Renewable Energy 216 (2023) 119096

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Measuring the diversification of energy sources: The energy mix

Francisco Triguero-Ruiz [a], Antonio Avila-Cano [b,*], Francisco Trujillo Aranda [c]

[a] Department of Languages and Computer Science, University of Malaga, 29071, Málaga, Spain
[b] Department of Economic Theory and Economic History, University of Malaga, 29013, Málaga, Spain
[c] University of Málaga, Spain

ARTICLE INFO

Keywords:
Concentration
Energy mix
Energy diversification index

ABSTRACT

In this paper we explore the concept of concentration/diversification of energy sources. Concentration is identified by a number that represents the energy mix, i.e., the vector of the share quota of each energy source in total consumptions. We construct a new concentration index which is a mathematical distance and economically interpretable and apply it to the world's major economies for the period between 1965 and 2018. We find that the diversification process has been considerable, intense, and widespread, and has presented different territorial patterns.

1. Introduction

In this article, we aim to address a knowledge gap: the measurement of the concentration of energy sources, or energy mix. We present the method and use to measuring diversification with an index that satisfies the cardinality property.

Concentration (or, conversely, energy *diversification*), considered from the perspective of suppliers, is a widely considered variable, indicative of the degree of *vulnerability* or *dependence* on a few suppliers and, therefore, related to supply *security* as well as other aspects related to diversification in the supply of equipment, raw materials, availability of storage, and possibilities of transmission of the energy generated. In any case, the energy mix can play as a proxy for the energy concept in models like the *Energy-Growth Nexus* and the *Energy-Environmental Nexus*. Could we improve the way we measure it?

Concentration indices are used in the literature to "summarise" the level of diversification. However, the commonly used indices are not mathematical distances and/or do not have the unit interval as range of variation (cardinality property). This means that their values do not keep the proportions, although researchers are generally not aware of this. In particular, the Herfindahl-Hirschman index is a quadratic function, and the Shannon index is a logarithmic function. This makes it difficult to draw quantitative conclusions when comparing results since neither of these indices complies with triangular inequality. We define and use a diversification index that fulfils the cardinality property (a mathematical distance that we define in the unit interval).

We want to know the meaning of the differences in magnitude of the concentration levels. This aspect is not usually considered and means that the measures of concentration or indices used are interpreted as percentages when they are not. Also, that the differences or ratios of their values are often calculated and interpreted without considering that the indices used are not mathematical distances and, therefore, do not maintain the scale, i.e., they do not keep the proportions. To this end, we present a new energy concentration/diversification index that can be interpreted as a percentage, where its differences are percentage points, and its ratios keep the proportions. This is unlike others normally used, such as that of Herfindahl-Hirschman or Shannon entropy. We believe that this point adds value to the work we present insofar as it is important for determining the cardinality of the measurement of the concentration/diversification variables. This is especially true in a scenario where modelling and the corresponding quantification of effects are shown.

The article is structured as follows. In Section 2, we explore the significance the energy concentration/diversification variable may have; we also introduce the discussion on its measurement. In Section 3 we construct a new energy concentration/diversification index. In Section 4, an application is made; specifically, we calculate the concentration/diversification of energy sources for countries or groups of countries worldwide and covering the period between 1965 and 2018. Finally, we present the conclusions. Data on energy source concentration by country and year, constructed from the original data, are presented as Supplementary Material.

* Corresponding author. Departamento de Teoría e Historia Económica, Universidad de Málaga, Plaza El Ejido, s/n, 29013, Málaga, Spain.
E-mail addresses: triguero@uma.es (F. Triguero-Ruiz), aavila@uma.es (A. Avila-Cano), trujillo@uma.es (F. Trujillo Aranda).

https://doi.org/10.1016/j.renene.2023.119096
Received 17 April 2022; Received in revised form 7 March 2023; Accepted 29 July 2023
Available online 30 July 2023
0960-1481/

2. Energy mix and measurement of diversification

The energy *mix* of an economy has been analysed in relation to national energy security (energy supply and diversification of consumption in the face of potential supply disruptions), and with the drive for cleaner, more efficient and secure means of production. Concern for energy security became increasingly prominent on the energy policy agendas of industrialised countries from 1973 onwards with the first oil crisis. This interest has gradually grown, so that the composition of the energy sources used by an economy has taken on greater importance as more attention is paid to environmental aspects as well as security. In any case, energy security is not only linked to the availability of a wide, balanced range of energy sources (only a necessary condition). Everything must also be available that makes it possible to take advantage of them from a technological and economic perspective. Therefore, diversification in the supply of equipment, the diversification of raw materials, the availability of storage and the possibilities of transmitting the energy generated are other relevant facets of a country's energy security. To the extent that we can distinguish the different facets in which diversification is embodied in each of these points, and to the extent that we can measure this diversification, the index we propose can be applied.

Thus, an energy *mix* that is highly concentrated in a certain type of generation shows vulnerability, regardless of the origin of the source (domestic or imported). In turn, concentration in polluting sources shows an additional weakness, insofar as the sustainability of that economy's growth process will have to undergo more pronounced structural adjustments. In any case, diversification of suppliers reduces the market power of any one supplier, and diversification in sources can reduce the vulnerability of supply cuts from a particular source.

In this respect, it is clear that, since The Paris Agreement [1], the share of fossil fuels in the energy *mix* has become a problem, especially when it unbalances a country's energy *mix*. Alterations in the long-term composition of the energy *mix* show the concept of *energy transition* [2].

To the extent that a country's energy profile can be understood with the notion of the energy *mix*, Csereklyei et al. [3] examine the *energy paths* or changes in the energy *mix*, i.e. the temporal developments of EU countries' energy *mixes* between 1971 and 2010 using a model based on clustering techniques. They identify seven clusters showing the typical dominant mixes in each country and observe that countries tend towards higher energy quality over time. Higher energy quality profiles are typically associated with higher national income and higher energy use per capita. They find convergence in energy intensity over time, with changes that may be voluntary, induced by market forces or political decisions, and that may develop slowly or abruptly.

