Environmental Science and Pollution Research (2022) 29:14036–14045
https://doi.org/10.1007/s11356-021-16771-5

RESEARCH ARTICLE

Introducing artificial intelligence to the radiation early warning system

Mohammed Al Saleh[1,2,3] · Béatrice Finance[1] · Yehia Taher[1] · Rafiqul Haque[4] · Ali Jaber[2] · Nourhan Bachir[2]

Received: 2 November 2020 / Accepted: 23 September 2021 / Published online: 2 October 2021

Abstract

Although radiation level is a serious concern which requires continuous monitoring, many existing systems are designed to perform this task. Radiation early warning system (REWS) is one of these systems which monitor the gamma radiation level in air. Such system requires high manual intervention, depends totally on experts' analysis, and has some shortcomings that can be risky sometimes. In this paper, the approach called RIMI (refining incoming monitored incidents) will be introduced which aims to improve this system while becoming more autonomous with keeping the final decision to the experts. A new method is presented which will help in changing this system to become more intelligent while learning from past incidents of each specific system.

Keywords Radiation · Early warning system · Data analytics · Anomaly detection · Refining incoming monitored incidents · Clustering · Prediction

Introduction

Radiation level is one of the most critical hazards that must be taken care of due to its catastrophic and persistent consequences on the environment, humans, and the other living things. Radioactive incidents and disasters such as Chernobyl (Jelewska and Krawczak 2018), Fukushima (International Atomic Energy Agency 2015), and the most recent one at Russian nuclear missile test site (Kramer 2019) raised a serious concern. These events have given rise to the need for continuous monitoring of the radiation level in the air. Since the artificial radionuclides can be transported with radioactive plume for long distances, it is important to monitor the radioactivity level within widespread geographical locations to detect any unwanted exposure. The continuous monitoring would greatly help in taking a proactive measure that would eventually raise an alert upon an occurrence of incidence. Therefore, many countries around the world raised the idea of developing several techniques for monitoring the radiation level in the environment to detect any abnormal release or discharge (Szegvary et al. 2007). These countries developed a national environmental radiation monitoring programs to establish radiation baseline level and determine trend of radiation level (Stöhlker et al. 2019). Air monitoring was one of the main scopes of these programs (El Samad et al. 2016).

Responsible Editor: Philippe Garrigues

✉ Mohammed Al Saleh
mhdalsaleh@gmail.com

Béatrice Finance
beatrice.finance@uvsq.fr

Yehia Taher
yehia.taher@uvsq.fr

Rafiqul Haque
rafiqul.haque@intelligencia.fr

Ali Jaber
ali.jaber@ul.edu.lb

Nourhan Bachir
n.bachir@st.ul.edu.lb

[1] David Laboratory, University of Versailles (UVSQ), 45 avenue des Etats-Unis, 78035 Versailles, France

[2] Lebanese University, Rafic Hariri University Campus, Al Hadath, Beirut, Lebanon

[3] Lebanese Atomic Energy Commission (LAEC), National Council for Scientific Research (CNRS), Airport Road, P.O.Box 11-8281, Beirut, Lebanon

[4] Intelligencia R&D, Paris, France

There exist different approaches to monitor and analyze the data of high radiation levels. Among them is the radiation early warning system (REWS) (Thieu et al. 2012) that is a widely used network system which exists in many countries around the world (Küçükarslan et al. 2004) (Biegalski et al. 2001) (El Samad et al. 2016). The REWS is composed of many radiation detection sensors that contain two Geiger Muller tubes: one for detecting high dose rate measurements and the other for low dose rate measurements. Both are incorporated in what is called probe. These probes are disseminated on a specific region that monitors the gamma radiation level. This system reacts as soon as possible to anomalies by raising an alert (Zähringer and Sempau 1996). Typically, the alerts are determined by predefined threshold values that are essentially chosen based on observations (i.e., experience) (Dombrowski et al. 2017). It is worth noting that there are different threshold values at different locations since the threshold value depends strictly on the normal reading of the radiation level (known as background level) which is in turn is not fixed due to many factors such as soil composition, soil moisture, and others. For example, in some locations, rocks could be rich in natural radionuclides which can increase the background value. In addition, altitude and gamma radiation level have incremental relation, as higher altitude results in higher gamma radiation readings due to cosmic rays. Once an alert is raised, it needs to be checked by an expert. Indeed, the expert needs to analyze the potential causes for the incident as some alerts refer to an authentic threat of high radiation level and others denote the rise of radiation level that has no hazardous impact on the environment or living beings. In order to do so, the expert will consult additional information such as readings from other probes in the vicinity, the weather broadcast, and the quality factors (also called *quality bits*) of the probe. For instance, the alert is *false* when the quality bits of the probe indicate that there is a defect in the probe (Casanovas et al. 2011), meaning that we cannot trust the collected gamma dose rate value. The alert is *innocent* when external factors have occurred such as rain, wind, and lightening. For example, the rain could cause wet deposition of radionuclides dispersed in the atmosphere and hence increased readings which will result in a peak value. These external factors are the more difficult to analyze, but they represent more than 90% of the alarms. Finally, if the alert is *real*, an emergency action needs to be taken by the authority immediately.

Existing REWS systems have various shortcomings. The most critical one is the manual intervention of the expert that is heavily time-consuming, labor-intensive, and risk-prone. Indeed, when an alarm is raised, a considerable amount of time is consumed, and efforts are exerted by the expert to analyze the parameters that are stemming from external data sets such as weather data sets in order to classify the alert as *false, innocent or real*. If there is no automated data collector, the experts must carry out data searching and data fetching operations manually. Moreover, most of the time, the expert cannot classify the alert immediately. This can take hours due to some parameters such as rain. Therefore, it is not possible to make a faster or real-time inference using the current methodology.

