Low-dimensional representation of monthly electricity demand profiles

Joaquin Luque[a*], *Enrique Personal*[a], *Francisco Perez*[b], *MariCarmen Romero-Ternero*[b], *Carlos Leon*[a]

[a]Dpto. Tecnología Electrónica. Escuela Politécnica Superior. Universidad de Sevilla. Virgen de África, 7. 41011-Sevilla.Spain.

[b]Dpto. Tecnología Electrónica. Escuela Técnica Superior de Ingeniería Informática. Universidad de Sevilla. Av. Reina Mercedes, s/n. 41012-Sevilla.Spain.

[*]Corresponding author

emails: jluque@us.es, epersonal@us.es, fperez@us.es, mcromerot@us.es, cleon@us.es

Abstract

This paper addresses the problem of reducing the number of values required to characterize an electricity demand profile, which is usually known as its dimensionality. This reduction may have a significant impact on the computational efforts and storage capacities required to analyze and process high volumes of electricity load curves. Also, the reduction to 2 or even 1 component enables its graphic representation. Specifically, this work is mainly focused on profiles defined by their monthly demand values, and where the clients are aggregated by locations and/or economic activities. This approach is of great interest for marketing analysis and decision-making of electricity retailers. In this sense, the use of dimensionality reduction techniques based on knowledge (calendar and temperature) along with the application of data-driven procedures (Principal Component Analysis and autoencoders), are explored in the paper. The results of this research show that autoencoders clearly outperform the other techniques, yielding errors in the reduction process between 15% to 40% lower and preserving distances between profiles in the low-dimensional spaces, with a correlation of 0.93 with the distances in high dimensional space. Additionally, the bidimensional graphical representation of a profile can easily be interpreted in a polar way, where the angle denotes the shape of the profile, and the radius reveals its scale. To reach these results, a very large dataset has been employed, with about half a million aggregated profiles corresponding to the electricity consumption during 3 years of more than 27 million clients in Spain.

Highlights

- Autoencoders are the best option to reduce the dimensionality of demand profiles.
- Data-driven techniques outperform knowledge-based models by reducing components.
- Low-dimensional characterization preserves the relative distances between profiles.
- The graphic representation of bidimensional profiles admits a polar interpretation.

Keywords

Electricity demand profile, dimensionality reduction, graphic representation, profile clustering, profile labelling.

1. Introduction.

During the last decades, most open economies have liberalized their electricity markets. This liberalization has provoked competition among different companies, with the goal of obtaining greater efficiency, that is, the ability to provide higher and better (cleaner, safer, resilient) electricity energy at lower costs (Khan et al., 2022). Also, the issue of electric energy involves not only costs, but also the policies of green development and carbon emissions that must be carefully addressed (Tian et al., 2022).

The different actors in these markets are mainly those that are focusing on: generation; high voltage transmission (transmission system operator, TSO); medium and low voltage distribution (distribution system operator, DSO); and electricity retailing that carry out marketing, invoicing, maintenance and client attention services. Usually, very few companies are involved in generation, transmission, and distribution; therefore, they operate with an oligopoly structure. However, a relatively high number of electricity retailers may compete in a certain territory. For example, Spain had 257 nationwide active retailers (2019 data), becoming the European country with the high number of them and which had experienced the highest increase in this number within the European Union (ACER, 2021).

This greater competition in the electricity retail has led to the development of different marketing strategies, many of them taking into account the client demand profile (Gong et al., 2019). These marketing strategies may consider hourly (Yang et al., 2019) and/or the monthly demand profiles (Lee et al., 2014). Therefore, robust demand profiling is an essential tool not only for retailers, but also for suppliers, DSOs, aggregators and energy service companies (Panapakidis et al., 2015).

For marketing purposes, it is very convenient to cluster the whole customer base into groups with similar characteristics where the location (Shaffer, 2019) and the economic sector of the client may play a key role (Guerrero et al., 2018) in companies' strategies. These two factors are so important that some retailers base their marketing decision on the demand profiles for a certain economic activity at a certain location (Luque et al., 2021).

Electricity demand profiles, both hourly and monthly, are defined using high-dimensional descriptions (for example, with 24 or 12 values). This relatively high number of components hinders both their automatic treatment and their interpretation by those who must design marketing strategies. Therefore, reducing the dimensionality of

these profiles has undoubted advantages in both the efficiency and the explainability of data analysis algorithms (Morán et al., 2013).

Especially useful is the two-dimensional representation of the profiles whereby a certain profile (e.g. 24 or 12 values) is represented by a point on a two-dimensional map (Yi Wang et al., 2016). On the other hand, reducing the profile to a single dimension allows it to be characterized by a single continuous value. This may be an advantage over classic profile segmentation techniques (either individual or aggregated) in which the output is a label that, if it is numerical, is discrete (instead of continuous) and does not necessarily imply an ordering of the profiles (Chhikara et al., 2022).

In this paper, the monthly demand profiles of a set of electricity consumers sharing the same location and economic activity are considered. Then, their mean profile is obtained and featured using 12 values, where each value measures the consumed energy by the customer throughout each one of the twelve months. In this sense, the main aim of this proposed paper is to represent these profiles with a lower number of values, exploring the relationship between the resulting number of components and the errors induced by this dimensionality reduction. As will be shown later, this reduction is a powerful tool for getting a significant reduction in storage needs, which could be critical due to the high number of handled profiles. Therefore, a reduction in the number of components required would significantly reduce the size of required memory.

Throughout the paper, the objective of lowering the error inherent to the process of dimensionality reduction will be combined with the aim of an easy interpretation of the low-dimensional profiles. So, the focus is not mainly on comparing dozens of dimensionality reduction techniques and selecting the best one, or even in proposing new methods, but in assuring that the results obtained using state of the art techniques can be adequately interpreted in the context of electricity demand.

Additionally, a diversity of analysis has been carried out. Specifically, special attention has been paid to reduce the profiles to just 2 components, allowing their graphic representation on a plane. The relationship between the position of a profile in the 2D plane and its high-dimensional representation will be extensively explored. The extreme particular case in which the original profile is reduced to a single component will also be addressed.

It is important to highlight among the novelties of the current work, the large number of profiles used (approximately half a million), the grouping of these profiles by economic activity and location, the introduction of reduction techniques based on models, the analysis of the recovery errors and preservation of distances, and finally, the interpretation of the representations in low-dimensional spaces. Thanks to all these analyses and their results, it is possible to say that the reduction in dimensionality through the proposed techniques is suitable, and that despite losing information, this loss is compensated by the reduction of storage requirements by reducing the number of

necessary components and, more importantly, by the easier representation and interpretation of the low-dimensional profiles. The main limitation of this research is derived by the fact that the customers' economic activity is omitted in the datasets for approximately half of them.

The remaining of the paper is organized as follows: first, some related works are described in section II; second, the used dataset and the employed methodology are formalized in section III; later, up to four different dimensionality reduction techniques are detailed in section IV; then, the main results obtained applying these techniques are summarized in section V; and finally, these results are discussed in section VI, presenting the main conclusions in section VII.

2. Related works.

There is a long tradition in the use of data analysis techniques in the different stages of the power electricity chain: generation, trading, transmission, distribution, and consumption (Scheidt et al., 2020). In the last decade, deep learning approaches have aroused huge attention in diverse research areas with applications to advanced forecasting support, power quality monitoring, microgrid energy administration, electric vehicle management, and many others (Mishra et al., 2020). Also, some methodologies based on fuzzy systems have been suggested for energy analysis for industrial processes (Tian et al., 2011) or for real-time control of complex systems (Xie et al., 2018).

Specifically, in the retailer arena, data-based decision making has gained increasing interest in methods for long-term load forecasting, energy procurement strategies, pricing schemes, and risk management in the retail market (Liu et al., 2020; Yang et al., 2018). Some authors have done some related works on the long-term evolution of electric demand (Fallahpour et al., 2021; Sánchez-Durán et al., 2019b) and the solar production (Sánchez-Durán et al., 2019a), and also in featuring electricity demand using spectral analysis (Luque et al., 2020).

The characterization of load curves has highly exploited the developments in the data mining of time series, with techniques generally categorized into representation and indexing, similarity measure, segmentation, visualization, and mining (Fu, 2011).

Although many related works consider the evolution of the electricity demand with a period of one hour, some other authors have focused their interest in load curves with a higher resolution, such as 1 second (Tjaden et al., 2015) or 15 minute (Kiesel and Paraschiv, 2017), while others take lower resolutions, such as one day (Zhou et al., 2017) or one month (Pełka and Dudek. Grzegorz, 2020).

The characterization of the electricity demand profiles is usually performed applying clustering techniques whose state-of-the-art has been summarized by different researchers (Ezugwu et al., 2022). Featuring and clustering load profiles have also been

described as the basis for establishing an annual framework for optimal price offering by a retailer (Mahmoudi-Kohan et al., 2010).

On the other hand, dimensionality reduction techniques have been extensively used in many areas, in particular in the electricity sector (Williams and Short, 2020). As an example of these applications, the reduction and 2D visual representation of a very small simulated network (IEEE 30-bus system) is analyzed in (Arechiga et al., 2017). Additionally, 15-minute daily demand profiles for 317 real consumers are reduced to 4 dimension and then clustered in (Shi et al., 2020). Another example is the dimensionality reduction and 3D visual representation of frequency measurements in the European Network of Transmission System Operators for Electricity (ENTSO-E) analyzed in (Sevilla et al., 2019).

In this sense and as it will be seen in the following sections, this paper proposes a detailed comparative analysis between different dimensionality reduction techniques and exhibits their results (in effectiveness terms) when them are applied to the consumer profiling problems.

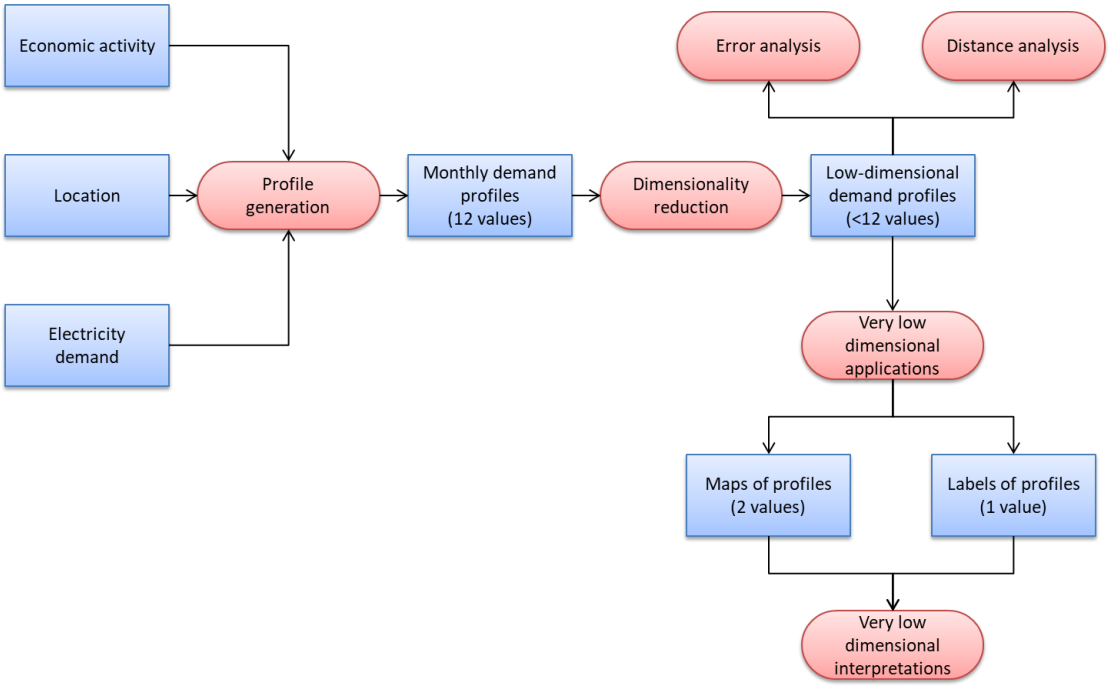

Fig. 1. Overall process for low-dimensional representation of electricity demand profiles.

3. Methodology.

The overall research ideas of the article are summarized in Fig. 1. Firstly, the monthly electricity demand of all the customers with the same economic activity sector at a certain location (for instance, hotels in the city of Alicante) are jointly characterized by a 12-valued profile. Then, several dimensionality reduction techniques are employed to shrink the number of values characterizing every profile. Next, the obtained low-

dimensional demand profiles are assessed analyzing the error introduced and to what extent is the topology preserved (distance analysis) during the dimensionality reduction process.

The paper pays special attention to very low dimensional profiles. By reducing the original profile to just 2 values, a bidimensional representation of the profiles can be obtained (map of profiles). The extreme case of shrinking the original profile to a single value is some sort of labelling them. The interpretation of the maps and labels is also of great interest for their use in actual engineering applications. These general ideas are developed and detailed in the following sections.

3.1. Energy demand profiles.

To improve services and reduce prices in the electricity market, many governments have forced electric companies to share their data on the monthly electricity demand for every customer. Although this information is usually anonymized before sharing, it still preserves some statistical data, such as, among others, the location and economic activity of the customer, identified by the NACE code (European Classification of Economic Activities, in French). In the case of the electricity market in Spain, 1011 different NACE codes are used, and 8203 locations are recorded. Taking all customers identified by the k-th NACE-location pair, a monthly electricity demand profile can be defined by the row-vector $\boldsymbol{p}^{(k)} = \begin{bmatrix} p_1^{(k)} & p_2^{(k)} & \cdots & p_{12}^{(k)} \end{bmatrix}$, where $p_j^{(k)}$ is the median electricity demand in the j-th month of all the customers belonging to the k-th NACE-location pair. The description to obtain the profile's dataset is detailed in (Luque et al., 2021).

Based on this previous analysis, a total of 8,293,233 (1011 x 8203) demand profiles could be registered. However, many of them do not contain any customers or they show null demand, as small or medium-size locations do not include all the economic activities. After these pairs are removed, 491,604 profiles remain. Additionally, as the process of dimensionality reduction may be affected by extreme profiles, they are also eliminated. These profiles are considered outliers and mainly correspond to a single customer with abnormal behavior. Formally, a $\boldsymbol{p}^{(k)}$ profile is considered an outlier if $\exists j: p_j^{(k)} \notin \left[\pi_{j,2.5}, \pi_{j,97.5}\right]$, where $\pi_{j,2.5}$ and $\pi_{j,97.5}$ are the 2.5-th and 97.5-th percentiles of $P_j = \left\{p_j^{(k)}\right\}, \forall k$. By definition of percentile, the probability that $p_j^{(k)} < \pi_{j,97.5}$ is 97.5%, and the probability that $p_j^{(k)} < \pi_{j,2.5}$ is 2.5%. Then the probability that $p_j^{(k)}$ is in the range $\pi_{j,2.5} < p_j^{(k)} < \pi_{j,97.5}$ is $97.5\% - 2.5\% = 95\%$. This magnitude is called the confidence interval, CI, being the use of $CI = 95\%$ a common practice in statistics. Finally, there are a total of 400,630 valid profiles, while 90% of them (360,567 profiles) are used to train the dimensionality reduction models (training dataset), and 10% (40,063 profiles) are employed to assess the models (testing dataset).

3.2. Profile transformations.

The process of reducing the dimensionality of an electricity demand profile can be seen as a transformation $T^{[pq]}$ operating on a 12-dimensional $\boldsymbol{p}^{(k)}$ profile, corresponding to the k-th NACE-location pair, to obtain a d-dimensional $\boldsymbol{q}^{(k)}$ profile, where $d < 12$, that is, $\boldsymbol{p}^{(k)} \xrightarrow{T^{[pq]}} \boldsymbol{q}^{(k)}$. This transformation is usually made up of several steps. In this research, up to four stages have been considered.

