Science of the Total Environment 521–522 (2015) 270–279

Contents lists available at ScienceDirect

Science of the Total Environment

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

Remediating radium contaminated legacy sites: Advances made through machine learning in routine monitoring of "hot" particles

Adam Varley [a,*], Andrew Tyler [a], Leslie Smith [b], Paul Dale [c], Mike Davies [d]

[a] Department of Biological and Environmental Sciences, University of Stirling, Stirling FK9 4LA, United Kingdom
[b] Department of Computing Science and Mathematics, University of Stirling, Stirling FK9 4LA, United Kingdom
[c] Scottish Environmental Protection Agency, Radioactive Substances, Strathallan House, Castle Business Park, Stirling FK9 4TZ, United Kingdom
[d] Nuvia Limited, The Library, Eight Street, Harwell Oxford, Didcot, Oxfordshire OX11 0RL, United Kingdom

HIGHLIGHTS

- Land contaminated with radium is hazardous to human health.
- Routine monitoring permits identification and removal of radioactive *hot particles*.
- Current alarm algorithms do not provide reliable *hot particle* detection.
- Spectral processing using Machine Learning significantly improves detection.

ARTICLE INFO

Article history:
Received 30 January 2015
Received in revised form 27 March 2015
Accepted 29 March 2015
Available online xxxx

Editor: Simon Pollard

Keywords:
Radium remediation
Gamma spectroscopy
"Hot" particles
Machine learning
Monte Carlo
Sodium iodide
Lanthanum bromide

ABSTRACT

The extensive use of radium during the 20th century for industrial, military and pharmaceutical purposes has led to a large number of contaminated legacy sites across Europe and North America. Sites that pose a high risk to the general public can present expensive and long-term remediation projects. Often the most pragmatic remediation approach is through routine monitoring operating gamma-ray detectors to identify, in real-time, the signal from the most hazardous heterogeneous contamination (*hot particles*); thus facilitating their removal and safe disposal. However, current detection systems do not fully utilise all spectral information resulting in low detection rates and ultimately an increased risk to the human health. The aim of this study was to establish an optimised detector-algorithm combination. To achieve this, field data was collected using two handheld detectors (sodium iodide and lanthanum bromide) and a number of Monte Carlo simulated *hot particles* were randomly injected into the field data. This allowed for the detection rate of conventional deterministic (gross counts) and machine learning (neural networks and support vector machines) algorithms to be assessed. The results demonstrated that a Neural Network operated on a sodium iodide detector provided the best detection capability. Compared to deterministic approaches, this optimised detection system could detect a *hot particle* on average 10 cm deeper into the soil column or with half of the activity at the same depth. It was also found that noise presented by internal contamination restricted lanthanum bromide for this application.

1. Introduction

1.1. Radium contamination

Radium (^{226}Ra) was used extensively during the 20th century predominantly in the form of luminescent paint. Waste generated from military, industrial and pharmaceutical products was regularly buried with little record of its location and inventory (Harvie, 1999). With a half-life of 1600 years, ^{226}Ra contamination is a multigenerational issue. In the UK, a recent government report conservatively estimated there to be 150 to 250 contaminated legacy sites, whilst acknowledging there could be as many as a 1000 (DECC, 2012). Similar extents of ^{226}Ra contamination have been found across Europe and North America (Harvie, 1999). Ultimately, the risk of human exposure at these sites is dependent on a number of potential pathways and the form of contamination and not exclusively on external dose (Dale et al., 2008).

One such pathway, that has the potential to cause significant radiological harm, is ingestion of small highly radioactive items often referred to as *hot particles* (Baker and Toque, 2005). One study explored the committed dose that could be received by a member of the public through simulated stomach acid digestions of a range of radium *hot particles*

* Corresponding author.
E-mail addresses: a.l.varley@stir.ac.uk (A. Varley), a.n.tyler@stir.ac.uk (A. Tyler), l.s.smith@cs.stir.ac.uk (L. Smith), paul.dale@sepa.org.uk (P. Dale), Mike.Davies@nuvia.co.uk (M. Davies).

http://dx.doi.org/10.1016/j.scitotenv.2015.03.131
0048-9697/

found in Scotland (Tyler et al., 2013). It was concluded that ingestion of a *hot particle* with an activity higher than 20 kBq could result in a committed dose to an infant exceeding the 100 mSv threshold deemed to cause significant radiological harm (ICRP, 2007). At particular sites in Scotland, it has been recognised by the Scottish Environment Protection Agency (SEPA) that there is the possibility of a member of the public coming into contact with such a *hot particle*. To safeguard against this at Dalgety Bay, Fife, Scotland, routine monitoring is undertaken to detect and retrieve any significant radioactive items (Dale et al., 2013). To confirm monitoring is undertaken with sufficient accuracy, SEPA have outlined the following criterion:

- A 20 kBq *hot particle* at a burial depth of 0.1 m must be detected 95% of the time.

This paper aims to develop an optimised detection system that provides better detection capability than systems currently available. This will allow more effective identification of *hot particles* at radium legacy sites, ultimately reducing the risk posed to the general public in both the short and long terms.