The presence of renewable sources adds quality to the consumption profile reflected in the energy mix. The intermittent nature of different renewable sources is inconvenient for their localised implementation (cloud effects in solar energy, changes in wind speed in wind energy, or, despite their greater predictability, phases of the Moon in tidal energy). Neto et al. [4] investigated the complementarity of different renewable sources to achieve higher efficiency of isolated microgrids, thereby preserving and extending the useful life of the storage systems.

Kibria et al. [5] study the relationship between the share of fossil fuels (coal, oil and natural gas) in the energy *mix* and real income using a panel dataset of 151 countries for the period between 1971 and 2013. They obtain a polynomial relationship, related to Kuznets' curve. Notably, they find that the fossil fuel share has an increasing relationship with the logarithm of real GDP per capita, with different paths between developed and less developed countries.

A particularly important area of energy, where the energy *mix* analysis is essential, is that of energy security. Cohen et al. [6] provide evidence for one facet of energy security in OECD countries: the extent of oil and natural gas diversification. Up to the previous decade they find little change in diversification among oil-importing countries, although they do find changes in gas-importing countries. They construct global diversification indices for oil and gas, where the shares are that of each country in world production, weight each share by country risk and calculate country diversification for each source globally. They report considerable heterogeneity in diversification.

Kisel et al. [7] discuss short- and long-term energy security methods and indicators and present the energy security matrix. This enables the incorporation of various approaches and the structuring of significant indicators such as technical resilience and vulnerability, economic dependence and political affinity, or transport.

Based on a survey of 104 studies from 2001 to 2014, Ang, Choong and Ng [8] analyse definitions of energy security, changes in such definitions, indices, specific areas and problems in the construction of the indices, as well as energy security in the field of energy policy. Moreover, they make recommendations on the construction of indices.

Radovanovic et al. [9] provided a new energy security indicator, the *Energy Security Index*, with long-term sustainability, and tested it on a sample of 28 EU countries for the period 1990–2012 to determine the impact of six indicators on energy security. Previous methodologies had focused on supply security, without considering environmental and social indicators. The *Energy Security Index*, as a weighted average of other indicators, does include them and shows a decline in recent years.

The above references are a short sample of some of the analyses in which the energy mix is used. In addition, we must consider the very wide use of the concept of energy diversification in different modelling. It is therefore of interest to look more closely at how this variable is measured.

2.1. Measuring diversification

The concept of diversification is the opposite of concentration. Concentration reflects the relative capacity of influence that the larger entities have. Inversely, diversification reflects the relative capacity of smaller entities. Both concepts have been applied to different areas to measure the level of control of some agents, for example, political parties in an election (concentration *versus* fractionalization), sport teams in a championship (concentration *versus* competitive balance) or firms in a market (concentration *versus* competition). Thus, diversification indices have been defined based on concentration indices.

The diversification of an economy's energy sources has been widely discussed and analysed in the literature from a dual perspective:

(i) On the one hand, in terms of the countries on which the supply depends, we can study the diversification of each source: On how many countries and in what relative quantities, for example, does my oil consumption depend? And what about my gas consumption? Thus, insofar as the world's exports as a whole do not depend on just a few countries, diversification at the global level should be greater. This concept of diversification is directly linked to that of supply security since diversification of suppliers reduces vulnerability to the possible interruption of supply by one or several suppliers. To measure diversification from this perspective, modifications have been proposed to the indices used, incorporating a factor that weights the share of each supplier by its country risk [6].

(ii) On the other hand, diversification can be understood in terms of the energy sources on which an economy depends. As there are more sources and the energy *mix* is more balanced, diversification will be greater. This alternative approach is also related to the concept of security, given that dependence on a certain type of generation implies vulnerability, even if the energy source comes from the same country. Moreover, while the first approach requires, if necessary, diplomatic, and commercial efforts to increase diversification, the second approach requires transformations in the energy-demanding economic structure. In a sense, it responds to a *more economically structural* concept of security.

Our interest is focused on the measurement of diversification, so that both ways of interpreting it are relevant for our theoretical purposes. However, we will focus the application developed in this article on the second interpretation.

In the field of energy economics, as well as in ecology, the most common concentration indices are based on the Herfindahl-Hirschman Index (HHI) and Entropy Index (EI) [10]. In ecology, the Simpson Index [11] and the Shannon-Wiener Index. In the field of energy security and energy economics, the HHI is more widespread. The International Energy Agency (IEA) [12] states that "*HHI is a well-established measure of market concentration commonly used by governments*". Among the numerous examples, we can mention Martínez [13]; or Rubio-Varas and Munoz-Delgado [14], who conclude convergence in concentration levels based on the *Energy Mix Concentration Index,* which is a quantitative indicator of concentration of the energy mix based on the HHI.

The above concentration indices, among others, consider the weight that each component has in the total (*countries* in a first interpretation of the concept, and *energy sources* in another) and carry out a weighted aggregation of such *shares*. As such, the HHI [15–17] entails aggregating the square of the shares and, therefore, each share is weighted by itself [13,14,18–20]. In contrast, the Shannon Entropy or Shannon Index, EI entails adding the weighting of the shares by the logarithm of the shares. Thus, in the first case, the weight of the largest shares is reinforced, and in the second case, the weight of the smallest shares is reinforced [6].

Consider a country whose consumption of a given energy source depends on imports from n countries (or, alternatively, which needs to use n energy sources). The contribution of each country (or each source), i, has a size $a_i \in \mathbb{R}_+$. Let $A = \sum_{i=1}^{n} a_i$ denote the country's total energy consumption. Let $\boldsymbol{a} = (a_1, \ldots, a_n)$ denote the representative profile of the n countries supplying energy to that country (or, alternatively, the profile reflecting the consumption from each of the n sources from which it gets its energy consumption). Let $s_i = a_i / A$ denote the share of the country i (or the share of the source i). Therefore, according to the second interpretation, $\mathbf{s} = (s_1, \ldots, s_n)$ will be the energy *mix* of the country.