Today, we assist to the explosion of machine learning techniques and complex algorithms in order to help experts or non-experts to analyze and understand more about their data. Machine learning techniques might help building predictive models in order to have a real-time proactive system (Alanazi et al. 2017). However, in order to apply these techniques, some preliminaries analysis should be done to better characterize the problem that needs to be solved. The main objective of this research is to analyze REWS and see if the expert can be partly removed from the picture and replaced by an autonomous REWS. There are many challenges to address before reaching this goal. The work described in this paper is the first attempt to do so, as to our knowledge it does not exist autonomous REWS in the literature.

The main objective of this research is to develop an end-to-end solution that will be integrated with running REWS systems without any disruption or without replacing its task completely. This work aims at the beginning to come up with a specific solution depending on each system old data. Later on, the suggested framework will be checked if it could be generalized on all the REWS systems around the world. Indeed, before replacing the expert, the system should prove its accuracy to predict the right answer. Thus, a supervised learning should take place at the beginning until it reaches its full potential and work on its own. In this paper, RIMI framework (refining incoming monitored incidents) is presented that highlight the different steps that need to take place before reaching an autonomous REWS solution. In this framework, a list of components will be developed from data acquisition and normalization, to building a predictive model on a real data set produced by a running REWS, then by using it to predict the right classification of the alarm on real-time data. In this work, we started to work with the data of each probe separately. However, the next step will be trying our approach on the data coming from a grouped several probes.

The remaining of this paper is organized as follows. The "Problem description" section will highlight the nature of the problems that need to be solved in REWS. In "The RIMI framework" section, RIMI framework is described and each of its components is detailed. Finally, our work is concluded in the "Conclusion" section. Notice, that there will not be a related work section in itself, as it does not exist similar work in the literature but rather, once a problem has been refined some hints of the approaches that have been proposed in the literature to solve this particular problem will be shared.

Problem description

In this section, some of the scenarios that an expert will encounter during his/her work are illustrated. These scenarios reflect the variety of the situations that occur most of the time.

The scenario shown in Figure 1 illustrates how an internal factor can affect the gamma dose rate. In this case, the low dose tube of the probe is broken. It affects the quality of the gamma dose rate value. Depending on the system, this type of scenario may produce a *false* alarm.

In the scenarios, respectively, shown in Figures 2 and 3, wind, electromagnetic, perturbations, or lightnings directly and immediately impact the signal coming from the probe. In these scenarios, we observe many peaks that do not last very long. We called them *hard parabola*. These types of scenarios may produce a *false* alarm.

On the opposite, the rain impacts the gamma dose rate in a completely different manner. Rain, for example, can cause wet deposition of the dispersed radionuclides in the atmosphere. It causes the soil to emit radioactive gases into the air resulting a true innocent high gamma radiation reading (Hõrrak et al. 2021). The increased gamma dose rate will return to normal values after specific time. Sometimes, even if it continues to rain after the peak, it will not affect the gamma dose rate anymore because the atmosphere is already washed out. This behavior is described in the Figure 4 and corresponds to what is called *soft parabola*. It is classified as an *innocent* alarm.

Fortunately, *real* alarms are very rare, but as you can imagine, the peak will not decrease after a short period of time but will stay at high level or even continue to increase. Many other scenarios can also be found in practice. For instance, some factors like earthquakes or a truck with radioactive materials load passing near a probe can cause the gamma dose rate level to increase immediately. Moreover, multiple factors can be combined together such as rain and wind making the recognition of the cause less easier.

Many data sources should be combined together. Some are collected in a continuous manner by the REWS and stored in an historical database. But many others data sources must be queried on demand when an investigation is launched by an expert. Combining all these heterogeneous data sources on the fly is also a difficult problem in itself.

Another dimension of the problem concerns the variability of the threshold values that evolve over time and that is also dependent on the location of the probe itself (Kessler et al. 2018) . As said earlier, these predefined threshold values are essentially chosen based on observations or experience at the beginning, but they evolve slightly over time, making the comparison of the time series over multiple months not an easy task.

All these examples illustrate the difficulty and the heterogeneity of analyzing the gamma dose rate shape and understanding its causes in order to classify properly the alarm in an automatic way. For all these reasons, we believe that the research problem is interesting to be tackled as it will require many different techniques or approaches to be used. This is the reason why we define the RIMI framework to offer an end-to-end solution towards an autonomous REWS.

The RIMI framework

In this section, a detailed description of our framework entitled, RIMI (refining incoming monitored incidents) is provided. The framework consists of three main components: (1) the data collector, (2) the building of the predictive model, and (3) the online detection and prediction. Figure 5 illustrates more

