

Article

Machine Learning-Based Extraction Method for Marine Load Cycles with Environmentally Sustainable Applications

Xiaojun Sun ^{1,*} , Yingbo Gao ¹, Qiao Zhang ¹ and Shunliang Ding ^{2,*} 

¹ School of Automobile and Traffic Engineering, Liaoning University of Technology, Jinzhou 121001, China; gaoyingbo0418@163.com (Y.G.); zq_625@163.com (Q.Z.)

² School of Mechanical and Power Engineering, Zhengzhou University, Zhengzhou 450001, China

* Correspondence: 18245904309@163.com (X.S.); dingshunliang@126.com (S.D.)

Abstract: The current lack of harmonized standard test conditions for marine shipping hinders the comparison of performance and compliance assessments for different types of ships. This article puts forward a method for extracting ship loading cycles using machine learning algorithms. Time-series data are extracted from real ships in operation, and a segmented linear approximation method and a data normalization technique are adopted. A hierarchical-clustering type of soft dynamic time-warping similarity analysis method is presented to efficiently analyze the similarity of different time-series data, using soft dynamic time warping (Soft-DTW) combined with hierarchical clustering algorithms from the field of machine learning. The problem of data bias caused by spatial and temporal offset characteristics is effectively solved in marine test condition data. The validity and reliability of the proposed method are validated through the analysis of case data. The results demonstrate that the hierarchically clustered soft dynamic time-warping similarity analysis method can be considered reliable for obtaining test cases with different characteristics. Furthermore, it provides input conditions for effectively identifying the operating conditions of different types of ships with high levels of energy consumption and high emissions, thus allowing for the establishment of energy-saving and emissions-reducing sailing strategies.

Keywords: time-series data; similarity; hierarchical clustering; loading cycles



Citation: Sun, X.; Gao, Y.; Zhang, Q.; Ding, S. Machine Learning-Based Extraction Method for Marine Load Cycles with Environmentally Sustainable Applications. *Sustainability* **2024**, *16*, 4840. <https://doi.org/10.3390/su16114840>

Academic Editor: Tingsong Wang

Received: 29 April 2024

Revised: 27 May 2024

Accepted: 31 May 2024

Published: 6 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Marine power plays an indispensable role in global maritime trade, logistics, and transportation [1,2]. For the first time, Michail et al. [3] explored the vital roles of different types of ships, such as dry bulk carriers and clean oil tankers, in improving economic indicators through the use of logistics and transportation service cost variables obtained from real ship operating conditions. However, a lack of standardized test conditions for ships makes it difficult to reasonably compare the quality characteristics of different types of ships in terms of improved power, economy, and emissions. On this basis, the focus of this study is to propose a marine loading cycle method that can effectively extract the operating conditions of real ships. Targeted marine cycling conditions are expected to assist in clarifying the performance differences between different ship types, providing a better understanding of ship performance and leading to ship power system upgrades and control strategy enhancements.

The shipping industry is affected by various MARPOL-like regulations, forcing the modernization of ship powertrains, especially diesel–electric hybrids [4,5], gas–electric hybrids [6,7], and energy storage hybrids [8,9]. The size of marine system components and the energy density of the power source affect energy efficiency improvements and power gains during ship operations, and a good management system can balance a ship's performance. As such, scholars have mainly focused on the selection, matching, and development of energy management strategies for hybrid powertrains.

In terms of ship power matching and selection, Wu et al. [10] proposed a component selection method based on multi-exponential evaluation and single-objective particle swarm optimization (PSO), which was applied to a marine hybrid power system consisting of a reciprocating gas engine and an energy storage system. The improvements in energy efficiency and economy after selection were tested using a 40,000 s power demand condition. The results indicated that fuel consumption and carbon emissions can be reduced by 17.1% and 14.8%, respectively, and that the proposed method improves the system's power performance. Wang et al. [11] established a scalable mathematical model of a diesel engine/battery/shore power hybrid propulsion system. Meanwhile, a three-objective collaborative optimization method containing key equipment parameters was designed. The reliability of the mathematical model and the effectiveness of the optimization method were verified using the real speed curves of a typical ship, including cruising and agitation modes. The results showed that the diesel engine, battery, and gear ratio parameters have a large influence on the performance of the hybrid propulsion system. Li et al. [12] introduced chaotic thinking and an adaptive crossover operator to improve the selection of parameters to optimize the capacity of energy storage equipment in electric ships. The method for optimizing energy storage equipment capacity was verified using the ship's total electric load, service load, and propulsion load, as well as other real engine working condition parameters. Haseltalab et al. [13] introduced solid oxide fuel cells into hybrid power systems for the first time and proposed a component selection method to determine the optimal power ratings of the gas engine, battery, and fuel cell with full consideration of the size and weight parameters of the power source. The results of their research demonstrated that the matched powertrain improves fuel economy by 21% and reduces CO₂ emissions by 53% as compared to a conventional diesel–electric-powered vessel. Chen et al. [14] put forward a novel energy management strategy incorporating a support vector machine and frequency control for a marine fuel cell hybrid energy storage system, which mainly focuses on the hybrid power structure and equipment capacity matching. The feasibility of the strategy was validated using the load profiles of a 360 s fuel cell vessel in constant and maneuvering states. The simulation results indicate that the optimal hybrid structure can save 5.4% of energy consumption while extending the equipment's lifetime.

Regarding energy management strategy research, Zhang et al. [15] proposed a two-stage predictive energy management method for a diesel–electric hybrid powertrain. A loaded power demand of 9000s was used to simulate the economic aspects of the energy management method, in order to improve its economy, and the results demonstrated that its performance is close to that obtained through global optimization. Ghimire et al. [16] presented a novel emissions assessment system for drillships that can be used in marine DC hybrid power systems. The reliability and validity of the proposed emissions assessment system was verified using 120 h of drilling power in a system consisting of transit, drilling, and reaming. Zhao et al. [17] integrated the working characteristics of a tourist ship to propose an energy allocation strategy based on an improved fuzzy strategy, in which the main devices contain solar energy, fuel cells, supercapacitors, and batteries. The proposed energy distribution strategy was validated using 1200 s of load power working conditions comprising docking and cruising states. The outcomes revealed that the voltage fluctuation of the battery and fuel cell can be suppressed, and the operating time of the battery can be extended. Xie et al. [18] presented an efficient and environmentally friendly solution for a hybrid ferry with a combination of hydrogen fuel cells and batteries and designed a two-tier energy management system to enhance the economic performance of the vessel. The effectiveness of the energy management system was verified in conjunction with a case study of the power demand of a 24 h sea cruise, and the fuel savings were shown to reach up to 28%. Planakis et al. [19,20] developed an energy management strategy considering the trade-off between fuel consumption and NO_x emissions based on model predictive control, and applied it to a diesel–electric parallel hybrid system. The effectiveness of the management strategy was verified using the propeller demand speed case, and the experimental results demonstrated that a trade-off between fuel consumption and NO_x emissions could be

reached, depending on the weights. Xu et al. [21] adopted waste heat recovery to further enhance the energy efficiency of a marine hybrid powertrain, while developing an energy management strategy with the objective of minimizing the total energy consumption and optimizing the engine efficiency. The effectiveness of the proposed scheme was verified using four driving cycles referenced to the ship's speed, in order to obtain quantitative conclusions. The experimental results indicated that the maximum energy savings of the marine hybrid power system with waste heat recovery reached 13.74%, when compared to the conventional engine propulsion system. Li et al. [22] designed an adaptive multi-objective joint optimization algorithm for hybrid energy storage electric propulsion ships to achieve cost minimization. The feasibility of their optimization algorithm was verified utilizing the Alsterwasser load demand condition. Kalikatzarakis et al. [23] applied an equivalent consumption minimization strategy to a hybrid power system for a multi-power source vessel. Experimental validation of the strategy using seven simulated operating profiles of a tugboat showed that the proposed strategy saved 6% of fuel in the case of unknown load demand. Sun et al. [24] came up with a novel layout for a gas–electric hybrid power system for ships, and proposed a model predictive control strategy with variable weighting coefficients to trade off between economy and power. The trade-off characteristics of the proposed energy management strategy were validated on a test rig using a 600 s test condition.

It can be seen, from the above literature, that different types of hybrid power systems for selection and matching analysis cannot be separated from the testing and research of loading cycle conditions. In order to improve the economy, power, and emission of ships through energy management strategies, different types of ships also need to be analyzed and verified according to their loading cycle conditions.

The testing conditions of tugboats [25], passenger ships [26], and engineering vessels [27] were obtained from the literature for a more intuitive reflection of the problem, as shown in Figures 1–3. From these figures, it can be seen that different types of ships utilize varying test conditions, and that there is no uniform test standard, even for the same type of ship. For example, passenger ships can include luxury cruise ships, ferries, and marine amusement rides, where the design and performance requirements can vary greatly between them. The lack of harmonized international calibration standards and the use of different test methods by different manufacturers and operators have resulted in non-comparable test results.

