Economics

An alternative approach to frequency of patent technology codes: the case of renewable energy generation

--Manuscript Draft--

Manuscript Number:	ECONJOURNAL-D-22-00106
Full Title:	An alternative approach to frequency of patent technology codes: the case of renewable energy generation
Article Type:	Research Article
Keywords:	Technological transitions; renewable energy generation; total variation distance; patent classification
Manuscript Region of Origin:	MEXICO
Abstract:	This paper proposes a methodology to detect Technological Transitions (TTs) by systematically using the Total Variation Distance (TVD) metric. We use a database of Renewable Energy Generation (REG) patents to exemplify the usefulness of TVD to uncover moments where a ‘big change’ in REG technology happened. To do this, we compare the observed frequency distribution of technology codes of REG patents filed between 1973 and 2015 in the US, spread across seven categories (e.g., wind and tidal). We identify two crucial TTs, one at the beginning of the 1980s and another in the late 1990s and early 2000s. In this manner, we reconcile qualitative evidence that registers major REG changes with a quantitative measure that reflects them. Policy evaluations or causality analyses often rely on identifying TTs accurately; therefore, this approach is not constrained to the REG technology or TTs but helps reveal such transition moments in a database whose characteristics are suitable for the use of TVD.
Manuscript Classifications:	3.4: Econometric and Statistical Methods and Methodology: General; 15.3: Innovation • Research and Development • Technological Change • Intellectual Property Rights

An alternative approach to frequency of patent technology codes: the case of renewable energy generation

November 2022

Abstract

This paper proposes a methodology to detect Technological Transitions (TTs) by systematically using the Total Variation Distance (TVD) metric. We use a database of Renewable Energy Generation (REG) patents to exemplify the usefulness of TVD to uncover moments where a 'big change' in REG technology happened. To do this, we compare the observed frequency distribution of technology codes of REG patents filed between 1973 and 2015 in the US, spread across seven categories (e.g., wind and tidal). We identify two crucial TTs, one at the beginning of the 1980s and another in the late 1990s and early 2000s. In this manner, we reconcile qualitative evidence that registers major REG changes with a quantitative measure that reflects them. Policy evaluations or causality analyses often rely on identifying TTs accurately; therefore, this approach is not constrained to the REG technology or TTs but helps reveal such transition moments in a database whose characteristics are suitable for the use of TVD.

Keywords: Technological transitions, renewable energy generation, total variation distance, patent classification.

1 Introduction

Technological change is one of the leading forces to explain economic growth and development. Technological changes happen continuously, and some moments during a time span are especially relevant since those changes are abrupt. As Perez-Molina and Loizides (2021) state, different parties, such as researchers or private firms, continuously look for tools to understand a specific technology's evolution and pinpoint moments of particular interest within it. In this context, we are interested in identifying technological transitions (TTs), which we refer to as those moments when the technology in a particular field exhibits a notable change.

TTs are complex phenomena with effects on innovative processes, organization of production factors, products, and new ideas. Not all these innovations can be legally protected, or owners may decide not to do so even if they can. Despite this limitation, patents serve as a proxy for successful innovation processes and allow us to use analytical approaches to study the complex dynamics behind them (Alkemade et al. 2015). Patent documents have been widely used in the academic literature since they provide valuable information about the state of a given technology at a given time. Specifically, patent classification has been a primordial instrument for creating a useful taxonomy that can be used by those interested in protecting and analyzing intellectual property (e.g. Ruijie et al. 2021, Meguro and Osabe 2019). Therefore, patents' technology codes can be used to identify technological novelty (Lobo et al. 2012).

This paper proposes a methodology to systematically identify TTs in a particular technological field using the distance between probability distributions among the technology codes in that field. We characterize a technological state as the probabilistic empirical density distribution over the technological patent categories, using the Cooperative Patent Classification (CPC), at a given time. Consequently, we approach technological change by measuring the distance between these distributions. As discussed in Frenken et al.

(2014), we conceive a TT as a tipping point in which the dominant technological state changes to an alternative one within a given system. Therefore, we claim that a disruptive change in the distribution of categories can be used as a signal of a TT.

To the best of our knowledge, the methodology of intertemporally measuring the distance between distributions within a given technology has not been used to identify historical moments where a notable technological change may have occurred. We measure it with the Total Variation Distance (TVD) metric based on its properties and explain why we use local maximums to identify TTs (Section 3). It is worth mentioning that there is literature on measure theory and its applications, mainly in computing and biology, and separately there is literature on technological transitions; however, it is almost non-existent the one that combines both.

Furthermore, our proposed methodology is easily executable and straightforwardly interpretable since it only uses the relative frequency of each category at a given time as input; thus, it is particularly valuable for exploratory analysis, providing a visualization tool to identify TTs. We apply this methodology to the REG patents case, which shows suitable data characteristics (Section 2), and we identify two TTs consistent with the REG historical evidence. Hence, the interest of the present paper is not to explain the drivers behind a TT or mathematically model them but to offer an alternative methodology to a systematic exploratory analysis of TTs identification.

The present paper is organized as follows. Section 2 compares the methodology to the existing literature and identification of TTs. Section 3 introduces the TVD and its properties and explains its calculus. In Section 4, we present the REG case and data set, for which we calculate the TVD under different specifications and discuss the interpretation of the identified two moments of TT. We conclude with Section 5.

2 Related literature

One way to study the evolution of a particular field of knowledge is through research topics (categories), which can be classified endogenously or exogenously. A classification is endogenous when its categories are not ex-ante defined but are the result of applying a methodology to come up with them, such as keywords, memes, bibliographic coupling, or co-citation (e.g. Behrouzi et al. 2020, Kuhn et al. 2014, Chang et al. 2015). Although an endogenous classification may be advantageous, particularly in the absence of a meaningful alternative, it might not be as effective as the classification carried out by experts.

A patent is classified in all relevant technology codes where the novelty fits, and these codes constitute a classification scheme (Lobo and Strumsky 2019). Thus, a critical task of patent offices worldwide is to create a taxonomy of technology codes able to organize and classify the technical information of patents (Lobo and Strumsky 2019) and, when required, to introduce new codes to capture the novelty embraced in patents. When the latter happens, offices retroactively reclassify all previous patents that may be affected (Lobo et al. 2012). Once the challenge of establishing a classification has been overcome (Makarov 2004), this can be used to uncover trends or statistically identify important events that shifted the innovative efforts using this information.

