JOURNAL OF MARITIME RESEARCH

SPECIAL ISSUE (2026) pp 101–117

ISSN: 1697-4840, www.jmr.unican.es

Evaluating Reliability in High-Dimensional Clustering: A Dimension-Aware and Stability-Based Analysis of Projection-Induced Misinterpretation in Offshore Wind Data

Hamid Reza Soltani Motlagh[1,2,*], Seyed Behbood Issa-Zadeh[3,4], Abdul Hameed Kalifullah[2], Md Redzuan Zoolfakar[1]

ARTICLE INFO

Special Issue:
1st International Maritime and Logistics Conference;
February 10-11, 2026.
National University of Science & Technology (NU), Oman

Article history:
Received 15 Feb 2026;
in revised from 24 Feb 2026;
accepted 05 Jun 2026.

Keywords:
Offshore wind analysis;
High-dimensional clustering; Cluster stability; Dimensionality reduction;
Projection-induced misinterpretation.

ABSTRACT

Analyses of variability of offshore wind resources using multi-dimensional atmospheric data are performed using clustering methods. Evaluating the quality of the clustering results becomes a challenging task when dealing with large number of features. The reasons behind this are that visual interpretation of the resulting low-dimensional representations of clusters is not clear and stable, because of the poor separation between clusters and the instability in the composition of the clusters. This work studies the effects of increasing the number of features in the data on the clustering performance, stability and interpretability in the context of offshore wind data, by applying a dimension-aware and stability-oriented methodology. A multi-dimensional offshore wind dataset from a controlled experimental testbed was derived from reanalysis-based atmospheric variables. The datasets were divided into four tiles that represented the diversity of offshore wind resources and were subjected to high-level clustering analyses. The progression of the clusters in decreasing orders of feature dimensions (from low to high) and their implications were investigated. In addition to widely used validity indices, other stability indices were introduced to analyse the stability of the clusters, the stability of cluster assignments, and the sensitivity of the clusters to new variables. The research investigates two aspects of cluster result analysis from high-dimensional data through studies of fixed quality evaluation methods and the reliability of low-dimensional projection techniques. The research shows that two- and three-dimensional visualizations produce incorrect cluster overlap and weak separation perceptions although the original feature space shows large distances between cluster centers. The methodological evaluation of high-dimensional clustering faces a major failure mode because of this phenomenon which scientists call projection-induced misinterpretation. Cluster analysis is not necessarily subject to the curse of dimensionality, as many clustering problems are a result of data visualization and dimensionality reduction limitations. This study addresses dimension-aware scenario design, stability-driven validation, use of the true high-dimensional structure of the data as a reference and the need for explicit projection auditing. The experimental results demonstrate that any reliable cluster analysis solution for high-dimensional data space has to at least include dimension-aware scenario design, stability-driven validation and explicit projection auditing. The purpose of this paper is to establish some rules of thumb for scientists for experimental design and presentation of results to facilitate cluster analysis investigations of offshore wind resources in high-dimensional data space.

[1]Institut Teknologi Malaysia Kejuruteraan Marin, Universiti Kuala Lumpur, Lumut 32200, Perak, Malaysia.
[2]International Maritime College Oman, National University of Science and Technology, Sohar P.O. Box 532, Oman.
[3]School of Maritime Science, University of Gibraltar, Campus Europa Point, Gibraltar GX11 1AA, UK.
[4]Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos, Universidad Politécnica de Madrid, 28040 Madrid, Spain.
*Corresponding author: H.R.S. Motlagh. E-mail: hamid@imco.edu.om.

1. Introduction.

Unsupervised clustering has become a widely adopted tool for analyzing offshore wind datasets, enabling the identification of wind regimes and dominant atmospheric patterns from large-scale reanalysis products. Research has proven that clustering methods work well for wind pattern identification which helps wind energy professionals conduct wind resource assessments through proper selection of wind characteristics and their combination[1], [2]. The increasing precision of atmospheric reanalysis data makes clustering methods more appropriate for offshore wind studies because new feature extraction methods allow researchers to describe wind patterns effectively in different wind conditions [2].

The analysis of offshore wind data now uses clustering methods which operate in feature spaces that contain wind speed data together with atmospheric variables including temperature and pressure and air density and wind characteristics. The use of higher-dimensional representations which should improve physical offshore wind system modeling creates new challenges for researchers who need to assess and understand these systems. In particular, conventional clustering studies often rely on a limited set of cluster validity indices to assess performance, without explicitly examining how clustering behavior changes as feature dimensionality increases [3], [4]. Researchers need to verify the stability and dependability of clustering results which occur when clustering methods operate in high - dimensional spaces.

The interpretation process becomes easier when researchers display their clustering results through two- or three - dimensional visualizations which stem from dimensionality reduction methods including principal component analysis (PCA) and t - distributed stochastic neighbor embedding (t-SNE) and uniform manifold approximation and projection (UMAP). The methods function as basic analytical instruments which help users generate basic visualizations from complex data sets in machine learning and data analysis. Visual analytics research demonstrates that high-dimensional data analysis through dimensionality reduction methods results in major modifications to both spatial relationships between points and their neighboring patterns [5], [6], [7]. The distortions create false visual representations which make it difficult to determine cluster separation or overlap even though the clusters maintain their correct positions in the original feature space [6].

Research results show that users must avoid depending on reduced-dimensional representation accuracy because these visualizations need to support analytical decision-making processes. Research using quantitative methods demonstrates that projection methods maintain specific local patterns but they alter the way different clusters connect to each other which results in reduced interpretability [5], [8]. The existing research about visual analytics and machine learning limitations contains well-known information but scientists have not studied these restrictions in detail for offshore wind clustering applications.

The research community now recognizes cluster stability as an essential factor which helps validate clustering results because it shows how well clustering solutions maintain their stability when researchers apply different types of data perturbations. Research studies show that stability analysis functions as an essential evaluation technique for unsupervised learning systems which need to be tested[9], [3], [4]. The assessment of stability occurs independently from visualization methods yet researchers have not studied how projection distortions affect cluster separation and robustness in high-dimensional wind data.

The current research investigates how increasing feature dimensionality and different cluster numbers affect clustering outcomes through its analysis of a controlled four-tile offshore wind dataset.The research investigates methods to evaluate and interpret clustering results which emerge when analyzing big datasets that contain various features. The evaluation of clustering results receives special focus because projection-induced misinterpretation stands as a methodological evaluation failure mode which distorts actual cluster relationships through simplified low-dimensional visualizations. The research combines dimension-aware clustering design with stability-oriented evaluation and physics-informed interpretability considerations to establish a framework which enables better analysis and reporting of offshore wind clustering results.

2. Background.

The analysis of offshore wind resources now depends on unsupervised learning methods which reduce extensive atmospheric data patterns into essential operational regimes for design work and weather prediction and energy production evaluation. Research indicates that clustering methods can identify wind states from reanalysis and hindcast data which scientists can use to study how these states affect seasonal patterns and daily cycles and their impact on wind energy production [1]. Offshore wind assessments now use multi-variable feature spaces to represent wind regimes because they need to describe wind conditions that single-variable wind speed measurements cannot separate[2].

Research on domain studies shows that regime discovery methods function well but requires proper execution for their successful application. Huang et al. demonstrated offshore wind energy assessment through clustering-based parameter estimation which proves that unsupervised methods can effectively perform regime-oriented resource assessment[1]. Schelbergen et al. The research demonstrates cluster high - dimensional wind profile shapes for airborne wind energy statistics which shows how wind applications now use profile-level and multi-feature structure instead of depending on scalar predictors[2]. The Argonne National Laboratory report by Mitra demonstrates an extensive offshore framework which uses PCA and SOMs and k-means to identify long-term offshore wind regimes and analyze their distinct operational patterns for offshore design and siting applications[10]. The SOM - based weather regime analysis of ERA5 reanalysis data for offshore operations revealed that mesoscale regime classification systems identify essential wind patterns which affect daily capacity-factor performance [11]. Research has applied synoptic-pattern clustering to link circulation patterns with combined offshore wind power output and wind speed changes which show that weather pattern shifts create significant patterns in offshore power generation and its peak changes [12].

The successful applications of these methods use feature dimensionality as a predetermined design parameter instead of an experimental factor which researchers can manipulate. Offshore regime papers define a set of features which they use to run a clustering process to identify regimes but they do not provide methods to measure how regime patterns change when new features are added step by step (e.g. from 2D to 3D to 5D and so on); the papers lack methods to determine when new variables enhance regime accuracy without creating unstable redundant patterns [10], [13], [12]. The current lack of research methods becomes more critical because offshore datasets now contain more detailed information (e.g. wind speed and direction data combined with thermodynamic variables and stability indicators). The behavior of high-dimensional clustering systems becomes significantly different when new features are added to the data while the data scales and correlations between variables change.

The interpretation of high-dimensional clustering data requires dimensionality reduction (DR) methods which produce low-dimensional projections through PCA and t-SNE and UMAP techniques. The projections serve as visual proof to show how clusters stay separate from each other and how they overlap with each other and their distance from each other which leads to conclusions about regime stability and distinctiveness. Visual analytics research shows that DR embeddings produce incorrect visualizations which lead users to misinterpret complex high-dimensional relationships because they rely on 2D/3D visualizations[6]. The survey about reliability in DR-based visual analytics shows how three factors which include distortion and workflow coupling and interpretability failure can damage scientific findings despite proper execution of analytics [14]. The research demonstrates through cross-domain empirical critique that DR applications produce numerous incorrect results which analysts incorporate into their analytical procedures because they treat projected views as absolute representations of global patterns [15]. The research results establish the need for projection-aware validation because scientists use low - dimensional visualizations to prove their high-dimensional clustering results.

The present methods used to evaluate embedding quality fail to assess the dependability of projection outcomes which occur during clustering operations. Jeon et al. The research develops methods to measure inter-cluster reliability which show how well cluster relationships maintain their connections between the original data space and the projected space while addressing a major weakness of neighbourhood-preservation metrics for cluster separation and configuration analysis [6]. The research by Machado et al. The research demonstrates that achieving high scores on standard projection quality assessment tools does not ensure that projections will produce accurate or meaningful results; the evaluation metrics fail to detect when projections produce confusing or deceptive visual patterns which indicates that metric-pass does not guarantee trustworthy interpretation [16]. Research into new methods for reliability assessment has developed through two main areas which detect missing data points in maps and establish mathematical tests to identify problematic interpretation issues in UMAP and t-SNE embeddings [17]. Visual analytics systems of today combine DR technology with cluster validation methods to conduct evidence-based statistical analysis of complex high-dimensional data environments. The field of cluster interpretation through projections now operates as an active research domain which exceeds its former status as a completed subject [18]. The existing research demonstrates that offshore wind regime studies experience a specific failure mode which occurs when projection methods create false impressions about cluster proximity in 2D/3D visualizations even though these clusters maintain their distinct positions in the original feature space.