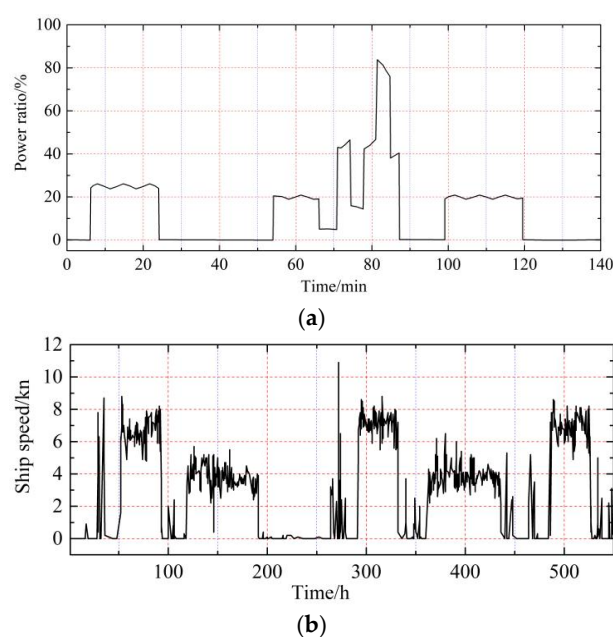


Figure 1. Tugboat working conditions. (a) Working condition 1; (b) Working condition 2.

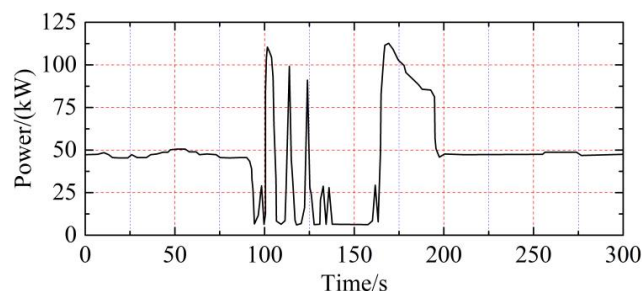


Figure 2. Passenger ship working conditions.

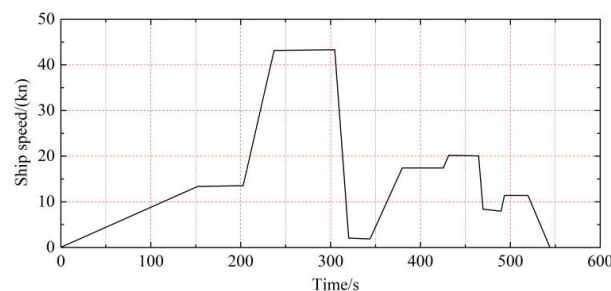


Figure 3. Engineering vessel working conditions.

Hence, the development of standardized ship performance testing protocols has emerged as a critical concern within the global shipping industry. These protocols encompass a set of regulations and procedures designed for the evaluation of ship performance, improvement of fuel efficiency, and reduction of environmental impacts, establishing an essential framework to achieve more efficient, sustainable, and safe vessel operations. This section provides an in-depth examination of the current progress in the development of standardized ship performance testing methods on an international scale, encompassing the latest advancements made by the International Organization for Standardization (ISO) and the International Maritime Organization (IMO).

- The International Organization for Standardization (ISO)

ISO 15016 [28]: ISO 15016, developed by the International Organization for Standardization (ISO), is an international standard designed for evaluating the propulsion performance of ships. This standard outlines comprehensive requirements for testing procedures, data processing, and uncertainty assessment, all aimed at ensuring the accuracy and comparability of the test results. The ongoing updates and improvements to this standard underscore the international community's sustained commitment to ship performance testing methods, forming a robust foundation for continuous advancements within the shipping industry [29].

ISO 19030 [30]: ISO 19030 is centered on standards for measuring a ship's displacement, speed, and fuel consumption. This standard offers guidelines for economically operating vessels, allowing operators to effectively monitor and enhance ship performance, ultimately leading to reduced fuel consumption and emissions [31].

- The International Maritime Organization (IMO)

The International Maritime Organization (IMO) has consistently championed the standardization of ship testing conditions, playing a pivotal role in advancing sustainability within the global shipping industry.

EEDI (Energy Efficiency Design Index): The IMO introduced the Energy Efficiency Design Index (EEDI) as a metric for assessment of the efficiency of new ship designs. This index mandates that ship manufacturers adhere to a set of performance standards when designing new vessels, ensuring heightened energy efficiency. The implementation of the

EEDI actively contributes to propelling the maritime industry toward greater environmental friendliness and enhanced energy efficiency [32,33].

SEEMP (Ship Energy Efficiency Management Plan): The IMO requires shipowners to implement Ship Energy Efficiency Management Plans (SEEMP) to oversee and improve the energy efficiency of existing vessels. The SEEMP encompasses various elements, including testing conditions, data recording, and analysis, among others, all aimed at optimizing ship performance and reducing energy consumption [34,35].

Above all, various kinds of test conditions provide input conditions for performance validation in terms of ship energy management and hybrid matching selection optimization. With regard to the source of test conditions, a single or limited number of operating conditions are mainly adopted, which prevents consideration of the operating characteristics in the water, leading to difficulties when comparing the performance of different types of vessels in practice. Especially in terms of data extraction, there is still a lack of systematic research that combines machine learning with time-series analysis, which limits comprehensive understanding of the real operating conditions of ships. Hence, in this study, a machine learning-based ship loading cycle extraction method is proposed to address the lack of standardized loading cycle testing conditions. The dynamic changes and spatio-temporal offset characteristics characterizing ship operations can be captured more accurately through combining dynamic time regularization (DTW) theory with hierarchical clustering in machine learning. Such an approach not only improves the accuracy and efficiency of data analysis, but also provides a unified standard test condition for the performance comparison of different types of ships.

On the basis of the above, the main structure of this article is as follows. Section 2 introduces basic knowledge on time-series data mining and presents the proposed novel method for data extraction based on machine learning. Furthermore, the proposed method is validated and discussed through a case study in Section 3. The concluding chapter is presented in Section 4.

2. Preliminaries and Methodology

2.1. Similarity Measurement

Time-series data are a common data type that includes numerical values obtained through continuous observations or regular measurements. Such data have applications across diverse domains, including finance, healthcare, telecommunications, and others [36]. Before commencing data mining, it is essential to evaluate the similarity between two time series to reveal potential relationships among distinct sequences. To achieve this goal, distance calculation methods are commonly used to quantify the similarity of time-series data. As a result, in the context of addressing challenges related to time-series similarity, distance metrics play a crucial role.

Similarity measures, often referred to as distance functions, serve to quantify the similarity between two compared data sets. Such data sets can assume various forms, encompassing equally or unequally sized raw numerical sequences, feature-rich vectors, or transformation matrices, among others. Frequently utilized similarity measurement methods include Euclidean distance and Dynamic Time Warping (DTW) [37].

2.1.1. Euclidean Distance

The Euclidean distance, commonly referred to as the Euclidean metric, serves as a widely employed distance measurement method for quantifying the separation between two points within a multi-dimensional space [38]. It is frequently utilized to gauge the similarity of numerical data, particularly in machine learning applications such as clustering, classification, regression, and others. The calculation formula is presented in Equation (1):

$$dist(A, B) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

where $A(x_1, x_2, \dots, x_n)$ and $B(y_1, y_2, \dots, y_n)$ denote distinct time series. It is essential to emphasize that, for this formula to be applicable, both time series must possess an identical number of data points. Despite its computational simplicity and intuitive nature, the Euclidean distance exhibits sensitivity to scale variations across different dimensions.

In comparison to more intricate approaches, the Euclidean distance exhibits faster computational speed when handling large-scale data. Nevertheless, the Euclidean distance possesses certain limitations. It proves ineffective in handling outliers or noise, and displays sensitivity to specific signal transformations, such as translation, uniform time scaling, uniform and non-uniform amplitude scaling, and time warping, among others. To mitigate the limitations of the Euclidean distance, researchers in the field have continuously introduced novel distance measurement methods.

2.1.2. Dynamic Time Warping

DTW, in contrast to the Euclidean distance, excels at accurately assessing the similarity between two time series that may not share the same frequency or sampling rate. These sequences undergo non-uniform time dimension warping (or stretching) to achieve the best possible alignment between the two sequences [39]. A notable advantage of DTW lies in its capacity to preserve the topological structure of time series through accommodating time stretching or compression, thereby enhancing its effectiveness in capturing the similarity between time series. With time series $A(x_1, x_2, \dots, x_n)$ and $B(y_1, y_2, \dots, y_m)$ as inputs, a distance matrix is generated, as depicted in Equation (2):

$$dist(A, B) = \begin{bmatrix} d(x_1, y_1) & d(x_1, y_2) & \cdots & d(x_1, y_m) \\ d(x_2, y_1) & d(x_2, y_2) & \cdots & d(x_2, y_m) \\ \vdots & \vdots & \ddots & \vdots \\ d(x_n, y_1) & d(x_n, y_2) & \cdots & d(x_n, y_m) \end{bmatrix}_{n \times m} \quad (2)$$

where $dist(i, j) = d(x_i, y_j) = \|x_i - y_j\|_2, i = 1, 2, \dots, n; j = 1, 2, \dots, m$ is the distance between x_i and y_j .

The objective of DTW is to identify the optimal path ($P^0(P_1, P_2, \dots, P_s, \dots, P_k)$) among all possible paths, thereby minimizing Equation (3):

$$DTW(A, B) = \Psi(n, m) = \frac{\sum_{s=1}^k d_{P_s} \omega_s}{\sum_{s=1}^k \omega_s} \quad (3)$$

where d_{P_s} represents the distance of path P_s , ω_s is the weighting coefficient, and $\Psi(n, m)$ is the cumulative distance matrix.

$\Psi(n, m)$ represents the minimum distance yielded by the DTW algorithm, and the optimal alignment path can be determined through employing the cumulative distance matrix, as illustrated in Equation (4).

$$\Psi(n, m) = dist(i, j) + \min^{DTW} \{\Psi(i-1, j-1), \Psi(i-1, j), \Psi(i, j-1)\} \quad (4)$$

DTW exhibits adaptability to temporal irregularities in time series, robust similarity measurement capabilities, and a keen sensitivity to temporal relationships [40]. DTW proficiently captures temporal features within signals, rendering it highly versatile in various domains, including speech recognition, handwriting character recognition, bioinformatics, and financial time-series analysis. Nonetheless, it encounters the challenge of high computational complexity, leading to notable computational overhead when employed on extensive data sets. Consequently, it is imperative to judiciously assess its strengths and limitations in particular application contexts.

