engineering proceedings

MDPI

Proceeding Paper

Time Series Clustering of High Gamma Dose Rate Incidents †

Mohammed Al Saleh [1,2,3,*], Beatrice Finance [1], Yehia Taher [1], Ali Jaber [2] and Roger Luff [4]

[1] David Laboratory, PARIS-SACLAY University, 45 Avenue des Etats-Unis, 78035 Versailles, France; beatrice.finance@uvsq.fr (B.F.); yehia.taher@uvsq.fr (Y.T.)
[2] Rafic Hariri University Campus, Lebanese University, Beirut 6573, Lebanon; ali.jaber@ul.edu.lb
[3] Lebanese Atomic Energy Commission (LAEC), National Council for Scientific Research (CNRS), Airport Road, P.O. Box 11-8281 Beirut, Lebanon
[4] Federal Office for Radiation Protection (BfS), D-24768 Rendsburg, Germany; rluff@bfs.de
* Correspondence: mhdalsaleh@gmail.com
† Presented at the 8th International Conference on Time Series and Forecasting, Gran Canaria, Spain, 27–30 June 2022.

Abstract: In this paper, we proposed an unsupervised machine-learning-based framework to automate the process of extracting suspicious gamma dose rate incidents from the real unlabeled raw historical data measured in the German Radiation Early Warning Network and identify the underlying events behind each. This raised the research problem of clustering unlabeled time series data with varying lengths and scales. Based on the many evaluations, we demonstrated that the state-of-the-art's most popular time series clustering models were not suitable to perform this task. This motivated us to introduce our own approach. Through this approach we were able to perform online classification for gamma dose rate incidents of varying lengths and scales.

Keywords: machine learning algorithms; predictive model; time series clustering; gamma dose rate; Radiation Early Warning Network

Citation: Al Saleh, M.; Finance, B.; Taher, Y.; Jaber, A.; Luff, R. Time Series Clustering of High Gamma Dose Rate Incidents. *Eng. Proc.* **2022**, *18*, 24. https://doi.org/10.3390/engproc2022018024

Academic Editors: Ignacio Rojas, Hector Pomares, Olga Valenzuela, Fernando Rojas and Luis Javier Herrera

Published: 22 June 2022

1. Introduction

Time series analysis is gaining more and more interest in so many domains. That is because, with the proliferation of the use of sensors and IoT (Internet of Things) devices that continuously produce massive amounts of real-time data, special care has been given for analyzing that data to understand past events and patterns and predict future ones. Medical heart monitor's data, stock market prices, weather conditions, etc., are all examples of such time series data.

In this paper, we are interested in analyzing the gamma dose rate (background radiation level) in the environment. Some incidents can cause an abrupt increase in the gamma dose rate, such as what happened in the Chernobyl accident where the biggest short-term leak of radioactive materials was ever recorded in history [1]. Such an event has to be intercepted at the earliest point possible to take the proper measures and precautions and notify the concerned authorities to minimize the effects of such a hazardous situation. It is a very critical task, as long term or acute exposure to a high gamma dose rate can have many hazardous consequences on humans as well as on the ecosystem.

Around the globe, there are thousands of probes (sensors) that collect gamma dose rates in real time. A Radiation Early Warning System (REWS) collects and analyses data while raising alarm in case of an increase in the local gamma dose rate. Whenever an event occurs (i.e., the gamma dose rate goes above the accepted threshold, provided by experts), an alarm is triggered, and a team of experts and personnel have to unite to investigate the reasons behind this rise. Currently, analyzing the incoming incidents is performed manually while relying on the expert efforts. Such a method is time-consuming and risky, knowing that the factors affecting the gamma dose rate are not always known immediately. Fortunately, most of the incoming incidents are mainly innocent as they remain in an

acceptable range value for humans health and this value returns to normal after a period of time.

The objective of our research is to propose an *Intelligent Radiation Early Warning System* that automatically finds the cause behind an incident and its classification into *real or innocent* in real time. In order to build such an intelligent system, we first need to understand and learn from the past by using the historical databases produced by REWS. These databases contain raw unlabeled data (i.e., time series) corresponding to the gamma dose rate monitoring at each probe. Second, we need to identify, in real time, situations already encountered or not in the past to predict the cause behind an incident as quick as possible.