Each category within a classification entails a collection of patents that should be similar to each other, not necessarily identical, but still sufficiently different from patents belonging to other categories. Perez-Molina and Loizides (2021) explain that for an individual patent, all its classification codes are proxies of its technological components; thus, the evolution of these codes conveys information about enhanced interest in a technical area or loss of it in a particular moment.

Schilling and Green (2011) show that vector-based distances on the taxonomy of patents concerning a patent classification system are an appropriate approach to measure patents similarity. However, this comparison has been limited to the individual level by classification or patent. Likewise, given a technological field determined by patent classes, its subsets' relative composition should indicate its current state since all relevant classes within a specific technology should be accounted for and should evolve when needed due to technological changes (Perez-Molina and Loizides 2021).

TTs have been identified by focusing on the increase of a particular subgroup of patents with respect to the total. Angelucci et al. (2012) use this approach to identify the effect of the Kyoto Agreement on the participation of two Climate Change and Mitigation Technologies (CCMTS) patent subclasses; however, their methodology requires an ex-ante identification of the patent subgroups that experiences the frequency increase. It is worth noting that we arrived at the same conclusion concerning identifying this particular TT, although we do not identify trends within specific patent subgroups but rather as a technology state. Unlike the approach used by Angelucci et al. (2012), if the frequency of the aggregated patent subgroups remained constant, but its composition changed, our methodology would be able to signal a potential TT.

Aminikhanghahi et al. (2018) show that distances between distributions can be used to detect tipping points, which they applied to high dimensional time series using the Separation distance. However, not only is the Separation distance non-symmetric, which limits the pairwise comparability, but it is also not sensitive to changes along the distribution. Therefore, if one wants to include the variations in all the categories of a distribution as a measure of a technological change in a given field, it is preferable to use metrics such as the Total Variation Distance or the Hellinger distance.

Other dynamic characteristics of technological change have been studied using entropy measures. For instance, Lin et al. (2021) use it to study technology life cycle (TLC) stages using an entropy measure on applicants for patents. Different patent indicators have been related to different phases that constitute the S-curve of a TLC (Gao et al. 2013). These methodologies can be grouped within the technology intelligence literature and focus on single technologies and the analysis of competitors instead of identifying TTs in a broader technological field (see Guderian 2019). Nevertheless, identifying TTs can complement the analysis offered by technology intelligence, which helps to point out emerging firms and countries within specific technologies.

Benner and Waldfogel (2008) demonstrate that patent-related measures are negatively affected by small sample sizes, and it is convenient to use large sample sizes and coarse patent classes. To avoid potential issues of deficient patent classification, we use the *Y-section* of the CPC, and on the sample size, we address it through the use of the complete database available in the USPTO of the relevant patents (Section 4.2).

In mathematics, a distance is a quantitative measure of how far apart two objects are, and measuring the distance between histograms has proven paramount in various areas to perform pattern recognition (Cha and Srihari 2002, Kurtz et al. 2013, Strelkov 2008). Histograms are structures to model data and study its statistical properties. There exist different types of histograms (nominal or ordinal) depending on the data nature; an ordinal one contains totally ordered values while a nominal one contains categorical values (Kurtz et al. 2013). If we consider a nominal histogram, which is our case, it represents the distribution of quantified labeled values (categories) where each bin contains the proportion of a category out of all categories, and the ordering is not important (shuffling invariance property)[1] (Sung-Hyuk 2007, Kurtz et al. 2013). Therefore, identifying the proper nature of the data is fundamental to applying statistical analysis.

The statistical comparison of two histograms provides a complementary tool over a merely visual comparison. As Strelkov (2008) points out, humans compare multiple histograms relatively slowly, and a comparison algorithm

[1]Nominal data distribution is invariant to "shuffling" the categories labels, and this property is not obvious for continuous variables(Duda et al. 2007).

is preferable; however, for such an algorithm to be reliable, it needs to reach conclusions close to those reached by experts using distinct methods. There are two types of histogram distance measures (vector and probabilistic). Cha and Srihari (2002) explain that, in the probabilistic distance approach, a histogram provides the empirical estimate of the probability density function (pdf) and compares corresponding and non-corresponding bins[2]. Thus, at a given year, the histogram that portrays the relative frequency of patents distributed in the different REG categories intrinsically provides information on the innovative interests at that time. Consequently, measuring the probabilistic distance between two nominal histograms of two points in time provides valuable information on how close or apart the inventive efforts were.

TVD is a probabilistic measure to quantify the difference between two pdf; it measures their similarity. The use of TVD has been widely used in several disciplines, but to the best of our knowledge, not in the patent literature or intertemporally as we propose. Examples of TVD usefulness are found in genomics (Garcia and Pinho 2011), criminology (Carte et al. 2020), epidemic models (Ball and Donnelly 1995) biology of sleep (Barger et al. 2019) and, more generally, time series with similar oscillations (Euán et al. 2018).

3 Methodology

3.1 Total Variation Distance

We explain below the TVD, its properties, and why it is preferable in our case compared to other probabilistic distances. If the sample space is discrete and categorical, as in our REG case, the TVD measures the difference of histograms by categories (Section 4).

[2] Cha and Srihari (2002) explain that a vector measure treats a histogram as a fixed-dimensional vector where standard vector norms measure the distance between two vectors, e.g., Euclidian distance.

TVD is a probabilistic measure to quantify the difference between two pdf; it measures their similarity. The use of TVD has been widely used in several disciplines, but to the best of our knowledge, not in the patent literature or intertemporally as we propose. Examples of TVD usefulness are found in genomics (Garcia and Pinho 2011), criminology (Carte et al. 2020), epidemic models (Ball and Donnelly 1995) biology of sleep (Barger et al. 2019) and, more generally, time series with similar oscillations (Euán et al. 2018).

