forecasting

MDPI

Review

Advances in Similar Day Methods for Short-Term Load Forecasting for Power Systems

Monica Borunda [1,*], Luis Conde-López [2], Gerardo Ruiz-Chavarría [3], Guadalupe Lopez Lopez [4], Victor M. Alvarado [4] and Edgardo de Jesús Carrera Avendaño [4]

[1] Secretaría de Ciencia, Humanidades, Tecnología e Innovación, Tecnológico Nacional de México, Centro Nacional de Investigación y Desarrollo Tecnológico, Cuernavaca 62490, Mexico
[2] Centro Nacional de Control de Energía, Gerencia de Control Regional Oriental, Heroica Puebla de Zaragoza 72307, Mexico; luis.conde@cenace.gob.mx
[3] Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico; gruiz@unam.mx
[4] Centro Nacional de Investigación y Desarrollo Tecnológico, Tecnológico Nacional de México (TecNM/CENIDET), Cuernavaca 62490, Mexico; guadalupe.ll@cenidet.tecnm.mx (G.L.L.); victor.am@cenidet.tecnm.mx (V.M.A.); edgardo.cap@cenidet.tecnm.mx (E.d.J.C.A.)
* Correspondence: monica.bp@cenidet.tecnm.mx; Tel.: +52-55-4009-9106

Highlights

What are the main findings?

- A study of the literature on short-term load forecasting was conducted using the similar day (SD) methodology. The study focused on the last 25 years and covered conventional, intelligent, and hybrid methods.
- The methods operate in three stages: (a) selecting similar days; (b) choosing forecasting techniques; and (c) integrating both steps to produce the short-term forecasted load demand profile.

What are the implications of the main findings?

- The operation of power systems is changing rapidly compared with just a few years ago. Modern power system operation is increasingly complex due to uneven grid expansion, high renewable penetration, and new technologies that change demand behavior unpredictably. Additionally, unusual events occur more frequently, contributing to this unpredictability. Therefore, forecasting tools need to be revisited and improved continuously if they are to remain useful in real operating conditions.
- The literature studied for this work indicated that very few methods have been developed for forecasting similar days, highlighting the need for forecasting tools that remain accurate under abnormal or disruptive conditions. Attempts in this direction could include generating synthetic data and a history of similar days only in extreme event scenarios; training intelligent methods only in extreme event scenarios; and using emergent technologies, such as explainable artificial intelligence (XAI).

Abstract

Short-term load forecasting is essential for the reliable, secure, efficient, and economic operation of modern power systems and electricity markets. Among many forecasting strategies, the similar day (SD) approach for short-term load forecasting was among the earliest used to assess power demand and remains one of the most intuitive and widely adopted techniques worldwide. However, over time, increasing system complexity, richer datasets, and advances in computational intelligence have led to the evolution of SD methodologies beyond heuristic-based rule formulations. This work presents a study of

Academic Editors: Antonio Gabaldón, María Carmen Ruiz-Abellón and Luis Alfredo Fernández-Jiménez

Received: 12 February 2026
Revised: 2 April 2026
Accepted: 8 April 2026
Published: 10 April 2026

the relevant literature on short-term load forecasting using SD methods reported between 2000 and 2025. This study analyzes how similarity is defined, how forecasts are generated, and how both stages interact within the complete forecasting process in the reviewed literature. Based on these criteria, a unified taxonomy is proposed to classify SD methods into conventional, intelligent, and hybrid formulations. This study provides insight into the methodologies, their performance, and the systems in which they have been tested. The results show that SD-based approaches remain competitive for short-term forecasting and that incorporating artificial intelligence techniques can further enhance their accuracy.

Keywords: artificial intelligence; energy demand; day-ahead load forecasting; load forecasting; machine learning; power systems; similar day; short-term load forecasting; time series

1. Introduction

Currently, decarbonization is an urgent priority that must be pursued from every possible front. To achieve carbon-neutral energy systems, cleaner energy sources must be utilized, along with improved energy efficiency and management in energy systems [1,2]. Energy systems consist of electrical loads ranging from a few watts to hundreds of gigawatts that must be supplied by electrical generation. A set of interconnected electrical components that generate, transmit, and distribute electrical energy from power plants to satisfy load demand is a power system. Thus, electrical load forecasting is essential for the success of power systems, from their design and installation to their operation and maintenance [3–5].

Electrical load forecasting is needed at multiple timescales [6,7]. The load forecasting horizons can be categorized as follows [8]:

- Very short-term forecasts, ranging from minutes to one hour ahead, are essential for managing contingencies using supplementary reserves or interruptible loads.
- Short-term forecasts, ranging from hours to a few days ahead, are required for unit commitment, the electric market, and reserve allocation [9]. In particular, in this time horizon lies the day-ahead load forecasting, which is crucial for economic dispatch, the electrical market, and the balance between all types of electric generation and storage.
- Medium-term forecasts, ranging from a few days to weeks, are useful for maintenance and fuel requirements.
- Long-term forecasts, ranging from months to years, are needed for investments and expansion studies.

Given that renewable penetration, particularly the intermittent forms, and demand variability continue to grow, accurate short-term load forecasting becomes more and more important every day for system operators [10]. Even early on, methods that extracted information from historical patterns were practical tools for anticipating next-day consumption, such as the SD method.

The SD method is one of the oldest, most intuitive, and most used methods currently in electrical utilities. In its simplest form, the SD method assumes that energy patterns repeat under comparable weather and calendar conditions, making historical analogs a valuable basis for forecasting. Although no publication documents the introduction of the SD method, it is believed that utilities have applied it since the 1970s, since it was documented in the 1980s [11]; it is described as an analogous day-selection procedure in utility practices of the North American Electric Reliability Council (NERC) [12]. Indeed, early implementations consisted of identifying similar days based on calendar attributes or basic weather variables, and their load curves were used, directly or combined with simple statistical calculations, to provide the day-ahead load curve. Since the methods were intuitive, easy to implement, and computationally low-cost, they were adopted by many utility company operators. Later works formalized this idea by interpreting SD forecasting as a pattern similarity problem in the space of daily load profiles, providing a theoretical foundation for modern SD-based methods [13,14].

In recent years, the development of larger and more complex power systems, coupled with the availability of vast amounts of high-resolution data and a growing demand for improved forecasting quality, has posed significant challenges. To address this issue, which is becoming a strategic focus in energy transition research, efforts are underway to develop forecasting methodologies that achieve the following:

(a) Capture the know-how of forecasting operators with data-driven information;

(b) Integrate more variables;

(c) Capture nonlinear behaviors;

(d) Account for structural changes in demand.

These methodologies include data-driven approaches, complex forecasting systems, and the use of sophisticated tools, such as AI, either alone or in combination [15]. In particular, load forecasting approaches can be broadly categorized into similarity-based methods, data-driven deep learning models, and hybrid approaches, as shown in Figure 1. While modern deep learning models, such as recurrent neural networks and Transformer-based architectures learn temporal dependencies directly from historical data, similar day methods rely on explicit similarity metrics to identify historical analog days.

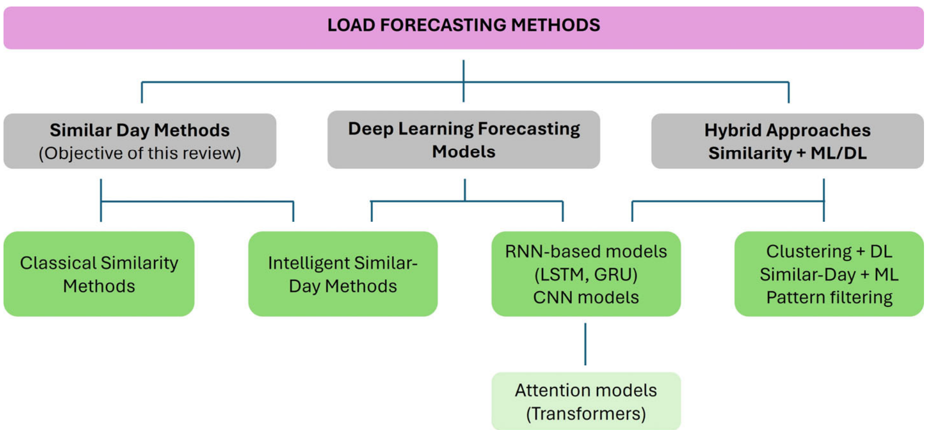

Figure 1. Conceptual taxonomy of load forecasting approaches, highlighting the position of similar day methods within the broader forecasting landscape.

In recent years, deep learning architectures, such as recurrent neural networks and Transformer-based models, have shown promising results in load forecasting due to their ability to learn complex temporal dependencies. However, unlike similar day methods, most of these approaches typically rely on data-driven representation learning rather than explicit similarity-based selection of historical analog days [16–18].

This review focuses on similar day forecasting approaches.

Despite this evolution, the literature lacks a consolidated technical synthesis that describes how SD methods have developed, which components drive their performance, and how recent machine learning (ML)-enhanced or hybrid approaches relate to the original formulation. Most existing studies focus on proposing new variations rather than situating them within a coherent methodological structure.

This review addresses that gap by offering a technical perspective on SD-based forecasting. Specifically, this review achieves the following:

1. Summarizes some recent advances in SD-based forecasting methods;

2. Presents a structured taxonomy organizing the literature into a clear classification based on similarity definition, forecasting model, and integration strategy;
3. Provides a comparison of approaches analyzing trends, performance ranges, and key technical insights reported over the past two decades.

It is important to clarify that, in this work, conventional methods refer to methods that do not use AI techniques, including heuristic, statistical, and other non-intelligent methods.

The goal of this work is to provide a structured view of SD-based methods for short-term load forecasting and to highlight the developments that have shaped their performance in contemporary power systems. The remainder of this work is structured as follows. Section 2 presents the conceptual background of SD methods for short-term load forecasting and introduces the analytical perspective and the taxonomy employed to classify SD forecasting methods according to their structural characteristics. Section 3 presents the results of the review, including a comparative analysis and identification of key technical findings. Section 4 provides a critical discussion of the reviewed approaches, highlighting cross-cutting comparisons, analytical insights, and practical implications for power system operators. Finally, Section 5 concludes the paper and outlines directions for future research.

2. Materials and Methods

This section presents the conceptual background of SD methods for short-term load forecasting together with the definition of the variables and the similarity metrics used in conventional SD selection and describing the approaches followed in intelligent SD selection. Additionally, the analytical perspective adopted in this work and the taxonomy that groups SD forecasting methods according to their structural characteristics are outlined at the end of this section.

2.1. Similar Day Method Background

The SD approach has been employed in short-term load forecasting for several decades and continues to be one of the most practical ways to incorporate historical information into prediction tasks. The SD method relies on the core assumption that if certain days in the past share similar conditions with the target day, then their load profiles may be very similar to that of the future target day. Early versions of the SD approach relied mainly on simple rules, such as matching calendar characteristics or weather conditions [19], and were widely adopted because they had low computational requirements and were easy for system operators to interpret.

As more data became available, forecasting demands became more complex, and the SD formulation expanded. Besides load data, additional variables, such as day type, meteorological conditions, and various pattern indicators, were incorporated to better define day-to-day similarity. Thus, the SD approaches expanded from a single, heuristic procedure to a diverse set of techniques for selecting similar days and integrating them into forecasting models. As part of the evolution of SD formulations, several studies focused on identifying and classifying daily load patterns using data-driven techniques. For instance, Zhang et al. (2012) proposed a clustering-based approach to extract typical load profiles from high-resolution daily load curves of large electricity customers [20]. Although this work does not perform load forecasting or SD selection, it illustrates how pattern recognition and clustering techniques can capture representative demand behaviors. Such approaches later influenced intelligent SD methodologies, where clustering is used intelligently as a preliminary step to identify candidate similar days or to enhance similarity definitions beyond simple calendar or weather rules.

It is worth noting that some studies do not adopt calendar-based SD rules in a strict sense but instead identify analogous historical instances through pattern-based similarity or data-driven partitioning of historical observations. Although these approaches differ from conventional SD formulations, they share the same underlying principle of retrieving historical analogs to support short-term load forecasting. For this reason, pattern-based similarity methods [13,14] are considered within the broader scope of SD-inspired approaches, particularly when intelligent or data-driven techniques, such as clustering, are used to define similarity beyond explicit calendar or weather features matching the target day, thus laying the foundation for intelligent and pattern-based SD methodologies [21].

Despite this diversity, as previously mentioned, most SD approaches share three elements that form their backbone. The first is the definition of similarity, which determines how candidate days are identified in the historical record as similar days. The second is the forecasting model that uses the selected days, ranging from classical statistical tools to more advanced machine learning algorithms. The third is the integration of the selection and modeling, since the SD step can shape both the structure and inputs of the forecasting model in different ways.

The evolution of SD approaches has evolved from conventional to intelligent methods, and, more recently, to hybrid methods, where similarity conditions coexist with clustering, optimization, or machine learning techniques. Therefore, a trend in load forecasting is emerging, combining domain knowledge with data-driven methods to develop more accurate and robust methodologies. Therefore, organizing SD approaches into a consistent framework is useful for comparing them effectively and understanding their relationships. The following subsections build this foundation, starting with the variables and distance metrics that define similarity in SD-based forecasting.

2.2. Variables and Similarity Metrics

The SD approach is based on the premise that daily electricity demand patterns repeat under comparable conditions. So, this approach relies on the definition of similarity between the target day and historical days. Thus, the first step of this process is the selection of the conditions or characteristics, also known as similarity-defining features, which are used to quantify similarity between days. However, similarity-defining features are time-varying quantities underlying time series, such as historical load profiles, meteorological observations, or calendar-related variables. For this reason, similarity-defining features can be regarded as explanatory variables, since they evolve over time and describe the factors that directly influence load dynamics, playing the role of explanatory variables in load forecasting models. The explanatory variables capture temporal, climatic, and/or behavioral patterns relevant to load demand. In many SD formulations, distances are computed directly over historical load profiles, which constitute a particular case of similarity-defining features. In addition to explicit distance-based similarity metrics, modern SD approaches increasingly rely on learning-based similarity measures, where similarity is implicitly learned through clustering, representation learning, or machine learning models rather than predefined distance functions.

Hence, in SD forecasting, the definition of similarity depends on two elements:

(a) The set of variables selected to characterize the demand conditions;
(b) The metric used to compare these variables and quantify similarity between the target day and historical days.

2.2.1. Variable Categories

SD methods for load forecasting typically use three main categories of explanatory variables:

1. Load variables, which incorporate short-term load history, such as the previous day's demand or recent load profiles representing temporal continuity and system inertia in consumption behavior;
2. Weather-related variables, which correspond to meteorological conditions that may influence electricity consumption, such as temperature, humidity, solar radiation, wind speed, and other meteorological parameters;
3. Temporal variables, which reflect calendar structure, such as the day of the week, hour, season, holidays, working or non-working period, or special events, which capture recurrent behavioral and operational patterns.

However, more categories could be considered, such as the following:

4. Economic variables, which measure regional economic activity or performance, such as regional gross domestic product, per capita income, electricity rates, inflation and exchange rates, energy regulation, and economic crises;
5. Political variables, which represent the political activity of the region, such as public policies, regulations, subsidies, tariff structures, international agreements, climate commitments, presidential or congressional elections, political instability, and governance;
6. Social variables, which characterize the region's society, such as holidays, social events and festivities, population density, level of urbanization, and social infrastructure.

2.2.2. Similarity Metrics

Once the variables are selected, the next step is to determine how to quantify the similarity of the chosen variables between the target day and historical days. Several distance metrics have been used for this purpose, ranging from simple to more elaborate formulations [22–24]. Table 1 shows the main metrics used in conventional SD approaches.

Table 1. Similarity metrics used to quantify the similarity of the chosen variables between the forecast day and historical days, where $v(t)$ is the explanatory variable, with $\overline{v}$ the corresponding mean value, chosen to find the similar days, at hour t, $i \epsilon$ H runs from 1 to n and corresponds to a historical day in the group of historical data H, and f corresponds to the forecast day, and i^* is the similar day value that, when maximized or minimized, determines the similar day. Additionally, ΔL_i is the deviation of the load of the forecast day and the historical day i, $\Delta L_{s,i}$ is the deviation between their load slopes, Δv_i is the deviation of the variable v chosen to find the similarity between the forecast and the historical day i, and w_1, w_2, and w_3 are the corresponding weight factors.