The third fundamental methodological base requires researchers to create proper evaluation methods which help them determine the quality of their cluster solutions. Static internal validity indices, particularly the silhouette coefficient, are widely used to report clustering "quality" and to select the number of clusters. The research approach for stability-based analysis demonstrates that one fixed score value does not work well for unsupervised learning because important cluster patterns need to survive different types of data modifications including resampling and feature addition or deletion and noise addition [4]. The complete assessment establishes stability as an evaluation method which works for unsupervised clustering when no external validation data exists and demonstrates how stability patterns reveal how models respond to different modeling decisions [3]. Hennig's cluster-wise stability framework provides an influential foundation for evaluating stability at the cluster level rather than treating stability as a single global number, supporting more diagnostic evaluation [9]. The concepts gain greater importance when working with high-dimensional offshore wind data because new variables enhance physical description but create stability problems when they increase noise levels and redundant information and scaling errors.

The research on offshore wind power shows that multivariate regime characterization methods work effectively for practical applications. The DR reliability research shows that low-dimensional projection methods face two major problems which include data distortion and reduced interpretability. The stability literature shows that unsupervised inference methods need to be resistant to perturbations for producing defensible results. This work constitutes a single unified contribution to the field of data clustering. The main contributions of this work are four fold: 1) demonstrating the importance of dimension aware scenario design in order to understand how clustering behavior changes with the dimensionality of the feature space; 2) demonstrating the importance of stability driven validation metrics in order to differentiate between meaningful cluster structures and trivial partitions; 3) demonstrating that the high dimensional centroid geometry should be considered as the primary structural metric for clustering; and 4) that any distortion measure provided in relation to projections should also be provided in relation to dimensionality reduction techniques used for data visualization, so that the distortion due to dimensionality reduction is understood and separated from the intrinsic structure of the data and not from its visualization. The usefulness of dimensionality reduction as a data visualization tool is not controversial. However, it is important to understand and quantify the distortion that is introduced, so that one can tell what in the data visualization is data and what is distortion.

Table 1: Comparative Evidence Map of Offshore Wind Clustering, Stability, and Projection Reliability.

Ref.	Study Focus	Data / Feature Dimensionality	Clustering & Evaluation Strategy	Use of Projection / Visualization	Identified Methodological Gap
[1]	Offshore wind assessment via clustering	Multivariate wind features; fixed feature design	Cluster-based parameter estimation	Visual inspection of clustering outputs	No staged analysis of dimensionality growth or stability trends
[2]	Wind profile-shape clustering for energy statistics	High-dimensional wind profiles	Silhouette + descriptive statistics	Profile plots for interpretation	Lacks stability analysis and projection reliability checks
[10]	Offshore regime characterization (ANL)	Wind speed/direction + shear/veer + LLJ index	PCA–SOM–k-means; regime diagnostics	PCA embedded in pipeline	No controlled feature-set growth; no projection auditing
[11]	Offshore wind regimes via SOM (ERA5)	Multivariate circulation–wind linkage	Regime relevance to offshore variability	Regime maps and composites	Interpretation relies on fixed projections
[12]	Synoptic regimes and offshore wind power ramps	Circulation-based regime descriptors	Regime-conditioned power analysis	2D regime visualizations	Projection distortion not assessed
[14]	Weather-pattern clustering for wind & wakes	Atmospheric pattern space	Regime → wake/power relationships	Pattern plots support narrative	No dimensional sensitivity or stability study
[15]	Long-term offshore wind event catalog	Multi-decade ERA5 wind features	Event classification	Statistical summaries	Not framed as clustering stability problem
[16]	Offshore ERA5 validation	Vertical wind profiles	Error/statistical evaluation	Profile visualization	No clustering or projection reliability
[19]	Offshore wind profiles (LiDAR/SCADA)	Tall-profile, multivariate	Case-based regime assessment	Profile plots	Not designed for dimension-aware clustering
[20]	Flow modification & blockage	LiDAR-based flow features	Diagnostic clustering	Visual diagnostics	No stability or projection auditing
[21]	Offshore wind datasets & energy estimation	ERA5-based multi-variable	Comparative assessment	Summary plots	Not clustering-centric
[22]	Tall-profile reanalysis validation	Multi-level wind structure	Validation metrics	Vertical plots	No regime or clustering analysis
[23]	Review of wind resource assessment	Multi-study synthesis	Methodological review	Conceptual figures	Does not address DR reliability
[24]	Extreme offshore wind events	High-dimensional ERA5 fields	Event classification	Statistical maps	Not clustering stability–focused
[25]	Wind profile clustering (Atmosphere)	Multivariate profile features	k-means clustering	2D/3D projections	Projection reliability not quantified
[26]	Wind farm aggregation via clustering	Multi-farm operational features	Probabilistic clustering	Aggregate plots	No projection or dimension sensitivity
[27]	Offshore wind regimes (North Sea)	Circulation + wind coupling	Regime classification	Regime composites	No stability or DR reliability analysis
[28]	Shape-based offshore wind profile clustering	Long-term LiDAR profiles	Shape clustering	Projection-based interpretation	No audit of projection distortion
[6]	Inter-cluster reliability metrics	Generic high-dimensional spaces	Reliability metrics (steadiness/cohesiveness)	Quantifies projection distortion	Not applied to offshore wind
[29]	Projection quality metric limits	Generic high-dimensional data	Metric critique	Demonstrates misleading projections	Not linked to clustering pipelines
[30]	Reliable visual analytics (survey)	Cross-domain	Workflow-level reliability	DR reliability synthesis	No domain-specific integration
[31]	Cross-domain DR misuse critique	Cross-domain	Usage analysis	Identifies misinterpretation	Not clustering-specific
[3]	Stability estimation (review)	Generic clustering	Stability as validation paradigm	Not visualization-focused	No coupling with DR reliability
[9]	Cluster-wise stability	Generic clustering	Cluster-level stability	Not visualization-focused	Predates modern DR-heavy workflows

Source: Authors.

The research evidence presented in Table 1 shows that scientists now use unsupervised clustering to analyze atmospheric complexity because this method produces meaningful wind regimes which help with wind resource evaluation and weather prediction and system performance monitoring. Research conducted in particular fields shows that clustering techniques extract important wind patterns from both reanalysis and observational data which connect to time-dependent patterns and production changes and wake dynamics. The offshore assessment process using clustering methods shows that multivariate regime discovery allows researchers to determine parameters and study resources by studying wind speed and direction and vertical structure as a single system [1], [2]. The Argonne National Laboratory regime classification system demonstrates how to use dimensionality reduction with clustering methods to analyze offshore wind data from multiple decades for identifying normal operating conditions and abnormal system behavior[10]. Research which unites self-organizing maps with ERA5 circulation data and synoptic-pattern clustering methods demonstrates that regime-based models remain essential for studying offshore wind power fluctuations which produce power ramp events [11], [12].

The evidence map shows that researchers do not commonly use feature dimensionality as an experimental factor. Offshore wind clustering research uses established multivariate features for analysis before it moves to identify regimes and interpret results without studying how new variables affect the established regime patterns. The offshore regime frameworks which employ multivariate descriptors fail to show how their systems react to changes in feature selection [10]–[14]. The existing research provides no clear direction about when using more complex models will enhance the accuracy of regime identification but will not create unnecessary information or system problems. The research on stability-focused clustering presents an opposing perspective because experts agree that data clusters need to preserve their structure to uncover vital unsupervised patterns when reference labels are unavailable [3]. The stability concepts of Classic cluster-wise stability demonstrate that stability diagnostics help identify which system regimes produce stable results and which ones produce unstable outcomes [9]. However, these stability principles are seldom operationalized as a primary evaluation axis in offshore wind regime studies.

A second observation from Table 1 is that low-dimensional projections are crucial to the regime classification. Research on offshore wind power uses two-dimensional and three - dimensional visualizations to evaluate how close wind turbines are to each other even though their features exist in much larger dimensional spaces [10]–[12],[25] , [28]. Research in visual analytics shows that this practice contains dangerous elements because dimensionality reduction techniques alter distance relationships which produce incorrect cluster structure representations. The research on inter-cluster reliability showed that standard embedding metrics do not work to evaluate how well cluster relationships stay intact which motivated scientists to create new reliability measures for projection-based reasoning systems [6]. Computer graphics evidence supports that projection methods can generate high quality scores which produce deceptive visual effects that demonstrate metric validation does not ensure correct interpretation [29]. Scientific workflows experience frequent misinterpretation of dimensionality-reduced views because researchers tend to rely too heavily on projected plots which they consider as evidence [30], [31].

The research findings confirm that projection-induced misinterpretation establishes a valid methodological challenge which affects offshore wind clustering operations. The original high-dimensional space contains well-defined clusters which become indistinguishable from each other when data is reduced to 2D or 3D embeddings. The incorrect analysis of regime characteristics becomes possible when analysts view these clusters because they appear to be close to each other or they seem to overlap with each other or their boundaries are not distinct. The failure mode functions independently from visualization artifacts because it emerges when high-dimensional clustering intersects with the methods which researchers use to evaluate stability and interpret results. Research on offshore wind power shows that multivariate regime discovery methods deliver useful results but most studies fail to implement protection mechanisms which prevent projection errors and they do not compare predicted relationships to actual high-dimensional distances [10]–[12].

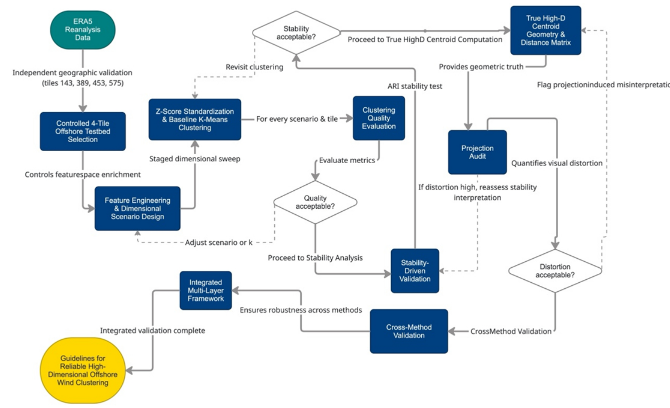
The results from Table 1 demonstrate that researchers lack a standard approach to their work. Research on offshore wind clustering has developed its ability to analyze multiple variables and interpret data through regime-based methods but it needs a complete system which handles feature-set expansion as an experimentally controlled variable and focuses on stability patterns to prove system reliability and performs projection distortion assessment for interpretation validation. The stability literature establishes effective methods for unsupervised validation which researchers can apply to their work [3], [9] while the visual analytics literature demonstrates through multiple studies that interpretability risks exist and systems need to meet reliability standards [6],[29], [30], [31]but these fields have not been integrated in offshore wind research applications. Physics-informed interpretability is introduced to strengthen the validity of results and interpretations obtained from high-dimensional offshore wind clustering analysis. The workflow begins with a unified statistical framework to obtain the cluster assignments and the estimated parameters of the wind models from the high-dimensional data. Then, dimension-aware scenario design, stability - driven evaluation and explicit projection audits are combined to validate the reliability of the obtained models and therefore the obtained cluster assignments and wind models.

