximates the degree of technology integration and combination. Thus, peripheral technologies (refrigerants, sensors, insulations panels) seem to be complementary and to serve the core technology (compressors and pumps).

On the other hand, the *within* distance averages the distance between the different SOM nodes, representing a measure of spatial density within each cluster. Since we used the number of CPC classes per patent as a proxy of technology variety, when we evaluate the distance between the nodes (which in turn includes different patents) the spatial density is to be interpreted as the technology specificity of a given cluster. Namely, lower *within* distances indicate more dense technological clusters characterised by a high technological specificity, as in cluster #7 (mechanical and electromechanical components) as well as in cluster #4 (chemical compositions for heat transfer and refrigerant liquids).

Table 3 - Cluster identification for refrigerators and freezers.

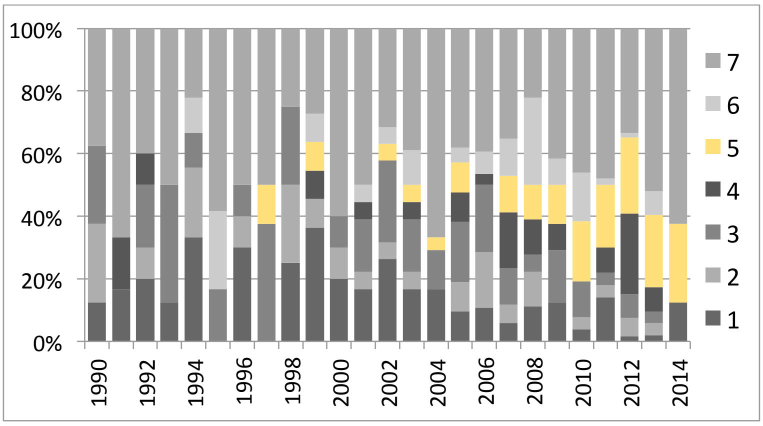
Refrigerators and Freezers				
Cluster #	Technology description	% patent share per cluster	Within distance (nodes)	Between distance (clusters)
#7	Mechanical and electrical components (compressors, pumps etc.)	43.15 (1)	0.32(1)	0.53 (1)
#3	Refrigerants circulation systems	12.27 (2)	0.51 (6)	0.82 (6)
#5	Components for power and control management	12.07 (3)	0.42 (3)	0.77 (5)
#1	New refrigerators and freezers (as a whole)	11.45 (4)	0.43 (4)	0.55 (2)
#4	Heat transfer and refrigerant compositions	8.18 (5)	0.35 (2)	0.71 (4)
#2	Insulation panels and foams	6.54 (6)	0.44 (5)	0.647 (3)
#6	Lubricant oils	6.34 (7)	0.74 (7)	1.26 (7)
	Total	100	Mean 0.46	Mean 0.76

Experiment II – Comparison over time

In the second experiment, we only take into account patents for cooling appliances (refrigerators and freezers) and divide the sample into two sub-samples by including the first and last 100 patents sorted by publication date. It is worth noting that the distribution of patenting activity across years is not homogenous since the first 100 patents are spread across eleven years (1990-2000), while the second 100 refer only to a four-year period (2012-2014).

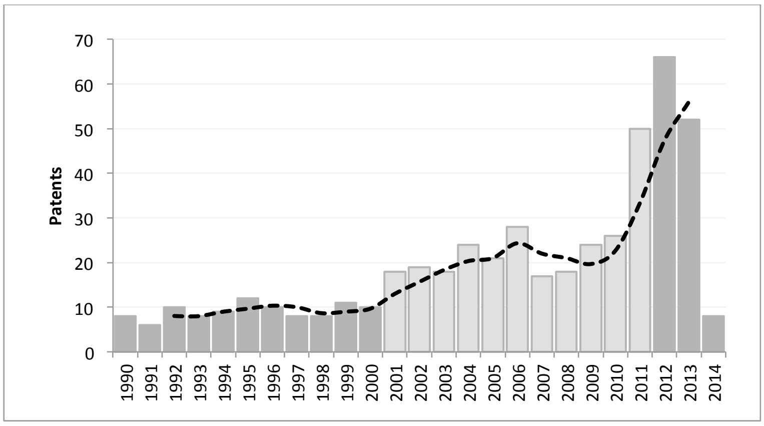
As shown in Figure 8, we moved to a more complex framework of technology variety, which provides first evidence of growing technological complexity due to the different content of clusters that increases over time, expressed as percentages. By repeating the Experiment I for the two sub-samples, we produce two temporal sections of the domestic cooling appliances aimed at comparing technological advances in EE components. Figure 9 shows the histogram of the total distribution of fridge patents in the years covered by the two sub-samples (in dark grey). The number of new patents belonging to energy-efficient refrigerators and freezers is clearly skewed toward more recent years, signalling increasing innovative efforts made by manufacturers to provide their appliances with more and more EE technologies. When we look at the patent maps (Figure 10), we note an equal number of clusters, but when the latter are under scrutiny, they unveil different technological content.