Assume a nonempty set $\Omega \subset \mathbb{R}$, which is the sample space, and let $\mathbb{P}(\Omega)$ be the set of probability distribution on Ω. Suppose that μ and η are in $\mathbb{P}(\Omega)$, the *total variation distance* $d(\mu, \eta)$ between μ and η is given by the function

$$d(\mu,\eta) := \sup \left\{ \left| \int_\Omega f(s)\mu(ds) - \int_\Omega f(s)\eta(ds) \right| : \quad \|f\| \leq 1 \right\} \tag{1}$$

where the supremum is taken over the set of real-valued bounded measurable functions on Ω, with the norm

$$\|f\| := \sup_{s \in \Omega} |f(s)|. \tag{2}$$

In probability theory, there are different ways to estimate the distance between two probability distributions. A few of these are TVD, Hellinger distance, Kullback-Leibler divergence, Wasserstein distance, Prokhorov distance, among others (see Massart 2007, for further discussion). The first two metrize strong topology, and the last two metrize weak convergence; Kullback-Leibler divergence is not a distance, but its proximity implies strong and weak convergence. Also, Lin et al. (2021) use the Kullback-Leibler divergence to identify technology life cycles.

We select TVD because:

- It is not sentient to the distance $|\omega - \omega'|$ between two any events $\omega, \omega' \in \Omega$. The distances that metrize weak topology, as Wasserstein and Prokhorov, are sentient to the distance $|\omega - \omega'|$ between two any events $\omega, \omega' \in \Omega$. Therefore, the TVD is not susceptible to changes in the order of categories and is suitable for unordered categories.

- It has properties similar to a norm in vector spaces (see Appendix A); in particular, it is homogeneous. Thus, for probabilities μ, η and the linear combination $\nu = \alpha\mu + (1-\alpha)\eta$ (where $\alpha \in [0,1]$), we have that $d(\mu,\nu) = (1-\alpha)d(\mu,\eta)$ and $d(\eta,\nu) = \alpha d(\eta,\mu)$, which means that if we consider an average of two distributions, the distance between the resulting probability and one of the original ones is proportional to the original distance. Weak topology metrics do not satisfy this property.

- Its calculation is more straightforward compared to other strong topology metrics, such as the Hellinger distance, and its interpretation is more intuitive.

- Since it is a distance, it allows us to measure the difference between several probability distributions, in contrast to the Kullback-Leibler divergence.[3]

- It can be calculated for probability distributions at which the probability of an event is 0, in contrast with the Jensen-Shannon divergence.

- It is sensitive to changes in the distribution concerning all the categories, as opposed to the Separation distance.

Due to the reasons above, we find TVD a suitable measure for qualitative categories where the order is not relevant, as is the case with patent

[3]The Kullback-Leibler divergence measures how a probability distribution is different from a second reference probability distribution. In this sense, if we fix one probability distribution, we can measure how the fixed distribution differs from others, which is not the purpose of this paper. We want to measure the difference between various distributions and be able to compare these without necessarily fixing one, so the Kullback-Leibler divergence is inappropriate for our purposes.

technology codes, and it allows for pairwise comparison between different distributions without keeping the same base.

3.2 TTs identification methodology

In the case where Ω is a finite set, (1) has a simpler expression (see Appendix A). Let $\Omega = \{\omega_1, ..., \omega_k\}$ be the set of patent subgroups and $\mu_t(\omega_i)$ be relative frequency of the subgroup $\omega_i \in \Omega$ at period t, where $\sum_{i=1}^{k} \mu_t(\omega_i) = 1$. Then, the relative frequencies distribution on patent subgroups, $\mu_t := \{\mu_t(\omega_i)\}_{i=1}^{k}$, is a technological state at period t. The TVD between the two discrete distributions of periods s and t, $d(\mu_s, \mu_t)$, is given by

$$d(\mu_s, \mu_t) = \sum_{i=1}^{k} |\mu_s(\omega_i) - \mu_t(\omega_i)|. \tag{3}$$

If (3) is strictly positive, then we have a technological state (or technological change) change between periods s and t.

While a distance only allows for pairwise comparison for a particular order of periods (r, s, t), if we have that

$$d(\mu_r, \mu_s) > d(\mu_s, \mu_t) > 0, \tag{4}$$

that indicates a greater technological change between r and s as compared to s and t. If this technological change is significant concerning other periods, we identify it as a potential TT.

Notice that TVD is invariant to the number of observations at each period since it depends on the relative distribution. However, we must acknowledge that the variability between the realization of patents and the actual technology distribution may occur with the number of observations and be affected by small samples. To avoid misguided interpretations, particularly related to a decrease of the TVD related to an increase in the number of observations, we consider two types of periods:

(a) *Natural Periods*, in which each observation is partitioned into periods according to the patent registration date.

(b) *Periods with the Same Number of Observations* (PSNO), in which the patent filing date, alone, orders the data. We partition the data into subsets with the same number of observations. Each subset is called a 'period'.

Thus, we use the TVD metric to distinguish a moment where abrupt changes arise in a given distribution of categories. Since TVD is used for probabilistic distributions, we apply it to the relative distribution of patents among REG seven subgroups used as categories in the TVD implementation (Section 4). TVD is not sensitive to category reordering, which is convenient in our setting because REG subgroups could be reordered.

TVD may vary to a certain degree in the absence of a TT; such variation depends on the nature of the technological field and moment, as well as the total number of observations in each period. Therefore, in absolute terms, we cannot claim any particular TVD level to be indicative of a TT. Instead, we focus on local maximums specific to the technological field under study and show a greater distance for both directions in time.

The application of the TVD as an identification strategy of TTs requires an adequate identification of the technological field to contain emergent technological innovations. A narrow field selection could exclude the potential emergent technological innovations. In our example, the sources from which REG takes their primary sources (e.g., wind, water, or sun) have not changed in at least the last century, and we can reasonably argue that those categories are fixed in our period of analysis. Suppose a new category appears at a particular moment in time. In that circumstance, we can either reorganize the previous observations according to the new categorization, assume that none of the previous (posterior) observations belong to the newly created (eliminated) category, or introduce an interpolation.

4 The REG case

4.1 The Cooperative Patent Classification

Much has been discussed on how patent offices prepare to tackle a sudden increase in patents due to the rapid development of emerging technologies. For instance, evidence suggests that public and private investment in R&D in nanotechnology led to a significant rise in related patent applications (Absalom et al. 2006, Angelucci et al. 2012). Absalom et al. (2006) mention that the European Patent Office (EPO) developed a new taxonomy in the early stages of nanotechnology R&D to ensure technical expertise within examiners and legal certainty to innovators. In this way, patents contain information about the birth and development of technologies.

One difficulty of reliably identifying whether a TT has occurred is to ensure consistent patent classifications throughout the period of study (Lacasa et al. 2003). On October 25, 2010, the EPO and the United States Patent and Trademark Office (USPTO) announced the CPC as a joint effort aimed at establishing a harmonized classification system for patent documents; the CPC is based on the International Patent Classification (IPC) system, managed by the World Intellectual Property Organization (WIPO)[4].

