Energy Strategy Reviews 50 (2023) 101241

Contents lists available at ScienceDirect

Energy Strategy Reviews

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

Methodology for selecting assessment periods of Logarithmic Mean Divisia Index decomposition techniques☆

Juan David Rivera-Niquepa [a,b,*], Daniela Rojas-Lozano [a], Paulo M. De Oliveira-De Jesus [a], Jose M. Yusta [b]

[a] Department of Electrical and Electronic Engineering, Los Andes University, Bogotá, 111711, Colombia
[b] Department of Electrical Engineering, University of Zaragoza, Zaragoza, 50009, Aragón, Spain

ARTICLE INFO

Keywords:
Divisa Index Decomposition Analysis
Logarithmic Mean Divisia Index Decomposition
Explanatory drivers
Mean square error

ABSTRACT

Logarithmic Mean Divisia Index Decomposition Analysis (IDA-LMDI) is a widely used statistical technique in the engineering, economics, energy, and environmental sciences. The main application of the IDA-LMDI method is to identify the drivers that explain the change in carbon dioxide emissions of a country or region over a given time span. Therefore, proper selection of each time period is fundamental to ensure that its implemented accurately and meaningful results are obtained. In the literature, decomposition periods have been defined based on a single-, n-, or one-year period. However, in all these cases, the duration of each period was fixed and depended on arbitrary criteria. The adoption of fixed periods did not capture specific changing trends in the time series under analysis. Therefore, this study presents a new methodology for defining the number of periods and their extent to obtain unbiased drivers using the IDA-LMDI technique. The period selection was defined according to the mean square error of all feasible combinations. We illustrated the application of this method in two cases. First, an illustrative example was successfully validated using multivariate linear regression. Second, we obtained the Kaya factors of the Organisation for Economic Co-operation and Development countries of Europe. The results were compared with the decomposition results provided by the International Energy Agency on 10–7 year fixed periods. The results showed that the application of fixed periods in IDA-LMDI dismissed critical drivers that could only be captured with the proper selection of each period under analysis. Therefore, the proposed methodology will aid in accurate information analysis and decision making in energy policy and other applications. The proposed method has a broad application and could be applied to any decomposition method in the Divisia family.

1. Introduction

Divisia Index decomposition techniques are used to identify the drivers that explain the change in an output indicator as a function of a set of input variables. Different methods have been reported in the literature since the early 1970 [1]. According to [1,2], the most notable and widely used methods include the Laspeyres Parametric Divisia Method, Paasche Parametric Divisia Method, Arithmetic Mean Divisia Index (AMDI), and Logarithmic Mean Divisia Index Decomposition Analysis (IDA-LMDI). According to [3], the IDA-LMDI analysis comprises two methods LMDI-I and LMDI-II with additive and multiplicative formulations each. Although the results of the two methods are similar the LMDI-I method is the most widely used because it has properties such as: consistency in aggregation and perfect decomposition. A review of the literature on the frequency of citations of the Divisia methods in the Scopus database is presented in Fig. 1. The results show that decomposition analysis originated in the 1970 with the development of the Laspeyres and Paasche parametric Divisia Methods [4]. On the other hand, the use of AMDI techniques is rare. The IDA-LMDI method has been used extensively in decomposition studies. This method was developed by [5] and has been used extensively since 2010. The fields in which IDA-LMDI decomposition techniques are applied are diverse, but mainly include economics, engineering, and energy, as well as environmental sciences.

A formulation, update, and review of IDA-LMDI methods are presented in [3]. The main indicators used in decomposition studies are energy consumption, energy intensity, carbon emissions, and carbon

☆ This study was funded by the Ministerio de Ciencia Tecnología e Innovación, Colombia, Convocatoria 1 de 2018.
* Corresponding author at: Department of Electrical and Electronic Engineering, Los Andes University, Bogotá, 111711, Colombia.
E-mail addresses: jd.rivera@uniandes.edu.co (J.D. Rivera-Niquepa), ddp.rojas11@uniandes.edu.co (D. Rojas-Lozano), pm.deoliveiradejes@uniandes.edu.co (P.M.D.O.-D. Jesus), jmyusta@unizar.es (J.M. Yusta).

https://doi.org/10.1016/j.esr.2023.101241
Received 28 June 2023; Received in revised form 9 October 2023; Accepted 15 October 2023
Available online 20 October 2023

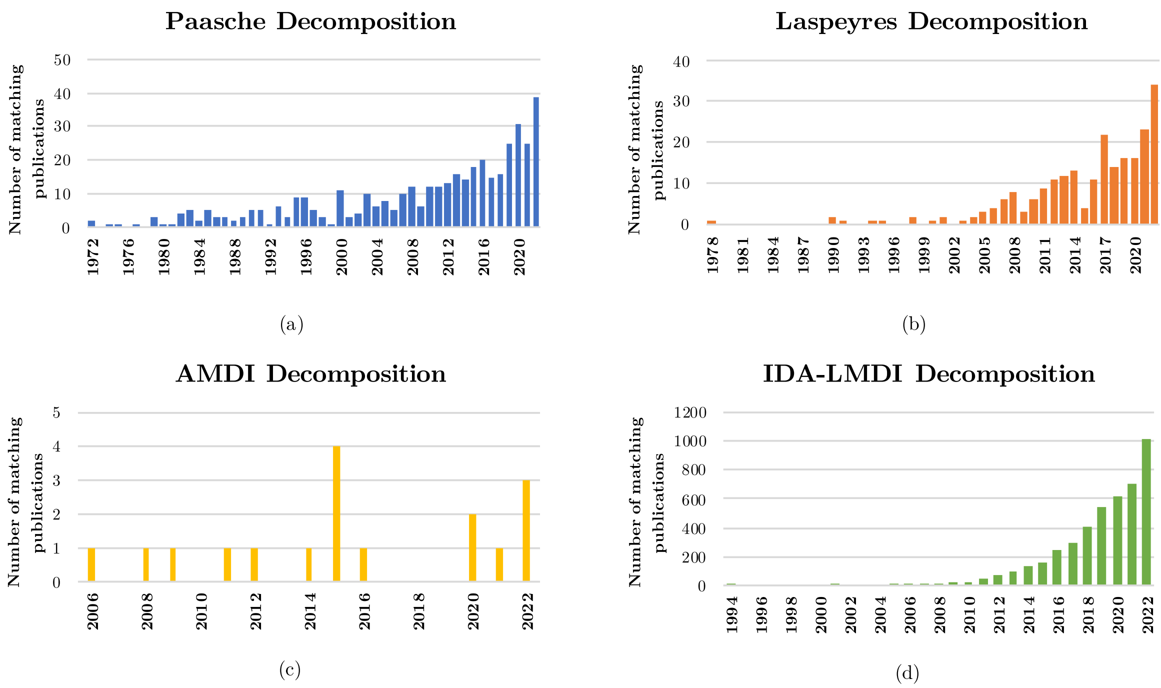

Fig. 1. Publication frequency on Divisia Index methods. Data retrieved from Scopus database on May 20, 2023.

intensity. The IDA-LMDI technique can be applied under three temporal frameworks: year-by-year (yy), fixed n-year single-period (sp) [1], and n-year multi-period (mp) approaches. Under the yy approach, decomposition is performed for each year and accumulated up to the final year. The sp approach only considers a fixed n-year period. Finally, in the mp approach, multiple n-year subperiods are fixed within one main period. In the latter case, subperiods of four or five years are usually considered. Literature reviews of IDA-LMDI methods can be found in [6] (up to 2000) and [2] (up to 2013). Detailed classifications of the mathematical models used and the areas of application in the analyses reported for carbon dioxide emissions decomposition are presented in [7,8].

Multiple methodologies have been reported in the literature to identify trend changes and segmentation in time series in different fields. Recently, the following can be highlighted: In [9] the Detecting Breakpoints and Estimating Segments in Trend (DBEST) algorithm is presented, this is used to analyze trend changes in vegetation in Iraq. In [10] different algorithms for detecting trend and seasonality changes in land surface temperature time series are compared. Methodologies combining fuzzy and clustering techniques for segmenting time series are introduced in [11]; in [12] this approach is used to analyze tunnel boring machine variables. Most of these reported cases are applied in time series with high presence of noise and reduced periodicity. However, in the field of annual time series decomposition analysis of carbon dioxide emissions with IDA methods, no segmentation and trend change methods were reported.

Classifications of the most recent and relevant publications on IDA-LMDI applications are presented in Table 1. The contributions are classified by region, application sector, time span, temporal framework approach, and whether the number of periods is defined arbitrarily. The review encompasses contributions to carbon dioxide emissions, Kaya factors, carbon intensity, and energy demand, as well as the power, transportation, and manufacturing sectors.

Table 1 that there is no unified criterion for defining the temporal framework required to perform IDA-LMDI decomposition analysis. The sp approach is easy to implement; however, because of its simplicity, it loses important interannual information. The cumulative yy approach is more detailed, but requires complete historical data and generates a large amount of output information; therefore, it has been used mainly in aggregate analyses. To compensate for these disadvantages, many authors use a mp approach with fixed periods of four or five years in most cases, which are arbitrary. In [34,45], the authors used a multi-period analysis considering the trend of the aggregate variable; however, there was no method to define the number of periods or the extent of each period. Instead, the specification of the assessment period depended on the analyst's criteria, which were generally arbitrary.