Figure 8 – Emergence of ICT clusters (#5) for power and control management. Percentage of patents per cluster (1990-2014).[18]



Source: own elaboration.

Figure 9 - Histogram of patent distribution for fridges (1990-2014) and three-year moving average trend.



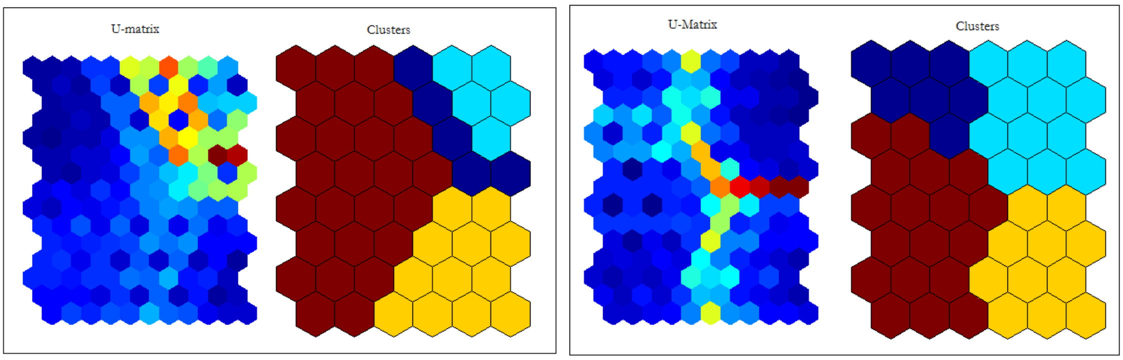
Source: own elaboration.

[18] Cluster numbers refer to Fig. 6.

Table 6 - Cluster identification for static temporal comparison.

Refrigerators and Freezers									
1990-1999					2010-2014				
Cluster	Technology description	% patent share per cluster	Average node distance within the cluster	Average distance between CLs	Cluster	Technology description	% patent share per cluster	Average node distance within the cluster	Average distance between CLs
#4	Mechanical and electromechanical components	67 (1)	0.40 (2)	0.98 (2)	#4	Insulation materials, mechanical components and new appliances.	56 (1)	0.44 (3)	0.76 (1)
#3	New appliances (as a whole)	26 (2)	0.34 (1)	1.05 (3)	#2	Components for power supply management	19 (2)	0.21 (1)	1.08 (3)
#2	Lubricant oils	4 (3)	0.43 (3)	1.47 (4)	#3	Refrigerant compositions	17 (3)	0.41 (2)	1.11 (4)
#1	Refrigerant compositions	3 (4)	0.57 (4)	0.84 (1)	#1	Control devices	8 (4)	0.46 (4)	0.76 (2)
	Total	100	Mean 0.43	Mean 1.08		Total	100	Mean 0.38	Mean 0.92

Figure 10 - Comparison of EE technology clusters
in two different periods (1990-1999 and 2010-2014).



Source: own elaboration on MatLab.

Indeed, in the first period, the technological space is mainly characterised by mechanical and electromechanical components and more efficient appliances as a whole, that together constitute almost 93% of total patents in the period analysed. On the contrary, in more recent years, the massive presence of digital components for energy management and motion control can be observed together with an increasing share of patents related to new compositions for refrigerants. The presence of this set of new EE components leads to a more complex technological space, thus enhancing the level of technological recombination[19] and accruing the resource pooling effect.

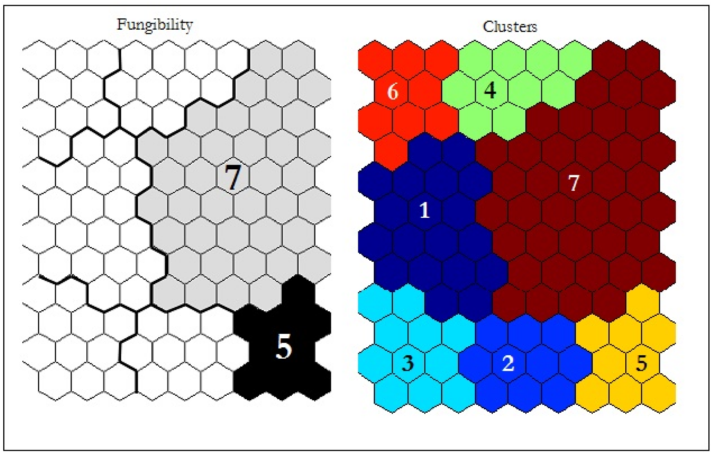
[19] According to Antonelli (1999), the technological recombination can operate both vertically (diachronic recombination) and horizontally (synchronic recombination). The first refers to a recombination of past elements of knowledge, while the second exploits contemporary acquisition of new pieces of knowledge (Antonelli, 1999).