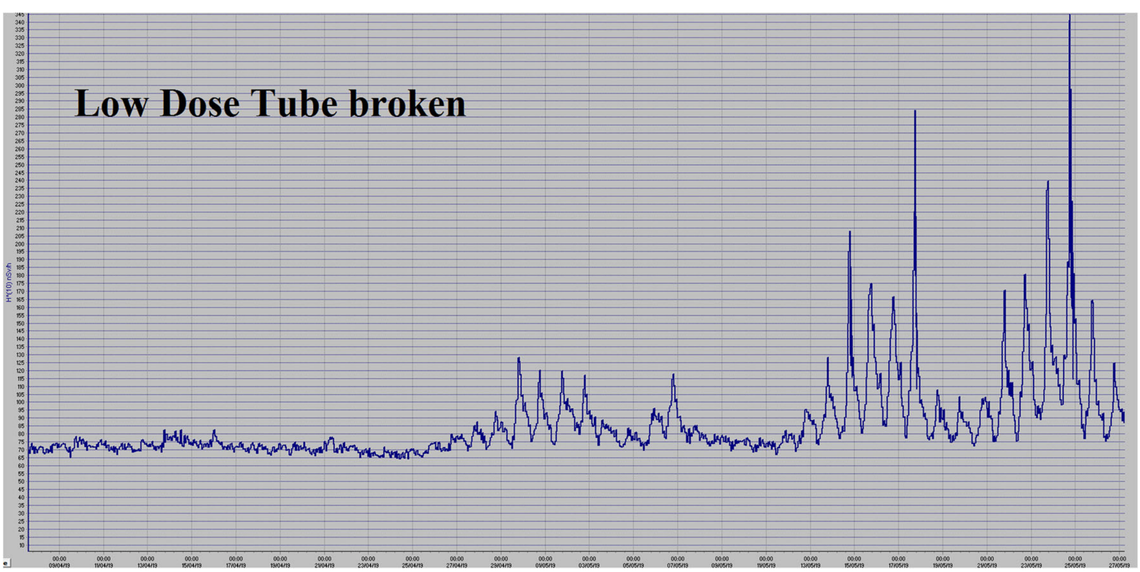


Fig. 1. Low-dose tube broken effect on gamma dose rate

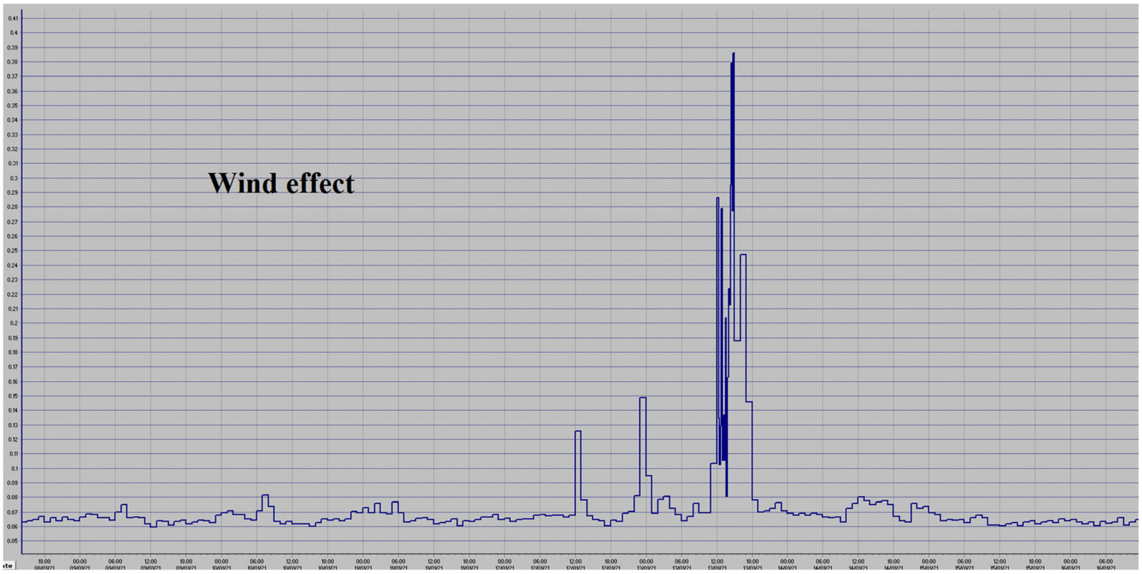


Fig. 2. Wind effect on gamma dose rate

in detail each of the main components which are incident extractor, incident clustering model, incident causes identifier, data enrichment, and incident pattern matching. As the incident is caused by a high gamma dose rate level which can be harmful for humans and environment, this framework aims to replace a human-driven verification system that refines the incoming incidents and alerts and detect its cause by doing it automatically with a high level of accuracy.

A. Data collector

As seen in the previous section, the raw data is heterogeneous (i.e., time series, quality bits, events) and comes either from the online REWS monitoring system automatically or from the external data sources, noting that the data coming from extremal data sources (such as weather forecast data) could be queried on demand by experts in the case of a triggered alarm. The data collected by the REWS system is stored in a historical database. The data acquisition is done on a regular basis through secure channels between the radiation detection sensors (i.e., probes) and the server. Usually, the probe sends a message containing the gamma dose rate average every minute. In addition to the gamma dose rate, the probes send other sensors data like temperature and quality bits as they are equipped with internal sensors that can detect the defectiveness of any of the system components. These sensors data are stored in the historical REWS database for later analysis. As seen earlier in one of the scenarios, the defect of a tube can cause a direct false high gamma radiation level.