3. Materials and Methods.

This contribution investigates how increasing the dimension of the feature space affects the structure, stability of clusters as well as the reliability of a visual analysis technique that relies on distances for the analysis of offshore wind data. An experiment is conducted following a scenario-based approach. All steps of the data analysis are scenario invariant and each scenario is defined based on a different dimensionality. Therefore, any effects that are observed can be directly attributed to increasing the dimensionality of the data. No effects are introduced by the clustering algorithm that is used for all scenarios. Finally, the data scenarios are defined based on data samples that were measured at highly dispersed offshore locations.

The proposed method consists of five stages: (i) the design of an offshore testbed of tiles using an automated method, (ii) the combination of features corresponding to staged dimensional scenarios using physical characteristics of wind and atmosphere, such as wind speed, turbulence intensity, temperature, and humidity, (iii) the baseline clustering for the desired number of clusters for each scenario and tile, (iv) the stability assessment of the clusters obtained for different initial seeds using a bootstrap resampling method, and (v) the true centroids in high dimensionality and their pairwise distances for auditing the 2D visualizations. The conceptual workflow of the proposed method is shown in Figure 1. The main stages are denoted by the solid arrows and the validation loops denoted by the dashed lines are triggered for cases of low quality, unstable clusters, or high distortion exceeding the predefined thresholds.

Figure 1: A conceptual overview of the workflow.



Source: Authors.

The basis for the underlying data is ERA5 atmospheric reanalysis data from the European Centre for Medium - Range Weather Forecasts (ECMWF). ERA5 provides uniformly and physically consistent global measurements of the parameters of the atmospheric wind field, which has been validated and used in many applications in the field of offshore wind energy. [32], [32].

As part of wider project, a global offshore dataset has been collated comprising 134,483 grid points and 12,552 timepoints of data, stored in the format of a rich NetCDF file. The offshore points are indexed using the "tile_the_globe" tiling scheme with a $10^\circ \times 10^\circ$ grid allowing for both data analysis of the grid points and efficient regional sampling.

This independent methodological study utilized a new controlled testbed that was not associated with any of the prior studies. For this analysis, four offshore tiles were selected for each of the 4 latitudinal bins (tropical, mid-latitudinal Northern Hemisphere, mid-latitudinal Southern Hemisphere, and high-latitude). The tiles were selected using the following criteria: 1. Candidates tiles were filtered such that a minimum of 500 offshore points were required. 2. A weighted score was calculated for each of the remaining tiles in each latitudinal bin. The weights for each of the three score components were the normalized point counts, the normalized mean wind speed, and the normalized wind speed variance. 3. The tiles were ranked within each latitudinal bin using the weighted score. From each latitudinal bin 4 tiles were selected such that the great circle distance between any two tiles in the testbed was as large as possible. The tiles for the independent testbed are listed in Table 2.

Features are constructed by taking physically meaningful combinations of wind and atmospheric features from the input data. All features are used in the same way for all tiles. The base set of wind and atmospheric features includes: - Wind speed and wind direction - Air density and temperature and pressure at 100m above ground level. The number of features is kept low by only taking one sample from the time series with

Table 2: Controlled independent offshore testbed tiles.

Tile ID	Latitudinal band / region label	Rationale (methodological)	n_points (offshore grid points)
143	MIDLAT_SH	Mid-latitude Southern Hemisphere regime	1,693
389	TROPICAL	Tropical wind regime	673
453	MIDLAT_NH	Mid-latitude Northern Hemisphere regime	569
575	HIGHLAT	High-latitude regime; large-sample stress-test	7,101

Source: Authors.

a fixed stride. This gives us an effective temporal resolution of one data point per day.

Seasonal means and standard deviations of the base variables are calculated at each of the offshore locations for the four climatological seasons: DJF (December, January, February), MAM (March, April, May), JJA (June, July, August)and SON (September, October, November). In addition, the robust wind-speed indicators (10th percentile, 90th percentile and interquartile range) and the directionality measure (variability of the wind direction) are derived from the subsampled time series.

By incorporating the design features into the configuration, a staged set of dimension scenarios is created to achieve a controlled dimension sweep. The dimension scenarios with the integrated design features are given in Table 3.

Table 3: Dimensional scenarios (conceptual feature sets) used for the dimension sweep.

Scenario	Conceptual dimensionality	Feature Tokens	Description
S2D	2	wspd, wdir	Seasonal mean and standard deviation of wind speed and wind direction
S3D	3	wspd, wdir, ρ	S2D + seasonal statistics of air density
S5D	5	wspd, wdir, ρ, T100, p100	S3D + seasonal statistics of temperature and pressure
S7D	7	S5D + wspd_p90, wspd_p10	S5D + wind-speed percentiles
S9D	9	S7D + wdir_var, wspd_iqr	S7D + direction variability and wind-speed IQR

Source: Authors.

The fact that the engineered features are in different physical units and scales means that for all distance based clustering and stability evaluation the data has to be z-score standardized. The reason for this is that a variable with a larger range should not dominate the distance computation because of its scale whereas other variables with smaller ranges but large effects in Euclidean distance based learning [33] should not have lesser effect. The z-score standardizes the value.

$$z_i = \frac{x_i - \mu}{\sigma} \tag{1}$$

Where $\boldsymbol{x_i}$is the raw feature value, μ is the feature mean, σ is the feature standard deviation and $\boldsymbol{z_i}$ is the standardized value.

Baseline clustering is performed for each tile, for each dimension, with cluster number fixed to K=3 and K=5. The centroids are calculated in the standardized feature space. The quality of clusters is evaluated using a number of internal validity criteria. Most of the criteria used are based onSilhouette coefficient [34], the Davies–Bouldin index [35], and the Calinski–Harabasz criterion [36]. These criteria are calculated on a fixed-size (deterministic) subset of the data, being proportional to the smallest sample size in the case of different sample sizes (i.e. in the case of unequal sample sizes, the criteria are calculated on a subset proportional to the smallest sample size in order to have the same number of data points to calculate the criteria for all cases and for all tiles). The baseline settings for clustering are shown in Table 4.

Table 4: Baseline clustering configuration.

Item	Value
Cluster numbers	(K = 3, 5)
Feature space	Scenario-specific standardized features
Distance metric	Euclidean
Quality indices	Silhouette, Davies–Bouldin, Calinski–Harabasz
Metric evaluation	Deterministic subsampling

Source: Authors.

The stability of clustering results obtained by different initializations is evaluated using the Adjusted Rand Index (ARI) [37], a measure that quantifies the degree of similarity between two partitions while correcting for random cluster assignment of instances. Two complementary stability modes are provided.

Seed stability measures the influence of the initializations by calculating the pairwise average of the adjusted Rand index (ARI) between the clusterings produced by different random seeds on the same dataset. On the other hand, the bootstrap stability measures the influence of the sampling noise in the data by iteratively performing the following steps a number of times: i) sampling a fixed proportion of instances with replacement, ii) applying a clustering algorithm to the new subset of instances, and iii) calculating the ARI between the obtained clustering and a reference clustering. Table 5 summarises the stability configuration employed in this work.

Table 5: Stability analysis configuration.

Item	Value
Stability metric	Adjusted Rand Index (ARI)
Seed runs	30
Bootstrap runs	30
Bootstrap fraction	0.80
Standardization	z-score

Source: Authors.

In order to obtain a stringent reference for distance-based analysis, the true high-dimensional cluster centroids for all (tile, scenario, K) configurations are computed. The centroid for each cluster c is given by:

$$\mu_c = \frac{1}{n_c} \sum_{i \in c} x_i \tag{2}$$

where $\boldsymbol{x_i}$ is the engineered feature vector of point i and $\boldsymbol{n_c}$ is the number of points in cluster c.

The pairwise distances between cluster centroids are computed using the Euclidean distance metric in the full engineered feature space.

$$d_{\text{true}}(c_a, c_b) = \sqrt{\sum_{m=1}^{D} (\mu_{c_a,m} - \mu_{c_b,m})^2 + \varepsilon} \quad (3)$$

where D is the dimensionality of the scenario at hand and $\epsilon = \mathbf{10^{-12}}$ is a fixed numerical constant introduced to ensure numerical stability.

For every tile, scenario and value of K, the distances between the centroids are calculated. The top and bottom closest/farthest pairs are stored for later use in the projection induced misinterpretation analysis.

This check performs a projection audit to verify that the low dimensional projections of the clusters do not alter the underlying relationships between them. The reference is the true high dimensional centroid distances. The audit is performed uniformly across all tiles and scenarios. The purpose of this audit is to identify any potential projection artifacts and biases that could lead to incorrect conclusions, and is not related to the quality of the clustering results.

Matrix of true high-dimensional centroids $\boldsymbol{C \in R^{K \times D}}$, where K denotes the number of clusters and D denotes the dimensionality of the scenario. The first step to compute the pairwise Euclidean distances between centroids in the original feature space. This yields the reference distance matrix $\boldsymbol{D_{true}}$.

Centroids are then projected into two- and three-dimensional spaces using principal component analysis (PCA), producing projected centroid matrices $\boldsymbol{C^{(2)}}$ and $\boldsymbol{C^{(3)}}$. The corresponding pairwise distance matrices in the projected spaces are denoted by $\boldsymbol{D^{(2)}_{proj}}$ and $\boldsymbol{D^{(3)}_{proj}}$, respectively.

Global geometric distortion introduced by projection is measured by Kruskal's stress which assesses the normalized distance between the true and the projected distances. Stress is calculated as [38]:

$$\text{Stress} = \sqrt{\frac{\sum_{i<j} \left(d_{ij}^{\text{true}} - d_{ij}^{\text{proj}}\right)^2}{\sum_{i<j} \left(d_{ij}^{\text{true}}\right)^2 + \varepsilon}} \quad (4)$$

where $\boldsymbol{d_{ij}^{true}}$ and $\boldsymbol{d_{ij}^{proj}}$ denote the true and projected centroid distances between clusters i and j, respectively, and $\epsilon = \mathbf{10^{-12}}$ is a fixed numerical constant introduced for numerical stability.

Relative inter-cluster ordering preservation under projection check To check if relative inter-cluster ordering is preserved under projection, rank-based and linear correlations are computed between the upper-triangular elements of $\boldsymbol{D_{rue}}$ and $\boldsymbol{D_{proj}}$. Specifically, Spearman's rank correlation coefficient and Pearson's correlation coefficient are employed [39], [40]:

$$\rho_s = \text{corr}_{\text{Spearman}} \left(d_{\text{true}}, d_{\text{proj}}\right) \quad (5)$$

$$\rho_p = \text{corr}_{\text{Pearson}} \left(d_{\text{true}}, d_{\text{proj}}\right) \quad (6)$$

where $\boldsymbol{d_{true}}$ and $\boldsymbol{d_{proj}}$ are vectors containing all unique inter-centroid distances.

Global performance criteria have been introduced and pairwise distortion for each centroid pair (i, j) has been measured by three related quantities.