In this paper, we are only interested in describing the challenges of the first phase. We aim at proposing an unsupervised machine learning model that will help us to automatically identify the reasons behind incidents. Since we do not have full knowledge regarding the incident causes (neither their number), or the incident behavior patterns, we solicited the help of experts to validate the result quality of our model and to label the obtained clusters. Another challenge we run into was finding the right unsupervised machine learning model, which is a highly challenging mission. We have run experiments on real data in order to come up with the most suitable approach that is described in this paper. Surprisingly, the most popular approaches found in the state-of-the-art were not the ones selected for our proposal.

The remaining sections of the paper are organized as follows. In Section 2, we address the context and problem statement. Section 3 recalls the background related to the state-of-the-art. Section 4 describes the methodology and the different evaluations we conducted to find the best model for our data. Finally, we conclude in Section 4.

2. Context and Problem Statement

For a long time now, time series data analysis has been a center of attention in research as it is used in different applications such as weather prediction, motion capture processing, analyzing insect behavior, pattern discovery on health-care data, and so on. Similarly, intelligence can be extracted from the gamma dose rate time series data in the radiation monitoring domain.

Although gamma dose rate is in theoretically affected by real incidents, it may also be affected by other factors such as weather conditions (rain, lightning, snow...), environmental factors (sun's cosmic radiations), and many other events as shown in Figure 1. Depending on the type of event, they cause the gamma dose rate to go above an acceptable value for a short or long period of time. These kinds of incidents, such as the ones caused by the rain, we call them *soft parabolas*, which are incidents that stay above the acceptable value of gamma dose rate for a considerable period of time. Incidents such as the ones caused by lightning are called *hard parabolas* and are hardly confused with real incidents or cause alarm. Usually, they are either caused by instantaneous events or a malfunctioning probe.

Note that even these *innocent* events trigger the system's alarm (depicted in red in Figure 1) and ensue a technically useless investigation; they have to be recognized and discarded by the experts after manually searching through and analyzing the multiple data sources (rain data, temperature data, wind data, even transportation data, etc.), which may not be available for inspection at any time and which elongates the useless process of evaluating the event as not alarming.

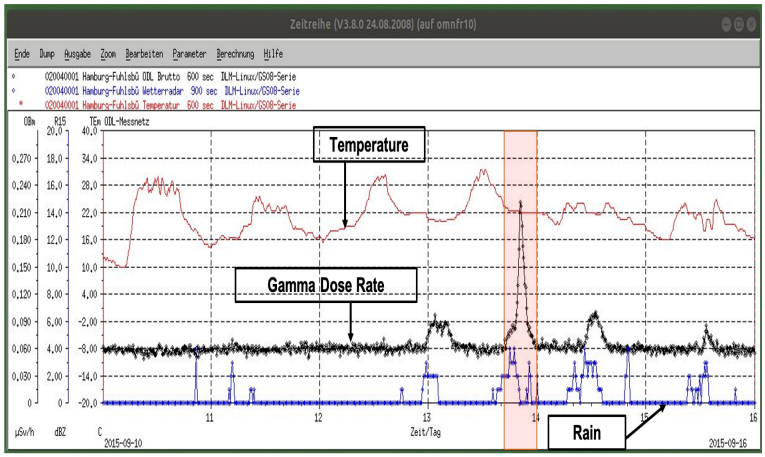

Figure 1. Increased gamma dose rate due to the deposits of Radon by-products from the atmosphere by rain.

To build our *Intelligent Radiation Early Warning System*, we faced the following problems. Firstly, the historical databases only maintain the gamma dose rate for each probe, unfortunately the experts' evaluations of the incidents are not maintained in the historical databases. We are not dealing with multivariate time series, as the different sources, such as precipitation and temperature, are not stored in some databases. Here, we deal with univariate time series, which are unlabeled as shown in the Figure 2. Secondly, incidents that are caused by the same event may not have a recognizable temporal trace or characteristics but more common behavior. For example, a particular event may cause peaks of increasing amplitudes that decrease over a longer period of time; another may cause an abrupt increase and maintain its amplitude for a period of time, and so on. Note that incidents caused by the same event can last for a varying length of time and reach different amplitudes. Thus, grouping the discovered incidents into similar patterns to explore their causes is deemed to be a big challenge.