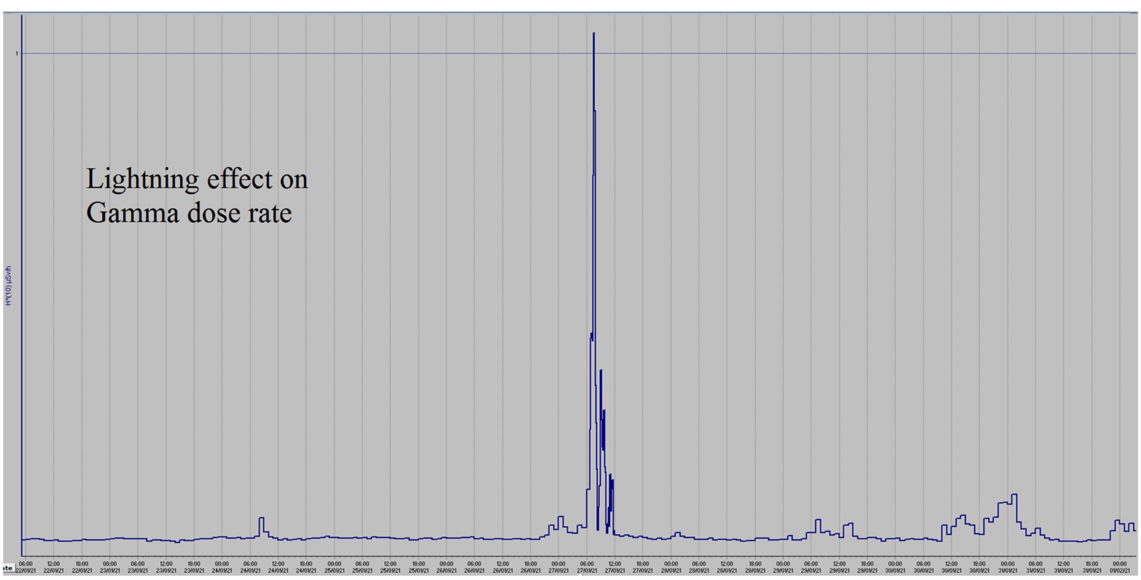


Fig. 3. Lightning effect on gamma dose rate

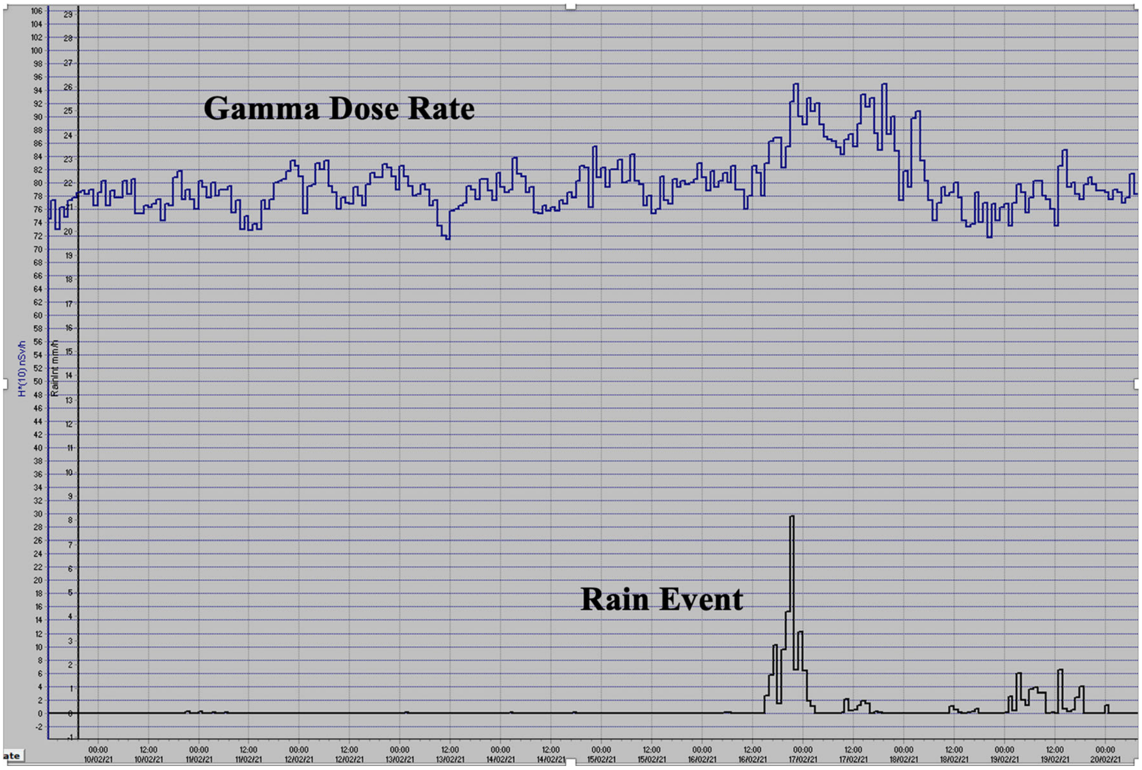


Fig. 4. Rain effect on gamma dose rate

For historical data, we decided to use the data collected by the German REWS. More than 15 years of historical data have been collected by the REWS that is used in Germany and controlled by the Federal Office for Radiation Protection (Bfs). These data are precious and can be used to train our predictive model, but these data need to be analyzed by machine learning algorithms in order to automatically form categories for similar incidents. Later, many data sources can be queried on demand to identify the causes for similar incidents. These numerous data sources can be a weather database, a radiation transportation database, etc. In order to be queried, the approximate timestamp of the alarm should be known in order to better understand the past context or situation in which the alarm was triggered.

B. Building the predictive model

The predictive model is built on the historical databases produced by the running REWS, which continuously collect sensors data produced on the different probes. First, there is a need to identify all the incidents. Second, similar incidents should be grouped together. This requires the data to be clustered into different classes that contain similar incidents, so that we may search later for the common cause behind these similar incidents. Finally, all this mass of information should be organized and classified in order to build our predictive model that will be used at run-time.

1) *Incident Extractor*: Incident extraction consists in analyzing the gamma dose rate time series data in order to identify a fragment (i.e., shape). A fragment corresponds in fact to a triggered alarm (i.e., incident). As said earlier, the threshold and the background values are not fixed, but they evolved over time. At the beginning of the system, a value is given, but it is refined over time to better suit the default gamma dose rate of the location on which the probe is installed. This value called the background can be different from one location to the next. At the end of each month, the average of the background values is calculated to find the background mean. This mean will be used for finding the background interval of each specific station. It is important here to mention that this mean will be calculated after removing the threshold values from the month data set. To find the threshold value, we noticed that experts in different countries depend on different methods. Some may consider that values that are equal and greater than 1.5 times the background mean as thresholds. Others refer to the values that are equal and greater than 2 or 3 times the background mean as thresholds (Farid et al. 2017) (Stöhlker et al. 2019). We decided to be more precise and rely on the 1.5 method knowing that this value can be changed to suit experts' expectations through different countries. Several methods were explored to find the most suitable one that determines the *lower* and the *upper* bounds of the background interval. Our study revealed that the *standard deviation* (Barde and