Under these circumstances, the aforementioned indices are defined as:

$$HHI = \sum\nolimits_{i=1}^{n} s_i^2$$

$$EI = \sum\nolimits_{i=1}^{n} s_i \cdot log\ s_i$$

Note that the concentration indices anonymise the energy mix; that is, they do not highlight the incidence of a given country (or a given energy source) in the final value of the index.

A real example is presented in Table 1, which reflects the percentages the different sources in the energy consumption of the European Union and China represented in 2018. Note that the EU's main source, oil, accounted for less than 40%, while China's was coal, which accounted for almost 60%. Therefore, there is apparently a higher concentration in China than in the EU, especially if we consider that the second source in the EU has more weight (natural gas, 23%) than in China (oil, 20%). Visually, Fig. 1 shows this higher concentration in China in terms of the higher *dependence* or *vulnerability* referred to above.

Table 1
Energy mix of the European Union and China, 2018 (percentages of energy consumed).

	European Union	China
Oil	38.3	19.6
Natural Gas	23.4	7.4
Coal	13.2	58.2
Nuclear energy	11.1	2.0
Hydroelectric	4.6	8.3
Renewables	9.5	4.4
Total	100.0	100.0

When making comparisons, our interest lies in identifying the meaning of the measurements, and the meaning of the differences between the diversification levels measured. This is the problem of cardinality: what is the meaning of the measurements given by a certain index? Some indices are interpretable, but others have less interpretability due to their sophistication. In this sense, a desirable property is that the index has a range of values between zero and one, because it can then be interpreted as a percentage. The unit interval is a space which has the same size as the set of real numbers. Therefore, the existence of a theoretical interval between zero and one, as the range of the index, ensures the greatest possible amplitude and facilitates the interpretation. By way of contrast, most indices (also HHI and EI) have an open or unbounded range.

Additionally, what is the *meaning of the differences* between the diversification levels measured? Cardinality implies that the differences between the diversification levels measured have a meaning. This is relevant because, for precision's sake, we have to compare the levels of diversification. Moreover, the diversification measurements are used for modelling purposes for which the parameters and their estimates must be interpretable. We can ensure this property if the function that represents the diversification index is based on a distance function, and, therefore, the index is defined in a metric space. However, most indices do not constitute a metric in the mathematical sense [21]. Herein lies our central interest.

In this regard, a twofold basic criticism of the use of either of these indices, and which underpins much of the contributions made in this paper, is that:

(i) The HHI is constructed as quadratic function and the EI index as logarithm function, so they are not based on mathematical distances, and therefore do not maintain the ratios, which makes them difficult to use in terms of comparatives.
(ii) In addition, neither has the unit interval as a theoretical range of variation, so the measurements are not percentages, and nor are their differences percentage points. In fact, the range of the HHI varies between the inverse of the number of agents and unity, and the range of the E index varies between zero and the logarithm of the number of agents (countries or energy sources). Both, therefore, present a significant interpretability problem that is sometimes overlooked.

The fulfilment of both conditions (mathematical distance and unit rank) ensures the cardinality of the measure [22]. Under these conditions, it seems reasonable to propose the need for a measure of diversification that meets this cardinality property, so that it is interpretable, and its differences and proportions are also interpretable. This is basically because the analyses are carried out in comparative terms: between countries or groups of countries, over time, or a combination of both. Therefore, a first contribution of this paper is the definition of a diversification index that satisfies the cardinality property. This index is an adaptation of the proposal set out by Triguero-Ruiz and Avila-Cano [22].

From this concentration index we can define the diversification index:

(i) If all energy sources have equal shares, the index will show a minimum concentration, diversification will be maximum, and it will have a unit value.
(ii) At the other extreme, if a country depends on one single energy source, the index will show a one hundred percent share for this single source, and it will have a null value: diversification will also be null.

We therefore propose that levels of diversification are measured by a distance function, whose range is the unit. Thus, cardinality is ensured, its meaning is a percentage, its differences are percentage points, and the index is interpretable.

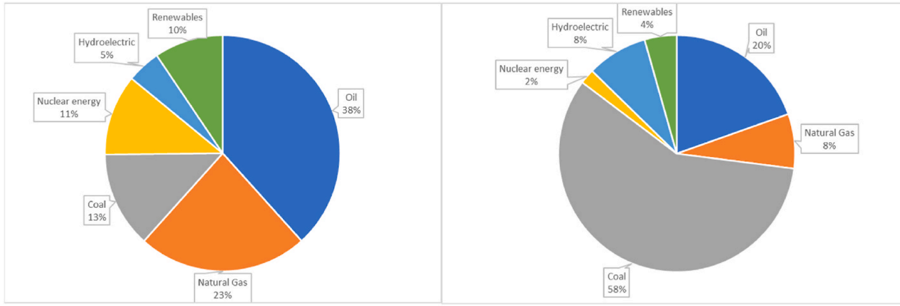

Fig. 1. Energy mix of the European Union and China in 2018.

The proposed concentration/diversification index can be flexibly used to study different phenomena in a complementary way. For example, in certain contexts, the incidence of isolated microgrids [4]. The isolate microgrids can be considered together as a component of the energy mix, separate from an eventual aggregate of "renewables". Indeed, the presence in an economy of isolated microgrids combining different primary renewable energies can be analysed from the perspective of the concentration of the energy mix: we only need to subtract the energy generated by each of the primary sources from the isolated microgrids and add to the mix an aggregate category. This would position isolated microgrids as another component of the energy mix. Similarly, as we will see below, the proposed index can also be used to study the concentration–diversification of other groups of primary energy types. Of interest may be the analysis of the diversification of renewable sources, that is, of the degree of concentration of the different renewables in a given economy.

3. The energy diversification index

Our interest is in measuring energy diversification across countries and over time by means of an index that satisfies the cardinality property. Therefore, the energy diversification variable (or, conversely, energy source concentration) will be represented by this index. As mentioned above, the existing indices do not fulfil this property (so it is erroneous to use them to make comparisons or to give meaning to measurements).