In our view, the arbitrary selection of periods can lead to bias in the interpretation of the resulting explanatory drivers [34], due to the lack of a rigorous and unified technical criterion that allows for contrasting and verifying the results. Thus, to the best of our knowledge, no methodology exists in the literature for appropriate selection of the evaluation periods in Divisia Index-based decomposition techniques.

To fill this research gap, this study presents a new methodology for adequate selection of assessment periods for Divisia Index-based decomposition techniques. To achieve this, we propose a novel technique based on the minimization of the mean squared error (MSE) of linear regressions in the feasible segmentation of the time series of the indicator to be decomposed.

The remainder of this paper is organized as follows. Section 2 describes the development of the proposed method. Section 3 presents the results of the methodology based on two case studies, Section 4 presents the discussion, and conclusions are drawn in Section 5.

2. Methodology

This section presents the proposed methodology for applying Divisia index methods through appropriate selection of assessment periods. In this paper a LMDI-I type decomposition model is used, however, the proposed methodology can be used with the decomposition methods of the Divisia family. First, the time-series partitioning problem is explained. Then, the mathematical fundamentals for the identification of partitions through an optimization problem are analyzed. Subsequently, decomposition analysis using the IDA-LMDI method is presented. Finally, we present a solution algorithm that aims to minimize the total mean squared error (TMSE) for a group of linear regression models performed on a set of feasible partitions. The results include

Table 1
Classification of relevant contributions in IDA-LMDI.

Reference	Region/Sector[a]	Time[b]		Period[c]		Non-Arbitrary
		span	mp	sp	yy	periods
Lin (2023) [13]	China(E)	05-15(5)	•			no
Huang (2023) [14]	China(E)	00-20(5)	•			no
Chun (2023) [15]	Henan-China(E)	06-19(4)	•			no
Wang (2022) [16]	China(E)	10-18(3)	•			no
Wang (2022) [17]	Asia(E)	00–14		•		no
Rivera (2022) [18]	Colombia(P)	90-20(4)	•			no
Li (2022) [19]	China(P)	00–17			•	no
Li (2022) [20]	China(E)	09–19		•		no
Fernández (2022) [21]	Spain(E)	08–18			•	no
Faridzad (2022) [22]	Netherlands(MI)	05–15		•		no
He (2022) [23]	China(P)	05-19(5)	•			no
Chen (2022) [24]	China(E)	07–16			•	no
An (2021) [25]	China(P)	09–16		•		no
Alajmi (2021) [26]	Saudi Arabia(E)	90–16			•	no
Patiño (2021) [27]	Colombia(E)	71–17			•	no
Liu (2021) [28]	China(T)	01–18			•	no
Papiez (2021) [29]	European Union(E)	96-04 05–17	•			no
Simbi (2021) [30]	Africa(E)	84-14(5)	•			no
Xia (2021) [31]	World(E)	00–17			•	no
Yang (2020) [32]	China(E)	96-16(5)	•			no
Hossain (2020) [33]	Cameroon(E)	71–14			•	no
De Oliveira (2020) [34]	Latin America(P)	90-95-03-08-15–17	•			no
Isik (2020) [35]	New York(P)	10–18			•	no
Kim (2020) [36]	OECD[d](P)	95-08-17	•			no
Sadorsky (2020) [37]	World(E)	07–17		•		no
De Oliveira (2019) [38]	Latin America(P)	90–15		•		no
Nan (2019) [39]	China(E)	00-15(5)	•			no
Chong (2019) [40]	Malaysia(E)	78-90-02-14	•			no
Chang (2019) [41]	World(E)	01–14			•	no
Ma (2018) [8]	China(E)	04–14		•		no
Zhu (2018) [42]	India(E)	07–14		•		no
Román (2018) [43]	Colombia(E)	00–15			•	no
Román (2018) [44]	Colombia(E)	90–12			•	no
Chen (2018) [45]	OECD[d](E)	01-07-09-10-15	•			no
Moutinho (2018) [46]	Europe-Top Countries(E)	86–11			•	no
Román (2018) [47]	Latin America(E)	90–13			•	no
Mousavi (2017) [48]	Iran(E)	03–14			•	no
Chong (2017) [49]	China(E)	04–14		•		no
Jiang (2017) [50]	China(E)	96-13(4)	•			no
Jiang (2017) [51]	United States(P)	90–14			•	no
Achour (2016) [52]	Tunes(T)	85–14			•	no
Zhang (2016) [53]	China(E)	95–12			•	no
Sumabat (2016) [54]	Philippine(P)	91–14			•	no
Torrie (2016) [55]	Canada(E)	95–10		•		no
Karmellos (2016) [56]	European Union(P)	00-07-12	•			no
Yang (2016) [57]	China(P)	85-10(5)	•			no
Tian (2016) [58]	Guangdong(P)	05–14			•	no
Ang (2016) [59]	Asia(P)	90–13		•		no
Ang (2016) [60]	World(P)	90–13		•		no
Mundaca (2016) [61]	World(E)	05–12		•		no
Andrés (2015) [62]	Spain(T)	97–12			•	no
Cansino (2015) [63]	Spain(E)	95–09			•	no
Chong (2015) [64]	China(E)	01-06-11	•			no
Moutinho (2015) [65]	European Union(E)	99-05-10	•			no
Fernández (2014) [66]	European Union(E)	01–08		•		no
This contribution	OECD[d] Europe(E)	00-04-06-14-17–19	•			yes

[a] Sector [E—Energy, P—Power, T—Transport, and MI—Manufacturing Industry].
[b] Time Periods - cutoff years, periodicity.
[c] Periods Selection [mp—multi-period, sp—single period, and yy—year-by-year].
[d] Organisation for Economic Co-operation and Development.

the breaking points, number of partitions, and extent of each partition. With the identification of partitions, a mp decomposition analysis was performed using the IDA-LMDI method. Finally, a validation methodology using multivariate linear regression is presented.

To start with the analysis, having a discrete time series of performance indicator of length n, $\{Y_t, t = 0, 1, 2, \ldots, n\}$, an ordered partition of m groups with size λ_i such that the sums of all λ_i are equal to n should be performed according to Eq. (1).

$$\sum_{i=1}^{m} \lambda_i = n, \{\lambda_i, m, n \in \mathbb{N}\} \tag{1}$$

According to [67], large number of breaking points for segmentation is not recommended because the model could become meaningless. Therefore, to avoid an excessive number of partitions, Eq. (2) limits the number and avoids the loss of mathematical meaning of the partition.

$$m_{max} = \frac{1}{4}(2n + (-1)^{(n+1)} - 3) \tag{2}$$

To obtain a linear regression model, the minimum size of a partition having an extension of three data, as shown in Eq. (3).

$$\lambda_{min} = 3 \tag{3}$$

According to the partitioning, the time series exhibits $m-1$ trend changes. Trend changes are identified using breakpoints τ, which are given by Eq. (4).

$$\tau = \{\tau_1, \tau_2, \dots, \tau_j, \dots, \tau_{m-1}\} = \{\{\lambda_1\}, \{\lambda_1+\lambda_2\}, \dots, \{\lambda_1+\lambda_2+\dots+\lambda_j\}, \dots, \{\lambda_1+\lambda_2+\dots+\lambda_{m-1}\}\} \quad (4)$$

Thus, the number of possible ordered combinations $q(n)$ for a time series of length n must satisfy Eqs. (1), (2), (3) and, (4). The problem of selecting appropriate periods was solved by determining the number of partitions m and the extent of each partition λ_i within a set of feasible combinations. The selected solution will have the minimum TMSE in the feasible solutions. The TMSE is calculated as the sum of the MSE of the linear regression models for each partition, as shown in Eq. (5).

$$TMSE = \sum_m MSE \quad (5)$$

The optimization problem is as follows:

$$\min_{m,\tau} TMSE = \frac{1}{n}\sum_{t=1}^{n}(\hat{Y}_t - Y_t)^2 \quad (6)$$

subject to:

$$3 \le n,\ n \in \mathbb{N} \quad (7)$$

$$3 \le \lambda_i \le n,\ \lambda_i \in \mathbb{N} \quad (8)$$

$$1 \le m \le \frac{1}{4}(2n + (-1)^{(n+1)} - 3),\ m \in \mathbb{N} \quad (9)$$

$$\sum_{i=1}^{m} \lambda_i = n \quad (10)$$

$$\hat{Y}_t = \left\{ \begin{array}{cc} \alpha_{11} + \alpha_{12}t + \xi_1 & t : t \le \tau_1 \\ \alpha_{21} + \alpha_{22}t + \xi_2 & t : \tau_1 \le t \le \tau_2 \\ \vdots & \vdots \\ \alpha_{i1} + \alpha_{i2}t + \xi_i & t : \tau_{i-2} \le t \le \tau_{i-1} \\ \vdots & \vdots \\ \alpha_{m1} + \alpha_{m2}t + \xi_m & t : \tau_{m-2} \le t \le \tau_{m-1} \\ \alpha_{(m+1)1} + \alpha_{(m+1)2}t + \xi_{m+1} & t : \tau_{m-1} \le t \end{array} \right\}. \quad (11)$$

Eq. (6) is the objective function to minimize the TMSE of the set of linear models of the time series partitions. Eq. (7) verifies that the time series has more than three data points. Eq. (8) checks that each partition has at least three time series data. Eq. (9) limits the maximum number of partitions to avoid the partition becoming meaningless. Eq. (10) guarantees that the sum of the partition lengths matches the length of the total time series. Finally, Eq. (11) represents the system of equations for the linear models of the partitions.