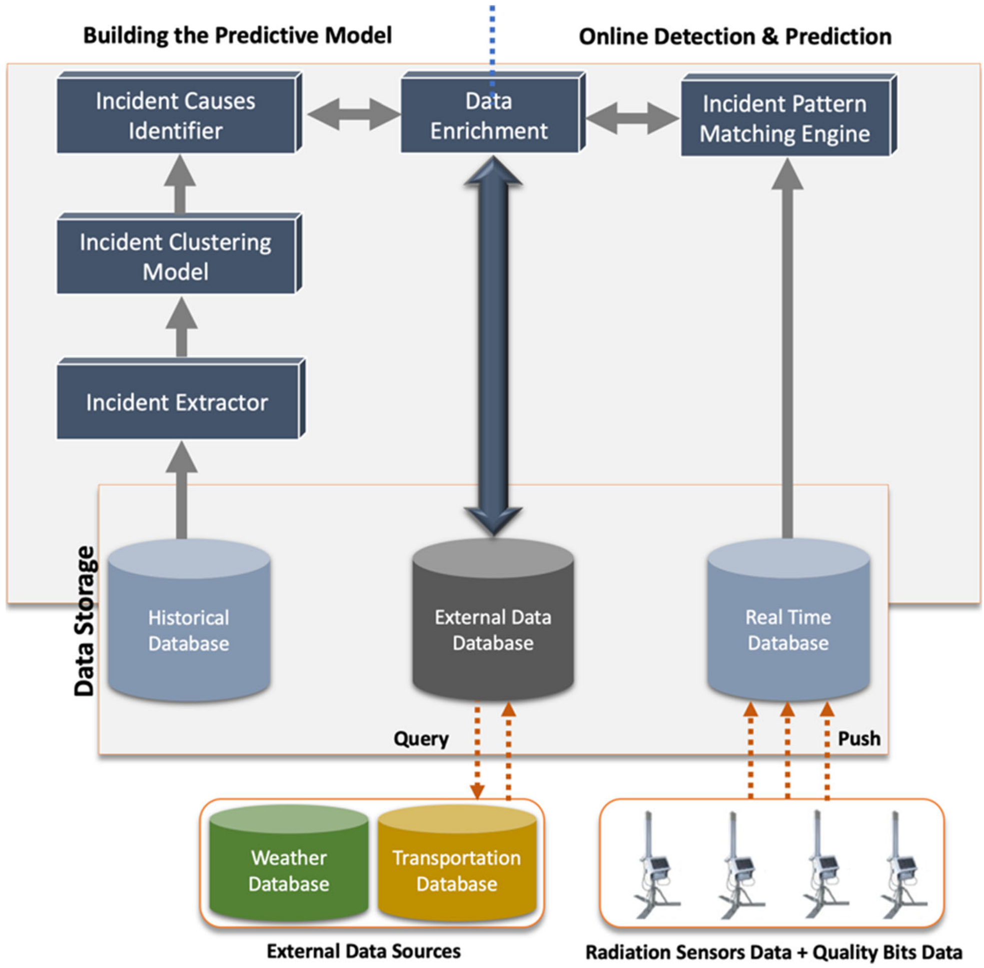


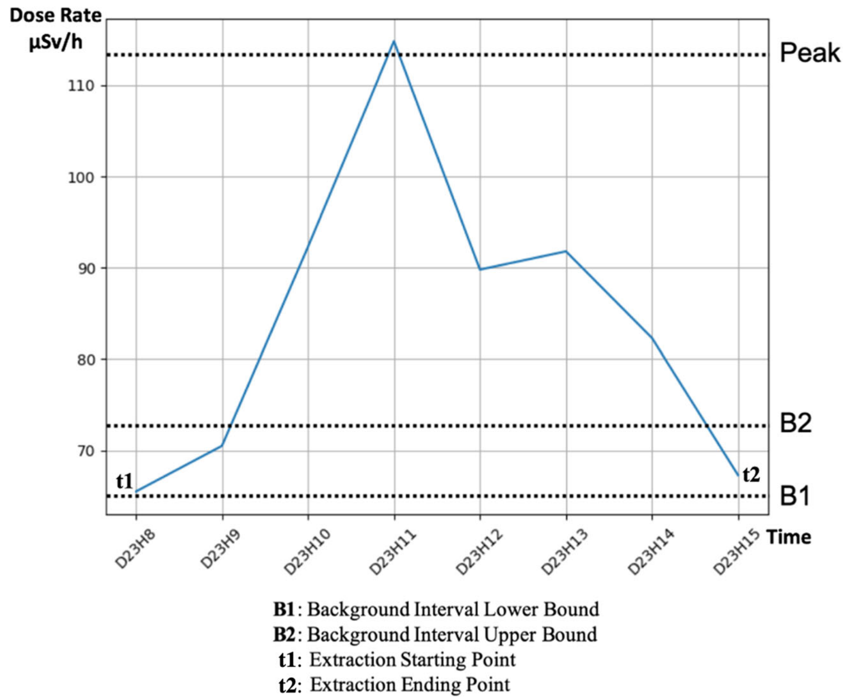
Fig. 5. A framework for real-time radiation detection (RIMI framework)

Barde 2012) is promising to find the background level interval. We choose standard deviation because of the nature of the distributions of data. According to our observation, radiation level data are uniformly distributed, and to the best of our understanding, standard deviation is a suitable technique for finding intervals when data are uniformly distributed in a two-dimensional graph. This background level interval is calculated by adding and subtracting the value resulted in by calculating the standard deviation to the mean of the background values in the current month. This computation model produces a catalog of parameters with the corresponding means, thresholds, and the background intervals values for each month. Thus, the incident extractor component relies on the catalog which defines the appropriate background interval values for each month. We assume that this catalog is fully computed on historical data before the extraction starts.