2.2. Clustering Method

Clustering algorithms, which are integral to the realm of time-series clustering, represent a profoundly significant research area within contemporary machine learning [41]. In contrast to classification, clustering stands as an unsupervised learning technique that operates without the need for labeled data during model training. Grounded in the concept of similarity, clustering endeavors to divide data into distinct clusters, emphasizing high similarity within each cluster and lower similarity between different clusters. Within the domain of time-series analysis, prevalent clustering techniques fall into two categories: partition-based methods and hierarchical methods.

2.2.1. Classification and Clustering Methods

Partition-based clustering methods constitute a prevalent category of clustering algorithms. They operate on the fundamental principle of dividing a data set into numerous non-overlapping subsets or clusters, characterized by high similarity among data points within each cluster and lower similarity between different clusters. These methods usually iteratively refine clustering outcomes until a predefined stopping condition is achieved. K-means stands as one of the most frequently employed partition-based clustering algorithms.

The goal of the K-means algorithm is to minimize the sum of squared distances between each data point within a cluster and its respective cluster center, thereby minimizing the total intra-cluster distance, as expressed by the formula in Equation (5):

$$J = \sum_{i=1}^K \sum_{j=1}^n \omega_{ij} \cdot ||x_j - \mu_i||^2 \quad (5)$$

where J denotes the total objective function value, K signifies the number of clusters, n stands for the total number of data points, x_j represents the j^{th} data point, μ_i corresponds to the center of the i^{th} cluster, and ω_{ij} denotes the binary indicator function.

The steps of the K-means clustering algorithm are as follows:

(1) Initialization

Randomly select K data points as the initial cluster centroids. These initial centroids can be chosen randomly from the data set or determined using other initialization methods.

(2) Iteration

The K-means algorithm iteratively optimizes cluster assignments and centroid positions through the following two steps.

(a) Cluster assignment step

For each data point, calculate its distance to each cluster centroid. Typically, the Euclidean distance is used as the distance metric. The distance equation is presented in Equation (6):

$$d_{ij} = ||x_j - \mu_i||^2 \quad (6)$$

where d_{ij} represents the distance between data point x_j and cluster centroid μ_i . Assign each data point to the cluster with the closest centroid (i.e., select the cluster associated with the minimum d_{ij}).

(b) Cluster update step

For each cluster, re-calculate its centroid by taking the mean (average) of all data points within that cluster. The update formula is as follows:

$$\mu_i = \frac{1}{|S_i|} \sum_{x_j \in S_i} x_j \quad (7)$$

where μ_i represents the new centroid of cluster i , S_i is the set of all data points belonging to cluster i , and $|S_i|$ is the size of cluster i (i.e., the number of data points it contains).

(3) Termination criteria

The K-means algorithm repeats the cluster assignment and cluster update steps until one of the following termination criteria is met:

- (a) Cluster centroids no longer change significantly (below a certain threshold).
- (b) The assignment of data points to clusters remains unchanged.
- (c) A maximum number of iterations is reached.

(4) Output

The output of the K-means algorithm consists of a set of clusters, each containing a group of similar data points and the centroids of these clusters. These clusters can be used for data clustering analysis or other subsequent tasks. In summary, the K-means algorithm's fundamental principle is to iteratively optimize cluster assignments and centroids to maximize the similarity of data points within clusters while minimizing the similarity between clusters.

2.2.2. Hierarchical Clustering Method

Hierarchical clustering is a technique employed to systematically analyze data through the utilization of either agglomerative or divisive algorithms. In an agglomerative algorithm, every data point begins as an independent cluster. Then, the closest clusters are progressively merged until all data points are encompassed within a single cluster. Conversely, a divisive algorithm initiates by grouping all data points into a single large cluster and subsequently divides it into more similar sub-clusters until each data point forms its own cluster [42]. Agglomerative algorithms are extensively employed in clustering analysis, due to their hierarchical and bottom-up construction, wide-ranging applicability, interpretability, and stability, among other advantages. Through treating each data point as an initial cluster, this algorithm systematically merges the most similar clusters over successive steps, yielding a multi-level clustering structure. This approach assists researchers in achieving a more in-depth comprehension of the intricate inherent relationships within the data. Additionally, it exhibits relative insensitivity to the selection of initial points, thereby enhancing the repeatability and interpretability of clustering [43].

The following outlines the standard procedure for agglomerative hierarchical clustering:

- (a) Initialization: Begin by treating each data point as a distinct cluster and calculate the distances between them, utilizing diverse distance metrics such as the Euclidean distance, Manhattan distance, cosine similarity, and more.
- (b) Merge the closest clusters: Identify the two clusters with the smallest distance separation and combine them into a new cluster. Various linkage methods can be employed for this purpose, with common options including complete linkage, average linkage, and Ward's linkage.
- (c) Update the distance matrix: Re-evaluate the distance matrix between clusters to account for the distances involving the newly merged cluster and the remaining clusters. This process usually entails computing the distances between the new cluster and other clusters.
- (d) Iteratively execute steps (b) and (c) until a stopping criterion is achieved, which can include reaching a predetermined number of clusters or the point at which the distance between clusters surpasses a specified threshold.

The principles and computational guidelines for complete linkage, average linkage, and Ward's linkage are discussed below, as illustrated in Equations (8)–(10), respectively.

Complete linkage is employed to compute the distance among all data points within a cluster and selects the maximum distance as the ultimate distance, as depicted in Equation (8):

$$L_{ij}^{Complete} = \max\{dist(x_\alpha, x_\beta), \forall x_\alpha \in C_i, x_\beta \in C_j\} \quad (8)$$

where C_i and C_j are two clusters and $dist(x_\alpha, x_\beta)$ represents the distance between data points x_α and x_β .

Ward's linkage serves to gauge the increase in variance upon merging two clusters, aiding in the selection of an optimal clustering structure. The calculation rule is presented in Equation (9).

$$L_{ij}^{Ward's} = \sum_{x_\alpha \in C_i} \sum_{x_\beta \in C_j} \|x_\alpha - x_\beta\|^2 \quad (9)$$

Calculating the average linkage is a relatively straightforward and intuitive process. It defines the distance between two clusters as the average of distances between all data points, as illustrated in Equation (10).

$$L_{ij}^{Average} = \frac{1}{|C_i||C_j|} \sum_{x_\alpha \in C_i} \sum_{x_\beta \in C_j} dist(x_\alpha - x_\beta) \quad (10)$$

2.3. Methodology

This section introduces a novel approach for extracting ship loading cycles through the utilization of the DTW-hierarchical clustering algorithm. Figure 4 illustrates the workflow of the proposed method, encompassing the subsequent steps.

Phase I: Data Acquisition and Preprocessing

- (1) Data Acquisition: Time-series data during the original voyage are collected through the ship's information acquisition system.
- (2) Piecewise Linear Approximation (PLA): To obtain complete time-series data, the PLA of the original data contributes to reducing the amount of data and preserving the main trends.
- (3) Data normalization: The time-series data are normalized to eliminate differences in magnitude between different data sources and to ensure consistency in subsequent analyses.
- (4) Missing data and outlier processing: Interpolation methods are used to fill in missing data, and outliers are detected and processed. Possible methods include mean interpolation, linear interpolation, or multiple interpolation.

Phase II: Similarity Calculation

The Dynamic Time Warping (DTW) formula (Equation (3)) is utilized to calculate the DTW values between different data time series, which measures the similarity or distance between them.

Phase III: Hierarchical Clustering

- (1) Distance matrix construction: Based on the calculated DTW similarity, the distance matrix of the time-series data is constructed.
- (2) Clustering process: A bottom-up hierarchical clustering approach is used to merge the two nearest clusters until all data points are clustered into one cluster.
- (3) Distance threshold selection: A suitable distance threshold is determined, in order to obtain the desired number of clusters while ensuring that the clustering results are reasonable and stable.

Phase IV: Template Profile Extraction

- (1) Mean profile construction: For the time-series data in each cluster, a template profile is extracted, which represents the average behavior of the time series in that cluster. Representative profiles are constructed using the time-series averaging method.
- (2) Average profile calculation: The time series for each cluster are aligned and their mean values are calculated to obtain a representative mean profile.

A depiction of the extraction of time-series data, DTW similarity calculation, and the hierarchical clustering method based on DTW is illustrated in Figure 4.