To date, it has been stated that it is necessary to identify the number of partitions m within a time span for each partition, considering that the trend changes in the time series. The following decomposition analysis was applied to these partitions.

We considered a time span that conveyed a given performance indicator Y. The indicator Y to be explained could be expressed as the product of several d explanatory drivers, as shown in Eq. (12).

$$Y = X_1 \cdot X_2 \cdot X_3 \cdot \dots \cdot X_k \cdot \dots \cdot X_d \quad (12)$$

where:

Y is an indicator that is decomposed or explained.

X_k denotes the ith explanatory driver,

The change in the indicator ΔY between times t_o and t_f with respect to d drivers is defined by the following additive formulation [5]:

$$\Delta Y = Y_{t_f} - Y_{t_o} = \Delta Y_{X_1} + \Delta Y_{X_2} + \Delta Y_{X_3} + \dots + \Delta Y_{X_k} + \dots + \Delta Y_{X_d} \quad (13)$$

where:

ΔY is the total change in indicator Y in the time span $t_o - t_f$ or a period.

t_o denotes the initial time.

t_f denotes the final time.

Y_{t_o} is the value of indicator Y at time t_o.

Y_{t_f} is the value of indicator Y at time t_f.

ΔY_{X_k} is the change in indicator Y related to driver X_k.

According to the IDA-LMDI method [3,5], the change in indicator ΔY_{X_k} related to the kth explanatory driver X_k is given by Eq. (14).

$$\Delta Y_{X_k} = L(Y_{t_f}, Y_{t_o}) \ln\left(\frac{X_{k,t_f}}{X_{k,t_o}}\right) \quad (14)$$

where:

X_{k,t_o} is the value of driver X_k at time t_o;

X_{k,t_f} is the value of driver X_k at time t_f.

Eq. (14) shows that the contributions of the driver X_k to the change in the indicator Y depends directly on the cut-off times t_o and t_f. The L operator for the indicator ΔY_{X_k} in Eq. (14) along the time span t_o and t_f is given by Eq. (15) [3,5].

$$L(Y_{t_f}, Y_{t_o}) = \frac{Y_{t_f} - Y_{t_o}}{\ln Y_{t_f} - \ln Y_{t_o}} \quad (15)$$

Here, the problem of partition identification and decomposition analysis for each partition in the time series of the Y indicator under analysis is explained. Notable, the objective function of Eq. (6) changes for each feasible combination. Therefore, to solve this problem, the following methodology was proposed, which combined the identification of assessment periods and their corresponding decomposition analysis.

Fig. 2 shows the proposed algorithm for the identification of a m number of assessment periods and the corresponding breaking points of the time series Y_t under analysis. The proposed method searches for feasible combinations of partitions within an analytical timeframe. Next, it performs a linear regression on each partition and calculates the MSE for each partition and the TMSE for each feasible combination. After evaluating all the feasible combinations, it selects the one with the lowest TMSE.

The algorithm shown in Fig. 2 comprises 12 steps, as follows.

Step 0: Define the time series Y_t with length of n elements (Indicator to decompose).

Step 1: Define the maximum number of partitions of Eq. (2).

Step 2: Find the set of feasible combinations $q(n)$, considering Eqs. (3), (7), (8), (9) and (10).

Step 3: Initialize the feasible combination evaluation.

Step 4: Select feasible combinations from the set.

Step 5: Solve the regression models given in Eq. (11).

Step 6: Calculate the MSE and compute the TMSE.

Step 7: A new combination is taken from the feasible set and the process is repeated until the set is complete.

Step 8: Add an iteration if the whole feasible set is not yet evaluated.

Step 9: Identify the partition combination with lowest TMSE of the feasible set.

Step 10: Identify the number of partition m, the breaking points τ_j, and the λ_i lengths of the solutions.

Step 11: Perform the multi-period IDA-LMDI decomposition using the identified partitions from Eqs. (12)–(15).

Multivariate linear regression was used to validate the results of the decomposition analysis. Consider d explanatory variables X_k, $k = 1, \dots, d$, and an output or explained variable Y; a multivariate linear regression model with consistent estimators is given Eq. (16):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + \dots + \beta_d X_d + \varepsilon \quad (16)$$

By evaluating the multivariate regression model using Eq. (16) at a time t_o and t_f, we obtain Eqs. (17) and (18):

$$Y_{t_o} = \beta_0 + \beta_1 X_{1,t_o} + \beta_2 X_{2,t_o} + \beta_3 X_{3,t_o} + \dots + \beta_k X_{k,t_o} + \dots + \beta_d X_{d,t_o} + \varepsilon_{t_o} \quad (17)$$

$$Y_{t_f} = \beta_0 + \beta_1 X_{1,t_f} + \beta_2 X_{2,t_f} + \beta_3 X_{3,t_f} + \dots + \beta_k X_{k,t_f} + \dots + \beta_d X_{d,t_f} + \varepsilon_{t_f} \quad (18)$$

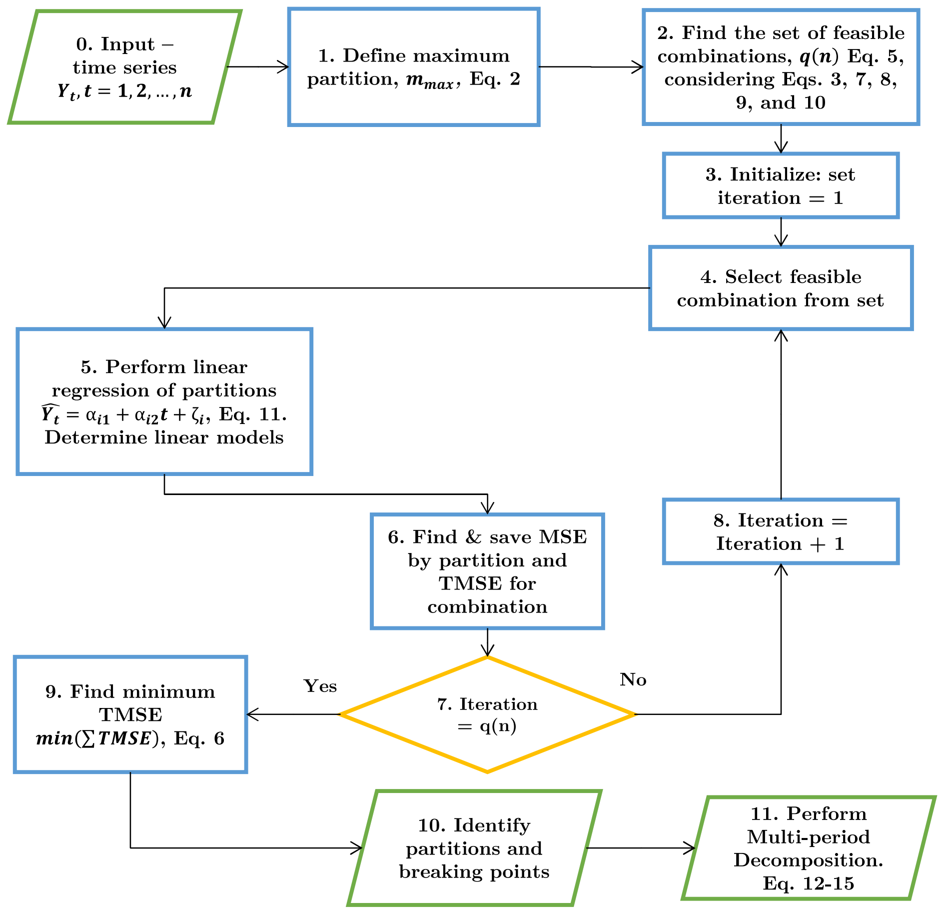

Fig. 2. Proposed algorithm to select the number of periods.

The change in the output variable Y is given by Eqs. (19) or (20):

$$\Delta Y = Y_{t_f} - Y_{t_o} = \beta_1 \Delta X_1 + \beta_2 \Delta X_2 + \beta_3 \Delta X_3 + \cdots + \beta_k \Delta X_k + \cdots + \beta_d \Delta X_d + \varepsilon \quad (19)$$

$$\Delta Y = Y_{t_f} - Y_{t_o} = \Delta Y_{X_1} + \Delta Y_{X_2} + \Delta Y_{X_3} + \cdots + \Delta Y_{X_k} + \cdots + \Delta Y_{X_d} + \varepsilon \quad (20)$$

Eqs. (19) and (20) show that the change in indicator Y is related to driver X_k given by Eq. (21).

$$\Delta Y_{X_k} = \beta_k \Delta X_k = \beta_k (X_{k,t_f} - X_{k,t_o}) \quad (21)$$

The terms of change, ΔY_{X_k} where X_k is an explanatory driver, could be approximated by multivariate linear regression analysis using Eq. (21) and also by IDA-LMDI decomposition analysis by using Eq. (14). Taking the regressors and changes ΔY_{X_k} obtained by multivariate regression analysis as the references, the mean squared error of the drivers (DMSE) between the drivers obtained with decomposition analysis and those obtained with regression analysis can be given by Eq. (22):

$$DMSE = \frac{1}{d} \sum_{k=1}^{d} (\Delta \hat{Y}_{X_k} - \Delta Y_{X_k})^2 \quad (22)$$

where:

d denotes the number of drivers obtained.

$\Delta \hat{Y}_{X_k}$ is the change in Y related to driver X_k obtained from the LMDI decomposition analysis.