Consider a group of economies in which there are various energy sources. Each economy demands diverse types and quantities of energy sources and uses a different variety and proportions of these. This is the *energy mix* of each country at any given moment.

Let $n \in \mathbb{N}$ be the maximum energy sources in any economy. We want to measure and compare diversification levels. Therefore, we need a reference. This reference will be an economy with n energy sources.

Consider the economy with n energy sources. Each source, i, has a size (quantity) $a_i \in \mathbb{R}_+$. Let $A = \sum_{i=1}^{n} a_i$ denote the size of the energy consumption of the economy or country. Let $\boldsymbol{a} = (a_1, \ldots, a_n)$ denote the profile of the energy source's size in the economy. Let $s_i = a_i / A$ denote the share of the size of source i.

An *energy mix* is a vector of the shares of the sources' size $s = (s_1, \ldots, s_n)$. Let $\mathbb{S}^{n-1} = \{(s_1, \ldots, s_n) \in \mathbb{R}_+ : s_i \in [0, 1]$ *for every* i, *and* $\sum_{i=1}^{n} s_i = 1\}$ be the set of admissible energy mixes, given n. Note that an economy with fewer sources may be thought of as having additional sources with zero shares. Note that the emergence of new energies can be incorporated into the energy mix in this way: by allocating zero initial quotas. So, a *maximum concentrated mix* is an energy mix $s^M \in \mathbb{S}^{n-1}$ such that $s_i = 1$ for source i, that is, i is the only source; this energy configuration has $n-1$ zeros. Furthermore, the *minimum concentrated mix* is $s^m \in \mathbb{S}^{n-1}$ such that $s_i^m = 1/n$ for every source i.

A *distance function* on the set of admissible energy mixes is a function $x : \mathbb{S}^{n-1} \times \mathbb{S}^{n-1} \to \mathbb{R}_+$ such that, for every $s, s', s'' \in \mathbb{S}^{n-1}$ the following conditions are satisfied:

(i) $x(s, s') = x(s', s)$.
(ii) $x(s, s') = 0$ if and only if $s = s'$, and
(iii) $x(s, s') \leq x(s, s'') + x(s'', s')$.

The distance to the minimum concentrated mix, s^m, can be measured from any energy mix, s, as $x(s, s^m) = X(s)$. Additionally, the maximum distance from any of the maximum concentrated mixes to the minimum concentration, $x(s^M, s^m) = X_{max}$, can also be measured with this. Note that, given n, X_{max} is constant.

Definition 1. An energy concentration index is a function $G : \mathbb{S}^{n-1} \to U \subseteq \mathbb{R}$ which assigns a real number to each energy mix.

Definition 2. An energy concentration index, G, complies with a cardinality property if (i) there exists a distance function, $x : \mathbb{S}^{n-1} \times \mathbb{S}^{n-1} \to \mathbb{R}_+$, such that $G = x(s, s^m)$ for every $s \in \mathbb{S}^{n-1}$, and (ii) $U = [0, 1]$ and for every $k \in U$, exists $z \in \mathbb{S}^{n-1} : G(z) = k$.

Remark 1. A concentration index, G, anonymises the components of the vector s that identifies any energy mix.

Definition 3. The family of energy concentration indices, $C(s)$, assigned to each energy configuration, s, is the ratio of the distance from s to the minimum concentrated mix, $X(s)$, and the maximum distance, X_{max}: $C(s) = X(s)/X_{max}$.

$C(s)$ represents a family of indices depending on the concept of distance. Thus, we must choose a certain concept of distance to apply it. Given that it is widely used, a good option is the Euclidean distance, defined as $x(s, s') = ||s - s'|| = [\sum_{i=1}^{n} (s_i - s_i')^2]^{1/2}$ for each $s, s' \in \mathbb{S}^{n-1}$. Thus, it is possible to measure the distance from each energy mix to the minimum concentrated mix as $X(s) = x(s, s^m) = ||s - s^m|| = \left[\sum_{i=1}^{n} \left(s_i - \frac{1}{n}\right)^2\right]^{1/2}$ [23]. Hereafter, we consider the Euclidean distance. Proposition 1 establishes the way in which to calculate $C(s)$ under these conditions.

Proposition 1. *For every energy mix,* $s \in \mathbb{S}^{n-1}$, *if* $x : \mathbb{S}^{n-1} \times \mathbb{S}^{n-1} \to \mathbb{R}_+$ *is the Euclidean distance, the energy concentration index* $C(s)$ *can be defined as*:

$$C(s) = \sqrt{\frac{n \sum_{i=1}^{n} s_i^2 - 1}{n - 1}}$$

Proof. Let $s \in \mathbb{S}^{n-1}$ be any energy configuration. Given that $X(s) = \left[\sum_{i=1}^{n} s_i^2 - \frac{1}{n}\right]^{1/2}$, and $X_{max} = \sqrt{(n-1)/n}$, we have the result. ∎.

$C(s)$ complies with the cardinality property, because it is based on a distance and, by construction, its range is the unit interval. Given $C(s) \in [0, 1]$, it is the percentage of the maximum distance if we multiply by one hundred. Furthermore, the difference between two configurations is measured in percentage points. Note that HHI index do not fulfil this property, because its range is not a unit interval and it is not based on a distance, and therefore do not comply with the triangular inequality. Neither does the E index.

By construction (as with other indices, such as the Herfindahl–Hirschman or entropy indices), our index anonymises the energy sources and 'summarises' their values in a single real number (Remark 1). Therefore, it is impossible to highlight the incidence of a particular energy source versus another. It is up to the researcher or analyst to perform the analysis and interpretation. Several examples can help to demonstrate the scope of this issue:

(i) Two different economies, A and B, that import their energy sources from a set of n countries may have the same level of concentration, although each economy depends on the exports made by a different country. It is up to the analyst to assess this circumstance.
(ii) Similarly, if a new source is incorporated that may pose availability problems, the greater diversity shown by the index value could be paradoxically associated with a greater security risk. It is also the analyst's responsibility to evaluate this circumstance.
(iii) Two economies, A and B, may have the same value for the energy source concentration index. However, economy A may have a high proportion of coal in its mix, and economy B may have a high proportion of renewables. This would indicate that if economy B has sufficient technology and supply capacity to continue to take advantage of the natural availability of its energy resources, economy B would be better positioned strategically in energy terms than economy A.