1) Normalization. Many dimensionality reduction techniques require or perform better if the values of the features are on the same scale (Obaid et al., 2019). So, in this work, where necessary, the first preprocessing step is the normalization of the profiles in such a way that $\boldsymbol{p}^{(k)} \xrightarrow{T^{[pn]}} \boldsymbol{n}^{(k)}$, where $\boldsymbol{n}^{(k)}$is a 12-dimensional profile and the value for its j-th month, $n_j^{(k)}$, is derived using the z-score given by

$$n_j^{(k)} = \frac{p_j^{(k)} - \mu_{Pj}}{\sigma_{Pj}}, \tag{1}$$

where μ_{Pj} and σ_{Pj} are the mean and standard deviation of the demand profile of every NACE-location pair at the j-th month. In matrix notation, $\boldsymbol{p}^{(k)} = \boldsymbol{\sigma}_P \circ \boldsymbol{n}^{(k)} + \boldsymbol{\mu}_P$, where the symbol $\circ$ represents the element-wise multiplication (Hadamard product), $\boldsymbol{\mu}_P = [\mu_{P1} \quad \mu_{P2} \quad \cdots \quad \mu_{P12}]$, and $\boldsymbol{\sigma}_P = [\sigma_{P1} \quad \sigma_{P2} \quad \cdots \quad \sigma_{P12}]$. If no normalization is applied, then $\boldsymbol{p}^{(k)} = \boldsymbol{n}^{(k)}$.

2) Dimensionality reduction. A normalized 12-dimensional profile $\boldsymbol{n}^{(k)}$ is then transformed into a lower dimensional vector $\boldsymbol{n}^{(k)} \xrightarrow{T^{[nl]}} \boldsymbol{l}^{(k)}$. The algorithms to perform this task will be described in Section 4.

3) Inverting axis. An effective low-dimensional representation of a profile should not only reduce d (to a number less than 12), but also give to each component an approximated physical interpretation. For this purpose, some values of the low-dimensional $\boldsymbol{l}^{(k)}$ profile can be inverted obtaining a mirrored low-dimensional profile, such that $\boldsymbol{l}^{(k)} \xrightarrow{T^{[lm]}} \boldsymbol{m}^{(k)}$.

4) Scaling axis. Physical interpretation of a low-dimensional profile $\boldsymbol{m}^{(k)}$ could be further improved if it is scaled and shifted on one or several of its d-axis. To perform this task, a new transformation $\boldsymbol{m}^{(k)} \xrightarrow{T^{[mq]}} \boldsymbol{q}^{(k)}$ is applied.

The overall transformation is then $\boldsymbol{p}^{(k)} \xrightarrow{T^{[pn]}} \boldsymbol{n}^{(k)} \xrightarrow{T^{[nl]}} \boldsymbol{l}^{(k)} \xrightarrow{T^{[lm]}} \boldsymbol{m}^{(k)} \xrightarrow{T^{[mq]}} \boldsymbol{q}^{(k)}$. The mean value of the profiles is transformed as $\boldsymbol{\mu}_P \xrightarrow{T^{[pn]}} \boldsymbol{\mu}_N \xrightarrow{T^{[nl]}} \boldsymbol{\mu}_L \xrightarrow{T^{[lm]}} \boldsymbol{\mu}_M \xrightarrow{T^{[mq]}} \boldsymbol{\mu}_Q$. This process follows the flowchart depicted in Fig. 2.

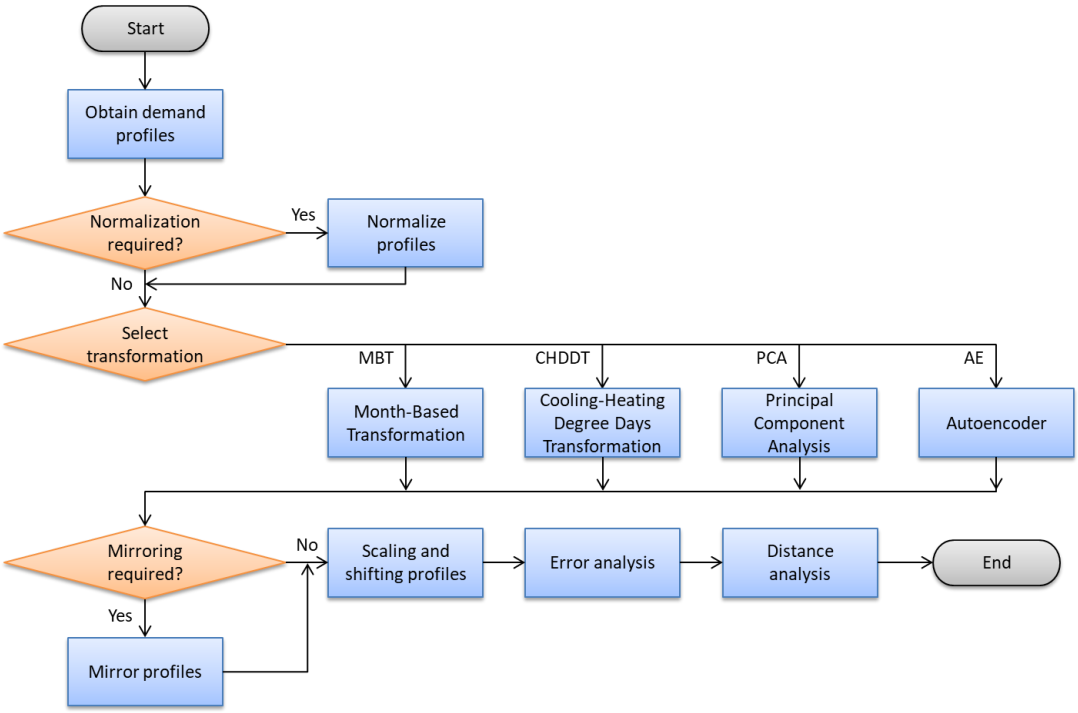

Fig. 2. Flowchart representing the process of obtaining and assessing low-dimensional electricity demand profiles.

3.3. Low dimensional representation.

In the 4-step previously described process, the normalization ($T^{[pn]}$) plays an instrumental role, while axis inversion ($T^{[lm]}$) and scaling ($T^{[mq]}$) have mainly a cosmetic purpose (enhancing low-dimensional interpretation). None of these steps yields a change in the number of components of the profile. Then, the kernel ingredient for the pursued objective is the dimensionality reduction step, $\boldsymbol{n}^{(k)} \xrightarrow{T^{[nl]}} \boldsymbol{l}^{(k)}$, which transforms a 12-component normalized profile $\boldsymbol{n}^{(k)}$ into a d-dimensional vector $\boldsymbol{l}^{(k)} = \begin{bmatrix} l_1^{(k)} & l_2^{(k)} & \cdots & l_d^{(k)} \end{bmatrix}$, where $l_c^{(k)}$ is its c-th component and $c \in [1, d]$.

Some dimensionality reduction techniques use a linear transformation to obtain $l_c^{(k)}$, that is, $l_c^{(k)} = T_c^{[nl]}\left(\boldsymbol{n}^{(k)}\right) = \boldsymbol{n}^{(k)}\boldsymbol{\alpha}_c^T + \beta_c$, where $\boldsymbol{\alpha}_c = [\alpha_{c1} \quad \alpha_{c2} \quad \cdots \quad \alpha_{c12}]$ is a row vector of constant factors, interpreted as the sensitivity vector of the c-th component. Besides, β_c is a constant scalar representing the bias of the transformation. Grouping the d sensitivity vectors in a $12 \times d$ sensitivity matrix $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{\alpha}_1^T & \boldsymbol{\alpha}_2^T & \cdots & \boldsymbol{\alpha}_d{}^T \end{bmatrix}$, and the d bias constants in a $1 \times d$ bias vector $\boldsymbol{\beta} = [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_d]$, the low-dimensional complete representation will be $\boldsymbol{l}^{(k)} = \boldsymbol{n}^{(k)}\boldsymbol{S} + \boldsymbol{\beta}$.

It can be seen that $\boldsymbol{\alpha}_c = \left[\frac{\partial l_c^{(k)}}{\partial n_1^{(k)}} \quad \frac{\partial l_c^{(k)}}{\partial n_2^{(k)}} \quad \cdots \quad \frac{\partial l_c^{(k)}}{\partial n_{12}^{(k)}}\right] = \boldsymbol{\nabla l_c^{(k)}}$. Then, the sensitivity matrix $\boldsymbol{S}$ can be interpreted as the Jacobian of the $T^{[nl]}$ transformation of $\boldsymbol{n}^{(k)}$, defined as $\boldsymbol{S} = \boldsymbol{J}_{T^{[nl]}} = \left[\boldsymbol{\nabla l_1^{(k)^T}} \quad \boldsymbol{\nabla l_2^{(k)^T}} \quad \cdots \quad \boldsymbol{\nabla l_d^{(k)^T}}\right]$. Due to the linearity of the transformation, this expression is valid for any profile, that is, $\forall k$. Each α_{cj} element of the sensitivity matrix can be thought as the influence of the j-th month normalized value on the construction of the c-th component of the low-dimensional demand profile.

The linear dimensionality reduction transformation $T_c^{[nl]}$ defined as $\boldsymbol{l}^{(k)} = \boldsymbol{n}^{(k)}\boldsymbol{S} + \boldsymbol{\beta}$ is valid for any value of k, that is, for any NACE-location pair. In particular, it is valid for the transformation of the normalized mean demand profile $\boldsymbol{\mu}_L = T^{[nl]}(\boldsymbol{\mu}_N) = \boldsymbol{\mu}_N\boldsymbol{S} + \boldsymbol{\beta}$. For this expression, the value of $\boldsymbol{\beta}$ can be derived as $\boldsymbol{\beta} = \boldsymbol{\mu}_L - \boldsymbol{\mu}_N\boldsymbol{S}$. Therefore, the linear transformation can be expressed as

$$\boldsymbol{l}^{(k)} = \left(\boldsymbol{n}^{(k)} - \boldsymbol{\mu}_N\right)\boldsymbol{S} + \boldsymbol{\mu}_L. \tag{2}$$

To reverse this process, the sensitivity matrix, $\boldsymbol{S}^{-1}$, will be required. But $\boldsymbol{S}$ is not a square matrix and, therefore, it is not invertible. An approximate solution can be found using the Moore-Penrose pseudo-inverse matrix (Golan, 2012), $\boldsymbol{S}^{+}$. Then, the inverse transformation, $\boldsymbol{l}^{(k)} \xrightarrow{T^{[ln]}} \tilde{\boldsymbol{n}}^{(k)}$, can be expressed as $\tilde{\boldsymbol{n}}^{(k)} = \left(\boldsymbol{l}^{(k)} - \boldsymbol{\mu}_L\right)\boldsymbol{S}^{+} + \boldsymbol{\mu}_N$, obtaining a 12-element row vector that estimates the normalized demand profile.

The $\boldsymbol{S}^{+}$ matrix has dimension $d \times 12$ and is also known as the explanation matrix $\boldsymbol{E} = \boldsymbol{S}^{+} = [\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \cdots \quad \boldsymbol{e}_{12}]$, where $\boldsymbol{e}_j = [e_{1j} \quad e_{2j} \quad \cdots \quad e_{dj}]^T$, and each $e_{cj} = \frac{\partial \tilde{n}_j^{(k)}}{\partial l_c^{(k)}}$ can be thought as the influence of the c-th component of the low-dimensional profile on the estimation of the j-th month normalized demand. The explanation matrix $\boldsymbol{E}$ can also be denoted as the Jacobian of the $T^{[ln]}$ reverse transformation of $\boldsymbol{l}^{(k)}$, defined as $\boldsymbol{E} = \boldsymbol{J}_{T^{[ln]}} = \left[\nabla \tilde{n}_1^{(k)} \quad \nabla\tilde{n}_2^{(k)} \quad \cdots \quad \nabla\tilde{n}_{12}^{(k)}\right]$, where $\nabla\tilde{n}_j^{(k)} = \boldsymbol{e}_j$. The inverse transformation is then written as

$$\tilde{\boldsymbol{n}}^{(k)} = \left(\boldsymbol{l}^{(k)} - \boldsymbol{\mu}_L\right)\boldsymbol{E} + \boldsymbol{\mu}_N. \tag{3}$$

If $T^{[nl]}$ is a linear transformation, the sensitivity is the same for every profile, that is, it does not depend on k. This is not the case for non-linear transformations where the Jacobian is different at every point in the space. Then, for an easy interpretation of the low-dimensional profiles, the $T^{[nl]}$ transformation is approximated by a linear one.

For that purpose, the expression to obtain the c-th component from the normalized demand (the $T^{[nl]}$ transformation), $l_c^{(k)} = T_c^{[nl]}\left(\boldsymbol{n}^{(k)}\right)$, is expanded in its Taylor series at the point $\boldsymbol{n}^{(*)} = \boldsymbol{\mu}_N$ and, later, only the two first terms are preserved:

$$l_c^{(k)} \approx T_c^{[nl]}(\boldsymbol{\mu}_N) + \sum_{j=1}^{12} \left[\frac{\partial l_c^{(k)}}{\partial n_j^{(k)}}\right]_{\boldsymbol{\mu}_N} \left(n_j^{(k)} - \mu_{Nj}\right), \tag{4}$$

that is, $l_c^{(k)} \approx T_c^{[nl]}(\boldsymbol{\mu}_N) + \left(\boldsymbol{n}^{(k)} - \boldsymbol{\mu}_N\right)\left[\nabla \boldsymbol{l}_c^{(k)}\right]_{\boldsymbol{\mu}_P}^T$. Recalling that $\boldsymbol{\mu}_L$ is the low-dimensional representation of the mean profile, that is $\boldsymbol{\mu}_L = T_c^{[nl]}(\boldsymbol{\mu}_N)$, and defining $\boldsymbol{s}_c$ as the sensitivity vector of the transformation at the point $\boldsymbol{\mu}_N$, that is, $\boldsymbol{s}_c = \left[\nabla l_c^{(k)}\right]_{\boldsymbol{\mu}_N}$, it can be written that $l_c^{(k)} \approx \boldsymbol{\mu}_L + \left(\boldsymbol{n}^{(k)} - \boldsymbol{\mu}_N\right)\boldsymbol{s}_c^T = \boldsymbol{\mu}_L - \boldsymbol{\mu}_N \boldsymbol{s}_c^T + \boldsymbol{n}^{(k)}\boldsymbol{s}_c^T$. Extending this result to the d components of $\boldsymbol{l}^{(k)}$, it can be written as $\boldsymbol{l}^{(k)} \approx \boldsymbol{\mu}_L - \boldsymbol{\mu}_N \boldsymbol{S} + \boldsymbol{n}^{(k)}\boldsymbol{S}$, or $\boldsymbol{l}^{(k)} \approx \left(\boldsymbol{n}^{(k)} - \boldsymbol{\mu}_N\right)\boldsymbol{S} + \boldsymbol{\mu}_L$, which is the same result obtained for the linear transformation. The inverse process also has the same expression as in the linear case, that is, $\tilde{\boldsymbol{n}}^{(k)} = \left(\boldsymbol{l}^{(k)} - \boldsymbol{\mu}_L\right)\boldsymbol{E} + \boldsymbol{\mu}_N$, where $\boldsymbol{E} = \boldsymbol{S}^+$.

3.4. Low-dimensional interpretation.

To enhance the physical interpretation of the low-dimensional profile, it can be further transformed, mirroring some axis $\boldsymbol{l}^{(k)} \xrightarrow{T^{[lm]}} \boldsymbol{m}^{(k)}$. In matrix notation, $\boldsymbol{m}^{(k)} = \boldsymbol{\gamma} \circ \boldsymbol{l}^{(k)}$, where $\boldsymbol{\gamma} = [\gamma_1 \quad \gamma_2 \quad \cdots \quad \gamma_d]$ whose c-th component is $\gamma_c = -1$ if it has been previously inverted; otherwise, $\gamma_c = 1$.