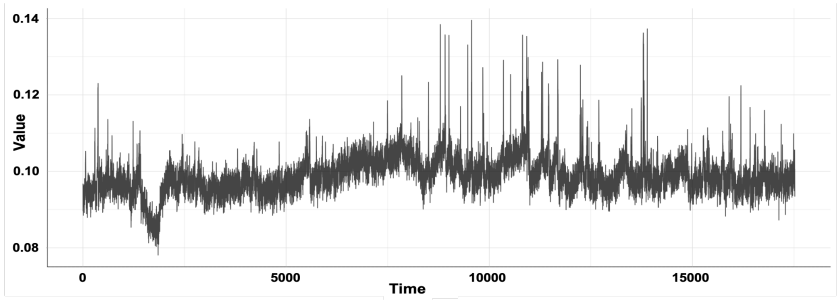

Figure 2. Typical gamma dose rate time series.

Although the literature on time series analysis is prolific, the aforementioned challenge still remains an open question that can be formulated as follows: *"What is the best-fit unsupervised machine learning model that should be used for clustering our time series of varying lengths and different scales*?*"*

We propose to answer this question in this paper; this can be achieved thanks to the Germany REWS System [2], which offer their data to run all experiments; we thank all the experts we are in contact with, to validate or invalidate the results found by applying the different time series clustering algorithms. The data used in this work comprise the past ten years minute-by-minute gamma dose rate real data for over a thousand probes.

3. Background and State-of-the-Art

In this section, we briefly recall the four main phases for defining a time series clustering approach. First, there is the time series data preprocessing. Second, the similarity

measure should be chosen. Then, the clustering algorithm must be selected. Finally, the optimal number of clusters should be determined. Based on these, we enumerate the approaches proposed in the state-of-the-art for univariate time series clustering, especially those for the ones of varying length.

3.1. Time Series Clustering Generic Model

1. Time Series Data Preprocessing Data normalization and missing data imputation are the basic data preprocessing activities to be applied on time series data. Both kinds of techniques have a significant impact on the performance of a model, and they should be chosen based on the problem and model at hand.
   Missing data imputation: Missing values may cause problems for machine learning algorithms as they will perform better with complete well-formed data. Some of the most popular approaches to deal with this problem are dropping rows with missing values, statistical imputation, and model imputation.
   Normalization: The most common normalization methods used during data transformation include min-max, decimal scaling, and z-normalization [3]. The first two methods rely entirely on the minimum and/or maximum values that should be predefined from the data and upon which normalization will be performed. This is not the case with gamma dose rate time series data because the minimum and the maximum values are unknown.
2. Similarity Measure Similarity measures are algorithms used to determine the resemblance between different samples. In time series clustering, it is the determining factor used by the clustering algorithm to decide which cluster each sample belongs to. Shape-based distances evaluate the similarity of samples based on the actual or the normalized values, whereas feature-based distances evaluate similarity based on extracted features. In our context, we are only interested in shape-based measures. They fall into one of two categories:
   Lock-step measures are metrics that evaluate the distance between two time series sequences as the overall difference between each point and its counterpart in the other sequence. These measures require data sequences to be of equal length. Minkowski (L_p norm) distances, specifically Euclidean [4], are the most favored lock-step metrics in machine learning. Their popularity is derived from their simplicity and success in machine learning literature as well as their being parameter-free.
   Elastic measures on the other hand, provide better flexibility as they permit one-to-many and one-to-none point evaluation. Due to this flexibility, these measures provide a better comparison. This flexibility, however, comes at the price of increased time complexity.
   Dynamic Time Warping (DTW) [5] is the most famous elastic measure in the literature, specifically introduced for time series analysis. As its name suggests, it warps the two considered sequences in time to deal with time shift and speed variations. DTW is a good similarity measure for comparing samples of varying lengths. It produces a scale-like effect, stretching and contracting, by accepting many-to-one matching; however, this also makes it sensitive to outliers.
3. Clustering Algorithm Despite the major role each part plays in the process, a recent study has shown that the choice of the proper similarity measure is considered to be more fundamental than that of the clustering algorithm itself in time series clustering. As a result, the majority of time series clustering fall back on classic clustering algorithms where either the choice of distance measure is modified as befits the time series data (raw-based methods), or the data are transformed to fit the clustering algorithm (feature and model-based) [6]. The raw-based approach is often preferable to the feature and model-based approaches. The latter are generally domain-dependent, where the features or models have to be altered depending on the application in the different domains. On the other hand, the main catch of