The result of this part of analysis provides a clear picture of technological evolution in which the technological space for domestic EE evolved, showing a dematerialisation process from mechanical to digital components and most likely improving the level of EE jointly operating by relying on different technological contributions. In addition, as far as the 'within' and 'between' distance means are concerned, we can observe that a decrease in both the measures provides an insight into growing density in the technological space across time, although the number of clusters remains the same.

Experiment III – Technological fungibility

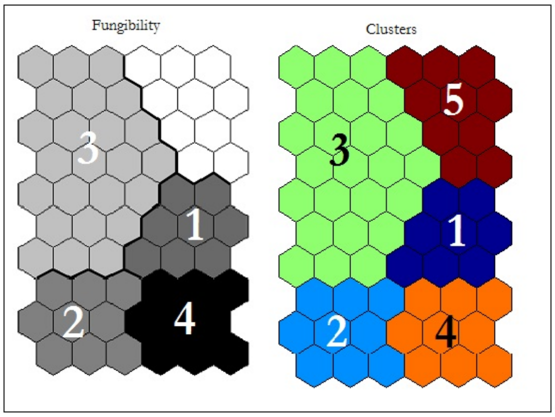
The third experiment investigates the presence of technological fungibility between two different groups of appliances by analysing the joint-use relationships of their EE technological components. We used co-occurrence analysis of CPC classes to mark those patents as *horizontal*, that is patents classified as employable both in: (i) refrigerators and washing machines; (ii) washing machines and dish washers. Then, we exploited the results of Experiment I for keeping information, in the SOM outputs, on each patent within nodes and clusters. It was thus possible to detect not only multi-appliance patents (that is patents employed in different appliances), but also the technological clusters that those patents refer to. As a further step, the percentage of multi-appliance patents was calculated for each cluster. Lastly, we used the SOM outputs to produce new K-means clustering maps, choosing this time a black-white visualization scheme in which the percentage of black is proportional to the percentage of multi-appliance patents in each cluster. The results in Figure 10 and 11 clearly show where the technological fungibility of EE of appliance components is nested. Specifically, both in the case of refrigerator vs. washing machines as well as in the case of washing machines vs. dishwashers, the cluster containing digital components for energy management and motion control (#5 and #4, for refrigerators and washing machines respectively) has been identified as the most pervasive and containing the highest level of fungible components.

Figure 10 - Technological fungibility of EE components between refrigerators and washing machines.



Source: own elaboration on MatLab.

Figure 11 - Technological fungibility of EE components between washing machines and dishwashers.



Source: own elaboration on MatLab.

In our case, power management and control systems include digital and communication devices as well as firmware and microprocessors which can be employed in a wide range of applications (general applicability), and characterised by continuous improvements (technological dynamism) considering the enormous growth of ICTs which has occurred recently (Cecere *et al*. 2014). In these dynamics, EE accommodates well and strongly exploits the interchangeable technological space of domestic electrical appliances.

4. Conclusions

The present work uses an original patent dataset belonging to four energy-efficient domestic appliances to develop, using SOMs, an analysis model in order to test a series of theoretical hypotheses. These hypotheses refer to: i) the pervasiveness of EE in different technological components; ii) the presence of resource pooling effect as a result of growing technological variety and development; iii) the fungibility of EE technological components in different appliances. In order to test the previous hypotheses, a set of three experiments was implemented and patents with EE implications in four large electrical appliances (refrigerators and freezers, washing machines and dishwashers) were analysed. The choice of this technological domain exploits the growing theoretical and empirical literature contributions on eco-innovation which are producing more statistical and economic analysis in complex and rapidly growing technological domains using patents.

The concept of technological relatedness is used to build a vector of CPC class co-occurrences for each patent and then use this vector as input for further analysis. In particular, with SOMs, highly multidimensional data deriving from the association of different CPC classes were reduced to bi-dimensional maps in which the patterns of different technological clusters clearly emerge. As a further identification process, a K-means clustering method were also applied to SOM outputs, producing clearer maps of EE technological clusters.

In the first experiment we find the technology clusters in which EE is nested. This experiment is thus repeated for each of the four appliances, providing a clear picture of the pervasive nesting of EE in different appliance components such as mechanical, electro-mechanical, digital and chemical ones as well as in operational processes. Our results confirm that EE is affected by pervasiveness and tends to be nested in many technological fields, thus admitting the hypothesis i). By also comparing the number of nodes belonging to each cluster, a measure of innovation effort in each particular technological niche has been also derived, making it possible to identify where most of the efforts for EE gains are addressed in each single appliance. Considering the entire sample of patents, dating from 1990 to 2014, such innovative efforts seem to be concentrated in mechanical and electrical components, without exceptions among the three groups of appliances.