First, the distance ratio is defined as:

$$\boldsymbol{Ratio_{ij} = \frac{d_{ij}^{proj}}{d_{ij}^{true} + \epsilon}} \quad (7)$$

Second, the relative error is computed as:

$$\boldsymbol{RelErr_{ij} = \frac{\left|d_{ij}^{proj} - d_{ij}^{true}\right|}{d_{ij}^{true} + \epsilon}} \quad (8)$$

Finally, to explicitly capture visual misinterpretation, a misleading closeness index is introduced:

$$\boldsymbol{Mislead_{ij} = \frac{d_{ij}^{true}}{d_{ij}^{proj} + \epsilon}} \quad (9)$$

Large values of $\boldsymbol{Mislead_{ij}}$ mean that there is large distance between two centroids in the high dimensional space, but the distance between the corresponding centroids in the projected space is relatively small, and this is an example of misinterpretation due to the projection effect. For each case, the centroid pairs are ranked based on the $\boldsymbol{Mislead_{ij}}$ measure and the most misleading cases are chosen for closer examination and visualization.

4. Results.

The subsampled silhouette score was used to evaluate the clustering quality for each of the staged dimensional scenarios (S2D–S9D) and for the four representative offshore tiles. The full-case matrix of results is not reported here, and is therefore only presented in the form of highlight cases in Table 6, along with the full results provided in the exported quality summaries. Table 6 presents the highlight cases, comprising: the best dimension for each tile; and the most diagnostically informative cases to fully capture the dimension-aware design of the study.

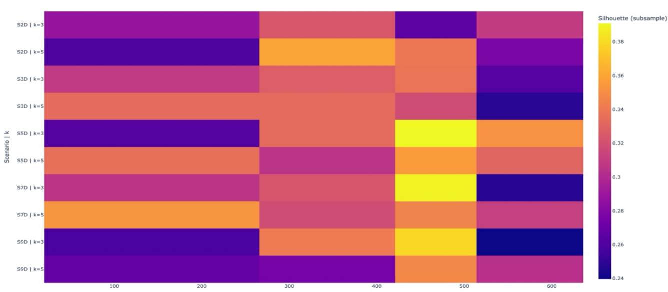
Table 6: Best-performing dimension-scenario configurations across tiles based on subsampled silhouette score.

Tile	Highlight case (scenario, k)	Silhouette (subsample)	Interpretation-relevant note
453	S5D, k=3	0.3909	Highest observed quality among stored highlights; indicates strong separation in an intermediate-dimensional setting.
453	S7D, k=3	0.3891	Comparable to S5D, supporting the "intermediate dimensionality" regime as consistently strong for this tile.
389	S2D, k=5	0.3594	Best stored highlight for the tropical tile; demonstrates that low-dimensional settings can be competitive in specific regimes.
143	S7D, k=5	0.3527	A strong case for the mid-latitude SH tile at higher k, later used in the projection-audit evidence.
575	S5D, k=3	0.3521	Best stored highlight for the high-latitude large-sample tile; supports intermediate dimensionality even under large n.
575	S3D, k=3	0.2630	Example of degradation relative to S5D within the same tile, motivating scenario sensitivity.
575	S9D, k=3	0.2400	Further decline at the highest-dimensional scenario (stored highlight), indicating that "more dimensions" does not guarantee better separability.

Source: Authors.

Looking at the full-resolution surface in Figure 2 provides further insight into the trends observed in the reduced images. First, Tile 453 is always located in the highest quality region of the surface except in the S5D and S7D scenarios. Second, Tile 575's surface is very sensitive to the choice of scenario and the quality of the outer regions in the lower- and higher-dimensional ends of the surfaces are much lower than the peak. Third, it is clear from the heatmap that the best clustering quality regions are not always located at the right end of the dimensionality sweep, suggesting that the highest dimensionality is not always the best.

Figure 2: Heatmap of subsampled silhouette scores across tiles, dimensional scenarios, and cluster-number settings.



Source: Authors.

To analyze the behavior of the clustering quality as the dimensionality increases, Figure 3 shows the silhouette evolution for each tile and k for each scenario-wise trend. This allows us to verify if the quality evolves in a monotonous fashion or if it has peaks at some intermediate dimensionality values.

Figure 3: Scenario-wise variation of subsampled silhouette score across tiles, with separate trends for k=3and k=5.

Source: Authors.

Figure 3 presents the same cross-scenario information as Figure 2, but in a more interpretable manner. For Tile 453 the line trends peak at intermediate dimensions in the S5D–S7D range, as one would intuitively expect for the optimal region for a good balance between the discriminatory power of the features and the robustness of the tree. For Tile 389 the best model for this tile in the sweep was S2D with k=5, which means that this tile was learned to a good level of accuracy in at least some of the tropical environments under study. For Tile 143 the strong value at S7D, k=5 shows that high dimensional feature spaces can lead to good clustering metrics, with the subsequent terms representing the impact of the projection-misinterpretation terms. The decline for Tile 575 towards S9D suggests that high dimensional feature spaces do not always lead to high discriminative power of the resulting representation.

Table 6 and Figures. 2–3 demonstrate the first core result of this paper. First, the performance of dimension-aware clustering is more influenced by the ordering of dimensions than by the number of dimensions. Second, the middle dimensional setting often achieves the best balance between accuracy and computational speed, while increasing the number of dimensions does not necessarily lead to a better clustering solution due to model overfitting. The findings of this experiment motivate the following stability analysis.

This paper investigates stability of clustering under controlled noise and whether stable clustering leads to non-trivial partitions in the data feature space. The stability of the clusterings is measured, obtained using the Adjusted Rand Index (ARI) metric under two types of perturbations. One is the sensitivity to the random seed initialization (seed stability) and the other is the robustness against different bootstrap resamples of the data (bootstrap stability). The analysis particularly focuses on the S7D data set which had the highest clustering quality in the intermediate dimensions (n = 7)

Seed stability is also calculated for each of the four representative tiles. The high values of the ARI in all cases suggest that solutions derived by applying the clustering algorithm to many different random initializations are very similar. In fact, seed stability is almost perfect for at least one value of k for the mid-latitude tiles. For the other tiles, the values are always very high indicating that the clusters derived from different initializations do not differ from each other. In other words, in S7D the clusters are stable and the resulting solutions are always the same.

Bootstrap stability analysis yields a closely aligned pattern. The stability of the classification obtained by repeatedly resampling a fixed fraction of the data and re-estimating clusters is also documented. The ARI values for the resulting sets of clusters are high for all tiles, with only modest drops relative to seed stability. The high values of the ARI for the highly stable inferred partitions that capture the sampling variability (i.e., stability under bootstrap perturbations) obtained for all tiles, which are very different in terms of sample size and climatological characteristics, further document the robustness of the results.

While the high stability of seeds under perturbations together with the stability of the corresponding bootstrap clusters is a good indication of the stability of the clusters, it does not mean necessarily that there is any non-trivial geometry between the clusters. The stability measures the robustness of the assignments of the samples (seeds) to the clusters (i.e., the stability of the clustering) under perturbations. However, it does not reflect the representativeness of the embedded mapping to the underlying data distribution or the accuracy of the distances between the clusters in the data space. A few examples illustrating such effects are provided in the following; they demonstrate that highly stable clustering results can be associated with strongly distorted 2D or 3D visualizations.

The results show that while stability is a necessary requirement for an interpretable clustering model, it is not sufficient. In particular, taking the same data set and feeding it to the same algorithm and obtaining the same partition at every split in the decision tree for S7D does not result in an interpretable clustering, except that the clustering is stable. The following analysis deals with the true high-dimensional centroid distances and projection distortions.

The S7D condition provides a high degree of stability in terms of reproducibility of clustering solutions for almost all the tiles and cluster-number conditions with a few exceptions that are described by the degree of stability in Table 7.

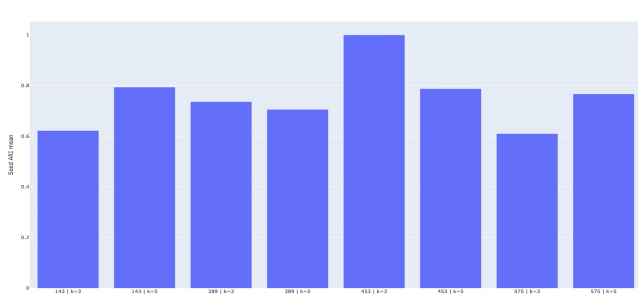
Table 7: Seed and bootstrap stability (ARI mean) for the intermediate-dimensional scenario (S7D).

Tile	(k)	Seed ARI mean	Bootstrap ARI mean	Interpretation-relevant note
143	3	0.6219	0.5603	Moderate-to-strong stability, but lower than the best-performing tile–(k) combinations.
143	5	0.7930	0.8395	Stability improves substantially at higher (k), indicating a more reproducible partition for this case.
389	3	0.7331	0.5754	Strong seed stability, with some reduction under resampling.
389	5	0.7099	0.6722	Stable under both perturbation types, though below the strongest cases.
453	3	0.9986	0.9819	Near-perfect reproducibility and robustness; the most stable stored S7D case.
453	5	0.7832	0.7177	High but clearly lower than the (k=3) solution for the same tile.
575	3	0.6078	0.5829	Moderate stability despite the large sample size.
575	5	0.7666	0.8900	Strong reproducibility and especially strong robustness under bootstrap perturbation.

Source: Authors.

The strongest stored case for this condition is for Tile 453 at seed ARI = 0.9986 and bootstrap ARI = 0.9819 (k=3), representing one of the almost perfectly reproducible clustering solutions under both types of perturbations. As this tile has also been shown to be one of the high-quality tiles, this also serves as a further example of a clustering condition that gives high-quality clustering solutions in addition to the stable solutions under the given perturbations. On the other hand, Tile 143 at k=3 and Tile 575 at k=3 have somewhat lower values indicating that there are a few more cases that, based on a purely visual or structural evaluation, seem also plausible, but which do not hold under the given perturbations. Figure 4 Seed-stability distribution for the S7D cases. Please refer to Table 7 for the actual numbers.

Figure 4: Mean seed stability (ARI) across tiles and cluster-number settings for the intermediate-dimensional scenario S7D.

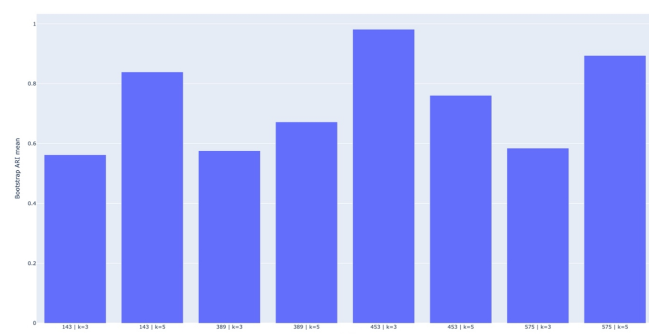


Source: Authors.

Seed Stability within the S7D metric for the cases investigated in Figure 4. The high values of the S7D metric that were derived from all the tile–k combinations have led to the conclusion that the seeds are stable for all tile–k combinations. While the large peak corresponding to Tile 453, k=3 is almost perfectly repeated, the small peak is found at the lower end of the S7D seed-stability range for Tile 143, k=3 and Tile 575, k=3. Hence k=3 leads to high repeatability in the former case, but not in the latter.