Additionally, a scaling and shifting transformation, $\boldsymbol{m}^{(k)} \xrightarrow{T^{[mq]}} \boldsymbol{q}^{(k)}$, usually enhances the low-dimensional profile meaning. In matrix notation this transformation can be expressed as, $\boldsymbol{q}^{(k)} = \boldsymbol{\lambda} \circ \boldsymbol{m}^{(k)} + \boldsymbol{\eta}$, where $\boldsymbol{\lambda} = [\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_d]$ is a d-dimesional scaling row vector, and $\boldsymbol{\eta} = [\eta_1 \quad \eta_2 \quad \cdots \quad \eta_d]$ is the corresponding shifting vector. The expression for the c-th component of the result will be $q_c^{(k)} = \lambda_c m_c^{(k)} + \eta_c$.

To determine the c-th scaling factor, λ_c, the physical interpretation of the c axis must be considered. Let us assume that the c-th component explains the demand profile at a certain set of months, A_c. For example, consider a certain 2D representation of the profiles where the horizontal axis ($c = 1$) mainly explains the demand in July and August (summer demand), while the vertical axis ($c = 2$) explains the demand in January and February (winter demand). Then, $A_1 = \{7,8\}$, and $A_2 = \{1,2\}$.

For the sake of low-dimension profiles interpretation, all the transformations are considered linear, or they are substituted by their linear approximations. Then, the inverse transformation of the c-th component profile $q_c^{(k)}$ can be approximated by its Taylor expansion at the point $q_c = \mu_{Qc}$, where μ_{Qc} is the c-th component of $\boldsymbol{\mu}_Q$. The j-th month demand explained by this c-th component is, $\tilde{p}_{jc}^{(k)} = T^{[qp]}(\mu_{Qc}) +$

$\left[\frac{\partial \tilde{p}_j}{\partial q_c}\right]_{q_c=\mu_{Qc}} \left(q_c^{(k)} - \mu_{Qc}\right)$. Calling $\mu_{\tilde{P}j}$ to the j-th month mean estimated demand, it can be written that $\tilde{p}_{jc}^{(k)} = \mu_{\tilde{P}j} + \left[\frac{\partial \tilde{p}_j}{\partial q_c}\right]_{q_c=\mu_{Qc}} \left(q_c^{(k)} - \mu_{Qc}\right)$ which is a straight line with a slope of value $\left[\frac{\partial \tilde{p}_j}{\partial q_c}\right]_{q_c=\mu_{Qc}}$. To explain $\tilde{p}_{jc}^{(k)}$ in terms of the value of $q_c^{(k)}$, the slope should be approximate 1. Then it can be written that $\left[\frac{\partial \tilde{p}_j}{\partial q_c}\right]_{q_c=\mu_{Qc}} = \left[\frac{\partial \tilde{p}_j}{\partial m_c}\frac{\partial m_c}{\partial q_c}\right]_{q_c=\mu_{Qc}} = \left[\frac{\partial \tilde{p}_j}{\partial m_c}\frac{\partial m_c}{\partial q_c}\right]_{m_c=\mu_{Mc}} = \left[\frac{\partial \tilde{p}_j}{\partial m_c}\right]_{m_c=\mu_{Mc}} \left[\frac{\partial q_c}{\partial m_c}\right]^{-1}_{m_c=\mu_{Mc}} \approx 1$.

But $\frac{\partial q_c}{\partial m_c} = \lambda_c$ at any point, so $\left[\frac{\partial \tilde{p}_j}{\partial m_c}\right]_{m_c=\mu_{Mc}} \lambda_c^{-1} \approx 1$ and, finally, $\lambda_c \approx \left[\frac{\partial \tilde{p}_j}{\partial m_c}\right]_{m_c=\mu_{Mc}} = \lambda_{cj}$. If there are several months that are explained by the c-th component, the scaling factor is computed as the mean of those corresponding to each month, that is $\lambda_c \approx \bar{\lambda}_{cj}, \forall j \in A_c$, that is,

$$\lambda_c \approx \underset{j \in A_c}{\text{mean}} \left[\frac{\partial \tilde{p}_j}{\partial m_c}\right]_{m_c=\mu_{Mc}}. \tag{5}$$

To obtain the values of the shifting vector, we will first consider that the c-th component should explain the j-th month demand, that is, $\tilde{p}_{jc}^{(k)} \approx q_c^{(k)}$. This equation is true for any point, in particular for the mean demand profile $\mu_{\tilde{P}j} \approx \mu_{Qc}$. But $\mu_{Qc} = \lambda_c \mu_{Mc} + \eta_c$, so $\eta_c \approx \mu_{\tilde{P}j} - \lambda_c \mu_{Mc} = \eta_{cj}$. If there are several months in A_c, then,

$$\eta_c \approx \underset{j \in A_c}{\text{mean}} \left(\mu_{\tilde{P}j} - \lambda_c \mu_{Mc}\right). \tag{6}$$

An alternative interpretation of the low-dimensional profile is in terms of over-demand (demand over the mean). In this case, the c-th component should explain the j-th month over-demand, that is, $q_c^{(k)} \approx \tilde{p}_{jc}^{(k)} - \mu_{Pj}$. This equation is true for any point, in particular for the mean demand profile $\mu_{Qc} \approx \mu_{\tilde{P}j} - \mu_{Pj} \approx 0$. But $\mu_{Qc} = \lambda_c \mu_{Mc} + \eta_c$, so $\eta_c \approx -\lambda_c \mu_{Mc} = \eta_{cj}$. If there are several months in A_c, then,

$$\eta_c \approx \underset{j \in A_c}{\text{mean}} (-\lambda_c \mu_{Mc}). \tag{7}$$

3.5. Mapping profiles in 2D.

A very common application of dimensionality reduction is the representation of a multidimensional point in a 2D plane. In the case of the monthly energy demand profile, it is a particular case, for $d = 2$, of the procedure described in the previous section. In this application, the $12 \times d$ sensitivity matrix $\boldsymbol{S}$, has a dimension 12×2, where each

row is called a biplot vector. The j-th biplot is defined as $\boldsymbol{b}_j = [s_{jx} \quad s_{jy}]$ and can be considered a vector in the 2D plane. Alternatively, each of the 2 columns of the sensitivity matrix $\boldsymbol{S}$ is called an axis-construction vector and, for the c-th column (c-th axis in the 2D plane) it is defined as $\boldsymbol{a}_c = [s_{1c} \quad s_{2c} \quad \cdots \quad s_{12c}], c \in \{x, y\}$.

Analogously, in this application the $d \times 12$ explanation matrix $\boldsymbol{E}$ has a dimension of 2×12, where each column is called a reverse biplot vector. The j-th reverse biplot is defined as $\boldsymbol{b}_j^r = [e_{jx} \quad e_{jy}]$ and can be considered a vector in the 2D plane. Alternatively, each of the 2 rows of the explanation matrix $\boldsymbol{E}$ is called a reverse axis-construction vector and, for the c-th row (c-th axis in the 2D plane), it is defined as $\boldsymbol{a}_c^r = [e_{1c} \quad e_{2c} \quad \cdots \quad e_{12c}], c \in \{x, y\}$.

Direct and reverse biplot vectors are usually depicted as a set of 12 vectors in a 2D plane. Moreover, direct and reverse axis-construction vectors are usually drawn as a set of two 12-valued lines. Some examples of these plots will be shown in Section 4.

3.6. Labelling profiles (1D).

Another interesting application of dimensionality reduction is to shrink the entire 12-valued profile to a single number, that is, the target space has dimension $d = 1$. In this case, the single component of the target profile can be used to label the profile. Now, the sensitivity matrix $\boldsymbol{S}$, has a dimension 12×1, where each row is called, by analogy, a uniplot vector. The j-th uniplot is defined as $\boldsymbol{u}_j = [s_{jx}]$ and can be considered a vector in the 1D space (line). Alternatively, the only column of the sensitivity matrix $\boldsymbol{S}$ is called the axis-construction vector and it is defined as $\boldsymbol{a} = [s_{1x} \quad s_{2x} \quad \cdots \quad s_{12x}]$.

Analogously, in this application the explanation matrix $\boldsymbol{E}$ has a dimension of 1×12, where each row is called a reverse uniplot vector. The j-th reverse uniplot is defined as $\boldsymbol{u}_j^r = [e_{jx}]$ and can be considered a vector in the 1D space (line). Alternatively, the only row of the explanation matrix $\boldsymbol{E}$ is called the reverse axis-construction vector and it is defined as $\boldsymbol{a}^r = [e_{1x} \quad e_{2x} \quad \cdots \quad e_{12x}]$.

The direct and reverse axis-construction vectors are usually drawn as a set of two 12-valued lines. Uniplot vectors are not commonly represented graphically.

4. Dimensionality reduction techniques.

Once the steps and the usefulness of dimensionality reduction have been defined in the previous sections, this section will focus on how to implement the main stage of this transformation, $\boldsymbol{n}^{(k)} \xrightarrow{T^{[nl]}} \boldsymbol{l}^{(k)}$. In particular, for this kernel phase, two approaches have been tried.

In the first approach, the transformation is obtained based on some engineering knowledge of the energy demand (Aoi and Pillow, 2018). This knowledge will be obtained by two techniques: relating the energy demand to the calendar (month of the year) (Li et al., 2019), or relating this energy with its temperature dependence (Apadula et al., 2012; Hor et al., 2005).

In the second approach, the transformation $T^{[nl]}$ is obtained with no prior knowledge about energy demand, but directly based on a bunch of actual consumer's demand profiles (training dataset). There are many dimensionality reduction data-driven techniques. They can be divided into two categories: feature selection, which selects some attributes while discarding the others (Cai et al., 2018); and feature extraction which combine the original attributes to obtain a lower number of components (Huang et al., 2019).

The results obtained using feature selection can be straightforwardly interpreted, as the resulting attributes are a subset of the original ones. However, they commonly show a significant higher error during the reduction process. For this reason, this paper is focused on feature extraction techniques which, can additionally be categorized into linear and non-linear methods (Ayesha et al., 2020). Among them, some of the most significant representantives of each category have been selected: Principal Component Analysis (PCA) as a linear transformation, and Autoencoders exemplifying the non-linear methods.

Therefore, four dimensionality reduction techniques will be described in the following paragraphs. Thus, the reduction and recovery processes (in blue), their interaction with the proposed two approaches (in red) are depicted in Fig. 3.

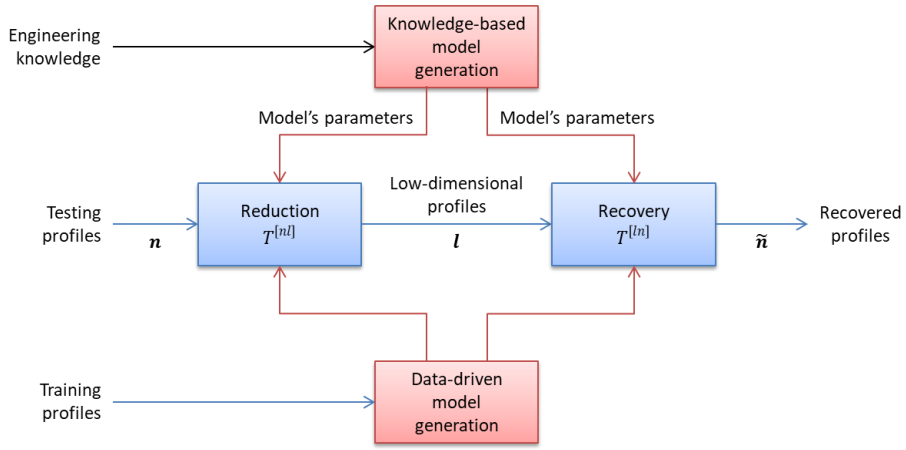

Fig. 3. Scheme of knowledge-based and data-driven dimensionality reduction techniques.

4.1. Month-Based Transformation.

A simple and intuitive way to represent profiles in two-dimensions is one that uses the month of the year to determine the interpretation of the horizontal and vertical axes. Thus, for example, the horizontal axis can be designated as "demand on summer" and

the vertical one as "demand in winter". This transformation includes a previous z-score normalization step.

This approach, which will be called “Month-Based Transformation” (MBT), can be characterized by a 12×2 explanation matrix $\boldsymbol{E} = \{e_{jc}\}$, with its horizontal element defined by $e_{j1} = \frac{1}{2}\left[1 - \cos\left(2\pi \frac{j-1}{12}\right)\right]$, while the vertical one is defined by $e_{j2} = \frac{1}{2}\left[1 + \cos\left(2\pi \frac{j-1}{12}\right)\right]$.

The reverse axis-construction vectors for this explanation matrix, $\boldsymbol{a}_c^r, c \in \{1,2\}$, are depicted in Fig. 4. They can be thought as the level of *summerness* (horizontal axis) and *winterness* (vertical axis) of each month.

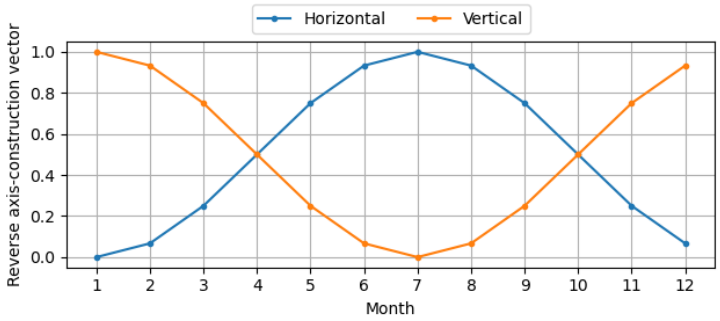

Fig. 4. Reverse axis-construction vectors of the Month-Based Transformation.

Instead of using 2 vectors of 12 elements, this information can alternatively be represented using 12 biplot vectors, $\boldsymbol{b}_j^r, j \in [1,12]$, as it is shown in Fig. 5. There, each vector corresponds to a month, from January (1) to December (12).

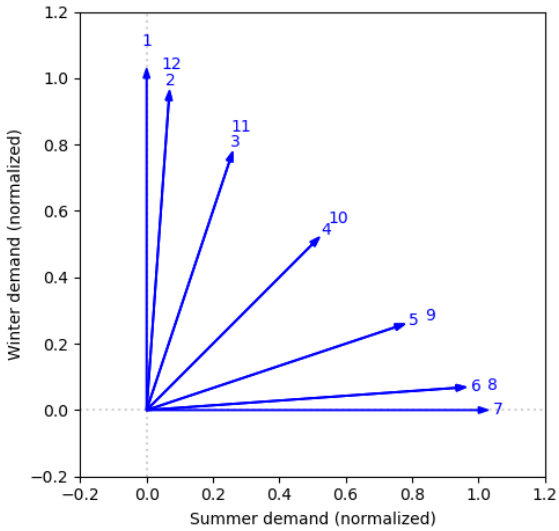

Fig. 5. Reverse biplot vectors of the Month-Based Transformation.

The horizontal/vertical component of the j-th vector represents to what extent the horizontal/vertical component of a bidimensional representation defines the electricity demand of the j-month. More formally, $\boldsymbol{b}_j^r$ is a bidimensional vector, corresponding to the j-month, with a horizontal component b_{jx}^r and a vertical component b_{jy}^r. The normalized demand of the k-th NACE-location pair during the j-th month is $n_j^{(k)}$. Its bidimensional transformation, $\boldsymbol{n}^{(k)} \xrightarrow{T^{[nl]}} \boldsymbol{l}^{(k)}$, obtains a vector $\boldsymbol{l}^{(k)} = \left(l_x^{(k)}, l_y^{(k)}\right)$ of dimension 2. Then, $n_j^{(k)} \approx \tilde{n}_j^{(k)} = b_{jx}^r l_x^{(k)} + b_{jy}^r l_y^{(k)}$.