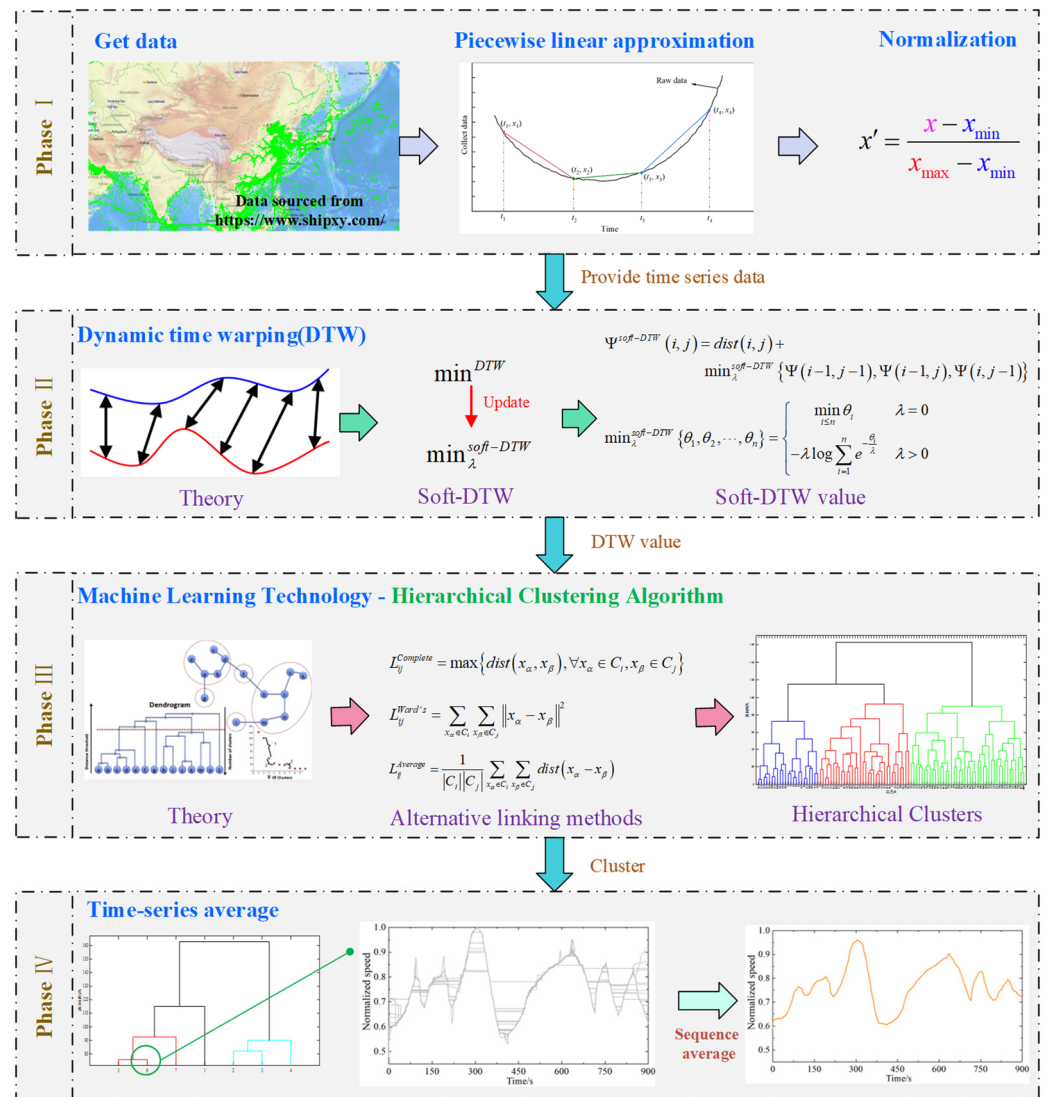


Figure 4. Flowchart of the proposed method.

2.3.1. Extraction of Time-Series Data

Time-series data comprise a sequential collection of records organized in chronological order. Each data point in this sequence is symbolized as $X_i = (t_i, x_i)$, where X_i denotes an individual sample point. Within this notation, x_i represents the numerical value of the data record, t_i corresponds to the time associated with the data record, and the subscript i designates the index of the sample point. Thus, the representation for time-series data is presented in Equation (11).

$$X = [(t_1, x_1), (t_2, x_2), \dots, (t_n, x_n)] \quad (11)$$

In this study, we utilize a piecewise linear approximation method to extract time-series data. This technique involves partitioning the time-series data into multiple segments and conducting linear fitting within each segment. Through connecting these consecutive line segments, the overarching characteristics of the time series can be captured, offering an intuitive depiction of its fluctuation patterns. In the following sections, we introduce and elaborate upon the concept of piecewise linear approximation.

A concise explanation of the piecewise linear approximation method, with reference to Figure 5, is provided below. In Figure 5, the continuous black line depicts the original collected time-series data. Time is partitioned into uniform intervals, and linear fitting is

conducted within each interval. As illustrated in Figure 5, this piecewise linear approximation process yields the final time-series data, as expressed in Equation (12).

$$X = [(t_1, x_1), (t_2, x_2), (t_3, x_3), (t_4, x_4)] \quad (12)$$

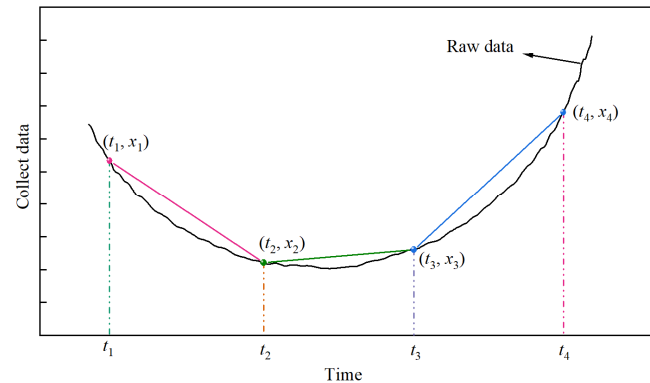


Figure 5. Time segmented linear method.

It is important to highlight that normalization is carried out using the min–max normalization technique, for which the calculation formula is presented in Equation (13):

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (13)$$

where x' denotes the normalized data, x denotes the time-series data, $x_{\min}$ is the minimum value of x , and $x_{\max}$ is the maximum value of x .

2.3.2. Time-Series Data Similarity Search

DTW provides solutions to various challenges related to time-series comparison, encompassing the management of varying lengths, shifts, and scaling transformations. Nevertheless, DTW has various limitations, notably in terms of its high computational complexity, susceptibility to noise, and constraints in global alignment. To address these issues, the introduction of Soft-DTW became imperative. Soft-DTW incorporates a smoothness property that enhances matching robustness, rendering it adept at handling noise, transformations, and discontinuous data. Moreover, it frequently offers enhanced computational efficiency. Consequently, Soft-DTW has emerged as a potent tool which is applicable to a broader spectrum of time-series analysis scenarios.

Consequently, the Soft-DTW computational formula is defined and presented in Equation (14):

$$\Psi^{\text{soft-DTW}}(i, j) = \text{dist}(i, j) + \min_{\lambda}^{\text{soft-DTW}} \{\Psi(i-1, j-1), \Psi(i-1, j), \Psi(i, j-1)\} \quad (14)$$

$$\text{where } \min_{\lambda}^{\text{soft-DTW}} \{\theta_1, \theta_2, \dots, \theta_n\} = \begin{cases} \min_{i \leq n} \theta_i & \lambda = 0 \\ -\lambda \log \sum_{i=1}^n e^{-\frac{\theta_i}{\lambda}} & \lambda > 0 \end{cases}.$$

In contrast to Equation (4), where $\min_{\lambda}^{\text{DTW}}$ is used, $\min_{\lambda}^{\text{soft-DTW}}$ is utilized in the computation. The parameter λ assumes a crucial role in controlling the smoothness of the result matrix, facilitating adjustment of the inherent discrete nature of the DTW distance. Within the framework of the Soft-DTW algorithm, λ serves the purpose of smoothing local minima, thereby creating a more conducive environment for optimization. Specifically, when $\lambda > 0$, the Soft-DTW algorithm introduces an additional smoothing term aimed at mitigating local minima, resulting in a smoother distance matrix. As the parameter λ approaches 0, Soft-DTW gradually converges towards the original DTW distance calculation.

The introduction of this smoothing parameter λ enhances the flexibility of the Soft-DTW method, enabling task-specific adjustments to strike a balance between precise matching and noise tolerance.

2.3.3. Hierarchical Clustering Method Based on Soft-DTW

Employing the piecewise linear approximation method and data normalization, we effectively extracted multiple time-series data sets. Subsequently, we employed the Soft-DTW distance to gauge the similarity among these time-series data sets and conduct hierarchical clustering grounded in this similarity metric. This iterative process was continued until the desired clustering structure was attained.

To be specific, our process commences with the utilization of Equation (12) to compute the Soft-DTW distance for each pair of time-series data. This distance metric incorporates the dynamic time warping between time series, enhancing its accuracy in capturing their similarity.

The Soft-DTW distance enables more accurate measurement of the similarity between time series, thus improving the quality of clustering. The agglomerative clustering method combines the most similar time series in each step and continuously optimizes the clustering structure, resulting in clusters with high internal similarity and low external similarity. When the desired cluster structure or a predefined number of clusters is reached, the cycle is stopped.

The pseudocode for the hierarchical clustering method based on Soft-DTW is presented in Table A1 in the Appendix A.

3. Case Study and Analysis

3.1. Time-Series Data Extraction Using the Modified Version

A total of 100 data points comprising actual ship operating conditions with a time range of 900 s, were acquired using the acquisition method of the website <https://www.shipxy.com> (accessed on 30 January 2023), and named condition 1 to condition 100, as illustrated in Figure 6. These data are categorized into six subgraphs, each containing 20 working conditions.