Figure 5 Initialization sensitivity alone is not sufficient to explain the sampling robustness of the solutions; The stability analysis under bootstrap resampling is needed to understand the stability of the solution partitions obtained under different initial conditions, i.e. for different realizations of the sampling procedure; The same S7D experiments as in Figure 5 are shown.

Figure 5: Mean bootstrap stability (ARI) across tiles and cluster-number settings for the intermediate-dimensional scenario S7D.



Source: Authors.

Figure 5 It appears that bootstrap robustness generally follows the pattern of seed stability with a few exceptions. Again, Tile 453, k=3 has a very stable clustering solution that is both stable to different seeds (initialisations of the algorithm) and stable to different resamplings of the data. Interestingly the bootstrap ARI is higher for the clusters in Tile 575, k=5 than the seed ARI indicating that while this clustering is not particularly stable to changes in the initialisation of the algorithm (cluster assignments are swapped when different seeds are used) it is very stable to resampling of the data. Similarly Tile 143, k=5 has a bootstrap ARI of 0.8395 and will be also be considered in the audit phase following the projection analysis.

The results in Table 7 and in Figures. 4 and 5 are the second core contribution of this paper. Besides the clustering performance, there are other criteria that may influence the choice of the optimal clusterization dimensionality. It has been shown in several experiments in Table 7 that for certain values of the parameters, i.e., for certain values of the dimensionality (i.e., certain intermediate values of the parameter), the results are repeatable and stable. For instance, for Tile 453, k=3 and for Tile 143, k=5. In this case, although the clusters are stable when the clusterization is perturbed by noise according to the ARI stability measure, the stability and ARI measures do not imply significant centroid distance between the clusters, nor a visually correct clustering. It is necessary to return to the discussion of the centroid distance concerning the clusterization performance. High clustering performance is expected when the clusterization assignments are stable and the ARI measure is high, however this does not exclude large distortions in the centroid distances of the clusters when projecting from the high dimensional space to a lower dimensional space for the specific pairs of clusters with high clustering performance and assignments stability.

All previous analyses showed that many of the intermediate dimensions lead to a high clustering quality and stability. However, the clustering quality and stability do not at all imply that the resulting cluster structure will also be preserved after projecting it to the 2D visual space of a PCA transformation. The following analysis will therefore take a closer look at the feature space engineered for an understanding of the cluster relationships and how much of this structure is preserved under a low dimensional PCA projection.

In order to determine the true geometric reference is obtained by calculating the inter-centroid distances in the full high-dimensional feature space for each (tile,scenario,k). These distances are then compared to their projections in PCA-2D and PCA-3D spaces. The distortion between the true and projected distances is measured using the projected-to-true distance ratio, relative error and the mislead_close factor. Table 8 shows the distortion cases in a compact way, listing the two most severe distortion cases for each of the two spaces, namely the most severe PCA-2D distortion and the most severe PCA-3D distortion.

Table 8: Extreme projection distortion cases: true high-dimensional vs projected centroid distances.

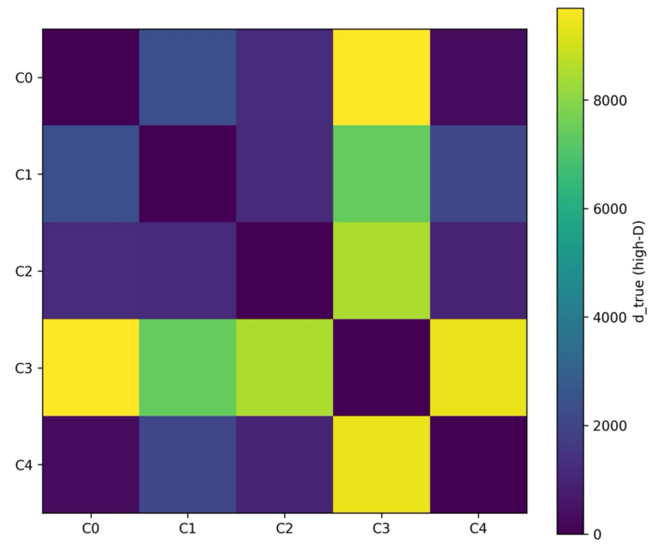
case_id	tile_id	scenario	k	cluster_i	cluster_j	d_true	d_proj	ratio	relerr	mislead_close	method
tile143 S9D k5	143	S9D	5	C0	C4	4.0334	1.2216	0.3028	0.6971	3.3017	PCA2
tile143 S7D k5	143	S7D	5	C0	C4	5.8665	3.5209	0.6001	0.3998	1.6661	PCA3

Source: Authors.

Table 8 shows the distortion index for the most distorted cluster pair. The pair in Tile 143, S9D, k=5 in PCA-2D was denoted as C0–C4, and is on average three plus times closer to each other than in the original feature space. Another representative example is Tile 143, S7D, k=5 in PCA-3D, where the distortion is not extreme, but is quite significant nonetheless. The true inter-centroid distance in the original space is also significantly compressed in these two examples. These two examples are provided to demonstrate that structural distances can be heavily distorted in low-dimensional projections while keeping the underlying clustering solution very strong.

Interpretation of distortion in a geometric context is provided in Figure 6, showing the true high-dimensional centroid-distance matrix for Tile 143, S9D, k=5. The purpose here is to confirm that the clusters which were highlighted in the bottom plot of Figure 5 are indeed well separated in the original feature space before any projection took place.

Figure 6: True high-dimensional centroid-distance matrix for Tile 143 under S9D with k=5.

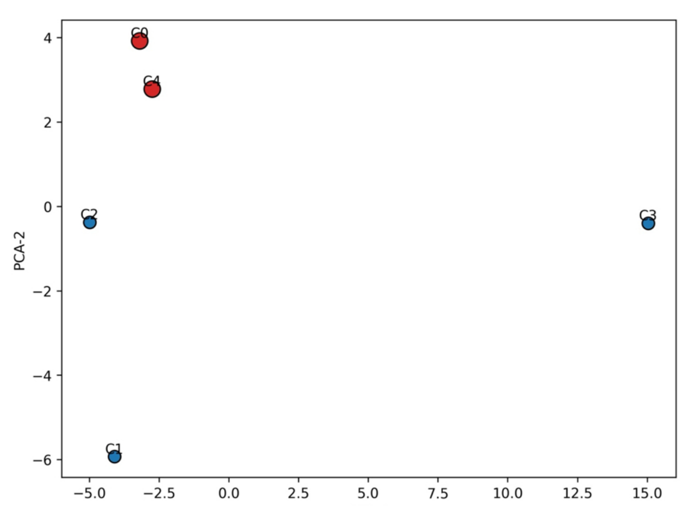


Source: Authors.

Figure 6 confirms that the selected centroid pair is not a visually chosen artifact; rather, it belongs to a configuration with substantial true inter-centroid separation in the original space. In particular, the heatmap makes clear that the highlighted structure is embedded within a broader distance pattern in which some cluster pairs are genuinely far apart. This establishes the high-dimensional geometric baseline needed to judge whether subsequent 2D or 3D projections remain faithful.

Having established the true-distance structure, Figure 7 shows the corresponding PCA-2D centroid map for the strongest 2D distortion case, again for Tile 143, S9D, k=5, with the pair C0–C4 highlighted.

Figure 7: PCA-2D centroid map for Tile 143 under S9D with k=5, highlighting the C0-C4 pair.

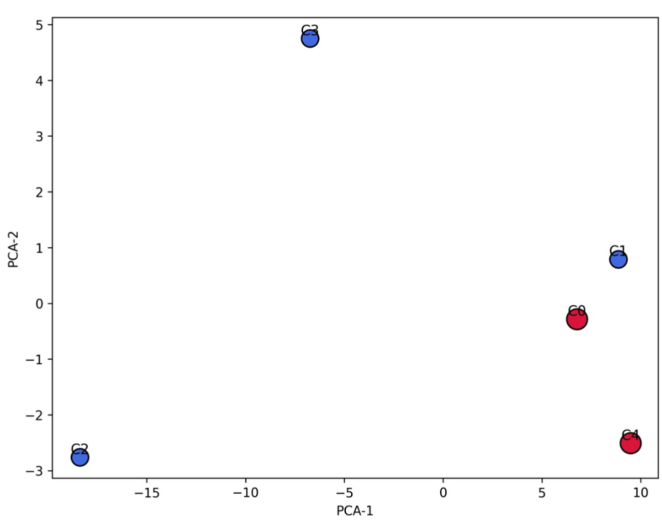


Source: Authors.

Figure 7 illustrates the core failure mode for this paper. As shown with highlighted pair C0–C4, true high-dimensional centroid distances between clusters are larger than what is derived from their two-dimensional (2D) PCA projections. In particular, it can occur that, although there is no indication of proximity between two clusters in the feature space, they appear to be near each other when visualizing the PCA-2D projection. Consequently, a reader relying solely on the PCA-2D projection would conclude that there is a high probability of partial overlap or only weak discrimination between the pairs of clusters that appear to be close to each other. This is an instance of projection-induced misinterpretation, identified here as a potential methodological pitfall in many clustering workflows defined in the high-dimensional feature space.

To address the potential objection that a three-dimensional projection should substantially resolve the compression observed in two-dimensional space, the matched S7D case is examined in both dimensions. Figure 8 presents the PCA-2D centroid map for Tile 143 under S7D with k=5, highlighting the centroid pair C0–C4, while Figure 9 presents the corresponding PCA-3D centroid map for the same configuration. This direct comparison reveals the extent to which increasing the projection dimension mitigates — but does not fully eliminate — the observed distortion.

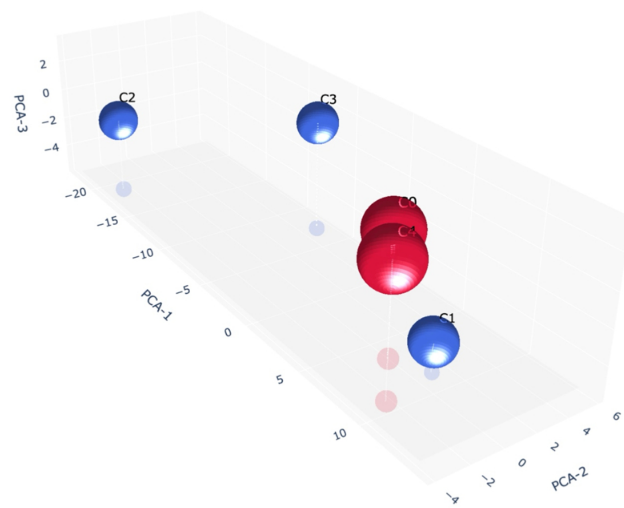
Figure 8: PCA-2D centroid map for Tile 143 under S7D with k=5, highlighting the C0-C4 pair.



Source: Authors.

Together these figures show that a 3D projection (or higher) of the data can better separate the points of a pair in the mapping feature space using the mapping feature space visualization tool, but still not enough to preserve the true separation in the original feature space. The pair C0–C4 in Figure 8 are almost indistinguishable from each other in the 2D projection, which could be an indication of weak separation between the underlying structures. The 3D version in Figure 9 alleviates the compression effect and shows better separation of the centroids. However, the pair is still more compact in the projected space than they are in the original high-dimensional space, as shown in Table 8. The S7D, k=5, result is therefore still of interest and

Figure 9: PCA-3D centroid map for Tile 143 under S7D with k=5, highlighting the C0-C4 pair.