The CPC is divided into nine sections (A-H and Y), where each section is divided into classes, sub-classes, groups, and subgroups and contains approximately 250,000 classification entries [5]. In our illustrative case, we use patents with at least one Current CPC Class within the subgroups of the *Y02E 10* group "Energy generation through renewable energy sources"[6] (Section 4).

[4]https://www.uspto.gov/about-us/news-updates/uspto-and-epo-work-toward-joint-patent-classification-system

[5]https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/classification/cpc.html

[6]It is worth noting that the CPC results from meticulous tagging work of various patent offices, ensuring that the *Y02E 10* subgroup contains only REG-related patents.

One important peculiarity of the *Y-section* is that the allocation of new patent applications to the relevant technology code(s) is done automatically by search strategies without affecting existing classifications outside it; thus, the classifiers intervention is constrained to developing and updating those strategies (EPO and USPTO 2015). Furthermore, there has been a dedicated effort to reclassify all applicable patents retroactively. The *Y-section* results from meticulous and continuous tagging work of various patent offices, ensuring that the *Y02E 10* subgroups contain only REG-related patents.

In this manner, we take advantage of the categorical property of the subgroups within the *Y02E 10* group "Energy generation through renewable energy sources" of the Current CPC Class, which contains only REG-related patents. Thus, each of the seven subgroups of the *Y02E 10* group, also called categories, belong to a different renewable energy source (e.g., wind or from the sea). We use categories that are known beforehand, where patents are sorted into seven separate categories. All patents classified under the "Wind" category are useful to generate energy through the wind but are sufficiently different from those "From the sea"[7].

Climate change mitigation technologies (CCMTs) involve several areas of knowledge. Thus, technologies using the existing alternative patent classifications can lead to an incomplete retrieval of documents as stated by the EPO "Y02 is a tagging scheme that enables documents relating to sustainable technologies to be retrieved quickly and accurately across classification categories"[8].

4.2 The data set

We illustrate the implementation of TDV using public patent data provided by the USPTO. In particular, we consider patents with at least one Cur-

[7] In our database, 93.58% of the patents belong to only one category, 5.72% to two categories, and the rest to three, four or five categories (0.7%).

[8] https://e-courses.epo.org/wbts/y02

rent CPC Class within the *Y02E 10* group of "Energy generation through renewable energy sources." The *Y02* class aims to identify CCMTs as stated in the Kyoto Agreement (Angelucci et al. 2012), which contains the *Y02E* subclass "Reduction of Greenhouse Gas (GHG) emissions, related to energy generation, transmission or distribution."

We searched all patents within each subgroup of *Y02E 10* (one, two, and three dots). For instance, in Geothermal energy, to one dot, we refer to a CPC subgroup of the form *Y02E 10/10* "Geothermal energy," to two dots, for *Y02E 10/12* "Earth coil heat exchangers," and three dots for *Y02E 10/125* "Compact tube assemblies, e.g., geothermal probes." The complete catalogue of CPC subgroups within the *Y02E 10* group is found in the USPTO website.[9]

Our database consists of 39,236 patents filed from 1973 to 2015. Although there is updated information on patent applications, we decided to exclude patents filed from 2016 onward from the analysis to avoid false TTs due to unobserved recent patent applications. Because according to Article 1122 of the Manual of Patent Examining Procedure, the USPTO allows for the possibility of requesting an application to be non-publishable until its acceptance, we may have not been able to observe a fair share of recent applications publicly. Considering patents granted until January 2020,[10] the mean lag between the issuance and filing date is 2.9 years with a standard deviation of 1.5 years. Then, we decided to exclude the patents corresponding to a period close to the mean lag plus a standard deviation[11]. Notice that the mean lag in our database is close to the 2.4 years obtained by Popp et al. (2004).

A patent could have more than one Current CPC Class, but we analyze only those patents recorded with at least one of the REG categories (Table

[9]https://www.uspto.gov/web/patents/classification/cpc/html/cpc-Y02E.html#Y02E

[10]January 31, 2020, is the last date we downloaded information from the USPTO website.

[11]Since we do not find any signal of a TT around the last years, increasing or decreasing one year of data does not affect the results.

1). When all subgroups within REG belong to only one subgroup, we assign the patent to that category. However, when the subgroups belong to two categories, the patent appears in two categories; the same principle applies to patents with three or more subgroups.

In Table 1, we show the seven REG categories within the *Y02E 10* group and the complete list of Current CPC Classes within each one.

	1	2	3	4	5	6	7
	Geothermal	Hydro	From the sea	Solar thermal	PV	Thermal-PV	Wind
	Y02E 10/10	Y02E 10/20	Y02E 10/30	Y02E 10/40	Y02E 10/50	Y02E 10/60	Y02E 10/70
	Y02E 10/12	Y02E 10/22	Y02E 10/32	Y02E 10/41	Y02E 10/52		Y02E 10/72
	Y02E 10/125	Y02E 10/223	Y02E 10/34	Y02E 10/42	Y02E 10/54		Y02E 10/721
	Y02E 10/14	Y02E 10/226	Y02E 10/36	Y02E 10/43	Y02E 10/541		Y02E 10/722
	Y02E 10/16		Y02E 10/38	Y02E 10/44	Y02E 10/542		Y02E 10/723
	Y02E 10/18			Y02E 10/45	Y02E 10/543		Y02E 10/725
				Y02E 10/46	Y02E 10/544		Y02E 10/726
				Y02E 10/465	Y02E 10/545		Y02E 10/727
				Y02E 10/47	Y02E 10/546		Y02E 10/728
					Y02E 10/547		Y02E 10/74
					Y02E 10/548		Y02E 10/76
					Y02E 10/549		Y02E 10/763
					Y02E 10/56		Y02E 10/766
					Y02E 10/563		
					Y02E 10/566		
					Y02E 10/58		
N	*770*	*2,253*	*1,671*	*8,033*	*20,055*	*246*	*9,009*

Table 1: Current CPC Classes considered in each REG category

Figure 1 shows the number of patents by REG category depending on the year they were filed. We observe that the mix of REG sources (subgroups) has changed since the 1970s. During the 1970s, the predominant source was solar thermal, followed by PV. In the early 1980s, PV patents surpassed solar thermal ones, and these two sources remained relevant for the next two decades. At the beginning of the 2000s, the wind source dramatically increased, exceeding solar thermal but remained below that of PV, which continues its rising trend. Since 2011, a general decline in patenting is observed. This phenomenon may be related to the maturity of some REG technologies, such as the PV and wind (International Energy Agency 2020).

