Home green home.
Unveiling eco-innovation in energy efficient domestic appliances.

Nicolò Barbieri†, Valeria Costantini‡, Francesco Crespi‡, Alessandro Palma‡ *

†*Department of Economics, University of Bologna, Bologna, Italy*

‡*Department of Economics, Roma Tre University, Roma, Italy*

Preliminary Draft (June 2014)

Abstract

The present study uses an original dataset on four large energy-efficient (EE) appliances and provide a methodology for: i) identifying specific clusters of EE technologies; ii) mapping their evolution over time; iii) discovering niches of technological fungibility. Our model exploits the well-known concept of technological relatedness using co-occurrences analysis of patent classes as an input for Self-Organising Maps, an unsupervised artificial neural network able to represent high-dimensional data in visually-attractive and low-dimensional maps. The results confirm the pervasive nature of EE to be nested in many technological niches. Moreover, it is shown that a de-materialisation process affected the evolution of EE technologies over time, in a technological space characterised by high level of complexity and variety. Lastly, we show that digital components of EE technology can be characterised as a case where downstream technology complementarity is relevant.

Keywords: energy efficiency, Self-Organizing Maps, innovation, patent analysis, electrical appliances.

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

1. Introduction

There is overwhelming consensus among economists on retaining technological change as a key driver in increasing resource efficiency. Such a broad agreement is shared by policy-makers who consider energy efficiency (EE) as a crucial economic and environmental long-run objective (IEA, 2012; EC, 2011). A clear specification of the jargon is here necessary, since energy saving and energy efficiency are not completely overlapping terms (Gillingham *et al*., 2009; Linares and Labandeira, 2010). Indeed, EE is a sub-set of the energy saving (or energy conservation) domain. This latter is a broader concept since energy conservation can be achieved by increasing the EE or by reducing the level of economic activity which may also mirror a change in consumers' behaviour. EE, on the other hand, is the relationship between output produced and energy consumed to produce it, often called energy services. Precisely, EE can be defined as the ratio of useful outputs to energy inputs for a system. The underlying system may be an individual energy conversion device (e.g., a boiler), a building, an industrial process, a firm, a sector or an entire economy. Thus, a general characteristic of EE is the use of less energy inputs for an equivalent (or even augmented) level of economic activity or service. In other words, gains in EE can raise the level of energy services, reduce the level of energy inputs or produce both effects. Given this definition, achieving higher EE performances intrinsically relies on technological innovation as a mean for improving productivity of the energy input (Rennings and Rammer, 2009). In this respect, the present study takes advantage of the increasing attention devoted to the analysis of eco-innovation (Kemp and Pearson, 2008; OECD, 2010) by the international scientific community, which is focusing on a growing number of technological domains, and it is performing more and more accurate statistical and economic analyses (Arundel *et al*. 2011; Berkhout 2011; Borghesi *et al*. 2013; Hascic *et al.* 2009; Horbach *et al*. 2012; Johnstone *et al*. 2012, 2010; Kemp and Oltra 2011; Lanjouw and Mody 1996; Markard *et al*. 2012; Nameroff *et al.* 2004; OECD 2011; Popp 2002). In the analysis of eco-innovation technological domains the role of patents has been largely exploited as they allow to analyse specific features as path-dependence, lock-in process (Arthur, 1989; Kemp, 1994; Rip and Kemp, 1998; Unruh, 2000)and the evolution of technological trajectories (Khun, 1969;Verspagen, 2007;Consoli and Mina, 2009).

Among all the residential EE technologies, domestic electrical appliances seem to be particularly interesting to investigate. Indeed, domestic appliances represent an important share in the final energy consumption, having at the same time high potential in terms of energy efficiency gainsthanks to their multiplication effect, i.e.themarginal contributions of each single appliance multiplied by the total number of appliances. Such appliances are systematically diffused in each dwelling, thus marginal energy efficiency gains can reduce[1] the level of energy demand (IEA, 2009) and, consequently, the level of GHG emissions deriving from electricity. Moreover, in recent times, electrical appliances show a growing level of technology integration, being the result of a wide set of industries and scientific branches. For instance, more and more appliances incorporate stand-by devices, digital displays, more sophisticated process for freezing and washing as well as many other functions.

All this considering, the aim of this paper is to find out how the aim of EEpervades the technological space of four large electrical appliances of common use, and precisely freezers and refrigerators, washing machines and dishwashers. By employing the means of patent maps through the use of SOMs - an unsupervised artificial neural network (Kohonen, 1988; 1990), we exploit an

1 Although the empirical literature indicates the existence of the so called "rebound effect" (Khazzom, 1980; 1987; 1989; Sorrell and Dimitropoulos, 2008). However, as far as the level of rebound effect maintains less than 100%, the positive contribution of energy efficiency is only partially offset.

original dataset of 688triadic patent families on EE electrical appliancesfor: i) identifying specific clusters of EE technologies; ii) mapping their evolution over time; iii) discovering fungibility in technological niches.

The rest of the paper is organised as follows. Section 2 describes the state-of-the-art in using patents for analysing technological domains in the flourishing literature of eco-innovation. Section 3introduces the particular technological domain of residential energy efficiency, with a focus on electrical appliances which constitutes the domain of analysis. In Section 4, the methodology for collecting our patents sample is described, along with the theoretical foundations of SOMs and the model design. Section 5providesthe experimentresults, while Section 6 concludes the paper.

2. Patents as a tool for analysing eco-innovation dynamics.

There are a number of possibilities for measuring innovation activities. As most economic variables, the problem of measurement is directly related to the availability and the quality of specific data. After several years, the international scientific community seems to have achieved a reasonable level of data standardization and reliability, even though the methodologies of analysis and related results are continuously in progress, leading to different metrics and interpretations of innovation performances (Archibugi and Pianta, 1996; OECD, 2005; Sirilli, 1997).

The most used innovation input and output indicators have been subject to much criticism (Sirilli, 1997). On the one hand, the growing literature on innovation indicators has shown that the resources devoted to research and development (R&D) represent only one source of innovation and that other innovation inputs might be relevant but are not easily measurable. On the other hand, although patents provide many information and allow for time-series analysis, they have been shown to be imperfect indicators of the inventive activity and their use strongly depends on some important limitations (see, among all, Griliches, 1998).

The first limitation is that patents are only one of the different mechanisms for protecting innovations, along with lead time, industrial secrecy or purposefully complex specifications (Frietsch and Schmoch, 2006). Indeed, the patenting behaviour can be different among firms. In particular, inventors may prefer secrecy to prevent public disclosure of the invention imposed by patent law, or to save the significant fees attached to patent filing (Archibugi and Pianta, 1996; Jaffe and Trajtemberg, 2004). Moreover, the possibility to innovate "around the patent" as well as the detailed description included in the patented inventionsconstitutes a source of diffusion of information and might translate in a limitation to patent (Oltra *et al.*, 2010). As a consequence, patented inventionsonly represent an incomplete share of the invention process, although there are very few examples of economically significant inventions which have not been patented (van Pottelsberghe *et al.*, 2001).

The propensity to patent may also differ among countries depending, respectively, on the nature of the technology and on the risk of imitation in one country (Cohen *et al.*, 2000). Hence, patenting activity is more likely to concern countries with technological capabilities as well as strict enforcement of intellectual property rights. Besides this, a number of empirical studies confirm that patenting propensity depends also on the specific industrial sector under scrutiny (Arundel and Kabla, 1998; Malerba and Orsenigo, 1996; Pavitt, 1984).