the raw-based algorithms is the *curse of dimensionality*, [7] where specs with high dimensionality are considered.
In our study, we focus on raw-based approaches as the results can be better generalized across the different applications and domains. We therefore only consider hierarchical and partitional clustering methods, as they are the most commonly used clustering methods in the literature on time series clustering [6,8].
Hierarchical clustering takes no parameters other than the linkage criteria [9]. Depending on the linkage criteria, a tree-like nested "hierarchy" of clusters is built, which can be visualized by a dendrogram. Hierarchical clustering's main advantage is that it does not require the number of clusters as input. Once the dendrogram is obtained, the clusters can be decided by making a cut at a certain point. On the other hand, it requires the distance matrix of all possible pairs of observations. This makes it very computationally expensive and not a favorable option for huge data sets.
Partitional clustering, as its name implies, partitional clustering partitions the data into k different clusters where k is specified a priori. Partitional clustering's aim is to minimize intra-cluster distance and maximize the inter-cluster distance. Partitional methods need the number of clusters k a priori. K-means [10] and K-medoids [11] heuristics are considered the front-men of the partitional methods. They are both based on the concept of finding the best cluster centers, minimizing the distance between each observation and the center of the cluster it is assigned to.

4. Determining optimal number of clusters Clustering methods require the number of clusters k as an input parameter in order to return a clustering. Non-hierarchical methods usually require k to be specified beforehand, whereas, for hierarchical methods, the value of k can be set afterward. Two of the main statistical approaches used for the evaluation of an optimal number of clusters are:

   - Elbow Method Is a method that estimates the number of clusters by comparing the within-cluster dispersion.
   - Silhouette Method The Silhouette index is proposed by Kaufman et al. [9] and is based on compactness and separation of clusters.

3.2. State-of-the-Art

In Table 1, we summarize the main approaches proposed in the literature to cluster time series of varying lengths. We found that the most favored similarity measure is DTW, and the most popular clustering algorithm is K-medoids. Combining DTW and K-means does not give valid clusters as stated in [12]. The only approach using DTW and K-Means is proposed by Petitjean et al. [13] who introduced a global averaging method called DTW barycenter averaging (DBA), which is a heuristic strategy; however, combining DTW with k-means seems to have a lot of complications, and even with the DBA averaging method, the verdict is left for the testing to see how the DBA fairs with a big length difference compared with the DTW with the k-medoids model.

Table 1. Combined Techniques in the Literature.

Similarity Measure	Clustering Algorithm	Literature
DTW	K-means (DBA)	Zhang et al., 2015 [14]
	K-medoids	Liao et al., 2002 [15] Liao et al., 2006 [16] Hautamaki et al., 2008 [17] Gao et al., 2020 [18]
LCSS	K-medoids	Soleimany et al., 2019 [19]

While this sounds good for the similarity measure (DTW), it is still not clear if this is still true when the similarity measure is used within a machine learning model. In a recent work, Tan et al. [20] explain that there was a little work published in the literature on the

classification of time series of varying lengths compared to the "time-warping" problem. They say the problem is comparatively "understudied and unappreciated". When looking at the UCR archive [21], we see also that there are a lot of datasets that are uniform and not much of varying length only very recently in 2018. That is why we believe that the context of our research will help to have a better understanding of the problem. Unfortunately, due to the nature of the data (radiation level), they cannot be rendered public to the UCR archive.

4. Our Approach and Experiments

In this section, we describe the different choices made in our model to cluster gamma dose rate time series. Knowing that we are not contributing to the clustering domain, we are proposing a kind of methodology where we are fine tuning the process of seeking for the best way to do clustering. This led us to the hardest part of our research where we tested all types of combinations between similarity measures and clustering algorithms. In the end, we present our contribution, which is the machine learning model we introduced that achieved the best results through testing. Our approach is depicted in Figure 3. As explained before, machine learning algorithms achieve better performance if the time series data have a consistent scale or distribution. Thus, an important attention has been on incident extraction and data preprocessing. Then, we detail the similarity measure and the clustering algorithm we retain.