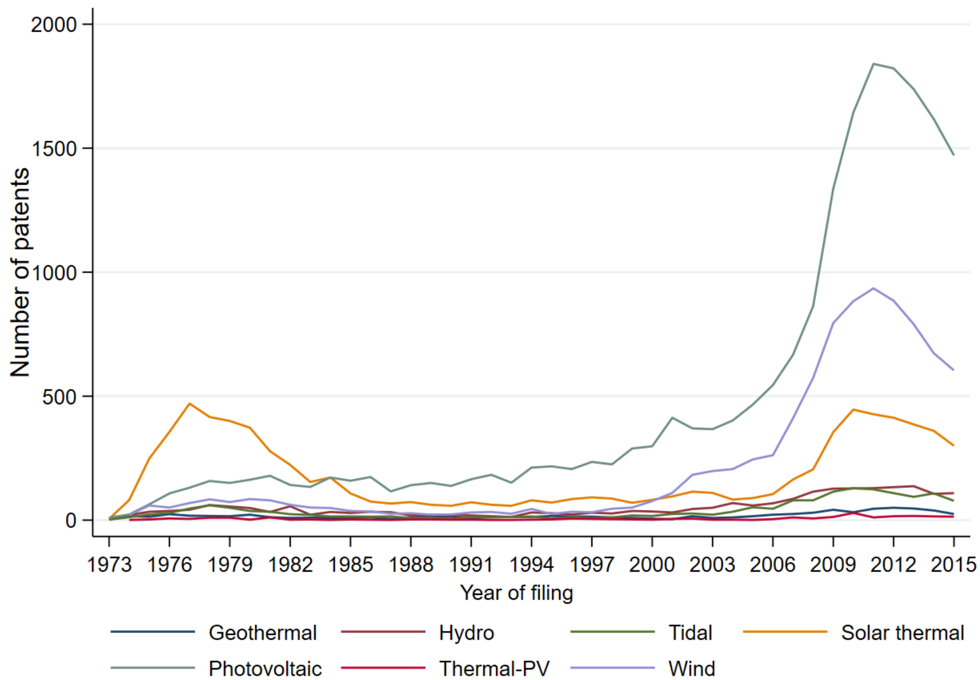


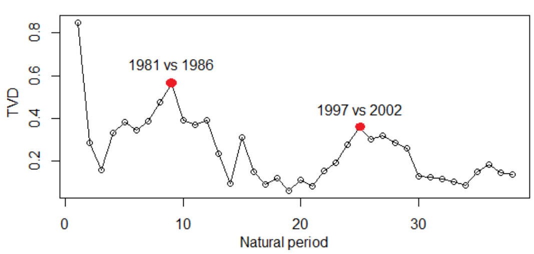
Figure 1: Number of patents filed by REG category

4.3 TTs identification

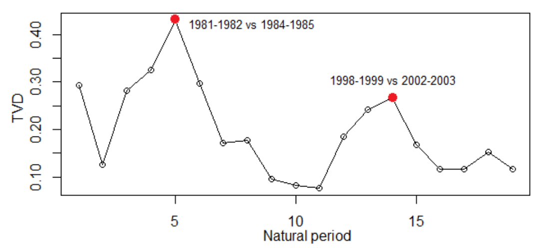
To use TVD, we need the number of periods within the whole sample and the distance between them. Thus, we define the periods using two criteria, natural periods and PSNO, which we describe below.

We first divide our data by natural years, each constituting a "natural period;" then, we select the distance of comparison between natural periods. The distributions may look very similar when periods are too short and distances too close. Therefore, shorter periods may need further distance calculations. We test a variety of combinations of years and periods of distance. Interestingly, we obtain high consistency in defining two TTs moments in the REG case: one in the 1980s and another between the end of the 1990s and the beginning of the 2000s.

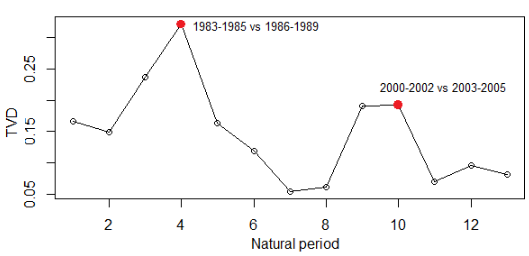
In Figures 2, we show different year-windows and distance between them. If we take one-year periods and calculate the distances between 5 periods, we identify two peaks (Figure 2a). The first peak compares the distribution in 1981 with the one in 1986. The second one signals a peak in the distance between the distribution in 1997 and 2002 and compares them. The graph with two-year periods and comparisons of a distance of 2 (Figure 2b) exhibits more clearly the two peaks: the first one compares the distribution of 1981-82 with 1984-85; the second one compares the 1998-99 distribution with 2002-03. In Figure 2c, again, we observe the two peaks in the three-year periods with consecutive (distance 1) comparisons: 1983-85 with 1986-89 and 2000-02 with 2003-05. The five-year periods with consecutive comparison, Figure 2d, gives a similar picture, with a first peak at 1978-82 comparing with 1983-87, and the second comparing 1998-2002 to 2003-2007.



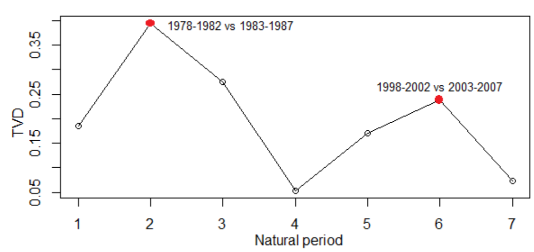
(a) One-year periods with five-period distance



