Journal of Magnetic Resonance 305 (2019) 104–111

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Spectral descriptors and supervised classifier for ammonium nitrate detection in landmines by nuclear quadrupole resonance

Lorena Cardona [a,*], Hideo Itozaki [b], Jovani Jiménez [c], Nelson Vanegas [d], Hideo Sato-Akaba [b]

[a] Department of Mechanics, School of Engineering, Institución Universitaria Pascual Bravo, Calle 73 No. 73A – 226, Medellín, Antioquia, Colombia
[b] Department of Systems Innovation, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan
[c] Department of Computation and Decision Science, School of Mines, Universidad Nacional de Colombia, Carrera 80 No. 65-223, Medellin, Antioquia, Colombia
[d] Department of Mechanical Engineering, School of Mines, Universidad Nacional de Colombia, Carrera 64 No. 63–120, Medellin, Antioquia, Colombia

ARTICLE INFO

Article history:
Received 27 February 2019
Revised 7 June 2019
Accepted 17 June 2019
Available online 20 June 2019

Keywords:
Nuclear quadrupole resonance
NQR
Landmine detection
Ammonium nitrate
Spectral descriptors
Supervised classification

ABSTRACT

The high specificity of Nuclear Quadrupole Resonance (NQR) makes it very suited for the detection of antipersonnel mines, where the intensity of the signal spectrum around the resonance frequency of the target substance is the standard decision parameter; however, radiofrequency interference, soil effects on the search coil, landmine size, burial depth, and target temperature affect signal intensity. To overcome this, the use of spectral descriptors and a supervised classifier are proposed in this work, where an assembly of decision trees was trained with NQR data collected on places where a target filled with ammonium nitrate was present and where it was not. A statistical test, comparing the proposed classifier and the solution based solely on the intensity of the signal spectrum, showed with significant evidence that the proposed classifier outperforms the traditional solution. A final blind experiment was conducted in a rural region of Colombia, where five landmines of different size filled with ammonium nitrate were shallowly buried in an area of 1.9 × 1.52 m, and the system with the proposed classifier detected four of them with three false alarms. This work is also novel in detecting ammonium nitrate in antipersonnel mines, which are typical in Colombia, the second most mined country in the world.

1. Introduction

Antipersonnel landmines are weapons, usually buried, that explode when a person steps on them and are designed to kill or injure people, leaving long-term physical and psychological effects [1]. Colombia ranks second in the world by the number of victims from landmines [2], mainly because landmines are used by groups outside the law to protect coca plantations and to counter the army [3]. Colombian mines are very varied, difficult to characterize, non-metallic, and, due to government controls, they are usually filled with ammonium nitrate (NH_4NO_3) mixed with different fuels instead of common military explosives [4].

Recently, a portable system for detection of Ammonium Nitrate (AN) by Nuclear Quadrupole Resonance (NQR) was developed [5], which could be used for landmine detection in Colombia. NQR is a spectroscopic technique that works by sending radio frequency pulses through a coil to excite quadrupolar nuclei (nuclei having spin number greater than ½) in the target substance; and after excited, the nuclei emit a signal which could be detected (usually, by the same coil). As the excitation frequency and the spectrum recorded from each substance is unique, NQR is considered a very specific technique [6].

NQR detection of AN poses a challenge: resonance frequency of AN is very low (423.6 kHz), leading to a weak NQR signal [7]. An improvement in SNR could be achieve by successively adding scans, as NQR signals add coherently while noise adds in random phase. However waiting time between scans for signal averaging and SNR improvement should be around$5T_1$ (where T_1 is the longitudinal relaxation time of the target substance) and for AN, T_1 is quite long ($T_1 = 16$ s). To overcome this issue, excitation sequences like Spin lock – spin echo (SLSE) and steady state free precession (SSFP) have been established that allow for increasing the SNR by averaging multiple acquisitions in times shorter thanT_1. This is particularly useful for substances with longT_1, like AN, for which the use of a pulse sequence is mandatory in order to get an acceptable SNR in a reasonable time. However, even with the use of pulse sequences, there are factors that harm the SNR when detecting landmines, like the landmine burial depth (as NQR signal strength highly decreases with distance); environmental noise; changes in the resonance frequency of the substance with temper-

* Corresponding author.
E-mail address: lorena.cardona@pascualbravo.edu.co (L. Cardona).

https://doi.org/10.1016/j.jmr.2019.06.009
1090-7807/

ature; and changes in the resonance frequency of the coil when near the soil surface.

Some algorithms have been proposed to improve NQR signal detection, most of them published by the authors Somasundaram, Jakobsson et al. [8–12]. They model each individual echo of the NQR signal (after excitation with a Spin-lock Spin-echo sequence) and match the acquired data to the model with a maximum likelihood algorithm. From this approach, the authors made subsequent improvements to include temperature effects [10], the presence of various polymorphs [8] and for dealing with radiofrequency interference [11]. A drawback on these algorithms is that they work over the whole echo train of a SLSE sequence, meaning that the echoes cannot add up. This makes them very difficult to apply for detecting landmines filled with AN, because the detection could not be achieved in a reasonable time without adding the echoes to compensate for a low SNR.

Currently, comparing the intensity of the NQR signal spectrum with a threshold seems to be the most employed strategy for landmine detection by NQR [13–17]; however, this strategy will not perform very well if the NQR signal is low, due to a small and/or deep landmine, to a high soil conductivity (affecting coil resonance frequency), or to changes on temperature of the substance. On those cases, the threshold should be low enough to detect the small NQR signals, opening the opportunity to signals from non-filtered radiofrequency interference (RFI) to cause false alarms. This work aimed evaluating landmine detection by NQR when treated as a binary classification task, training a classifier with NQR data acquired when the target substance is present and when it is not. This way, more than one distinctive feature could count in the decision, instead of just the signal intensity, to improve the system sensitivity and specificity. Features could also add value to a fusion of NQR with another technology, e.g. GPR, at the feature level.