A further source of limitation is that a national patent only grants the exclusive right to use the technology in a given country; this does not mean that the patent owner will actually do so but could significantly bias researcher's results if the protection was not costly, so that inventors might patent widely and indiscriminately. But this is not the case in practice. First of all, patenting activity is costly – including the costs of preparation for the application process and the administrative costs and fees associated with the approval procedure. Moreover, if the enforcement is weak, the publication of the patent in the local language can increase vulnerability to imitation. As a result, inventors are unlikely to apply for patent protections in a country, unless they have the relatively certainty of the potential market associated to the discovered technology. The possibility to filepatents in the international patent offices (such as EPO or WIPO) allows firms for a multi-country protection of their inventions (van Zeebroeck *et al*., 2006).

Lastly, the value of individual patents is heterogeneous since patents differ greatly in their technical and economic significance. Many of them reflect minor improvements of little economic value so that the distribution of patents value is skewed (Griliches, 1998). The OECD Triadic Patent Family[2] database assumes relevant importance in recognizing high-quality patents, since the use of patent families enables to focus on the most valuable inventions. Besides this, several methods have been developed to enrich the level of information related to patents value (Lanjouw *et al*., 1998) such as the use of weights based on the number of times a given patent is cited in subsequent ones (Hall *et al*., 2005) or indexes based on multiple indicators (Lanjouw and Schankerman, 2004).

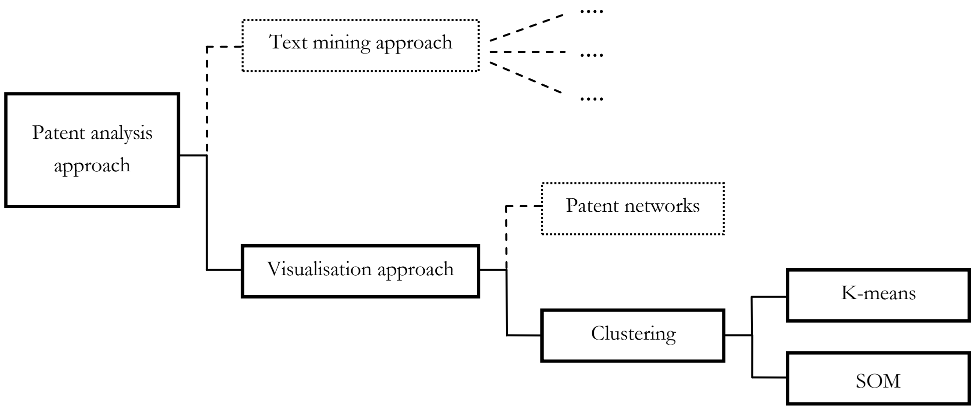
Despite these limitations, the use of patent data is widespread in the economics of innovation literature. Patents provide a public wealth of information on the nature of the invention and of the applicant for rather long time series, indicating not only the countries where inventions are developed, but also where these new technologies are used and derive from. Patent data frequently represent the direct result of R&D processes, a further step toward the final output of innovation that is useful knowledge through which firms are able to generate new income. Patent applications are usually filed early, hence they can be interpreted not only as a measure of innovative output, but also as a proxy ofinnovative activity (Popp, 2005).Moreover, patent data provide a detailed description of the technical contents of each invention and are subject to an extensive updating process of their informative content, continuously enriched by national and international patent offices. These latter, for administrative purposes and according to international procedures and agreements, classify patent data in specific technological areas (classes). The most important classification system is the International Patent Classification (IPC) developed by WIPO in 1971. Such classification system should allow researchers to properly identify relevant patents in specific technological domains.

2.1. Patent analysis.

The number of patents has been increasing steadily and experts of patent analysis can rely on a broad set of methods and techniques aimed at identifying coherent information for different purposes; these include, among all, identification of potential research areas and technological trends (Narin and Noma, 1987; Mogee, 1991) as well as strategic support for R&D through Technology Road Map (Lee and Park, 2005; Lee *et al*., 2009). Basically, patent analysis techniques automatically reduce the large amount of information provided by patent documents to useful low-dimensional information. In classifying the numerous approaches for patent analysis, Abbas *et al*. (2014), propose ataxonomy which distinguishesbetweendata processing using text-mining techniques and visualisation maps, as showed in Fig. 1.

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.

Figure 1 - Techniques taxonomy for patent analysis.



Source: Abbas et al. (2014).

In our view, such a classification mirrorstheconsequentiality of the outputs obtained by processing patent data rather than the mere differences between the two approaches. Focusing on the first stage of analysis (data processing), Lee *et al*. (2009) classifies patent data analysis between 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 citations count, which are analysed mostly exploiting bibliometric techniques such as citations analysis (Yoon and Park, 2004, among others). On the other hand, unstructured patent items (UI) are composed by free text strings, in some cases very long as in the case of patent description or claims. Since the innovative content of a patent is enclosed in its UIs, unstructured data analysis using text-mining techniques represents a powerful alternative. According to Kim *et al*. (2008), TMtechniques assign a label to each document's word and process the output to extract a set of keywords by using text-mining algorithms (see also Kostoff *et al*., 2001). As a result, raw patent data written in natural language are transformed into structured data by which useful information such as similarity patterns or technology trends can be retrieved. Despite TM provides a valid toolset for extracting patent knowledge, it is not exempt from some limitations. For instance, Abbas *et al*. (2014) point out that TMcan lead to incorrect and misleading interpretation of the texts, together to limited classification of synonyms in large text documents. Moreover, the use of stop words[3] (or negative dictionaries) largely employed by analysts for cleaning UIs,"can introduce biases and somehow jeopardize the meaning and usefulness of the [output] map" (Blanchard, 2007, p. 315).

2.2. Patent mapping

The visualization process constitutes an important part in explorative patent data analysis, especially when the level of information complexity and data dimensionality are high (Vesanto, 1999). As a consequence, the representation of patent information content by the means of maps is widespread in the literature of patent analysis and numerous are the techniques devoted to this aim.

---

3 Lists of non information-bearing words. Such a list is based on the principle that "the frequency of word occurrence in an article furnishes a useful measurement of word significance" (Luhn, 1958, p. 60).

Broadly speaking, a patent map[4] is able to show complex and invisible relationships among different patent documents as well as their peculiar features exploiting a simpler low-dimensional visualization. There exist many typologies of patent maps, which also differ according to the techniques chosen for pre-processing raw data. Abbas *et al*. (2014) classifies maps in patent networks and cluster-based maps. In a patent network, the relationships between objects are investigated by analyzing 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 represents now 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 all). A patent map can also derive from a clustering process, reducing observations intogroups ''internally homogeneous (internal cohesion) and heterogeneousfrom group to group (external separation) [...] reducing the space dimensionality'' (Giudici, 2003, pp.76). For instance, Kim *et al*. (2008) clustered patent documents using the K-means algorithm including both SIs and UIs, visualizing results in a semantic network of keywords.

Patent maps may also differ in the output map that they produce. Bibliometric analysis uses two common mapping techniquesto detect information, i.e. graph-based mapand distance-based map (Van Eck and Waltman, 2010). While in the former the focus is on whether the items are linked or not, the lattercapturesthe strength of these relationships reflecting themthrough a spatial order based on the relatedness between objects. Thisclass of mapsmeasure the similarity between the input data and represent the output in a low dimensional space where, usually, the lower the distance between items the more similar they are.