(b) Two-year periods with two-period distance



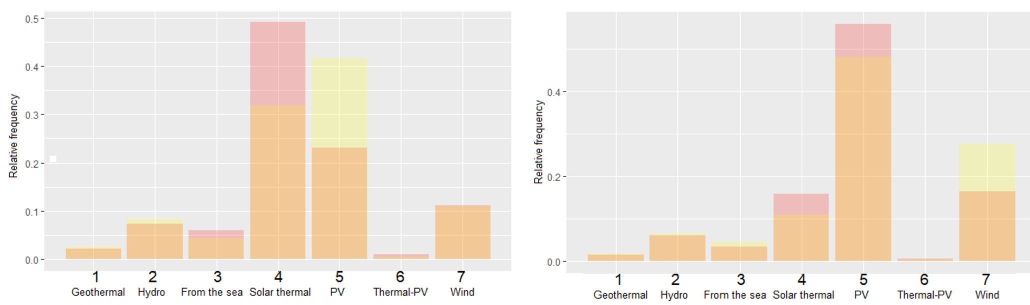
(c) Three-year periods consecutive



(d) Five-year periods consecutive

Figure 2: TVD for Natural Periods

For illustration purposes, we show in Figure 3 the histograms for five-year natural periods comparing 1978-1982 and 1983-1987 (Figure 3a) and 1998-2002 and 2003-2007 (Figure 3b), which are associated to the two red points of Figure 2d. The most notorious change from 1978-1982 to 1983-1987 is a relative reduction of subgroup 4 (Solar-thermal) and an increase in subgroup 5 (PV). During the second peak, from 1998-2002 to 2003-2007, we observe a relative reduction of subgroups 4 and 5, which derive into an increase of subgroup 7 (Wind).



(a) Distributions during 1978-1982 (red) and 1983-1987 (yellow) and common (orange)

(b) Distributions during 1998-2002 (red) and 2003-2007 (yellow) and common (orange)

Figure 3: Distributions during TT for five-year natural periods

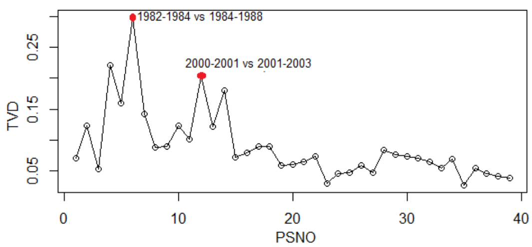
We repeat the exercise for Periods with the Same Number of Observations (PSNO), Figures 4. In this way, we correct for a variability bias due to a greater number of patents in recent years. Similar to the case of natural periods, a greater number of periods would mean shorter periods, which would ask for a comparison conceding a greater distance.

If we consider 40 PSNO and a consecutive comparison (Figure 4a), we find two notable peaks: one between June 1982 and March 1988 [12] and the second between March 2000 and April 2003.[13] For the same 40 PSNO

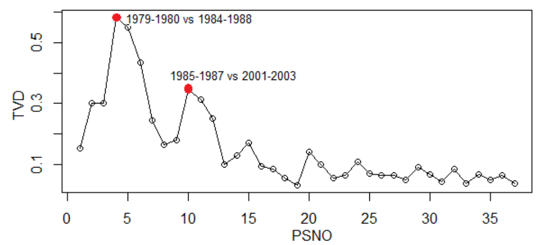
[12]The first distribution has patents filed between June 30, 1982, and December 13, 1984; the second one has patents filed between December 14, 1984, and March 28, 1988.

[13]The first distribution has patents filed between March 13, 2000, and November 28, 2001; the second one has patents filed between November 28, 2001, and April 21, 2003.

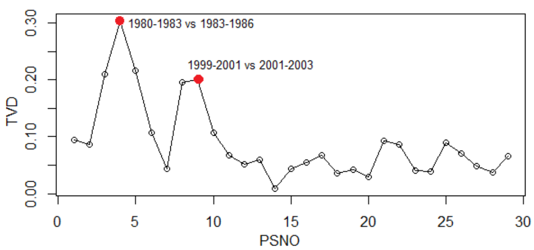
and a comparison of distance 3 (Figure 4b), we obtain the same two peaks: one between June 1979 and March 1988[14] and the second between January 1995 and April 2003[15]. For 30 PSNO and a consecutive comparison (Figure 4c), we obtain two main peaks between October 1980 and December 1986[16] and between June 1999 and October 2003[17]. For 20 PSNO (Figure 4d), the identification is less precise: one between October 1980 and December 1991[18] and the second between October 1997 and August 2004[19].



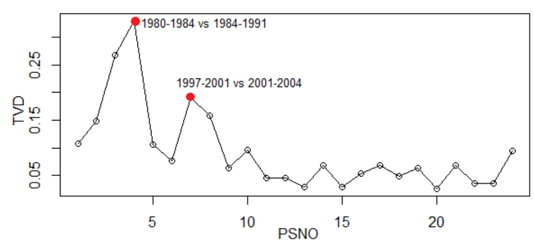
(a) 40 PSNO consecutive



(b) 40 PSNO with three-period distance



(c) 30 PSNO consecutive



(d) 40 PSNO consecutive

Figure 4: TVD for PSNO

[14]The first distribution has patents filed between June 11, 1979, and October 28, 1980; the second one has patents filed between December 14, 1984, and March 28, 1988.

[15]The first distribution has patents filed between January 31, 1995, and October 23, 1997; the second one has patents filed between November 28, 2001, and April 21, 2003.

[16]The first distribution has patents filed between October 31, 1980, and May 5, 1983; the second one has patents filed between May 6, 1983, and December 17, 1986.

[17]The first distribution has patents filed between June 10, 1999, and November 30, 2001; the second one has patents filed between December 3, 2001, and October 1, 2003.

[18]The first distribution has patents filed between October 29, 1980, and December 18, 1984; the second one has patents filed between December 19, 1984, and December 13, 1991.

[19]The first distribution has patents filed between October 30, 1997, and November 30, 2001; the second one has patents filed between November 30, 2001, and August 13, 2004.

In this Section, we show that TVD approach has consistently identified two TTs, despite considering a wide range of period sizes and comparison distances. Notice that since we have not considered any external information, the measure is ideal for an initial exploratory analysis and only uses the relative frequency of each category within REG field.

4.4 Interpretation

Our methodology systematically indicates two moments of notable change in the distribution of patents for REG. Although TTs could have occurred naturally, we aim to acknowledge some policies linked to the two moments identified through the TVD calculation, which qualitatively reinforce the correct identification of the TTs. The first peak in the early 1980s could result from the research that would have started in the late 1970s. Between the late 1990s and the beginning of the 2000s, the second peak could relate to Kyoto's Treaty. However, we do not pretend to establish any causality between the policies mentioned and the TTs.