Name of the Metric	Analytical Expression	Condition for SD Selection
Euclidean distance	$d_E(f,i) = \sqrt{\sum_{t=1}^{24} \left(v(t)^f - v(t)^i \right)^2}$	$i^* = \underset{i\,\epsilon\,H}{\text{argmin}}\, d_E(f,i)$
Weighted Euclidean distance	$d_w(f,i) = \sqrt{w_1(\Delta L_i)^2 + w_2(\Delta L_{s,i})^2 + w_3(\Delta v_t)^2}$	$i^* = \underset{i\,\epsilon\,H}{\text{argmin}}\, d_w(f,i)$
Manhattan distance	$d_M(f,i) = \sum_{t=1}^{24} \left| v(t)^f - v(t)^i \right|$	$i^* = \underset{i\,\epsilon\,H}{\text{argmin}}\, d_M(f,i)$
Cosine similarity	$cos(f,i) = \frac{\sum_{t=1}^{24} \left(v(t)^f v(t)^i \right)}{\sqrt{\sum_{t=1}^{24} \left(v(t)^f \right)^2} \sqrt{\sum_{t=1}^{24} \left(v(t)^i \right)^2}}$	$i^* = \underset{i\,\epsilon\,H}{\text{argmax}}\, cos(f,i)$
Pearson correlation coefficient	$r(f,i) = \frac{\sum_{t=1}^{24} \left(v(t)^f - \overline{v}^f \right) \left(v(t)^i - \overline{v}^i \right)}{\sqrt{\sum_{t=1}^{24} \left(v(t)^f - \overline{v}^f \right)^2 \sum_{t=1}^{24} \left(v(t)^i - \overline{v}^i \right)^2}}$	$i^* = \underset{i\,\epsilon\,H}{\text{argmax}}\, r(f,i)$

As shown below, the Euclidean and weighted distances are the most used, followed by the Manhattan distance. However, there are also a few works that use the Cosine metric and the Pearson coefficient. Nevertheless, there are probabilistic or fuzzy similarity indices, which allow partial membership when weather uncertainty is significant [25]. More recent work explores distances learned through machine learning models, with clustering being the most common one, which capture structures not directly visible in raw variables alone or in big datasets.

Since SD approaches involve the choice of variables and the distance metric, their sensitivity to weather anomalies, structural breaks, or changes in user behavior is different. Some approaches explicitly tune the metric or adjust the variable set to reduce these effects. Others embed the similarity computation within optimization or hybrid schemes, using learned weights to emphasize the variables most relevant to demand variability.

Thus, the combination of variables and distance metrics forms the foundation for defining and evaluating the concept of similarity and strongly influences SD performance. Indeed, this foundation is essential for interpreting how conventional, AI-based, or hybrid SD methods behave and sets the stage for the analytical and taxonomic perspectives introduced in the next section.

It is worth noting that not all similarity-based load forecasting approaches operate at the level of individual days. Some recent studies explore similarity at an aggregated or regional scale, where entire load profiles (daily, weekly, or yearly) are compared across different systems rather than selecting analogous historical days for a specific target date. For instance, clustering-based regional similarity combined with transfer learning has been proposed to exploit structural similarities among power systems, enabling knowledge transfer across regions with limited historical data [26]. Although these approaches do not follow the traditional or pattern-based similar day paradigm, they reflect a broader evolution of similarity concepts in load forecasting and highlight alternative ways of leveraging historical information beyond day-level analogies.

2.3. Analytical Approach

In this work, the examination of SD methods follows a technical perspective, focusing on how these approaches are built and how they have changed over time. This analysis focuses on understanding the structure behind SD techniques rather than compiling an exhaustive list of publications.

While reviewing the literature, three aspects appeared consistently across SD approaches, even when described in different forms:

1. The choice of variables and similarity criteria, which determines how historical days are selected to form the set of similar days;
2. The forecasting model that uses information from the selected days;
3. The integration of the previous parts to complete the forecasting process.

By focusing on these shared structural elements, different SD variants can be compared under a unified framework that highlights the progressive integration of machine learning, optimization techniques, and hybrid architectures, and more elaborate definitions of similarity. Understanding this evolution is essential for formulating a coherent taxonomy, which is presented in the following subsection to organize SD methods according to their operational structure and modeling strategy.

Finally, forecasting performance is evaluated using the Mean Absolute Percentage Error (MAPE), given by

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - \hat{x}_i}{x_i}\right| \cdot 100\%, \tag{1}$$

where x_i and $\hat{x}_i$ are the real and predicted N values of the variable, respectively. Even though there are other errors such as Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), the MAPE is one of the most common metrics to measure the accuracy of a

https://doi.org/10.3390/forecast8020032

forecasting model due to its interpretability and suitability for comparing systems with different load magnitudes [27].

2.4. Proposed Taxonomy Diagram

The taxonomy diagram, shown in Figure 2, presents a structured classification of SD-based short-term load forecasting approaches by decomposing them into their fundamental building blocks. Rather than cataloging individual studies, the taxonomy focuses on the structural logic that defines how SD methods operate, evolve, and relate to one another. It organizes SD approaches into three conceptual layers, each representing a key decision in the forecasting process.

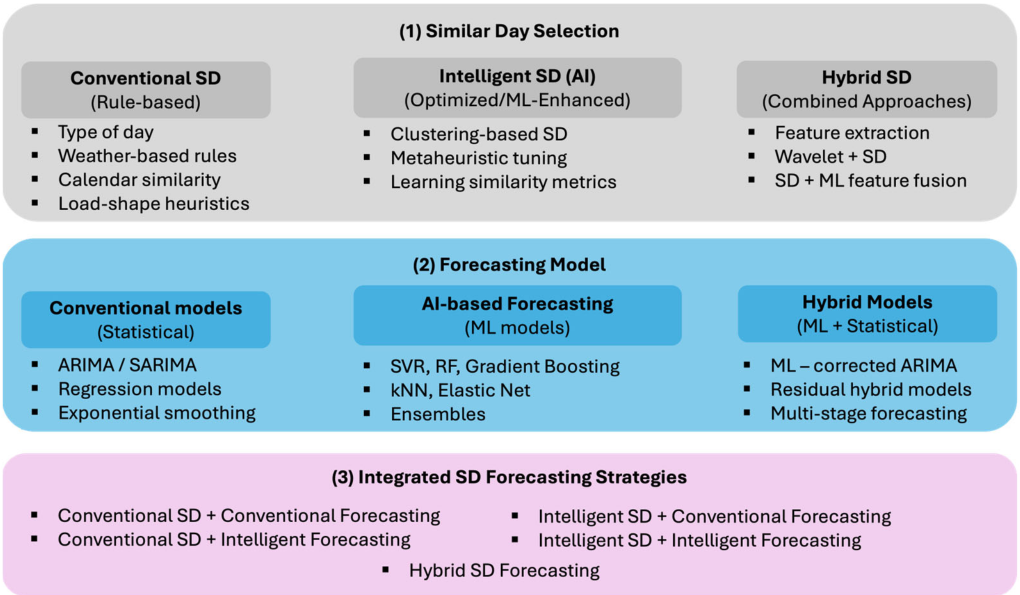

Figure 2. Taxonomy of the SD-based methods for short-term load forecasting.

The following section presents the reviewed literature organized according to this taxonomy together with an analysis of the methodologies highlighting the most important findings.

- First Layer: Similarity Selection Mechanism

The first layer classifies SD methods according to how similar days are identified from historical data. This layer reflects the knowledge paradigm used to define similarity.

- Conventional SD methods rely on explicit, rule-based criteria derived from domain expertise, such as calendar attributes (day type, season, holidays) and basic weather conditions. These approaches do not use AI techniques and have been widely adopted in utilities due to their simplicity, transparency, and low computational cost.
- Intelligent SD methods employ data-driven techniques—such as clustering, optimization algorithms, or machine learning—to identify similar days. These methods are particularly effective when large datasets are available, as they can uncover latent structures and complex similarity patterns not captured by simple rules.
- Hybrid SD methods combine conventional similarity rules with intelligent or signal-processing techniques. By merging explicit domain knowledge with data-driven similarity learning, hybrid approaches aim to capture more complex and nonlinear demand behaviors.

This layer highlights the evolution of SD similarity definition from heuristic rules to intelligent and hybrid formulations.

- Second Layer: Forecasting Model

The second layer categorizes the forecasting models applied after the similar days have been selected. It reflects how information extracted from historical analogs is transformed into load predictions.

- Conventional forecasting models include classical statistical techniques such as linear regression, exponential smoothing, and ARIMA. These models are interpretable and computationally efficient but limited in capturing nonlinear relationships.
- AI-based forecasting models use machine learning and data-driven techniques, such as Support Vector Regression, Random Forests, gradient boosting, neural networks, and ensemble learners. These models are capable of modeling nonlinearities and complex interactions among variables.
- Hybrid forecasting models integrate multiple forecasting techniques, often through multi-stage architectures or residual correction schemes, to improve accuracy and robustness.

This layer emphasizes that forecasting accuracy depends not only on similarity selection but also on the modeling strategy applied to the selected data.

- Third Layer: Integrated Forecasting Strategy

The third layer represents the interaction between similarity selection and forecasting models, defining the overall SD-based forecasting architecture. Different combinations are possible:

- Conventional SD combined with conventional or AI-based forecasting models.
- Intelligent SD combined with conventional or AI-based forecasting models.
- Fully hybrid systems, where both similarity selection and forecasting models are hybridized.

The choice of integration strategy depends on data availability, system complexity, and operational constraints. This layer underscores that SD forecasting is not a single algorithm but a modular framework, where similarity identification and prediction modeling jointly determine performance.

3. Results

3.1. Structured Organization of the Literature

The reviewed literature spans more than two decades (approximately 2001–2025), revealing a clear methodological evolution in SD-based short-term load forecasting. Early studies predominantly relied on rule-based similarity selection combined with conventional statistical forecasting models, while more recent works increasingly incorporate artificial intelligence either in the similarity selection stage, the forecasting model, or both. This evolution culminates in hybrid SD frameworks that integrate multi-stage or fused methodologies. Figure 3 summarizes the percentage of the literature studied according to the proposed taxonomy of a total of 83 reviewed works. The results are as follows:

- Conventional SD approaches correspond to 39% (outer ring): 11% utilize conventional forecasting methods, while 28% employ intelligent forecasting methods (inner ring).
- Intelligent SD approaches correspond to 14% (outer ring): 6% use conventional forecasting methods and 8% apply intelligent forecasting methods (inner ring).
- Hybrid approaches, either on SD or forecasting methodologies, correspond to the remaining 47% (outer and inner ring).

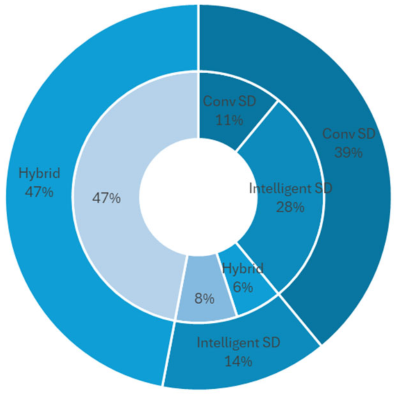

Figure 3. Distribution of the literature with methodologies according to the classifications in the proposed taxonomy.

Most studies utilize data from various countries worldwide. China leads in job numbers, followed by the USA, Australia, Japan, and India. Figure 4 illustrates the participation rates for the various countries. Despite regional differences, the explanatory variables used to determine similarity remain fairly consistent across the studies. Load, weather, and calendar variables dominate the selection process, while economic or social indicators appear in a small number of works.

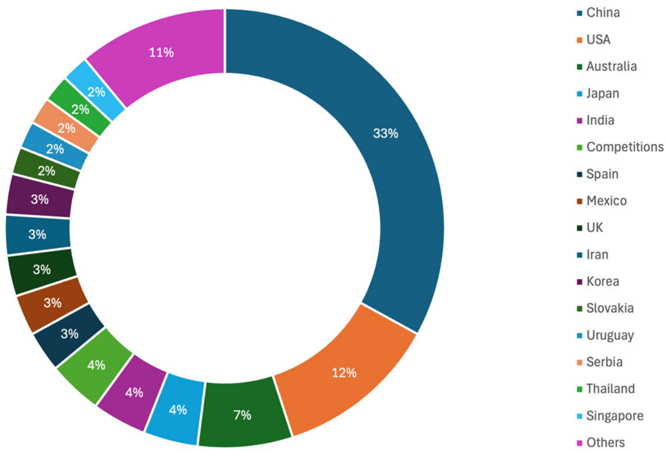

Figure 4. Percentage of countries where the different SD forecasting approaches from the revised literature have been applied. The level "competitions" refer to organized networks to compete on electricity load forecasting.

3.2. Literature Search and Selection Criteria

The literature search was conducted using major scientific databases, including IEEE Xplore Digital Library, Scopus, Web of Science, ScienceDirect, and SpringerLink. These databases were selected to ensure comprehensive coverage of the peer-reviewed literature in engineering and energy systems.

The literature search was conducted using a set of predefined keywords associated with short-term load forecasting and similar day approaches. The primary search terms included "short-term load forecasting", "day-ahead load forecasting", "similar days", and "power systems". These terms were combined using logical operators (AND, OR) and applied across multiple scientific databases to ensure a comprehensive and systematic identification of relevant studies. Additional variations and related terms were also considered to capture different formulations used in the literature.

In addition to the inclusion criteria, exclusion criteria were applied to ensure the relevance of the selected studies. Studies focusing on load forecasting in domains other than power systems—such as buildings, electric vehicles, residential households, or micro-level energy consumption—were excluded. The review specifically targets forecasting methods applied to power system-level demand, ensuring consistency in the scope and applicability of the analyzed studies.

3.3. Taxonomy of Similar Day Forecasting Approaches

In this subsection, we summarize some of the most recent works published. The works are organized in tables according to the taxonomy described above, which is also illustrated in the last layer of Figure 2. The main information extracted from these studies includes (a) the SD selection method, (b) the forecasting technique applied, (c) the input variables used, (d) the reported forecasting error, and (e) the country where the power system data was collected. As mentioned at the end of Section 2.3, the forecasted error is provided by the MAPE. In general, we include the minimum and maximum values reported in the literature, but in the cases where this information is missing, we specify the reported forecasting error. Finally, unless otherwise stated, the target variable in all studies is the electricity load demand (hourly, daily, or peak load, depending on the references) of power systems.

It is important to highlight that although multiple performance metrics are used in the literature, including the RMSE and the MAE, the MAPE remains the most widely adopted metric in short-term load forecasting due to its interpretability and scale-independence. For this reason, the comparative analysis in this review is primarily based on MAPE values. However, it is noted that only 4 out of 85 revised references do not report the MAPE, as explicitly stated in the tables.

3.3.1. Conventional Similar Day Methods and Conventional Forecasting Approaches

The summary of some main works employing conventional approaches for similar day classification and load forecasting is presented in Table 2.

Therefore, the short-term load forecasting methodologies under the category of conventional SD with conventional forecasting methods mainly use day-type rules and Euclidean distance metrics on load and temperature data for similar day selection. Moreover, as forecasting methodologies they employ regression, ARIMA or exponential smoothing, among others.

3.3.2. Conventional Similar Day Methods and Intelligent Forecasting Approaches

The summary of major works using conventional methods for similar day classification and intelligent techniques for load forecasting is provided in Table 3.

Thus, the short-term load forecasting methodologies under the category of conventional SD with intelligent forecasting methods mainly use day-type/calendar rules or distance metrics on load and temperature data for similar day selection. However, as forecasting methodologies, they employ machine learning techniques such as ANN, SVM and Random Forest, among others, or shallow deep learning models.