In this work, some distinctive features based on spectral descriptors were proposed. Data was collected in the countryside of Colombia, using the portable system of [5] with and without AN near the coil. Part of the data was used to train an ensemble of classifiers (assembly of 20 decision trees trained by bagging) and the rest, to test it and compare it with the traditional strategy (based only on intensity of signal spectrum). A statistical test, comparing the area under the ROC curves of each solution, showed that the proposed classifier outperforms the traditional one. A final blind experiment was conducted where five landmines of different size were buried in an area of 1.9 × 1.52 m, and it was possible to detect four of them with three false alarms. From our knowledge, this is the first work showing field results of NQR detection of landmines filled with AN.

2. Materials and methods

2.1. NQR device

The NQR device as presented in [5] was used in this work, which is capable of detecting AN remotely. The system can send pulses at the resonance frequency of AN (423.4 kHz), and can record the NQR signals coming from AN samples. It also has a Q switching system that allows lowering the Q factor of the coil after pulsing and increasing it again for signal acquisition. The gradiometer coil used with the system is shown in Fig. 1. This coil was designed and tuned to the resonance frequency of AN, and in conjunction with the NQR device, it was capable of remote excitation of AN and of picking up the NQR signal produced after the excitation. The gradiometer coil works fine with the Q switching circuit of the device.

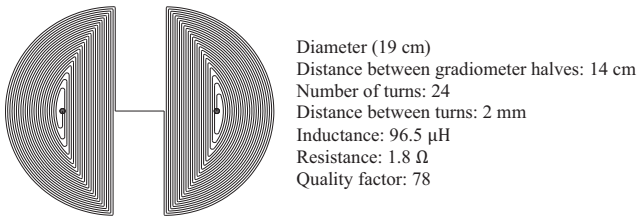

Fig. 1. Gradiometer coil. Black dots are connection ports to capacitors and the NQR system.

2.2. Ammonium nitrate and pulse sequence

AN has a melting point of 169 °C at atmospheric pressure, below which it can exist in five different crystallographic forms (phases I-V) [18]. The stable polymorph at room temperature (from −16 °C to 32 °C) is the fourth (known as AN-IV). Since there are two nitrogen atoms in the AN molecule, each distinct ^{14}N nucleus give rise to three NQR transition frequencies [18]. In this work, the ν_- resonance frequency at nitrate ion is used (423.4 kHz @ 25 °C) as it is higher than those for ammonium ion, and its temperature coefficient (+91 Hz/K) is markedly lower than that of ν_+ at nitrate ion (−300 Hz/K) [7].

As a multi-pulse sequence is needed to overcome the long longitudinal relaxation time T_1 (16 s) of AN, a steady-state free precession sequence (SSFP) was used. In SSFP, evenly spaced RF pulses are used to excite the sample, all of them having the same duration. After a few pulses, a new steady state is established where continuous signal of varying amplitude is formed, and coherent NQR signals generated after each pulse could be averaged to increase the SNR.

The parameters for the SSFP sequence were previously optimized for AN detection at the laboratory. Pulse width of 200 μs was estimated from the expected target depth and from the range of the NQR equipment, (we measured magnetic field intensity at different distances from the coil). A repetition time of 2.2 ms was selected, considering the dead time after pulsing (dictated by the design of the NQR device); the number of pulses (3000) which in turns determined sequence duration (6.6 s) was selected because it gave the higher SNR for a single pulse sequence. The parameters of the SSFP pulse sequence are summarized on Table 1.

2.3. Proposed features

The proposed features for discriminating NQR signals (from AN) were based on spectral descriptors, as they have been successfully employed in analysis of spectral data from spectroscopic techniques in chemistry [19]. They provided the benefit of incorporating spectroscopic knowledge about the substance. The proposed spectral descriptors are:

C_1: **spectrum magnitude at reference frequency**. The reference frequency is the AN resonance frequency (at the estimated temperature), and this feature is expected to be high if AN is present. Although the intensity at the signal peak is included in other descriptor, a peak could be absent if other signals at different frequencies are too wide, as shown in Fig. 2 (right).

Table 1
Parameters of the SSFP sequence for landmine detection.

Pulse width	200 μs
Repetition time	2.2 ms
Number of pulses	3000
Total time of the sequence	6.6 s

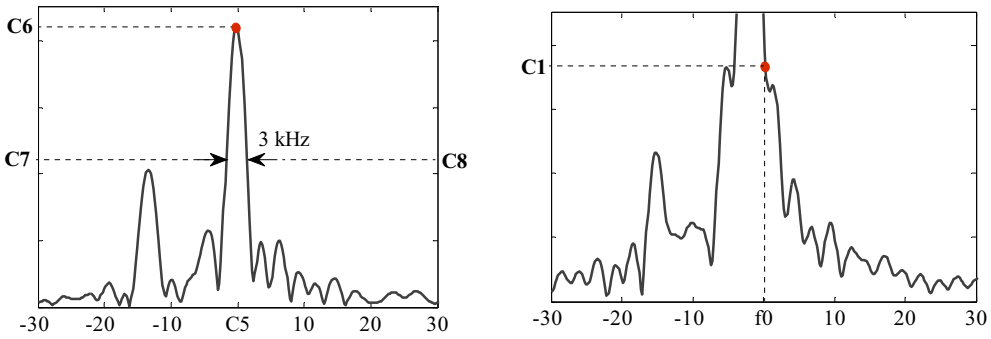

Fig. 2. A high NQR signal shape of AN in the frequency domain is shown on the left, showing the proposed spectral descriptors $\boldsymbol{C}_5$–$\boldsymbol{C}_8$. On the right, an illustration of another signal that conceals the shape of the NQR signal, justifying the descriptor $\boldsymbol{C}_1$.

C_2: **sum of spectrum intensities between 422.5 and 423.8 kHz**. This is the range where AN signal is expected.

C_3: **sum of spectrum intensities in the range from 420 to 426.3 kHz**. This feature is related to the previous one, but 2.5 kHz are added to the right and to the left of the expected range to account for the whole shape of the signal and not just for the peak.