For the k-th normalized demand profile, represented as a 12-element row vector $\boldsymbol{n}^{(k)}$, the MBT transformation, $\boldsymbol{n}^{(k)} \xrightarrow{T^{[nl]}} \boldsymbol{l}^{(k)}$, can be expressed, recalling the Eq. (2), as $\boldsymbol{l}^{(k)} = \left(\boldsymbol{n}^{(k)} - \boldsymbol{\mu}_N\right)\boldsymbol{S} + \boldsymbol{\mu}_L$, obtaining a 2-element row vector. This transformation is preceded by a z-score normalization, so $\boldsymbol{\mu}_N = \boldsymbol{0}$. Furthermore, the bias vector for this transformation is chosen as $\boldsymbol{\beta} = \boldsymbol{0}$, then $\boldsymbol{\mu}_L = \boldsymbol{0}$. So, MBT can finally be expressed as $\boldsymbol{l}^{(k)} = \boldsymbol{n}^{(k)}\boldsymbol{S}$. The inverse transformation, $\boldsymbol{l}^{(k)} \xrightarrow{T^{[nl]}} \tilde{\boldsymbol{n}}^{(k)}$, can be expressed as $\tilde{\boldsymbol{n}}^{(k)} = \boldsymbol{l}^{(k)}\boldsymbol{S}^{+} = \boldsymbol{l}^{(k)}\boldsymbol{E}$, obtaining a 12-element row vector that estimates the normalized demand profile.

In this technique, the qualitative interpretation of the low-dimensional axis is straightforward, and therefore, it does not require axis inversion: the horizontal and vertical axis represent, respectively, the summer and the winter demand. Then, $\boldsymbol{m}^{(k)} = \boldsymbol{l}^{(k)}$.

For the scaling and shifting transformation, the results obtained in Section 3.4 are applied, considering that the set of months that are explained by each axis are, respectively, $A_x = \{7\}$ and $A_y = \{1\}$. The corresponding constants are derived using Eq. (5) and Eq. (6).

As it was commented above, the Month-Based Transformation can be also adapted for the representation of profiles in just one-dimension. The one-dimensional MBT can be characterized by a 12×1 explanation matrix $\boldsymbol{E} = \{e_j\}$, with its elements defined by $e_j = \frac{1}{2}\left[1 - \cos\left(2\pi\frac{j-1}{12}\right)\right]$. The corresponding reverse axis-construction vector is the same that the horizontal component depicted in Fig. 4. The values in this single-axis can be thought as the level of *summerness* of each month, and it is equivalent to the horizontal axis of the 2D representation.

The qualitative interpretation of the only axis is straightforward and does not require axis inversion: $\boldsymbol{m}^{(k)} = \boldsymbol{l}^{(k)}$. And, as in the previous description of the 2D case, the constants associated with the scaling and shifting of the only axis are derived for $A_x =$

{7} using Eq. (5) and Eq. (7). All the constants associated with these transformations are summarized in Table 1.

Table 1. Interpretation of the MBT and its related constants.

Dimensions	Axis	γ	A	λ	η	Meaning
2	Horizontal	1	{7}	0.264	1.069	Demand on summer
	Vertical	1	{1}	0.266	1.115	Demand on winter
1	Single-axis	1	{7}	0.264	1.069	Demand on summer

The MBT algorithm can be extended to reductions in dimensionality where the number of the resulting components is greater than 2, $d > 2$. For this purpose, the cosine function used to model this transformation is modified assigning a frequency f_c and phase φ_c to the each c-th component. More formally,

$$e_{jc} = \frac{1}{2}\left\{1 - \cos\left[f_c\left(2\pi\frac{j-1}{12} + \varphi_c\right)\right]\right\}, \forall c \in [1, d], j \in [1,12]. \quad (8)$$

where $\varphi_c = 2\pi\frac{c-1}{d}$, and $f_c = \left\lfloor\frac{c-1}{2}\right\rfloor + 1$, while $\lfloor\cdot\rfloor$ represents the function "floor". An example of this generalization for $d = 5$ is shown in Fig. 6.

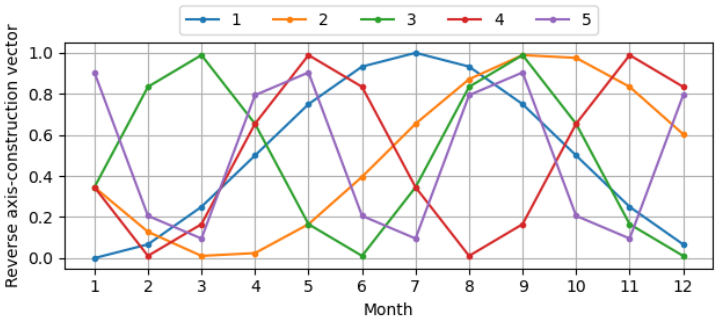

Fig. 6. Reverse axis-construction vectors of the Month-Based Transformation for a reduction to 5 components ($d = 5$).

4.2. Cooling-Heating Degree Days Transformation.

In the Month-Based Transformation, the values of the explanation matrix are derived from the month of the year. This makes sense because the electricity demand depends, among other factors, on the temperature and its variations throughout the year. It is a well-known fact that the demand grows in summer for cooling needs and in winter for heating requirements. Then, instead of the indirect association of the demand with the months, a more direct relationship with the temperature can be explored. This transformation includes a previous z-score normalization step.

For a formal description of these ideas, it is typical to use the concepts of "Heating Degree Day" (HDD) and "Cooling Degree Day" (CDD) (D'Amico et al., 2019). The i-th day of the j-th month is said to require heating (it is a heating day) if its mean

temperature, $T_m^{(i)[j]}$, is lower than a certain threshold temperature (T_h). Analogously, this day is to be a cooling day if $T_m^{(i)[j]} > T_h$. The threshold temperature in Spain has been set to $15ºC$ (UNE, 1988). The Heating Degree Day of the j-th month is defined as

$$\begin{cases} HDD_j = \sum_{i=1}^{d_j} \left(T_h - T_m^{(i)[j]}\right) & \forall i, T_m^{(i)[j]} < T_h \\ CDD_j = \sum_{i=1}^{d_j} \left(T_m^{(i)[j]} - T_h\right) & \forall i, T_m^{(i)[j]} > T_h \end{cases}, \tag{9}$$

where d_j is the number of days of the j-th month.

Averaging the 3-year meteorological data of the four main cities in Spain (Madrid, Barcelona, Valencia and Seville) (BizEE Software, 2022), the distribution of the HDD and CDD for each month is represented in Fig. 7.

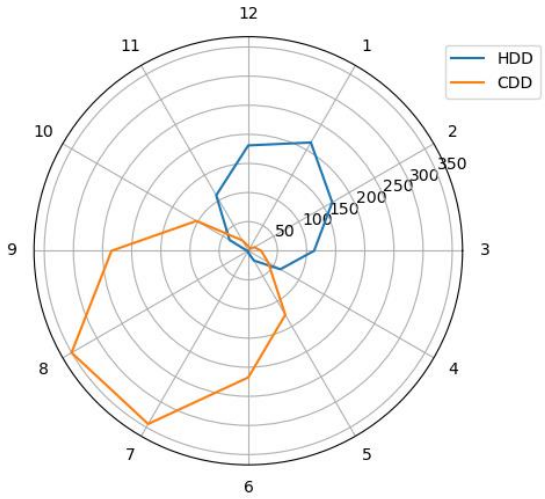

Fig. 7. Heating and Cooling Degree days in Spain.

This graph shows a relationship between the months and the electricity demand which can be used to measure *summerness* and *winterness* in a more precise way that the formula used in MBT. Following this idea, it is proposed a Cooling-Heating Degree Days Transformation (CHDDT) which is characterized by the explanation matrix $\boldsymbol{E} = \{e_{jc}\}$, with its elements horizontal component is defined as $e_{j1} = \text{range}_{01}(CDD_j)$, and the vertical one as $e_{j2} = \text{range}_{01}(HDD_j)$. The function $\text{range}_{01}(\cdot)$ scale the values of the argument to the $[0,1]$ range.

The reverse axis-construction vectors corresponding to this explanation matrix are depicted in Fig. 8. They can be thought of as the level of cooling-related (horizontal axis) and heating-related (vertical axis) demand of each month.

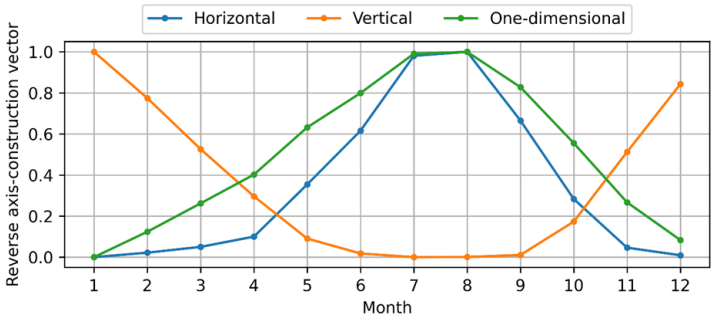

Fig. 8. Reverse axis-construction vectors of the Cooling-Heating Degree Days Transformation.

The k-th normalized demand profile is reduced to 2 dimensions using the expression $\boldsymbol{l}^{(k)} = \boldsymbol{n}^{(k)}\boldsymbol{S}$, where $\boldsymbol{S}$ is the sensitivity matrix which can be obtained as the pseudoinverse of the explanation matrix, $\boldsymbol{S} = \boldsymbol{E}^{+}$.

Given the previous definitions, the qualitative interpretation of the low-dimensional axis is straightforward, and does not require axis inversion: The horizontal and vertical axis represent the summer and the winter demand respectively. Then, $\boldsymbol{m}^{(k)} = \boldsymbol{l}^{(k)}$.

For the scaling and shifting transformation, the results obtained in Section 3.4 are applied, considering that the set of months that are explained by each axis are, respectively, $A_x = \{7,8\}$ and $A_y = \{1,2,12\}$. The corresponding constants are derived using Eq. (5) and Eq. (6).

The adaptation of the CHDDT to obtain one-dimensional profiles is made using a 12×1 explanation matrix, $\boldsymbol{E} = \{e_j\}$, with its elements defined by the expression $e_j = \text{range}_{01}(CDD_j - HDD_j)$. The corresponding reverse axis-construction vector is depicted in Fig. 8. The values in this axis can be thought as the level of electricity over-demand related to cooling/heating, where higher values represent cooling over-demand and lower values mean heating over-demand.

The qualitative interpretation of the axis is straightforward and does not require axis inversion: $\boldsymbol{m}^{(k)} = \boldsymbol{l}^{(k)}$. And, as in the previous description of the 2D case, the constants associated with the scaling and shifting of the only axis are derived for $A_x = \{7,8\}$ using Eq. (5) and Eq. (7). All the constants associated with this transformation are summarized in Table 2.

Table 2. Interpretation of the CHDDT and its related constants.

Dimensions	**Axis**	$\boldsymbol{\gamma}$	$\boldsymbol{A}$	$\boldsymbol{\lambda}$	$\boldsymbol{\eta}$	**Meaning**
2	Horizontal	1	{7,8}	0.285	1.063	Demand related to cooling
	Vertical	1	{1,2,12}	0.215	1.098	Demand related to heating
1	Single-axis	1	{7,8}	0.286	1.063	Demand related to cooling(↑) or heating(↓)

The CHDDT algorithm can also be extended to reductions in dimensionality where $d > 2$. For this purpose, the cosine function used to model this transformation is modified assigning a frequency f_c and phase φ_c to the each c-th component. More formally,

$$e_{jc} = \text{chdd}\{\lfloor f_c(j-1) - \varphi_c \rfloor \bmod 12\}, \forall c \in [1,d], j \in [1,12]. \quad (10)$$

where $\varphi_c = \left\lfloor 12\frac{c}{d} + 4 \right\rfloor$, and $f_c = \left\lfloor \frac{c}{2} \right\rfloor + 1$, while $\lfloor \cdot \rfloor$ represents the function "floor", and the function $\text{chdd}(k) = \text{range}_{01}(CDD_k - HDD_k)$. An example of this generalization for d = 5 is shown in Fig. 9.

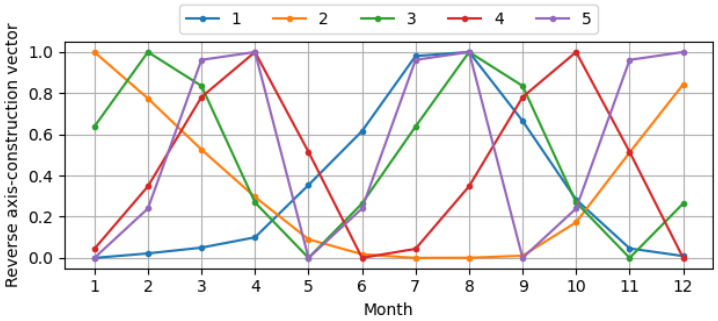

Fig. 9. Reverse axis-construction vectors of the Cooling-Heating Degree Day Transformation (CHDDT) for a reduction to 5 components ($d = 5$).

4.3. Principal Component Analysis.

The dimensionality reduction methods described above have the advantage that they allow for an easy explanation of the results obtained, either as a function of the month or the temperature, mainly for 2D and 1D representations. However, if the explainability requirement is relaxed (at least in a first step), it is possible to use other data-driven dimensionality reduction techniques, which will show better performance. Among these techniques, Principal Component Analysis (PCA) is one of the most widely used (Vidal et al., 2016). Briefly, PCA transforms the 12 dimensions of a z-score normalized electricity demand profile into another profile of 12 dimensions (orthogonal to each other), called components. These components are then ordered from their highest to lowest contribution (usually called explained variance) to recreate the original profile. Finally, only the d first components are retained, thus, reducing the profile dimensionality.

The result of applying the PCA to the profiles in the training dataset is shown in Fig. 10, where, the explained variance of each component is represented by a set of blue bars and the accumulated variance is drawn using a red line (Luque et al., 2021). As can be seen, using just two components (out of 12, 17% of them), up to 63% of the variance is explained, while using only one component (an 8% of them) a 50% is explained.

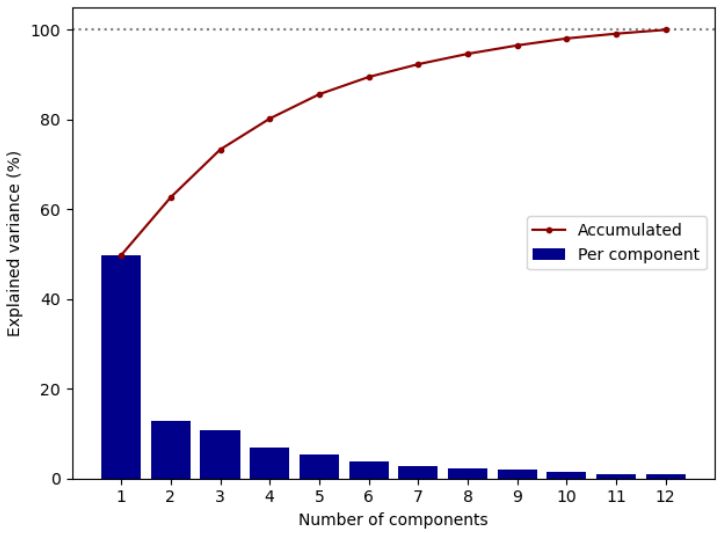

Fig. 10. PCA explained variance of electricity demand profiles.