3.3.3. Intelligent Similar Day Methods and Conventional Forecasting Approaches

The summary of key works utilizing intelligent methods for similar day classification and conventional techniques for load forecasting is presented in Table 4.

https://doi.org/10.3390/forecast8020032

Table 2. Summary of key information on forecasting methods using conventional similar day classification and conventional load forecasting techniques.

Author, Year, and Reference	Conventional Similar Day Selection Method	Conventional Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Rahman et al., 1988 [28]	Similar days are calculated using heuristic rules based on season, day type, time of day, and dry-bulb temperature similarity	A base load forecast with the weighted load from similar days is calculated, and then corrections from a Base Factor and a Dynamic Temperature Factor associated with temperature are applied	Temperature, day type, season, time of day, and load data from 1983	Weekdays, weekends	Daily: (2.43, 3.30); Hourly: (0.43, 5.43)	Virginia Electricity Utility, USA
Ruzic et al., 2003 [29]	Euclidean distance considering the daily energy, temperature, day type, and time distance factor	First, the total daily energy is forecasted, and then the daily load shape definition is determined based on the type of day	Temperature, day type, and load data from 1988 to 1998	Weekdays, Weekends, holidays	Monthly: (1.21, 2.58); Daily: (1.48, 4.18)	Power Utility of Serbia, Serbia
Mu et al., 2010 [30]	Constructs an index-mapping database considering day type, week, and temperature, and then uses Cosine similarity to choose the similar days	The forecasted load is calculated from the weighted average of the similar load profiles using enhanced similarity weights	Temperature, day type, week type, and load data from 2008	Weekdays, weekends, holidays	Daily: (0.24, 1.67)	Hainan Power Grid, China
Jin, 2011 [31]	Similar days are constructed by minimizing the Euclidean distance of weather vectors	The total load is decoupled into the secular trend load, which is forecasted using a nonlinear time polynomial regression and meteorological load, which is forecasted using linear regression with multiple weather variables; the final load forecast is obtained, recombining both components	Temperature, wind speed, humidity, rainfall, and load data from 2010	Generic day	Daily: (0.01, 7.66)	Regional Chinese power company, China

https://doi.org/10.3390/forecast8020032

Table 2. *Cont.*

Author, Year, and Reference	Conventional Similar Day Selection Method	Conventional Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Dudek, 2015 [13]	Similarities between daily patterns of the seasonal cycles of load time series using Euclidean and Manhattan distances	Average of similar daily load patterns	Load data from 2002 to 2004	Weekdays	Daily: (1.90–9.44)	Polish Power System, Poland
Jian-Kai et al., 2015 [32]	Similar days are selected using the box-counting fractal dimension of the daily load curves, using the fractal measure as the similarity	The forecasted load is determined by a Fractal Interpolation Function, which is constructed through an Iterated Function System	Daily load curves	Weekdays, weekends, holidays	Daily relative error: Maximum—3.7; Average—1.6	Eastern Slovakian Electricity Corporation, Slovakia
Karimi et al., 2018 [33]	A priority index of similar days involving the weighted Euclidean norm of the temperature of the cities in the region and date proximity	The system is partitioned into smaller regions whose forecasted load is obtained according to the priority index and the forecasted load; the forecasted load is the sum of the former ones	Load data from 2011 to 2014	Weekdays, weekends	Daily: (0.91, 1.27) Average: 1.01	Iran’s National Power Network, Iran
Son et al., 2022 [34]	Similar days are selected according to the holiday pattern and the day of the week	The Behind-the-Meter Photovoltaic effect of the similar day groups is removed; then, genetic algorithms are used to optimize a weather threshold and a sensitivity matrix for each similar load profile; and finally, the forecasted load profile is obtained by the weighted aggregation of the modified profiles	Holiday type, day of the week, temperature, Behind-The-Meter Photovoltaic generation, and the load data from 2018 to 2019	Holidays	Daily: (0.79, 4.73) Average in a year: 2018, 1.76 2019, 2.62	Korean power system operator, Republic of Korea

Table 2. *Cont.*

Author, Year, and Reference	Conventional Similar Day Selection Method	Conventional Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Conde-López et al., 2025 [35]	Similar days based on calendar day and similar temperature conditions	The forecasted maximum load is obtained by pattern matching and heuristic correction from four recent similar days	Temperature, day type, and load data from 2025	Weekdays, weekends, holidays, anomalous days	Monthly: (0.50, 2.14); Blackout day: 44.00; Cold front: 0.60; Flood, weekly: 2.19	National Center for Energy Control, Mexico

Table 3. Summary of key information on forecasting methods using the conventional similar day classification and intelligent load forecasting techniques.

Author, Year, and Reference	Conventional Similar Day Selection Method	Intelligent Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Senjyu, 2002 [36]	Euclidean distance with weighted factors considering the load and temperature data	The forecasted load is obtained as the output of an ANN trained with similar days	Temperature and load data, from 1995 to 1997, and a load correction factor	Weekdays, weekends, holidays	Weekly: (0.9, 1.63); Maximum daily: 14	Okinawa Electric Power Company, Japan
Mandal et al., 2004 [37]	Euclidean distance with weighted factors considering temperature, day type, and load data	The forecasted load is obtained as the corrected output of mean similar day data obtained from an artificial neural network	Load and temperature data, from 1999 to 2000, and the type of day	Weekdays, weekends	Monthly: (0.78, 3.01); Seasonal: (0.80, 3.15)	Okinawa Electric Power Company, Japan
He et al., 2005 [38]	(a) The similarity degree is provided by the sum of the trend similarity degree of the load data; (b) the day similarity degree is measured by the Cosine similarity metric of the correlated factors	The forecasted load is the output of a neural network with a back-propagation momentum training algorithm	Temperature, humidity, wind speed, day type, and load data from 1999	Weekdays, weekends	Daily (one sample): 2.64	Hebei province, China

Table 3. *Cont.*

Author, Year, and Reference	Conventional Similar Day Selection Method	Intelligent Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Mandal et al., 2006 [39]	Weighted Euclidean distance considering temperature and load data	The forecasted load is obtained as the output of an ANN with averaged load data from similar load curves	Temperature and load data from 1999 to 2000	Weekdays, weekends, holidays	Hourly: (0.72, 0.80) Monthly: (3.95, 6.16)	Okinawa Electric Power Company, Japan
Mandal et al., 2006 [40]	Euclidean distance with weighted factors considering load, temperature data, and the type of day	The forecasted load is obtained as the output of an ANN with averaged load data from similar days, type of day, and temperature data	Type of day, load, and temperature data from 2000 to 2003	Weekdays, weekends	Weekly: (0.56, 1.30); Monthly: (0.77, 2.06)	NEMMCO, Australia
Chang-Chun et.al., 2008 [41]	The selection of the similar day is based on the load periodic variation rule, the day type difference, and the Pearson correlation coefficient for the temperature, visibility, and rainfall	The forecasted load is calculated using a Support Vector Machine	Load data, type of day, temperature, visibility, and rainfall from 2001	Weekdays, weekends	Daily: (1.73, 5.48)	Henan province, China
Sun et al., 2008 [42]	The similarity is provided by the time interval and the Euclidean distance of the temperatures between the historical and target days	The forecasted day is computed using either linear extrapolation or a fuzzy Support Vector Machine, depending on the day type	Load and temperature data from 2004	Weekdays, weekends	Daily: (1.01, 2.23)	Chinese Nanjing Interconnected Power System, China
Chen et al., 2010 [43]	The similar day is selected by choosing the similar day type, load data, and minimizing the Manhattan distance of weather features	The similar day's load is decomposed into low- and high-frequency components, and then a separate neural network is used to forecast both components	Day type, wind-chill temperature, wind speed, dew point, humidity, cloud cover, and precipitation and load data from 2006	Weekdays, weekends	Monthly: (1.32, 3.24)	ISO New England, USA

Table 3. *Cont.*

Author, Year, and Reference	Conventional Similar Day Selection Method	Intelligent Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Jain and Jain, 2012 [44]	The similar day is selected by minimizing the Euclidean distance considering temperature, humidity, and day type with weight factors	Fuzzy logic, whose parameters are determined by evolutionary s optimization, is used to modify the load curves of the selected days	Day type, temperature, humidity, and load data from 1997	Weekdays, weekends	Daily: (1.09, 2.09)	EUNITE Network Competition
Ebrahimi et al., 2013 [45]	Similar days are identified using 3 essential load characteristics: average, standard deviation, and self-correlation	The forecasted load profile is constructed from 3 load profiles with the closest value of the 3 essential characteristics, each corrected with a set of fuzzy rules	Load data from 2002 to 2005	Holidays	Daily: (1.06, 16.20)	Isfahan, Iran
Eapen and Simon, 2018 [46]	The similar day is defined as the same day of the week one week earlier	Two independent back-propagation neural networks are trained separately for similar day and day-ahead load forecasting; the forecasted load is the output of a second-stage sequential neural network	Load data from 2004 (GEFCOM) and 2014 (ERCOT)	Weekdays, weekends	Over the dataset: ERCOT- Max 19.56; GEFCOM-Max 17.95	Electric Reliability Council of Texas (ERCOT), USA; Global Energy Forecasting Competition (GEFCOM) of 2012
Song et al., 2018 [47]	Similar days are chosen by minimizing the Euclidean distance of the load, hour by hour, of the past Lunar New Year's Day	Forecasting is performed by estimating the daily peak load using fuzzy linear regression and reconstructing the hourly load profile by scaling normalized similar day load patterns	Load data of Lunar New Year's Day from 1991 to 2018	Week of Lunar New Year's Day	Daily: (1.98, 7.04)	Republic of Korea
Barman et al., 2020 [48]	Similar days only for regional special event days and Euclidean distance for temperature	The forecasted load is determined using a Support Vector Machine with hyperparameters optimized with a Grey Wolf Optimizer	Temperature, load data, and hour for the regional special event days from 2012 to 2016	Regional special event days: Rongali Bihu, Durga Puja, and Diwali	Daily: Rongali Bihu: 1.79, Durga Puja: 1.72, Diwali:1.68	Assam Power Distribution Company Limited, India

Table 3. *Cont.*

Author, Year, and Reference	Conventional Similar Day Selection Method	Intelligent Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Gundu et al., 2022 [49]	Euclidean distance considering load data	Particle Swarm Optimization and Gated Recurrent Unit network trained with day-ahead and similar day data	Load data from February to March 2007	Weekdays, Weekends	Weekly: (0.49, 0.68)	Global Energy Forecasting Competition
Xiang et al., 2022 [50]	Euclidean norm of load data with sample weight assignment considering the type of day	Output of the deep neural network	Temperature, humidity, wind speed, precipitation, type of day, and load data from 2016	Weekdays, Weekends, Holidays	Weekly: (3.30, 10.80)	China
Zhang et al., 2022 [51]	The similar day is defined by a weighted Manhattan distance involving dry-bulb temperature, dew point, wet-bulb temperature, humidity, and load data	The load data is decomposed into low- and high-frequency components with wavelet transform; prediction of low- and high-frequency components with a long short-term neural network is conducted separately	Temperature, humidity, and load data from 2006 to 2010	Weekdays, Weekends	Weekly average: 1.66	Australia Electricity Market Operator, New South Wales, Australia
Wei et al., 2023 [52]	Pearson correlation coefficient method for meteorological factors, week, and time factor similarity for week type and type of day, and state regulations define the similarity day	Output of a deep belief network trained with similar days	Temperature, relative humidity, and load data from 2019 to 2022	Weekdays, Weekends, Holidays	Daily average: 3.06	China
Ortega et al., 2023 [53]	Similarity based on the average temperature of the main load zones and day type	The forecasted load is the outcome of a long short-term memory neural network	Temperature, type of day, and load data from 2016 to 2023	Weekdays	Hourly: (0.07, 9.44)	National Center for Energy Control, Mexico

Table 3. *Cont.*

Author, Year, and Reference	Conventional Similar Day Selection Method	Intelligent Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Jiao et al., 2023 [54]	The similarity is determined by calculating the Gaussian index between two feature vectors containing the temperature, irradiance, and weekday type	The predicted load is the output of a long short-term memory and radial basis function neural network	Temperature, irradiance, wind speed, wind direction, and day type	Weekdays, Weekends	Daily average: 2.33	China
Linzi et al., 2023 [55]	Minimization of the Euclidean distance of meteorological factors weighted by their grey relation degree	A long short-term memory neural network with input parameters optimized using a Firefly algorithm	Air temperature, relative humidity, rainfall, and load data from 2018 to 2020	Weekdays	Daily average: 0.0181	Regional grid in Western China
Aswanuwath et al., 2023 [56]	Weighted K-dimension Euclidean distance, where the K dimensions and the weights are the significant variables and coefficients obtained in a stepwise regression	The load data is decomposed and seasonality is obtained using variational and empirical mode decomposition and fast Fourier transform; this information is the input of an artificial neural network with optimized hyperparameter configuration to forecast the peak load	Day type and load data from 2016 to 2018	Weekdays, Weekends, Holidays	Average peak load: All days—4.89 Holiday days—7.09 Week and weekend days—2.80	Electricity Generating Authority of Thailand, Thailand
Aswanuwath et al., 2024 [57]	Similar day selected by day type for long-term input set	The input data are split into long, medium, and short-term; stepwise regression selects medium-term input data; short-term input data are used in the grid-optimized long short-term memory neural network; and long and medium-term input data are used for an optimized neural network	Load data from 2018 to 2020	Weekdays, Weekends, Holidays in disrupted situation	Monthly: average: 3.07	Electricity Generating Authority of Thailand, Thailand

Table 3. *Cont.*

Author, Year, and Reference	Conventional Similar Day Selection Method	Intelligent Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Wang et al., 2025 [58]	Similar day load profiles used as an input feature derived from historical load patterns	Load forecasting is conducted with a Bi-directional Gated Recurrent Unit neural network with multi-input multi-output architecture	Temperature, day type, and load data from 2018 to 2019	Weekdays, Weekends	Hourly: (3.67, 3.78)	Singapore's National Electricity Market System, Singapore

Table 4. Summary of key information on forecasting methods using intelligent similar day classification and conventional load forecasting techniques.

Author, Year, and Reference	Intelligent Similar Day Classification Method	Conventional Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Martínez-Álvarez et al., 2011 [59]	Day pattern similarity is found by clustering	The forecasted load is calculated by averaging the samples corresponding to the same pattern	Load data from 2005 and 2006	Weekdays, Weekends	Monthly: NYISO: (3.41, 6.88); ANEM: (3.91, 6.12); OMEL: (2.34, 6.21)	New York Independent System Operator, USA; Australia National Electricity Market, Australia; Electricity Market Operator, Spain
Feng and Ryan, 2016 [60]	K-means clustering is applied twice, first to find seasons and then to group days according to temperature	The load is decomposed into non-weather dependent, temperature-dependent, and humidity-dependent components, and then the forecasting is performed per component via epi-splines	Temperature, humidity, day type, hour of the day, season, and load data from 2011 to 2012	Weekdays, Weekends, Holidays	Daily weighted MAPE: (2.00, 4.00)	Independent System Operator New England, USA

Table 4. *Cont.*

Author, Year, and Reference	Intelligent Similar Day Classification Method	Conventional Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Martínez-Álvarez et al., 2019 [61]	Pattern-based similar day-inspired method where pattern similarity is found by clustering to discretize the time series in a sequence of labels	The sequence of labels matching the sample to be forecasted is obtained and weighted averaging is performed to obtain the forecasted load	Load data from 2007 to 2014	Weekdays, Weekends, Holidays	Monthly: (3.87, 8.75); Daily: (5.06, 6.01); Hourly (5.09, 6.99)	Uruguay
Pérez-Chacón et al., 2020 [62]	Pattern-based similarity is found by clustering of big data load time series using K-means, obtaining a sequence of labels	The forecasted load is obtained via weighted averaging of the sequence of labels similar to the target day	Big data of load time series from 2007 to 2014	Weekdays, Weekends, Holidays	Daily: (4.70, 6.96)	Uruguay
Jiang et al., 2023 [63]	A neural network that assigns similarity weights for similar day screening is employed	The forecasted day is computed with a linear model, auto regression, based on the similar day screening results	Load data from 2009 to 2015	Weekdays, Weekends, Holidays	Daily: (2.83, 10.98)	A region of China

https://doi.org/10.3390/forecast8020032

Therefore, the short-term load forecasting methodologies under the category of intelligent SD with conventional forecasting methods mainly use clustering, with different algorithms, SOM, regression trees or grey relational analysis to obtain similar days. Then, they employ conventional forecasting techniques such as regression or ARIMA models within each group of similar days.