A further promising approach in this sensewere mapsbased on neural networks, which showed a high level of efficiency in managing high-dimensional observed data, incomplete information, errors or inaccuracies. There are many types of artificial neural networks (for a comprehensive classifications, see Giudici, 2003) but a first important distinction can be made between supervised and unsupervised ANNs. Differing from supervised[5] ANNs, more suitable for prediction analysis, a SOM is based on unsupervised learning processesable to map every dimensional observation in a spatial grid of output; such feature makes it particularly effective for classification and clustering analysis. Indeed, the nodes are placed in such a way that those adjacent will be more similar than distant output nodes, thus introducing a topological dependence between clusters while preserving the spatial correlation among the input vectors and the clusters.

While the use of SOMs has been increasingly adopted in several applications, as for instance scientific journal networks (Campanario, 1995), author co-citation data (White *et al.*, 1998) or mapping of industrial districts (Carlei and Nuccio, 2014), there are only a limited number of applications of patent analysis using SOMs and, to ourknowledge, all of them usedUIs for analysing technical patents content.

A first paper by Yoon *et al*. (2002)applied SOMs in order to show complex relationships and dynamic patternsamong different technologies. In particular, they built technology vacuum, claim point and technology portfolio maps, for the identification oftechnology missing areas, potential infringements and technology classifications, respectively.Afurther recent contribution using TM techniques as an input for SOMs is given by Segev and Kantola (2012). They used Term Frequency-Inverse Document Frequency (TF-IDF) algorithm to extract patent knowledge, then represented by the means of SOMs, Theyalso compareSOM performances with K-means (MacQueen, 1967) and

4 Map is here used as a generic term, being synonym of diagram, chart or graph.

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

with Density-Based Spatial Clustering of Applications with Noise (DBSCAN, Ester *et al.*, 1996) cluster classifications. By doing so, they provided an affective methodology for representingexisting knowledge boundaries and research trends using PMs.

3. The technological domain of residential energy efficiency.

Although their importance has been recognised only in recent times, EE technologies constitute now a relevant share of climate change mitigation technologies. Notwithstanding, due to their nature, EE technologies are only partially and roughly represented in the set of international patent classification. Indeed, when one looks at the relationship between EE and its technological content, the former shows two important features. First of all, EE appears as a latent technological domain since the improvements in EE are not always explicitly mentioned by the main non-structured items of patent documents, namely title, abstract and claims. Hence, it is necessary to analyse the full text of the document, including also the patent description. This implies time-demanding analysis, high calculation capacity and efficient text-mining algorithms. Such a difficulty in capturing the latent EE technologies is also a reason of delay for developing a comprehensive patent classification by the most important patent offices, although the recent new Cooperative Patent Classification (CPC) includes, since 2013, explicitly EE patent classes for some domestic electrical appliance. The second important feature of EE is its pervasiveness to be embodied in the devices, since EE not only operates in the most advanced technologies but, comparatively, in the entire panorama of technologies using energy. Hence, within a single technology paradigm, it is possible to find different EE levels, so that it is possible to compare – in terms of energy use – different appliances using the same technologies. In other words, there is not always full identification between technology improvements and EE gains, since this latter can or cannot be included in the innovative content of a patent. In some other cases, the identification is very weak and gains in EE have few links with patented technological improvements because it can simply derive by the use of better quality of material (e.g. windows insulation) or concepts difficult to patent (e.g. buildings orientation or design). In these last cases, we would fail in believing that EE benefits from eco-innovation and the actual effect of technology in boosting EE might be underestimated (see Costantini *et al.*, 2014).

3.1. Energy efficient electrical appliances.

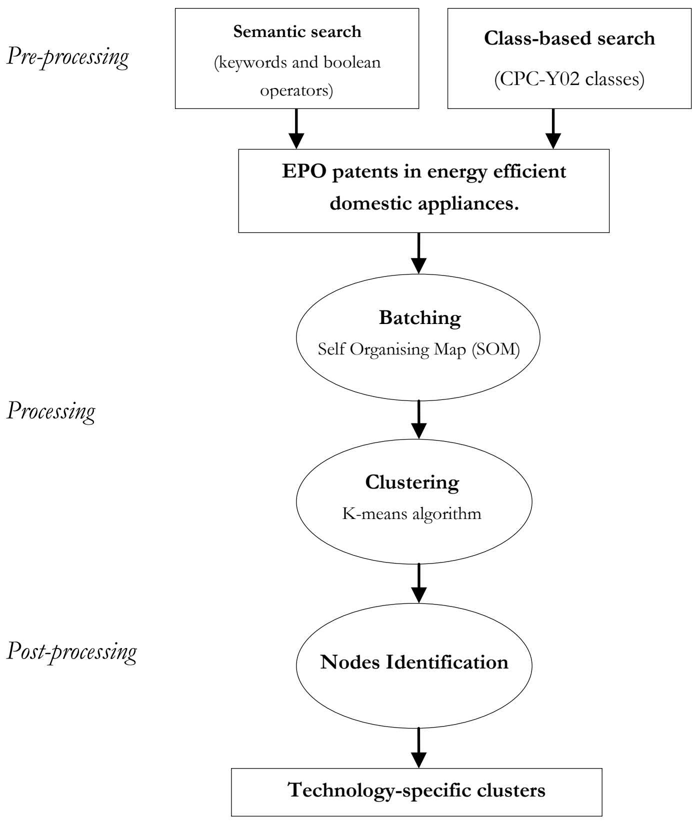
Among all the residential EE technologies, domestic electrical appliances seem to be particularly interesting to investigate. Indeed, domestic appliances are showing one of the world's fastest-growing segments of total energy consumption, representing an increasing share in the end-use energy consumption.According to OECD (2006), the growth of refrigerators, washing machines, lighting, water heaters, air conditioners, computers, fax and photocopying was around 3,7 per cent in the decade 1992-2002 and such a growth is projected to increase in the future with growth rates much higher in developing countries.At the same time,residential electrical appliances show a high potential in terms of energy efficiency gains, although the differences in the energy performance of similar appliances can be large among different countries. For instance, in the EU area theleast-energy-efficient refrigerators on sale consume up to three times the electricity of the most efficient despite Minimum Energy Performance Standards being in place since 1999. In the United States and Canada, energy-efficiency programmes are credited with helping to reduce the amount of energy used to power new models of refrigerator-freezers by over two-thirds between 1973 (OECD, 2006).

Such performances are mainly due to the multiplication effect of marginal contributions of each single appliance by the total number of appliances. Since domestic appliances are systematically diffused in each dwelling, this can produce a large impact in terms of energy reduction (IEA, 2009).These characteristics make the diffusion of efficient domestic electrical appliances an important policy goal as a source for reducing final energy consumption, thus contributing to the energy independence, to mitigate harmful GHG emissions as well as to improve the economic efficiency as a whole. On the same time, a methodology for the identification of EE technologies is required and the present work fills this gap and provides empirical results for mapping and analysing the structure and evolution of EE technologies, confirming the hypothesis of latent and pervasive technological domain.

4. Patent analysis model

From a methodological perspective, the objective of the present work is to provide a set of instrument and procedures that allow, through the creation of a patent map (PM), to retrieve the complex technical relationship among patents related to energy efficiency.