The PCA transformation $\boldsymbol{n}^{(k)} \xrightarrow{T^{[nl]}} \boldsymbol{l}^{(k)}$ is modeled using a 12×12 square matrix $\boldsymbol{W} = \{w_{uv}\}, u, v \in [1,12]$, where only the d first rows are retained. The resulting $d \times 12$ matrix, $\boldsymbol{W_d} = \{w_{uv}\}, u \in [1,12], v \in [1, d]$, plays the role of the explanation matrix, that is, $\boldsymbol{E} = \boldsymbol{W_d}$. It can be written that $\boldsymbol{n}^{(k)} = \boldsymbol{l}^{(k)}\ \boldsymbol{E} = \boldsymbol{l}^{(k)}\boldsymbol{W_d}$. The reverse axis-construction vectors corresponding to this explanation matrix for the case $d = 2$ are depicted in Fig. 11.

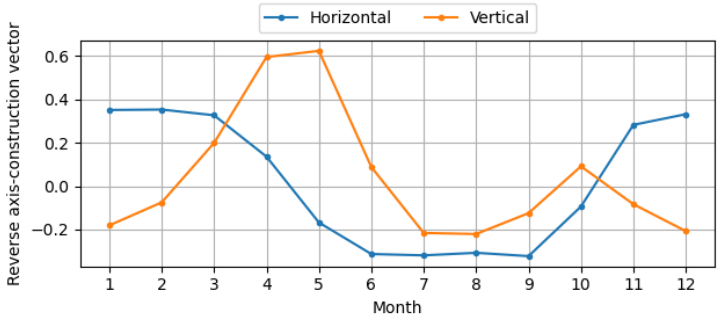

Fig. 11. Reverse axis-construction vectors of the PCA transformation.

As it can be seen in this figure, the horizontal component can be thought of as the winter-related (normalized) demand because it indicates an over-demand in winter and an infra-demand in summer. To follow the same convention as in the MBT and CHDDT, the horizontal axis is inverted to show positive values for summer and negative ones for winter.

Analogously, the vertical component shows the maximum positive values in spring while the most negative values are in summer and winter. As the electricity demand usually increases in summer and winter due to the cooling and heating requirements, the vertical axis is also inverted, now having the meaning of a cooling/heating-related (normalized) demand.

With these inversions, $\boldsymbol{m}^{(k)} = -\boldsymbol{l}^{(k)}$. For the scaling and shifting transformation, the results obtained in Section 3.4 are applied, considering that the set of months that are explained by each axis are, respectively, $A_x = \{6,7,8,9\}$ and $A_y = \{1,7,8,12\}$. The corresponding constants are derived using Eq. (5) and Eq. (6).

The reverse axis-construction vector for the one-dimensional case corresponds to the horizontal axis, as it is shown in Fig. 11. The values in this single dimension can be thought of as the winter-related (normalized) demand, as it indicates an over-demand in winter and an infra-demand in summer. To follow the same convention as in MBT and CHDDT, the horizontal axis is inverted to show positive values for summer and negative ones for winter, $\boldsymbol{m}^{(k)} = -\boldsymbol{l}^{(k)}$.

As in the previous description of the 2D case, the constants associated with the scaling and shifting of the only axis are derived for $A_x = \{6,7,8,9\}$ using Eq. (5) and Eq. (7). All the constants associated with this transformation are summarized in Table 3.

Table 3. Interpretation of the PCA transformation and its related constants.

Dimensions	**Axis**	$\boldsymbol{\gamma}$	$\boldsymbol{A}$	$\boldsymbol{\lambda}$	$\boldsymbol{\eta}$	**Meaning**
2	Horizontal	-1	{6,7,8,9}	0.072	1.027	Demand related to summer
	Vertical	-1	{1,7,8,12}	0.055	1.082	Demand related to cooling/heating
1	Single-axis	1	{6,7,8,9}	0.072	1.027	Demand related to summer(+) or winter(-)

4.4. Autoencoders.

The three previous techniques for dimensionality reduction are based on a sequence of linear transformations. To test whether a nonlinear approach could enhance the performance of the dimensionality reduction, an autoencoder has also been tested. An autoencoder is a sequence of two neural networks, an encoder followed by a decoder (Yasi Wang et al., 2016). The encoder section accepts a vector of 12 values, $\boldsymbol{p}^{(k)}$, (the monthly demand profile) and converts it into a vector of dimension $d < 12$, $\boldsymbol{l}^{(k)}$, its low-dimensional representation. The decoder section accepts a vector of d values, $\boldsymbol{l}^{(k)}$ (the low-dimensional representation of the demand profile) and converts it into the corresponding vector of dimension 12, $\widetilde{\boldsymbol{p}}^{(k)}$. The 2 neural networks are trained to minimize the mean error between $\boldsymbol{p}^{(k)}$ and $\widetilde{\boldsymbol{p}}^{(k)}$. Then, once trained, the decoder section is discarded, and the coder section is used to reduce dimensionality. The typical architecture of an autoencoder is depicted in Fig. 12.

An autoencoder may have many possible architectures varying, among other considerations, the number of hidden layers and the number of neurons at each layer. Then, to properly choose these hyperparameters, the training dataset (containing 360,567 profiles) is split in a more reduced training dataset (320,504 profiles, an 80% of

the total valid ones), and a validation dataset (40,063 profiles, a 10% of the total). The training dataset is used to train the different neural network architectures, while the validation dataset is employed to select the best architecture (and other related hyperparameters).

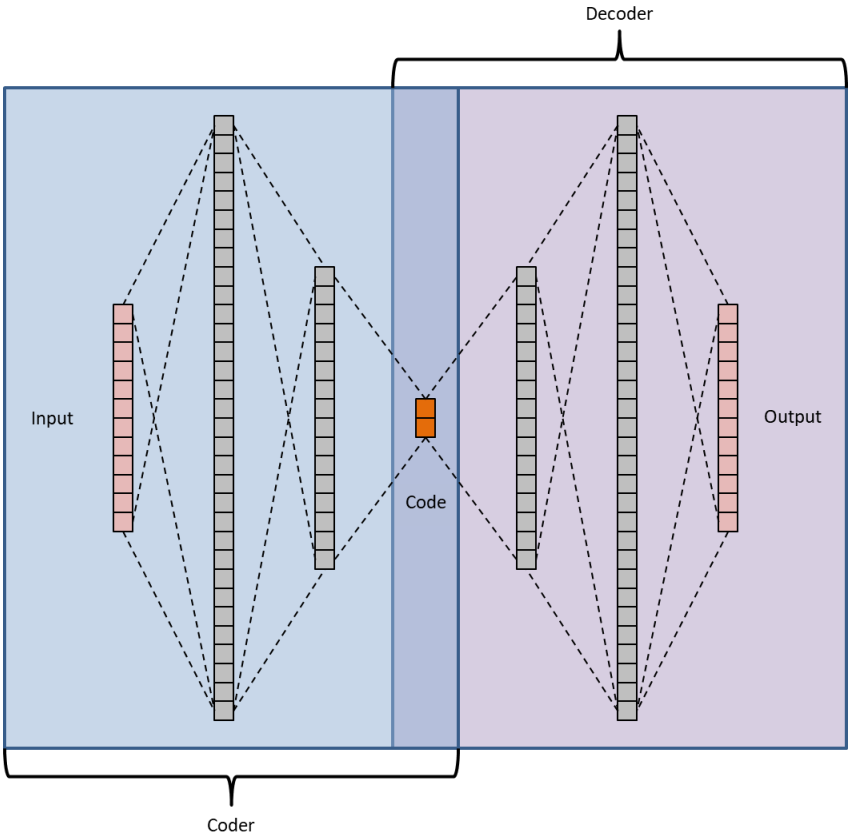

Fig. 12. Typical architecture of an autoencoder neural network.

Different architectures have been tried, from 1 to 5 hidden layers in each section, and from 8 to 512 neurons per layer, performing a grid search for 2 or less hidden layers and a random search for 3 or more. The effect of the z-score normalization as a preprocessing step has also been tested. The resulting architecture has been a single hidden layer with 128 neurons and no normalization preprocessing.

The reverse axis-construction vectors corresponding to the trained autoencoder for the 2D case, $d = 2$, are shown in Fig. 13.

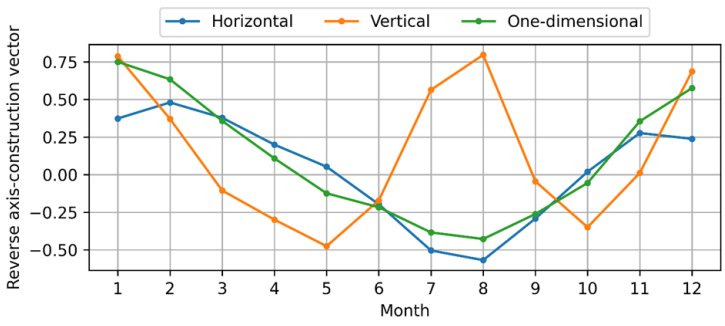

Fig. 13. Reverse axis-construction vectors of the autoencoder transformation.

As it can be seen in this figure, and in a similar way as it happened in the transformation based on PCA, the horizontal component can be thought as the winter-related demand because it indicates an over-demand in winter and an infra-demand in summer. To follow the same convention as in the MBT and CHDDT, the horizontal axis is inverted to show positive values for summer and negative ones for winter.

Analogously, for this transformation, the vertical component shows the maximum positive values in summer and winter. In this sense, as the electricity demand usually increases in summer and winter due to the cooling and heating requirements, the vertical axis is not inverted in this case, having the meaning of a cooling/heating-related demand.

With these inversions, $\boldsymbol{m}^{(k)} = \boldsymbol{\gamma} \circ \boldsymbol{l}^{(k)}$, where $\boldsymbol{\gamma} = [-1 \quad 1]$. In this sense, on the one hand, the reverse axis-construction vectors are very connected to the concepts of CDD and HDD. The horizontal axis highly resembles the evolution of CDD by months. On the other hand, the vertical axis has high values in the months that require cooling or heating. So, it should resemble the CDD+HDD evolution. These evolutions are depicted in the Fig. 14 where, for comparison, all the values have been converted to the [0,1] range.

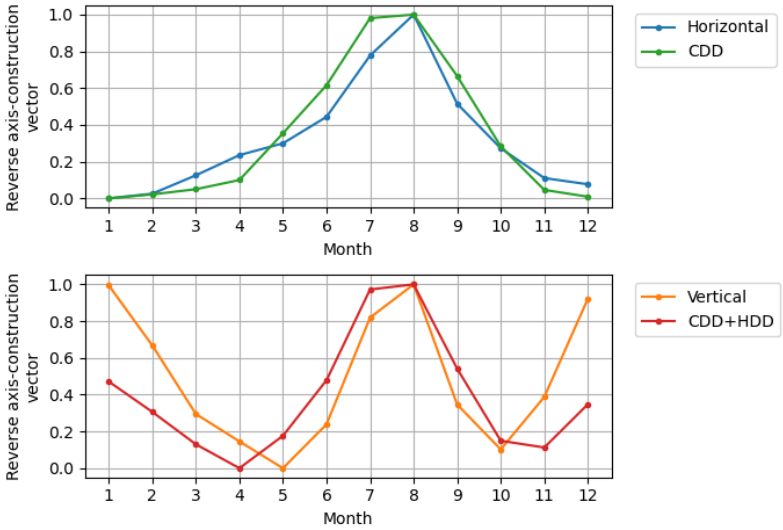

Fig. 14. The reverse axis-construction vectors of the autoencoder and their similarity to the CDD (horizontal component) and the summation of the CDD and HDD (vertical component).

For the scaling and shifting transformation, the results obtained in Section 3.4 are applied, considering that the set of months that are explained by each axis are, respectively, $A_x = \{8\}$ and $A_y = \{1,8,12\}$. The corresponding constants are derived using Eq. (5) and Eq. (6).

The reverse axis-construction vector for the one-dimensional case is depicted in Fig. 13. The values on this axis can be thought of as the level of electricity demand related to summer/winter, where positive values represent heating demand and negative values

mean cooling demand. This is the opposite convention used in the previous transformation, so an inversion of the single axis is advised. Then $\boldsymbol{m}^{(k)} = -\boldsymbol{l}^{(k)}$.

As in the previous description of the 2D case, the constants associated with the scaling and shifting of the single axis are derived for $A_x = \{7,8\}$ using Eq. (5) and Eq. (7). All the constants associated with this transformation are summarized in Table 4.

Table 4. Interpretation of the autoencoder transformation and its related constants.

Dimensions	Axis	γ	A	λ	η	Meaning
2	Horizontal	-1	{8}	0.843	1.002	Demand related to summer
	Vertical	-1	{1,8,12}	0.751	0.108	Demand related to cooling/heating
1	Single-axis	1	{7,8}	0.405	1.916	Demand related to summer(+) or winter(-)

5. Results.

The process described in the methodology Section and depicted in Fig. 1 has been fully implemented using a desktop computer with a processor based on an Intel® Core™ i7-11700 @ 2.50GHz processor, with 64 GB of RAM and a Solid State Disk (SSD) of 2TB. It has been developed in Python 3.9 under a Windows 11 operating system. For PCA analysis the scikit-learn 1.0 library has been used, while the development of autoencoders relies on routines defined in TensorFlow 2.10.

As for the profile transformation, each of the four dimensionality reduction techniques described in the previous section has been applied to a set of six actual profiles. These profiles have been selected as representative cases (with different illustrative shapes) which depict several types of electricity demand and each of them corresponds to a certain NACE-location pair:

A. Sewerage activities (NACE: E37) in the city of Águilas, Murcia. It has a single maximum demand in summer.
B. Retail sale of medical and orthopedic goods in specialized stores (NACE: G4774) in the city of Riola, Valencia. It has a maximum demand in summer and a lower relative maximum demand in winter.
C. Other manufacturing not elsewhere classified (NACE: C3299) in the city of Carmena, Toledo. It has a maximum demand in winter and a lower relative maximum demand in summer.
D. Development of building projects (NACE: F411) in the city of Andoain, Guipúzcoa. It has a single minimum demand in summer.
E. Manufacture of lime and plaster (NACE: C2352) in the city of Bilbao, Vizcaya. It has a minimum demand in summer and a lower relative minimum demand in winter.
F. Restaurants and mobile food service activities (NACE: I561) in the city of Aduna, Guipúzcoa. It has a minimum demand in winter and a lower relative minimum demand in summer.

These profiles are summarized in Table 5 and depicted in Fig. 15.

Table 5. Summary of the profiles used as examples.

Profile	NACE	Location	Demand on				Shape
			Summer	Winter	Spring	Autumn	
A	E37	Águilas Murcia	Max	Min	-	-	Summit
B	G4774	Riola Valencia	Max	Relative max	Min	Min	Mexican hat
C	C3299	Carmena Toledo	Relative max	Max	Min	Min	W-shaped
D	F411	Andoain Guipúzcoa	Min	Max	-	-	V-shaped
E	C2352	Bilbao Vizcaya	Min	Relative min	Max	Max	Seagull
F	I561	Aduna Guipúzcoa	Relative min	Min	Max	Max	M-shaped

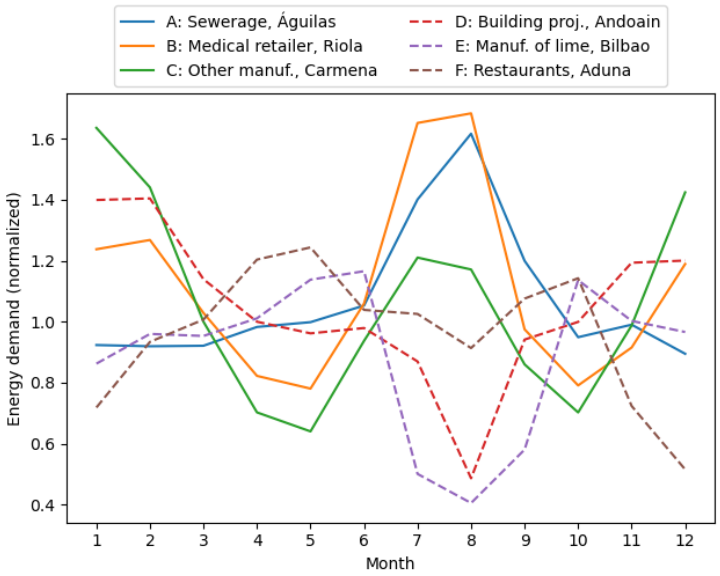

Fig. 15. Examples of electricity demand profiles.