ΔY_{X_k} is the change in Y related to driver X_k obtained using multivariate linear regression analysis.

The analysis of Eq. (17) to Eq. (22) could be extended for all time periods of the partitions of a time series. Finally, the total mean square error of the drives (TDMSE) for the time series was calculated using Eq. (23).

$$TDMSE = \sum_{m} DMSE \quad (23)$$

3. Results

This section presents the results of the proposed methodology and the validation results. The results of the identification of the partitions and their corresponding decomposition results using the IDA-LMDI method are presented. Two case studies were analyzed: the first was an illustrative case observing the differences in the results when both arbitrary and non-arbitrary selection of assessment periods is taken. In addition, this case showed statistical consistency in the data, and therefore allowed us to obtain a valid linear regression model. The linear model enabled the validation of the results and the comparison of arbitrary and non-arbitrary selection of assessment periods. The second case was a real case of the decomposition of Kaya factors for total carbon dioxide emissions of Organisation for Economic Co-operation and Development (OECD) countries in Europe. This case showed the importance of the proper selection of assessment periods for decomposition analysis using the IDA-LMDI method, which was key to identifying the drivers that guide energy policy decisions.

3.1. Results—illustrative test case

This case was an illustrative example developed to apply and validate the proposed methodology, and the results were obtained once the period selection was performed. The data in the illustrative case are self-compiled and with no practical or economic significance, but with

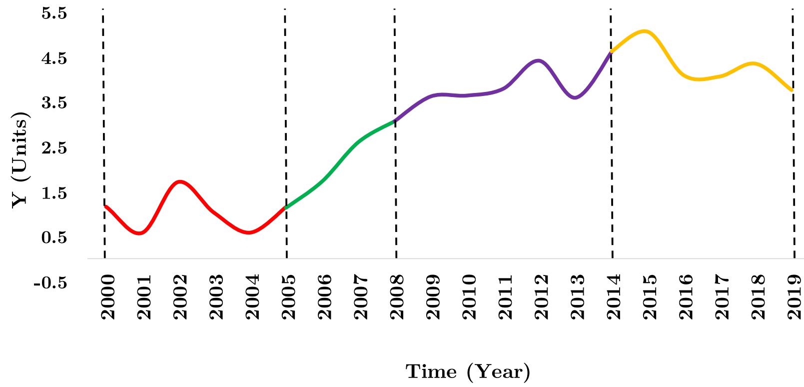

Fig. 3. Partition algorithm results—illustrative case.

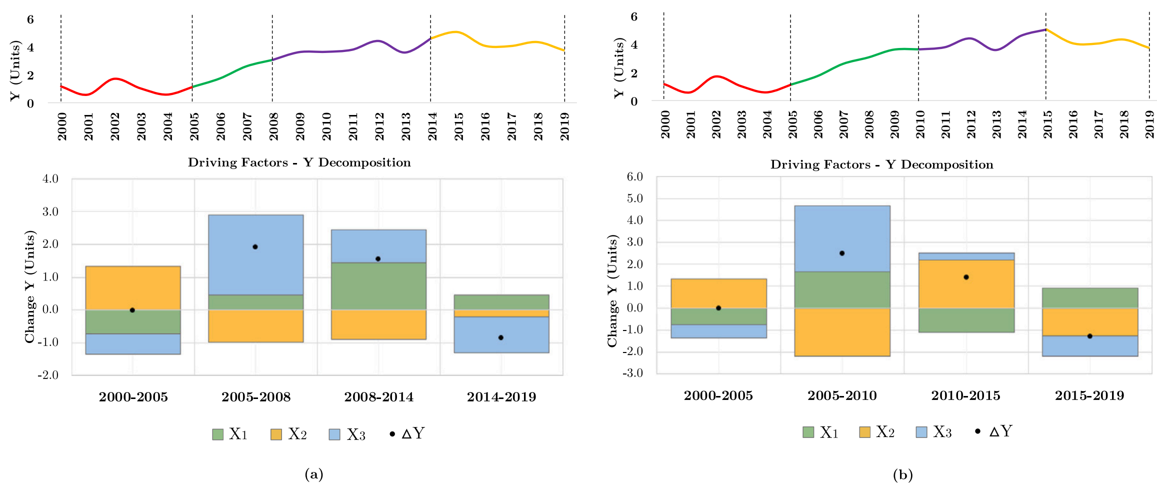

Fig. 4. Decomposition results—illustrative case. (a) Non-arbitrary selection periods; (b) arbitrary selection periods.

consistency for unbiased decomposition and statistical analysis. The use of consistent data allows the validation by econometric analysis, since in this case it is required that estimators are unbiased, reliable, and meet the assumptions of the linear model. In this experiment, the results were obtained using decomposition analysis and compared with those obtained using linear regression analysis. A sample of 20 observations considering one dependent variable Y and three independent variables X_1, X_2 and X_3 was used. The year 2000 was set as the base year. The data set is presented in Appendix C.

The proposed algorithm was applied with a maximum number of partitions of $m_{max} = 8$. The set of feasible combinations identified is $q(20) = 2584$. The algorithm was performed on a PC computer with the following characteristics: 2.90 GHz Intel Core i7-5500U CPU and 8.00 GB (RAM). The algorithm was coded in Python, and the CPU time was 11.2 s. The results obtained for the partitions with the lowest TMSE are shown in Fig. 3.

According to Fig. 3, $m = 4$ partitions with lengths of $\lambda_1 = 5$, $\lambda_2 = 3$, $\lambda_3 = 6$, and $\lambda_4 = 5$ have been obtained. The breaking points identified were $\tau = \{2005, 2008, 2014\}$. The TMSE obtained for the selected partition is $TMSE = 0.32901$.

IDA-LMDI decomposition, as specified by Eqs. (12) to (15) was carried out using the selected periods shown in Fig. 3. The resulting explanatory drivers are presented in Figs. 4(a). The resulting IDA-LMDI decomposition considering four 5-year fixed arbitrary periods, is shown in Fig. 4(b).

We observed that a slight variation in the selection of periods led to different results for the drivers. Notably, the drivers changed in magnitude and in their coupling and decoupling with the change in the Y variable. This was evident when the last three decomposition periods were evaluated. However, one method of selecting the periods was better than the other. An analysis was performed using the classical multivariate linear regression method.

Applying Eqs. (16) to (23), the estimators were obtained to calculate the contributions of each driver to the change in the variable Y. The estimators must be statistically consistent and the model must be homoscedastic, with no autocorrelation of the residuals and low multicollinearity. Statistical hypothesis tests were conducted to verify these characteristics. In particular, the Breusch–Pagan test was used to verify homoscedasticity, the Durbin–Watson test for correlation, and the Variance Inflation Factor test to evaluate multicollinearity. As this example satisfied these assumptions that statistical confidence in the estimators was obtained. In addition to model consistency, the estimators must have statistical significance. For this analysis, R software was used to obtain the linear model and carry out the hypothesis and inference tests. The dataset included in Appendix C was used. The results are shown in Table 2.

The intercepts and estimators were statistically significant. In all cases, a p-value less than 5% of the significance level was obtained. The F-statistic and its p-value showed that the model as a whole had statistical significance. The model also showed a high degree of fit, with an adjusted R^2 coefficient of 0.9549. The Breusch–Pagan and Durbin–Watson statistics and their respective p-values indicated that the model was homoscedastic and showed no correlation in the residuals. The null hypothesis of each test is indicated for ease of understanding. In the multicollinearity test, variance inflation factors of less than 10 indicated the presence of low multicollinearity. Based on these results, a consistent linear model for the data is given by Eq. (24).

$$Y = -3.34697 + 2.10830X_1 + 1.24276X_2 + 0.92735X_3 + \varepsilon \tag{24}$$

Table 2
Results for linear model—illustrative case.

Residuals:				
Min	1Q	Median	3Q	Max
−0.57510	−0.17072	0.01078	0.20209	0.43980
Coefficients:				
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	−3.34697	0.87332	−3.832	0.00147
X_1	2.10830	0.98903	2.132	0.04888
X_2	1.24273	0.27747	4.479	0.00038
X_3	0.92735	0.06567	14.122	1.88E−10
Residual standard error: 0.312 on 16 degrees of freedom				
Multiple R-squared: 0.9621, Adjusted R-squared: 0.9549				
F-statistic: 135.2 on 3 and 16 DF, p-value: 1.412e−11				
Assumption	Test	Value	p-value	H_0
Homocedasticity	Breusch–Pagan	$BP = 0.73372$	0.8652	Homoscedastic Model
Auto-correlation	Durbin–Watson	$DW = 2.5738$	0.8657	No Auto-correlation
Multicollinearity	Variance Inflation Factor	$X_1 = 2.820535$	$X_2 = 2.277111$	$X_3 = 1.589151$

Table 3
Driver calculation results – four 5-year periods – illustrative case.

Period	Decomposition			Regression		
	$\Delta\hat{Y}_{X_1}$	$\Delta\hat{Y}_{X_2}$	$\Delta\hat{Y}_{X_3}$	ΔY_{X_1}	ΔY_{X_2}	ΔY_{X_3}
2000–2005	−0.738	1.337	−0.609	−0.826	1.908	−0.751
2005–2010	1.656	−2.193	3.028	1.064	−1.775	3.313
2010–2015	−1.102	2.192	0.325	−0.449	0.670	0.344
2015–2019	0.903	−1.254	−0.951	0.353	−0.419	−0.927
TDMSE	4.703					

Table 4
Driver calculation results – proposed selection periods method – illustrative case.