In the Figure 6, the background interval that corresponds to the acceptable background values is defined. It is represented by a *lower B1* and an *upper B2* bounds. We also show the threshold value which is 1.5 times the background mean.

A fragment is defined by a beginning *t1* and an end *t2* timestamps. In other words, once a threshold value is found, the incident extractor will search for the nearest points *t1* and *t2* that represent the preceding and succeeding values of the threshold and extract the current fragment from *t1* till *t2*. Note that these two values must lie within the background level interval. They respectively identify the time when the gamma dose rate starts to increase in an abnormal way and when its return to a normal state. However, the incident extractor should monitor the incident for 15 min after returning to the background level before extracting it. This will assure that the incident is totally ended since the data is collected every 1 min. Moreover, it is worth noting that a locking mechanism is designed so that it does not allow the incident extractor to start a new extraction operation unless the previous one is completed. The locking mechanism was used because a graph may contain more than one fragment exceeding the threshold value. Thus, incident extractor extracts these fragments sequentially and the endpoint of the preceding fragment may become the starting point for succeeding fragment.

Fig. 6. Fragment extractor



Shapelet extraction has drawn significant research attention, in recent years. Many algorithms have been proposed in the literature such as Piecewise Aggregate Approximation (PAA) (Wang et al. 2018) and the Multivariate Shapelets Detection (MSD) (Ghalwash and Obradovic 2012). These approaches search for shapelets that are similar to a referent shapelet. These approaches are not suitable to our problem as we do not have any referent shapelet, and due to the evolving of the background level, the duration of our fragment can be multiple. Moreover, some of the approaches go beyond that and discuss extracting shapelets based on predefined key points (Guiling et al. 2019). These methods aim at detecting key points in the time series and then extract the shapelets referring to these key points. Such approaches need to be investigated more in order to check their compatibility with our evolving background interval.

2) *Data Preparation*: Before any analysis phase, a lot of time have to be spent studying the data to find the proper preparation algorithms to apply (Tan et al. 2005) (Han and Kamber 2006; Pyle 1999). The preparation algorithms are to ensure the quality and cleanliness of the data before inserting it into any analysis algorithm.

Inspecting the incidents, the two major observations were that the incidents are of vastly varying length and the incidents are not at the same level. The incidents' length varies very widely even when extracted from the same probe's data. The variation is because each incident's length is dependent upon the duration of the underlying event's effect; it could be as short as three minutes long (i.e., lightening) and as long as 6 h long (i.e., storm). On the other hand, different probes have different acceptable range of background radiation rate. Each probe's normal readings' range depends upon its environment, the nature of the soil, and many other factors. Additionally, with time, if the latter factors change, the range of normal values evolves too on the same probe. Because of that, the extracted incidents are not all at the same "level."

Researching these characteristics, we found that processing samples of different length have little to no mention in the literature. Only one very recent work (Tan et al. 2019) has been done on classification of time series of varying length. For preprocessing, they suggest either scaling the samples to be of similar length for comparison or prep-ending/appending inconsequential data points like zero or the mean value depending on the nature of the data. On the other hand, regarding the other characteristic, studies focused on applying normalization methods. Specifically, z-normalization (Rakthanmanon et al. 2012) appears to be the most popular (Gujarati and Porter 2009; Kendall 1976) when considering time series samples from different devices having different scales as the samples become all of mean zero with a variation between − 1 and 1.

Another problem with the data would be the missing data when some data points fail to be saved over a certain period of time. Proper data re-sampling methods have to be applied taking into consideration not inserting any biased data points. More investigation and research have to be done for the specific characteristics of the data.

3) *Clustering*: Once the fragments are extracted and pre-processed, we need to classify them in order to build and train our predictive model. We will start by comparing the different incidents extracted in order to set different classes. The clustering process will start forming several classes by following a model-based approach. Through this approach, the incidents extracted from the gamma dose rate time series will be compared based on the graphical shapes representing the incidents. Several incident shapes will be gathered and classified as soft or hard parabola shapes. Our approach is to use time series clustering to be able to infer the different classes (clusters) behind the incidents' causes. Some expected problems will be faced through the clustering process because we have samples of different length which is not dealt with very much in the literature of time series clustering.

Existing time series clustering approaches proposed in literature use different techniques. In (Liao 2005), Warren Liao surveys the different approaches in time series clustering literature. The author describes the most popular similarity measures as well as the clustering algorithms that are most frequently used in the literature. Euclidean (Julazadeh et al. 2012) and Mahalanobis (Arathi and Govardhan 2014) distances are lock-step similarity measures that are very much successful in clustering literature; however, in time series clustering specifically, they have their drawbacks such as their inability to encompass the temporal aspect of time series data. These drawbacks draw attention to elastic measures which are not as rigid. For instance, the DTW (dynamic time warping) algorithm (Huiqing et al. 2018), which was introduced specifically for time series analysis, focuses on reducing the computing complexity, and improves efficiency. On the other hand, time series clustering approaches are majorly based on the classic clustering algorithms such as k-means (MacQueen 1967), k-medoids (Kaufmann and Rousseeuw 1987), and hierarchical (Sarda-Espinosa 2018). In addition, the state-of-the-art approaches in time series clustering will be investigated to check all potential similarity measures and clustering algorithms and models to find what best befits our data.

After applying the clustering process, we will end up with a wide range of classes that we may need to identify their relevant causes. These classes will be referred to in the online detection of the incident.

C. Online detection and prediction

Once the classes are defined, the online detection and prediction phase can take place. Its job starts when an incident alarm is triggered. As said earlier, the alarm can be categorized into three types: false, innocent, or true. The aim behind this phase is to check if the current incident occurring with the specific annotated information corresponds to a predefined class as quickly as possible. The main job of the online predictive model will be calculating the context by checking what is the current situation once an incident is captured. Based on the discovered situation, the model will start searching for similar incidents in the related classes.