3.3.4. Intelligent Similar Day Methods and Intelligent Forecasting Approaches

A summary of key studies using intelligent methods for similar day classification and load forecasting techniques is presented in Table 5.

Thus, the short-term load forecasting methodologies under the category of intelligent SD with intelligent forecasting methods use a machine learning technique to select SDs, such as clustering, SOM, grey relational analysis, etc., and to forecast the load, such as ANN, LSTM, GRU, etc.

3.3.5. Hybrid Similar Day Metrics Forecasting Approaches

The summary of the main works employing hybrid methods for similar day classification or load forecasting techniques is presented in Table 6.

Thus, the short-term load forecasting methodologies under the category of hybrid methodologies for SD and forecasting models often use multi-criteria, such as calendar rules, distance, correlations and grey relation analysis, for the selection of SD. Concerning load forecasting, load decomposition, using wavelets or EMD/VMD, is often used and forecasting is then performed per component. Additionally, machine and deep learning techniques are also used to add residual corrections. Other methodologies present ensembles of models combining SD-based forecasts with non-SD models.

3.4. Distribution of Studies by Methodological Category

Table 7 provides a taxonomy-based quantitative synthesis of these methodologies identified across the reviewed studies, which are classified into conventional, intelligent, and hybrid SD approaches.

The results indicate that conventional SD methods are predominantly distance-based, with Euclidean distance (26%) and weighted Euclidean distance (18%) being the most frequently employed metrics. Calendar-based selection (18%) and Pearson correlation (10%) also remain widely used, confirming the continued relevance of deterministic similarity measures in SD modeling.

In contrast, intelligent SD approaches are primarily driven by clustering techniques, particularly K-means (43%) and other algorithms (14%), followed by regression tree methods (14%) and neural-network-based or metaheuristic techniques (29%). This distribution reflects a methodological transition toward data-driven similarity construction, where implicit grouping mechanisms replace explicit metric definitions.

Hybrid SD methodologies integrate deterministic distance measures with clustering or learning-based components. Within this category, K-means-based strategies (22%) and Euclidean-type distances (18%) remain central, followed by correlation analysis (8%), grey relational analysis (4%), or neural network modules (4%).

Overall, the quantitative distribution highlights a progressive evolution from purely metric-based similarity definitions toward clustering-enhanced and hybrid architectures, evidencing the growing methodological sophistication of SD-based forecasting frameworks.

Table 5. Summary of key information on forecasting methods using intelligent similar day classification and intelligent load forecasting techniques.

Author, Year, and Reference	Intelligent Similar Day Selection Method	Intelligent Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Mori et al., 2001 [21]	Deterministic annealing clustering of attributes, such as temperature, humidity, discomfort index, and load, is used to form groups of similar days	The maximum load is the output of a multilayer perceptron neural network, which is trained for each cluster	Temperature, humidity, discomfort index and load data from 5 years	Weekdays	Monthly: (3.27, 4.91) Daily: (2.96, 3.66)	Japan
Wang, 2006 [64]	Case-based reasoning with self-organizing maps is used to cluster similar days	To extract the principal case attributes, a fuzzy set method is used; the peak load of the forecasted day is obtained from the ones of the most similar case	Temperature, humidity, precipitation, day type, and load data from 2000 to 2002	Weekdays, Weekends, Holidays	Daily average: 0.05	Hang Zhou Electric Power Company, China
Guo and Niu, 2008 [65]	Using classification and regression tree recognizes different patterns of daily load temperature, humidity, rainfall, air pressure, cloud quantity, wind speed and direction, and day type	The forecasted load is obtained as the output of an artificial neural network	Weather and date attributes and load data from 2007	Weekdays, Weekends, Holidays	Not reported	HeBei Electric Power Company, China
Guo and Niu, 2009 [66]	Using classification and regression tree recognizes different daily patterns based on temperature, humidity, rainfall, air pressure, cloud cover, and wind speed	The forecasted load is obtained using Support Vector Regression	Weather and date attributes and load data from 2007	Weekdays, Weekends, Holidays	Weekly: (1.73, 3.18);	HeBei Electric Power Company, China
Duan et al., 2011 [67]	Similar samples are obtained via fuzzy C-means clustering optimized with Particle Swarm Optimization	The load forecasting value is obtained using a Support Vector Regression model whose parameters are optimized with Particle Swarm Optimization	Temperature and load data	Weekdays, Weekends	Hourly: (1.06, 2.79)	Chongqin, China

https://doi.org/10.3390/forecast8020032

Table 5. *Cont.*

Author, Year, and Reference	Intelligent Similar Day Selection Method	Intelligent Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Martins et al., 2022 [68]	Similar day groups are found by K-medoids clustering to identify daily load patterns using a convex distance, which is a combination of Euclidean, Pearson-base, periodicity, and autocorrelation-based distances	A temperature-based k-nearest neighbor filtering is applied and a weighted mean of the selected neighbor day's load profiles are performed	Temperature, day type, and load data from 2014 to 2017	Weekdays, weekends, Holidays	Monthly: (2.7, 5.1)	Portuguese Transmission System Operator, Portugal
Janković et al., 2022 [69]	Genetic algorithm-based smart similar day selection that calculates weight coefficients for similarity factors	A feed-forward multilayer artificial neural network with optimized hyperparameters with genetic algorithms forecasts the hourly load	Temperature, wind speed, sky cover, type of day, and load data from 2015 to 2020	Weekdays, Weekends	Monthly: (1.59, 2.20)	Norwegian Transmission System Operator, Norway

Table 6. Summary of key information on forecasting methods using hybrid approaches.

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Chen et al., 2008 [70]	Euclidean distance considering wind chill temperature and load data	First, low and high load decomposition is conducted using wavelet decomposition; then, multilayer perceptron networks are used to forecast the low- and high-frequency load	Load, wind speed and direction, cloud cover, wind chill temperature data, and dew point from 2003 to 2005	Weekdays, Weekends, Holidays	Weekly: (1.32, 3.24)	New England, USA.
Lu and Liu, 2011 [71]	Similar days are selected by clustering using K-means based on temperature similarity	The daily load curve is divided into 3 time zones according to temperature sensitivity, and the forecast uses either linear regression or Support Vector Regression; a linear programming smoothing stage is applied to the final forecasted load curve	Temperature, day type, and load data from 2001 to 2003	Weekdays	Daily: (1.02, 2.67)	Ningbo Electricity Corporation, Zhejiang province, China

Table 6. *Cont.*

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Al-Qahtani and Crone, 2013 [72]	K-nearest neighbor regression based on Euclidean distance in a multivariate feature space	The forecasted load is obtained by averaging the values of the k-means clustering of the most similar historical days	Day type and load data from 2001 to 2008	Weekdays, Weekends, and bank holidays	Daily average: 1.81	UK electricity load data, UK
Moazzami et al., 2013 [73]	Multiple correlation coefficient analysis of load with weather variables, day type, and season	Daily peak load signal is decomposed into low- and high-frequency components with Discrete Wavelet Transform; artificial neural networks, optimized with genetic algorithms, are trained for each frequency; and the final forecast is obtained via wavelet reconstruction	Temperature, humidity, wind speed, and load data from 2006 to 2011	Weekdays	Daily: (1.20, 1.58)	Iran's National Grid, Iran
Zheng et al., 2017 [74]	An extreme gradient boosting-based weighted K-means algorithm is used to cluster the similar days	The load profiles are decomposed into intrinsic mode functions and residual errors, which are then forecasted using a long short-term memory neural network separately	Temperature, humidity, precipitation, wind speed, day type, and load data from 2003 to 2016	Weekdays, Weekends, Holidays	Monthly: (0.75, 1.87)	Independent System Operator New England, USA
Sun et al., 2018 [75]	Factor Analysis is used to obtain a set of latent factors from 22 variables related to load behavior; then, ant colony clustering is used to obtain similar days according to similar latent factors	The forecasted load is the output of an extreme learning machine, whose input weights and bias threshold are optimized with a bat algorithm	Temperature, humidity, visibility, air pressure, wind speed, day type, date, and load data from 2013	Weekdays, Weekends, Holidays	Daily average: 0.49	Yangquan City, China

https://doi.org/10.3390/forecast8020032

Table 6. *Cont.*

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Barman et al., 2018 [76]	Similar days are selected from minimizing the weighted Euclidean distance considering temperature and humidity	Support Vector Regression model with radial basis function kernel, whose hyperparameters are optimized using the Grasshopper Optimization Algorithm, trained with similar day data	Temperature, humidity, day type, and load data from 2013 to 2015	Weekdays, weekends, holidays	Weekly: (1.40, 1.50)	State Load Dispatch Centre, Assam, India
Liu et al., 2018 [77]	Similar days are selected using a fuzzy combination weights approach that integrates entropy-weighted temperature similarity and weighted K-means clustering of load profiles, ensuring consistency in both weather conditions and load pattern characteristics	Selected similar day load series are decomposed using empirical mode decomposition; each component is forecasted using a bat algorithm-optimized Support Vector Machine and refined via a Kalman filter, and the final load forecast is obtained by reconstructing all predicted components	Temperature, day type, and load data from 2014 to 2015	Weekdays, weekends	Daily: (1.35, 2.05)	South China, China
Nazar et al., 2018 [78]	Similar days are selected by clustering wavelet-decomposed daily load patterns using the Kohonen self-organizing map	Forecasting is performed using a three-stage hybrid framework combining wavelet decomposition and Kalman filtering for initial prediction, followed by multilayer perceptron neural networks for load forecasting and a final refinement stage to reduce residual errors	Temperature, electricity price, and load data from 2011 to 2012	Weekdays, weekends	Monthly: (1.31, 3.32)	Nord Pool electricity market, Europe; Mainland Spain electricity market, Spain

Table 6. *Cont.*

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Liao et al., 2019 [79]	Similar days are selected using clustering based on day type and time distance and the degree of normalized temperature correlation between the day to be predicted and the historical days	The forecasted load is calculated with extreme gradient boosting regression trained on the similar day samples	Temperature, day type, and load data from 2014 to 2018	Weekdays, weekends	Calculated over a period of two weeks of testing: 0.088	State Grid Corporation of China, China
Sun et al., 2019 [80]	Euclidean distance considering the heat index, day type, and calendar constraints	A seasonal Auto Regressive Integrated Moving Average is used for a first prediction, and then a neural network is used to capture the nonlinearity for a refined prediction	Temperature, cloud cover, wind speed, heat index, and day type from 2012	Weekdays	Weekly: (4.83, 8.42)	State Load Dispatch Center, Assam, India
Barman and Choudhury, 2019 [81]	Introduce a season-specific similarity based on a weighted Euclidean distance using season-specific meteorological variables	The forecasted load is the output of a Support Vector Machine with hyperparameters optimized with the Firefly Algorithm	Temperature, cloud cover, wind speed, day type, and load data from 2013 to 2016	Weekdays, weekends	Weekly: (1.56, 1.79)	State Load Dispatch Center, Assam, India
Huang et al., 2019 [82]	Grey relational analysis is used to identify the most influential external factors, and then fuzzy C-means clustering, including load and calendar characteristics, is used to group the historical data into similar days	An Analytic Hierarchy Process with weights is used with a back-propagation neural network, a Support Vector Machine optimized with Particle Swarm Optimization, and an Auto Regressive Integrated Moving Average time series model to obtain the load forecast	Temperature, weekday index, month, date, holiday indicators, and load data	Weekdays, weekends, holidays	Daily: (0.04, 9.09); Overall: 3.42	Regional system power, China

https://doi.org/10.3390/forecast8020032

Table 6. *Cont.*

Author, Year, and Reference	**Similar Day Classification Method**	**Technique for Load Forecasting**	**Input Model Variables**	**Type of Target Day**	**Achieved MAPE (%) (min, max)**	**Data Source (Location)**
Kong et al., 2020 [83]	Starting from the nearest historical days, a multi-factor correlation-based similarity function is applied to select the similar days	Hybrid technique: (1) apply empirical mode decomposition to the load and price series; (2) convert them into images and apply them to a long short-term convolutional neural network for prediction; and (3) apply them to a fully connected neural network	Electricity price, day of the week, hour of the day, holiday index, and hourly load data from 2014 to 2016	Weekdays, weekends, holidays	Hourly average: 1.10	Singapore Electricity Market, Singapore
Park et al., 2020 [84]	Similar days are selected using a Reinforcement Learning-based model (deep Q-network), which learns an optimal similarity policy by maximizing a reward provided by a Euclidean distance	The forecasted load is the output of a back-propagation neural network	Temperature, solar irradiation, rainfall, day distance, weekdays, special days, and load data from 2018	Weekdays, weekends	Monthly: (1.17, 1.83); Daily: (0.93, 1.80)	Korea Power Exchange, Repulic of Korea
Lu et al., 2020 [85]	Similar days are selected by minimizing a weighted meteorological similarity measure; the season-derived weights are derived from partial correlation coefficients	The load forecast is performed with a Wavelet Neural Network whose output is corrected using a probabilistic error based on historical forecast errors from the meteorological similar days	Temperature, relative humidity, rainfall, season indicators, and load data from 2013 to 2014	Generic day of each season	Maximum hourly MAPE: 1.38	China

Table 6. *Cont.*

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Gunawan and Huang, 2021 [86]	Dynamic nearest neighbor with compensation consisting of selecting the nearest workday and compensatory holidays and the use of dynamic time warping to compare daily load curves with select similar days	The nearest workday and a compensatory holiday load is forecasted by means of a long short-term memory neural network, and these outputs are the inputs of another network to forecast the load of the target holiday	Temperature, day of the week, holiday occurrence, hour of the day, and load for 9 holidays	Holidays	Daily: (2.00, 4.00)	Asia
Wang et al., 2021 [87]	An improved maximum deviation similarity criterion is used; daily load curves are clustered according to shape similarity, with deviation thresholds, and the target day is assigned to the most similar cluster	The day is segmented in 7-time intervals, and a back-propagation neural network is used for each segment; the daily forecast is obtained by integrating all the segment predictions	Temperature, humidity, wind speed, precipitation, and load data from 2017 to 2019	Weekdays, weekends, holidays	Daily: (1.62, 1.98)	State Grid, Henan Province, China
Dudek, 2021 [88]	Daily load sequences are represented as normalized patterns, and then the Euclidean distance/dot product is computed in them to select the similar days	The load is predicted using a neural network, where the training error is weighted based on pattern similarity to the target day	Load data from 2012 to 2015	Weekdays, Weekends	Daily: Poland (1.28, 1.48); Great Britain (2.95, 3.15); France (1.80, 1.87); Germany (1.50, 1.64)	European Network of Transmission System Operators for Electricity, Europe