C_4: **frequency of the highest peak in the range from 400 to 450 kHz.** A peak is defined as a point where intensity value is greater than the immediate left and right frequency neighbors. It would be likely, in many acquired signals from AN, that this feature will be equal to the expected resonance frequency of AN, or very close to it (although sources of strong RFI can make this to be different). In the case where signal comes just from noise and RFI, this feature is expected to have higher variability.

C_5: **frequency of the peak closer to the reference frequency**. There could be changes in the peak of the spectrum due to differences in temperature estimation, however, in the case of a NQR signal, this feature should range from 422.5 to 423.8 kHz

C_6: **spectrum intensity at $\boldsymbol{C}_5$ frequency**. This is used to compute the next two features, to account for the shape of the NQR signal (see Fig. 2, left).

C_7: **spectrum intensity 1.5 kHz to the left of frequency $\boldsymbol{C}_5$, divided by $\boldsymbol{C}_6$**. As stated above, this descriptor is used to account for the shape of the signal (Fig. 2).

C_8: spectrum intensity 1.5 kHz to the right of frequency $\boldsymbol{C}_5$, divided by $\boldsymbol{C}_6$.

The next three descriptors are given to relate the descriptors based on signal intensity, to the average intensity at other frequencies of the spectrum.

$$C_9: \quad \boldsymbol{C}_1/\boldsymbol{S}_{ref}$$

$$C_{10}: \quad \boldsymbol{C}_2/\boldsymbol{S}_{ref}$$

$$C_{11}: \quad \boldsymbol{C}_3/\boldsymbol{S}_{ref}$$

where S_{ref} is the average of spectrum intensity on the ranges from 400 to 420 kHz, and from 426.5 to 450 kHz.

As the proposed features give redundant information, principal component analysis (PCA) was carried out to derive fewer independent common features.

2.4. Data collection and feature extraction

It has been suggested in [20] that meaningful qualitative conclusions can be drawn from classification experiments performed with a total of about 100 observations; requiring a minimum of 50 cases in each class for one case to represent more than 2% of the observations. Thus, a dataset consisting of 200 data samples was collected in a rural area in Girardota, a municipality of Colombia. On the place, five samples of ammonium nitrate on varied enclosures were shallowly buried (0 cm from soil surface) and excited with the SSFP sequence, using the parameters on Table 1. For each sample, 10 data were collected, varying the position of the coil as shown in Fig. 3. The characteristics of the enclosures are presented in Table 2. From Table 2, only targets M1 to M5 were used (target M6 was used in the field test).

For each data acquired with AN, another data was collected without sample on different locations, with the same pulsing parameters. In all the cases, the coil was on the ground.

The data was processed in Matlab for feature extraction. FFT was applied to the time domain data, and the spectrum data on the range between 400 and 450 kHz was taken. This spectrum range consisted of 240 data points, as the FFT was made with 4096 points and the sampling frequency was 846 kHz. A program was coded in Matlab to extract the features from each sample, and they were stored in a Matrix having 200 rows (one for each data sample) and 12 columns (one for each feature plus another one marking with 1 the samples that had AN signal, and with 0 the samples that did not have AN).

2.5. Proposed classifier

The task of detecting landmines was considered as a binary classification problem. The class of data samples that have NQR signal from AN was named "mine" class, and the class of the samples without signal from AN was named "other" class. The proposed classifier on this work was an ensemble of decision trees. Decision trees are systems that split the feature space into regions,

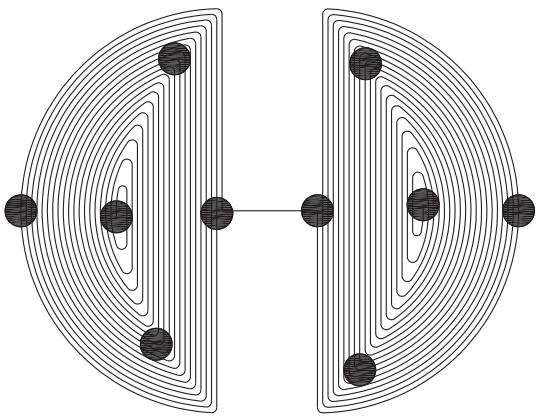

Fig. 3. Approximated positions of the center of the sample when the coil was placed over it when acquiring data.

Table 2
Information of each of ammonium nitrate samples.

Label	Amount of AN (g)	Diameter (cm)	Height (cm)	Enclosure
M1	300	9.5	7.0	Plastic
M2	55	4.5	5.0	Plastic
M3	214	7.0	9.0	Glass
M4	170	6.4	8.5	Plastic
M5	190	6.0	11.0	Plastic
M6	200	9.5	4.5	Plastic

corresponding to the classes, in a sequential manner. Upon a data sample, the searching of the region to which the sample will be assigned is achieved via a sequence of decisions along a path of nodes of an appropriately constructed tree. An ensemble of classifiers is a set of classifiers whose individual decisions are combined (typically by weighted or unweighted voting) to classify new data samples. There are many strategies for assembling classifiers; a common one is Bagging, a method that trains each classifier on a random redistribution of the training set. Each classifier (on this case, each decision tree) may result in higher classification error; but when combined, these classifiers often produce lower error than that of the single classifiers. Assembling classifiers is a known strategy to seek for generalization (the ability of a classifier to correctly classify new data).

For training the ensemble of decision trees, the 200 samples on the data set were split in two groups, a training data set (T_{tr}) and a test data set (T_{ts}) each having 100 samples with 50% data belonging to the "mine" class and 50% data belonging to the "other" class. The T_{ts} set was not used in any moment for training the classifier. The proposed classifier was trained and tested in Matlab, using the Statistics and Machine Learning toolbox. The training process was repeated using the data after applying PCA, for comparison. An ensemble of decision trees was used as it gave better performance in a previous test where it was compared to other classifiers (kNN, Support vector machine, neural network and naïve Bayes).

2.6. Classifier performance

A common measure of a classifier performance is given by its Receiver Operating Characteristic (ROC) curve, a two-dimensional plot that represents true positive rate vs false positive rate of the classifier.