Period	Decomposition			Regression		
	$\Delta\hat{Y}_{X_1}$	$\Delta\hat{Y}_{X_2}$	$\Delta\hat{Y}_{X_3}$	ΔY_{X_1}	ΔY_{X_2}	ΔY_{X_3}
2000–2005	−0.738	1.337	−0.609	−0.826	1.908	−0.751
2005–2008	0.453	−0.987	2.449	0.241	−1.103	2.711
2008–2014	1.434	−0.897	1.015	0.537	−0.354	1.165
2014–2019	0.443	−0.212	−1.096	0.191	−0.066	−1.147
TDMSE	1.692					

Table 5
Driver calculation results – five 4-year periods – illustrative case.

Period	Decomposition			Regression		
	$\Delta\hat{Y}_{X_1}$	$\Delta\hat{Y}_{X_2}$	$\Delta\hat{Y}_{X_3}$	ΔY_{X_1}	ΔY_{X_2}	ΔY_{X_3}
2000–2004	−0.136	−0.280	−0.164	−0.262	−0.249	−0.325
2004–2008	−0.365	1.472	1.378	−0.323	1.053	2.285
2008–2012	1.074	−0.161	0.441	0.393	−0.072	0.481
2012–2016	0.424	−1.443	0.683	0.164	−0.463	0.746
2016–2019	0.372	0.374	−1.077	0.172	0.115	−1.207
TDMSE	2.671					

Because the linear model was consistent, the contributions of each driver to the selection of arbitrary and non-arbitrary periods could be calculated using Eqs. (14) and (21). Using the drivers obtained by the decomposition analysis and the estimators obtained by the multivariate regression analysis, the TDMSE was calculated using Eqs. (22) and (23) and separately for the arbitrary and non-arbitrary periods. The results are summarized in Tables 3 and 4. Table 5 presents additional results obtained for the arbitrary selection of five 4-year periods.

From the results obtained, we must note that the resulting TDMSE values were significantly reduced with the non-arbitrary selection of decomposition periods. This meant that the drivers thus obtained were closer to those obtained by classical econometric methods, implying a lower bias in the results. This result highlighted the importance of proper period selection. Notably, with small changes in the periods, the results of the drivers and coupling varied significantly. Using classical regression analysis, we demonstrated that the selection of decomposition periods helped bias reduction.

3.2. Results - OECD-Europe case

The proposed method was also applied to a case study reported by the International Energy Agency (IEA) [68]. Carbon dioxide CO_2 emissions from the 2000–2019 time series for European OECD countries were analyzed by determining the Kaya factors and using the IDA-LMDI decomposition methodology presented in Appendix B [69]. The validation approach with econometric analysis used with the illustrative case is not used for this case because the data present inconsistency in the fulfillment of the assumptions of the linear model. The data have autocorrelation in the residuals and high multicollinearity; as shown in Appendix D. Then, the identification of the decomposition periods is performed directly with the proposed methodology. In [68] the IEA presented yy cumulative decomposition and mp results. For the mp analysis, arbitrary sub-periods of 2000–2010 and 2010–2017 were taken. The results of the IEA's decomposition analysis are presented in Fig. 5 and Table 6.

Table 6
Driver calculation results – arbitrary selection periods – OECD-Europe case [$MtCO_2$].

Period	Population	GDP PPP[a]/Population	TPES[b]/GDP PPP[a]	CO_2/TPES[b]	Change CO_2
2000–2010	208.49	408.40	−453.43	−254.51	−91.06
2010–2017	117.25	369.46	−635.99	−140.01	−289.29

[a] Purchasing power parity-adjusted gross domestic product.
[b] Total primary energy source.

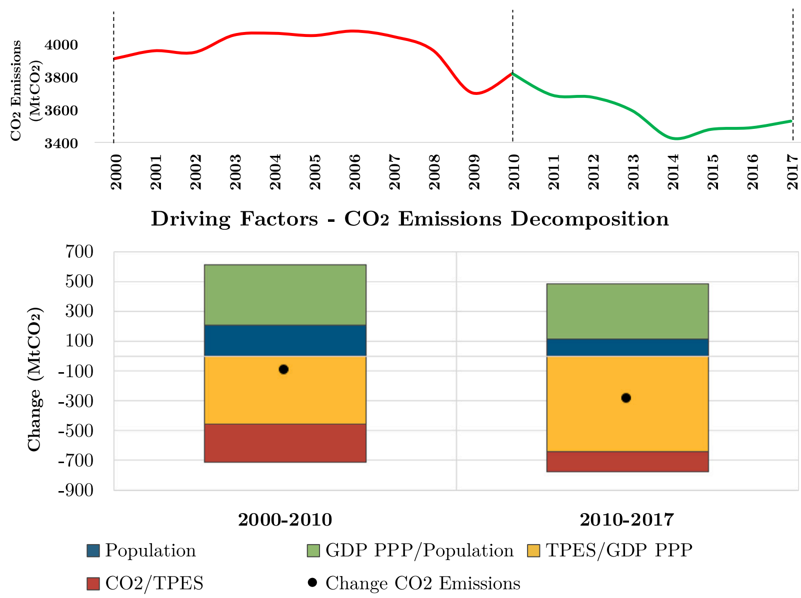

Fig. 5. Decomposition analysis results – arbitrary selection periods – OECD-Europe case.

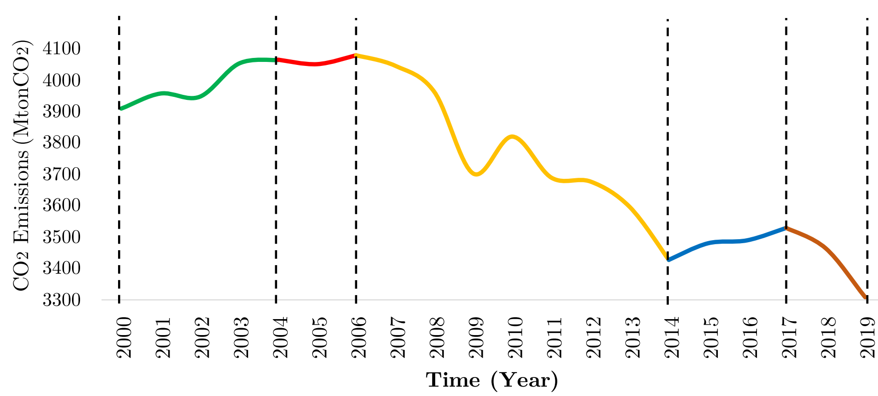

Fig. 6. Partition algorithm results—OECD-Europe case.

The proposed algorithm was applied to the OECD-Europe case on the time series of the total CO_2 emissions with $n = 20$ observations; the maximum number of partitions was $m_{max} = 8$. The set of feasible combinations identified was $q(20) = 2584$. The CPU time was 10.9 s. The results obtained for the partitions with the lowest TMSE are shown in Fig. 6.

According to Fig. 6, $m = 5$ partitions with lengths of $\lambda_1 = 4$, $\lambda_2 = 2$, $\lambda_3 = 8$, $\lambda_4 = 3$ and $\lambda_5 = 2$ have been obtained. The years in the identified breaking points are $\tau = \{2004, 2006, 2014, 2017\}$. The TMSE obtained for the selected partition was $TMSE = 4.95$. The following five periods were identified:

1. 2000–2004, increase in CO_2 emissions from 3910.0 to 4064.2 [$MtCO_2$].
2. 2004–2006, slight increase in CO_2 emissions from 4064.2 to 4077.8 [$MtCO_2$].
3. 2006–2014, strong reduction in CO_2 emissions from 4077.8 to 3426.1 [$MtCO_2$].
4. 2014–2017, increase in CO_2 emissions from 3426.1 to 3529.6 [$MtCO_2$].
5. 2017–2019, reduction in CO_2 emissions from 3529.6 to 3307.6 [$MtCO_2$].

The results of the IDA-LMDI decomposition analysis (Eq. (25) to (31) in Appendix B) in the selected five periods are shown in Fig. 7 and Table 7.

Fig. 7 shows the differences between the proposed method and the arbitrary selection of decomposition periods defined by the IEA. This solution provides much more relevant information in the non-arbitrary case.

The period selection methodology captured more relevant information for identifying the drivers of change. The most significant changes in the behavior of carbon dioxide emissions were observed within the

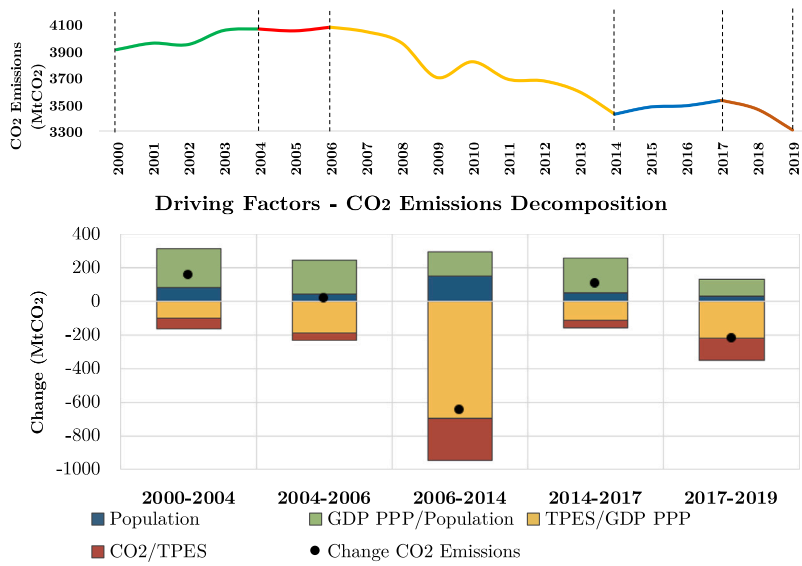

Fig. 7. Decomposition analysis results – non-arbitrary selection periods – OECD-Europe case.