Figure 2 – Patent Analysis Model



To this aim, we implement a Patent Analysis Model (Figure 2), characterised by three main phases: pre-processing, processing and post-processing.The first phase is devoted to obtain a set of patents belonging to energy efficient electrical appliances, while the last two – which constitute the core of our work – identify the technological niches in which energy efficiency is nested among numerous appliance components, also analysing their technological evolution over time. In doing so, we provide an efficient model to analyse patents in any technological fields in which the technical boundaries between the subfields are latent or even unknown.

4.1. Sample selection

As previously stated, EE technologies for electrical appliances has been only recently incorporated in the international patent classifications. In order to overcome such a lack and to build a coherent set of patent documents to be analysed, both a top-down and bottom-up approach is proposedfor catching patentsfiled at several patent offices, adopting a multi-stage process.We combined the new CPC-Y02B classes[6] with a set of keywords by an ENEA'steam of experts[7].Then,we proceed to the patent extraction using the Thomson Innovationpatents search engine.

The top-down search method employs the CPC-Y02B ''Climate Change Mitigation Technologies'' classes, as shown in Table A1.

The bottom-up approach is characterized by two levels of searching, using keywords together with Boolean operators (AND, OR, NOT). Both the levels of search have been implemented on the full patent text, including the most important UIs such as title, abstract, claims and description of patent applications. The first level search 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 EE technologies applied to residential sector, using words like ''refrigerator'', ''washing machine'' and so on (search strings are provided in Table A2). By doing so, we reduced the set of applications to three electrical appliances: refrigerators and freezers, washing machines and dishwashers.

Subsequently, we collected the patent family of each patent. This procedure allows to drop double counting of patents that refer to the same technology but whose protection had been extended to many patent offices. Furthermore, to increase the quality of our dataset, we focus on triadic patent families (Martinez, 2010). The use oftriadic patent families reduces the sample further cleaningthe patent dataset, and permits to focus on high-valued patents, since the cost to file for patent protection in more countries is associated with a higher costs.

As a final step, weeliminated patent classes not belonging to the domain of electrical appliances and tested a sample of 15% as further manual validation process, obtaining a unique dataset of 688patents over the time span 1990-2014and divided by the three domestic appliances, as specified in Table 1

Table 1- Patents sample, by appliance.

	Patents	Share
Dishwashers	66	9,60%
Refrigerators and Freezers	489	71,08%
Washingmachines	133	19,33%
	688	100,00%

Source: own elaboration.

6 CPC-Y02B classes are those related to “climate change mitigation technologies related to buildings, e.g. including housing and appliances or related end-user applications”. For more information, see: http://www.cooperativepatentclassification.org/cpc/scheme/Y/scheme-Y02B.pdf

7 ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development.

4.2. Building the patent maps

Despite the increasing number of studies that use citation data and connectivity analysis (Verspagen, 2007) to investigate the technological importance of inventions and to detect technological trajectories, the present paper applies PMs to unveil the domains of energy efficient technologies. Our choice is twofold. First, the characteristics of energy efficient technologies may permeate many technological domains decreasing the usefulness of using citation data to detect those domains (see Section 3). Secondly, the use of citations would have produced biased information since recent patents show lower probability to be cited with respect to older patents (see, among others, Jaffe and Trajtenberg, 2004). In our case, this is particularly true if we consider that our patents sample is referred to a rather long time series (25 years).

In order to create the PM, the Self-Organising Map (SOM), a topological ordered mapping technique firstly introduced by Kohonen (1988; 1990), has been applied. The process is inspired by the system of cells that compose ''brain maps'', in which some neural cells respond selectively to specific external sensory stimuli. In addition, the topological position of these cells assemblies, within the brain structure, behaves in some coherent ways to specific stimuli. A similar process can be imitated using the SOM, a two-layer unsupervised competitive Artificial Neural Network (NN) able to represent multidimensional data onto a two dimensional topological grid (Kohonen, 2001). This technique is a nonlinearity projecting mapping in which the input data become spatially and globally ordered (Kohonen, 2012). Indeed a SOM reduces complex nonlinear statistical relationship into a more simple, easy to understand and graphically attractive low-dimensional display (Kohonen *et al.*, 1996), in which the topological relationship between input data, that tend to be clustered, is preserved.

The SOM resembles the Vector Quantisation (VQ) process, 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. What makes the SOM a suitable tool for multidimensional reduction, is its capability to provide a spatial and global order within the output map. Thus, similar input data are placed closer in the map, while different input data gradually farther away (Kohonen, 2012). Such a feature is provided by the adaptive units (that compose the map) able to modify their response in such a way that the position of the node in the map becomes peculiar to the features in the set of input signals (Yoon *et al*, 2002). In addition, the SOMs have the capability to learn, from input data, how to represent them in the more effective way performing dimensionality reduction. Given the peculiarities of EE technologies (see Section 3), this is a desirable feature for a deep analysis of our technological domain by the means of EE patent documents, where the complexity of data may hide the information to be extrapolated.

To build the SOM we used the SOM Toolbox[8], a free Matlab© function package developed by the SOM Toolbox Team at the Helsinki University of Technology (Vesanto *et al.*, 1999).

4.2.1. The SOM algorithm

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

8 The SOM Toolbox is downloadable under GNU General Public License at the website: http://www.cis.hut.fi/projects/somtoolbox/

SOM Toolbox is Copyright (C) 2000-2005 by Esa Alhoniemi, Johan Himberg, Juha Parhankangas and Juha Vesanto.

lattice can be rectangular, hexagonal or irregular and its shape can be plane, cylindrical or a toroid. A SOM requires an input vector of information. Assuming that such an input vector is define 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 weights 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's 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 mapis modified to make the nodes more similar to the winning neuron and this 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 continuing 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-calledsmoothing kernel, defined as:

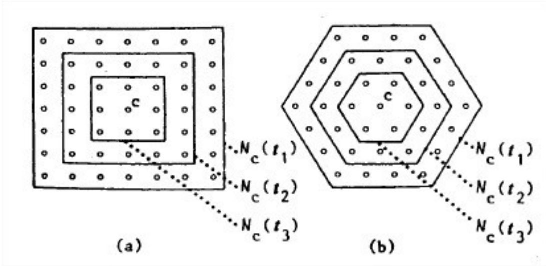
$$h_{ci}(t) = \alpha(t) \cdot \gamma$$

where

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

The smoothing kernel h_{ci} is composed by $\alpha(t)$, the monotonically decreasing learning rate factor rangingfrom [0,1], and by γ, the (Gaussian) neighbourhood function that determines the strength of the relation between map's nodes. The term $\|r_c - r_i\|$ defines the spatial relationships among the nodes in terms of 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 decreaseswithtime as shown in Figure 3.

Figure 3 – Monotonically decreasing σ^2.



Source: Kohonen (2001)

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 summing up, the SOM's algorithm can be summarisedin two stages:

(a) assignment ofthemap's nodes' weight vectors (initialization phase);

(b) selecting an input vector from the dataset;

(c) calculate for each node in the map the Euclidean distance to find similarity between the input vector and the map's nodes 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.

Kohonen (2012) suggests using a Batch algorithm (BA) instead of the sequential algorithm illustrated above, producingmore results accuracy 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 present 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 firstly 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.