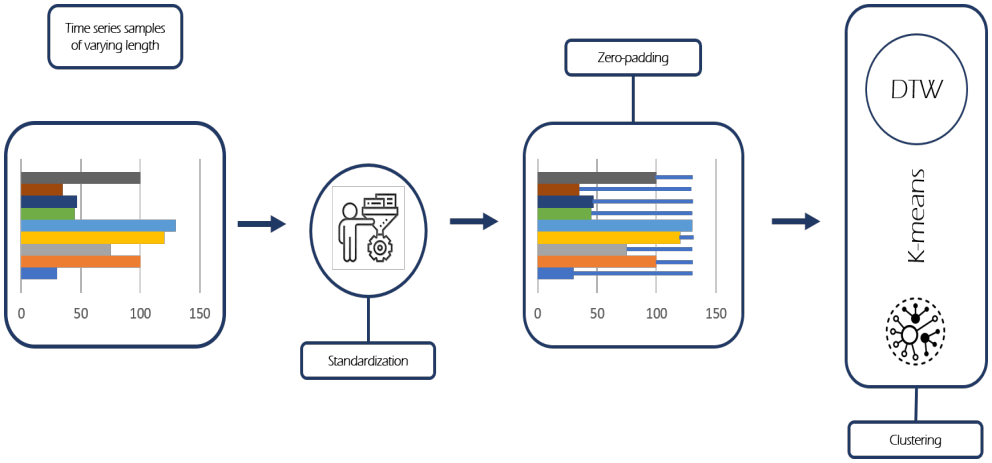

Figure 3. Specific Model for Clustering Sequences of Varying Length.

4.1. Incident Extraction and Preprocessing

In the beginning, we considered all subsequences of the time series where the gamma dose rate went above the threshold as incidents; however, we found that short incidents added much noise to the data set after experimentation. The clustering could not achieve any satisfactory results, so we re-consulted the experts. In the remaining work, we discarded all incidents that did not last at least 30 min above the peak threshold.

1. **Missing Data Imputation:** As we explained previously, the gamma dose rate data are very well susceptible to the missing data problem. That is because we are dealing with data coming from sensors, and these sensors are most probably going to malfunction at one time or another. Because the data we are dealing with are relatively huge and based on the intelligence obtained from experts, we decided to deal with the problem of having missing data by: (1) dropping the whole time series data (one year worth of data) if the missing data are distributed in big patches throughout it; (2) dropping the extracted incident if it encounters a missing data point because this means that the probe is malfunctioning at the time and hence it cannot be trusted.
2. **Scale Standardization:** The extracted incidents resulting from the extraction approach are of varying scale and amplitudes. The gamma dose rate can reach unpre-

dictable levels when affected by a radiation event; we cannot know the maximum and minimum values in order to perform min-max or decimal-scaling normalization. For this reason, we had to discard them. On the other hand, **z-normalization** is highly applied in the time series literature. Its strong point is that it normalizes the samples, so only the *shape* of them is left to compare to each other. A value a of A is standardized to a' by computing:

$$a' = \frac{a - \mu(A)}{\sigma(A)}$$

The fact that it normalized the data to be of a mean equal to zero and standard deviation between 1 and -1 has great advantages, as explained in the next section.

3. **Length Standardization**. As mentioned before in the state-of-the-art, the elastic measure DTW is very sensitive to outliers, which means that if the variation in length between samples is too high, the clustering is not performed well, as we see later in the evaluation. To solve the varying length problem, a **padding** technique has been used as proposed by Tan et al. [20]. Samples are padded with in-consequential data points such as zero or the mean or the median depending on the data distribution. By padding with zero to the z-normalized data, neither the mean (0) nor the standard deviation was affected since zero is indeed in-consequential for this distribution of data. Notice that without the z-normalization, it would have been impossible to apply the z-padding. Thus, resulting in having all the incidents in the dataset of equal length and without interfering in the characteristics of the data.
The standardization applied in the preprocessing phase was critical for the approach. Without this preprocessing phase, the padding could not have been performed and the other experiments would not yield meaningful clusters.