To construct a ROC curve for a classifier, it is assumed that the output of the classifier (giving a data sample) is a numeric value $x$that can be compared with a threshold (u) to assign a class to the given sample. It is also assumed that higher values of the output are in favor of the positive class, while the lower ones are in favor of the negative class. There are four possible outputs for the classification. If the sample is of the positive class and it is classified as such, then it is denoted as a "true positive". If the classifier makes a mistake and classifies the sample as negative, it is called a "false negative". Similarly, if the sample is from the negative class and it is classified as such, it is a true negative, otherwise, it is a "false positive". With the total amount of true positives and false negatives, the true detection rate (TDR) is computed as the total number of "true positives", divided by the total number of samples from the "positive" group; and the false alarm rate (FAR) is computed as the total number of "false positives" divided by the total samples from the negative group. This pair of FAR (on the horizontal axis) and TDR (on the vertical axis) is a point in the ROC curve. By assessing the value of FAR and TDR for different values of u, the ROC curve can be completed.

On a ROC curve, a diagonal line from the point 0.0 to 1.1 describes what would be a classifier that is incapable of discriminating between the two classes because, for any threshold, the classifier gives the same proportion of true positives and false positives. A classifier will have a greater discriminative capacity to the extent that its ROC curve is as far as possible to the line of non-discrimination; in other words, as close as possible to the left and top sides of the graph.

Although the performance of a classifier can be assessed by observing the ROC curve, a good measure in just one number is the area under the curve [21]. This parameter, for this specific problem, corresponds to the probability of correctly identifying whether a sample belongs to the "mine" or to the "other" class.

After training, the ROC curve of the ensemble of decision trees was created using only the test data set, which consisted of 100 observations. Then, the area under the ROC curve (AUC) was computed.

Next step was getting the ROC curve of the solution based on comparing the spectrum intensity with a predefined threshold. However, the spectrum variable that is used for comparison is not clearly specified in the literature, it is called just "signal intensity" in [13–15], and "sum over the measurement interval" in [16]. Due to this ambiguity, the ROC curves of the decisions based on the featuresC_1,C_2, C_3 and C_6(related to signal intensity) were plotted using Matlab and the one showing the best performance was chosen.

2.7. Comparison of the proposed classifier to the common classification approach

A comparison between two ROC curves is usually made by assessing the statistical significance of the difference between two AUCs, using az-test, as described in [22]. Thus, the null hypothesis of no difference in the $AUC$of both classifiers was tested:

$$H_0: \ AUC_{ens} - AUC_{int} = 0$$

$$H_1: \ AUC_{ens} - A_{int} > 0$$

where AUC_{ens} and AUC_{int}refer to the estimated AUC associated with the ensemble of decision trees and with the classifier based on signal intensity, respectively.

The approach for assessing whether the difference in the areas under two ROC curves derived from the same set of observations is random or real consisted on computing the standard error of the difference of the two areas, using Eq. (1) [22].

$$SE(AUC_{ens} - AUC_{int}) = \sqrt{SE_{ens}^2 + SE_{int}^2 - 2rSE_{ens}SE_{int}} \quad (1)$$

Where SE_{ens} and SE_{int} refer to the standard error of AUC_{ens} and AUC_{int}, respectively; and r represents the estimated correlation between AUC_{ens} and AUC_{int}.

To compute r, two intermediate correlation coefficients were required, which were converted into a correlation between AUC_{ens} and AUC_{int} via a table supplied in [22]. The first coefficient was the correlation coefficient for the ratings given to samples from the "other" class by the two classifiers. The second coefficient was the correlation coefficient for the ratings given to samples

from the "mine" class by the two classifiers. Each of them was calculated using the Pearson product-moment [22].

To test the null hypothesis, the z statistics was employed [22]:

$$z = \frac{AUC_{bag} - AUC_{int}}{SE(AUC_{bag} - AUC_{int})} \quad (2)$$

An α level test would reject H_0 when z exceeds 1.645, the $1 - \alpha$quantile of the standard normal distribution.

2.8. Threshold selection

A threshold u is needed completely define the classifier. A common approach to choose the threshold is to employ the Youden index Y, which is the maximum vertical distance between the ROC curve and the diagonal line, and the threshold will be the one giving that point on the curve. The main aim of Youden index is to maximize the difference between rate of true positives and the rate of false alarms. Youden index is a commonly used criterion because it reflects the purpose of maximizing the correct classification rate. However, the threshold value corresponding with the Youden index is the optimal criterion value only when equal weight is given to sensitivity and specificity, and costs of various decisions are ignored. In landmine detection, mostly if it is made by hand, the cost of missing a target is higher than the cost of having a false alarm. In that sense, to stablish a threshold for the classifier, the Youden index and the corresponding threshold were computed by bootstrap [23], with 1000 iterations, and from a 95% confidence interval, the threshold giving the higher TDR was selected.

2.9. Field test of the NQR system

To make a final test of the system and of the proposed classifier, five targets filled with pure ammonium nitrate were shallowly buried in an area of 1.9 m in length by 1.52 m wide (2.9 m^2). This was done in a farm located in Santa Helena, a municipality of Colombia. First, a rectangular area was surrounded with a strip. Then, a third person shallowly buried (0 cm depth) the targets inside the fenced area and covered the area with a plastic tarp to hide any soil disturbance. The area was overlaid with a rectangular grid having 8 columns (named with letters A to H) and 10 rows (named with numbers 1–10). Each cell on the grid had an area of 19x19 cm and was referred by the corresponding column's letter and row's number of its position. The grid was needed to scan the area completely and in a regular manner, and to help with the construction of an "alarm map" in post processing. The characteristics of the buried targets were summarized in table 2 (all targets in the table except M5).