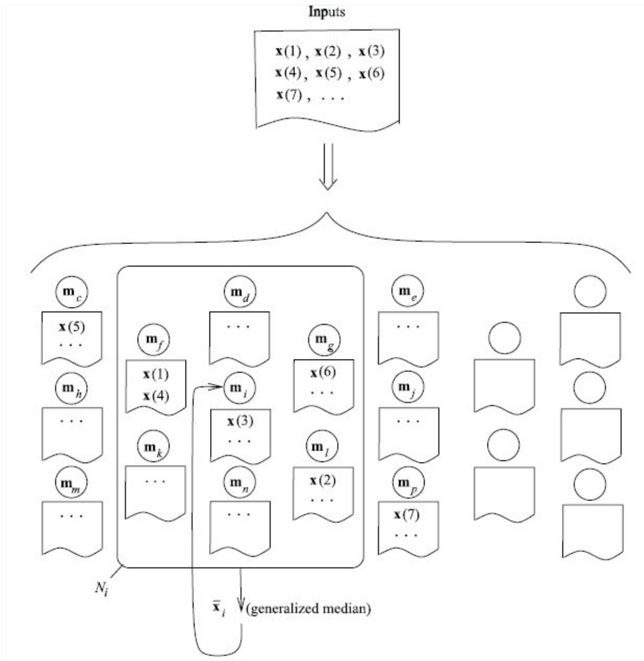
Figure 4 illustrates how this process works. The input data $x = [1,2, \dots 7]$ are assigned to their BMUs, whose nodes weights minimise the Euclidean distance.

The map's weights are updated in the subsequent step. The neighbourhood function defines the area of the map that impacts on the node's weights N_i. Finally, for each m_i the mean between the input data assigned to the neurons within the neighbourhood function is calculated, and the weights of the nodes are modified 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.

Figure 4 – Illustration of batch algorithm weights update.



Source: Kohonen (2012)

4.2.2. SOM input matrix.

In order to define technological niches that characterise energy efficiency inventions, we apply the SOM to create a PM that will return the technological niches in which inventive efforts are nested when the incorporation of EE technologies in four large electrical appliances is investigated.Instead of using TM techniques for extracting patent knowledge as in Yoon *et al*. (2002) and in Segev and Kantola (2012), the PM is developed using 8-digit CPC (Cooperative Patent Classification) classes assigned to each patent. Such classes label patents according to their technological content through a hierarchical, language-independent classification system. The CPC has been established in 2010 as a joint partnership between the USTPO and EPO to provide a harmonization between the two classification systems developed by each office (European Classification and United States Patent Classification, ECLA and USPC hereafter). We assume that the presence of some CPC classes within single patents characterises the technological domaintowhichtheinvention refers. 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 and Verspagen, 1997, among others).

In the related literature, many efforts have been pursued in order to measure the technological relatedness among patents. For instance, Leydesdorff *et al*. (2014) built a matrix for measuring the number of time 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) have used patent classification co-occurrences to measure the strength between technological fields.

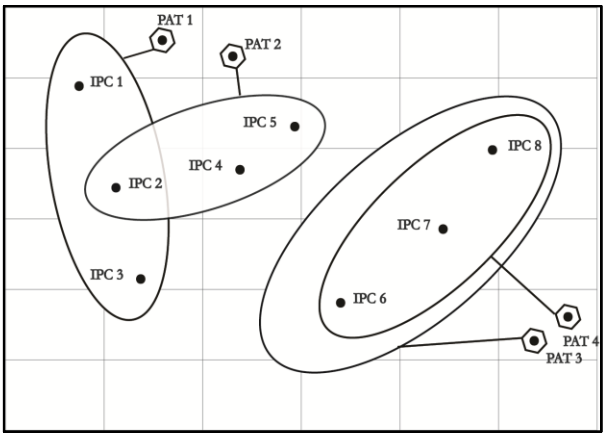
Starting from the methodological contributions that these studies have proposed, we use co-occurrences among CPC classes to define patents similarities. In addition, by employing SOM to map the patent dataset, we can define a space where the rate of similarity between technological fields decreases gradually according to the distance among patents. Therefore, the input data of the SOM are, in each column, the frequency of 8-digit CPC classes assigned to each patent, while in the rows the patent ID (Application Number), as showed 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 5illustrates the mechanisms, showing that Patent 1 and Patent 2 share only one CPC class (CPC 2), while Patent 3 and patent 4 show the same set of CPC code. Therefore, due to their technological similarities, Patent 3 and 4 are placed in the same position, and far away from Patent 1 and 2 that refer to different technical developments. Finally, Patent 1 and 2 are located close, but not in the same position (due the fact that they share just one CPC class). This bearing in mind, our choice for applying the SOM relies on the achieved local spatial order that distinguishes this methodology from the others. Indeed, as previously explained, when the input data are processed and placed into the map, the portion of the map around the BMU is further modified for better matching. Subsequently, the local order propagates on the whole map with the increasing number of iterations, reaching a final global order.

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



Source: own elaboration.

4.2.3. SOM visualisation

Once the SOM has been trained, a visualisation process is required in order to show the resulting output. Such a process basically consists in a method for locating the BMUs in an effective and visually-attractive map (projection). Among different techniques (see Vesanto, 1999), the Unified Distance Matrix (U-Matrix) is here proposed, choosing a grid with hexagonal lattice[9] (Ultsch and Siemon, 1990). By assigning different colour hues according to the distance between each map's node, distance matrices showsimilarity among SOM's nodes.

9 The choice of the lattice mirrors only esthetical reasons, not producing bias in the data representation.

4.2.4. Clustering

Based on the spatial order of input data, obtained using the SOM, map's nodes are clustered through the non-hierarchical K-means algorithm (MacQueen, 1967). The synergy that arises from the use of this two-stage method, generally produces more powerful results than using them singularly (Chi and Yang, 2008; Kuo *et al.*, 2002). Through the application of the K-means method, the nodes are partitioned into k groups. The clustering process is based on the concept of centroid, i.e. points with a low distance between them and the other elements of the cluster. The number of clusters is defined by choosing k in order to minimise the Davis-Bouldin (DB) index (Davis and Bouldin, 1979), a clustering performance index that measures compactness and separation between nodes and clusters.

4.2.5. Benefits and limitations of SOMs

One of the main benefits of the SOMs is the local and global order that they provides to thefinal map.Differently from other mapping techniques where each observation is represented in a single unit in the final output (e.g. citation networks), the SOM algorithmassignseach input datato the neurons, whose number – determined before running the SOM,is usually lower than the number of items to be mapped, providing a first reduction of the output complexity (especially for large datasets). Thus, the highest similarity between input data is represented through their location in the same neuron of the final map.Moreover, the neurons can be mapped through a Unified Distance Matrix (U-Matrix), which includes the neuron distances to the adjacent neurons.

However, the heuristic choice of the number of neurons may increase the computation efforts.In addition, a too small (or too large) map maynegatively impact data visualisationand cluster structure represented in the map[10].

5. Unveiling energy efficient technologies: experiments design and results.

Experiment I – Patent Maps

As a first step, weuse the entire patents sampleto discover technology niches in which EE plays a role. Such niches are presented on the maps as clusters that include nodes with similar technological content, previously identified by SOM. We repeat this experiment for each of the three appliances, providing a clear picture of the pervasive nesting of EE in different natures of appliance components such as mechanical, electro-mechanical, digital, chemical as well as in operational processes. The resulting framework is thus characterised by high technological complexity, generated by numerous components deriving from a wide range of scientific and industrial contributions. According to Antonelli (2003), ''the opportunities to generate new knowledgeare conditional on the capability to draw togetherbits of knowledge that are actually diverse and yetcomplementary''(page 598), a concept known in literature as resource pooling (see also Bresnahan and Trajtenberg, 1995; Lypsey *et al.*, 1998; Bresnahan, 2010).