4.2. The Time Series Clustering Specific Model

1. **The Similarity Measure**. Among the two elastic measures, we chose DTW as it tolerates slight time axis misalignment. Moreover, DTW is tolerant to samples of varying lengths. The same can be said about LCSS; however, between the two similarity measures, we found that DTW performed better with our data set than LCSS as the latter is more likely to ignore significant data points in the time series, considering them as outliers. You will see in our experiments that our samples are basically made of outliers as they are abnormal behavior of the gamma dose rate, showing up in a stochastic behavior.
2. **The Clustering Algorithm**. Due to the preprocessing of the data with z-normalization and zero-padding, we opted for the **K-means with DTW Barycenter Averaging:** algorithm for the clustering. Although according to the state-of-the-art, K-medoids is the most popular technique to be used with DTW, we will see that, in our context, K-means performs better as with the zero-padding, samples become of equal length. DBA, which stands for DTW barycenter averaging [14], evaluates the mean of a set of sequences by iteratively refining the potential average sequence to reach the minimum DTW distance between it and the sequences.
3. **Choosing Optimal Number of Clusters**. Now that we have our clustering model, we have to choose the optimal number of clusters k, which is the maximum number of clusters with no redundancy. In order to do this we had to experiment with different ks and evaluate the results of each. We first tried to do this using the indices mentioned before for determining the optimal number of clusters; however, the results obtained from the algorithms were not helpful and sometime not logical. In our approach, we presented to the experts the computed cluster centers from our experiments for $1 < k < 10$, and, together, we saw that for $k > 3$ we started to have redundant clusters (as shown in Figure 4 for $k = 4$), so we decided that the optimal k for this dataset is 3.

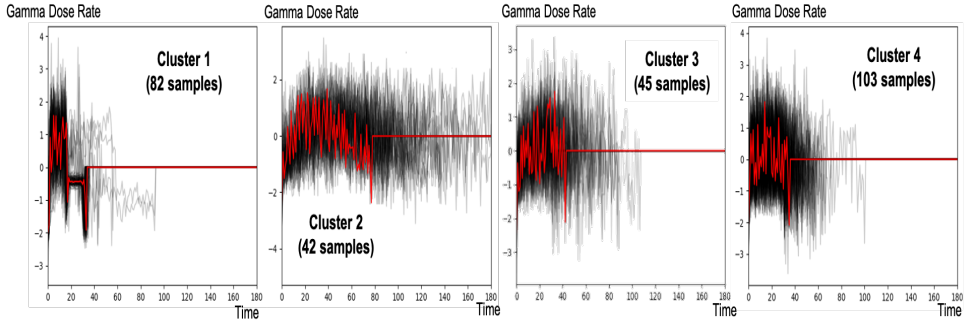

Figure 4. Our model's cluster results for $k = 4$.

4.3. Experimentation

In order to compare the different approaches of the state-of-the-art, as well as to see the benefit of our proposed model, we decide to evaluate, in a systematic way, different combinations of preprocessing phases (with or without z-normalization or zero-padded) with different clustering models (K-means, K-medoids, K-shape) as synthesized in Table 2. The overall number of experiments performed was 24, including the 16 described experiments in Table 2. Due to space limitations, we only give the results obtained with K-medoids or K-means with the same similarity measure and the same preprocessing; however, all evaluations are available by contacting the authors.

Table 2. Model Experiments.

Clustering Algorithm	Similarity Measure	Z-Normalized	Zero-Padded	
			Yes	No
K-means	DTW	Yes	✓	✓
		No	✓	✓
	LCSS	Yes	✓	✓
		No	✓	✓
	Euclidean	Yes	✓	
K-medoids	DTW	Yes	✓	✓
		No	✓	✓
	DTW with length factor	Yes		✓
		No		✓
K-shape	SBD	Yes	✓	

In Figure 5, using K-medoids with DTW with/without padding, we faced the same problem caused by the fact that the K-medoids algorithm tolerates outliers, so the obtained clusters have a lot of misplaced incidents and the centroid of the clusters does not clearly represent the observations in the cluster. On the other hand, observing the results of K-means clustering, we can see how adding up each preprocessing step brought us closer to the best cluster results, shown in Figure 6, which were approved by the experts who found that indeed each cluster (from left to right) can be explained by a different underlying event. Cluster 1's incidents are caused by a **calibration event** performed on the probes. Cluster 2's incidents are caused by a **stormy** rain where the wind causes the very sensitive probes to be affected by vibrations. Cluster 3's incidents are caused by a **normal rain** that causes the elements to go straight down and affect the probe with an immediate sharp increase. Notice that we also tried to increase the number of k, but we found that, when above 3, we started to see redundant clusters.