To have a high coverage while scanning the area, sequence illustrated on Fig. 4 was followed for each of the 80 cells. Fig. 4 shows the Cells A1 and B1 when cell A1 is being scanned. First, left side of the coil is "aligned" with left of the cell under scan, and then, with the right. The two data were stored in separated files, named as A1-1 and A1-2, as shown in Fig. 4. As two data is taken from each cell, a total of 160 measurements were taken. The field was scanned row by row (cells in row 2 were scanned only after all cells in row 1 were completed, and so on). Before starting the experiment, the tuning of the coil was checked.

After completing the experiment, each data file was processed with the trained classifier. By the position of the coil on the soil, diagrams were constructed, showing the sequence of coil positions on the cells having the targets. For assessing the true detection rate and false alarm rate, true classes of each of the 160 acquisitions were needed. Thus, acquisitions where coil surface was over the target (or over a great part of it) were tagged to the "mine" class. The rest of the measurements were assigned to the "other" class.

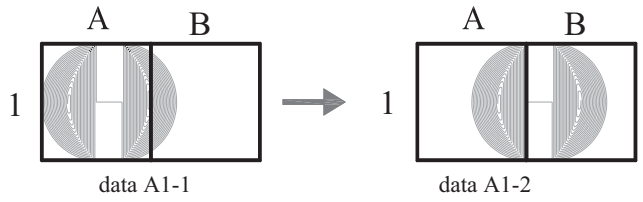

Fig. 4. The cells A1 and B1 are shown while cell A1 is being scanned. The nomenclature of the data file is shown on the bottom.

3. Results

3.1. ROC curves of the proposed classifier and of the common approach

From the results, the classifier trained with all the features of the data set (instead of the features after PCA) was chosen. Fig. 5 shows the ROC curve of the classifier. The selected feature for detection based on intensity of signal spectrum was the sum of intensities in the range between 422.5 and 423.8 kHz (C_2). Its ROC curve is also shown in Fig. 5.

3.2. Comparison of the classifiers

The difference between the two *AUC* and the results of the hypothesis test are shown in table 3. As z is greater than 1.645, the null hypothesis is rejected, concluding that the evidence support the belief that the difference between the two *AUC* is not zero, and thus, that the area under the ROC curve of the proposed ensemble of decision trees is greater than the area under the ROC curve of the classical solution. As the significance level P is less than 0.05 (the α level), the test is statistically significant, and the hypothesis H_0 can be rejected to a significance level of 0.0031.

3.3. Threshold selection

The true detection rate and false alarm rate for the threshold associated to the Youden index of the proposed classifier are presented on table 4, as well as the confidence interval for both the index and the threshold. The selected threshold was 0.286, the lower limit of the 95% confidence interval. The point on the ROC curve corresponding to this threshold is shown in Fig. 5; this threshold resulted in 100% of true detection rate and 3.33% of false alarm rate.

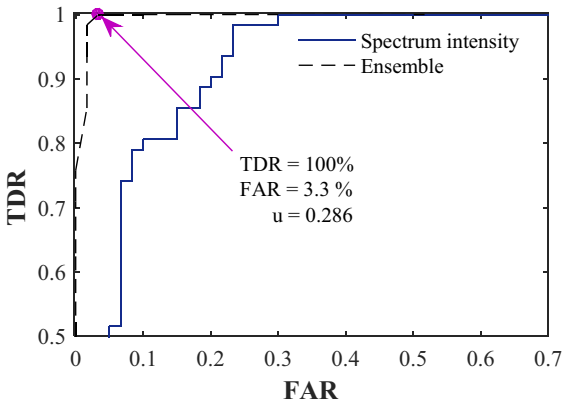

Fig. 5. In the same plot, the ROC curves of the ensemble of decision trees, and of the classical solution based in the sum of spectrum intensities between 422.5 and 423.8 kHz, obtained with the test dataset (100 observations).

Table 3
Results of the statistical analysis for the difference of the areas under the ROC curves of the ensemble of decision trees and the solution based on the sum of spectrum intensities.

$AUC_{bagging} - AUC_{intensity}$	0.0827
Standard Error	0.0301
95% confidence interval	0.0236–0.142
z statistic	2.742
Significance level	$P = 0.0031$

Table 4
Youden index and related threshold for the ensemble of decision trees.

Youden index Y	0.9672
95% confidence interval for Y	0.9003–1
Threshold U	0.333
95% confidence interval for U	0.286–0.333
True detection rate	98.39%
False alarm rate	1.67%

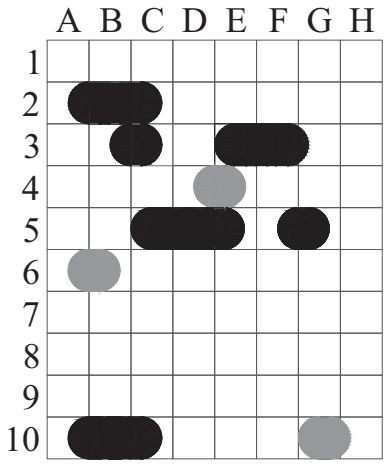

Fig. 6. In black, areas covered by the coil corresponding to the data that the classifier indexed on the "mine" class. Areas in grey correspond to three measurements that were repeated on the field for showing very high intensity peaks, and that were discarded by the classifier from the second measurement.

3.4. Results from field test

When tuning the coil, capacitance had to be reduced to move the spectrum frequency about 3 kHz, suggesting the presence of iron oxides on the soil (also suspected from the reddish color of the soil), and the measured Q factor was 74, showing a small reduction (from 78) due to soil conductivity.

At the end of the experiment, spectra of the acquisitions were reviewed in LabVIEW, and four measurements having a very high intensity peak were repeated a second time. On three of those cases, the peak disappeared on the second measurement. The second sequence employed to re-scan those places was not a phase-alternated pulse sequence, but a SSFP sequence; thus, it was probable that the intensity peaks seen on the first measurement but not on the second one were false alarms caused by RFI. The fourth case could have come from piezoelectric ringing. After completing the experiment, each data file was processed with the trained classifier, using the selected threshold ($u = 0.286$). The map in Fig. 6 shows the positions of the coil corresponding to detections made by the classifier, considering that the detection could have come from any of the two halves of the coil. Shadows in grey correspond to three of the measurements that were repeated a second time. On those cases, both signals were passed to the classifier algorithm and it gave an alarm for the first acquired data but not for the second one.