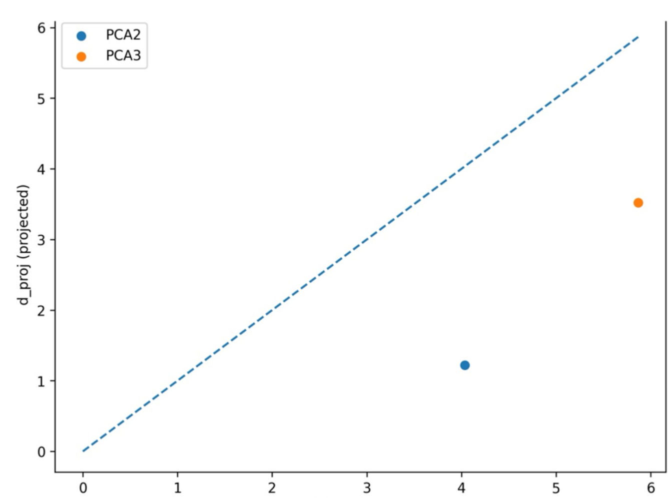
Source: Authors.

needed to compare the PCA-2D and PCA-3D projections.

This comparison supports two of the methodology arguments in this study. First, the projection-induced misinterpretation shown in the PCAs is not unique to the 2D plot format. Second, even though 3D representations may offer better visualization of the data, they should not be considered as being equal to the true data geometry. So PCA-3D is only a better graphical tool than PCA-2D, but it is still far from representing the true centroid distances in the original high dimensional space, and therefore is not an appropriate metric to be used in this context.

For a more direct comparison, the strongest retained evidence cases are plotted in the (d_{true},d_{proj}) space in Figure 10, together with the identity line that represents ideal preservation of distances.

Figure 10: True versus projected centroid distances for the strongest retained distortion cases.



Source: Authors.

Figure 10 summarizes in a concise form the quantitative visualization of the phenomenon shown in Table 8. Both pieces of retained evidence lie below the identity line which confirms that the estimated distances are under-estimated. The under-estimation is more pronounced for PCA-2D than for PCA-3D. However, the estimated distances in the case of PCA-3D are closer to the identity line than those obtained for PCA-2D and are still under-estimated. Figure 10 links the quantitative information in Table 8 to the failure of distance preservation.

In particular, the fact that there are two examples of retained distortion implies that these distortions happen in good data – i.e., data that were not flagged as weak by the previous analysis. For example, the S7D data cluster was flagged as being one of the strongest data clusters in terms of quality–stability of the stored data and the S9D data cluster was one of the outputs of the dimension-aware evaluation framework for controlled clustering. This means that the distortion in these cases is not a byproduct of poor clustering or unstable partitioning. Rather, it indicates a deeper, inherent property of the technique – that high quality clustering and high clustering stability do not necessarily imply faithful representations of the data in lower dimensionality.

This analysis concludes the fourth core contribution of this study. Projection-induced misinterpretation is a non-trivial visualization failure mode in dimension aware clustering workflows. It remains to confirm the validity of the assertion that interpreting high-dimensional atmospheric regimes by means of low-dimensional projections lead to visually appealing, but feature space geometry misleading conclusions. This will be achieved by collecting cross-method evidence to verify whether the distortion caused by projection-induced misinterpretation is exclusive to the employed clustering methodology.

The purpose of the final validation step was to verify if the distortions, introduced by the projection, detected in the previous step are method dependent, i.e. if they were present in the baseline clustering method or if they will also be present in other clustering methods. For Tile 143 and for the higher dimensional spaces S7D and S9D, a cross-method audit was performed using two other density based clustering methods, DBSCAN and OPTICS. The new clusters found by applying these methods were then validated by applying the projection audit metrics used for the baseline clustering method.

Summary of cross-method results for all four cases is given in Table 9 (a, b), where clustering and projection-audit metrics are condensed. The table shows the number of clusters, the noise fraction, the average internal clustering coefficient, the projection stress, the rank preservation, and the largest mislead_close value of PCA-2D and PCA-3D.

Table 9 & 10 are Cross-method projection audit results for Tile 143 using DBSCAN and OPTICS under S7D and S9D.

Table 9: Core clustering results.

Tile ID	Scenario	Method	Points (n)	Dimensions	Clusters	Noise fraction	Silhouette (non-noise)
143	S9D	DBSCAN	1693	44	9	0.051	0.241
143	S9D	OPTICS	1693	44	5	0.096	0.277
143	S7D	DBSCAN	1693	42	9	0.048	0.243
143	S7D	OPTICS	1693	42	5	0.073	0.281

Source: Authors.

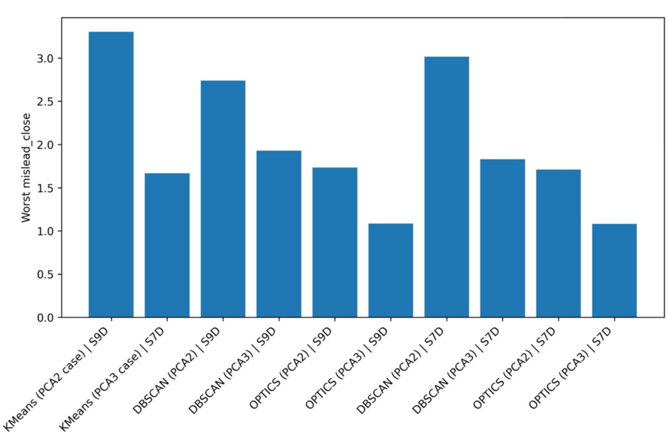
Table 10: Projection distortion audit metrics.

Tile ID	Scenario	Method	PCA-2 Stress	PCA-2 Spearman	PCA-2 Worst mislead_close	PCA-3 Stress	PCA-3 Spearman	PCA-3 Worst mislead_close
143	S9D	DBSCAN	0.134	0.993	2.737	0.078	0.996	1.928
143	S9D	OPTICS	0.106	0.964	1.733	0.027	1.000	1.085
143	S7D	DBSCAN	0.136	0.991	3.014	0.078	0.990	1.830
143	S7D	OPTICS	0.104	0.952	1.707	0.026	0.988	1.081

Source: Authors.

Table 9 & 10 confirms that projection distortion remains clearly visible across all clustering algorithms. Even when density - based methods (DBSCAN and OPTICS) are employed instead of the centroid-based clustering used in the baseline, the high-dimensional distances continue to be substantially compressed in PCA-2D, with all mislead_close values reaching the worst levels in each case and remaining far from 1. Furthermore, the cross-method worst mislead_close value exceeds 3, consistent with the baseline results, thereby indicating a comparable degree of underestimation. Figure 11 provides a direct visual representation of the strongest projection-distortion magnitude that occurred between any two methods in the baseline and cross-method evaluations. It is important to note that distortion is not an artefact of a single clustering algorithm.

Figure 11: Worst projection-induced compression across baseline, DBSCAN, and OPTICS cases for Tile 143.



Source: Authors.

Figure 11 confirms that the compression caused by the projection is not eliminated by changing the method. The baseline distortions are among the largest distortions, but the density-based methods all have larger distortion values, and large distortion magnitudes are also observed in the PCA-2D method. The largest cross-method distortion value in S7D is caused by DBSCAN. While OPTICS can reduce the distortion caused to the points by DBSCAN, it does not eliminate it. Therefore the distortion is not an artefact of the particular partitioning method used, but a real effect due to the low-dimensional representation that can cause unusual geometric situations to occur.

A second observation from the table is that higher clustering quality does not directly translate to lower projection distortion. Although silhouette values for the OPTICS algorithm are higher than those for DBSCAN for almost all the datasets, the latter still undergoes relatively low dimensional compression. This again shows that the internal clustering quality and the projection distortion are not necessarily correlated in a monotonic fashion. In particular, it may happen that a clustering algorithm produces a cluster assignments for which the clustering criteria indicate a good clustering, yet the visual representation of the data provided by the projections is highly distorted.

The comparison between PCA-2D and PCA-3D is still relevant. In all the cases, PCA-3D has lower stress and weaker (though not always the strongest) distortion compared to PCA-2D. However, the projection error is not yet eliminated. Consequently, inter-cluster distances continue to be collapsed within the visual embedding owing to the inherent limitations of the low-dimensional layout, even with the addition of a third projection dimension.

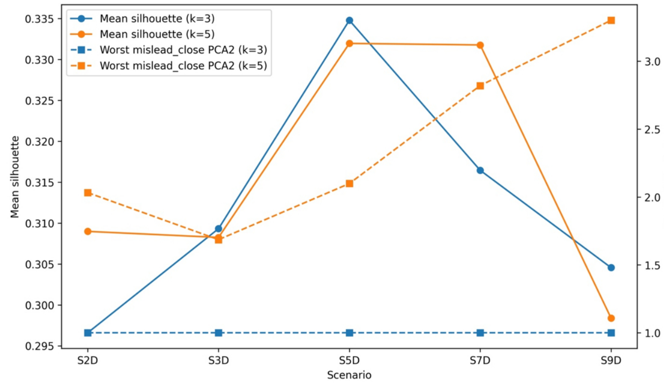
Table 9 & 10 and Figure 11 together provide the cross-method evidence required to support the final methodological claim of this analysis. Projection-induced misinterpretation is not an artifact of a specific clustering algorithm. Rather, it arises from the inherent difficulty of projecting structured high-dimensional atmospheric feature spaces into two or three dimensions without losing important geometric relationships. This finding strengthens the central argument of the paper by demonstrating that the distortion persists across both centroid-based and density-based clustering paradigms.

The cross-method evidence for the 10th methodological claim is provided in Table 9 and Figure 11. The distortion caused by the projection is not caused by the clustering algorithm used. The dimensionality reduction technique causes loss of information when the centroids or density estimates in the original high-dimensional feature space are being projected into the lower dimensional space. This has been verified by applying both centroid-based and density-based clustering methods in the present study.

Figure 12 shows the summary relation between the dimension - aware quality trends and distortion evidence derive from analyzing the scenario evolution, in terms of the average clustering quality and the strongest retained projection distortion.

Figure 12 illustrates the methodological trade - off that is discussed in this paper. The lower dimensional clusters generally lead to higher clustering quality on average, but high dimensional data does not necessarily translate to more meaningful lower dimensional representations. The dimension-aware approach improves the expressiveness of the clustering analysis, while there is a need to understand the effect of the higher dimensionality on the preservation of the shape properties of the clusters in the projection. In summary, dimension-aware clustering, stability, and projection audit all require a holistic consideration.

Figure 12: Dimension?scenario trends linking average clustering quality and worst projection distortion.



Source: Authors.