In the second experiment, we compared the technological space between two different periods. In order to maximise the sample size for analysis, we only take into account patents for refrigerators and freezers and divide the sample into two sub-samples. These include the first and last one hundred patents, sorted by priority date. By moving from the first to the second patenting period, identified by the two sub-samples, we found higher levels of complexity along with a strong presence of the resource pooling effect due to the growing technological variety. The evolution of EE technology niches shows a dematerialisation process, initially characterised by the almost complete presence of mechanical components and, over time, moving toward an increasing complexity dominated by a massive presence of digital components of a different nature.

Considering such a high level of complexity in the technological space of EE electrical appliances, the third experiment is devoted to investigating the hypothesis of technological fungibility and finding that technology clusters containing patents with horizontal usefulness were particularly evident between two different couples of appliances, namely between refrigerators-freezers and washing machines, and between washing machines and dishwashers. As a result, in both cases we identified a single cluster that includes patents for energy management and digital motion control. This technological cluster, characterised by the highest level of fungible components and rapid growth, can be referred to the group of ICTs, defined by many authors as a GPT, that is, a technology showing general applicability and technological dynamism thus able to generate lower reproducibility costs for manufacturers. In light of this, we conclude that this technological cluster, including the set of components referring to power management and digital motion controllers appears not only particularly able to embrace the aim of EE, but it constitutes an interesting case of technological fungibility when domestic electrical appliances are under scrutiny.

References

Abbas A., Zhang L., Khan S. U. (2014). A literature review on the state-of-the-art in patent analysis. World Patent Information, available online (in press).

Antonelli C. (1992) The Economics of Information Networks (Ed). Elsevier, Amsterdam.

Antonelli C. (1999). The Microdynamics of Technological Change. Routledge, London.

Antonelli C. (2003). Knowledge Complementarity and Fungability: Implications for Regional Strategy. Regional Studies, 37(6-7), pp. 595-606.

Antonelli C. (2008). Localised Technological Change: Towards the Economics of Complexity. Routledge Studies in Global Competition. Taylor & Francis.

Archibugi D., Pianta M. (1996). Measuring technological change through patents and innovation surveys. Technovation, 16 (9), pp. 451-468.

Arthur B. W. (1989). Competing Technologies, Increasing Returns, and Lock-In by Historical Events. The Economic Journal, vol. 99 (394), pp. 116-131.

Arundel A., Kanerva M., Kemp R. (2011). Integrated innovation policy for an integrated problem: Addressing climate change, resource scarcity and demographic change to 2030. Technical report, European Commission, DG Enterprise and Industry.

Arundel A., Kabla W. E. (1998). What percentage of innovations are patented? Empirical estimates for European firms. Research Policy, 27 (2), pp. 127-141.

Berkhout F. (2011). Eco-innovation: reflections on an evolving research agenda. International Journal of Technology, Policy and Management 11 (3-4).

Blanchard A. (2007). Understanding and customizing stopword lists for enhanced patent mapping. World Patent Information vol. 29, pp. 308-316.

Borghesi S., Costantini V., Crespi F., Mazzanti M. (2013). Environmental innovation and socio-economic dynamics in institutional and policy contexts. Journal of Evolutionary Economics 23 (2), pp. 241-245.

Breschi S., Lissoni F., Malerba F. (2003). Knowledge-relatedness in firm technological diversification. Research Policy, 32(1), pp. 69-87.

Bresnahan T. F. (2010). General purpose technologies, in: Hall B., Rosenberg N. (eds.) Handbook of the Economics of Innovation 2, North Holland, 2010, pp. 761–791.

Bresnahan T. F., Traitenberg M. (1995). General purpose technologies: 'engines of growth'? Journal of Econometrics (65), pp. 83–108.

Campanario J. M. (1995). Using neural networks to study networks of scientific journals. Scientometrics, Vol. 33 (1), pp. 23-40.

Carlei V., Nuccio M. (2014). Mapping industrial patterns in spatial agglomeration: A SOM approach to Italian industrial districts. Pattern Recognition Letters, vol. 40(1), pp. 1-10.

Chi S. C., Yang C. C. (2008). A Two-stage Clustering Method Combining Ant Colony SOM and K-means. Journal of Information Science & Engineering, 24(5), pp. 1445-1460.

Cohen W. M., Nelson R. R., Walsh J. P. (2000). Protecting Their Intellectual Assets: Appropriability Conditions and why U.S. Manufacturing Firms Patent. NBER Working Paper No. 7552.

Costantini V., Crespi, F., Palma, A. (2014), Mapping innovation systems through patent analysis. The case of technologies for energy efficiency in the residential sector, in Patrucco P. (Ed.) The Economics of Knowledge Generation and Distribution: The Role of Interactions in the System Dynamics of Innovation and Growth, Routledge.