1) *Incident Pattern Matching Engine*: The incident pattern matching engine is the analytical engine deployed to recognize the new incidents by performing a matching operation over the annotated patterns produced at the incident clustering step. This analytical engine is designed based on Kappa architectural style (Ounacer et al. 2017) which means that the incident pattern matching operations will be performed in real-time. At the same time, a pre-designed algorithm will be running in the background to calculate the parameters that will be used for next month's evaluation. On incident detection, the framework will search for the current situation, since it will be always connected to external and internal factors databases, and will run the incident matcher phase, which will look to the nearest point before the threshold belonging to the background level, and start matching the beginning of the current shapelet with those represented in the predefined classes.

For example, Figure 7 shows the analyzing process for incoming data. As we can notice, the readings started within the background level interval which means a normal situation with accepted values. Once the readings exceed the upper bound of the background interval (B2) and reached the threshold value, then the alarm will be triggered detecting incident case and the matching process starts. This will provide different possibilities for the continuity of the current shapelet referring to the already obtained shape after comparing it to the previously classified incidents. It will repeat this process until the possibilities become so limited that the cause can be detected. Thus, the framework will be able as soon as possible to detect the cause behind the incident and alert the experts if special procedures must be taken. To perform this task, the techniques defined in the incident extractor module can be used.

2) *Accuracy and Verification*: The objective of our research is to propose a fully automated framework. However, we strongly believe that at the final stage, the solution needs an expert opinion to validate the results produced by the system. This validation is important due to the sensitivity of the use cases that will be implemented using this solution. This will help in increasing the accuracy rate of the proposed framework. Moreover, in case of exceptional

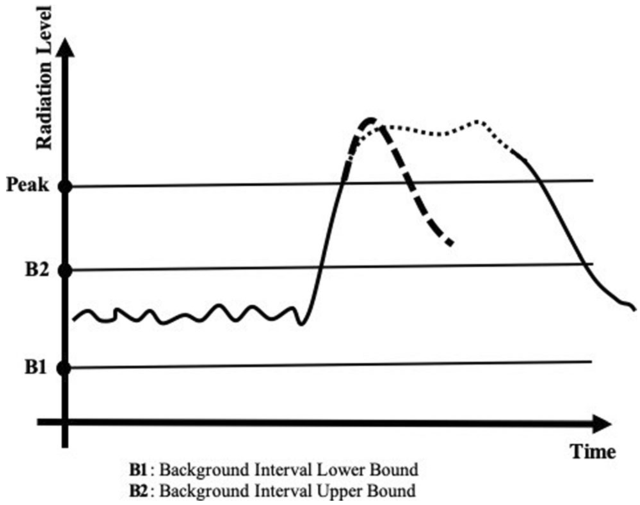


Fig. 7. Fragment matcher

use cases that were not known, the involvement of the experts would help to enhance the solution by training the classifier over the data and make it capable of recognizing incident patterns that were unknown before.

Conclusion

This paper presented an end-to-end framework for (pre)-processing, processing, and analysis of ambient gamma dose rate in air. The objective of developing this framework is to reduce the manual intervention in radiation early warning systems. In this paper, the key components of the framework including data pre-processor, incident extractor, incident classifier, and data enhancer were explained. A detailed description of an analytical engine, which matches the fragment patterns in real-time and helps the experts in faster decision-making regarding verification of an alarm, was provided. Several works have been lined up for future. In the near future, we planned to develop techniques for classifying the incidents. The classifier will be trained, tested, and optimized to guarantee that accuracy of classification. Also, a real-time analytical engine using advanced tools for performing classification in real-time was developed.

Acknowledgements We would like to express our thankful to 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 giving us the permission for using the data collected by their REWS for more than 15 years ago.

Availability of data and materials The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contribution MS analyzed and interpreted the radiation early warning system data and discovered some shortcomings. He also introduced the RIMI framework and was a major contributor in writing the manuscript. RH and NB queried the AI methodologies and techniques to find some approaches that can address the shortcomings. BF, YT, and AJ verified the tested techniques and functions for analyzing the data and were major contributors in writing the manuscript. All authors read and approved the final 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 PhD degree including this research.

Declarations

Ethics approval and consent to participate Not applicable

Consent for publication Not applicable

Competing interests The authors declare no competing interests.

References

Alanazi H, Abdullah A, Qureshi K (2017) A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care. Journal of Medical Systems 41:69. https://doi.org/10.1007/s10916-017-0715-6

Arathi M, Govardhan A (2014) Accurate time series classification using shapelets. International Journal of Data Mining & Knowledge Management Process 4:39–47. https://doi.org/10.5121/ijdkp.2014.4204

Barde P, Barde P (2012) What to use to express the variability of data: standard deviation or standard error of mean? Perspectives in Clinical Research 3:113–116. https://doi.org/10.4103/2229-3485.100662

Biegalski S, Shipman G, Mason L, Pickering M, Altman S (2001) Caribbean radiation early warning system (CREWS). Journal of Radioanalytical and Nuclear Chemistry 248:637–642. https://doi.org/10.1023/A%3A1010624224587

Casanovas R, Morant J, López M, Hernández-Girón I, Batalla E, Salvadó M (2011) Performance of data acceptance criteria over 50 months from an automatic real-time environmental radiation surveillance network. Journal of Environmental Radioactivity 102(8):742–748. https://doi.org/10.1016/j.jenvrad.2011.04.001

Dombrowski H, Bleher M, De Cort M, Dabrowski R, Neumaier S, Stöhlker U (2017) Recommendations to harmonize European early warning dosimetry network systems. Journal of Instrumentation 12: 12024. https://doi.org/10.1088/1748-0221/12/12/P12024