Therefore, the analyst must provide added value that the index, like any other that anonymises energy sources, cannot directly provide.

Remark 2. The consideration in the index of the number of energy types, n, assumes that the distance in the corresponding metric space is being measured. This makes it possible to quantify the effect of the appearance, for example, of a new or future primary energy, k, which can be incorporated into the mix and which initially might have a zero share: $s_k = 0$. In this way, the dynamics of diversification can be contemplated. Strategic planning would have an index that fulfills the cardinality property and that allows us to correctly measure the objectives set and the levels achieved throughout intermediate evaluations.

Note that the concentration in the mix depends on the number of categories or types of energy and the inequality of the shares they have. The fact that we foresee the future presence of more types of primary energy makes it possible to identify this share in the present as zero, which 'increases' the concentration value. As the new types of primary energy become present, and their shares in the mix become greater than zero, the concentration will decrease.

The dynamic analysis must anticipate the emergence or even disappearance of energy sources. This requires that the concentration index be normalised [24], that is, relativised to its achievable minimum and maximum. In this sense, Remark 3 tells us that our concentration index is normalised while providing an easy method of calculating it and reproducing results of previous studies.

Remark 3. Because the maximum of *HHI* is unity (a single energy source cups the total mix) and its minimum is the inverse of the number of energy sources, $^{1}/_{n}$, our concentration index, C, can be reinterpreted as the square root of the standardised *HHI*, HHI_{norm}: $C = \sqrt{HHI_{norm}}$.

By construction, $C(s)$ is the ratio between two distances and represents how different a particular energy mix is with respect to a maximum concentrated mix. Additionally, we can offer an interesting interpretation of $C(s)$. Consider that one of the sources is the preponderant source. This energy mix is $\boldsymbol{d} = \left(d_1, \frac{1-d_1}{n-1}, \ldots, \frac{1-d_1}{n-1}\right)$, where d_1 is the share of the preponderant source and the rest, $1 - d_1$, is distributed among the rest of $n - 1$ sources. So, we have the following proposition.

Proposition 2. *The value of the energy concentration index $C(\boldsymbol{d})$ in an economy with a preponderant source whose energy share is d_1 is:*

$$C(\boldsymbol{d}) = \frac{nd_1 - 1}{n - 1}$$

Proof. Given an energy mix $\boldsymbol{d} \in \mathbb{S}^{n-1}$, we have that $\sum_{i=1}^{n} d_i^2 = d_1^2 + \frac{(1-d_1)^2}{n-1}$ and, by substituting for $C(\boldsymbol{d})$, we can obtain the result. ∎.

Corollary 1. *When the number of sources reaches infinite, the index value of $C(\boldsymbol{d})$ tends towards d_1.*

Proof. Immediate: $\lim_{n\to\infty} C(\boldsymbol{d}) = d_1$. ∎.

Therefore, $C(s)$ can also be interpreted as the relative size of the predominant source in an economy with many sources in which all the others are of equal size. This may be a useful interpretation, because the value of $C(s)$ in an economy would be the "*effective share*" of the predominant source in this economy. Therefore, when we compare different values of $C(s)$ for different economies, we are comparing the levels of effective shares of predominant sources measured in percentages.

We understand the phenomenon of diversification as complementary to concentration. In these terms, Definition 4 provides us with a measuring function of diversification levels.

Definition 4. An energy diversification index is a function $F : \mathbb{S}^{n-1} \to U \subseteq \mathbb{R}$ which assigns a real number to each energy mix.

Definition 5 gives us an index constructed as a complement to the unit of the concentration index.

Definition 5. The family of diversification indices, $D(s)$, assigned to each energy mix, $s \in \mathbb{S}^{n-1}$, is the ratio from the maximum distance, X_{max}, to $X(s)$, and X_{max}:

$$D(s) = \frac{X_{max} - X(s)}{X_{max}} = 1 - C(s)$$

We understand that $D(s)$ represents a family of indices based on the concept of distance. As we did for the concentration index, hereafter we consider the Euclidean distance.

$D(s)$ measures the degree of diversification, that is, the share of $n - 1$ divided into infinitesimal sources as opposed to the predominant source share, which is $C(s)$.

Under these conditions, a simple example serves to illustrate the indices constructed and the relevance of the cardinality property. Table 1 shows the energy mixes of European Union and China in 2018. Table 2 shows the concentration and diversification indices.

Note that the HHI and E indices show lower energy concentration differences in China than our $C(s)$ index. If these differences are measured as percentage points (for which we must multiply the index values by one hundred), as is sometimes mistakenly done, the first two indices would indicate that the Chinese economy's energy consumption is around 15% points more concentrated in certain sources than that of the European economy. The reality is that, if we consider $C(s)$ which as

Table 2
Concentration and diversification energy indices. The European Union and China, 2018.

	European Union (1)	China (2)	Differences	
			Absolute (2-1)	Relative (2/1)
HHI	0.242	0.392	0.150	1.62
EI	−0.688	−0.543	0.145	0.79
$C(s)$	0.301	0.520	0.220	1.73
1-HHI	0.758	0.608	−0.150	0.80
- EI	0.688	0.543	−0.145	0.79
$D(s)$	0.699	0.480	−0.220	0.69

has been shown satisfies the cardinality property, the energy consumption of the Chinese economy is 22% points more concentrated than that of the European economy. The inverse analysis is valid for the comparative analysis of diversification levels.

Also note that, in relative terms, the energy diversification shown by the European economy is also substantially higher than that of the Chinese economy, which shows an index value below 70% of that of the EU. However, using HHI and E, we would interpret China's energy diversification levels to be around 80% of those of the EU.