5. Discussion.

The result demonstrates that the dimensionality of the feature space is a design parameter which strongly affects the clustering of the offshore wind. Increasing the dimensionality of the space does not necessarily result in an increase of the clustering performance. Higher dimensional spaces always contain all lower dimensional features of physical meaning such as seasonal means and wind speed percentiles as well as turbulence characteristics. However, also looking at the four representative tiles shows that the quality of the clustering performance is different in various regimes of the space. While some of them show improvements at some intermediate dimensionality, others have poor clustering performances at higher dimensions. Therefore, the increase of the dimensionality of the space is not an unescapable modeling choice. It has to be considered as an experimental design parameter.

The non-monotonic behavior observed in the study is crucial for the dimension-aware design. The strongest values of stored quality were found for the intermediate dimensional conditions S5D and S7D compared to the highest dimensional condition S9D. In Tile 453 the stored clustering quality is found to be the highest in S5D and only slightly decreased in S7D, which indicates that moderate increases in the feature space can lead to more discriminative models without over-explaining them in terms of geometry. Conversely for Tile 575 the value of stored quality is found to be decreased in S9D compared to S5D indicating that increased dimensionality can lead to increased dimensionality-related effects such as dimensionality increase of redundancy, noise, and cluster fragmentation. Therefore, the design of the feature space should be validated at the structural level, rather than at the level of the number of features.

The second finding is that stability provides more information than that contained in any of the individual criteria for internal clustering. The seed and bootstrap analyses indicated that a few of the S7D configurations were stable when a configuration was reinitialized and reoptimized many times. In these cases, the clustering behaviour of different realisations of such configurations has now also been examined. Several of the stable configurations, such as the very stable Tile 453 with k=3 and the quite stable Tiles 143 with k=5 and 575 with k=5, provided further confirmation of the clustering behaviour in the intermediate-dimensional space. While the stability of the clustering assignments obtained in different runs and realisations guarantees that the assignments will be the same (up to permutations of indices) under somewhat different initialisations and different resamplings, the fact that assignments can be stable for many different choices of initialisations and resamplings does not guarantee that the structure of the clusters will be the same, and that the two- and three-dimensional representations of the data provided by dimensionality reduction are therefore true representations of the data. Stability is a prerequisite for examining and using the resulting data visualisations obtained by dimensionality reduction. However, stability is only a necessary condition; it is not a sufficient condition.

This is a very important conclusion for the methodological contribution of this work. A visually appealing 2D or 3D layout of a subset of the objects in the data does not guarantee that the corresponding structural interpretation is accurate. The true Euclidean distance between centroids must serve as the reference frame in order to assess the quality of any low-dimensional projection. This is illustrated by the centroid-distance audit in Figure 6. The low-dimensional projections strongly compress the true between-cluster distances, as illustrated with the strongest retained distortion case of Tile 143 in S9D for k = 5. The two clusters are clearly separated by large Euclidean distance in the engineered feature space (top of the figure), but have a much smaller distance in PCA-2D (middle of the figure). The distortion reduces but does not disappear when projecting up to PCA-3D.

This finding is important, as all distortion cases that were removed as a result of inspecting the retained evidence are not due to weak clustering. Rather, they correspond to clusters that started with high silhouette scores, high stability, or both. This adds another methodological contribution to this study. The projection misinterpretation is not restricted to cases of poor clustering performance, since it can occur for some of the strongest clustering solutions. High silhouette scores and high ARI scores, which are generally regarded as indicative of good clustering solutions, are therefore not sufficient guarantees of a correct and/or visually valid projection. Auditing projections should therefore be regarded as a separate step in the evaluation of the clustering solution rather than as an optional plot.

This conclusion was verified by a cross-method analysis. The Tile 143 cases were examined with DBSCAN and OPTICS in S7D and S9D again and the projection distortion was verified. The distortions differed a bit from each other depending on the method used and were generally smaller in PCA-3D than in PCA-2D. This cross-method validation is important as it removed the possibility that the distortion was caused only by the baseline centroid-based partitioning method used to derive the initial results. Now it has been verified that the distortion occurs also with a centroid-based method and with a density-based method and therefore it must be caused by the dimensional reduction, not by the clustering method.

One more consequence is derived from the relationship between data statistics and physical representation. Since this study does not lead to any energy or exergy application, it simply shows that potentially physically different wind regimes can be represented in the same way in the reduced space of data when no analysis of the clusters geometric structure is performed prior to data reduction. Consequently, a stable partition that is statistically sound may not have a physically adequate representation in the reduced space, which is by no means a minor issue for offshore wind regime studies. Indeed, the representation of clusters (as regime types or archetypes) is often used to derive other variables of interest (such as resources, trends or decision support tools). The present study only contributes to the physics-informed interpretability of regime analysis. A stable and statistically valid partition does not necessarily mean that the representation in the reduced space is physically meaningful, unless the geometric structure of the clusters is verified as adequate.

Clustering in the dimension aware setting, requires evaluation within a multi-layer methodological framework. Internal measures are adopted to describe the distributions of clusters for the purposes of compactness and separation. Measures which evaluate the stability of the clustering against noise and support the fact that true high dimensional centroid distances are a geometrical standard against which the distortion caused to the data by the projection should be measured, in order to evaluate the projection auditing. The main methodological contribution of this paper is to demonstrate that no individual metric layer of the multi-layer evaluation framework of dimension aware clustering is sufficient to credibly assess the structural validity of the obtained clustering, or to guarantee the interpretability of the clustering results. It is rather the multi-layered combination of them, that enables the assessment of the validity and interpretability of a clustering result.

In a broader context, the implications of this study may be significant for practitioners dealing with high-dimensional clustering analyses in the context of offshore wind and other geophysical applications. In particular, it may be advisable to consider dimensional design as a modelling assumption rather than as a data preprocessing step, that stability assessment is at least as important as assessment of the quality of the cluster solution, that projected visualizations do not necessarily reflect the actual data structure, that high-dimensional centroid distances and distortion metrics are always presented in conjunction with the lower-dimensional visualizations in order to validate the visual impression derived from the latter, and that the findings of this study serve as a warning against an excessive use of clustering analyses which are based on little more than personal feeling and which can easily lead to nonsensical results.

This paper provides a unified contribution. In particular, it discusses the following important points: (i) designing scenarios when understanding how cluster quality evolves with increasing data dimensionality, (ii) stability-driven validation to distinguish between stable and fragile clusters, (iii) the high dimensional centroid structure should be considered as the reference geometry instead of its projected version, and (iv) that the projection-induced distortion should be quantified before any interpretation is made from the projected high dimensional data. Accordingly, while dimensional reduction can be a convenient tool for visualization, the proposed framework highlights the important difference between visualization and structural analysis.

Conclusions.

A dimension-aware methodological framework for assessing the clustering properties of high-dimensional features associated with offshore wind data is introduced and validated. The validation testbed was a controlled four-tile offshore wind resource testbed with data generated in a staged sequence of increasing dimensions from S2D to S9D. This validated study explores the impact of high-dimensional feature space complexity on cluster quality and clustering interpretation stability and reliability. This validation study moves beyond the cluster validity index-based clustering assessment practices. The assessments are based on dimensional design, stability and geometric faithfulness.

These experiments show that the increase in dimension does not lead to a monotonic improvement of the clustering. Most often the intermediate dimensionality settings S5D and S7D result in the strongest clustering, while the most complex S9D setting does not lead to a higher clustering performance and, in particular, loses some of the Euclidean properties which are necessary for such a performance. In conclusion, one has to be aware that increasing the dimension of the feature space is not always a good idea. The validation of the effect of adding new dimensions to the feature space has to be done for a specific wind regime. It is therefore very important to have in mind the structure of the cluster of points describing this regime, and how this regime evolves when moving to more complex Euclidean spaces.

This paper investigates how the quality and stability of clustering solutions relate to each other and confirmed that the seed and bootstrap stability analysis methods can provide some insight into which clustering solutions are stable (i.e., almost identical) and robust (i.e., resistant to some level of noise) and which are just more sensitive to different initializations or resamplings. Therefore, besides the classical validity indices which provide a measurement of the intrinsic quality of a clustering solution, stability-based metrics are needed to understand which clusterings can be considered as stable and robust. Indeed, a good solution according to the classical criteria may not be robust to a certain level of noise and therefore cannot be considered as a robust clustering solution, and so metrics that relate to the intrinsic structure of the solution as well as to the robustness of the solution are needed.

The most important methodological result of the present study is related to the projection-induced misinterpretation. Comparing the true high-dimensional centroid distances between different clusters with their corresponding values in the lower dimensional space derived by PCA, it was found that the true inter-cluster distances can be strongly reduced by projecting them to lower dimensional spaces. While the discrepancies are clearly not large for the weak solutions, the most of the cases where this effect is significant deal with very stable solutions as well: in some cases the projected 2D distances between centroids of two truly disjoint clusters were found to be about only 70% of their true values in the high dimensional space, and even in the few cases where 3D projections were derived, the full restoration of the original distances was not achieved in all cases. This means that the visually highly persuasive lower dimensional picture of the data does not necessarily give an accurate description of the structure of the high dimensional feature space.

Cross-method validation of the distortion effect resulted in the same conclusion, and the distortion pattern was clearly visible in both DBSCAN and OPTICS clustering algorithms. This means that the compression is an inherent aspect of the clustering process, and that it is independent of the clustering algorithm used (centroid-based or density-based). This means that the distortion is not a method-specific artefact, but rather a necessary by-product of the dimensionality reduction that must be carried out in order to perform clustering in the high-dimensional feature spaces of the atmosphere. This poses a significant methodological challenge to the clustering of high-dimensional data, such as that which describes the complex offshore wind resources.

This paper is a single submission to the data clustering area. The contributions of this paper are fourfold: 1) the need for dimension-aware clustering problems is argued to study how different dimensions in a feature space affect clustering behavior in distinct ways; 2) the need for stability-driven validity indices is argued to prevent trivial clustering solutions by comparing the relative importance of the underlying data structure and data distribution associated with different clustering solutions; 3) high dimensional centroid geometry of data clusters is argued to be an important structural clustering metric; 4) the need for distortion measures associated with projections is argued to match those associated with dimensionality reduction techniques used in data clustering applications, so that the distortion due to data projection is decoupled from that due to dimensionality reduction and data visualization. The need and merits of dimensionality reduction for data visualization are not disputed. However, it is also important to measure the distortion due to data projection so that it is not obscured by the data visualization.

The four-tile testbed is used in order to limit the complexity of the validation and to focus on method reproducibility. A validation over a more diverse and representative set of offshore wind resources is highly recommended. The audit methodology here introduced may be used for validating other wind resource assessments, as well as for the assessment of the impact of slightly different assumptions and for features selected based on more physics-based criteria, e.g. including further metrics for the description of the wind persistence, turbulence and extreme values.

Dimensionality is not a factor that can be ignored when it comes to the clustering of data in terms of quality, stability or interpretability. This paper shows that for a given dataset, even though good and stable clusterings can be obtained, low dimensional projections can lead to such a distortion of the original data set topology that the high dimensional properties of the complex structures of offshore wind data are lost. Therefore, the clustering of offshore wind data is an example of a design-stability-audit-interpretation cycle that should be considered with dimensionality as a central aspect.