Table 6. *Cont.*

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Janković et al., 2021 [89]	Similar day selection is carried out with a multi-stage filtering with weighting coefficients optimized with Particle Swarm Optimization using weather factors, daylight duration, and load	The forecasted load is the outcome of a feed-forward artificial neural network with parameters optimized with Particle Swarm Optimization	Temperature, wind speed, daylight duration, day type, and load data from 2018	Weekdays, Weekends	Hourly: (0.06, 2.29)	Elektromreža Srbije, Transmission System, Serbia
Zhao et al., 2021 [90]	Light gradient boosting machine-based feature importance analysis followed by K-means clustering for similar load pattern selection	The forecasted load is obtained with an encoder–decoder Transformer neural network with a self-attention mechanism	Temperature, humidity, wind speed, hour of the day, day index, and load data from 2006 to 2010	Weekdays, Weekends, Holidays	Monthly: (0.93, 1.23)	New South Wales System, Australia
Xiong et al., 2021 [91]	Similar days are dynamically identified by calculating the Manhattan distance between feature vectors of historical and forecast days, with weights assigned using a SoftMax-based attention mechanism	Load forecasting is performed using an attention-based encoder–decoder bi-directional long short-term memory architecture that integrates feature-selection attention in the encoder and hierarchical temporal attention in the decoder	Temperature, humidity, wind speed and direction, holiday indicator, hour of the day, day of the week, month of the year, utility discount programs, and load data from 2006 to 2014	Weekdays, Weekends, Holidays	Overall annual: 1.93	Global Energy Forecasting Competition 2014 Extended dataset
Dong et al., 2022 [92]	K-means clustering applied to historical load data using temperature, day type, and seasonal attributes	The forecasted load is obtained with Support Vector Regression with radial basis function kernel trained with the cluster corresponding to the target day	Temperature, day type, and seasons from 3 periods spanning 1998 to 2020	Weekdays, Holidays	MAPE values not reported	Nantong Power Supply Company, China; EUNITE Load Forecasting Competition, Slovakia; REFIT, UK

Table 6. *Cont.*

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Zhao et al., 2022 [93]	K-means clustering considering the Euclidean distance of the load data	Quantile regression long short-term memory neural network forecasts the load for different quantiles and then the probability density function is obtained for the target day	Load data from 2012 to 2013	Weekdays, Weekends	Weekly: Average: 9.19	New South Wales, Australia
Pang et al., 2022 [94]	The Pearson correlation coefficient method is used to identify the meteorological factors and day type influencing the load; the similar day sets are constructed using gray relational analysis with the Pearson coefficient as the weight; and finally, the set of similar days is refined with K-means clustering	Support Vector Machine optimized by an artificial bee colony algorithm	Temperature, wind, radiation, rainfall, type of day, and load data from 2003	Weekdays, Weekends, Holidays	Weekly: Average: 1.79	Nanjing, China
Shi and He, 2022 [95]	Daily load series clusters based on shape selection are made by the K-shape method; then, the extreme gradient boosting classifier is used to classify the data according to meteorological and date information	A neural network with two inputs: the shape like-similar days and the meteorological-like similar days is used to forecast the load profile	Temperature, sunshine hours and other meteorological facts; date information; and load data from 2017 to 2021	Weekdays, Weekends	Weekly in one sample: 2.74	Southeast China

Table 6. *Cont.*

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Yang et al., 2022 [96]	Weighted difference function of load, day type, and meteorological factors	Fuzzy Support Vector Machine to forecast daily maximum and minimum load and then linear extrapolation of similar days for the load curve	Temperature, humidity, day type, and load data from 2017	Weekday, Weekend	Daily: Maximum-4.07; Average-1.55	China
Wang et al., 2023 [97]	Comprehensive grey relational analysis of meteorological factors, load data, and type of day	A recurrent neural network with Gated Recurrent Units	Load data from 2006 to 2018	Weekdays, Weekends	Weekly: Average: 1.85	ENTSO-E, Europe
Pajić et al., 2023 [98]	Similar days are found by using a multilayer encoder neural network that minimizes the Euclidean distance considering load and temperature data	The forecasted load is the output of a fully connected feed-forward artificial neural network	Temperature, calendar, seasonality, and load data from 6 years (period not specified)	Weekdays	Daily value for the forecasted day: 2.21	Southeast European utility, Europe
Criado-Ramón et al., 2023 [99]	Similar days are found using a pattern sequence-based method that applies self-organizing maps, optimized with genetic algorithms, to find clusters with similar characteristic load shapes	A neural network, whose parameters are optimized with genetic algorithms, is trained with similar pattern sequence clusters	Type of day, month, and load data from 2009 to 2019	Weekdays, Weekends	Over all test period: REE—3.62; AEMO—4.74; NYISO—5.22	Spanish Electric Network (REE), Spain; Australian Energy Market Operator (AEMO), Australia; New York Independent System Operator (NYISO), USA

Table 6. *Cont.*

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Bu et al., 2023 [100]	The grey weighted correlation method is used to generate similar days by calculating the Pearson correlation coefficient between load and relevant factors	Variational Mode Decomposition is performed and a conditional generative adversarial network, with a convolutional neural network, captures features of each mode and extracts the nonlinear and dynamic behavior to forecast each mode	Temperature, humidity, rainfall, day type, and load data from 2014	Weekdays, Weekends	Monthly: (1.70, 4.21)	Shanxi Province, China
Xiong and Zhang, 2023 [101]	Feature-distance similarity with soft similar day weighting implemented through hierarchical temporal attention	Attention-based encoder–decoder within a bi-directional long short-term memory and Gated Recurrent Unit framework with feature weighting and error correction	Temperature, humidity, wind speed and direction; holiday, season, and weekend indicators; and hour of the day, day of the week, month of the year, and load data in 2 periods: 2015 to 2017 and 1985 to 1992	Weekdays, Weekends, Holidays	Yearly: (1.66, 2.15)	Independent System Operator—New England, USA, (2015–2019) and North American Unity, USA (1987–1991)
Luo et al., 2024 [102]	Adaptive hierarchical clustering considering temperature, day type, weather type, and load data	Stacking integrated algorithm based on convolutional neural networks, bi-directional long short-term memory attention, and XGBoost	Temperature, weather type, day type, and load data from 2018	Weekdays, Weekends	Daily: (4.74, 5.20)	Quanzhou, China
Amini et al., 2024 [103]	Similar days are selected using fuzzy clustering and consider the weight of the factor's influence on the load, such as day type, temperature, season, and TV peak-up	The forecasted load is the output of a long short-term memory neural network with the normalized energy of similar days as the input	Temperature, day type, wind, humidity, and load data from 2019 to 2023	Weekdays, Weekends, Holidays	Daily: (1.32, 3.63)	Pennsylvania-New Jersey-Maryland Regional Transmission Organization, USA

Table 6. *Cont.*

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Wang et al., 2024 [104]	Similar days are chosen according to day type and calendar-based distance	A convolutional neural network with a long short-term neural network optimized using a Sparrow Search Algorithm forecasts the load	Temperature, humidity, day type, and load data from 2010	Weekdays, Weekends, Holidays	Daily: Weekday-1.2. Weekend-2.0. Holiday-1.6	Zhejiang Province, China
Borunda et al., 2025 [105]	K-means clustering of daily load profiles is conducted to find seasonality; for each cluster (season), similar day selection is carried out based on day type	The forecasted load is the output of a long short-term memory neural network trained for each season, including the representative load and meteorological conditions of the region	Temperature, relative humidity, irradiance, atmospheric pressure, wind speed, and load data from 2016 to 2022	Weekdays, Weekends	Daily: (0.63, 1.48); Weekly: (1.14, 1.25)	National Center for Energy Control, Mexico
Luo et al., 2025 [106]	Gaussian mixture clustering with an improved crested porcupine optimizer using weather and weekday type	Load data is decomposed into multi-scale components using optimized feature mode decomposition; the quantile interval forecasting is performed with a convolutional Bi-directional Gated Recurrent Unit attention model; and the final forecast is obtained by aggregating the component values and their intervals	Temperature, humidity, wind speed, seal level pressure, and load data from 2018	Weekdays, Weekends	Daily: (1.40, 2.18)	Elia Company, Belgium
Cheng et al., 2025 [107]	A weighted eigenspace coordinate system considering temperature, relative humidity, precipitation, weekday index, and load data is used to measure similarity	Decomposition of low-, medium-, and high-frequency components using Variational Mode Decomposition with Beluga Whale Optimization; high-frequency modes are predicted with a similar day average and low- and medium-frequency modes with a neural network	Temperature, relative humidity, precipitation, type of day, and load data from 2012 to 2014	Weekdays, Holidays	Daily: (2.2, 20) Average: 7.81	China

https://doi.org/10.3390/forecast8020032

Table 6. *Cont.*

Author, Year, and Reference	Similar Day Classification Method	Technique for Load Forecasting	Input Model Variables	Type of Target Day	Achieved MAPE (%) (min, max)	Data Source (Location)
Zhao et al., 2025 [108]	Days are classified into weekdays and weekends, and then similar days are identified with K-means clustering based on load profiles and weather type	Auto Regressive Integrated Moving Average and Support Vector Regression generate forecasts, which are combined through an adaptive weighted ensemble optimized with a Particle Swarm Optimization algorithm with elastic momentum	Temperature, humidity, day type, and load data from 2015	Weekdays, Weekends	Not reported	Microgrid demonstration project in northwest China, China

Table 7. Percentage of the main SD methodologies used in the conventional, intelligent, and hybrid load short-term forecasting approaches studied in the literature.

Type of SD Method	SD Method		Percentage (%)
Conventional (32 references)	Distance	Euclidean	27
		Weighted Euclidean	21
		Manhattan	9
	Calendar day/day type		13
	Pearson coefficient		6
	Cosine similarity		6
	Pattern of the day		9
	Others	Fractal measure	9
		Mean temperature	
		Gaussian coefficient	
Intelligent (14 references)	Clustering	K-means	15
		K-medoids	7
		Fuzzy C-means	7
		Annealing	7
		Pattern similarity	22
	Regression tree		14
	Others	Neural network	28
		Self-organizing maps	
		Genetic algorithms	
		Particle Swarm Optimization	
Hybrid (39 references)	Distance	Euclidean	11
		Weighted Euclidean	9
		Manhattan	2
	Clustering	K-means	13
		K-means with weights	2
		Fuzzy	5
		Other algorithms	9
	Pearson correlation coefficient		4
	Correlation analysis		7
	Grey relational analysis		7
	Calendar day type		5
	Attention-based		4
	Neural networks		4
	Others	Multi-stage filtering	18
		K-shape method	
		Gradient boosting	
		Self-organizing maps	
		Genetic algorithm	
		Particle Swarm Optimization	
		Weighted eigenspace	

3.5. Temporal Evolution

Overall, the literature reveals a gradual shift from simple, rule-based SD formulations toward intelligent and hybrid SD frameworks that integrate clustering, machine learning, and feature-extraction techniques. This structure supports the taxonomy presented in the next subsection. Figure 5 shows the methodological evolution of SD load forecasting approaches. Early studies were predominantly based on conventional similar day approaches combined with classical forecasting methods. Over time, a transition toward intelligent and hybrid approaches can be observed, reflecting the integration of machine learning and advanced data-driven techniques.

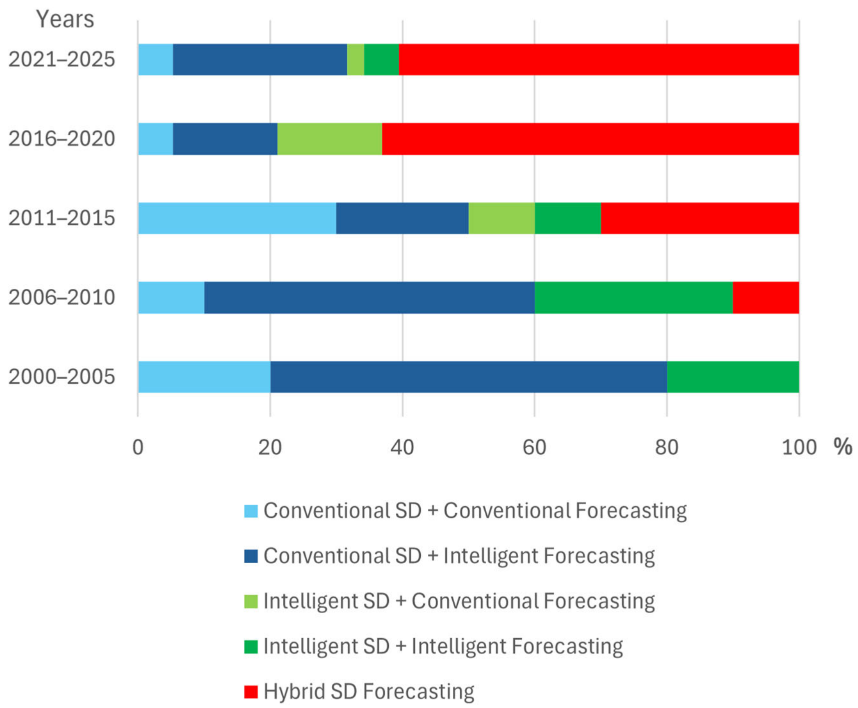

Figure 5. Methodological evolution of similar day load forecasting approaches across different time periods. The stacked bar chart illustrates the relative proportion of methodological categories, highlighting a clear transition from conventional similar day approaches toward intelligent and hybrid forecasting frameworks in recent years.

3.6. Technical Characteristics of Reviewed Studies

The analysis of the reviewed literature reveals several consistent patterns in terms of input variables, forecasting configurations, and methodological preferences.

Regarding input variables, most studies rely primarily on historical load data as the core predictor. In addition, weather-related variables—particularly temperature—are frequently incorporated due to their strong influence on electricity demand. Other variables, such as humidity, wind speed, and solar radiation, are included less consistently. Temporal features, including hour of the day, day of the week, and seasonal indicators, are also commonly used to capture periodic patterns in load behavior.

In terms of forecasting scenarios, the majority of studies focus on regular operating conditions, particularly weekday and weekend load forecasting. However, several works explicitly address special cases such as holidays, which are characterized by more irregular demand patterns. These special day scenarios are generally more challenging due to reduced predictability and deviations from typical consumption behavior.

From a methodological perspective, a clear predominance of hybrid and intelligent approaches can be observed in recent studies. While conventional SD methods remain relevant, especially in simpler or more stable contexts, there is a growing trend toward combining SD selection with machine learning or deep learning techniques. These hybrid approaches aim to improve forecasting accuracy by integrating pattern-based reasoning with data-driven modeling capabilities.

https://doi.org/10.3390/forecast8020032

Overall, these findings highlight a transition from traditional, rule-based SD approaches toward more flexible and adaptive frameworks that incorporate additional variables and advanced modeling techniques. These technical characteristics provide important context for understanding the performance results presented in the following subsection.

3.7. Performance Overview Based on MAPE

The performance of the reviewed methods was analyzed using the MAPE. Figure 6 presents the distribution of MAPE values across the proposed taxonomical categories.

The subsequent results follow:

1. The best and most consistent results are obtained with Hybrid approaches (Figure 6e), despite some outliers; the results show the lowest variability;
2. The second-best results (second-lowest variability across all approaches) were obtained from conventional SD + intelligent forecasting models (Figure 6b), with some outliers;
3. The third-best results belong to conventional SD + conventional forecasting models (Figure 6a). In this approach, weekly MAPE results were not reported, but the variability in hourly, daily, and monthly MAPEs is lower than that in the intelligent approaches shown in Figure 6c,d;
4. The fourth-best results are for intelligent SD + conventional forecasting models (Figure 6c). In this approach, weekly MAPE results were not reported, but the variability in hourly, daily, and monthly MAPEs is lower than the intelligent approaches of Figure 6d;
5. Resultant MAPE value for intelligent SD + intelligent forecasting models (Figure 6d) shows the widest interquartile range (IQR) for daily and monthly MAPEs, but with minimum variability between the lowest and highest MAPE values, that is, very short whisker length;
6. The largest number of results reported across all methodologies corresponds to the daily MAPE. For this reason, Figure 6f shows a comparison of the daily MAPE values obtained by all methodologies. The best results are achieved with conventional SD + intelligent forecasting models and hybrid SD approaches (combined similarity + combined forecasting), with low MAPE values and low variability. It is important to highlight that conventional SD + conventional forecasting models achieve excellent results comparable to those of intelligent methods, demonstrating the effectiveness of the method.