In the case of refrigerators and freezers,whose sample is based on 489 patents, we identified seven clusters (Figure 6). A first, generic cluster (#1) including patents on new refrigerators and freezers as a whole can be identified. Besides this, if we look at the remaining clusters, a set of

[10] We employed an heuristic formula proposed within the SOM Toolbox: $n = 5\sqrt{dlen}$, where n is the number of units that compose 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 eigenvectors of the training data, during the initialization phase.

specific technology niches can be unveiled, constituting a technological decomposition of the appliance under scrutiny, in this case a domestic refrigerator or freezer. As showed in Table 3, clusters 7, 3, 5, 4, 2 and 6, refer to various technology fields, ranging from electrical components (as in energy management systems, which include sensors, microprocessors, displays etc.) to chemical components (e.g. refrigerant compositions, insulating foams and lubricant oils).

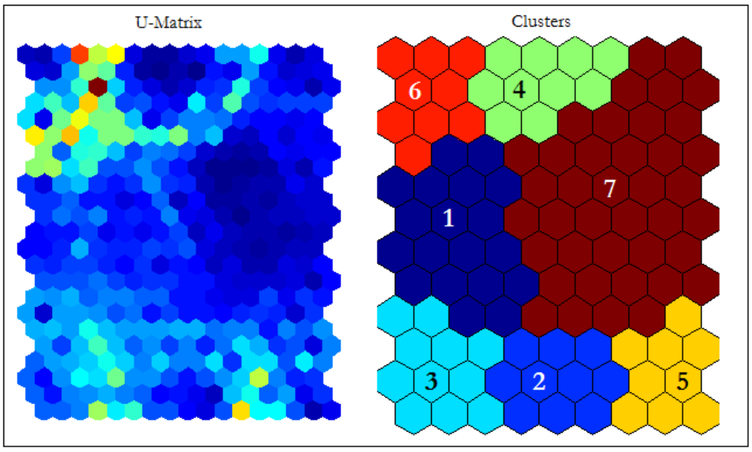
The presence of EE appears also when washing machines and dishwashers are under scrutiny (Fig. 7 and 8). In fact, although the number of patents related to these two appliances is lower, the pervasiveness of EE affect, as in the previous case, a number of different clusters. In particular, we identified 5 clusters (133 patents in total) in the case of washing machines and only 4 (66 patents in total) in the case of dishwashers (Table 4 and 5, respectively). The identification and study of these clusters confirms the hypothesis of EE pervasiveness and the presence of resource pooling effect also in these two electrical appliances, since 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 derive also a measure of innovation effort in each particular technological niche. In other words, it is possible to identify where most of the efforts for EE gains have been addressed in each appliance over the entire period of analysis, specifying that such a rank assumes only qualitative nature since the assessment of the technological value of each cluster(and hence of patents) is not taken into account in this paper.

Table 3 - Cluster identification for refrigerators and freezers.

Refrigerators and Freezers		
Cluster #	Technology description	% patents share per cluster
7	Mechanical and electrical components (compressors, pumps etc.)	43,15
3	Refrigerants circulation systems	12,27
5	Components for power and control management	12,07
1	New refrigerators and freezers (as a whole)	11,45
4	Heat transfer and refrigerant compositions	8,18
2	Insulation panels and foams	6,54
6	Lubricant oils	6,34
	Total	100

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

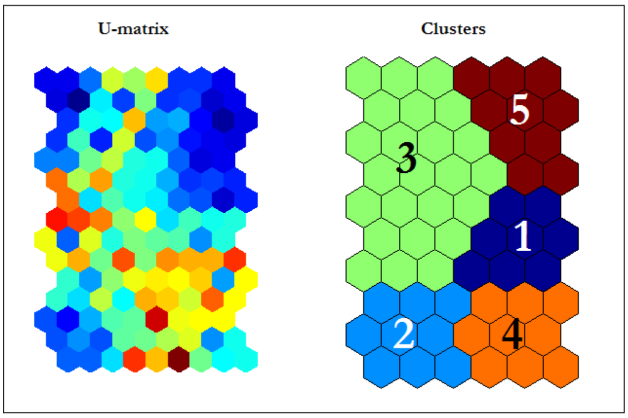


Source: own elaboration on MatLab.

Table 4- Cluster identification for washing machines.

Cluster #	Technology description	% patents share per cluster
Washing Machines		
3	Mechanical and electromechanical components.	47,37%
5	Washing process/methods and washing machine as a whole	21,80%
4	Digital components for energy management	12,03%
2	Sensors	11,28%
1	Motion and heating electrical controllers.	7,52%
	Total	100

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

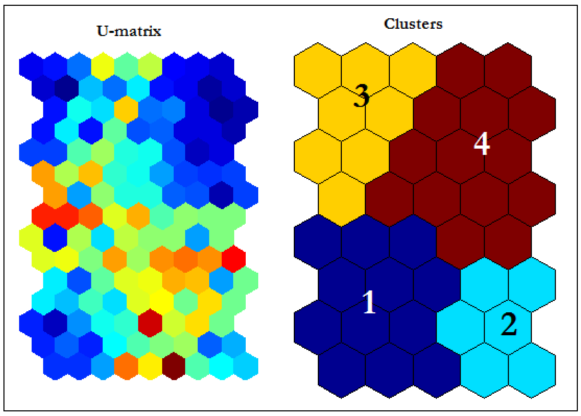


Source: own elaboration on MatLab.

Table 5- Cluster identification for dishwashers.

Cluster #	Technology description	% patents share per cluster
Dishwashers		
2	Controllers and sensors	12,12%
3	New washing methods and dishwashers as a whole	24,24%
1	Components for energy management	22,73%
4	Other components (mechanical, electromechanical and chemical)	40,91%
	Total	100

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

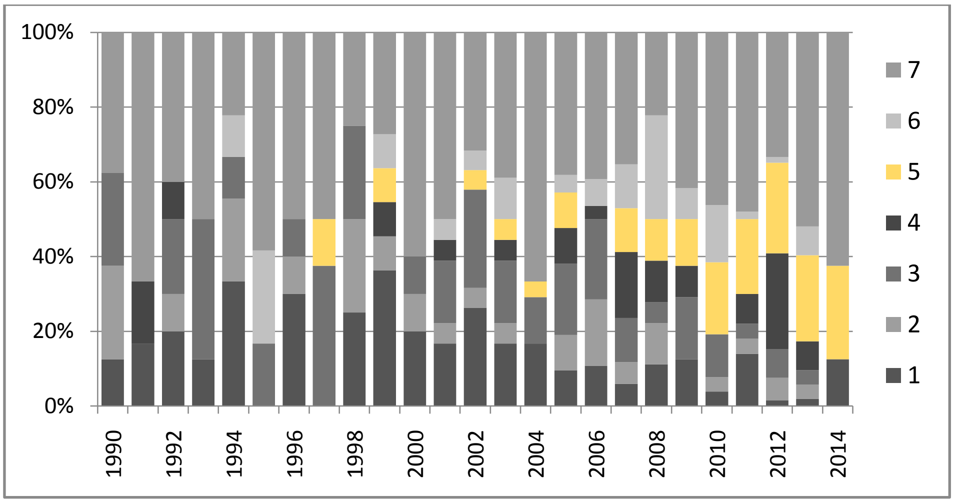


Source: own elaboration on MatLab.