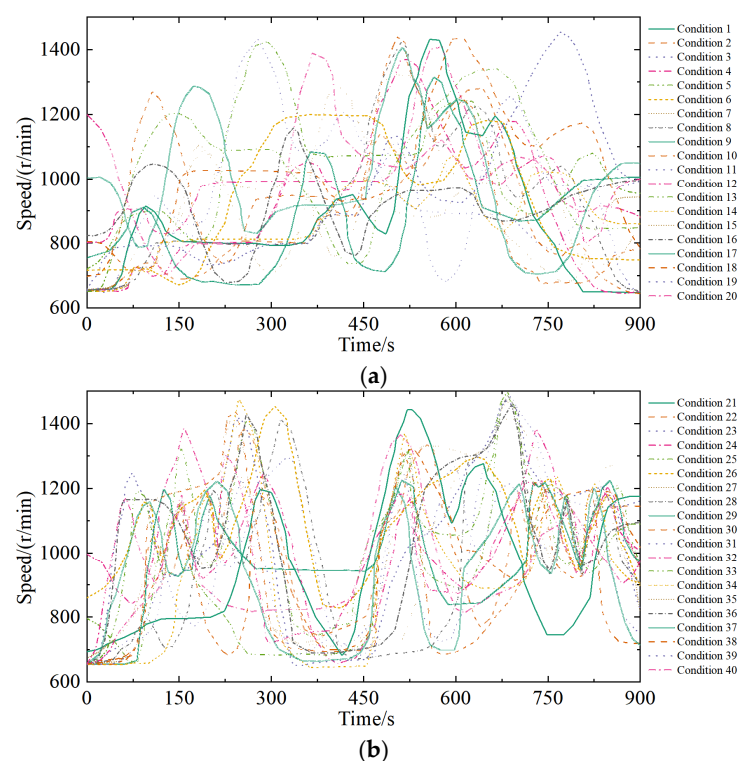


Figure 6. Cont.

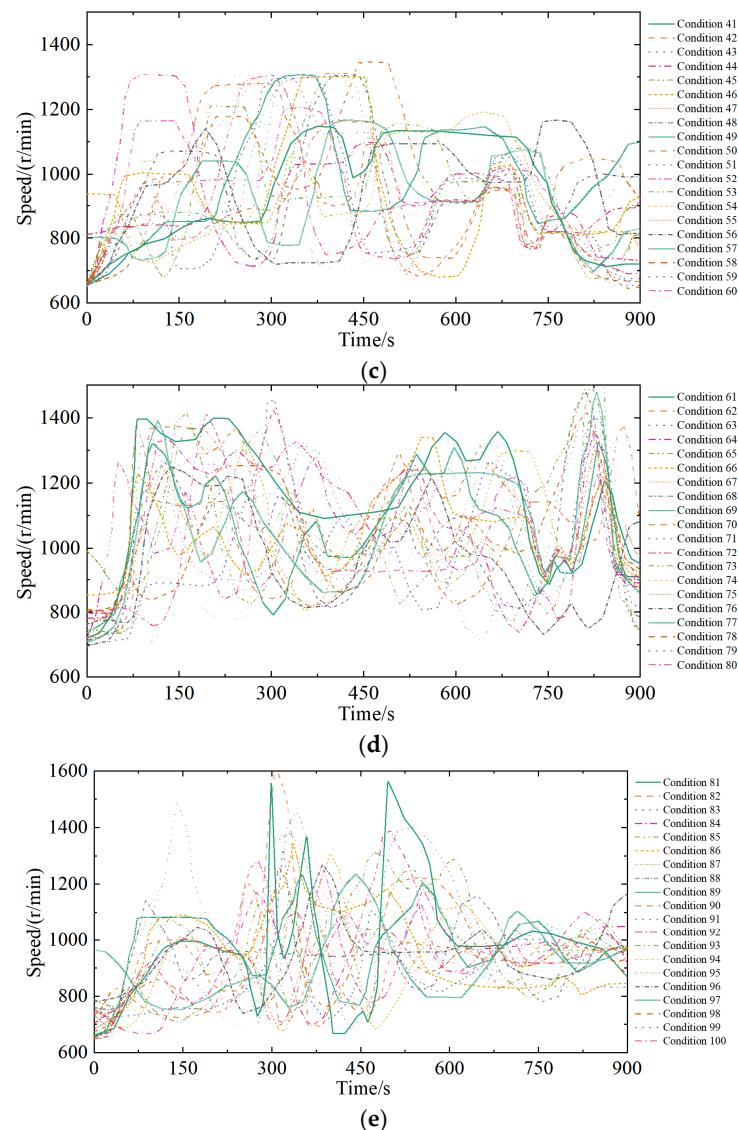


Figure 6. Acquisition of working conditions. (a) Condition 1–Condition 20; (b) Condition 21–Condition 40; (c) Condition 41–Condition 60; (d) Condition 61–Condition 80; (e) Condition 81–Condition 100.

3.2. Results and Discussion

In order to conduct hierarchical clustering analysis, we computed the DTW distances between distinct operating conditions, which are visually represented in Figure 7.

Cluster analysis—an unsupervised machine learning approach—endeavors to group data samples into clusters or classes contingent upon their shared characteristics. During the clustering process, a similarity metric referred to as “ $R^n \rightarrow R$ ” quantifies the resemblance between diverse samples, with ‘ n ’ signifying the count of samples within each operating condition cycle. When comparing the DTW distances across time series, those with analogous temporal profiles were amalgamated into the same cluster. This approach culminated in a clustering of the data, as visually represented in Figure 8.

Utilizing Soft-DTW values, the time-series data were then classified through the complete linkage clustering method, culminating in the creation of a hierarchical clustering tree employing the ‘complete’ method, as depicted in Figure 8a. Throughout the clustering process, a critical decision involves the selection of an appropriate inter-cluster distance or determining the desired number of clusters to uphold cluster diversity. The chosen threshold for the clustering cutoff distance was 71.4, resulting in the identification of seven clusters, as exemplified in Figure 8b.

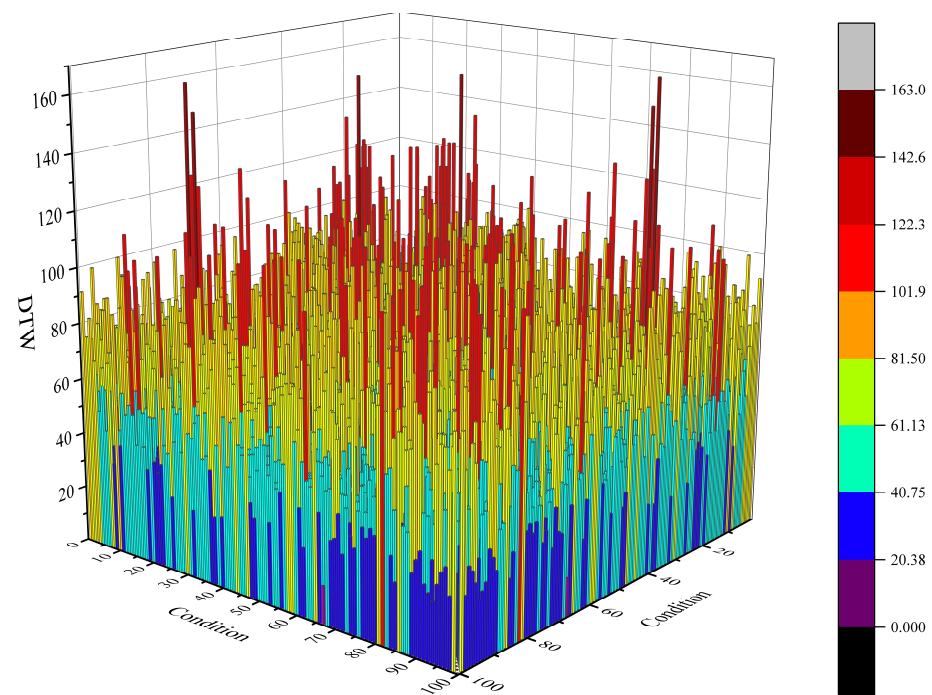


Figure 7. DTW distribution.

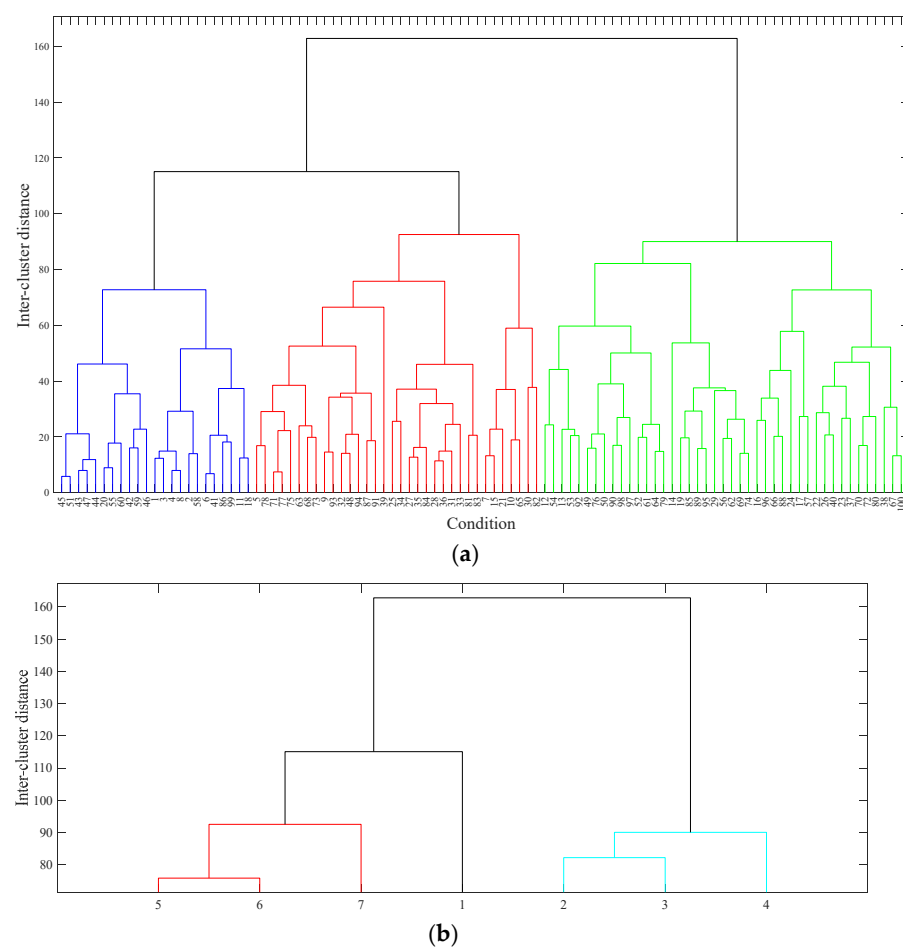


Figure 8. Hierarchical clustering dendrogram: (a) The hierarchy tree of the 'complete' method; (b) Cluster tree with a threshold of 71.4.