Costantini V., Crespi F, Curci Y. (2013). BioPat: An Investigation Tool for Analysis of Industry Evolution, Technological Paths and Policy Impact in the Biofuels Sector, in Costantini V.,

Mazzanti, M. (Eds.) The Dynamics of Environmental and Economic Systems. Innovation, Environmental Policy and Competitiveness. Springer Netherlands.
Davies D. L., Bouldin D. W. (1979). A cluster separation measure, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 1, 1979, pp. 224-227.
EC (2011a). Communication from the Commission to the European parliament. Energy Efficiency Plan 2011.Com. (2011) 109.Technical report, European Commission.
EC (2011b). Energy Efficiency Plan 2011.Commission Staff Working Document.Com/2011/0109.
EC (2012). Energy Efficiency Status Report 2012. Electricity Consumption and Efficiency Trends in the EU-27. Joint Research Center Scientific and Policy Reports.
Ester M., Kriegel H. P., Sander J., Xu X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the second international conference on knowledge discovery and data mining (KDD-96)(pp. 226–231)
Gillingham K., Newell R., Palmer K. (2009).Energy efficiency economics and policy. Annual Review of Resource Economics, 2(1), pp. 597-620.
Giudici P. (2003). Applied Data Mining. Statistical Methods for Business and Industry. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, England.
Griliches Z. (1998). Patent Statistics as Economic Indicators: A Survey. In R&D and Productivity: The Econometric Evidence. Ed. Zvi Griliches. National Bureau of Economic Research. University of Chicago Press.
Hall B. H., Jaffe A., Trajtenberg M. (2005). Market Value and Patent Citations. Rand Journal of Economics, vol. 36 (1), pp. 16-38.
Hascic W. E., de Vries F., Johnstone N. (2009). Effects of environmental policy on the type of innovation: The case of automotive emission-control technologies. OECD Journal of Economic Studies 2009 (1), pp. 1-18.
Horbach J., Remmer C., Rennings K. (2012). Determinants of eco-innovations by type of environmental impact and the role of regulatory push/pull, technology push and market pull. Ecological Economics 78 (June), pp. 112-122.
Huang Z., Chen H., Chen Z.-H., Roco M. C. (2004). International nanotechnology development in 2003: Country, institution, and technology field analysis based on USPTO patent database. Journal of Nanoparticle Research (6), pp. 325-354.
IEA (2009). Gadgets and Gigawatts. Policies for energy efficient electronics. OECD/International Energy Agency.
IEA (2012). Energy Technology Perspectives 2012: Pathways to a Clean Energy System. OECD/International Energy Agency.
Jaffe A. (1986). Technological opportunity and spillovers of R&D. American Economic Review 76, 984–1001.
Jaffe A., Trajtenberg M. (2004). Patents, Citations, and Innovations: A Window on the Knowledge Economy. Journal of Economic Literature, 42 (4), pp. 1158-1160.
Johnstone N., Hascic W. E., Poirier J., Hemar M. (2012). Environmental policy stringency and technological innovation: Evidence from survey data and patent counts. Applied Economics 17, pp. 2157–2170.
Johnstone N., Hascic W. E., Popp D. (2010). Renewable energy policies and technological innovation: Evidence based on patent counts. Environmental and Resource Economics 45, pp. pp. 133-155.
Kemp R. (1994). Technology and the transition to environmental sustainability. Futures, 26 (10), pp. 1023-1046.

Kemp R., Oltra V. (2011). Research insights and challenges on eco-innovation dynamics. Industry & Innovation 18 (3), pp. 249-253.

Kemp R., Pearson P. (2008). Measuring eco-innovation. Final report MEI project. Technical report, UNU-MERIT Maastricht.

Khazzoom J. D. (1980). Economic implications of mandated efficiency in standards for household appliances. Energy Journal, 1(4), pp. 21-39.

Khazzoom J. D. (1987).Energy savings resulting from the adoption of more efficient appliances. Energy Journal, 8(4), pp. 85-89.

Khazzoom J. D. (1989). Energy savings from more efficient appliances: a rejoinder. Energy Journal 10(1).pp. 157-166.

Kim Y.G., Suh J.H., Park S.C. (2008).Visualization of patent analysis for emerging technology. Expert Systems with Applications 34 (3), pp. 1804-1812.

Kohonen T. (1988). Self-organization and associative memory.3rd edition. Springer Series in Information Sciences, vol. 8.Springer-Verlag, Berlin Heidelberg, New York.

Kohonen T. (1990). The self-organizing map. Proceedings of the IEEE 78 (9), 14641480.

Kohonen T. (2001). Self-Organizing Maps (3rd ed.). Secaucus, NJ, USA: Springer-Verlag New York, Inc.

Kohonen T. (2012). Essentials of the self-organizing map. Neural Networks 37, pp. 52–65.