4. Application to energy source diversification (1965–2018)

Given that we now have two energy source concentration/diversification indices that satisfy the cardinality property and therefore allow comparisons to be made, in this section we first present an analysis of the annual evolution of this variable, by country and geographical area, over the last half century (1965–2018).

The information comes from *BP Statistical Review of World Energy*, 68th edition, corresponding to June 2019, and can be located at:

https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html.

This information is compiled from government sources and published data. The database used refers to primary energy consumption, and distinguishes between (i) oil, (ii) gas, (iii) coal, (iv) nuclear, (v) hydroelectric, and (vi) renewables. From this information, shares and diversification indices have been constructed, resulting in a database consisting of 4212 pieces of data, corresponding to the 54 annuities of the period 1965–2018, for 78 countries or groups of countries. This information is provided in Supplementary Material.

Note that the consumption data used are based on gross generation and not accounting for cross-border electricity supply. Therefore, in this application, the analysis focuses on consumption from domestic primary sources without incorporating imports. In any case, the availability of information on the primary sources of imported electricity can be incorporated into the index. On the other hand, the characterisation and analysis of the diversification of energy imports and exports, performed with traditional indices such as the entropy index [25], can also be conducted using the index we have proposed.

In each large geographical area, individual data are not available for smaller countries, meaning only aggregated data can be accessed. Thus, there are observations referring to Central America, Other Caribbean, Other South America, Other Europe, Other Middle East, Eastern Africa, Middle Africa, Western Africa, Other Northern Africa, Other Southern Africa, and Other Asia Pacific. Each of these aggregates is dealt with as a separate piece of data. In these cases, consumption data refers to the aggregate of small countries, so we know the total consumption in the large geographic area.

Given the geopolitical changes that have occurred over such a long period of time, we have had to make some choices to homogenise the data series. Specifically, we must bear in mind the following:

(i) Bangladesh, for which data are available from 1971, does not appear.
(ii) The original data for Ukraine are available from 1985. It was decided to reconstruct them up to 1965. For this purpose, the data for the former USSR have been assigned. We believe that this is not a serious error insofar as the data for the Commonwealth of Independent States (CIS) and the USSR are the same between 1965 and 1984.
(iii) Data for Estonia, Latvia, Lithuania are not included in the analysis as they are not available until 1985. The same applies for Azerbaijan, Belarus, Kazakhstan, the Russian Federation, Turkmenistan, Uzbekistan and the rest of the CIS countries, which are aggregated under the CIS heading.
(iv) Croatia, North Macedonia, and Slovenia, for which data are available from 1990 onwards, do not appear.

The minimum value of the $C(s)$ index (maximum diversification) is 0.23, corresponding to Slovakia in 2014, 2015 and 2016. The highest value is equal to 1.00 and is repeated for several years and countries between 1965 and 1990, indicating that only one energy source was available.

The arithmetic mean of the annual data has decreased significantly. In the late 1960s and early 1970s, before the first oil crises, the concentration of energy sources was around 70%. Between 1974 and 1993 there was a gradual reduction from 68% to 58%. In the last twenty-five years it has fallen to 50%. It is noteworthy that in no year has the average concentration increased in comparison with the previous one: the process of diversification of energy sources has been slow and gradual, but inexorable. In fact, world oil and gas consumption has increased linearly due to growth in non-OECD countries. Coal consumption is stable over the last decade, and clearly declining in the OECD and EU. Renewable energies are growing exponentially. So, the landscape of the energy industry is altering, and most likely renewable energy will replace fossil fuel and dominate the worldwide-energy mix [26].

This is true if we look at the overall concentration at the global level, which would correspond to a weighted average. Fig. 2 shows a linear fit with statistical significance of the negative slope: Student's $t = -23.4$ and $p = 2.842E-29$ (***).

Until the energy crises of the first half of the 1980s, concentration was above 40% and as high as 47%. Since then, there has been a steady decline until reaching the current figure of 32%.

The magnitude of the extent of diversification can be seen in Fig. 3. The number of countries or groups of countries with a concentration level below (green) and above the global arithmetic average (red) is identified for each year of the period. Note that from the mid-1980s onwards, the majority of countries (or groups of countries) are more energy diversified.

Furthermore, Fig. 4 shows the countries and areas ordered from lowest to highest according to the number of years where they are above the average in terms of diversification.

Although the process of increasing diversification of energy sources is considerable, intense, and widespread, it is also true that it has presented different territorial patterns, as can be seen in Fig. 5. In Fig. 5, concentration indices for energy sources calculated specifically for each major geographic area and for the world total are represented (Table A2 in the Supplementary Material). These indices are calculated from

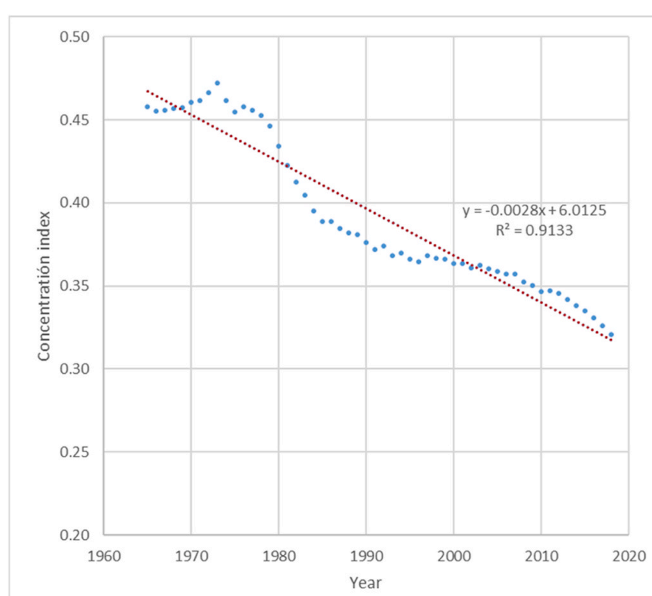

Fig. 2. Evolution of world energy concentration (1965–2018).