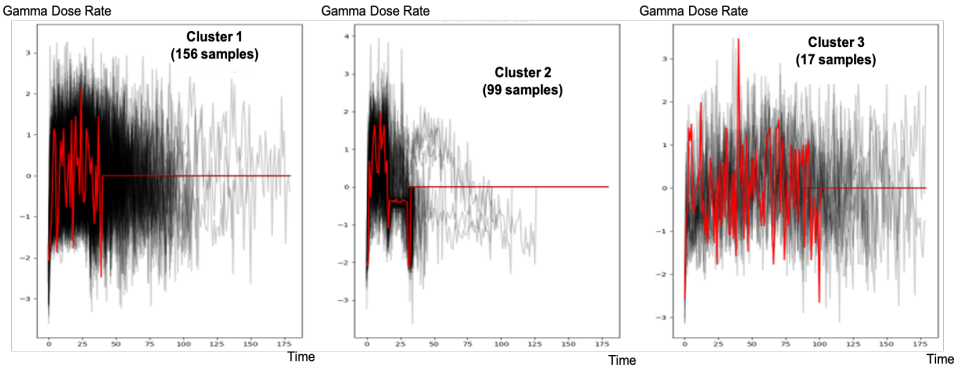

Figure 5. K-medoids with DTW(DBA) and z-normalized data with padding.

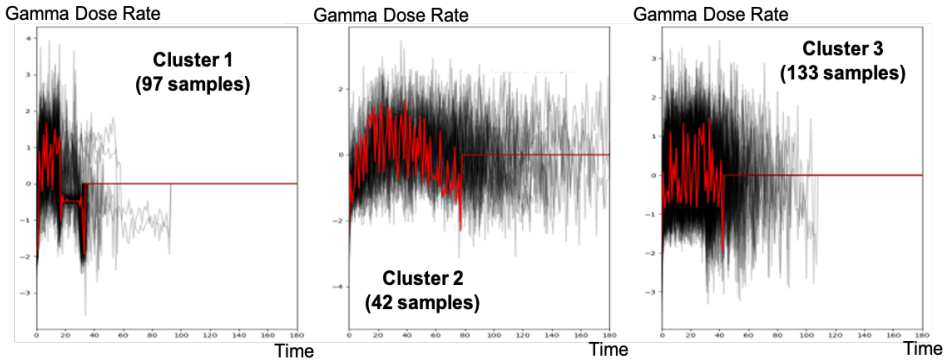

Figure 6. K-means with DTW(DBA) and z-normalized data with padding.

5. Conclusions

In this work, we presented our unsupervised machine-learning-based framework for autonomously identifying underlying events behind high gamma dose rate historical incidents. After extracting and preprocessing the extracted incidents, our machine learning model groups similarly behaving incidents caused by the same underlying event. The experts evaluated the groups, recognized the events, and labeled the incidents. The model that we have proposed is the result of an intense period of evaluations. The systematic methodology has convinced us of the foremost importance of the preprocessing phase. We believe that our proposal could be applied to other application domains, dealing with incidents of varying scale and length. To complete our *Intelligent Radiation Early Warning System*, online incidents should be classified to the proper cluster in real time. We will present our proposed solution in a future publication.

Author Contributions: M.A.S. analyzed and interpreted the need for an intelligent radiation early warning system, introduced the first phase of RIMI framework, and was a major contributor in writing the manuscript. He also queried the AI methodologies and techniques to introduce an approach that can address the shortcomings behind the RIMI first phase. B.F., Y.T., A.J. and R.L. verified the tested techniques and functions for analyzing the data and were major contributors in writing the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: The first and corresponding author "Mohammed Al Saleh" has a scholarship from the National Council for Scientific Research in Lebanon (CNRS) to continue his Ph.D. degree, including this research.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: We would like to thank the National Council for Scientific Research (CNRS) in Lebanon for supporting this work. We would like to express our gratitude to the Federal Office for Radiation Protection (BfS) in Germany for allowing us to use the data collected by their REWS for more than 15 years ago. We would also like to thank Roy Issa and Nourhan Bachir for their support in implementing the code behind this research.

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

Abbreviations

The following abbreviations are used in this manuscript:

REWS	Radiation Early Warning System
DTW	Dynamic Time Warping
DBA	DTW Barycenter Averaging
LCSS	Longest Common Subsequence