Table 7
Driver calculation results – non-arbitrary selection periods - OECD-Europe case [$MtCO_2$].

Period	Population	GDP PPP[a]/Population	TPES[b]/GDP PPP[a]	CO_2/TPES[b]	Change CO_2
2000–2004	80.19	235.33	−102.14	−59.14	154.24
2004–2006	45.53	198.57	−186.69	−43.80	13.60
2006–2014	150.61	143.44	−695.73	−250.01	−651.69
2014–2017	49.87	209.21	−115.26	−40.31	103.51
2017–2019	29.66	100.99	−221.03	−131.60	−221.98

[a] Purchasing power parity-adjusted gross domestic product.
[b] Total primary energy source.

arbitrary periods of 2000–2010 and 2010–2017. In the first years of 2000–2010, there was an increase in CO_2 emissions, and from 2016 onwards, there was a significant decrease; however, the change reported for that period was only −91.06 $MtCO_2$. For 2010–2017, there was a decrease in CO_2 emissions. However, from 2014, there was an increase in emissions that was not captured within the arbitrary period. With respect to the drivers, it was observed that energy intensity and carbon intensity boosted the reduction in CO_2 emissions, whereas population and purchasing power parity-adjusted gross domestic product (GDP-PPP) per capita boosted the increase. However, none of the drivers showed a dominant behavior in the change in CO_2 emissions.

In contrast, the results obtained with a non-arbitrary selection of periods allowed us to observe important facts. First, in three of the five identified periods, an increase in CO_2 emissions was observed, specifically during 2000–2004, 2004–2006, and 2014–2017. In all three cases, the driver of growth was per capita income. These increases could not have been identified if the periods were not appropriately selected. Another significant result was for 2006–2014; this period shows a significant reduction in CO_2 emissions of −651.69 $MtCO_2$. This change was primarily related to the energy intensity driver and contributed significantly to the change in carbon intensity. This important result could not be observed with the arbitrary selection of periods provided by the IEA, as the change was diluted between two arbitrarily selected periods. Finally, data were added for 2017–2019, and the algorithm took this as a new period in which there was a significant trend change in CO_2 emissions, again due to energy intensity as the main driver of change. In this context, it was evident that the arbitrary selection of the assessment periods led to an overestimation or underestimation of the explanatory drivers, and to reduce bias, it was necessary to select the assessment periods appropriately.

From a general point of view, it could be observed that the behavior of CO_2 emissions in the OECD countries of Europe has been climbing, with a significant change observed for the period 2006–2014. In recent years, there has been an increasing trend, but the overall trend has been towards a reduction. The drivers of this reduction have mainly been energy and carbon intensity, which counterbalanced the increases associated with population and economic growth. Thus, it can be concluded that energy and climate change policies in this region should continue to focus on reducing the carbon intensity by integrating clean and low-carbon technologies that allow the production of more energy with lower CO_2 emissions. However, in addition to decarbonization, it is important to focus on energy intensity and efficiency, as the results showed that the great progress in reducing CO_2 emissions in OECD countries in Europe was due to these factors. For this reason, the design of public policies should be oriented not only towards the decarbonization of the energy sector but also towards a more efficient use of energy resources and the production of wealth with less energy consumption.

4. Discussion

It is observed that the time frame of the proposed methodology has advantages over the other reported approaches. The proposed methodology can capture significant inter-annual information from the drivers unlike the fixed n-year single-period approach. The proposed approach does not need complete detailed information for all years like the year-by-year (yy) approach, it only requires detailed information at identified breakpoints. In addition, defining a smaller number of analyzed periods facilitates the interpretation of the results and this is an advantage when performing analyses with different levels of disaggregation. Compared to the n-year multi-period (mp) approach, the two cases analyzed in this paper showed that the arbitrary selection of fixed periods of four or five years is not the most appropriate. This is because biases can occur and there is still loss of important inter-annual information that can lead to overestimate or underestimate drivers and decoupling.

The proposed methodology showed to be a complement to improve the decomposition analysis with IDA techniques in the methods of the Divisia family; mainly the IDA-LMDI techniques. It can be used to analyze case studies at city, country, regional and global levels. In addition, it is suitable for analyzing cases of emissions, energy, and intensity decomposition of different sectors. In the case of countries or cities, the technique can capture information in cases of aggregated or disaggregated analysis at the sectoral level. In the regional case it is suitable for analyzing aggregated information such as the example reported in this paper of CO_2 emissions from the OECD countries of Europe. However, for disaggregated analysis at the regional level it is required to consider the time series of each country that address specific economic and energy policy conditions.

5. Conclusions

This paper presents a new methodology for determining the appropriate selection of assessment periods to perform decomposition analysis in Divisia Index studies. The proposed method is based on the selection of non-arbitrary periods using an MSE minimization algorithm in a linear regression model. This method allows the identification of breaking points, number of partitions, and partition extensions in a time series. This selection considers trend changes and allows for the determination of significant periods prior to the decomposition analysis using any Divisia Index-based decomposition method. In this study, the IDA-LMDI was employed to test the proposed method. A multivariate linear regression model with consistent estimators was used to validate the results.

The proposed methodology was illustrated in a simple test case that showed the differences in the resulting drivers when arbitrary and nonarbitrary selections of assessment periods were made. The results and methodology were validated using a multivariate linear regression model with consistent estimators that fulfilled homoscedasticity, had no autocorrelation of residuals, and had low multicollinearity. This validation allowed the verification of the calculation of the drivers and showed that the non-arbitrary selection of the evaluation periods presented a lower MSE of the drivers than in the case of arbitrary selection. This implies that bias in the results due to the arbitrary selection of periods would be reduced if selected in a non-arbitrary manner.

The methodology was also applied to the Kaya factors of carbon dioxide emissions of OECD countries in Europe in the time-series period 2000–2019. The results were compared with those reported by the IEA for two fixed periods. They showed that it was necessary to define five different periods to capture drivers that could truly explain the changes in carbon emissions during the assessment period. The use of only two periods reported by the IEA led to the overestimation and underestimation of explanatory drivers, with important policy implications, such as the promotion of laws to adopt new technologies. Instead of the two periods selected by the IEA, the adoption of five periods provided a more accurate explanation of the changes in carbon dioxide emissions in OECD-European countries. It could be concluded that energy policy efforts should continue decarbonization, while also placing greater focus on energy intensity and efficiency.

For future work, the use of the methodology can be explored with disaggregated application cases at the sectoral level. On the other hand, the proposed technique evaluates combinations of feasible solutions, so the computational cost will increase as the length of the analyzed time series increases. For this reason, other methodological approaches for time series segmentation can also be explored in the framework of decomposition analysis with IDA methods.

CRediT authorship contribution statement

Juan David Rivera-Niquepa: Conceptualization, Research, Methodology, Writing – original draft. **Daniela Rojas-Lozano:** Conceptualization, Methodology. **Paulo M. De Oliveira-De Jesus:** Conceptualization, Supervision, Writing – review & editing. **Jose M. Yusta:** Supervision, 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.

Data availability

Data will be made available on request.

Appendix A. Nomenclature

See Table 8.

Appendix B. Detailed IDA-LMDI decomposition formulation for the additive approach

The total carbon emissions CO_2 as a function of the four drivers in Kaya identity are given by (25).

$$C = P \cdot \frac{GDP}{P} \cdot \frac{TPES}{GDP} \cdot \frac{C}{TPES} = p \cdot g \cdot u \cdot e \tag{25}$$

where:

C is the total CO_2 emissions from fuel combustion [$MtCO_2$].

P is the total population [Num people].

GDP is the Gross domestic product of purchasing power [USD 2015].

$TPES$ is the total primary energy supply [TJ].

The decomposition drivers are given by

$p = P$ is the total population [Num people].

$g = GDP/P$ is the Per capita gross domestic product of purchasing power [USD 2015].

$u = TPES/GDP$ is the energy intensity [TJ/USD in 2015].

$e = C/TPES$ is the carbon intensity of fuel combustion [$MtCO_2$/TJ].

The change in total carbon emissions CO_2 related to the change in drivers is given by (26).

$$\Delta C = C_{t_f} - C_{t_o} = \Delta C_p + \Delta C_g + \Delta C_u + \Delta C_e \tag{26}$$

where:

t_o is the initial year of the analysis sub-period.

t_f is the final year of the analysis sub-period.

The contribution of each driver to the change in total carbon emissions CO_2 in the analysis sub-period is obtained using (27)–(30).

$$\Delta C_p = L(C_{t_f}, C_{t_o}) \ln(\frac{p_{t_f}}{p_{t_o}}) \tag{27}$$

$$\Delta C_g = L(C_{t_f}, C_{t_o}) \ln(\frac{g_{t_f}}{g_{t_o}}) \tag{28}$$

$$\Delta C_u = L(C_{t_f}, C_{t_o}) \ln(\frac{u_{t_f}}{u_{t_o}}) \tag{29}$$

$$\Delta C_e = L(C_{t_f}, C_{t_o}) \ln(\frac{e_{t_f}}{e_{t_o}}) \tag{30}$$

The operator L is given by (31).

$$L(a,b) = \frac{a-b}{\ln a - \ln b} \tag{31}$$

Appendix C. Data for the illustrative validation case

See Table 9.