The hierarchical structure of the inter-cluster distances was leveraged to construct a dendrogram, thus delineating seven clusters, as detailed in Table 1. Figure 9 portrays the average outcomes of the time-series data for distinct cluster centers, with darker regions denoting the average loading cycles. Upon scrutinizing Figure 9a–g, it becomes evident that the operating conditions within each cluster exhibit analogous fluctuation characteristics. This observation strongly validates the feasibility and effectiveness of hierarchical clustering analysis predicated on Soft-DTW for discerning loading cycles. This outcome underscores the proficiency of our approach in effectively grouping similar operating conditions and accurately capturing their shared attributes. Such findings carry substantial significance in the realm of loading cycle recognition, and they hold promise for providing robust support in practical applications.

Table 1. Results of Soft-DTW-based method to obtain typical operation clusters.

No.	Naming of Typical Loading Conditions	Included Working Conditions
1	Operating profiles I	Condition 25,27,28,31,33,34,35,36,81,83,84
2	Operating profiles II	Condition 5,9,32,39,48,63,68,71,73,75,77,78,87,91,93,94
3	Operating profiles III	Condition 14,19,29,56,62,74,85,89,95
4	Operating profiles IV	Condition 12,13,49,50,52,53,54,61,64,76,79,90,92,97,98
5	Operating profiles V	Condition 16,17,22,23,24,26,37,38,40,57,66,67,70,72,80,88,96,100
6	Operating profiles VI	Condition 7,10,15,21,30,65,82
7	Operating profiles VII	Condition 1,2,3,4,6,8,11,18,20,41,42,43,44,45,46,47,51,55,58,59,60,86,99

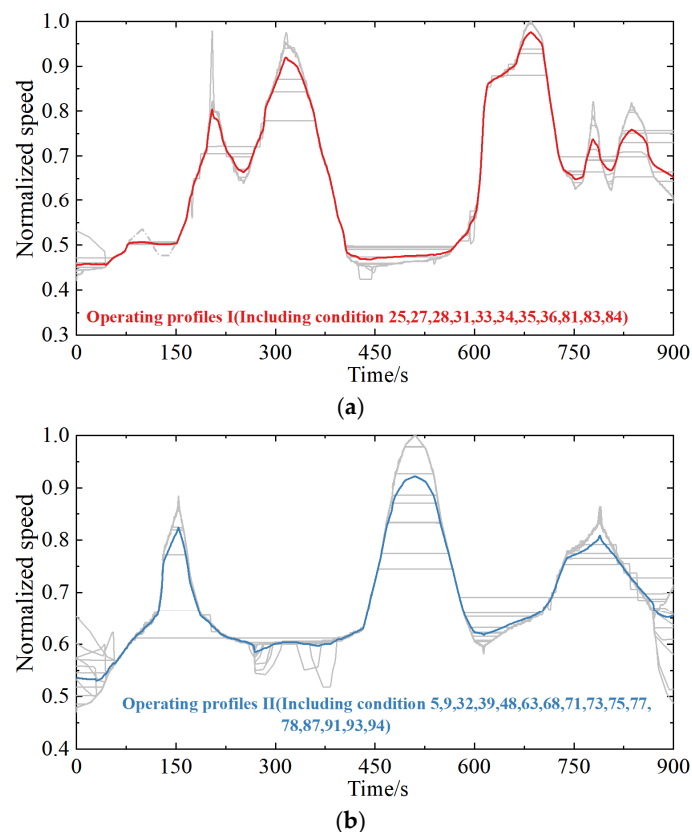
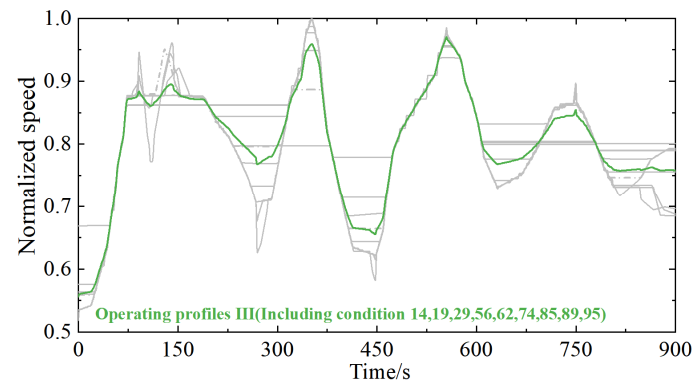
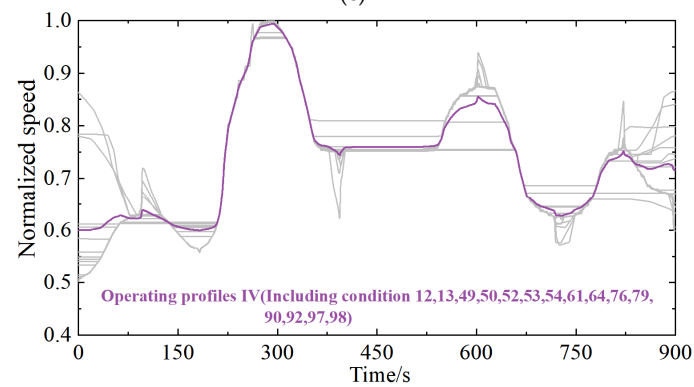


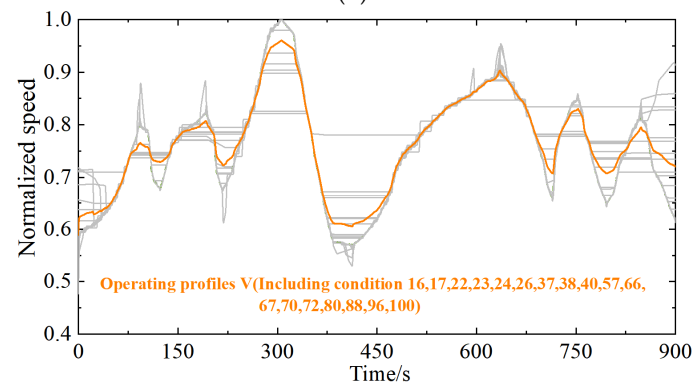
Figure 9. Cont.



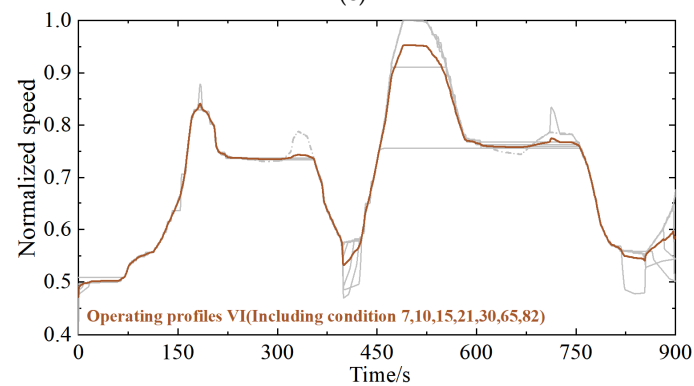
(c)



(d)



(e)



(f)

Figure 9. Cont.

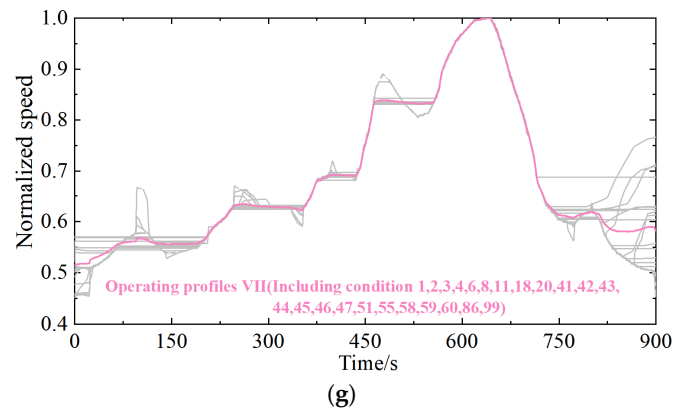


Figure 9. Typical operating conditions. (a) Operating profiles I; (b) Operating profiles II; (c) Operating profiles III; (d) Operating profiles IV; (e) Operating profiles V; (f) Operating profiles VI; (g) Operating profiles VII.

Drawing upon the outcomes of hierarchical clustering and the time-series averaging results illustrated in Figure 9, a comprehensive summary of the loading cycles was meticulously assembled, as depicted in Figure 10. It becomes evident that distinct test conditions offer valuable insights into evaluating the effectiveness of energy management strategies across a range of facets.