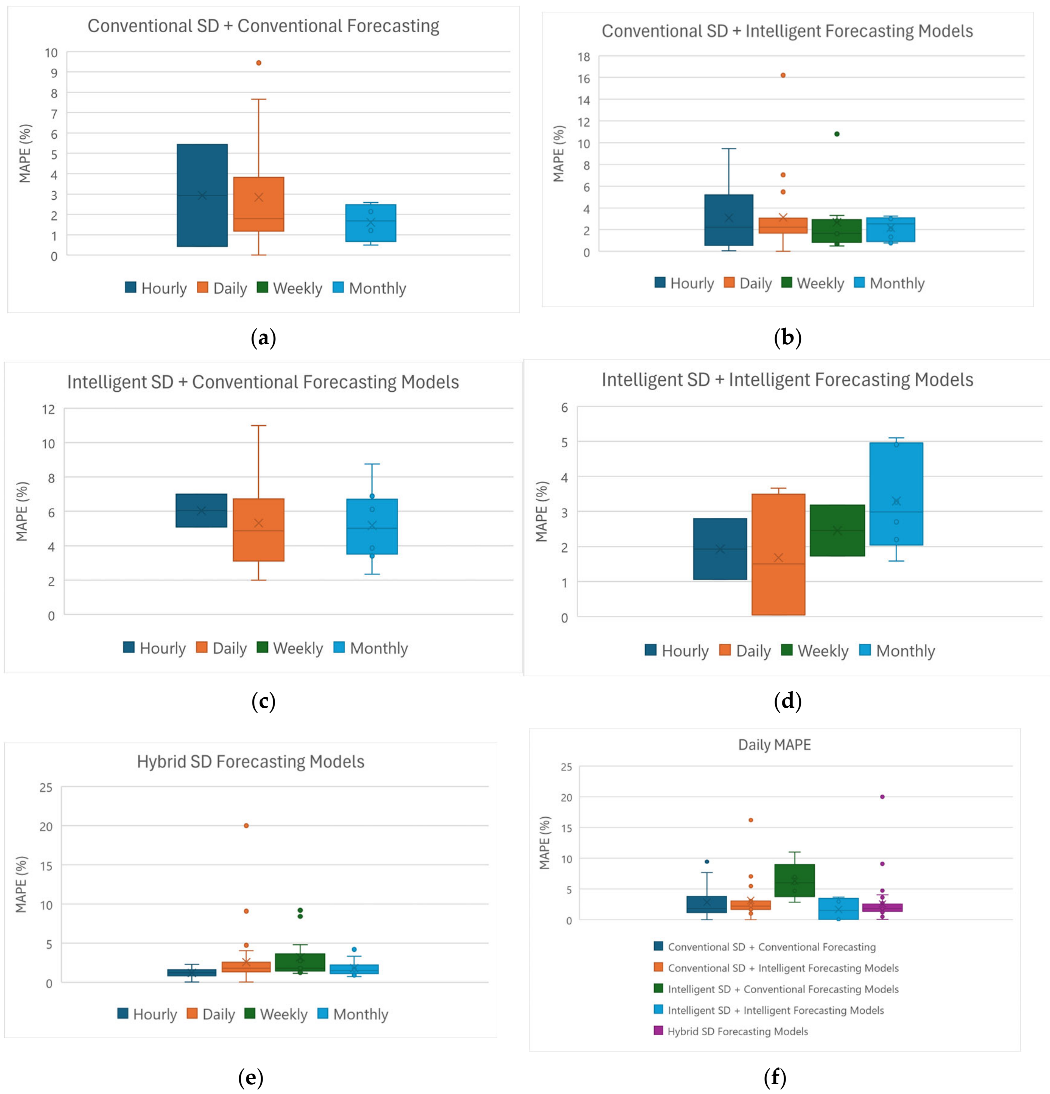

Figure 6. Box plot charts for the different MAPE values—hourly, daily, weekly, and monthly—reported in the reviewed works. The graphs were produced according to the classification of the proposed methodology: (**a**) conventional SD + conventional forecasting models; (**b**) conventional SD + intelligent forecasting models; (**c**) intelligent SD + conventional forecasting models; (**d**) intelligent SD + intelligent forecasting models; and (**e**) hybrid SD forecasting models. Figure (**f**) shows the comparative boxplot of the daily MAPE values for the 5 types of models. Note that there were no reported data for weekly MAPES for figures (**a**,**c**).

3.8. Key Findings from the Reviewed Studies

Building on the technical characteristics described previously, several key findings emerge from the reviewed studies.

First, the selection of similarity variables is a critical factor for forecasting performance, with historical load and temperature consistently identified as the most influential predictors.

Second, methods that employ more advanced similarity-definition techniques–such as clustering, weighted metrics, or intelligence-based procedures—generally outperform those based on fixed distance rules.

Third, the integration of intelligent techniques in either the SD selection or forecasting stage enhances accuracy, particularly under nonlinear and evolving demand conditions.

Finally, hybrid approaches—where similarity and forecasting are tightly coupled—offer a practical balance between interpretability and accuracy, emerging as one of the most promising directions for SD-based short-term load forecasting.

4. Discussion

While the previous section provides a structured overview of the reviewed studies, a deeper analysis is required to interpret the observed patterns and their implications. This section presents a cross-cutting discussion of methodological performance, limitations, and emerging trends, with particular emphasis on the conditions under which Similar-Day approaches are effective or fail, and their practical relevance for power system applications.

4.1. Cross-Cutting Comparison

A cross-cutting comparison of the reviewed methodologies reveals clear trade-offs across several key dimensions, including accuracy, robustness, interpretability, and implementation complexity.

In terms of accuracy, hybrid Similar-Day approaches and intelligent SD combined with advanced forecasting models consistently achieve the best performance, while conventional SD combined with classical forecasting methods tends to show lower accuracy, particularly in complex or highly dynamic systems.

Regarding robustness, conventional approaches are more vulnerable to regime changes and extreme events, such as those caused by behavioral shifts, electrification trends, or climate variability. Although intelligent and hybrid approaches generally exhibit improved adaptability, they still require additional mechanisms—such as data augmentation or specialized modules—to effectively handle rare or anomalous conditions.

From the perspective of interpretability, conventional Similar-Day methods remain the most transparent, as their decision-making process is directly linked to observable historical patterns. Intelligent SD approaches based on clustering offer a moderate level of interpretability, while fully intelligent and hybrid models tend to behave as black-box systems unless explainable artificial intelligence (XAI) techniques [109] are incorporated.

These methodological differences are also reflected in their data and infrastructure requirements. Conventional approaches require minimal data and computational resources, whereas intelligent and hybrid methods demand more extensive datasets, higher computational capacity, and more sophisticated model tuning. Similarly, implementation complexity increases progressively from conventional to intelligent and hybrid approaches.

Overall, these results indicate that no single methodology is universally optimal. Instead, the selection of an appropriate approach depends on the specific operational context, including data availability, system variability, and the required balance between accuracy, interpretability, and computational cost.

To provide a structured comparison of the reviewed methodologies, Table 8 summarizes their main advantages, limitations, and typical conditions under which each approach performs best.

Table 8. Comparative analysis of methodologies, including examples with their advantages, disadvantages, and the situations in which they perform best.

Methodologies	Advantages	Disadvantages	Best Performance When:
Conventional SD + Conventional Forecasting Models	High interpretability, simplicity, low computational requirements, easy implementation, historically proven in operational environments	Limited ability to capture nonlinearities, poor adaptability to structural changes, weak performance under extreme events, limited feature space	Stable systems, predictable weather patterns, environments prioritizing interpretability and robustness over maximum accuracy
Conventional SD + Intelligent Forecasting Models	Improved nonlinear modeling, better accuracy with minimal change in SD logic, easier integration into existing workflows	Performance depends on similarity definition, risk of overfitting, reduced interpretability	Scenarios requiring moderate accuracy improvement with available data and ML expertise
Intelligent SD + Conventional Forecasting Models	Improved similarity definition via clustering, better data homogeneity, enhanced performance with simple models, balanced interpretability	Requires careful preprocessing, complex parameter tuning, limited performance in extreme or rare events	Systems with clear regimes (seasonal, weather-driven), where improving similarity is key
Intelligent SD + Intelligent Forecasting Models	High accuracy potential, ability to capture nonlinear and dynamic patterns, flexible architectures, suitable for probabilistic forecasting	High complexity, reduced transparency, risk of overfitting, high computational cost	Complex and evolving systems where accuracy is a priority and advanced infrastructure is available
Hybrid SD Forecasting Models	Best overall performance, robustness across conditions, modular design, supports probabilistic forecasting	Very high complexity, integration challenges, difficult error attribution	Highly complex environments requiring state-of-the-art performance and robust forecasting frameworks

https://doi.org/10.3390/forecast8020032

This comparison highlights that methodological selection involves clear trade-offs between accuracy, interpretability, robustness, and implementation complexity. Building on this comparison, the following section provides a deeper analysis of the underlying mechanisms driving these differences, as well as the emerging trends shaping the evolution of Similar-Day approaches.

4.2. Analytical Insights and Emerging Trends

Beyond methodological categorization, the reviewed literature reveals a set of recurring patterns that explain not only how these approaches perform but also why certain strategies consistently outperform others under specific conditions.

In particular, the evolution of SD methods highlights the growing importance of similarity definition, data structure, and system dynamics over the mere selection of forecasting models. At the same time, emerging challenges such as structural changes in load behavior, increasing system complexity, and the need for uncertainty quantification are reshaping the way forecasting methodologies are designed and applied.

The following subsections synthesize these insights, identifying key trends and conceptual advances that characterize the current state and future direction of SD approaches in power system load forecasting.

4.2.1. SD Is Not Obsolete; It Is Being "Upgraded"

Despite the rapid growth of machine learning and deep learning techniques, the SD principle remains a fundamental component of operational load forecasting. Its persistence is explained by its intuitive nature and strong alignment with operator experience.

The continued relevance of SD can be attributed to the following:

(a) Its intuitive logic based on identifying historical days with similar conditions;
(b) Its strong alignment with operator expertise and calendar/weather heuristics;
(c) Its ability to provide transparent and interpretable forecasting frameworks.

However, what is evolving is not the SD principle itself, but the way in which similarity and forecasting are implemented. Traditional approaches based on simple averages or regression models are being progressively replaced by more advanced techniques, including machine learning and deep learning models such as LSTM, GRU, support vector machines, ensemble methods, and probabilistic forecasting frameworks.

This evolution is reflected in the increasing adoption of hybrid and intelligent SD approaches, indicating that SD is not being replaced, but rather integrated into modern data-driven forecasting frameworks.

4.2.2. Why the Similarity Definition Matters More than "Just a Better Model"

The taxonomy and reported percentages indicate that differences in Similar-Day approaches are largely driven by the choice of distance metrics and input variables used to define similarity. Even when powerful machine learning models are employed, their performance is constrained by the quality of the input data, as similarity definition effectively determines what the model can learn.

The performance results show that simply replacing classical models (e.g., ARIMA or regression) with ANN or LSTM yields limited improvements if similarity is not properly defined. In contrast, the largest gains are achieved when Similar-Day sets are carefully constructed using techniques such as clustering, weighted metrics, or grey relational analysis, and combined with ML/DL models trained on structurally consistent data.

This supports a two-stage design philosophy: first, define similarity through appropriate data selection and feature engineering; and second, apply a suitable forecasting model.

4.2.3. Structural Breaks: Where SD Is Weakest and Where Research Is Still Needed

Several structural disruptors where classical SD approaches fail can be identified:

- Social shifts: COVID-19, hybrid work, informal bridge holidays, etc.
- Climate extremes: heat waves, cold snaps, hurricanes, floods and, wildfires.
- New loads: data centers, crypto mining, EV charging and, behind-the-meter PV.

From a quantitative perspective, very few methods are specifically designed for anomalous or extreme conditions, and forecasting errors during such periods can increase significantly (e.g., MAPE values exceeding 40% during blackout days). Moreover, most SD definitions are calibrated on "normal" years and fail to capture emerging regimes, such as:

- Weekend peaks exceeding weekday peaks;
- Pronounced midday troughs due to PV generation;
- Sudden EV-charging surges.

These limitations reveal critical gaps:

1. Data scarcity for extreme events;
2. Nonstationary conditions driven by climate and behavioral changes;
3. Poor representation of new demand drivers (e.g., Behind-the-Meter PV, EVs, large industrial loads, market price signals).

Addressing these challenges requires approaches such as synthetic data generation, extreme-event–specific models, and the integration of XAI for anomaly diagnosis, which are aligned with current ML trends and are feasible for TSOs and ISOs.

4.2.4. The Rise of Clustering and Hybrid Architectures: Why This Makes Sense

The reported percentages indicate that clustering-oriented methods (e.g., K-means, K-medoids, SOM, and fuzzy C-means) represent a significant fraction of intelligent and hybrid SD approaches. In this context, clustering can be understood as a form of structural segmentation of the load time series based on season, day type, temperature regime, or load shape.

Once clustering is applied, several advantages emerge:

- Each cluster becomes more homogeneous, making it easier to model with simpler or specialized algorithms.
- Cluster-specific models can be designed to capture local behavior (e.g., high-temperature vs. mild conditions).
- Clusters can often be interpreted in physical terms (e.g., summer cooling regime, winter heating regime).

This approach helps reconcile the trade-off between accuracy and interpretability. On one hand, the Similar-Day concept remains intuitive—forecasting is based on historical days that exhibit similar behavior. On the other hand, machine learning models, which are often opaque, can be complemented by clustering and XAI techniques to provide insight into which cluster and variables drive the forecast.

In operational environments, this interpretability is critical. System operators are more likely to trust and adopt AI-based forecasting when the model can be explained in terms of recognizable conditions, such as a high-temperature weekday in summer with a characteristic peak demand pattern.

4.2.5. Deterministic vs. Probabilistic SD: The Next Frontier

The reviewed literature shows that most SD approaches are deterministic, focusing on point forecasts and reporting performance through metrics such as MAPE. However, a smaller subset of more recent studies has begun to explore probabilistic formulations, including quantile regression and interval forecasting based on deep learning models.

https://doi.org/10.3390/forecast8020032

In the current context of increasing uncertainty—driven by extreme events, renewable integration, and the growth of behind-the-meter generation—probabilistic forecasting becomes essential. In particular, it enables:

- Risk-aware unit commitment and reserve allocation.
- Better alignment with security-constrained economic dispatch under uncertainty.
- The provision of realistic confidence intervals instead of a single deterministic trajectory.

From a methodological perspective, Similar-Day approaches are naturally well-suited for probabilistic forecasting. Each similar day can be interpreted as a sample of possible system behavior under comparable conditions, allowing the direct estimation of:

- Load distributions or quantiles.
- Scenario sets for stochastic optimization.

Despite these advantages, the development of fully probabilistic SD frameworks remains limited compared to traditional point-forecasting approaches, representing a clear and promising direction for future research.

4.2.6. Interpretable and Explainable SD

Beyond the methodological advances shown in this work, another opportunity to strengthen similar day forecasting lies in integrating XAI. So far, SHapley Additive exPlanations (SHAP) have been employed to interpret a forecasting model by quantifying the contribution of each input variable for load forecasting [58]. However, it can also be used in extreme conditions—such as heat waves, sudden behavioral shifts, or atypical system events—that often disrupt the assumptions under which traditional SD models operate. Under these non-standard conditions the importance of input variables can shift dynamically, making model interpretation essential. Recent studies have shown that XAI tools can find changes in variable importance during anomalous periods [110]; thus, they could help diagnose when a forecasting model relies on unstable or non-representative patterns [111]. Techniques such as SHAP [112] provide transparent insight into why certain days are selected as similar and how their features influence the forecast. By improving interpretability under stress conditions, XAI can make SD-based short-term forecasting more robust and trustworthy, representing a promising and practical direction for future research.

In this context, XAI techniques can:

- Reveal changes in variable importance (e.g., humidity or wind-chill becoming more influential).
- Identify when models rely on unstable or outdated features.
- Distinguish between data quality issues and genuine regime shifts.

For SD approaches specifically, XAI could provide answers to key operational questions:

- Why were certain days selected as similar?
- Which variables are driving the similarity or clustering process?
- When large errors occur, which features contributed to the misprediction?

This level of transparency is particularly valuable in operational environments, as it transforms SD-based forecasting from a black-box system into a tool that can be interpreted, audited, and improved collaboratively by engineers and data scientists.