Fig. 7 shows the true position of the targets. From comparison with Fig. 6, target M2 was not detected and looks like there was one false alarm (or four, if the repeated measurements added up).

To estimate the true detection rate and false alarm rate of the classifier, true classes were needed for the 160 data acquired on the field. For assigning the true classes, diagrams on Fig. 8 were made, showing the sequence of coil positions on all the cells having the targets. The name of the file for each coil position is also shown in the figure. The true class for the data corresponding to file names marked with asterisk in Fig. 8 was decided to be "mine", as on those cases the coil surface was overlying a great part of the target. The true class for the rest of the measurements (measurements not marked with asterisk in Fig. 8 and measurements not shown in Fig. 8) was decided to be "other". As a result, of the 160 data samples, 14 were on the "mine" class, and 146 on the "other" class.

Having the true classes of the data, the ROC curve in Fig. 9 was plotted, considering as false alarms the three repeated measurements that were discarded by the classifier the second time. The highlighted point on the ROC curve corresponds to the previously

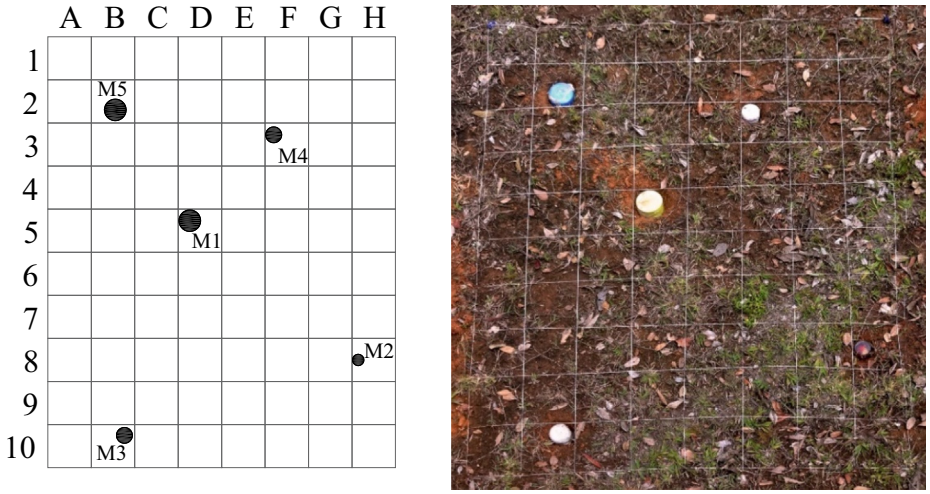

Fig. 7. Diagram (left) and picture (right) showing the true position of the targets.

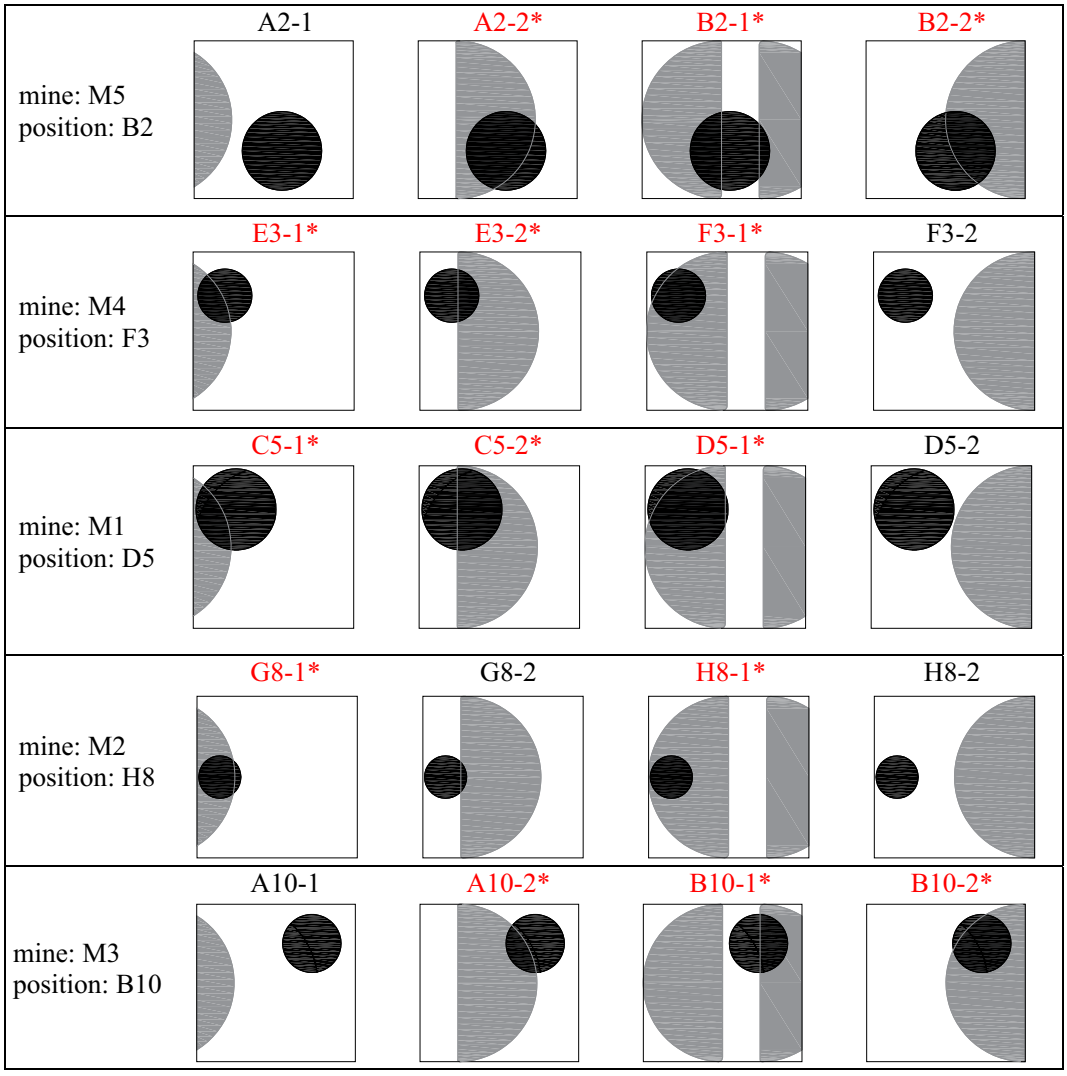

Fig. 8. The cells having a target are shown while the coil is passing through. The gray areas correspond to the areas under the coil that are inside the cell, and the black areas correspond to the targets. The name of the scan is also displayed on the top.