5.1. ExperimentII – Static temporal comparison

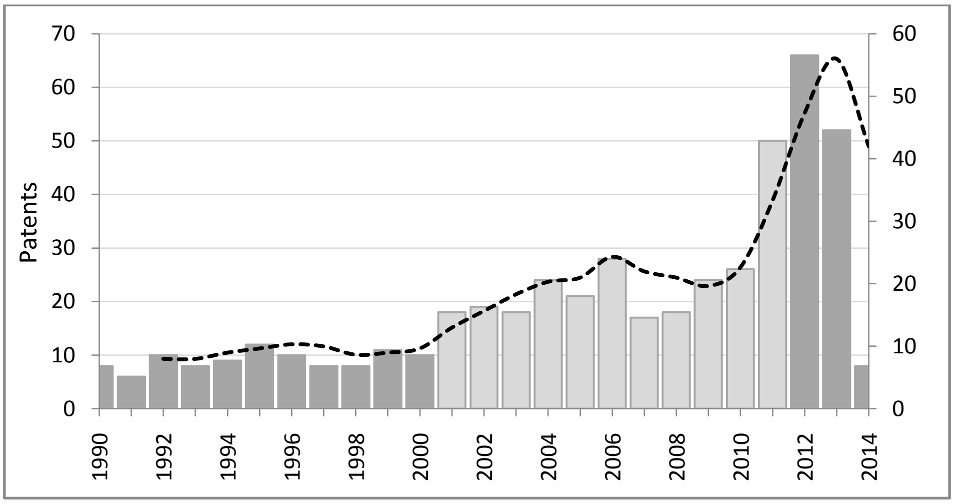
In the second experiment, we take into account only patents for refrigerators and freezers, dividing the sample in two sub-samples. The two sub-samples include the first and last 100 patents, sorted by publicationdate. It is worth noting that the distribution of the two samples is not homogeneous with respect to the time, since the first sample includes a longer period of time (1990-1999), while the first is the result of only two years of EE invention activity (2010-2014).

Figure 8 –Percentage of patents per cluster (1990-2014).[11]



Source: own elaboration.

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



Source: own elaboration.

Here,we moved to a more complex framework of technology variety, as showed in Figure 8, which provides a first evidence of growing technological complexity due to the different content of clusters that increases over time.

---

11 Clusters' number refer to Fig. 6.

This bearing in mind, by repeating the experiment for each temporal sub-sample, weproduce two temporal sections of the technological space of refrigerators which allow us to compare the evolution of technology niches in which EE nested.

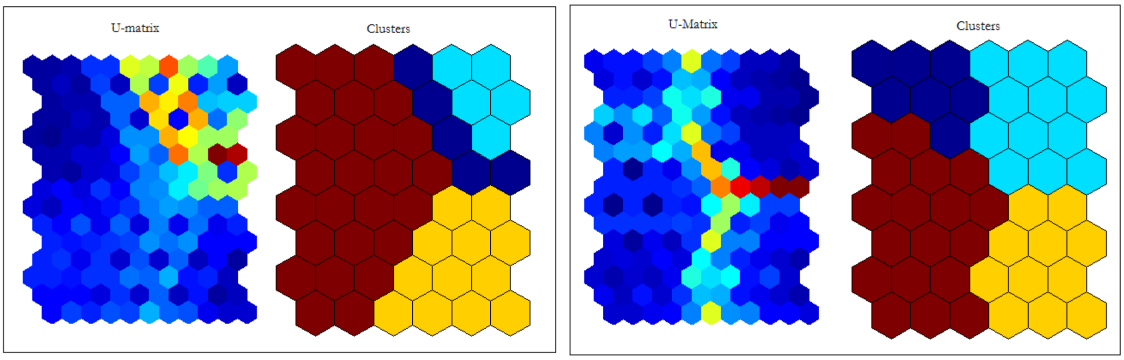
Figure 9 shows the histogram of the total distribution of fridge patents and the years interested by the two sub-samples (in dark grey). The number of new patents belonging toenergy efficient refrigerators and freezers is clearly skewed toward more recent years, signalling increasing innovative efforts made by firms for providing their appliances with more and more EE technologies. By looking at the patent maps (Figure 10), we note an equal number of clusters, but when these latter are under scrutiny, they unveil different technological content.

Table 6 - Cluster identification for static temporal comparison.

Refrigerators and Freezers					
1990-1999			2010-2014		
Cluster	Technology description	% patents share per cluster	Cluster	Technology description	% patents share per cluster
4	Mechanical and electromechanical components	67	4	Insulation materials, mechanical components and new appliances.	56
3	New appliances (as a whole)	26	3	Refrigerant compositions	17
2	Lubricant oils	4	2	Components for power supply management	19
1	Refrigerant compositions	3	1	Control devices	8

Figure 10 - Comparison of EE technologiesniches

in two different periods (1990-1999 and 2010-2014).



Source: own elaboration on MatLab.

Indeed, while in the first period the technological space is mainly affected by mechanical and electromechanical components such as more efficient compressors, lubricant oils and new compositions for refrigerants, in the second periodthe massive presence of digital components for energy management and motion control can be observed. The presence of such new EE niches lead to a more complex technological space, thus enhancing the level of technological recombination[12] and accruing the resource pooling effect. The result of this part of analysis provides a clear picture of technology evolution in which the technological space for EE evolved, showing a de-materialisation process from mechanical toward digital components and most likely improving the level of EE jointly operating by different technological contributions.

5.2. Experiment III – Technological fungibility

A further element to be taken into account when technological complexity grows, is the level of technological fungibility. The concept of fungibility applies to technologies having applications and usefulness "to a great array of new products and processes" (Antonelli, 2003, page 598). In the case of electrical appliances, the identification of technological fungibility deserves further investigation, also considering the previous results, which demonstrates that the level of technological complexity matters when EE aims are under scrutiny. Moreover, as far as technological fungibility is concerned, some economic implications arise as lower production costs due to the joint use of appliance's components as well as gains in production efficiency through scope economies (Panzar and Willig, 1981).

In light of this, the third experiment investigates the presence of technological fungibility between two sets of different appliances analysing the joint-use relationships of their EE technological components.