Kohonen T., Oja E., Simula O., Visa A., Kangas J. (1996). Engineering applications of the self-organizing map. Proceedings of the IEEE 84 (10), pp. 1358–1384.

Kostoff R., Toothman D., Eberhart H., Humenik J. (2001). Text mining using database tomography and bibliometrics: a review. Technological Forecasting and Social Change, 68, pp. 223-252.

Kuo R. J., Ho L. M., Hu C. M. (2002). Integration of self-organizing feature map and K-means algorithm for market segmentation. Computers & Operations Research, 29-11, pp.1475-1493.

Lanjouw J. O., Mody A. (1996). Innovation and the international diffusion of environmentally responsive technology. Research Policy 25 (4), pp. 549-571.

Lanjouw J., Pakes A., Putnam J. (1998). How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Applications Data. Journal of Industrial Economics, 46, 4, pp. 405-433.

Lanjouw J., Schankerman M. (2004). Patent quality and research productivity: measuring innovation with multiple indicators. Economic Journal, 114 (495), pp. 441-465.

Lee S., Park Y. (2005) Customization of technology roadmaps according to roadmapping purposes: Overall process and detailed modules. Technological Forecasting & Social Change 72, pp. 567-583.

Lee S., Yoon B., Park Y. (2009). An approach to discovering new technology opportunities: keyword-based patent map approach. Technovation, Vol. 29 (6-7), pp. 481-497.

Leydesdorff L., Kushnir D., Rafols W.E. (2014). Interactive overlay maps for US patent (USPTO) data based on International Patent Classification (IPC). Scientometrics, Vol. 98 (3), pp. 1583-1599.

Linares P., Labandeira X. (2010). Energy efficiency: Economics and policy. Journal of Economic Surveys 24 (3), pp. 573-592.

Luhn H.P. (1958). The automatic creation of literature abstracts. IBM Journal of Res. Dev., vol.2, pp. 59-65.

Lypsey R., Bekar C., Carlaw K. (1998). General purpose technologies: what requires explanation, in Helpman E. (Ed) General Purpose Technologies and Economic Growth. MIT Press, Cambridge, MA.

MacQueen J. B. (1967). Some methods for classification and analysis of multivariate observations. In Proceedings of 5th Berkeley symposium on mathematical statistics and probability, pp. 281–297.

Malerba F., Orsenigo L. (1996). The dynamics and evolution of industries. Industrial and Corporate Change, 5 (1), pp. 51-87.

Markard J., R. Raven, Truffer B. (2012). Sustainability transitions: An emerging field of research and its prospects. Research Policy 41 (6), pp. 955- 967.

Martinez C. (2010). Insight into Different Types of Patent Families, OECD Science, Technology and Industry Working Papers, 2010/02, OECD Publishing.

McKinsey & Company (2009). Pathways to a low-carbon Economy: Version 2 of the Global Greenhouse Gas Abatement Cost Curve.

Mogee M. E. (1991). Using Patent Data for Technology Analysis and Planning. Research-Technology Management , Vol. 34 (4), pp. 43-49.

Nameroff T. J., Garant R., Albert M. (2004). Adoption of green chemistry: an analysis based on us patents. Research Policy 33 (67), pp. 959-974.

Narin F. (2000). Tech-Line® background Paper. In J. Tidd, Measuring Strategic Competence. London: Imperial College.

Narin F., Noma E. (1987). Patents as indicators of corporate technological strength. Research Policy 16 (2-4), pp. 143-155.

Nesta L., Saviotti P. P. (2005). Coherence of the knowledge base and the firm's innovative performance: evidence from the US pharmaceutical industry. The Journal of Industrial Economics, 53(1), pp. 123-142.

OECD (2005). Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data, 3rd Edition, OECD Publishing, Paris.

OECD (2006).Can Energy-efficient Electrical Appliances be considered "Environmental Goods"? OECD Trade and Environment Working Paper No. 2006-04.

OECD (2010). Eco-Innovation in Industry: Enabling Green Growth. OECD Publishing.

OECD (2011). Fostering Innovation for Green Growth. OECD Green Growth Studies.

Oltra V., Kemp R., De Vries F.P. (2010). Patents as a measure for eco-innovation. International Journal of Environmental Technology and Management 13 (2), pp. 130-148.

Panzar J. C., Willig R. D. (1981). Economies of Scope. American Economic Review, vol. 71(2), pp. 268-272.

Park H., Yoon J., Kim K. (2013). Identification and evaluation of corporations for merger and acquisition strategies using patent information and text mining. Scientometrics, Vol. 97(3), pp. 883-909.

Pavitt K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. Research Policy, vol. 13 (6), pp. 343-373.

Polanco X., François C., Lamirel J. C. (2001). Using artificial neural networks for mapping of science and technology: A multi-self-organizing-maps approach. Scientometrics, Vol. 51(1), pp. 267-292.