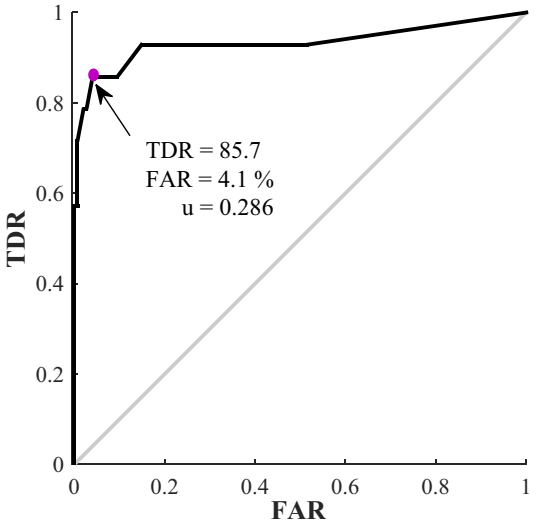

Fig. 9. ROC curve of the ensemble of decision trees for the 160 data files. The marked point corresponds to the previously selected threshold (0.286).

selected threshold (2.86). For this threshold, the false alarm rate was 4.1% and the true detection rate was 85.7%. This FAR was not very different from the one in Fig. 5; however, the TDR for this test was noticeably lower. This was mainly due to the differences in sample size. On this field test, only 14 data samples of 160 (8.75%) belonged to the positive class, thus, each positive sample represented 7.14% of the class and each target represented about 14.28% or more (since a single target caused more than one positive data sample). Thus, the non-detected target on this experiment (the smallest one) caused a decrease of 14.28% on the true detection rate. In contrast, on the previous test, each sample represented only 2% of the class.

On each position, measuring time was 8 s. For the 164 acquired data (160 data samples plus 4 repeated measurements), the total measurement time was 1312 s or 21.9 min; thus the measurement time per unit area was 21.9 min/2.9 m^2 = 7.5 min/m^2.

4. Conclusions

A set of spectral descriptors, encoding information of the NQR frequency domain signal, were used for training a classifier in order to detect NQR signals from AN. An ensemble of decision trees was

trained and tested, giving a true detection rate of 100% and false alarm rate of 3.33%. It was also shown that the ensemble of decision trees outperformed the classical approach, based solely on the intensity of the signal spectrum, in terms of the area under the ROC curve.

A final more realistic experiment where five AN targets were hidden in a small area, at 0 cm depth from the surface, showed that the system could detect four of the five targets, missing the smaller one, and having four false alarms. Three of the four false alarms were discarded after repeating the measurement a second time, indicating that they proceed from a non-steady source of RFI. The reason for missing the target that was already detected in previous experiments used for training the proposed classifier could be the soil conductivity or an error in estimating the temperature of the target; factors that are detrimental for NQR signal detection, and that probably were devastating for the smallest mine. From a constructed ROC curve, considering the 160 data samples of the experiment, a true detection rate of 85.7% was achieved with 4.1% of false alarms.

It was noticed that false alarms, coming from RFI can produce high intensity peaks at the frequency of ammonium nitrate, or at a neighbor frequency but with a spectral shape that could hide the NQR signal coming from a target. It is recommended to use an algorithm that first checks for the presence of very strong signals around the AN resonance frequency, to indicate the need for a second measurement. A second measurement may solve many of these alarms, before presenting the observation to the classifier.

Although the spectral descriptors proposed on this paper as well as the classification methodology employed (an ensemble of decision trees) can be used with other NQR equipment and for detecting other substances; in those cases, the classifiers must be trained with new data, since the values of the features depend both on the substance being detected and on the equipment. By training the classifier, it will adapt decision criteria to the new data and, from the results of this work, better performance would be expected when using this approach than when using the intensity of the signal spectrum as the only parameter for decision, which has been conventionally used in the literature of landmine detection by NQR. There are other areas of magnetic resonance that can benefit from the use of spectral descriptors and ensemble of decision trees, mainly on those applications where a diagnosis or a detection is required, as in the detection of subsurface water; detection of narcotics; classification of brain tumors from MR spectra; and in metabonomics, for examining biofluids by NMR spectroscopy to detect biochemical changes. Spectral descriptors can be applied to describe different metabolites to deal with the NMR spectrum complexity and variations across samples.

Results showed that the NQR technology has good potential for detecting landmines in Colombia; however, the system is not ready for real landmine detection tasks since its detection range is still low. There are still some improvements that can be made to increase the system sensitivity. Future work would focus on improving sensitivity of the system, working on coil and Q-switching design.

The proposed spectral descriptors opens the door for a feature-level fusion of the NQR detection system with another technology. Future research could focus on the fusion, at feature level, of the NQR technology with ground penetrating radar (GPR), as GPR is a mature technology for detecting nonmetallic mines.

Acknowledgments

This work was mainly supported by a grant from the Japan Society for the Promotion of Science and by the Administrative Department of Science, Technology and Innovation of Colombia (COLCIENCIAS).