El Samad O, Baydoun R, Aoun M, Zaidan W, El Jeaid H (2016) Public exposure to radioactivity levels in the Lebanese environment. Environmental Science and Pollution Research 24(2):2010–2018. https://doi.org/10.1007/s11356-016-7911-7

Farid M, Prawito SI, Yuniarto A (2017) Design of early warning system for nuclear preparedness case study at Serpong. AIP Conference Proceedings. https://doi.org/10.1063/1.4991171

Ghalwash M, Obradovic Z (2012) Early classification of multivariate temporal observations by extraction of interpretable shapelets. BMC Bioinformatics 13:195–206. https://doi.org/10.1186/1471-2105-13-195

Guiling L, Wenhe Y, Zongda W (2019) Discovering shapelets with key points in time series classification. Expert Systems with

Applications 132:76–86. https://doi.org/10.1016/j.eswa.2019.04.062

Gujarati D, Porter D (2009) Basic econometrics. McGraw-Hill Irwin, USA

Han J, Kamber M (2006) Data mining concepts and techniques, 2nd edn. Morgan Kaufmann, USA

Hõrrak U, Chen X, Hõrrak K, Rand U, Komsaare K, Vana M, Luts A, Junninen H (2021) The rain induced gamma-radiation dose rate enhancement at Järvselja SMEAR station. EGU General Assembly 2021:EGU21–EGU15644. https://doi.org/10.5194/egusphere-egu21-15644

Huiqing W, Li C, Sun H, Guo Z, Bai Y (2018) Shapelet classification algorithm based on efficient subsequence matching. Data Science Journal 17:1–12. https://doi.org/10.5334/dsj-2018-006

International Atomic Energy Agency (2015) The FukushimaDaiichi accident. IAEA, Vienna

Jelewska A, Krawczak M (2018) The spectrality of nuclear catastrophe: the case of Chernobyl. politics of the machines - art and after. https://doi.org/10.14236/ewic/EVAC18.30

Julazadeh A, Marsousi M, Alirezaie J (2012) Classification based on sparse representation and Euclidian distance. IEEE Visual Communications and Image Processing (VCIP). https://doi.org/10.1109/VCIP.2012.6410815

Kaufmann L, Rousseeuw P (1987) Clustering by means of medoids. data analysis based on the L1-norm and related methods 405-416

Kendall M (1976) Time-series. Charles Griffin, London

Kessler P, Behnke B, Dabrowski R, Dombrowski H, Röttger A, Neumaier S (2018) Novel spectrometers for environmental dose rate monitoring. Journal of Environmental Radioactivity 187:115–121. https://doi.org/10.1016/j.jenvrad.2018.01.020

Kramer A (2019) Russia orders evacuation of village near site of nuclear explosion. The New YorkTimes. https://www.nytimes.com/2019/08/13/world/europe/russia-nuclear-explosion-accident.html. Accessed 13 August 2019

Küçükarslan N, Erdoğan A, Güven A, Gülay Y (2004) Early warning environmental radiation monitoring system. In: Zaidi MK, Mustafaev I (eds) Radiation safety problems in the caspian region. Nato Science Series: IV: Earth and Environmental Sciences, vol 41. Springer, Dordrecht. https://doi.org/10.1007/1-4020-2378-2_7

Liao T (2005) Clustering time series data – a survey. Pattern Recognition 38(11):1857–1874. https://doi.org/10.1016/j.patcog.2005.01.025

MacQueen, J (1967) Some methods for classification and analysis of multivariate observations. In: BSMSP 281–297.

Ounacer S, Talhaoui M, Ardchir S, Daif A, Azouazi M (2017) A new architecture for real time data stream processing. International Journal of Advanced Computer Science and Applications 8(11). https://doi.org/10.14569/IJACSA.2017.081106

Pyle D (1999) Data preparation for data mining. Morgan Kaufmann, USA

Rakthanmanon T, Campana B, Mueen A, Batista G, Westover B, Zhu Q, Zakaria J, Keogh E (2012) Searching and mining trillions of time series subsequences under dynamic time warping. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. https://doi.org/10.1145/2339530.2339576

Sarda-Espinosa A (2018) Comparing time-series clustering algorithms in r using the dtwclust package. In: Manual of the R package dtwclust

Stöhlker U, Bleher M, Doll H, Dombrowski H, Harms W, Hellmann I, Luff R, Prommer B, Seifert S, Weiler F (2019) The German dose rate monitoring network and implemented data harmonization techniques. Radiation Protection Dosimetry 183(4):405–417. https://doi.org/10.1093/rpd/ncy154

Szegvary T, Conen F, Stöhlker U, Dubois G, Bossew P, Vries G (2007) Mapping terrestrial gamma-dose rate in Europe based on routine monitoring data. Radiation Measurements 42(9):1561–1572. https://doi.org/10.1016/j.radmeas.2007.09.002

Tan P, Steinbach M, Kumar V (2005) Introduction to data mining, First edn. Addison Wesley Longman, USA

Tan C, Petitjean F, Keogh E, Webb G (2019) Time series classification for varying length series. Researchgate. https://www.researchgate.net/publication/336411021_Time_series_classification_for_varying_length_series. Accessed 10 October 2019

Thieu D, Toan T, My N, Sy N, Tien V, Mai N, Cuong L (2012) Study, design and construction of an early warning environmental radiation monitoring station. Communications in Physics 22:375–382. https://doi.org/10.15625/0868-3166/22/4/2665

Wang H, Li C, Sun H, Guo Z, Bai Y (2018) Shapelet classification algorithm based on efficient subsequence matching. Data Science Journal 17:1–12. https://doi.org/10.5334/dsj-2018-006

Zähringer M, Sempau J (1996) The assessment of the representativeness of data from dose rate monitoring stations. Radiation Protection Dosimetry 64:275–288. https://doi.org/10.1093/oxfordjournals.rpd.a031584

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.