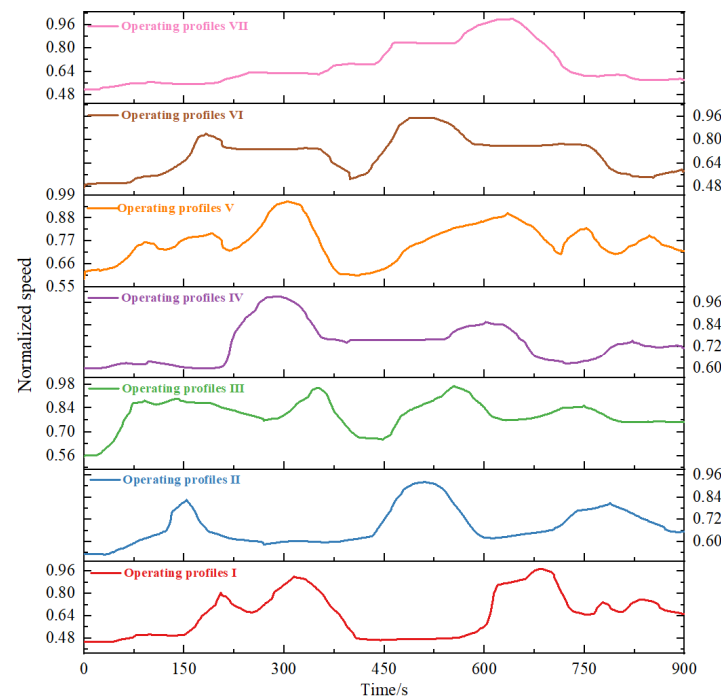


Figure 10. Typical loading cycles.

It can be observed, from Figure 10, that the proposed method effectively divided the time series of 100 different working conditions to obtain seven cyclic working conditions with different characteristics. More specifically, the similarity of the time-series data was calculated using the Soft-DTW method, the results of which were combined with the hierarchical clustering method for clustering. This combination was able to effectively deal with the dynamic and spatio-temporal deviation characteristics in the ship test operating condition data, and the obtained results demonstrate that the Soft-DTW and hierarchical clustering methods can be used to effectively extract the operation cycles with similar characteristics. The finally obtained loading cycle conditions can be used for energy man-

agement performance assessment, which provides a scientific basis for ship performance comparison and energy management strategy optimization.

The methods described above for obtaining data of different loading cycles allow for the identification of types of ships with higher energy consumption and more serious carbon emissions and air pollution. Consequently, ship operators can use this information to develop energy-saving and emissions-reducing navigation strategies to reduce fuel consumption and carbon emissions, thus minimizing their impact on the atmosphere and promoting the environmental sustainability of ship operations. Practical applications of the proposed approach for the advancement of sustainable development are set out below.

(1) Identifying high-energy-consumption and high-emissions operating conditions.

Through obtaining data from different load cycles, it is possible to accurately identify the operating conditions in which different types of ships operate with high energy consumption and severe carbon emissions. Experimental simulations and bench tests can be leveraged to help ship operators identify those operating conditions under which the ship is least energy efficient and has the highest emissions, such that optimization adjustments can be defined.

(2) Developing navigation strategies for energy conservation and emissions reduction.

Ship operators can use this information to develop more accurate navigation strategies and reduce unnecessary energy consumption. For example, it is possible to significantly reduce fuel consumption and emissions through optimizing speeds, adjusting routes, or improving operational procedures under high energy consumption conditions. Through applying these strategies, not only can operational costs be reduced, but environmental pollution can also be significantly reduced.

The above measures assist in promoting the environmental sustainability of ship operations. The ships can maintain efficient operations while reducing carbon emissions and air pollution through standardized test conditions and optimized energy management strategies. As such, the proposed approach is not only compliant with the environmental regulations of the IMO and other regulatory bodies, but also promotes the development of green shipping.

4. Conclusions

In this paper, an advanced marine loading cycle extraction method was presented, which aims to solve the problem related to the lack of standardized test conditions for ships. A Soft-DTW-based hierarchical clustering ship loading cycle extraction method was proposed in combination with a hierarchical clustering method. The DTW theory is skillfully integrated with the hierarchical clustering theory to effectively deal with the dynamic and spatio-temporal deviations of ship test condition data. The proposed extraction method was tested and validated utilizing 100 sets of ship conditions, which yielded successful clustering of 100 sets of time-series data with similar ship condition characteristics. The time-series data in each cluster were averaged to obtain seven loading cycle conditions for matching, ship sizing, and energy management tests, which are characterized by different features and can be used for different functions according to the actual needs of a given application. The contribution of this research is the application of machine learning techniques to marine load cycle extraction, providing a systematic and repeatable extraction method. This not only facilitates comparison and evaluation of the performance of different types of ships, but also provides input conditions for the development of energy saving and emissions reduction strategies for ships.

Despite the important progress made, there are still some limitations to this research. First, the proposed method is highly dependent on a large amount of high-quality operational data, which may be limited by data access concerns in practical applications. Second, the variations between different ship types and operating environments may affect the universality of the method, which requires further validation and optimization. Hence, future research can be expanded in terms of the applied data set, covering more types of

ships and different hydrographic environments, and optimizing the algorithms to help ship operators more easily apply the method for individual performance evaluation and strategy development.

Author Contributions: Conceptualization, X.S. and Q.Z.; methodology, X.S.; software, X.S.; validation, X.S. and Q.Z.; formal analysis, Y.G.; investigation, Y.G.; resources, X.S. and S.D; data curation, X.S.; writing—original draft preparation, X.S.; writing—review and editing, X.S. and S.D; visualization, X.S. and Q.Z.; supervision, X.S., S.D. and Y.G.; project administration, X.S. and Q.Z.; funding acquisition, X.S. and Q.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Initiation Project of Doctoral Research Fund of Liaoning University of Technology, Liaoning Province, China (XB2023027) and The Ministry of Education’s “Chunhui Plan” Cooperative Research Project grant (HZKY20220405).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

Appendix A

Table A1. Pseudocode for Soft-DTW.

Algorithm	Soft-DTW-Hierarchical Clustering Method
1	Inputs: Data combination $X = \{X_1, X_2, \dots, X_m\}$ extracted from time series; Introduce the calculation formula $\Psi^{soft-DTW}$ for Soft-DTW.
2	for $i = 1$ to m
3	$C_i = \{X_i\}$;
4	end for
5	for $i = 1$ to m
6	for $j = 1$ to m
7	$\varphi(i, j) = \Psi^{soft-DTW}(i, j)$;
8	$\varphi(j, i) = \varphi(i, j)$;
9	end for
10	end for
11	Set the current number of clusters $M = m$;
12	$k = 1$;
13	While $M > k$
	Find the two clusters closest to Soft-DTW through equation
14	$(C_i, C_j) = \min L_{ij}^{Complete}$ $= \minmax \{ \varphi(i, j), \forall x_i \in C_i, x_j \in C_j \}$.
15	Merge Cluster C_i, C_j ;
16	Renumber the cluster;
17	for $i = 1$ to $M - 1$
18	$\varphi(i, j) = \Psi^{soft-DTW}(i, j)$;
19	$\varphi(j, i) = \varphi(i, j)$;
20	end for
21	$M = M - 1$;
22	end While
23	Outputs: $C_i = \{C_1, C_2, \dots, C_M\}$

References

1. Zhou, Y.; Li, X.; Yuen, K.F. Sustainable Shipping: A Critical Review for a Unified Framework and Future Research Agenda. *Mar. Policy* **2023**, *148*, 105478. [\[CrossRef\]](#)
2. Dey, A.; Ejohwomu, O.A.; Chan, P.W. Sustainability Challenges and Enablers in Resource Recovery Industries: A Systematic Review of the Ship-Recycling Studies and Future Directions. *J. Clean. Prod.* **2021**, *329*, 129787. [\[CrossRef\]](#)