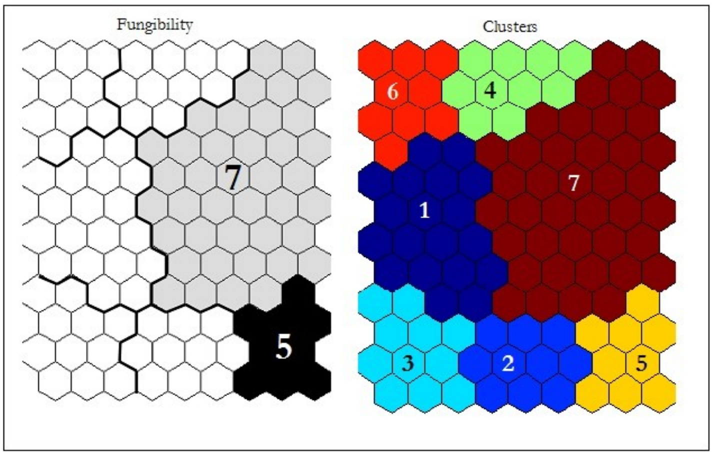
Departing from the patents sample, we used the co-occurrence analysis of CPC classes for marking those patents as "horizontal", that is patents classified as employable both in i) refrigerators and washing machines; ii) washing machines and dish washers. Then, we exploited the results of Experiment I for keeping track, in the SOM outputs, of each patent within nodes and clusters. By doing this, it was possible to detect not only multi-appliance patents (that is patents employed in different appliances), but also the technological clusters to which those patents refer to. As a further step, the percentage of multi-appliance patents has been calculated for each cluster. Lastly, we used the SOMs' outputs for producing new K-means clustering maps, choosing a black-white colour scheme, in which the percentage of grey mirrors the percentage of multi-appliance patents in each cluster, as showed in Figure 10 and 11.

By looking at the figures, it is possible to clearly unveil where the technological fungibility of EE of appliance's components is nested. Specifically, both in the case of refrigerator vs. washing machines as well as in the case of washing machines and 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. This result is clearly related to the role of ICTs as general purpose technologies (GPT), that is a technology characterised by general applicability and technological dynamism (see Bresnahan and Trajtenberg, 1995; Rosenberg and Trajtenberg, 2004 and Bresnahan, 2010,among others. As a matter of fact, ''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, page 598; Antonelli 1992).

---

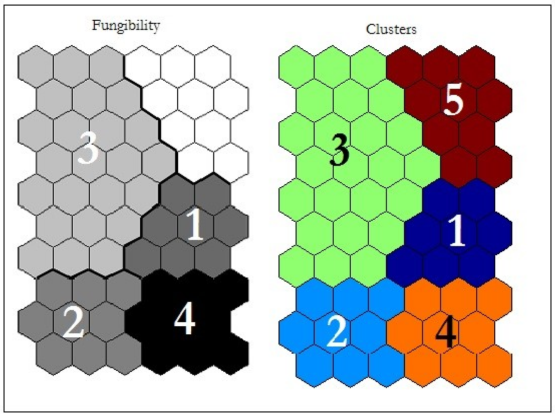
12 This latter can operate both vertically (diachronic recombination) and horizontally (synchronic recombination). The first refers to recombination of past elements of knowledge, while the second exploits contemporary acquisition of new bits of knowledge (Antonelli, 1999).

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



Source: own elaboration on MatLab.

Figure 11 - Technological fungibilityof 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 able to be set and employed in a wide range of applications (general applicability), and characterised by continuous improvements (technological dynamism) considering to the enormous growth of the ICTs in the last thirty years. In such a dynamics, EE well accommodates and strongly exploits the interchangeable technological space of domestic electrical appliances.

6. Conclusions

The present work uses an original patents dataset on EE technologies to develop, by the means of SOMs, an analysis model in order to test a set of theoretical hypothesis. Such hypothesis refer to: i) the pervasivity of EE in different technological niches; ii) the presence of technological variety and resource pooling effect as a result of growing technological complexity and technologies evolution over time; iii) the fungibility of EE technologies in different appliances. In order to test the previous hypotheses, a set of three experiments has been implemented, analysing patents with EE implications in four large electrical appliances (refrigerators and freezers, washing machines and dishwashers). The choice of this technological domain allows for exploiting the growing theoretical and empirical literature contributions on eco-innovation.

We use the well-known concept of technological relatedness building a vector of CPC class co-occurrences for each patent and then using such a vector as an input for further analysis. By using SOMs, highly multidimensional data deriving from the association of different CPC classes have been reduced to bi-dimensional maps in which the patterns of different technological niches clearly emerges. As a further identification process, a K-means clustering method has been also applied to SOM outputs, producing clearer maps of EE technological clusters.

The first experiment is devoted to find out technology niches in which EE is nested. This experiment is thus repeated for each of the three appliances, providing a clear picture of the pervasive nesting of EE in different appliance's components such as mechanical, electro-mechanical, digital, chemical as well as in operational processes. Our results confirm that EE is affected by pervasivity and tends to be nested in many technological niches, thus admitting the hypothesis i). By mapping also the number of nodes belonging to each cluster, a measure of innovation effort in each particular technological niche has been also derived, making 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 2010, such innovative efforts seem to be concentrated in insulation devices and mechanical components, without exceptions among the three appliances.

In the second experiment, we take into account only patents for refrigerators and freezers, dividing the sample in two temporal sub-samples. The two sub-samples include the first and last one hundred patents, sorted by priority date. Here, we found higher levels of complexity when moving from the first to the second sub-sample, along with a strong presence of the resource pooling effect due to the growing technological variety. The evolution of EE technology niches shows a de-materialisation process, initially characterised by the almost complete presence of mechanical components and, over time, moving toward an increasing complexity dominated by the massive presence of digital components.

Considering such a high level of complexity in the technological space of energy efficient electrical appliances, the third experiment is devoted to investigate the hypothesis of technological fungibility, finding out technology niches characterised by horizontal usefulness particularly evident between two different couples of appliances, namely between refrigerators-freezers and washing machines, and between washing machines and dishwashers. As a result, both in the case of refrigerators-freezers/washing machines as well as in the case of washing machines/dishwashers, we identified a single cluster that includes patents for energy management and digital motion controlling. This technological niche, characterised by the highest level of fungible components, belongs to the domain of information and communication technologies, defined by many authors as a general purpose technology, that is a technology showing general applicability and technological dynamism. In light of this, we conclude that the technological niche including the set of components referring to power management and digital motion controllers represents an interesting case of technological fungibility when domestic electrical appliances are under scrutiny. This latter, not only may implyincreasing returns due to lower replicability costs, but it seems particularly able to embrace the aim of EE.

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 Fungeability: 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., CostantiniV., 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 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.

Consoli D., Mina A. (2009). An evolutionary perspective on health innovation systems. Journal of Evolutionary Economics, 19 (2), pp. 297-319.

Costantini V., Crespi F., Orsatti G., Palma A. (2014). Policy Inducement Effects in Energy Efficiency Technologies. An Empirical Analysis on the Residential Sector. In *Green Energy and Efficiency. An Economic Perspective*. Eds. Ansuategi A., Delgado J., Galarraga. I. Springer (in press).

Davies D. L., BouldinD. W. (1979). A cluster separation measure, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 1, 1979, pp. 224-227.

EC (2011). Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions. Energy Efficiency Plan 2011.Com. (2011) 109.Technical report, European Commission.

Ester, M., Kriegel, H. P., Sander, J., Xu, X. (1996). A density-based algorithm fordiscovering clusters in large spatial databases with noise. In Proceedings of thesecond 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. EnergyJournal, 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: are view. Technological Forecasting and Social Change, 68, pp. 223-252.

Kuhn T. (1962). The structure of scientific revolutions. University of Chicago Press, Chicago.

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., ModyA. (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 (29), 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 (98), 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.

Mac Queen 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.

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.

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

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

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.

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.

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'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.

Verspagen B. (2007). Mapping technological trajectories as patent citation networks: A study on the history of fuel cell research. Advances in Complex Systems, 10(01), pp. 93-115.

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 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

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	energy sav* 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		washing machine*
Dishwashers		dishwash*