The oil embargo of 1973 triggered interest in alternative energies, which motivated a wide range of energy conversion techniques from renewable sources. Among these, PV panels started to be considered a viable option for commercial use (Sørensen 1991). The oil crisis in the 1970s triggered public funding for R&D programs aimed to advance in PV generation, mainly from Japan and the USA, and it is calculated that around 60 percent of cost reduction in this technology was due to public and private R&D (International Energy Agency 2020). President Nixon's Project Independence in 1973, President Ford's Energy Policy and Conservation Act in 1975, and the Public Utility Regulatory Policies Act in 1978 also promoted renewables and have been related to technological changes that led to cost reductions in wind and solar PV (Solomon and Krishna 2011, Smith 2004, Clayton 2004).

In 1997, the Kyoto Protocol was adopted by almost 200 countries, committing the signers to reduce their carbon emissions by 2012. Evidence suggests that the Protocol triggered a rapid increase in the patenting activity in CCMTs. A positive effect has been found on applications for renewable technology patents in countries with emission targets (Miyamoto and Takeuchi 2019). This increase has significantly affected solar and wind technologies patents in European countries (Johnstone et al. 2010). Even though the USA signed the Protocol on November 12, 1998, the treaty was never ratified by the US Senate, as needed.

However, the Protocol could have awakened economic incentives in other countries that had signed and ratified the Protocol, which led to increased interest in patenting in the US. Also, US polls on political concerns and political initiatives, such as President Clinton's tax on BTUs proposal or the Climate Change Action Plan, indicated potential changes in energy regulation in favor of renewable sources. Chalvatzis et al. (2020) argues that the Kyoto Protocol redirected innovation efforts towards REG, promoting the interest of the corresponding R&D to meet the growing demand in that sector. The prevailing dominant positions in patenting of PV and wind could reflect the extensive use of these technologies in the marketplace (UNEP et al. 2010).

5 Final Comments

As we have seen, TVD is commonly used in measure theory, computer and natural sciences. However, this paper shows evidence that TVD is a useful exploratory tool for measuring and detecting TTs based solely on patent data. We use the information of the entire probability density distribution of a particular technology instead of focusing on specific subgroup trends. Since the methodology does not depend on exogenous information, it is ideal for exploratory analysis of the evolution of a given technology.

This paper provides two main results. On the one hand, we first apply a distribution distance on the categories of a technological field to measure technological change. On the other hand, since we focus on temporal changes, we propose a methodology to identify tipping points in different systems, not only on technology.

Our proposed methodology might be applied to identify ecosystem regime changes in different fields. The optimal application requires keeping the same categories within the whole evaluation period or reorganizing the data for the new categorization, as in the REG case.

We acknowledge that a TT is a complex phenomenon and that the distribution of categories within a particular technological field may not completely describe it. For instance, Antal et al. (2017) suggest that energy transitions involve highly complex processes involving different actors (government, private firms, and research institutions) and that regime shifts depending on energy resources and infrastructure to utilize and benefit from them. However, since the methodology has low requirements for information, it is helpful for exploratory and visual analysis.

We do not intend to forecast technology changes or study the causality of phenomena around the TTs; our primary purpose is to avoid subjective approaches to identifying such moments. Policy evaluations or causality analysis rely on correctly identifying tipping points; in various technologies, it is unclear whether such changes happened or when. We do not distinguish between incremental or radical innovations since our interest is to detect possible tipping points within the REG technology's recent history.

The properties of TVD make it a desirable distance for the case we have considered, at which the order of categories can be changed without affecting the results. However, if the order of categories was relevant, the same methodology could be applied using a different distance. In that case, it would be better to use weak topology metrics, for example, the Wasserstein distance. Furthermore, TVD does not rely merely on a visual procedure, which is inconvenient in technology with many subgroups or over an extended period. Nevertheless, we obtain a graphic visualization of the TVD for the analysis period.

Finally, we notice that we have focused on the relative frequency of patents. However, each of those patents may have a different value, which can be approximated using forward citations or their international scope. Alternatively, those measures could be used to assign different weights to each patent so that the pdf approximated a value distribution across the correspondent subgroups.

References

Absalom, R., Förster, W., Galan, E. M., Scheu, M., Veefkind, V., and Verbandt, Y. (2006). Mapping nanotechnology patents: The EPO approach. *World Patent Information*, 28(3):204–211.

Alkemade, F., Heimeriks, G., Hoekstra, R., and Leydesdorff, L. (2015). Patents as instruments for exploring innovation dynamics: Geographic and technological perspectives on "photovoltaic cells". *Scientometrics*, 102(1):629–651.

Aminikhanghahi, S., Wang, T., and Cook, D. J. (2018). Real-time change point detection with application to smart home time series data. *IEEE Transactions on Knowledge and Data Engineering*, 31(5):1010–1023.

Angelucci, S., Hurtado-Albir, J., Karachalios, K., Thumm, N., and Veefkind, V. (2012). A new EPO classification scheme for climate change mitigation technologies. *World Patent Information*, 34(2):106–111.

Antal, M., Cherp, A., Jewell, J., Suzuki, M., and Vinichenko, V. (2017). Comparing electricity transitions: A historical analysis of nuclear, wind and solar power in Germany and Japan. *Energy Policy*, 101(May 2016):612–628.

Ball, F. and Donnelly, P. (1995). Strong approximations for epidemic models. *Stochastic processes and their applications*, 55(1):1–21.

Barger, Z., Frye, C. G., Liu, D., Dan, Y., and Bouchard, K. E. (2019). Robust, automated sleep scoring by a compact neural network with distributional shift correction. *PLoS ONE*, 14(12):1–18.

Behrouzi, S., Sarmoor, Z. S., Hajsadeghi, K., and Kavousi, K. (2020). Predicting scientific research trends based on link prediction in keyword networks. *Journal of Informetrics*, 14(4):101079.

Benner, M. and Waldfogel, J. (2008). Close to you? Bias and precision in patent-based measures of technological proximity. *Research Policy*, 37(9):1556–1567.

Carte, K. M., Lu, M., Luo, Q., Jiang, H., and An, L. (2020). Microbial community dissimilarity for source tracking with application in forensic studies. *PLoS ONE*, 15(7 July):1–17.

Cha, S. H. and Srihari, S. N. (2002). On measuring the distance between histograms. *Pattern Recognition*, 35(6):1355–1370.