3. Michail, N.A.; Melas, K.D. Shipping Markets in Turmoil: An Analysis of the COVID-19 Outbreak and Its Implications. *Transp. Res. Interdiscip. Perspect.* **2020**, *7*, 100178. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Acanfora, M.; Balsamo, F.; Fantauzzi, M.; Lauria, D.; Proto, D. Design of an Electrical Energy Storage System for Hybrid Diesel Electric Ship Propulsion Aimed at Load Levelling in Irregular Wave Conditions. *Appl. Energy* **2023**, *350*, 121728. [\[CrossRef\]](#)
5. Planakis, N.; Papalambrou, G.; Kyrtatos, N. A Real-Time Power-Split Strategy for a Hybrid Marine Power Plant Using MPC. *Int. J. Model. Identif. Control.* **2020**, *34*, 147–157. [\[CrossRef\]](#)
6. Sun, X.; Yao, C.; Song, E.; Yang, Q.; Yang, X. Optimal Control of Transient Processes in Marine Hybrid Propulsion Systems: Modeling, Optimization and Performance Enhancement. *Appl. Energy* **2022**, *321*, 119404. [\[CrossRef\]](#)
7. Sun, X.; Yao, C.; Song, E.; Liu, Z.; Yang, X.; Yang, Q. Analysis of Operating Characteristics for Marine Gas-Electric Hybrid Power System. *J. Inst. Eng. Ser. C* **2023**, *104*, 1–13. [\[CrossRef\]](#)
8. Wang, X.; Zhu, H.; Luo, X.; Chang, S.; Guan, X. A Novel Optimal Dispatch Strategy for Hybrid Energy Ship Power System Based on the Improved NSGA-II Algorithm. *Electr. Power Syst. Res.* **2024**, *232*, 110385. [\[CrossRef\]](#)
9. Ganjian, M.; Bagherian Farahabadi, H.; Alirezapouri, M.A.; Rezaei Firuzjaei, M. Optimal Design Strategy for Fuel Cell-Based Hybrid Power System of All-Electric Ships. *Int. J. Hydrogen Energy* **2024**, *50*, 1558–1571. [\[CrossRef\]](#)
10. Wu, S.; Li, T.; Xu, F.; Chen, R.; Zhou, X.; Wang, B. Configuration Size Optimization of Gas-Electric Hybrid Power Systems on Ships Considering Energy Density and Engine Load Response. *Energy Convers. Manag.* **2024**, *301*, 118069. [\[CrossRef\]](#)
11. Wang, Z.; Chen, L.; Wang, B. Tri-Objective Optimal Design of a Hybrid Electric Propulsion System for a Polar Mini-Cruise Ship. *Ocean Eng.* **2024**, *300*, 117355. [\[CrossRef\]](#)
12. Li, Y.; Liu, X.; Zhao, Y.; He, T.; Zeng, H. Optimization Design of Hybrid Energy Storage Capacity Configuration for Electric Ship. *Energy Rep.* **2024**, *11*, 887–894. [\[CrossRef\]](#)
13. Haseltalab, A.; van Biert, L.; Sapra, H.; Mestemaker, B.; Negenborn, R.R. Component Sizing and Energy Management for SOFC-Based Ship Power Systems. *Energy Convers. Manag.* **2021**, *245*, 114625. [\[CrossRef\]](#)
14. Chen, H.; Zhang, Z.; Guan, C.; Gao, H. Optimization of Sizing and Frequency Control in Battery/Supercapacitor Hybrid Energy Storage System for Fuel Cell Ship. *Energy* **2020**, *197*, 117285. [\[CrossRef\]](#)
15. Zhang, Y.; Xue, Q.; Gao, D.; Shi, W.; Yu, W. Two-Level Model Predictive Control Energy Management Strategy for Hybrid Power Ships with Hybrid Energy Storage System. *J. Energy Storage* **2022**, *52*, 104763. [\[CrossRef\]](#)
16. Ghimire, P.; Karimi, S.; Zadeh, M.; Nagalingam, K.K.; Pedersen, E. Model-Based Efficiency and Emissions Evaluation of a Marine Hybrid Power System with Load Profile. *Electr. Power Syst. Res.* **2022**, *212*, 108530. [\[CrossRef\]](#)
17. Zhao, Z.-H. Improved Fuzzy Logic Control-Based Energy Management Strategy for Hybrid Power System of FC/PV/Battery/SC on Tourist Ship. *Int. J. Hydrogen Energy* **2022**, *47*, 9719–9734. [\[CrossRef\]](#)
18. Xie, P.; Asgharian, H.; Guerrero, J.M.; Vasquez, J.C.; Araya, S.S.; Liso, V. A Two-Layer Energy Management System for a Hybrid Electrical Passenger Ship with Multi-PEM Fuel Cell Stack. *Int. J. Hydrogen Energy* **2024**, *50*, 1005–1019. [\[CrossRef\]](#)
19. Planakis, N.; Papalambrou, G.; Kyrtatos, N. Predictive Power-Split System of Hybrid Ship Propulsion for Energy Management and Emissions Reduction. *Control Eng. Pract.* **2021**, *111*, 104795. [\[CrossRef\]](#)
20. Planakis, N.; Papalambrou, G.; Kyrtatos, N. Ship Energy Management System Development and Experimental Evaluation Utilizing Marine Loading Cycles Based on Machine Learning Techniques. *Appl. Energy* **2022**, *307*, 118085. [\[CrossRef\]](#)
21. Xu, C.; Fan, L.; Feng, Y.; Zhu, Y.; Shen, C.; Jiang, Z. A Multi-Objective Optimization Energy Management Strategy for Marine Hybrid Propulsion with Waste Heat Recovery System. *Appl. Therm. Eng.* **2024**, *236*, 121548. [\[CrossRef\]](#)
22. Li, X.; Huang, J.; Zhang, J.; Zhou, M.; Wang, T.; Tang, X.; Lai, J.; Yang, X. An Adaptive Multi-Objective Joint Optimization Framework for Marine Hybrid Energy Storage System Design Considering Energy Management Strategy. *J. Energy Storage* **2023**, *68*, 107689. [\[CrossRef\]](#)
23. Kalikatzarakis, M.; Geertsma, R.D.; Boonen, E.J.; Visser, K.; Negenborn, R.R. Ship Energy Management for Hybrid Propulsion and Power Supply with Shore Charging. *Control Eng. Pract.* **2018**, *76*, 133–154. [\[CrossRef\]](#)
24. Sun, X.; Yao, C.; Song, E.; Liu, Z.; Ke, Y.; Ding, S. Novel Enhancement of Energy Distribution for Marine Hybrid Propulsion Systems by an Advanced Variable Weight Decision Model Predictive Control. *Energy* **2023**, *274*, 127269. [\[CrossRef\]](#)
25. Yuan, L.C.W.; Tjahjowidodo, T.; Lee, G.S.G.; Chan, R.; Adnanes, A.K. Equivalent Consumption Minimization Strategy for Hybrid All-Electric Tugboats to Optimize Fuel Savings. In Proceedings of the 2016 American Control Conference (ACC), Boston, MA, USA, 6–8 July 2016; pp. 6803–6808.
26. Han, J.; Charpentier, J.-F.; Tang, T. An Energy Management System of a Fuel Cell/Battery Hybrid Boat. *Energies* **2014**, *7*, 2799–2820. [\[CrossRef\]](#)
27. Soleymani, M.; Yoosofi, A.; Kandi-D, M. Sizing and Energy Management of a Medium Hybrid Electric Boat. *J. Mar. Sci. Technol.* **2015**, *20*, 739–751. [\[CrossRef\]](#)
28. ISO 15016; Ships and Marine Technology—Guidelines for the Assessment of Speed and Power Performance by Analysis of Speed Trial Data. International Organization for Standardization: Geneva, Switzerland, 2015.
29. Kim, K.-S.; Roh, M.-I. ISO 15016:2015-Based Method for Estimating the Fuel Oil Consumption of a Ship. *J. Mar. Sci. Eng.* **2020**, *8*, 791. [\[CrossRef\]](#)
30. ISO 19030; Ships and Marine Technology—Measurement of Changes in Hull and Propeller Performance. International Organization for Standardization: Geneva, Switzerland, 2016.

31. Shaw, H.-J.; Lin, C.-K. Marine Big Data Analysis of Ships for the Energy Efficiency Changes of the Hull and Maintenance Evaluation Based on the ISO 19030 Standard. *Ocean Eng.* **2021**, *232*, 108953. [\[CrossRef\]](#)
32. Ančić, I.; Theotokatos, G.; Vladimir, N. Towards Improving Energy Efficiency Regulations of Bulk Carriers. *Ocean Eng.* **2018**, *148*, 193–201. [\[CrossRef\]](#)
33. Böckmann, E.; Steen, S. Calculation of EEDIweather for a General Cargo Vessel. *Ocean Eng.* **2016**, *122*, 68–73. [\[CrossRef\]](#)
34. Hansen, E.K.; Rasmussen, H.B.; Lützen, M. Making Shipping More Carbon-Friendly? Exploring Ship Energy Efficiency Management Plans in Legislation and Practice. *Energy Res. Soc. Sci.* **2020**, *65*, 101459. [\[CrossRef\]](#)
35. Nguyen, S. Development of an MCDM Framework to Facilitate Low Carbon Shipping Technology Application. *Asian J. Shipp. Logist.* **2018**, *34*, 317–327. [\[CrossRef\]](#)
36. Liu, Z.; Zhu, Z.; Gao, J.; Xu, C. Forecast Methods for Time Series Data: A Survey. *IEEE Access* **2021**, *9*, 91896–91912. [\[CrossRef\]](#)
37. Miljkovic, D.; Vatsa, P. On the Linkages between Energy and Agricultural Commodity Prices: A Dynamic Time Warping Analysis. *Int. Rev. Financ. Anal.* **2023**, *90*, 102834. [\[CrossRef\]](#)
38. Qin, H.; Wang, Y.; Ma, X.; Wang, J.; Jiang, C. A Euclidean Distance-Based Parameter Reduction Algorithm for Interval-Valued Fuzzy Soft Sets. *Expert Syst. Appl.* **2023**, *234*, 121106. [\[CrossRef\]](#)
39. Cai, J.; Fei, W.; Fu, S.; Lu, M.; Yang, H. High Reliability Damage Imaging under Non-Uniform Environmental Temperature Variations Based on Modified Dynamic Time Warping. *Mech. Syst. Signal Process.* **2023**, *203*, 110737. [\[CrossRef\]](#)
40. Kumar, P.S.; Laha, S.K.; Kumaraswamidhas, L.A. Assessment of Rolling Element Bearing Degradation Based on Dynamic Time Warping, Kernel Ridge Regression and Support Vector Regression. *Appl. Acoust.* **2023**, *208*, 109389. [\[CrossRef\]](#)
41. Hu, H.; Liu, J.; Zhang, X.; Fang, M. An Effective and Adaptable K-Means Algorithm for Big Data Cluster Analysis. *Pattern Recognit.* **2023**, *139*, 109404. [\[CrossRef\]](#)
42. Varshney, A.K.; Muhuri, P.K.; Danish Lohani, Q.M. PIFHC: The Probabilistic Intuitionistic Fuzzy Hierarchical Clustering Algorithm. *Appl. Soft Comput.* **2022**, *120*, 108584. [\[CrossRef\]](#)
43. Habib, A.; Akram, M.; Kahraman, C. Minimum Spanning Tree Hierarchical Clustering Algorithm: A New Pythagorean Fuzzy Similarity Measure for the Analysis of Functional Brain Networks. *Expert Syst. Appl.* **2022**, *201*, 117016. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.