4.3. Practical Implications for Power System Operators

Bringing all of this together, the quantitative evidence and analysis suggest the following practical takeaways for control-center and TSO/ISO environments, where load forecasting is a critical operational process:

1. Keep SD approaches, but modernize them:
   SD methods should not be abandoned, but integrated into clustering, decomposi-

tion, and machine learning pipelines to enhance performance while preserving their intuitive foundation.

2. The first step is to invest in similarity definition:
   It is essential to identify the key variables that define similarity in the system—such as load, temperature, humidity, photovoltaic generation, electric vehicle penetration, and calendar effects—and refine SD selection using clustering, grey relational analysis, or correlation-based methods.
3. Use hybrid architectures as a default strategy:
   Combining SD selection with signal decomposition techniques (e.g., wavelets, EMD, VMD), machine learning or deep learning models, and residual correction improves both accuracy and robustness.
4. Design explicit strategies for extreme events:
   Dedicated SD models for anomalous conditions and the use of synthetic data or scenario generation are necessary to address data scarcity and improve performance under rare but critical events.
5. Move toward probabilistic SD-based forecasting:
   The use of similar-day ensembles, quantile methods, and interval forecasts enables more robust decision-making under uncertainty.
6. Adopt XAI for trust and diagnosis:
   Integrating explainable AI techniques, such as SHAP, allows operators to interpret feature contributions, understand clustering behavior, and diagnose forecasting errors.

5. Conclusions

This review synthesized more than two decades of Similar Day-based short-term load forecasting research and proposed a unified taxonomy that organizes existing methods according to the similarity-definition mechanism, the forecasting model, and their integration strategy. This framework enables a clearer understanding of the evolution of SD approaches, from rule-based formulations to intelligent and hybrid architectures.

The analysis confirms that, despite the increasing adoption of artificial intelligence techniques, SD-based methods remain widely used due to their intuitive nature, low implementation cost, and strong connection to operational knowledge. However, their traditional formulations are increasingly challenged by structural changes in power systems, including social shifts (e.g., post-COVID behavioral patterns and hybrid work), climate-driven extreme events, the emergence of new large and distributed loads, and the growing impact of behind-the-meter generation. These factors introduce nonstationarity and increased variability that cannot be fully captured by conventional similarity rules.

In this context, the evidence indicates that the future of SD-based forecasting lies not in its replacement, but in its transformation. Enhancing similarity definitions through richer explanatory variables and adaptive strategies, together with the integration of clustering, machine learning, and hybrid modeling approaches, enables more accurate and flexible forecasting frameworks. In particular, hybrid architectures provide a practical balance between interpretability and accuracy, while probabilistic formulations offer a pathway to address uncertainty under increasingly complex operating conditions.

Finally, this work identifies key directions for future research:

- Systematic evaluation of SD methods under extreme conditions and other atypical operating conditions.
- Standardization of similarity metrics and variable sets to facilitate method comparison and reproducibility across systems and regions.
- Integration of high-resolution meteorological forecasts and climate information into SD selection and forecasting models.

- Improved representation of emerging demand drivers,
- Broader adoption of probabilistic and explainable AI techniques to quantify uncertainty and improve model transparency, particularly under stress scenarios.

These developments are essential to ensure reliable, transparent, and robust forecasting in modern power systems.

Author Contributions: Conceptualization, M.B.; methodology, M.B. and L.C.-L.; software, G.R.-C. and V.M.A.; validation, G.L.L. and E.d.J.C.A.; formal analysis, M.B.; investigation, M.B. and L.C.-L.; resources, G.R.-C., G.L.L. and E.d.J.C.A.; data curation, G.R.-C. and E.d.J.C.A.; writing—original draft preparation, M.B. and L.C.-L.; writing—review and editing, M.B. and L.C.-L.; visualization, M.B.; supervision, M.B.; project administration, M.B.; funding acquisition, M.B., L.C.-L., G.R.-C., G.L.L., V.M.A. and E.d.J.C.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data used in this work is available online.

Acknowledgments: M.B. thanks SECIHTI for her "Investigadoras e Investigadores por México" research position with I.D. 71557 and CENIDET-TECNM for its hospitality and support. E.J.C.A. acknowledges the support provided by SECIHTI during his postdoctoral period. He also acknowledges the support of CENIDET for the development of this work. The authors thank Raul Garduno for his invaluable support.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ANNs	Artificial Neural Networks
ARIMA	Auto Regressive Integrated Moving Average
BTM	Behind-the-Meter (Generation)
DL	Deep Learning
EMD	Empirical Mode Decomposition
GBoost	Gradient Boosting
GRU	Gated Recurrent Unit
ISO	Independent System Operator
LSTM	Long Short-Term Memory
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MLOps	Machine Learning Operations
NSDAR	Neural-based Similar Day Auto Regression
NERC	North American Electric Reliability Council
NN	Neural Network
PSO	Particle Swarm Optimization
PV	Photovoltaic
RBF	Radial Basis Function
RF	Random Forest
SDs	Similar Days
SHAP	SHapley Additive exPlanations
SOM	Self-Organizing Map
SVR	Support Vector Regression
SVM	Support Vector Machine
STLF	Short-Term Load Forecast
TSO	Transmission System Operator
VMD	Variational Mode Decomposition

XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

References

1. Fragkos, P.; Siskos, P. Energy systems analysis and modeling towards decarbonization. *Energies* **2022**, *15*, 1971. [CrossRef]
2. Russo, M.A.; Carvalho, D.; Martins, N.; Monteiro, A. Forecasting the inevitable: A review on the impacts of climate change on renewable energy resources. *Sustain. Energy Technol. Assess.* **2022**, *52*, 102283. [CrossRef]
3. Cengiz, E.; Yilmaz, C.; Yilmaz, E.N.; Tolga Kahraman, H. Importance of power flow and load analysis in pre-installation power systems. In *Proceedings of the 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 11–13 October 2019*; IEEE: Piscataway, NJ, USA, 2019; pp. 1–4.
4. Soliman, S.A.H.; Al-Kandari, A.M. *Electrical Load Forecasting: Modeling and Model Construction*; Elsevier: Amsterdam, The Netherlands, 2010.
5. Hahn, H.; Meyer-Nieberg, S.; Pickl, S. Electric load forecasting methods: Tools for decision making. *Eur. J. Oper. Res.* **2009**, *199*, 902–907. [CrossRef]
6. Iskandarnia, E.; Al-Amal, H.; Al-Madany, W. Load forecasting in different scale and horizon—A review. In Proceedings of the 3rd Smart Cities Symposium (SCS 2020), Online, 21–23 September 2020; pp. 464–470.
7. Kuster, C.; Rezgui, Y.; Mourshed, M. Electrical load forecasting models: A critical systematic review. *Sustain. Cities Soc.* **2017**, *35*, 257–270. [CrossRef]
8. Akhtar, S.; Shahzad, S.; Zaheer, A.; Ullah, H.S.; Killic, H.; Gono, R.; Jasinski, M.; Leonowicz, Z. Short-Term Load Forecasting Models: A Review of Challenges, Progress, and the Road Ahead. *Energies* **2023**, *16*, 4060. [CrossRef]
9. Nti, I.K.; Teimeh, M.; Nyarko-Boateng, O.; Adekoya, A.F. Electricity load forecasting: A systematic review. *J. Electr. Syst. Inf. Technol.* **2020**, *7*, 13. [CrossRef]
10. Mamun, A.A.; Sohel, M.; Mohammad, N.; Haque Sunny, M.S.M.; Dipta, D.R.; Hossain, E. A comprehensive review of the load forecasting techniques using single and hybrid predictive models. *IEEE Access* **2020**, *8*, 134911–134939. [CrossRef]
11. Abu-El-Magd, M.; Shina, N. Short-term load demand modeling and forecasting: A Review. *IEEE Trans. Syst. Man Cybern.* **1982**, *12*, 370–382. [CrossRef]
12. Hong, T. Short Term Electric Load Forecasting. Ph.D. Dissertation, North Carolina State University, Raleigh, NC, USA, 2010.
13. Dudek, G. Pattern similarity-based methods for short-term load forecasting—Part 1: Principles. *Appl. Soft Comput.* **2015**, *37*, 277–287. [CrossRef]
14. Dudek, G. Pattern similarity-based methods for short-term load forecasting—Part 2: Models. *Appl. Soft Comput.* **2015**, *36*, 422–441. [CrossRef]
15. Fallah, S.N.; Ganjkhani, M.; Shamshirband, S.; Chau, K.W. Computational intelligence on short-term load forecasting: A methodological overview. *Energies* **2019**, *12*, 393. [CrossRef]
16. L'Heureux, A.; Grolinger, K.; Capretz, M.A.M. Transformer-based model for electrical load forecasting. *Energies* **2022**, *15*, 4993. [CrossRef]
17. Bouhamed, O.; Dissem, M.; Amayri, M.; Bouguila, N. Transformer-based deep probabilistic network for load forecasting. *Eng. Appl. Artif. Intell.* **2025**, *152*, 110781. [CrossRef]
18. Wu, M.; Feng, W.; Liu, Y.; Cao, C. Short-term power load forecasting using an improved model integrating GCN and transformer. *Appl. Sci.* **2025**, *15*, 7003. [CrossRef]
19. Staffell, I.; Pfenninger, S. The increasing impact of weather on electricity supply and demand. *Energy* **2018**, *145*, 65–78. [CrossRef]
20. Zhang, T.; Zhang, G.; Lu, J.; Feng, X.; Yang, W. A new index and classification approach for load pattern analysis of large electricity customers. *IEEE Trans. Power Syst.* **2012**, *27*, 153–160. [CrossRef]
21. Mori, H.; Yuihara, A. Deterministic annealing clustering for ANN-based Short-Term load forecasting. *IEEE Trans. Power Syst.* **2001**, *16*, 541–551.
22. Voronin, S.; Partanen, J.; Kauranne, T. A hybrid electricity price forecasting model for the Nordic electricity spot market. *Int. Trans. Electr. Energy Syst.* **2014**, *24*, 736–760. [CrossRef]
23. Nozawa, T.; Konda, M.; Fujibayashi, M.; Imai, M.; Kotani, K.; Sugawa, S.; Ohmi, T. A parallel vector-quantization processor eliminating redundant calculations for real-time motion picture compression. *IEEE J. Solid-State Circuits* **2000**, *35*, 1744–1751. [CrossRef]
24. Lee, C.-Y.; Wu, C.-E. Short-term electricity price forecasting based on similar day-based neural network. *Energies* **2020**, *13*, 4408.
25. Jain, A.; Jain, M.B. Fuzzy modelling and similarity based short term load forecasting using swarm intelligence—A step towards smart grid. In *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012)*; Advances in Intelligent Systems and Computing; Bansal, J., Singh, P., Deep, K., Pant, M., Nagar, A., Eds.; Springer: New Delhi, India, 2013; Volume 202.

26. Tzortzis, A.M.; Pelekis, S.; Spiliotis, E.; Karakolis, E.; Mouzakitis, S.; Psarras, J.; Askounis, D. Transfer learning for day-ahead load forecasting: A case study on European National Electricity demand time series. *Mathematics* **2024**, *12*, 19. [CrossRef]
27. Klyuev, R.V.; Morgoev, I.D.; Morgoeva, A.D.; Gavrina, O.A.; Martyushev, N.V.; Efremenkov, E.A.; Mengxu, A. Methods of forecasting electric energy consumption: A literature review. *Energies* **2022**, *15*, 8919. [CrossRef]
28. Rahman, S.; Bhatnagar, R. An expert system based algorithm for short term load forecast. *IEEE Trans. Power Syst.* **1988**, *3*, 392–399. [CrossRef]
29. Ruzic, S.; Vuckovic, A.; Nikolic, N. Weather sensitive method for short term load forecasting in Electric Power Utility of Serbia. *IEEE Trans. Power Syst.* **2003**, *18*, 1581–1586. [CrossRef]
30. Mu, Q.; Wu, Y.; Pan, X.; Huang, L.; Li, X. Short-term load forecasting using improved similar days method. In Proceedings of the 2010 Asia-Pacific Power and Energy Engineering Conference, Chengdu, China, 28–31 March 2010; pp. 1–4.
31. Jin, X. Meteorological factors considered load decoupling forecasting techniques. *Adv. Mater. Res.* **2011**, *354–355*, 922–926. [CrossRef]
32. Jian-Kai, L.; Cattani, C.; Wan-Qing, S. Power load prediction based on fractal theory. *Adv. Math. Phys.* **2015**, *2015*, 827238. [CrossRef]
33. Karimi, M.; Karami, H.; Gholami, M.; Khatibzadehazad, H.; Moslemi, N. Priority index considering temperature and date proximity for selection of similar days in knowledge-based short term load forecasting method. *Energy* **2018**, *144*, 928–940. [CrossRef]
34. Son, J.; Cha, J.; Kim, H.; Wi, Y. Day-ahead short-term load forecasting for holidays based on modification of similar days' load profiles. *IEEE Access* **2022**, *10*, 17864–17880. [CrossRef]
35. Conde-López, L.; Borunda, M.; Ruiz-Chavarría, G.; Aparicio-Cárdenas, T. A highly accurate and efficient statistical framework for short-term load forecasting: A case study for Mexico. *Forecasting* **2026**, *8*, 3. [CrossRef]
36. Senjyu, T.; Takara, H.; Uezato, K.; Funabashi, T. One-hour-ahead load forecasting using Neural Networks. *IEEE Trans. Power Syst.* **2002**, *17*, 113–118. [CrossRef]
37. Mandal, P.; Senjyu, T.; Uezato, E.; Funabashi, T. Forecasting several-hours-ahead electricity demand using neural network. In Proceedings of the 2004 IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies, Hong Kong, China, 5–8 April 2004; pp. 515–521.
38. He, Y.J.; Zhu, Y.C.; Gu, J.C.; Yin, C.Q. Similar day selection based neural network model and its application in short-term load forecasting. In Proceedings of the International Conference on Machine Learning and Cybernetics, Guangzhou, China, 18–21 August 2005; Volume 8, pp. 4760–4763.
39. Mandal, P.; Senjyu, T.; Urasaki, N.; Funabashi, T. A neural network based several-hour-ahead electric load forecasting using similar days approach. *Int. J. Electr. Power Energy Syst.* **2006**, *28*, 367–373. [CrossRef]
40. Mandal, P.; Senjyu, T.; Funabashi, T. Neural networks approach to forecast several hour ahead in electricity prices and loads in deregulated market. *Energy Convers. Manag.* **2006**, *47*, 2128–2142. [CrossRef]
41. Chang-Chun, C.; Min, W. Support vector machines with similar day's training sample application in short-term load forecasting. In Proceedings of the 2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, Nanjing, China, 6–9 April 2008; pp. 1221–1225.
42. Sun, C.; Song, J.; Li, L.; Ju, P. Implementation of hybrid short-term load forecasting system with analysis of temperature sensitivities. *Soft Comput.* **2008**, *12*, 633–638. [CrossRef]
43. Chen, Y.; Luh, P.B.; Guan, C.; Zhao, Y.; Michel, L.D.; Coolbeth, M.A.; Friedland, P.B.; Rourke, S.J. Short-term load forecasting: Similar day-based wavelet neural networks. *IEEE Trans. Power Syst.* **2010**, *25*, 322–330. [CrossRef]
44. Jain, A.; Jain, B. Fuzzy modeling and similarity based short-term load forecasting using swarm intelligence-a-step towards smart grid. In Proceedings of the 2012 IEEE Power and Energy Society General Meeting, San Diego, CA, USA, 22–26 July 2012; p. 1.
45. Ebrahimi, A.; Moshari, A. Holidays short-term forecasting using fuzzy improved similar day method. *Int. Trans. Electr. Energy Syst.* **2013**, *23*, 1254–1271. [CrossRef]
46. Eapen, R.R.; Simon, S.P. Performance analysis of combined similar day and day ahead short-term electrical load forecasting using sequential hybrid neural networks. *IETE J. Res.* **2018**, *65*, 216–226. [CrossRef]
47. Song, K.B.; Park, J.D.; Park, R.J. Short term load forecasting algorithm for Lunar New Year's Day. *J. Electr. Eng. Technol.* **2018**, *13*, 591–598.
48. Barman, M.; Choudhury, N.B.D. A similarity based hybrid GWO-SVM method of power system load forecasting for regional special event days in anomalous load situations in Assam, India. *Sustain. Cities Soc.* **2020**, *61*, 102311. [CrossRef]
49. Gundu, V.; Simon, S.P. Quality analysis of combined similar day and day ahead short-term load forecasting using recurrent neural networks. *Sadhana* **2022**, *47*, 43. [CrossRef]
50. Xiang, M.; Du, S.; Yu, J.; Xu, Y. Short-term load forecasting based on sample weights assignment. *Energy Rep.* **2022**, *8*, 783–791. [CrossRef]