Chalvatzis, K., Pitelis, A., and Vasilakos, N. (2020). Fostering innovation in renewable energy technologies: Choice of policy instruments and effectiveness. *Renewable Energy*, 151:1163–1172.

Chang, Y.-W., Huang, M.-H., and Lin, C.-W. (2015). Evolution of research subjects in library and information science based on keyword, bibliographical coupling, and co-citation analyses. *Scientometrics*, 105(3):2071–2087.

Clayton, M. (2004). Solar power hits suburbia. *The Christian Science Monitor*.

Duda, R. O., Hart, P. E., and Stork, D. G. (2007). *Pattern Classification*. Wiley, 2nd edition.

EPO and USPTO (2015). Guide to the CPC (Cooperative Patent Classification).

Euán, C., Ombao, H., and Ortega, J. (2018). The hierarchical spectral merger algorithm: a new time series clustering procedure. *Journal of Classification*, 35(1):71–99.

Frenken, K., Kupers, R., and Zeppini, P. (2014). Thresholds models of technological transitions. *Environmental Innovation and Societal Transitions*, 11:54–70.

Gao, L., Porter, A. L., Wang, J., Fang, S., Zhang, X., Ma, T., Wang, W., and Huang, L. (2013). Technology life cycle analysis method based on patent documents. *Technological Forecasting and Social Change*, 80(3):398–407.

Garcia, S. P. and Pinho, A. J. (2011). Minimal absent words in four human genome assemblies. *PLoS ONE*, 6(12):1–11.

Gibbs, A. L. and Su, F. E. (2002). On choosing and bounding probability metrics. *International statistical review*, 70(3):419–435.

Guderian, C. C. (2019). Identifying Emerging Technologies with Smart Patent Indicators: The Example of Smart Houses. *International Journal of Innovation and Technology Management*, 16(2):1–24.

International Energy Agency (2020). Special Report on Clean Energy Innovation: Accelerating technology progress for a sustainable future. Technical report.

Johnstone, N., Haščič, I., and Popp, D. (2010). Renewable energy policies and technological innovation: evidence based on patent counts. *Environmental and resource economics*, 45(1):133–155.

Kuhn, T., Perc, M., and Helbing, D. (2014). Inheritance patterns in citation networks reveal scientific memes. *Physical Review X*, 4(4):041036.

Kurtz, C., Gançarski, P., Passat, N., and Puissant, A. (2013). A hierarchical semantic-based distance for nominal histogram comparison. *Data and Knowledge Engineering*, 87:206–225.

Lacasa, I. D., Grupp, H., and Schmoch, U. (2003). Tracing technological change over long periods in Germany in chemicals using patent statistics. *Scientometrics*, 57(2):175–195.

Lin, D., Liu, W., Guo, Y., and Meyer, M. (2021). Using technological entropy to identify technology life cycle. *Journal of Informetrics*, 15(2):101137.

Lobo, J. and Strumsky, D. (2019). Sources of inventive novelty: two patent classification schemas, same story. *Scientometrics*, 120(1):19–37.

Lobo, J., Strumsky, D., and van der Leeuw, S. (2012). Using patent technology codes to study technological change. *Economics of Innovation and New Technology*, 21(3):267–286.

Makarov, M. (2004). The process of reforming the international patent classification. *World Patent Information*, 26(2):137–141.

Massart, P. (2007). *Concentration inequalities and model selection*, volume 6. Springer.

Meguro, K. and Osabe, Y. (2019). Lost in patent classification. *World Patent Information*, 57:70–76.

Miyamoto, M. and Takeuchi, K. (2019). Climate agreement and technology diffusion: Impact of the kyoto protocol on international patent applications for renewable energy technologies. *Energy Policy*, 129:1331–1338.

Perez-Molina, E. and Loizides, F. (2021). Novel data structure and visualization tool for studying technology evolution based on patent information: The DTFootprint and the TechSpectrogram. *World Patent Information*, 64(April 2020):102009.

Popp, D., Juhl, T., and Johnson, D. K. (2004). Time in purgatory: Examining the grant lag for us patent applications. *The BE Journal of Economic Analysis and Policy*, 4(1).

Ruijie, Z., Ying, X., Shuaichen, J., and Yonghe, L. (2021). Patent text modeling strategy and its classification based on structural features. *World Patent Information*, 67:102084.

Schilling, M. A. and Green, E. (2011). Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences. *Research Policy*, 40(10):1321–1331.

Smith, R. (2004). Not just tilting anymore: higher fuel costs, tax credits, better technology whip up hopes for wind power again. *The Wall Street Journal*, 1.

Solomon, B. D. and Krishna, K. (2011). The coming sustainable energy transition: History, strategies, and outlook. *Energy Policy*, 39(11):7422–7431.

Sørensen, B. (1991). A history of renewable energy technology. *Energy policy*, 19(1):8–12.

Strelkov, V. V. (2008). A new similarity measure for histogram comparison and its application in time series analysis. *Pattern Recognition Letters*, 29(13):1768–1774.

Sung-Hyuk, C. (2007). Comprehensive survey on distance/similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Sciences*, 1(4):300–307.

UNEP, EPO, and ICTSD (2010). Patents and clean energy: bridging the gap between evidence and policy. Technical report.

Appendix A Total Variation Distance properties

The TVD 1 $d(\cdot,\cdot)$ satisfies the following properties for any μ, η and ν in $\mathbb{P}(\Omega)$

i) $0 \leq d(\mu,\eta) \leq 2$

ii) $d(\mu,\eta) = 0$, if and only if $\mu = \eta$

iii) $d(\mu,\eta) \leq d(\mu,\nu) + d(\nu,\eta)$

iv) $d(a\mu, a\eta) = |a| d(\mu,\eta)$ for any $a \in R$

If μ and η have density functions f and g, the TVD distance between them can be computed as (see Massart 2007):

$$d(\mu,\eta) := \int_{-\infty}^{\infty} |f(s) - g(s)| ds. \tag{5}$$

Moreover, if $\Omega := \{\omega_1, \ldots, \omega_k\}$ is a finite set then (5) can be written as (see Gibbs and Su 2002):

$$d(\mu,\eta) := \sum_{i=1}^{k} |\mu(\omega_i) - \eta(\omega_i)| \tag{6}$$

where $\mu(\omega_i) = \mu(\{\omega_i\})$ and $\eta(\omega_i) = \eta(\{\omega_i\})$ are the μ-probability and the η-probability of the event $\{\omega_i\} \subset \Omega$.