As it can be seen, profiles A and D show approximately one cycle per year, while the remaining profiles have two cycles per year. Moreover, profiles D, E, and F (dashed lines in the figure), have an approximately inverse shape than profiles A, B, and C, respectively (continuous lines).

Each of these profiles has been reduced to two-dimensions (2D) using every dimensionality reduction technique. The result is depicted in the upper part of Fig. 16 where the example profiles are coded by colors, and the reduction technique is identified by different symbols. The reduction to a single dimension (1D) is also depicted in the lower part of the figure.

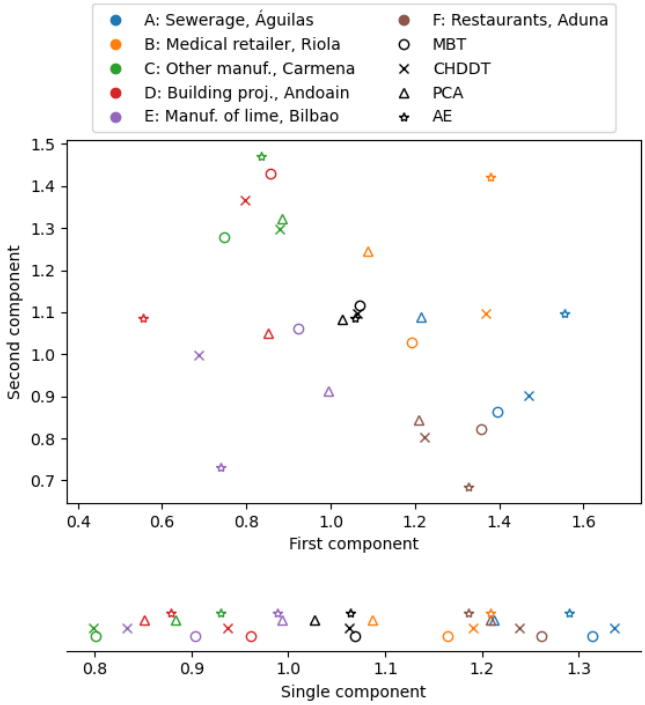

Fig. 16. Representing the examples of electricity demand profiles in two dimensions (upper graph) and one single dimension (lower graph).

The interpretation of the axis is summarized in Table 6. In 2D representations, the horizontal axis always means demand in summer, while the vertical axis may represent demand in winter (for MBT and CHDDT) or the combined demand in summer and winter (for PCA and AE). In the 1D representations, the single component increases its value according to the demand in summer, and decreases it in relation to the demand in winter.

The 2D representation of profile A (blue markers) has horizontal values over the mean (high demand in summer), vertical values below the mean for MBT and CHDDT (lower demand in winter), and vertical values on the mean for PCA and AE (combined demand in summer and winter). The 1D representation of profile A (blue markers in the lower part of the figure) has horizontal values over the mean (high demand in summer combined with low demand in winter). The symmetric behavior is observed in the 2D and 1D representations of profile D (red markers). Analogous interpretations can be derived for the remaining profiles.

Table 6. Interpretation of the axis in low-dimensional representation of demand profiles.

		Technique	
		MBT/CHDDT	**PCA/AE**
Axis	**Horizontal**	Demand in summer	
	Vertical	Demand in winter	Demand in summer and winter
	Single-axis	Demand in summer(↑) and winter(↓)	

Extending the results of reducing profiles to 2D, a heatmap showing the density of profiles in each region of the plane is depicted in Fig. 17. As it can be observed in this

figure, MBT and CHDDT have a high concentration of profiles in the anti-diagonal direction, meaning that high demands in summer correlates with low demands in winter and vice versa (negative correlation). Conversely, PCA and AE density maps are more regular-shaped, resembling some kind of ellipse, with a greater variance in the horizontal axis (demand in summer) than in the vertical one (combined demand in summer and winter).

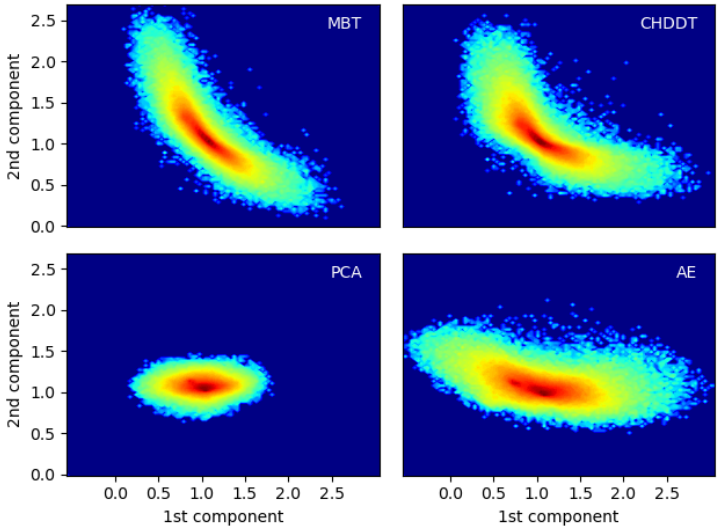

Fig. 17. Density of profiles in the 2D representation according to each dimensionality reduction technique.

The equivalent information after reducing profiles to 1D is shown in Fig. 18, where the probability density functions (pdf) for each technique are plotted in log scale. As it can be seen, all four methods have very similar pdf as they share the same interpretation of the single component. However, MBT and CHDDT reach higher values, while PCA may have slightly lower values.

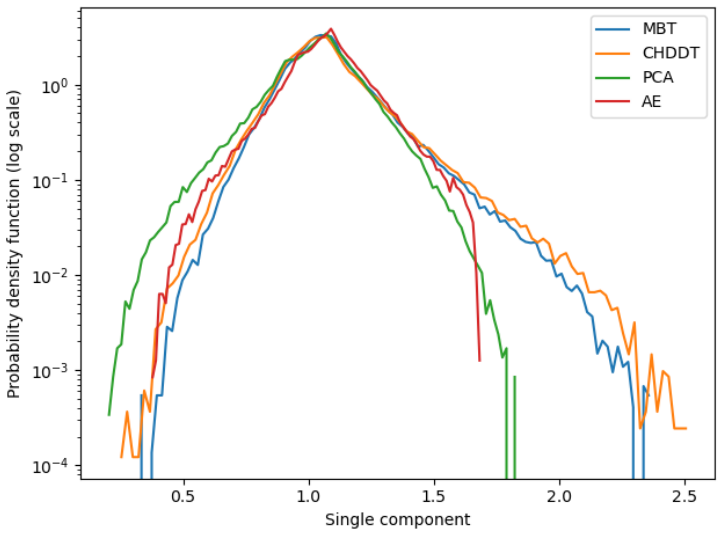

Fig. 18. Probability density functions of the 1D representation of profiles, according to each dimensionality reduction technique.

6. Discussion.

6.1. Error analysis.

Due to the process of reducing dimensionality, some information is inevitably lost. So, the complete round-trip process, $\boldsymbol{p}^{(k)} \xrightarrow{T^{[pq]}} \boldsymbol{q}^{(k)} \xrightarrow{T^{[qp]}} \tilde{\boldsymbol{p}}^{(k)}$, does not recover the original profile $\boldsymbol{p}^{(k)}$, but an estimation of it, $\tilde{\boldsymbol{p}}^{(k)}$, that is, $\tilde{\boldsymbol{p}}^{(k)} \approx \boldsymbol{p}^{(k)}$. Information lost in the dimensionality reduction process can be measured by the error of the round-trip transformation, that is, by the difference $\boldsymbol{e}^{(k)} = \boldsymbol{p}^{(k)} - \tilde{\boldsymbol{p}}^{(k)}$, which is a 12-valued vector.

That is, the performance of a dimensionality reduction technique is revealed by comparing the low-dimensionality profiles, $\tilde{\boldsymbol{p}}^{(k)}$, with the original 12-dimensional profile $\boldsymbol{p}^{(k)}$ (without dimensionality reduction). The analysis of that comparison (the analysis of the error $\boldsymbol{e}^{(k)} = \boldsymbol{p}^{(k)} - \tilde{\boldsymbol{p}}^{(k)}$) is carried out in detail in this subsection.

Specifically, the error vectors corresponding to the example profiles (analyzed in previous section) are depicted in Fig. 19.

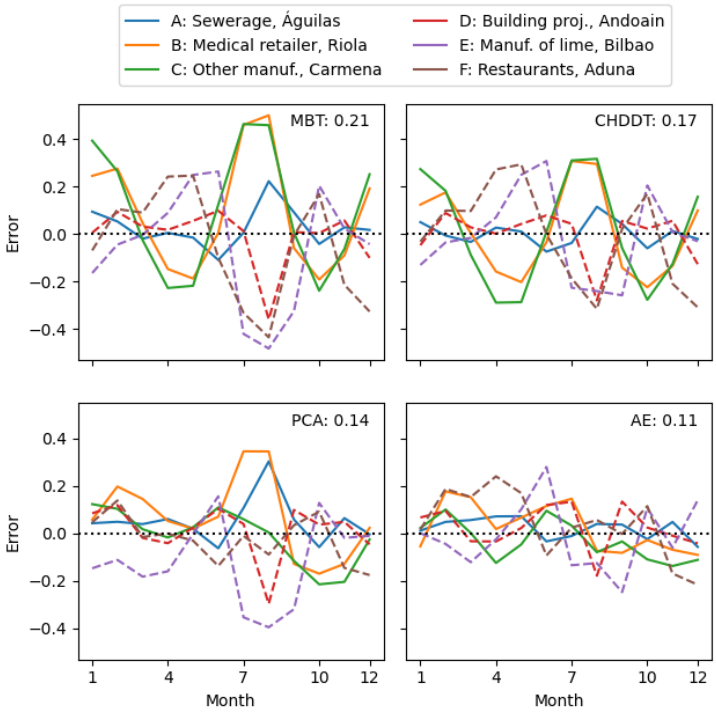

Fig. 19. Error of reducing the example monthly profiles to 2D. The mean error for each technique is also shown.

A single metric for the recovery process of the k-th pair, can be obtained using the Root Mean Square Error (RMSE) defined as

$$RMSE^{(k)} = \sqrt{\frac{1}{12}\sum_{j=1}^{12}\left(e_j^{(k)}\right)^2} = \sqrt{\frac{1}{12}\sum_{j=1}^{12}\left(p_j^{(k)} - \tilde{p}_j^{(k)}\right)^2} \tag{11}$$

This metric can be extended to the entire testing dataset using the expression

$$RMSE = \sqrt{\frac{1}{12\,n}\sum_{k=1}^{n}\sum_{j=1}^{12}\left(e_j^{(k)}\right)^2}. \tag{12}$$

The values of this metric for dimensionality reduction to 1D and 2D, are distributed according to the boxplots depicted in Fig. 20. As it can be seen in this figure, the reduction to 2D offers better results than to 1D, which is a logical result as more information is kept in 2D. Comparing the boxplots in the figure, it can also be derived that autoencoders offer the most efficient reduction technique, while the MBT obtains the worst results.

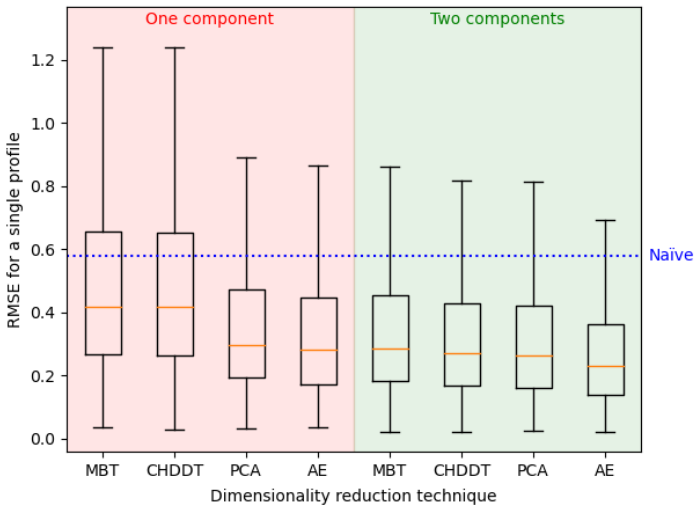

Fig. 20. Error (RMSE) of applying 1D and 2D dimensionality reduction techniques to the profiles in the testing dataset.

To have a clearer insight about the significance of these errors, they will be compared to those errors obtained in the cases where the recovered profiles are, either randomly generated, or estimated in a naïve way.

For the random generation of a profile, the value of the energy demand corresponding to the j-th month, $\tilde{p}_j^{(k)}$, is obtained by a random sampling which considers the statistical distribution of the demands in this month. The probability density function corresponding to several months is depicted in Fig. 22. Applying this random generation of the profiles, the error is characterized by a value of the $RMSE = 2.511$.

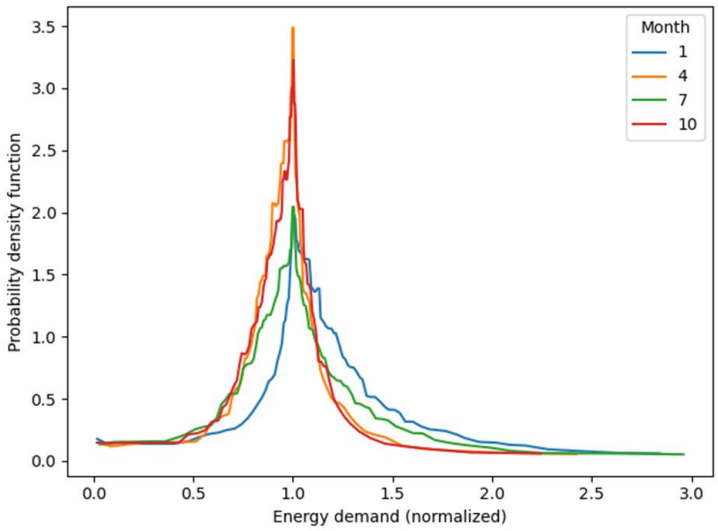

Fig. 21. Probability density functions of the energy demand corresponding to several months.

Also, in the naïve approach, the recovered profile is estimated to be equal to the mean demand, $\tilde{\boldsymbol{p}}^{(k)} = \boldsymbol{\mu}_P$, regardless of the value of $\boldsymbol{p}^{(k)}$. In this case the error is characterized by a value of the $RMSE = 0.581$, a value which is also depicted in the Fig. 20.

As it can be seen, any of the four dimensionality reduction techniques explored in the research clearly outperforms the random and naïve approaches. The best results are obtained using an autoencoder, showing a recovery error for the two-components case of $RMSE = 0.278$. As the electricity demand profiles are normalized, its mean value is 1, so the absolute and relative values of $RMSE$ are the same.