Popp D. (2002). Induced innovation and energy prices. American Economic Review, vol. 92 (1), pp. 160-180.

Popp D. (2005). Lessons from patents: using patents to measure technological change in environmental models. Ecological Economics, 54 (2), pp.209-226.

Rennings K., Rammer C. (2009). Increasing energy and resource efficiency through innovation: An explorative analysis using innovation survey data. Czech Journal of Economics and Finance 59 (5), pp. 442-459.

Rosenberg N., Trajtenberg M. (2004). A general-purpose technology at work: The Corliss steam engine in the late-nineteenth-century United States. The Journal of Economic History 64(1), pp. 61-99.

Rip A., R. Kemp (1998). Technological change, in Human choice and climate change: resources and technology (S. Rayner and E. L. Malone, eds.), Battelle Press, Columbus, pp. 327-399.

Saidur H. H., Masjuki, Jamaluddin M.Y., Ahmed S. (2007). Energy and associated greenhouse gas emissions from household appliances in Malaysia. Energy Policy, 35 (3) pp. 1648-57.

Segev A., Kantola J. (2012). Identification of trends from patents using self-organizing maps. Expert Systems with Applications 39 (18), pp. 13235-13242.

Scherer F. (1982). Interindustry technology flows in the United States, Research Policy 11, 227–245.

Silva Almendra V., Enăchescu D., Enăchescu C. (2014). Ranking computer science conferences using self-organizing maps with dynamic node splitting. Scientometrics, Vol. 102 (17), pp. 1-17.

Sirilli G. (1997). Science and technology indicators: the state of the art and prospects for the future, in: Antonelli G. and De Liso N. (eds) Economics of structural and technological change. London, Routledge.

Sorrell S., Dimitropoulos J. (2008). The rebound effect: Microeconomic definitions, limitations and extensions. Ecological Economics 65 (3), pp. 636–649.

Sternitzke C., Bartkowski A., Schramm, R. (2008). Visualizing patent statistics by means of social network analysis tools. World Patent Information 30, pp. 115-131.

Ultsch A., Siemon H.P. (1990). Kohonen' s self-organizing feature maps for exploratory data analysis, in: Proceedings of International Neural Network Conference (INNC'90), Kluwer academic Publishers, Dordrecht, pp. 305-308.

Unruh G. C. (2000).Understanding carbon lock-in. Energy Policy, 28 (12), pp. 817-830.

van Pottelsberghe B., Dernis H., Guellec D. (2001). Using Patent Counts for Cross-Country Comparisons of Technology Output. STI Review, 27.

van Zeebroeck N., van Pottelsberghe De La Potterie B., Han W. (2006). Issues in measuring the degree of technological specialisation with patent data. Scientometrics, 66 (3), pp. 481-492.

Verspagen B. (1997). Measuring intersectoral technology spillovers: estimates from the European and US Patent Office Databases. Economic Systems Research, vol. 9, pp.47-65.

Vesanto J. (1999). SOM-based data visualisation methods. Intelligent Data Analysis, vol. 3, pp. 111-126.

Vesanto J., Himberg J., Alhoniemi E., Parhankangas J. (1999). Self-organizing map in Matlab: the SOM Toolbox. In Proceedings of the Matlab DSP conference (Vol. 99, pp. 16-17).

White H. D., Lin X., McCain K. W. (1998). Two modes of automated domain analysis: multidimensional scaling vs Kohonen feature mapping of information science authors. Advances in Knowledge Organization 6, pp. 57-63.

Yoon B. P. Y. (2004). A text-mining-based patent network: analytical tool for high-technology trend. Journal of high technology management research, Vol. 15, pp. 37-50.

Yoon B.-U., C.-B. Yoon, Y.-T. Park (2002). On the development and application of a self-organizing feature map-based patent map. RD Management 32 (4), pp. 291-300.

Appendix A

Table A1 - List of CPC-Y02B classes and related descriptions.

Y02B 40 - "Climate Change Mitigation Technologies"	
Y02B 40/30	**Refrigerators or freezers**
	Y02B 40/32
	Y02B 40/34
Y02B 40/40	**Dishwashers**
	Y02B 40/42
	Y02B 40/44
Y02B 40/50	**Washing machines**
	Y02B 40/52
	Y02B 40/54
	Y02B 40/56
	Y02B 40/58

Table A2 - List of search strings.