51. Zhang, R.; Zhang, C.; Yu, M. A similar day based short term load forecasting method using wavelet transform and LSTM. *IEEJ Trans. Electr. Electron. Eng.* **2022**, *17*, 506–513. [CrossRef]
52. Wei, S.; He, C.; Zhang, J.; Sun, F.; Jiang, X. Short-term load forecast based on similar daily retrieval and deep learning. In Proceedings of the IEEE IAS Industrial and Commercial Power system Asia (I&CPS Asia), Chongqing, China, 7–9 July 2023.
53. Ortega, A.; Borunda, M.; Conde, L.; Garcia-Beltran, C. Load demand forecasting using a long-short term memory neural network. In *Advances in Computational Intelligence, MICAI 2023*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2023; Volume 14391.
54. Jiao, M.; Qi, X.; Mneg, F.; Wang, T.; Hou, Q.; Zheng, X.; Zhou, H. LSTM-RBF short-term load forecasting method based on Gaussian similar days. In Proceedings of the 5th Asia Energy and Electrical Engineering Symposium, Chengdu, China, 23–26 March 2023; pp. 1189–1194. [CrossRef]
55. Linzi, Z. Short-term power load forecasting based on FA-LSTM with Similar Day Selection. In Proceedings of the IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI), Changchun, China, 26–28 May 2023; pp. 1110–1115. [CrossRef]
56. Aswanuwath, L.; Pannakkong, W.; Buddhakulsomsiri, J.; Karnjana, J.; Huynh, V.N. A hybrid model of VMD-EMD-FFT, Similar Days Selection Method, Stepwise Regression and Artificial Neural Network for Daily Electricity Peak Load Forecasting. *Energies* **2023**, *16*, 1860. [CrossRef]
57. Aswanuwath, L.; Pannakkong, W.; Buddhakulsomsiri, J.; Karnjana, J.; Huynh, V.N. An improved hybrid approach for daily electricity peak demand forecasting during disrupted situations: A case study of COVID-19 impact in Thailand. *Energies* **2024**, *17*, 78. [CrossRef]
58. Wang, Z.; Xu, M.; Hao, J.; Li, W.; Cheng, S. Research on explainable BiGRU deep learning framework for short term load forecasting in smart power systems. *Discov. Artif. Intell.* **2025**, *5*, 407. [CrossRef]
59. Martínez-Álvarez, F.; Troncoso, A.; Riquelme, J.C.; Aguilar-Ruiz, J.S. Energy time series forecasting based on pattern sequence similarity. *IEEE Trans. Knowl. Data Eng.* **2011**, *23*, 1230–1243. [CrossRef]
60. Feng, Y.; Ryan, S.M. Day-ahead hourly electricity load modeling by functional regression. *Appl. Energy* **2016**, *170*, 455–465. [CrossRef]
61. Martínez-Álvarez, F.; Schmutz, A.; Asencio-Cortés, G.; Jacques, J. A novel hybrid algorithm to forecast functional time series based on patterns sequence similarity with application to electricity demand. *Energies* **2019**, *12*, 94. [CrossRef]
62. Pérez-Chacón, R.; Asencio-Cortés, G.; Martínez-Álvarez, F.; Troncoso, A. Big data time series forecasting based on pattern sequence similarity and its application to the electricity demand. *Inf. Sci.* **2020**, *540*, 160–174. [CrossRef]
63. Jiang, Z.; Zhang, L.; Ji, T. NSDAR: A neural network-based model for similar day screening and electric load forecasting. *Appl. Energy* **2023**, *349*, 121647. [CrossRef]
64. Wang, Z.Y. Development case-based reasoning system for short-term load forecasting. In Proceedings of the 2006 Russia Power Engineering Society General Meeting, Montreal, QC, Canada, 18–22 June 2006; pp. 1–6.
65. Guo, Y.C.; Niu, D.X. Intelligent short-term forecasting based on pattern-base. In Proceedings of the 2008 International Conference on Machine Learning and Cybernetics, Kunming, China, 12–15 July 2008; pp. 1282–1287.
66. Guo, Y.C.; Niu, D.X. Short-term load forecasting using support vector regression based on pattern-base. In Proceedings of the 2009 First Asian Conference on Intelligent Information and Database Systems, Dong Hoi, Vietnam, 1–3 April 2009; pp. 336–340.
67. Duan, P.; Xie, K.; Guo, T.; Huang, X. Short-term load forecasting for electric power systems using the PSO-SVR and FCM clustering techniques. *Energies* **2011**, *4*, 173–184. [CrossRef]
68. Martins, A.; Lagarto, J.; Canacsinh, H.; Reis, F.; Cardoso, M.G.M.S. Short-term load forecasting using time series clustering. *Optim. Eng.* **2022**, *23*, 2293–2314. [CrossRef]
69. Janković, Z.; Vesin, B.; Selakov, A.; Berntzen, L. Gab-SSDS: An AI-Based Similar Days Selection Method for Load Forecast. *Front. Energy Res.* **2022**, *10*, 844838. [CrossRef]
70. Chen, Y.; Luh, P.B.; Rourke, S.J. Short-term load forecasting: Similar day-based wavelet neural networks. In Proceedings of the 2008 7th World Congress on Intelligent Control and Automation, Chongqing, China, 25–27 June 2008; pp. 3353–3358.
71. Lu, H.; Liu, K. Next day load forecasting using local temperature sensitive information. *Key Eng. Mater.* **2011**, *460–461*, 786–791. [CrossRef]
72. Al-Qahtani, F.H.; Crone, S.F. Multivariate k-nearest neighbor regression for time series data-A novel algorithm for forecasting UK electricity demand. In Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN), Dallas, TX, USA, 4–9 August 2013; pp. 1–8.
73. Moazzami, M.; Khodabakhshian, A.; Hooshmand, R. A new hybrid day-ahead peak load forecasting method for Iran's National Grid. *Appl. Energy* **2013**, *101*, 489–501. [CrossRef]
74. Zheng, H.; Yuan, J.; Chen, L. Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation. *Energies* **2017**, *10*, 1168. [CrossRef]

75. Sun, W.; Zhang, C. A hybrid BA-ELM model based on factor analysis and similar-day approach for short-term load forecasting. *Energies* **2018**, *11*, 1282.
76. Barman, M.; Choudhury, N.B.D.; Sutradhar, S. A regional hybrid GOA-SVM model based on similar day approach for short-term load forecasting in Assam, India. *Energy* **2018**, *145*, 710–720.
77. Liu, Q.; Shen, Y.; Wu, L. A hybrid FCW-EMD and KF-BA-SVM based model for short-term load forecasting. *CSEE J. Power Energy Syst.* **2018**, *4*, 226–237. [CrossRef]
78. Nazar, M.S.; Fard, A.E.; Heidari, A.; Shafie-khah, M.; Catalao, J.P.S. Hybrid model using three-stage algorithm for simultaneous load and price forecasting. *Electr. Power Syst. Res.* **2018**, *165*, 214–228. [CrossRef]
79. Liao, X.; Cao, N.; Li, M.; Kang, X. Research on short-term load forecasting using XGBoost based on similar days. In Proceedings of the International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), Changsha, China, 12–13 January 2019; pp. 675–678. [CrossRef]
80. Sun, G.; Cheng, P.; Liu, R. A hybrid short-term load prediction with input selection for the distribution system. In Proceedings of the IEEE 2nd International Conference on Electronic Information and Communication Technology (ICEICT), Harbin, China, 20–22 January 2019; pp. 826–830.
81. Barman, M.; Choudhury, N.B.D. Season specific approach for short-term load forecasting based on hybrid FA-SVM and similarity concept. *Energy* **2019**, *174*, 886–896. [CrossRef]
82. Huang, W.; Zhang, D.; Song, Z.; Wang, H.; Liu, H. Short-term load forecasting based on similar day approach and intelligent algorithm using analytic hierarchy process. In Proceedings of the IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chengdu, China, 15–17 March 2019; pp. 2507–2512.
83. Kong, Z.; Zhang, C.; Lv, H.; Xiong, F.; Fu, Z. Multimodal feature extraction and fusion deep neural network for short-term load forecasting. *IEEE Access* **2020**, *8*, 185373–185383. [CrossRef]
84. Park, R.J.; Song, K.B.; Kwon, B.S. Short-term load forecasting algorithm using a similar day selection method based on reinforcement learning. *Energies* **2020**, *13*, 2640. [CrossRef]
85. Lu, Z.; Zheng, Z.; Wang, H.; Gao, Y.; Lei, Y.; Zhao, M. Study on short-term load forecasting considering meteorological similar days. In *Advances in Intelligent Systems and Interactive Applications, ISSA 2019*; Advances in Intelligent Systems and Computing; Xhafa, F., Patnaik, S., Tavana, M., Eds.; Springer: Cham, Switzerland, 2020; Volume 1084.
86. Gunawan, F.; Huang, C.Y. An extensible framework for short-term holiday load forecasting combining dynamic time warping and LSTM network. *IEEE Access* **2021**, *9*, 106885–106894. [CrossRef]
87. Wang, R.; Chen, S.; Lu, J. Electric short-term load forecast integrated method based on time-segment and improved MDSC-BP. *Syst. Sci. Control Eng.* **2021**, *9*, 80–86. [CrossRef]
88. Dudek, G. Short-term load forecasting using neural networks with pattern similarity-based error weights. *Energies* **2021**, *14*, 3224. [CrossRef]
89. Janković, Z.; Selakov, A.; Bekut, D.; Đorđević, M. Day similarity metric model for short-term load forecasting supported by PSO and artificial neural network. *Electr. Eng.* **2021**, *103*, 2971–2988. [CrossRef]
90. Zhao, Z.; Xia, C.; Chi, L.; Chang, X.; Li, W.; Yang, T.; Zomaya, A.Y. Short-term load forecasting based on the transformer model. *Information* **2021**, *12*, 516. [CrossRef]
91. Xiong, J.; Zhou, P.; Chen, A.; Zhang, Y. Attention-based neural load forecasting: A dynamic feature selection approach. In Proceedings of the 2021 IEEE Power & Energy Society General Meeting (PESGM), Washington, DC, USA, 26–29 July 2021; pp. 1–5.
92. Dong, X.; Deng, S.; Wang, D. A short-term power load forecasting method based on k-means and SVM. *J. Ambient. Intell. Humaniz. Comput.* **2022**, *13*, 5253–5267. [CrossRef]
93. Zhao, Z.; Tang, J.; Liu, J.; Ge, G.; Xiong, B.; Li, Y. Short-term microgrid load probability density forecasting method based on k-means deep learning quantile regression. *Energy Rep.* **2022**, *8*, 1386–1397. [CrossRef]
94. Pang, X.; Sun, W.; Li, H.; Wang, Y.; Luan, C. Short-term power load forecasting based on gray relational analysis and support vector machine optimized by artificial bee colony algorithm. *PeerJ Comput. Sci.* **2022**, *8*, e1108. [CrossRef]
95. Shi, M.; He, X. Shape-aided similar day selection method for short term load forecasting. In *Proceedings of 2021 Chinese Intelligent Automation Conference*; Lecture Notes in Electrical Engineering; Den, Z., Ed.; Springer: Singapore, 2022; Volume 801, pp. 195–202.
96. Yang, J.; Tang, Y.; Duan, H. Application of fuzzy support vector machine in short-term power load forecasting. *J. Cases Inf. Technol.* **2022**, *24*, 5. [CrossRef]
97. Wang, X.; Wu, B.; Wu, D.; Wang, W.; Ma, X. Enhanced ES-adRNNe load forecasting with contextual augmentation on similar load days. *IEEE Access* **2023**, *11*, 93727–93738. [CrossRef]
98. Pajić, Z.; Selakov, A. A flexible approach for selection of the training set for ANN-based load forecast using autoencoder and similar day method. In Proceedings of the 2023 IEEE Belgrade Power Tech, Belgrade, Serbia, 25–29 June 2023; pp. 1–5.
99. Criado-Ramón, D.; Ruiz, L.G.B.; Pegalajar, M.C. An improved patter sequence-based energy load forecast algorithm based on Self-Organizing Maps and artificial neural networks. *Big Data Cogn. Comput.* **2023**, *7*, 92. [CrossRef]

100. Bu, X.; Wu, Q.; Zhou, B.; Li, C. Hybrid short-term load forecasting using CGAN with CNN and semi-supervised regression. *Appl. Energy* **2023**, *338*, 120920. [CrossRef]
101. Xiong, J.; Zhang, T. A unifying framework of attention-based neural load forecasting. *IEEE Access* **2023**, *11*, 51606–51616. [CrossRef]
102. Luo, S.; Wang, B.; Gao, Q.; Wang, Y.; Pang, X. Stacking integration algorithm based on CNN-BiLSTM-Attention with XGBoost for short-term electricity load forecasting. *Energy Rep.* **2024**, *12*, 2676–2689. [CrossRef]
103. Amini, H.; Mehrizi-Sani, A. Similar days fuzzy clustering load forecasting. In Proceedings of the IECON 2024—50th Annual Conference of the IEEE Industrial Electronics Society, Chicago, IL, USA, 3–6 November 2024; pp. 1–5.
104. Wang, Y.; Hao, Y.; Zhang, B.; Zhang, N. Short-term power load forecasting using SSA-CNN-LSTM method. *Syst. Sci. Control Eng. Open Access J.* **2024**, *12*, 2343297. [CrossRef]
105. Borunda, M.; Ortega Vega, A.; Garduno, R.; Conde, L.; Adam Medina, M.; Ramírez Aparicio, J.; Magallón-Cacho, L.; Jaramillo, O.A. An intelligent method for day-ahead regional load demand forecasting via machine-learning analysis of energy consumption patterns across daily, weekly, and annual scales. *Appl. Sci.* **2025**, *15*, 4717. [CrossRef]
106. Luo, S.; Meng, X.; Pang, X.; Li, H.; Zheng, Z. Optimized hybrid deep learning framework for short-term power load interval forecasting via improved crowned crested porcupine optimization and feature mode decomposition. *Algorithms* **2025**, *18*, 659. [CrossRef]
107. Cheng, Q.; Shi, J.; Cheng, S. Short-term load forecasting based on similar day theory and BWO-VMD. *Energies* **2025**, *18*, 2358. [CrossRef]
108. Zhao, W.; Xu, H.; Chen, P.; Zhang, J.; Li, J.; Cai, T. Elastic momentum-enhanced adaptive hybrid method for short-term load forecasting. *Energies* **2025**, *18*, 3263. [CrossRef]
109. Ali, S.; Abuhmed, T.; El-Sappagh, S.; Muhammad, K.; Alonso-Moral, J.M.; Confalonieri, R.; Guidotti, R.; Del Ser, J.; Díaz-Rodríguez, N.; Herrera, F. Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. *Inf. Fusion* **2023**, *99*, 101805. [CrossRef]
110. Alam, K.; Kifayat, K.; Sampedro, G.A.; Karovic, V.; Naeem, T. SXAD: Shapely eXplainable AI-based anomaly detection using log data. *IEEE Access* **2024**, *12*, 95659–95672. [CrossRef]
111. Xiao, Y.; Hu, X.; Lin, Y.; Lu, Y.; Jing, R.; Zhao, Y. Interpretable short-term electricity load forecasting considering small sample heatwaves. *Appl. Energy* **2025**, *398*, 126417. [CrossRef]
112. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In Proceedings of the NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 4768–47777.