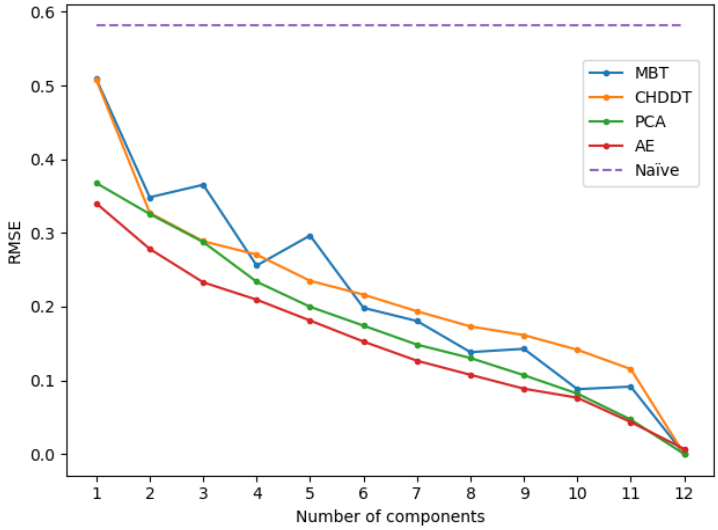

Fig. 22. Error (RMSE) of applying dimensionality reduction techniques to the profiles in the testing dataset for different number of components in the target space.

This error analysis can be extended to dimensionality reduction techniques with other number of target components (not just 1D or 2D). The results of this extension are shown in Fig. 22 where cases based on autoencoders clearly outperform the other techniques, except for the degenerate case where $d = 12$. Additionally, it can be observed that PCA is also better (or equal) than the more intuitive MBT and CHDDT techniques.

To assess the robustness of the previous error analysis, a bootstrap technique (Zoubir and Iskandler, 2007) has been applied to the values of the RMSE. Up to 1000 testing datasets are generated from the original testing datasets by a random sampling with replacement process. The RMSE values for each testing dataset are then statistically distributed. Its probability density function is depicted in Fig. 23 for different techniques and number of components. As it can be seen, in all cases, the RMSE values are distributed in a narrow region, which is mainly due to the high number of profiles in the testing datasets (40,063 profiles).

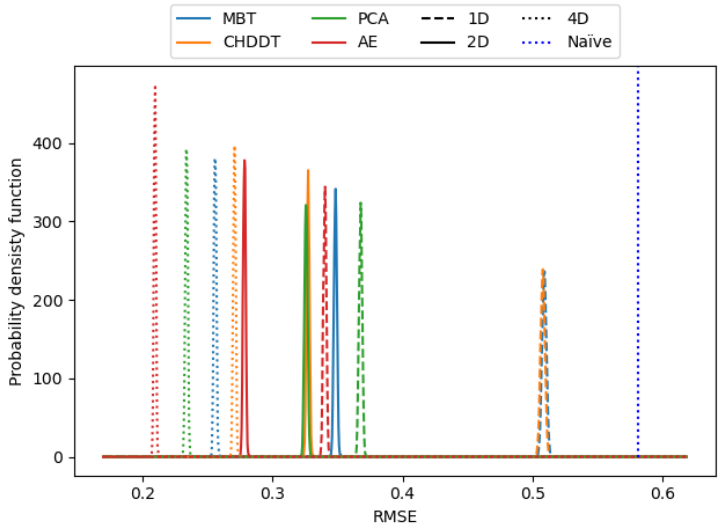

Fig. 23. Bootstrap analysis of the recovery error: probability density function of the RMSE after applying dimensionality reduction techniques to the profiles in the testing dataset.

From these distributions, the confidence intervals (CI) for the values of the RMSE can be derived (DiCiccio and Science, 1996). The results for a 95% confidence level are shown in Fig. 24, where the values of the CI are expressed in relation to the RMSE, that is, as a relative value of the CI in %. For example, the recovery error of an autoencoder using a 2D target space is $RMSE = 0.278 \pm 0.7\%$.

A different approach to assess the confidence of the conclusions obtained in the error analysis is to apply an analysis of variance (ANOVA) technique (Tabachnick and Fidell, 2007). For every value of the number of components in the target space (d), and for every pair of dimensionality reduction technique, an ANOVA test is performed. Each case tests if the results obtained for a certain pair of techniques have been generated by the same statistical distribution (null-hypothesis, H_0). Then the probability

of H_0 (p_{H_0}) is computed and the null-hypothesis is rejected if this probability is lower than a certain significance level, α, usually chosen as $\alpha = 5\%$. The results of the ANOVA tests show that the null hypothesis can be rejected in all cases except for the pair MBT-CHDDT with one component ($p_{H_0} = 62\%$), and the pair CHDDT-PCA with two components ($p_{H_0} = 31\%$) and with three components ($p_{H_0} = 49\%$). This result is consistent with the non-distinguishable points in Fig. 22.

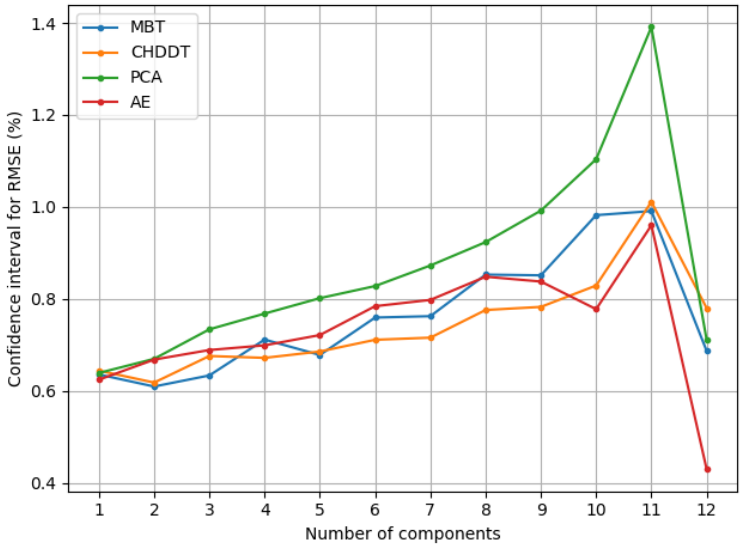

Fig. 24. Confidence interval of the RMSE (in % of the RMSE) for a confidence level of 95%.

6.2. Distance analysis.

A different approach to assess the dimensionality reduction process is to measure its ability to preserve relative distances (Yang, 2011). The intuitive idea is that points that are close in the original 12D space should also be close in the low-dimensional space, while points that are originally distant should be kept apart in the target space.

Let us consider two points (profiles), $\boldsymbol{p}^{(A)}$ and $\boldsymbol{p}^{(B)}$, in the original space which are transformed in the corresponding low-dimensional points $\boldsymbol{q}^{(A)}$ and $\boldsymbol{q}^{(B)}$. The distance between the original points can be measured by the Euclidean norm of their difference, that is, $\delta_p^{(A,B)} = \left\|\boldsymbol{p}^{(A)} - \boldsymbol{p}^{(B)}\right\|_2$. Analogously, the distance between these two points in the target space is defined as $\delta_q^{(A,B)} = \left\|\boldsymbol{q}^{(A)} - \boldsymbol{q}^{(B)}\right\|_2$.

Let us now apply these definitions to the case of reducing dimensionality to two components using an autoencoder. Then, for every pair of points (A, B), their distances in the original and target spaces are obtained and located in a plane with δ_p in the horizontal axis and δ_q in the vertical one. Finally, the density of pairs in each region of the (δ_p, δ_q) plane is shown as a heatmap in Fig. 25. As it can be seen in this figure, the densest regions are approximately around a straight line, indicating that, in most cases,

relative distances are preserved. This visual correlation among distances can be formalized using, for example, the Pearson's correlation coefficient (Edelmann et al., 2021) which, in this case, has a value $\rho = 0.934$, a high value that confirms that the relative distances are mostly preserved.

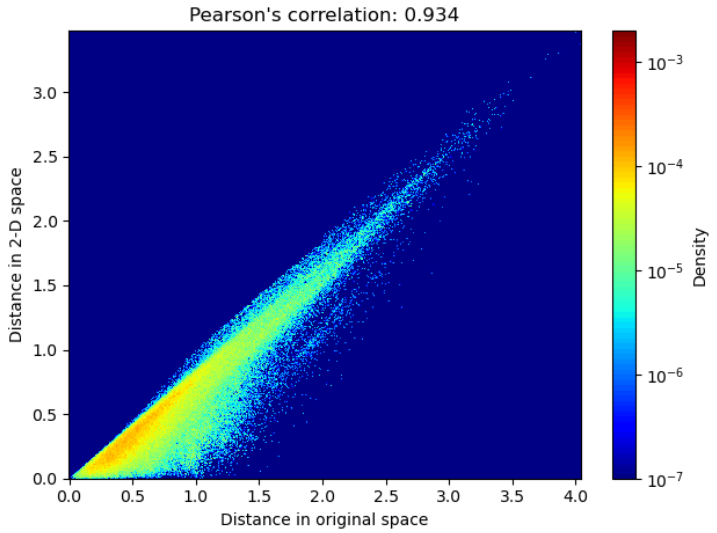

Fig. 25. Density of profiles pairs in the distances plane. A reduction in dimensionality to two components is performed using an autoencoder.

The previously defined correlation among distances can be extended to other techniques using a different number of components in the low-dimensional space. The results are shown in Fig. 26 where the correlation among distances increases with the number of components for PCA and AE techniques, which was an expected result.

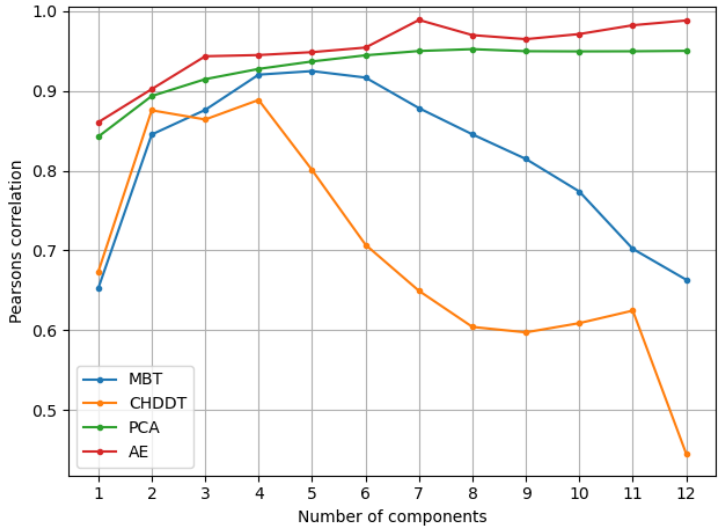

Fig. 26. Correlation between distances in the high and low-dimensional spaces.

However, the correlation for the MBT and CHDDT does not show a monotonous increasing behavior. This means that, for instance, in the case of CHDDT, although reducing profiles to eight components obtains smaller errors than the 1D or 2D

reductions (see Fig. 22), the distances are worse preserved. This, somehow disturbing result, could probably be explained because the MBT and CHDDT have an intuitive meaning for one- and two-dimensional reductions, while the reduction to a higher number of components have been obtained through a non-intuitive (artificial) extension.

6.3. Interpretation of maps and labels.

By reducing dimensionality to just two components, the resulting profiles can be drawn in a 2D plane (map of profiles). Also, reducing profiles to a one component permits assigning them a single figure (labelling profiles). Then, the meaning of these maps and labels must be elucidated. This interpretation will be focused on autoencoders as they have clearly outperformed the remaining techniques. For that reason, the shapes of the profiles corresponding to several points in the 2D maps and 1D labels obtained using autoencoders are depicted in Fig. 27.

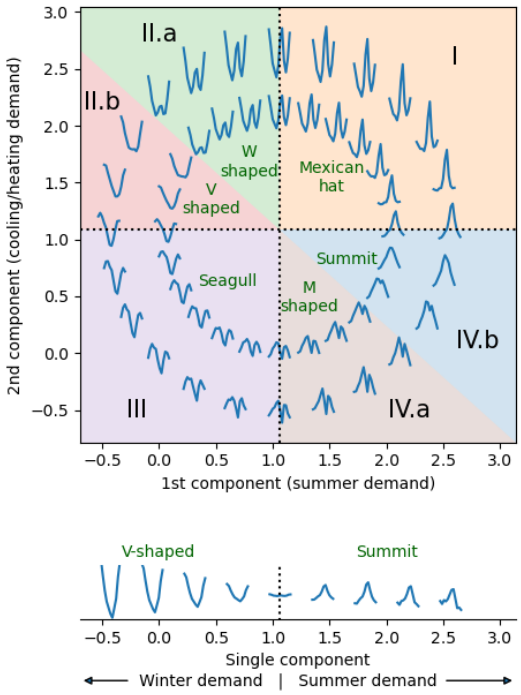

Fig. 27. Energy demand profiles corresponding to different positions in the 2D maps and 1D labels.

The maps of profiles show a noteworthy polar interpretation centered around the 2D representation of the mean profile. Indeed, the radius defines the scale of the profile, while the polar angle determines its shape. Up to six different angular sectors can be identified:

- Sector I (first quadrant), with a maximum demand in summer and a lower relative maximum demand in winter (similar to profile B in Table 5). The profile resembles the shape of a Mexican hat.

- Sector II.a (right part of the second quadrant), with a maximum demand in winter and a lower relative maximum demand in summer (similar to profile C in Table 5). The profile has the shape of a “W”.
- Sector II.b (left part of the second quadrant), with a single minimum demand in summer (similar to profile D in Table 5). The profile has the shape of a “V”.
- Sector III (third quadrant), with a minimum demand in summer and a lower relative minimum demand in winter (similar to profile E in Table 5). The profile resembles the shape of a seagull.
- Sector III.a (left part of the fourth quadrant), with a minimum demand in winter and a lower relative minimum demand in summer (similar to profile C in Table 5). The profile has the shape of an “M”.
- Sector III.b (right part of the fourth quadrant), with a single maximum demand in summer (similar to profile A in Table 5). The profile resembles the shape of a summit.

For a clearer interpretation of the map of profiles (2D), a video animation can be found in the supplementary material of this paper.

Besides, the single-dimensional representation of profiles shows an approximate symmetric behavior with respect to the 1D representation of the mean profile. The left part (values below the mean) corresponds to V-shaped profiles, while the right part (values above the mean) drives to summit-shaped profiles.

6.4. Assessing results.

In comparison with other state-of-art works referenced previously, the research described in this paper presents several novelties and strengths. In the first place, the large number of profiles that have been used (approximately half a million) corresponding to the actual demand during 3 years of a very high number of customers (more than 27 millions) must be mentioned. Most of the previous works are based either on synthetic profiles or in much more limited datasets.

Additionally, other authors have focused mostly on 24 hour load curves while the current research is centered in 12 months demand profiles, permitting a medium to long term insight in the demand characterization.

One of the most defining characteristics of our proposal is probably the fact that the profiles do not correspond to a single customer but to a cluster of them sharing the same location and/or economic activity. Then, although the available public datasets are anonymized, customers are still indirectly reachable by their NACE and location. In the increasingly competitive electricity retailer arena, it is easier to define marketing strategies with this grouped demand profiles than using load curves of a single but unknown customer.

In our research, two novel model-based dimensionality techniques have been proposed: MBT and CHDDT. While they have lower performance than state-of-art off-the-shelf reduction techniques (like PCA and autoencoders), their results are more easily explained, either in terms of the demand associated to certain months, or to certain temperatures. In the case of reducing profiles to 2D (map of profiles), CHDDT shows similar performance than state-of-art PCA (although lower than autoencoder).

Another novelty of the current research is the study of distances in the original and target spaces, usually omitted in previous works. Finally, a great effort, uncommon in the related literature, has been dedicated to interpreting the engineering meaning, in terms of electricity demand profile, of the 2D maps and 1D labels obtained using the proposed methodology.

As for the applications of the described methodology, it can be noted that reducing the dimensionality of electricity demand profiles can be directly used, in the first place, as a pre-processing step that improves the performance of forecasting, classification or clustering algorithms.