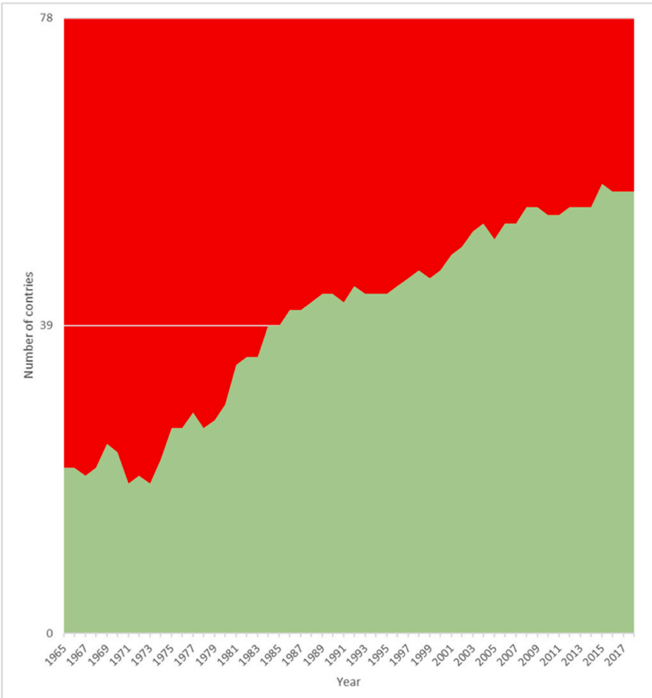

Fig. 3. Evolution of the number of countries or groups of countries with higher and lower levels of concentration of energy sources than the average (1965–2018).

aggregated primary energy sources consumption in these areas and globally. Therefore, they are not averages of concentration index values by countries or areas. Specifically, diversification by area has increased across the board, except in the CIS countries where, since the mid-1980s, after the collapse of the USSR, there has been an increase in concentration. The same is true of the Asia-Pacific region since the turn of the century, although in recent years it has shown a sharp decline in concentration.

The remaining areas show a reduction in concentration, with differences in levels between the highest values in the Middle East and the lowest in Europe and North America. In this regard, it should be stressed that, during the first twenty years of the period analysed, the former USSR had the highest levels of diversification, together with Europe and North America, and that this position has dropped since the mid-1980s, to currently reach second position in terms of concentration levels after the Middle East.

5. Conclusions

An energy source concentration index "summarises" the energy mix in a number, i.e., the vector of the shares of each energy source in total consumption is reduced to a number. We have resolved an issue strictly linked to the measurement of concentration/diversification levels, which we want to use in comparative terms. To this end, we have constructed an index which, unlike those habitually used (Herfindahl-Hirschman index or Shannon index), complies with the cardinality property. This property ensures that the index values are in the unit interval, that they can be interpreted as percentages and, since the index is defined as a mathematical distance, they maintain the proportions. Our index can be interpreted as the share that a main source would have if all other sources were plentiful, and they made a small contribution.

We have applied the index to the main world economies, or groups of them, on an annual basis between 1965 and 2018, so that we have a global view of the evolution of this significant variable.

The diversification process of energy sources at the global level can

Fig. 4. Number of years between 1965 and 2018 in which each country has an energy source concentration index below the world average.

be summarised by three features. First, over the last half century, the diversification process has been extremely pronounced on a global level, such that the actual trend in the evolution of the index has considerable significance and the corresponding coefficient of determination, R^2, is

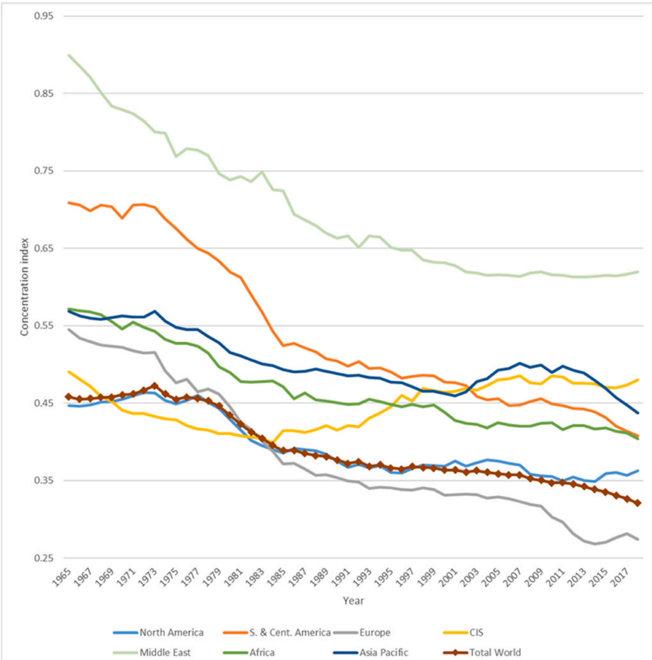

Fig. 5. Evolution of the concentration of energy sources by geographical area (1965–2018).

above 0.90. Second, the diversification process has, moreover, been generally extensive in terms of the countries involved, so that since the mid-1980s most countries have shown energy diversification levels above the world average. And third, the diversification process of energy sources has been considerable, intense, and widespread, but it has presented different territorial patterns: after the collapse of the USSR, the countries of the former CIS have increased in concentration, as has the Asia-Pacific area, although in this case the process has been reversed in recent years. The highest levels of concentration are found in the Middle East, while the highest levels of diversification are found in Europe.

CRediT authorship contribution statement

Francisco Triguero-Ruiz: Data curation, Formal analysis, Software, Methodology, Visualization, Writing – review & editing. **Antonio Avila-Cano:** Conceptualization, Investigation, Funding acquisition, Methodology, Writing – original draft. **Francisco Trujillo Aranda:** Formal analysis, Software, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are very grateful to three anonymous reviewers for extremely useful comments and suggestions. This work was supported by the Ministerio de Ciencia e Innovación (Spain) PID2020-114309GB-I00. Funding for open access charge: Universidad de Málaga / CBUA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.renene.2023.119096.