References

1. Mann, W.B. The international chernobyl project technical report: Assessment of radiological consequences and evaluation of protective measures. *Appl. Radiat. Isot.* **1993**, *44*, 985–988. [CrossRef]
2. Stöhlker, U.; Bleher, M.; Doll, H.; Dombrowski, H.; Harms, W.; Hellmann, I.; Luff, R.; Prommer, B.; Seifert, S.; Weiler, F. The German Dose Rate Monitoring Network And Implemented Data Harmonization Techniques. *Radiat. Prot. Dosim.* **2019**, *183*, 405–417. [CrossRef] [PubMed]
3. Z-Normalization of Time Series. Available online: https://jmotif.github.io/sax-vsm_site/morea/algorithm/znorm.html. (accessed on 27 January 2022)
4. Gower, J.C.: Properties of Euclidean and non-Euclidean distance matrices. *Linear Algebra Its Appl.* **1985**, *67*, 81–97. [CrossRef]
5. Myers, C.S.; Rabiner, L.R. Connected digit recognition using a level-building DTW algorithm. *IEEE Trans. Acoust. Speech Signal Process.* **1981**, *29*, 351–363 [CrossRef]
6. Liao, T. Clustering of time series data—A survey. *Pattern Recognit.* **2005**, *38*, 1857–1874. [CrossRef]
7. Zervas, G.; Ruger, S. The Curse of Dimensionality and Document Clustering. In Proceedings of the 1999 IEE Colloquium on Microengineering in Optics and Optoelectronics, Glasgow, UK, 11–12 November 1999; Volume 187, pp. 19:1–19:3.
8. Aghabozorgi, S.; Shirkhorshidi, A.S.; Wah, T.Y. Time-series clustering—A decade review. *Inf. Syst.* **2015**, *53*, 16–38. [CrossRef]
9. Kaufman, L.; Rousseeuw, P.J. *Finding Groups in Data: An Introduction to Cluster Analysis*, 1st ed.; John Wiley: New York, NY, USA, 1990.
10. MacQueen, J. Some methods for classification and analysis of multivariate observations. *Comput. Chem.* **1967**, *4*, 257–272.
11. Kaufman, L.; Rousseeuw, P.J. Clustering by means of Medoids. In *Data Analysis based on the L1-Norm and Related Methods*; Springer: Neuchatel, Switzerland, 1987; pp. 405–416.
12. Niennattrakul, V.; Ratanamahatana, C.A. On clustering multimedia time series data using K-means and dynamic time warping. In Proceedings of the International Conference on Multimedia and Ubiquitous Engineering, Seoul, Korea, 26–28 April 2007; pp. 733–738.
13. Petitjean, F., Ketterlin, A., Gançarski, P. A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognit.* **2011**, *44*, 678–693. [CrossRef]
14. Anh, D.T.; Thanh, L. An efficient implementation of k-means clustering for time series data with DTW distance. *Int. J. Bus. Intell. Data Min.* **2015**, *10*, 213–232. [CrossRef]
15. Liao, T.; Bolt, B.; Forester, J.; Hailman, E.; Hansen, C.; Kaste, R.; O'May, J. Understanding and projecting the battle state. In Proceedings of the 23rd Army Science Conference, Orlando, FL, USA, 2–5 December 2002.
16. Liao, T.W.; Ting, C.F.; Chang, P.-C. An adaptive genetic clustering method for exploratory mining of feature vector and time series data. *Int. J. Prod. Res.* **2006**, *44*, 2731–2748. [CrossRef]
17. Hautamaki, V.; Nykanen, P.; Franti, P. Time-series clustering by approximate prototypes. In Proceedings of the 19th International Conference on Pattern Recognition, Tampa, FL, USA, 8–11 December 2008; pp. 1–4. [CrossRef]
18. Gao, Y., Duan, Y., Liu, Z., Ma, C. Improved K-medoids algorithm-based clustering analysis for handle driving force in automotive manual sliding door closing process. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2020**, *235*, 871–880. [CrossRef]
19. Soleimany, G.; Abessi, M. A New Similarity Measure for Time Series Data Mining Based on Longest Common Subsequence. *Am. J. Data Min. Knowl. Discov.* **2019**, *4*, 32–45. [CrossRef]
20. Tan, C.W.; Petitjean, F.; Keogh, E.; Webb, G. Time series classification for varying length series. 2019. *Preprints and early-stage research may not have been peer reviewed yet*.
21. Hoang, A.D.; Eamonn, K.; Kaveh, K.; Chin-Chia, M.Y.; Yan, Z.; Shaghayegh, G.; Chotirat, A.R.; Yanping, C.; Bing, H.; Nurjahan, B.; et al. The UCR Time Series Classification Archive. Available online: https://www.cs.ucr.edu/~eamonn/time_series_data_2018/ (accessed on 27 Janurary 2022)