Additionally, as the proposed technique has paid special attention to the explainability of the low-dimensional profiles, 2D maps can be applied to visualize the electricity demand of a customer (or set of customers with the same NACE-location) as a simple point in a diagram, where its position reveals its profile. Several demand profiles can be simultaneously plotted in the 2D graphs using a dot for each profile. As the dimensionality reduction process preserves distances, the proximity of dots on the 2D maps indicates similarity in the demand profiles.

Also, 1D labelling can be applied to assign a continuously valued score to demand profiles. Recalling that classical clustering techniques group profiles in a non-ordered discrete number of categories, it is clear that 1D labelling can improve the interpretation of the tag assigned to a profile and, therefore, it can enlighten the characterization of customers. A practical example of the application of these labels can be found in (Luque et al., 2021).

The current research also has some limitations. Firstly, the analysis can be extended to other data-driven dimensionality reduction techniques. However, since autoencoders are among the most efficient algorithms in this kind of applications, no remarkable improvement should be expected. Also, there are limitations due to some lack of information in the dataset. Although the Spanish authorities have fostered competition by making the demand profiles public, they have declared the NACE field as optional, which means that retailers are not forced to share this information. So, only 55% of records contain those values, reducing the extent of the research.

7. Conclusions.

As for the dimensionality reduction techniques, the main conclusions reached in the research are:

- A small number of low-dimensional components yields a decrease in the reduction efficiency, that is, higher recovery errors.
- The relationship between number of components and recovery errors is approximately linear for all the techniques examined in the paper.
- These techniques have shown a high capacity to preserve profile information. Even for the most extreme reduction (keeping just one component) the methods employed exhibit an 80% (for MBT) to 86% (for autoencoder) decreasing in the recovery error compared to the random estimation and a 10% (for MBT) to 40% (for autoencoder) reduction over the naïve prediction.
- Data-driven techniques (PCA and autoencoder) clearly outperform model-based transformations, either the ones based on calendar (MBT) or those relying on temperature (CHDDT), with about a 15% lower recovery error.
- The explored techniques also have good behavior preserving the distances after the transformation. That is, profiles that are close in the high-dimensional space are also mostly close in the low-dimensional representation.
- Distance preservation has been measured using the correlation between distances before and after the transformation, and the figures obtained are about 85% correlation in 1D and 90% in 2D.
- Among the techniques explored, a properly trained autoencoder has clearly outperformed any other algorithm for any dimensionality of the target profile. In this technique, the 2D representation, which discards 83% of the components (10 out of 12) has a mean recovery error of 28%.
- Analogously, the 1D representation exhibits a mean error of 34%, while it discards 92% of the components.
- Finally, the 2D representation of a profile enables a polar interpretation according to its position in any of the four quadrants of the plane (with two additional sub-quadrants), where its angular position determines the shape of the high-dimensional profile, while the radius indicates its scale.

Acknowledgement

This work was supported by the “Ministerio de Ciencia, Innovación y Universidades”, Government of Spain under the project “Bigdata Analitycs e Instrumentación Cyberfísica para Soporte de Operaciones de Distribución en la Smart Grid”, number RTI2018-094917-B-I00. Additionally, the authors thank the TSO Company for granting access to their datasets.

References

ACER, 2021. Annual Report on the Results of Monitoring the Internal Electricity and Natural Gas Markets in 2020 - Energy Retail Markets and Consumer Protection Volume.

Aoi, M., Pillow, J.W., 2018. Model-based targeted dimensionality reduction for neuronal population data. Adv Neural Inf Process Syst 31.

Apadula, F., Bassini, A., Elli, A., Scapin, S., 2012. Relationships between meteorological variables and monthly electricity demand. Appl Energy 98, 346–356. https://doi.org/10.1016/j.apenergy.2012.03.053

Arechiga, A., Barocio, E., Ayon, J.J., Garcia-Baleon, H.A., 2017. Comparison of Dimensionality Reduction Techniques for Clustering and Visualization of Load Profiles, in: 2016 IEEE PES Transmission and Distribution Conference and Exposition-Latin America, PES T and D-LA 2016. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/TDC-LA.2016.7805661

Ayesha, S., Hanif, M.K., Talib, R., 2020. Overview and comparative study of dimensionality reduction techniques for high dimensional data. Information Fusion 59, 44–58. https://doi.org/10.1016/J.INFFUS.2020.01.005

BizEE Software, 2022. Heating & Cooling Degree Days – Free Worldwide Data Calculation [WWW Document]. URL https://www.degreedays.net/ (accessed 4.23.22).

Cai, J., Luo, J., Wang, S., Yang, S., 2018. Feature selection in machine learning: A new perspective. Neurocomputing 300, 70–79. https://doi.org/10.1016/J.NEUCOM.2017.11.077

Chhikara, P., Jain, N., Tekchandani, R., Kumar, N., 2022. Data dimensionality reduction techniques for Industry 4.0: Research results, challenges, and future research directions. Softw Pract Exp 52, 658–688. https://doi.org/10.1002/SPE.2876

D'Amico, A., Ciulla, G., Panno, D., Ferrari, S., 2019. Building energy demand assessment through heating degree days: The importance of a climatic dataset. Appl Energy 242, 1285–1306. https://doi.org/10.1016/J.APENERGY.2019.03.167

DiCiccio, T., Science, B.E., 1996. Bootstrap confidence intervals. Statistical science 11, 189–228.

Edelmann, D., Móri, T.F., Székely, G.J., 2021. On relationships between the Pearson and the distance correlation coefficients. Stat Probab Lett 169, 108960. https://doi.org/10.1016/J.SPL.2020.108960

Ezugwu, A.E., Ikotun, A.M., Oyelade, O.O., Abualigah, L., Agushaka, J.O., Eke, C.I., Akinyelu, A.A., 2022. A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. Eng Appl Artif Intell 110, 104743. https://doi.org/10.1016/J.ENGAPPAI.2022.104743

Fallahpour, A., Wong, K.Y., Rajoo, S., Tian, G., 2021. An evolutionary-based predictive soft computing model for the prediction of electricity consumption using multi expression programming. J Clean Prod 283, 125287. https://doi.org/10.1016/J.JCLEPRO.2020.125287

Fu, T.C., 2011. A review on time series data mining. Eng Appl Artif Intell 24, 164–181. https://doi.org/10.1016/J.ENGAPPAI.2010.09.007

Golan, J.S., 2012. Moore–Penrose Pseudoinverses, in: The Linear Algebra a Beginning Graduate Student Ought to Know. Springer, Dordrecht, pp. 441–452. https://doi.org/10.1007/978-94-007-2636-9_19

Gong, T., Li, D., Liu, Y., Wang, G., Zhu, H., 2019. Analysis of Marketing Strategy of Electricity Selling Companies in the New Situation. J Phys Conf Ser 1187, 022043.

Guerrero, J.I., Monedero, I., Biscarri, F., Biscarri, J., Millan, R., Leon, C., 2018. Non-Technical Losses Reduction by Improving the Inspections Accuracy in a Power Utility. IEEE Transactions on Power Systems 33, 1209–1218. https://doi.org/10.1109/TPWRS.2017.2721435

Hor, C.L., Watson, S.J., Majithia, S., 2005. Analyzing the impact of weather variables on monthly electricity demand. IEEE Transactions on Power Systems 20, 2078–2085. https://doi.org/10.1109/TPWRS.2005.857397

Huang, X., Wu, L., Ye, Y., 2019. A Review on Dimensionality Reduction Techniques. International Journal of Pattern Recognition and Artificial Intelligence 33, 1950017. https://doi.org/10.1142/S0218001419500174

Khan, B., Mahela, O.P., Padmanaban, S., Alhelou, H.H., 2022. Deregulated Electricity Structures and Smart Grids. CRC Press.

Kiesel, R., Paraschiv, F., 2017. Econometric analysis of 15-minute intraday electricity prices. Energy Econ 64, 77–90. https://doi.org/10.1016/J.ENECO.2017.03.002

Lee, S., Whaley, D., Saman, W., 2014. Electricity Demand Profile of Australian Low Energy Houses. Energy Procedia 62, 91–100. https://doi.org/10.1016/J.EGYPRO.2014.12.370

Li, K., Cursio, J.D., Jiang, M., Liang, X., 2019. The significance of calendar effects in the electricity market. Appl Energy 235, 487–494. https://doi.org/10.1016/J.APENERGY.2018.10.124

Liu, Y., Zhang, D., Gooi, H.B., 2020. Data-driven decision-making strategies for electricity retailers: Deep reinforcement learning approach. CSEE Journal of Power and Energy Systems. https://doi.org/10.17775/cseejpes.2019.02510

Luque, J., Anguita, D., Pérez, F., Denda, R., 2020. Spectral Analysis of Electricity Demand Using Hilbert–Huang Transform. Sensors 20, 2912. https://doi.org/10.3390/s20102912

Luque, J., Personal, E., Garcia-Delgado, A., Leon, C., 2021. Monthly Electricity Demand Patterns and Their Relationship with the Economic Sector and Geographic Location. IEEE Access 9, 86254–86267. https://doi.org/10.1109/ACCESS.2021.3089443

Mahmoudi-Kohan, N., Parsa Moghaddam, M., Sheikh-El-Eslami, M.K., 2010. An annual framework for clustering-based pricing for an electricity retailer. Electric Power Systems Research 80, 1042–1048. https://doi.org/10.1016/j.epsr.2010.01.010

Mishra, M., Nayak, J., Naik, B., Abraham, A., 2020. Deep learning in electrical utility industry: A comprehensive review of a decade of research. Eng Appl Artif Intell 96, 104000. https://doi.org/10.1016/J.ENGAPPAI.2020.104000

Morán, A., Fuertes, J.J., Prada, M.A., Alonso, S., Barrientos, P., Díaz, I., Domínguez, M., 2013. Analysis of electricity consumption profiles in public buildings with dimensionality reduction techniques. Eng Appl Artif Intell 26, 1872–1880. https://doi.org/10.1016/J.ENGAPPAI.2013.04.005

Obaid, H.S., Dheyab, S.A., Sabry, S.S., 2019. The impact of data pre-processing techniques and dimensionality reduction on the accuracy of machine learning, in: IEMECON 2019 - 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference. Institute of Electrical and Electronics Engineers Inc., pp. 279–283. https://doi.org/10.1109/IEMECONX.2019.8877011

Panapakidis, I., Alexiadis, M., Papagiannis, G., 2015. Evaluation of the performance of clustering algorithms for a high voltage industrial consumer. Eng Appl Artif Intell 38, 1–13. https://doi.org/10.1016/J.ENGAPPAI.2014.10.013

Pełka, P., Dudek. Grzegorz, 2020. Pattern-based Long Short-term Memory for Mid-term Electrical Load Forecasting, in: IEEE Computational Intelligence Society, International Neural Network Society, Institute of Electrical and Electronics Engineers (Eds.), 2020 International Joint Conference on Neural Networks (IJCNN).

Sánchez-Durán, R., Barbancho, J., Luque, J., 2019a. Solar Energy Production for a Decarbonization Scenario in Spain. Sustainability 11, 7112. https://doi.org/10.3390/su11247112

Sánchez-Durán, R., Luque, J., Barbancho, J., 2019b. Long-Term Demand Forecasting in a Scenario of Energy Transition. Energies (Basel) 12, 3095. https://doi.org/10.3390/en12163095

Scheidt, F. vom, Medinová, H., Ludwig, N., Richter, B., Staudt, P., Weinhardt, C., 2020. Data analytics in the electricity sector – A quantitative and qualitative literature review. Energy and AI 1, 100009. https://doi.org/10.1016/j.egyai.2020.100009

Sevilla, F.R.S., Korba, P., Barocio, E., Chavez, H., Sattinger, W., 2019. Data analytic tool for clustering identification based on dimensionality reduction of frequency measurements, in: 2019 International Conference on Smart Grid Synchronized Measurements and Analytics, SGSMA 2019. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/SGSMA.2019.8784626

Shaffer, B., 2019. Location matters: Daylight saving time and electricity demand. Canadian Journal of Economics/Revue canadienne d'économique 52, 1374–1400. https://doi.org/10.1111/CAJE.12407

Shi, Y., Yu, T., Liu, Q., Zhu, H., Li, F., Wu, Y., 2020. An Approach of Electrical Load Profile Analysis Based on Time Series Data Mining. IEEE Access 8, 209915–209925. https://doi.org/10.1109/ACCESS.2020.3019698

Tabachnick, B., Fidell, L., 2007. Experimental designs using ANOVA. Thomson/Brooks/Cole.

Tian, G., Chu, J., Liu, Y., Ke, H., Zhao, X., Xu, G., 2011. Expected energy analysis for industrial process planning problem with fuzzy time parameters. Comput Chem Eng 35, 2905–2912. https://doi.org/10.1016/J.COMPCHEMENG.2011.05.012

Tian, G., Yuan, G., Aleksandrov, A., Zhang, T., Li, Z., Fathollahi-Fard, A.M., Ivanov, M., 2022. Recycling of spent Lithium-ion Batteries: A comprehensive review for identification of main challenges and future research trends. Sustainable Energy Technologies and Assessments 53, 102447. https://doi.org/10.1016/J.SETA.2022.102447

Tjaden, T., Bergner, J., … J.W., 2015. Representative electrical load profiles of residential buildings in Germany with a temporal resolution of one second, ResearchGate. ResearchGate.

UNE, 1988. 100002: 1988. Air conditioning. Degrees day based at 15ºC.

Vidal, R., Ma, Y., Sastry, S.S., 2016. Principal Component Analysis. Interdisciplinary Applied Mathematics 40, 25–62. https://doi.org/10.1007/978-0-387-87811-9_2

Wang, Yi, Chen, Q., Kang, C., Xia, Q., 2016. Clustering of Electricity Consumption Behavior Dynamics Toward Big Data Applications. IEEE Trans Smart Grid 7, 2437–2447. https://doi.org/10.1109/TSG.2016.2548565

Wang, Yasi, Yao, H., Zhao, S., 2016. Auto-encoder based dimensionality reduction. Neurocomputing 184, 232–242. https://doi.org/10.1016/J.NEUCOM.2015.08.104

Williams, S., Short, M., 2020. Electricity demand forecasting for decentralised energy management. Energy and Built Environment 1, 178–186. https://doi.org/10.1016/J.ENBENV.2020.01.001

Xie, X., Yue, D., Peng, C., 2018. Relaxed real-time scheduling stabilization of discrete-time Takagi–Sugeno fuzzy systems via an alterable-weights-based ranking switching mechanism. IEEE Transactions on Fuzzy Systems 26, 3808–3819.

Yang, J., Zhao, J., Luo, F., Wen, F., Dong, Z.Y., 2018. Decision-Making for Electricity Retailers: A Brief Survey. IEEE Trans Smart Grid 9, 4140–4153. https://doi.org/10.1109/TSG.2017.2651499

Yang, J., Zhao, J., Wen, F., Dong, Z., 2019. A Model of Customizing Electricity Retail Prices Based on Load Profile Clustering Analysis. IEEE Trans Smart Grid 10, 3374–3386. https://doi.org/10.1109/TSG.2018.2825335

Yang, L., 2011. Distance-preserving dimensionality reduction. Wiley Interdiscip Rev Data Min Knowl Discov 1, 369–380. https://doi.org/10.1002/WIDM.39

Zhou, K., Yang, C., Shen, J., 2017. Discovering residential electricity consumption patterns through smart-meter data mining: A case study from China. Util Policy 44, 73–84. https://doi.org/10.1016/j.jup.2017.01.004

Zoubir, A.M., Iskandler, D.R., 2007. Bootstrap methods and applications. IEEE Signal Process Mag 24, 10–19. https://doi.org/10.1109/MSP.2007.4286560