References

[1] The Paris Agreement | UNFCCC, (2016). https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement (accessed August 4, 2021).
[2] M. Rubio-Varas, B. Muñoz-Delgado, Long-term diversification paths and energy transitions in Europe, Ecol. Econ. 163 (2019) 158–168, https://doi.org/10.1016/j.ecolecon.2019.04.025.
[3] Z. Csereklyei, P.W. Thurner, J. Langer, H. Küchenhoff, Energy paths in the European Union: a model-based clustering approach, Energy Econ. 65 (2017) 442–457, https://doi.org/10.1016/j.eneco.2017.05.014.
[4] P.B.L. Neto, O.R. Saavedra, D.Q. Oliveira, The effect of complementarity between solar, wind and tidal energy in isolated hybrid microgrids, Renew. Energy 147 (2020) 339–355, https://doi.org/10.1016/j.renene.2019.08.134.
[5] A. Kibria, S.B. Akhundjanov, R. Oladi, Fossil fuel share in the energy mix and economic growth, Int. Rev. Econ. Finance 59 (2019) 253–264, https://doi.org/10.1016/j.iref.2018.09.002.
[6] G. Cohen, F. Joutz, P. Loungani, Measuring energy security: trends in the diversification of oil and natural gas supplies, Energy Pol. 39 (2011) 4860–4869, https://doi.org/10.1016/j.enpol.2011.06.034.
[7] E. Kisel, A. Hamburg, M. Härm, A. Leppiman, M. Ots, Concept for energy security matrix, Energy Pol. 95 (2016) 1–9, https://doi.org/10.1016/j.enpol.2016.04.034.
[8] B.W. Ang, W.L. Choong, T.S. Ng, Energy security: definitions, dimensions and indexes, Renew. Sustain. Energy Rev. 42 (2015) 1077–1093, https://doi.org/10.1016/j.rser.2014.10.064.
[9] M. Radovanović, S. Filipović, D. Pavlović, Energy security measurement – a sustainable approach, Renew. Sustain. Energy Rev. 68 (2017) 1020–1032, https://doi.org/10.1016/j.rser.2016.02.010.
[10] A. Stirling, A general framework for analysing diversity in science, technology and society, J. R. Soc., Interface 4 (2007) 707–719, https://doi.org/10.1098/rsif.2007.0213.
[11] E.H. Simpson, Measurement of diversity, Nature 163 (1949), https://doi.org/10.1038/163688a0, 688–688.
[12] (IEA) International Energy Agency, Energy Security and Climate Policy, Assessing Interactions — European Environment Agency, Organisation for Economic Co-operation and Development, Paris, 2007. https://www.iea.org/news/energy-security-and-climate-policy-assessing-interactions. (Accessed 4 August 2021).
[13] P. Martínez Fernández, An Application of the Modern Portfolio Theory to the Optimization of the European Union Power Generation Mix from an Environmental Perspective, Universidade da Coruña, 2019. https://ruc.udc.es/dspace/handle/2183/23805. (Accessed 4 August 2021).
[14] M. Rubio-Varas, B. Muñoz-Delgado, The energy mix concentration index (EMCI): methodological considerations for implementation, MethodsX 6 (2019) 1228–1237, https://doi.org/10.1016/j.mex.2019.05.023.
[15] O.C. Herfindahl, Concentration in the US Steel Industry, Unpublished PhD Dissertation, Columbia University, Columbia, 1950.
[16] A.O. Hirschman, National Power and the Structure of Foreign Trade, University of California Press, Berkeley, 1945.
[17] A.O. Hirschman, The paternity of an index, Am. Econ. Rev. 54 (1964) 761–762. http://www.jstor.org/stable/1818582. (Accessed 6 August 2018).
[18] E. Gupta, Oil vulnerability index of oil-importing countries, Energy Pol. 36 (2008) 1195–1211, https://doi.org/10.1016/j.enpol.2007.11.011.
[19] C. Le Coq, E. Paltseva, Measuring the security of external energy supply in the European Union, Energy Pol. 37 (2009) 4474–4481, https://doi.org/10.1016/j.enpol.2009.05.069.
[20] A. Löschel, U. Moslener, D.T.G. Rübbelke, Indicators of energy security in industrialised countries, Energy Pol. 38 (2010) 1665–1671, https://doi.org/10.1016/j.enpol.2009.03.061.
[21] K.G. Binmore, The Foundations of Topological Analysis: A Straightforward Introduction: Book 2 Topological Ideas, Cambridge University Press, 1981.
[22] F. Triguero-Ruiz, A. Avila-Cano, The distance to competitive balance: a cardinal measure, Appl. Econ. 51 (2019) 698–710, https://doi.org/10.1080/00036846.2018.1512743.
[23] G.G. Szpiro, Hirschman versus Herfindahl: some topological properties for the use of concentration indexes, Math. Soc. Sci. 14 (1987) 299–302, https://doi.org/10.1016/0165-4896(87)90008-4.
[24] E.P. Santos Júnior, M.V.B. da Silva, F.J. Simioni, P. Rotella Junior, R.S.C. Menezes, L.M. Coelho Junior, Location and concentration of the forest bioelectricity supply in Brazil: a space-time analysis, Renew. Energy 199 (2022) 710–719, https://doi.org/10.1016/j.renene.2022.09.001.
[25] C.-C. Lee, S.-J. Ho, Impacts of export diversification on energy intensity, renewable energy, and waste energy in 121 countries: do environmental regulations matter? Renew. Energy 199 (2022) 1510–1522, https://doi.org/10.1016/j.renene.2022.09.079.
[26] Y. Chen, R. Mamon, F. Spagnolo, N. Spagnolo, Renewable energy and economic growth: a Markov-switching approach, Energy 244 (2022), 123089, https://doi.org/10.1016/j.energy.2021.123089.