Electrical appliance	First Level Keywords	Second level keywords
Freezers and Refrigerators	energysav* OR energy efficien* OR energy conservation OR high efficien* OR low energy OR low-energy OR low electricity consumption OR energy reduction OR energy economis* OR energy economiz* OR energy performanc* OR less electric energy OR less electricity OR less energy OR energy use manage* OR energy ADJ use control* OR energy manage*) AND (residen* OR hous* OR domestic OR hom* OR dwellin* OR famil*)	refrigerator OR refrigerators OR fridge OR fridges
Washing machines		washingmachine*
Dishwashers		dishwash*

Appendix B

The theoretical backbone of the SOM resides on a lattice of interconnected nodes (neurons) to which input data are assigned through the similarity pattern that the process retrieves in the sample. The SOM is a lattice of nodes (map) where each neuron is connected to its neighbours. The lattice can be rectangular, hexagonal or irregular and its shape can be a plane, cylindrical or toroidal. A SOM requires an input vector of information. Assuming that such an input vector is defined as $x = [\vartheta_1, \vartheta_2, \dots, \vartheta_n]$, during the initialisation phase, the process assigns to each node (i) a corresponding weight vector $m_i = [\mu_{i1}, \mu_{i2}, \dots, \mu_{in}]$. Note that the two vectors must have the same length n. The weight assignment can follow a random process (random initialisation) or, as in this case, a ''regular, two-dimensional sequence of vectors taken along a hyper plane spanned by the two largest principal components'' of the input data (linear initialisation) (Kohonen, 2012 pp.6). Subsequently, the initialised map is trained with the multidimensional input data. Using a measure of distance (typically the Euclidean distance), the algorithm identifies for each x the most similar neuron (m_c) among the map nodes, minimising the vector distance between x and m_i. This neuron is labelled as the Best-Matching Unit (BMU) and calculated as follows:

$$\|x - m_c\| = \min_i\{\|x - m_i\|\}$$

for each neuron i. At this point, the SOM differs from the other VQ techniques by exploiting the learning process that is implemented in order to modify the weight of the BMU and the nodes close to it. Therefore, this portion of the map is modified to make the nodes more similar to the winning neuron and the latter more similar to the input vector. This smoothing effect in which each neuron in the neighbourhood of the BMU ''learns'' something from the input vector x, if protracted, leads to global ordering (Kohonen, 2001), which occurs when the algorithm converges. The basic updating procedure of the i-th node weight follows the formula:

$$m_i(t+1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)]$$

in which $m_i(t+1)$ is the node weight at time $t+1$,$m_i(t)$ is the weight assigned in the previous step, while $[x(t) - m_i(t)]$ is the Euclidean distance between input and node vectors. Finally, $h_{ci}(t)$ is the so-called smoothing kernel, defined as:

$$h_{ci}(t) = \alpha(t) \cdot \gamma$$

where

$$\gamma = exp\left(-\frac{\|r_c - r_i\|^2}{2\sigma^2(t)}\right)$$

The smoothing kernel h_{ci} consists of $\alpha(t)$, the monotonically decreasing learning rate factor ranging from [0,1], and of γ, the (Gaussian) neighbourhood function that determines the strength of the relation between the map nodes. The term $\|r_c - r_i\|$ defines the spatial relationships between the nodes in terms of the Euclidean distance between the location of the BMU (r_c) and other nodes (r_i). The radius σ^2 defines the width of the kernel around the BMU and decreases with time.

Such a process is iterated N times. In each interaction the radius determining the size of the BMU neighbourhood shrinks until only the best-matching neuron is included in it. To sum up, the SOM algorithm can be summarised in the following stages:

(a) assignment of the map node weight vectors (initialization phase);

(b) selecting an input vector from the dataset;

(c) calculating the Euclidean distance for each node in the map to find similarity between the input vector and the map node weight vector;

(d) tracking the node with the smallest distance as the best matching unit (BMU);

(e) updating the nodes in the neighbourhood of BMU by pulling them closer to the input vector through the learning formula;

(f) incrementing$t$ and repeating from (b) while $t < \lambda$.

The algorithm stops after a λ number of cycles where in each cycle the process is repeated for each input vector.

More recently, the use of a Batch algorithm (BA) instead of the sequential algorithm illustrated above, has been recommended (Kohonen, 2012). This is due to the fact that the BA produces more accurate results as well as less computational time. The BA differs from the sequential algorithm in the way input data are presented to the grid of neurons (step (b)). In particular, the whole set of input data is presented to the map at the same time (epoch) and only subsequently the nodes' weights are adjusted to reproduce the similarity between them. In this way, the order in which input data are presented to the map does not influence the final output. Therefore, after the initialisation phase, instead of modifying the weights of the nodes after each input data, the process first defines the BMU for each input vector. When all the inputs are associated to a node, the weight of each neuron is updated computing the mean of the $x(t)$ assigned (in the previous step) to the neurons placed in the kernel defined by the neighbourhood function as follows:

$$m_i(t+1) = \frac{\sum_{j=1}^{n} h_{ic(j)}(t) x_j}{\sum_{j=1}^{n} h_{ic(j)}(t)}$$

where $c(j)$ represents the BMU for the input data x_j and $h_{ic(j)}$ the neighbourhood function previously described.