Table 8
Nomenclature.

Indices		ξ_i	Regression error of the ith partition
i	Partition number	τ_j	jth breaking point
j	Breaking point number	ΔX_k	Change in the kth driver X
k	Explanatory driver number	ΔY	Total change in the indicator or dependent variable Y
t	Time	ΔY_{X_k}	Change in the indicator related to the kth driver X
		d	Number of explanatory drivers
Variables		m	Number of partitions
α_{i1}	Intercept of linear model of the ith partition	m_{max}	Maximum number of partitions
α_{i2}	Slope of the linear model of the ith partition	n	Length of time series under analysis
β_k	kth estimator of the multivariate linear model	$\mathbb{N}$	Natural numbers set
ε	Regression error of linear model	q	Number of feasible combinations
λ_i	Length of ith partition	Y_i	ith dependent or explained variable
λ_{min}	Minimum number of elements allowed	$\hat{Y}_i$	ith estimated from linear model

Table 9
Data for the illustrative validation case.

n	Year	Y	X_1	X_2	X_3
1	2000	1.16268529	0.8308618	0.70633709	1.98116831
2	2001	0.57676082	0.79972266	0.34029694	2.11932866
3	2002	1.71222583	0.76584856	1.05700528	2.91810797
4	2003	1.02963212	0.73569378	1.00062766	1.39866109
5	2004	0.58306535	0.7066331	0.50609042	1.63040363
6	2005	1.15267147	0.43915731	2.2417318	1.17085139
7	2006	1.73704614	0.74448907	0.89090252	2.61892365
8	2007	2.60606449	0.69443857	1.07593012	3.48792598
9	2008	3.06776717	0.55350844	1.35378348	4.09401035
10	2009	3.62330808	0.85519586	1.05798292	4.00461726
11	2010	3.64344345	0.94395719	0.81380006	4.74287868
12	2011	3.79382567	0.70465189	1.40766889	3.8247428
13	2012	4.42195243	0.73969248	1.2961226	4.61229185
14	2013	3.59473157	0.73653368	0.94885909	4.43393811
15	2014	4.62008214	0.80807058	1.06855066	5.35063448
16	2015	5.05751001	0.73105311	1.35285598	5.11371191
17	2016	4.08693243	0.81730291	0.92324437	5.41623806
18	2017	4.05870979	0.76517657	1.18748173	4.46682973
19	2018	4.34449198	0.97924208	0.85282303	4.96274908
20	2019	3.75527757	0.89865847	1.01568613	4.11422324

Table 10
Linear model results- OECD-Europe case.

Coefficients:				
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.863E+03	4.307E+03	0.897	0.375
Population	−8.000E−06	6.367E−06	−1.257	0.216
GDPP	1.129E−01	2.052E−02	5.503	1.80E−06
Ene_Int	1.024E+09	2.166E+08	4.728	2.36E−05
Carb_Int	−7.195E+07	1.517E+07	−4.743	2.24E−05
Residual standard error: 121.9 on 44 degrees of freedom Multiple R-squared: 0.649, Adjusted R-squared: 0.6171 F-statistic: 20.34 on 4 and 44 DF, p-value: 1.514e−09				
Assumption	Test	Value	p-value	H_0
Homoscedasticity	Breusch–Pagan	BP = 8.5398	0.07369	Homocedastic
Autocorrelation	Durbin–Watson	DW = 0.28644	2.20E−16	No Auto-corr
Multicollinearity	Variance Inflation Factor	Population = 198.679 Ene_Int = 160.589	GDPP = 65.543 Carb_Int = 35.918	

Appendix D. Linear model results - OECD-Europe case

See Table 10.

References

[1] P.F. González, M. Landajo, M.J. Presno, The Driving Forces of Change in Environmental Indicators: An Analysis Based on Divisia Index Decomposition Techniques, Springer, 2014.

[2] X. Xu, B.W. Ang, Index decomposition analysis applied to CO2 emission studies, Ecol. Econom. 93 (2013) 313–329.

[3] B.W. Ang, LMDI decomposition approach: A guide for implementation, Energy Policy 86 (2015) 233–238.

[4] B.W. Ang, Decomposition methodology in industrial energy demand analysis, Energy 20 (11) (1995) 1081–1095.

[5] B.W. Ang, F.Q. Zhang, K.-H. Choi, Factorizing changes in energy and environmental indicators through decomposition, Energy 23 (6) (1998) 489–495.

[6] B.W. Ang, F.Q. Zhang, A survey of index decomposition analysis in energy and environmental studies, Energy 25 (12) (2000) 1149–1176.

[7] T. Goh, B. Ang, X. Xu, Quantifying drivers of CO2 emissions from electricity generation–Current practices and future extensions, Appl. Energy 231 (2018) 1191–1204.

[8] L. Ma, C. Chong, X. Zhang, P. Liu, W. Li, Z. Li, W. Ni, LMDI decomposition of energy-related CO2 emissions based on energy and CO2 allocation Sankey diagrams: The method and an application to China, Sustainability 10 (2) (2018) 344.

[9] S. Jamali, P. Jönsson, L. Eklundh, J. Ardö, J. Seaquist, Detecting changes in vegetation trends using time series segmentation, Remote Sens. Environ. 156 (2015) 182–195.