References

[1] A. Duttine, E. Hottentot, Landmines and explosive remnants of war: a health threat not to be ignored 160–160A Bull. World Health Organ. 91 (2013), https://doi.org/10.2471/BLT.13.118885.
[2] A. EFE, Colombia, segundo país con mayor número de víctimas de minas, Periódico El País. (2011) Seccion Judicial. http://www.elpais.com.co/elpais/judicial/noticias/colombia-segundo-pais-con-mayor-numero-victimas-minas.
[3] P.E. Parra-Gallego, IEDs : A Major T hreat for a Struggling Society, J. ERW Mine Action. 13 (2009) 1–5. http://maic.jmu.edu/journal/13.3/specialreport/gallego/gallego.htm.
[4] Y.A. Pino, Determinación de técnicas de detección de explosivos óptimas para el departamento de antioquia, Universidad Nacional de Colombia, sede Medellín, 2009.
[5] L. Cardona, Y. Miyato, H. Itozaki, J. Jimenez, N. Vanegas, H. Sato-Akaba, Remote detection of ammonium nitrate by nuclear quadrupole resonance using a portable system, Appl. Magn. Reson. 46 (2015) 295–307, https://doi.org/10.1007/s00723-014-0623-6.
[6] Z. Bielecki, J. Janucki, A. Kawalec, J. Mikolajczyk, N. Palka, M. Pasternak, T. Pustelny, T. Stacewicz, J. Wojtas, Sensors and systems for the detection of explosive devices – An overview, Metrol. Meas. Syst. XIX (2012) 3–28, https://doi.org/10.2478/v10178-012-0001-3.
[7] J. Barras, M.J. Gaskell, N. Hunt, R.I. Jenkinson, K.R. Mann, D.A.G. Pedder, G.N. Shilstone, J.A.S. Smith, Detection of ammonium nitrate inside vehicles by nuclear quadrupole resonance, Appl. Magn. Reson. 25 (2004) 411–437, https://doi.org/10.1007/BF03166538.
[8] S.D. Somasundaram, A. Jakobsson, J.A.S. Smith, Frequency selective detection of NQR signals in the presence of multiple polymorphic forms, in: 14th Eur. Signal Process. Conf., Florence, Italy, 2006.
[9] S.D. Somasundaram, A. Jakobsson, J.A.S. Smith, K. Althoefer, Exploiting spin echo decay in the detection of nuclear quadrupole resonance signals, IEEE Trans. Geosci. Remote Sens. 45 (2007) 925–933, https://doi.org/10.1109/TGRS.2006.890413.
[10] A. Jakobsson, M. Mossberg, M.D. Rowe, J.A.S. Smith, Exploiting temperature dependency in the detection of NQR signals, IEEE Trans. Signal Process. 54 (2006) 1610–1616, https://doi.org/10.1109/TSP.2006.871969.
[11] S.D. Somasundaram, A. Jakobsson, N.R. Butt, Countering radio frequency interference in single-sensor quadrupole resonance, IEEE Geosci. Remote Sens. Lett. 6 (2009) 62–66, https://doi.org/10.1109/LGRS.2008.2005792.
[12] N.R. Butt, A. Jakobsson, S.D. Somasundaram, J.A.S. Smith, Robust multichannel detection of mixtures using nuclear quadrupole resonance, IEEE Trans. Signal Process. 56 (2008) 5042–5050, https://doi.org/10.1109/TSP.2008.928704.
[13] H. Itozaki, Nuclear quadrupole resonance for explosive detection, in: K. Furuta, J. Ishikawa (Eds.), Anti-Personnel Landmine Detect. Humanit. Demining Curr. Situat. Futur. Dir. Japanese Res. Dev., Springer - Verlag, London, 2009, pp. 148–155, https://doi.org/10.1007/978-1-84882-346-4_9.
[14] A.N. Garroway, M.L. Buess, J.B. Miller, B.H. Suits, A.D. Hibbs, G.A. Barrall, R. Matthews, L.J. Burnett, Remote sensing by nuclear quadrupole resonance, IEEE Trans. Geosci. Remote Sens. 39 (2001) 1108–1118, https://doi.org/10.1109/36.927420.
[15] A.D. Hibbs, G.A. Barrall, P.V. Czipott, A.J. Drew, D.K. Lathrop, Y.K. Lee, E.E. Magnuson, R. Matthews, D.C. Skvoretz, S.A. Vierkotter, Man portable mine detector using nuclear quadrupole resonance - first year progress and test results, in: Second Int. Conf. Detect. Abandon. L. Mines, Edinburgh, UK, 1998, pp. 138–141, https://doi.org/10.1049/cp:19980706.
[16] A.D. Hibbs, In-situ real time detection of explosive-chemical compounds in mines using Nuclear Quadrupole Resonance (NQR), San Diego, CA, 2001.
[17] M. Ostafin, B. Nogaj, 14N-NQR based device for detection of explosives in landmines, Measurement 40 (2007) 43–54, https://doi.org/10.1016/j.measurement.2006.04.003.
[18] J. Seliger, V. Zagar, R. Blinc, 14N NQR study of the structural phase transitions in NH4NO3, Zeitschrift Für Phys. B Condens. Matter. 77 (1989) 439–443, https://doi.org/10.1007/BF01453795.
[19] H. Lohninger, J. Ofner, Multisensor hyperspectral imaging as a versatile tool for image-based chemical structure determination, Spectrosc. Eur. 26 (2014) 6–10.
[20] C.E. Metz, Basic principles of ROC analysis, Semin. Nucl. Med. 8 (1978) 283–298, https://doi.org/10.1016/S0001-2998(78)80014-2.
[21] A.J. Hanley, J.B. McNeil, The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve, Radiology 143 (1982) 29–36, https://doi.org/10.1148/radiology.143.1.7063747.
[22] J.A. Hanley, B.J. McNeil, A method of comparing the areas under receiver operating characteristic curves derived from the same cases, Radiology 148 (1983) 839–843, https://doi.org/10.1148/radiology.148.3.6878708.
[23] T.J. DiCiccio, B. Efron, Bootstrap confidence intervals, Stat. Sci. 11 (1996) 189–228, https://doi.org/10.1214/ss/1032280214.