Author Contributions:

Conceptualisation and methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualisation, project administration: H.R.S.M. supervision: S.B.I.-Z., A.H.K.,and M.R.Z. All authors have read and agreed to the published version of the manuscript.

Funding:

This research received no external funding.

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Not applicable.

Data Availability Statement:

The original contributions presented in this study are included in the article. The ERA5 data are publicly available from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu). Further inquiries can be directed to the corresponding authors.

Acknowledgements

: The authors acknowledge the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing the ERA5 reanalysis dataset through the Copernicus Climate Data Store, which served as the primary data source for this research.

Conflicts of Interest:

The authors declare no conflicts of interest.

References.

[1] W. Huang, X. Zhu, H. Xia, and K. Wu, “Offshore Wind Energy Assessment with a Clustering Approach to Mixture Model Parameter Estimation,” *J. Mar. Sci. Eng.*, vol. 11, no. 11, p. 2060, Oct. 2023, doi: 10.3390/jmse11112060.

[2] M. Schelbergen, P. C. Kalverla, R. Schmehl, and S. J. Watson, “Clustering wind profile shapes to estimate airborne wind energy production,” *Wind Energy Science*, vol. 5, no. 3, pp. 1097–1120, Aug. 2020, doi: 10.5194/wes-5-1097-2020.

[3] T. Liu, H. Yu, and R. H. Blair, “Stability estimation for unsupervised clustering: A review,” *WIREs Computational Statistics*, vol. 14, no. 6, Nov. 2022, doi: 10.1002/wics.1575.

[4] von L. Ulrike, “Clustering Stability: An Overview,” *Foundations and Trends® in Machine Learning*, vol. 2, no. 3, pp. 236–272, Apr. 2010, doi: 10.1561/2200000008.

[5] S. S. Abdullah, N. Rostamzadeh, K. Sedig, A. X. Garg, and E. McArthur, “Visual Analytics for Dimension Reduction and Cluster Analysis of High Dimensional Electronic Health Records,” *Informatics*, vol. 7, no. 2, p. 17, May 2020, doi: 10.3390/informatics7020017.

[6] H. Jeon, H.-K. Ko, J. Jo, Y. Kim, and J. Seo, “Measuring and Explaining the Inter-Cluster Reliability of Multidimensional Projections,” *IEEE Trans. Vis. Comput. Graph.*, vol. 28, no. 1, pp. 551–561, Jan. 2022, doi: 10.1109/TVCG.2021.311-4833.

[7] W. Yi, S. Bu, H.-H. Lee, and C.-H. Chan, “Comparative Analysis of Manifold Learning-Based Dimension Reduction Methods: A Mathematical Perspective,” *Mathematics*, vol. 12, no. 15, p. 2388, Jul. 2024, doi: 10.3390/math12152388.

[8] T. Munz-Körner and D. Weiskopf, “Exploring visual quality of multidimensional time series projections,” *Visual Informatics*, vol. 8, no. 2, pp. 27–42, Jun. 2024, doi: 10.1016/j.-visinf.2024.04.004.

[9] C. Hennig, “Cluster-wise assessment of cluster stability,” *Comput. Stat. Data Anal.*, vol. 52, no. 1, pp. 258–271, Sep. 2007, doi: 10.1016/j.csda.2006.11.025.

[10] A. Mitra, “Regime Characterization of Offshore Wind Resource Using Unsupervised Learning,” Sep. 2025. doi: 10.2-172/2589309.

[11] S. Yang, H. Yuan, and L. Dong, “Offshore wind resource assessment by characterizing weather regimes based on self-organizing map,” *Environmental Research Letters*, vol. 17, no. 12, p. 124009, Dec. 2022, doi: 10.1088/1748-9326/aca2c2.

[12] M. C. A. Clare, S. C. Warder, R. Neal, B. Bhaskaran, and M. D. Piggott, “An Unsupervised Learning Approach for Predicting Wind Farm Power and Downstream Wakes Using Weather Patterns,” *J. Adv. Model. Earth Syst.*, vol. 16, no. 2, Feb. 2024, doi: 10.1029/2023MS003947.

[13] B. R. Cheneka, S. J. Watson, and S. Basu, “Associating Synoptic-Scale Weather Patterns with Aggregated Offshore Wind Power Production and Ramps,” *Energies (Basel).*, vol. 14, no. 13, p. 3903, Jun. 2021, doi: 10.3390/en14133903.

[14] Y. Liu, T. W. Juliano, R. Krishnamurthy, B. J. Gaudet, and J. Lee, “Linking large-scale weather patterns to observed and modeled turbine hub-height winds offshore of the US West

Coast," *Wind Energy Science*, vol. 10, no. 2, pp. 483–495, Feb. 2025, doi: 10.5194/wes-10-483-2025.

[15] S. Rubinetti, V. Fofonova, E. Arnone, and K. H. Wiltshire, "A Complete 60?Year Catalog of Wind Events in the German Bight (North Sea) Derived From ERA5 Reanalysis Data," *Earth and Space Science*, vol. 10, no. 10, Oct. 2023, doi: 10.1029/2023EA003020.

[16] W. Ji *et al.*, "Evaluation of ERA5 wind parameter with in-situ data offshore China," *PLoS One*, vol. 20, no. 5, p. e0317751, May 2025, doi: 10.1371/journal.pone.0317751.

[17] Z. Liu, R. Ma, and Y. Zhong, "Assessing and improving reliability of neighbor embedding methods: a map-continuity perspective," *Nat. Commun.*, vol. 16, no. 1, p. 5037, May 2025, doi: 10.1038/s41467-025-60434-9.

[18] D. Y. Badawood, "StatVis: A visual analytics framework for statistical cluster validation in high dimensions," *Array*, vol. 28, p. 100560, Dec. 2025, doi: 10.1016/j.array.2025.100560.

[19] R. Marini *et al.*, "Assessing the impact of wind profiles at offshore wind farm sites for field data-enabled design," Feb. 11, 2025. doi: 10.5194/wes-2025-9.

[20] C. Moss, M. Puccioni, R. Maulik, C. Jacquet, D. Apgar, and G. Valerio Iungo, "Profiling wind LiDAR measurements to quantify blockage for onshore wind turbines," *Wind Energy*, vol. 27, no. 11, pp. 1268–1285, Nov. 2024, doi: 10.1002/we.2877.

[21] Y. Ma *et al.*, "Wind dataset assessment and energy estimation for potential future offshore wind farm development areas on the Scotian Shelf," *Wind Energy Science*, vol. 10, no. 12, pp. 2965–2999, Dec. 2025, doi: 10.5194/wes-10-2965-2025.

[22] E. Cheynet, J. M. Diezel, H. Haakenstad, Ø. Breivik, A. Peña, and J. Reuder, "Tall wind profile validation of ERA5, NORA3, and NEWA datasets using lidar observations," *Wind Energy Science*, vol. 10, no. 4, pp. 733–754, Apr. 2025, doi: 10.5194/wes-10-733-2025.

[23] T. Pelser, J. M. Weinand, P. Kuckertz, R. McKenna, J. Linssen, and D. Stolten, "Reviewing accuracy & reproducibility of large-scale wind resource assessments," *Advances in Applied Energy*, vol. 13, p. 100158, Feb. 2024, doi: 10.1016/j.adapen.2023.100158.

[24] Y. Zhao, Y. Tao, Y. Chen, J. Yan, and Z. Zeng, "Increasing extreme winds challenge offshore wind energy resilience," *Nat. Commun.*, vol. 16, no. 1, p. 9529, Nov. 2025, doi: 10.1038/s41467-025-65105-3.

[25] Y. Wang *et al.*, "Study on Wind Profile Characteristics Using Cluster Analysis," *Atmosphere (Basel).*, vol. 15, no. 6, p. 708, Jun. 2024, doi: 10.3390/atmos15060708.

[26] M. T. Rahman, K. N. Hasan, and P. Sokolowski, "Evaluation of wind farm aggregation using probabilistic clustering algorithms for power system stability assessment," *Sustainable Energy, Grids and Networks*, vol. 30, p. 100678, Jun. 2022, doi: 10.1016/j.segan.2022.100678.

[27] K. Ali, D. M. Schultz, A. Revell, T. Stallard, and P. Ouro, "Assessment of Five Wind-Farm Parameterizations in the Weather Research and Forecasting Model: A Case Study of Wind Farms in the North Sea," *Mon. Weather Rev.*, vol. 151, no. 9, pp. 2333–2359, Sep. 2023, doi: 10.1175/MWR-D-23-0006.1.

[28] Z. Shu, P. W. Chan, and X. He, "Characterization and classification of offshore wind profiles via a shape-based cluster analysis framework," *Physics of Fluids*, vol. 37, no. 7, Jul. 2025, doi: 10.1063/5.0278542.

[29] A. Machado, M. Behrisch, and A. Telea, "Necessary but not Sufficient: Limitations of Projection Quality Metrics," *Computer Graphics Forum*, vol. 44, no. 3, Jun. 2025, doi: 10.1111/cgf.70101.

[30] H. Jeon *et al.*, "Unveiling High-dimensional Backstage: A Survey for Reliable Visual Analytics with Dimensionality Reduction," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: ACM, Apr. 2025, pp. 1–24. doi: 10.1145/3706598.3713551.

[31] D. Cashman, M. Keller, H. Jeon, B. C. Kwon, and Q. Wang, "A Critical Analysis of the Usage of Dimensionality Reduction in Four Domains," *IEEE Trans. Vis. Comput. Graph.*, vol. 31, no. 10, pp. 9405–9423, Oct. 2025, doi: 10.1109/TVCG.2025.3567989.

[32] H. Hersbach *et al.*, "The ERA5 global reanalysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 730, pp. 1999–2049, Jul. 2020, doi: 10.1002/qj.3803.

[33] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20150202, Apr. 2016, doi: 10.1098/rsta.2015.0202.

[34] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, Nov. 1987, doi: 10.1016/0377-0427(87)90125-7.

[35] D. L. Davies and D. W. Bouldin, "A Cluster Separation Measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, Apr. 1979, doi: 10.1109/TPAMI.1979.4766909.

[36] T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Commun. Stat. Theory Methods*, vol. 3, no. 1, pp. 1–27, 1974, doi: 10.1080/03610927408827101.

[37] L. Hubert and P. Arabie, "Comparing partitions," *J. Classif.*, vol. 2, no. 1, pp. 193–218, Dec. 1985, doi: 10.1007/BF01908075.

[38] J. B. Kruskal, "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," *Psychometrika*, vol. 29, no. 1, pp. 1–27, Mar. 1964, doi: 10.1007/BF02289565.

[39] C. Spearman, "The Proof and Measurement of Association between Two Things," *Am. J. Psychol.*, vol. 15, no. 1, p. 72, Jan. 1904, doi: 10.2307/1412159.

[40] K. Pearson, "VII. Note on regression and inheritance in the case of two parents," *Proceedings of the Royal Society of London*, vol. 58, no. 347–352, pp. 240–242, Dec. 1895, doi: 10.1098/rspl.1895.0041.