[10] J. Li, Z.-L. Li, H. Wu, N. You, Trend, seasonality, and abrupt change detection method for land surface temperature time-series analysis: Evaluation and improvement, Remote Sens. Environ. 280 (2022) 113222.
[11] X. Yang, F. Yu, W. Pedrycz, Z. Li, Clustering time series under trend-oriented fuzzy information granulation, Appl. Soft Comput. 141 (2023) 110284.
[12] X. Song, M. Shi, J. Wu, W. Sun, A new fuzzy c-means clustering-based time series segmentation approach and its application on tunnel boring machine analysis, Mech. Syst. Signal Process. 133 (2019) 106279.
[13] Y. Lin, L. Ma, Z. Li, W. Ni, The carbon reduction potential by improving technical efficiency from energy sources to final services in China: An extended Kaya identity analysis, Energy 263 (2023) 125963.
[14] Y. Huang, Y. Wang, J. Peng, F. Li, L. Zhu, H. Zhao, R. Shi, Can China achieve its 2030 and 2060 CO2 commitments? Scenario analysis based on the integration of LEAP model with LMDI decomposition, Sci. Total Environ. 888 (2023) 164151.
[15] T. Chun, S. Wang, X. Xue, H. Xin, G. Gao, N. Wang, X. Tian, R. Zhang, Decomposition and decoupling analysis of multi-sector CO2 emissions based on LMDI and Tapio models: Case study of Henan Province, China, Environ. Sci. Pollut. Res. (2023) 1–16.
[16] Y. Wang, S. Jeong, Y. Hang, Q. Wang, Determinants of net energy-related CO2 emissions in China: A source-to-sink decomposition analysis, Environ. Impact Assess. Rev. 98 (2023) 106979.
[17] J. Wang, K. Dong, G. Hochman, G.R. Timilsina, Factors driving aggregate service sector energy intensities in Asia and Eastern Europe: A LMDI analysis, Energy Policy 172 (2023) 113315.
[18] J.D. Rivera-Niquepa, D. Rojas-Lozano, P.M. De Oliveira-De Jesus, J.M. Yusta, Decomposition analysis of the aggregate carbon intensity (ACI) of the power sector in Colombia—A multi-temporal analysis, Sustainability 14 (20) (2022) 13634.
[19] R. Li, Z. Chen, J. Xiang, A region-scale decoupling effort analysis of carbon dioxide emissions from the perspective of electric power industry: A case study of China, Environ. Dev. Sustain. (2022) 1–26.
[20] W. Li, Z. Ji, F. Dong, Spatio-temporal analysis of decoupling and spatial clustering decomposition of CO2 emissions in 335 Chinese cities, Sustainable Cities Soc. 86 (2022) 104156.
[21] P.F. González, M.J. Presno, M. Landajo, Tracking the change in Spanish greenhouse gas emissions through an LMDI decomposition model: A global and sectoral approach, J. Environ. Sci. (2022).
[22] A. Faridzad, M.G. Ghadim, CO2 intensity decomposition analysis in the Netherlands' manufacturing industry: An application of monetary and physical indicators, Environ. Dev. Sustain. (2022) 1–19.
[23] Y. He, Y. Xing, X. Zeng, Y. Ji, H. Hou, Y. Zhang, Z. Zhu, Factors influencing carbon emissions from China's electricity industry: Analysis using the combination of LMDI and K-means clustering, Environ. Impact Assess. Rev. 93 (2022) 106724.
[24] W. Chen, Q. Zhang, Z. Gao, Y. Geng, Y. Cheng, X. Tian, Exploring the drivers of energy-related CO2 emissions in western China: a case study of Haixi, Environ. Dev. Sustain. (2022) 1–15.
[25] Y. An, D. Zhou, Q. Wang, Carbon emission reduction potential and its influencing factors in China's coal-fired power industry: a cost optimization and decomposition analysis, Environ. Dev. Sustain. 24 (3) (2022) 3619–3639.
[26] R.G. Alajmi, Factors that impact greenhouse gas emissions in Saudi Arabia: Decomposition analysis using LMDI, Energy Policy 156 (2021) 112454.
[27] L.I. Patiño, V. Alcántara, E. Padilla, Driving forces of CO2 emissions and energy intensity in Colombia, Energy Policy 151 (2021) 112130.
[28] M. Liu, X. Zhang, M. Zhang, Y. Feng, Y. Liu, J. Wen, L. Liu, Influencing factors of carbon emissions in transportation industry based on CD function and LMDI decomposition model: China as an example, Environ. Impact Assess. Rev. 90 (2021) 106623.
[29] M. Papież, S. Śmiech, K. Frodyma, The role of energy policy on the decoupling processes in the European union countries, J. Clean. Prod. 318 (2021) 128484.
[30] C.H. Simbi, J. Lin, D. Yang, J.C. Ndayishimiye, Y. Liu, H. Li, L. Xu, W. Ma, Decomposition and decoupling analysis of carbon dioxide emissions in african countries during 1984–2014, J. Environ. Sci. 102 (2021) 85–98.
[31] Q. Xia, H. Wang, X. Liu, X. Pan, Drivers of global and national CO2 emissions changes 2000–2017, Clim. Policy 21 (5) (2021) 604–615.
[32] J. Yang, W. Cai, M. Ma, L. Li, C. Liu, X. Ma, L. Li, X. Chen, Driving forces of China's CO2 emissions from energy consumption based on Kaya-LMDI methods, Sci. Total Environ. 711 (2020) 134569.
[33] M.A. Hossain, J. Engo, S. Chen, The main factors behind Cameroon's CO 2 emissions before, during and after the economic crisis of the 1980s, Environ. Dev. Sustain. 23 (2021) 4500–4520.
[34] P.M. De Oliveira-De Jesus, J.J. Galvis, D. Rojas-Lozano, J.M. Yusta, Multitemporal LMDI index decomposition analysis to explain the changes of ACI by the power sector in Latin America and the Caribbean between 1990–2017, Energies 13 (9) (2020) 2328.
[35] M. Isik, P.O. Kaplan, Understanding technology, fuel, market and policy drivers for New York State's power sector transformation, Sustainability 13 (1) (2020) 265.
[36] H. Kim, M. Kim, H. Kim, S. Park, Decomposition analysis of CO2 emission from electricity generation: Comparison of OECD countries before and after the financial crisis, Energies 13 (14) (2020) 3522.
[37] P. Sadorsky, Energy related CO2 emissions before and after the financial crisis, Sustainability 12 (9) (2020) 3867.
[38] P.M. De Oliveira-De Jesus, Effect of generation capacity factors on carbon emission intensity of electricity of Latin America & the Caribbean, a temporal IDA-LMDI analysis, Renew. Sustain. Energy Rev. 101 (2019) 516–526.
[39] N. Liu, Z. Ma, J. Kang, B. Su, A multi-region multi-sector decomposition and attribution analysis of aggregate carbon intensity in China from 2000 to 2015, Energy Policy 129 (2019) 410–421.
[40] C.H. Chong, W.X. Tan, Z.J. Ting, P. Liu, L. Ma, Z. Li, W. Ni, The driving factors of energy-related CO2 emission growth in Malaysia: The LMDI decomposition method based on energy allocation analysis, Renew. Sustain. Energy Rev. 115 (2019) 109356.
[41] C.-P. Chang, M. Dong, B. Sui, Y. Chu, Driving forces of global carbon emissions: From time-and spatial-dynamic perspectives, Econ. Model. 77 (2019) 70–80.
[42] B. Zhu, B. Su, Y. Li, Input-output and structural decomposition analysis of India's carbon emissions and intensity, 2007/08–2013/14, Appl. Energy 230 (2018) 1545–1556.
[43] R. Román-Collado, J.M. Cansino, C. Botia, How far is Colombia from decoupling? Two-level decomposition analysis of energy consumption changes, Energy 148 (2018) 687–700.
[44] R. Román, J.M. Cansino, J.A. Rodas, Analysis of the main drivers of CO2 emissions changes in Colombia (1990–2012) and its political implications, Renew. Energy 116 (2018) 402–411.
[45] J. Chen, P. Wang, L. Cui, S. Huang, M. Song, Decomposition and decoupling analysis of CO2 emissions in OECD, Appl. Energy 231 (2018) 937–950.
[46] V. Moutinho, M. Madaleno, R. Inglesi-Lotz, E. Dogan, Factors affecting CO2 emissions in top countries on renewable energies: A LMDI decomposition application, Renew. Sustain. Energy Rev. 90 (2018) 605–622.
[47] R. Román-Collado, A.V. Morales-Carrión, Towards a sustainable growth in Latin America: A multiregional spatial decomposition analysis of the driving forces behind CO2 emissions changes, Energy Policy 115 (2018) 273–280.
[48] B. Mousavi, N.S.A. Lopez, J.B.M. Biona, A.S. Chiu, M. Blesl, Driving forces of Iran's CO2 emissions from energy consumption: An LMDI decomposition approach, Appl. Energy 206 (2017) 804–814.
[49] C. Chong, P. Liu, L. Ma, Z. Li, W. Ni, X. Li, S. Song, LMDI decomposition of energy consumption in Guangdong Province, China, based on an energy allocation diagram, Energy 133 (2017) 525–544.
[50] J. Jiang, B. Ye, D. Xie, J. Tang, Provincial-level carbon emission drivers and emission reduction strategies in China: Combining multi-layer LMDI decomposition with hierarchical clustering, J. Clean. Prod. 169 (2017) 178–190.
[51] X.-T. Jiang, R. Li, Decoupling and decomposition analysis of carbon emissions from electric output in the United States, Sustainability 9 (6) (2017) 886.
[52] H. Achour, M. Belloumi, Decomposing the influencing factors of energy consumption in Tunisian transportation sector using the LMDI method, Transp. Policy 52 (2016) 64–71.
[53] W. Zhang, K. Li, D. Zhou, W. Zhang, H. Gao, Decomposition of intensity of energy-related CO2 emission in Chinese Provinces using the LMDI method, Energy Policy 92 (2016) 369–381.
[54] A.K. Sumabat, N.S. Lopez, K.D. Yu, H. Hao, R. Li, Y. Geng, A.S. Chiu, Decomposition analysis of Philippine CO2 emissions from fuel combustion and electricity generation, Appl. Energy 164 (2016) 795–804.
[55] R.D. Torrie, C. Stone, D.B. Layzell, Understanding energy systems change in Canada: 1. Decomposition of total energy intensity, Energy Econ. 56 (2016) 101–106.
[56] M. Karmellos, D. Kopidou, D. Diakoulaki, A decomposition analysis of the driving factors of CO2 (carbon dioxide) emissions from the power sector in the European Union countries, Energy 94 (2016) 680–692.
[57] L. Yang, B. Lin, Carbon dioxide-emission in Chinas power industry: Evidence and policy implications, Renew. Sustain. Energy Rev. 60 (2016) 258–267.
[58] Z.-H. Tian, Z.-L. Yang, Scenarios of carbon emissions from the power sector in Guangdong Province, Sustainability 8 (9) (2016) 863.
[59] B. Ang, T. Goh, Carbon intensity of electricity in ASEAN: Drivers, performance and outlook, Energy Policy 98 (2016) 170–179.
[60] B.W. Ang, B. Su, Carbon emission intensity in electricity production: A global analysis, Energy Policy 94 (2016) 56–63.
[61] L. Mundaca, A. Markandya, Assessing regional progress towards a 'Green Energy Economy', Appl. Energy 179 (2016) 1372–1394.
[62] L. Andrés, E. Padilla, Energy intensity in road freight transport of heavy goods vehicles in Spain, Energy Policy 85 (2015) 309–321.

[63] J.M. Cansino, A. Sánchez-Braza, M.L. Rodríguez-Arévalo, Driving forces of Spain's CO2 emissions: A LMDI decomposition approach, Renew. Sustain. Energy Rev. 48 (2015) 749–759.

[64] C. Chong, L. Ma, Z. Li, W. Ni, S. Song, Logarithmic mean Divisia index (LMDI) decomposition of coal consumption in China based on the energy allocation diagram of coal flows, Energy 85 (2015) 366–378.

[65] V. Moutinho, A.C. Moreira, P.M. Silva, The driving forces of change in energy-related CO2 emissions in Eastern, Western, Northern and Southern Europe: The LMDI approach to decomposition analysis, Renew. Sustain. Energy Rev. 50 (2015) 1485–1499.

[66] P.F. González, M. Landajo, M. Presno, The driving forces behind changes in CO2 emission levels in EU-27. differences between member states, Environ. Sci. Policy 38 (2014) 11–16.

[67] V.M. Muggeo, Estimating regression models with unknown break-points, Stat. Med. 22 (19) (2003) 3055–3071.

[68] I.E. Agency, CO2 Emissions from Fuel Combustion 2019, 2019, p. 514, http://dx.doi.org/10.1787/2a701673-en, URL: https://www.oecd-ilibrary.org/content/publication/2a701673-en.

[69] S. Jung, K.-J. An, G. Dodbiba, T. Fujita, Regional energy-related carbon emission characteristics and potential mitigation in eco-industrial parks in South Korea: Logarithmic mean Divisia index analysis based on the Kaya identity, Energy 46 (1) (2012) 231–241.