1.2. Challenges of "hot" particle detection

The most effective method of detecting *hot particles* in real-time is through a series of mobile measurements using either handheld or vehicle mounted gamma-ray sensors (Tyler, 2008). At many sites where vehicular access is limited handheld detectors can be the only option. Handheld detectors produce gamma-ray spectra, the shape and magnitude of which will provide information of the localised radiation field that the detector has passed through during acquisition. To ensure that an area is adequately covered by a survey in a reasonable time frame, typically a spectrum is acquired every second and a walking speed of 0.5 m s^{-1} is maintained. This maximises the spatial density of measurements and ultimately the probability of detecting any *hot particles*. To initiate the immediate identification of a *hot particle* real-time analysis of the spectral time series is critical (Kock et al., 2012).

The spectral response or signal quality of a detector will largely be governed by the composition of the detector's active volume. Lightweight scintillators tend to be used for handheld detectors, of these sodium iodide (NaI:Tl) is the standard as it is relatively cheap and robust (Knoll, 2010). Nevertheless alternatives are available. One such detector, lanthanum bromide (LaBr:Ce), has recently received much attention (Guss et al., 2010). It has better energy resolution (~2.5% at 662 keV) than the NaI:Tl (~7% at 662 keV) (Fig. 1), greater photon efficiencies and better temperature stability. However, a relatively large intrinsic background signal attributed to internal isotopes (^{138}La and ^{227}Ac) can be found distributed throughout the spectrum (Iltis et al., 2006). Notice considerable contamination contributions at 1468 keV (gamma and x-ray summation peak), 786–1100 keV (beta continuum) and over 1700 keV (alpha) (Menge et al., 2007). This raises concerns about its potential in low source signal situations as these are frequently encountered in the environment.

The second element of a detection system is the algorithm used to process the detector's signal in real-time to determine whether there are signal contributions from a *hot particle* (Fig. 1). If the algorithm is tripped, an alarm is sounded, allowing the operator to locate and retrieve any potential radioactive items (Jarman et al., 2008). Yet processing environmental gamma-ray spectra is not straightforward.

First of all the vast majority of spectral changes are benign and can occur over the scale of a few metres (Fagan et al., 2012). This spatial variation is brought about by changes in the natural radioelements (^{40} K, and the ^{238}U and ^{232}Th series) contained within the local geology. Changes can also occur temporally due to variations in the density and chemical composition of the geological matrix and radon exhalation (IAEA, 2003). Temporal fluctuations can be very challenging to account for particularly as they can occur on varying time scales, for example deviations in density over a tidal cycle and radon exhalation caused by atmospheric pressure changes (Ball et al., 1991; De Groot et al., 2009). Furthermore, weak source signals can appear very similar to background signals given that ^{226}Ra forms part of the ^{238}U series. Consequently, in an attempt to isolate source signal from benign signal, it is often appropriate to use the most recent observations in the time series as estimates of background (Ely et al., 2004).

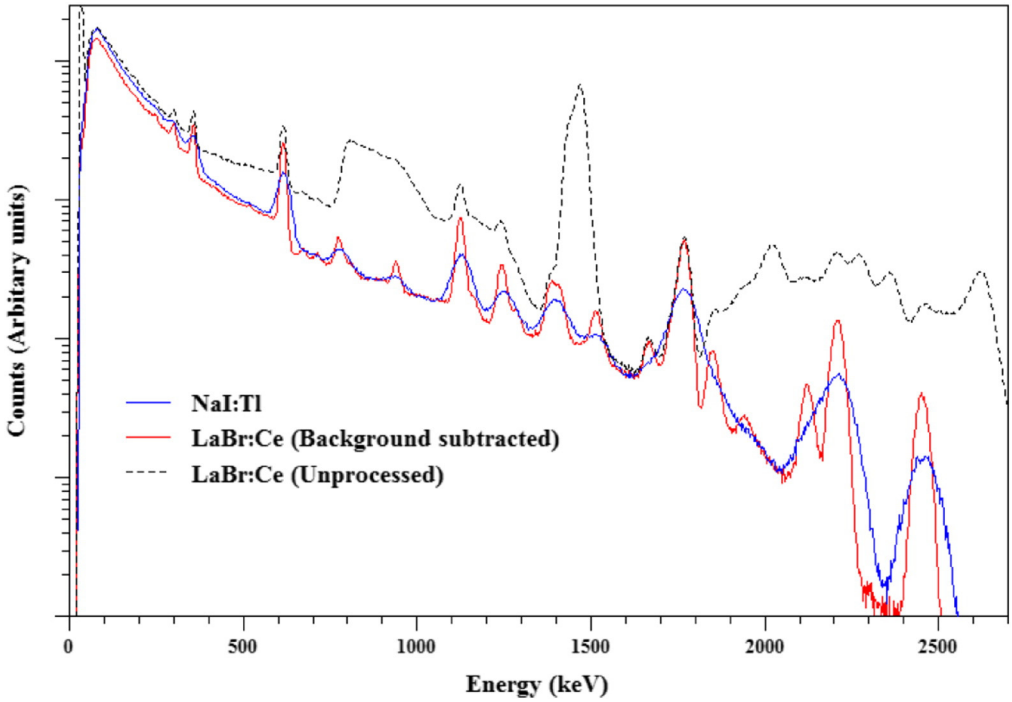

Fig. 1. Detector responses to a ^{226}Ra *hot particle*: NaI:Tl (blue), LaBr:Ce background subtracted (red) and LaBr:Ce unprocessed (broken black).

Another issue is that source signal will not remain constant. Firstly, the signal will decrease significantly with increasing distance between the detector and source. Secondly, non-linear changes will occur across the spectrum as the amount of shielding (depth in soil) is increased and often the majority of photons that reach the detector are scattered to low energy (Wei et al., 2010). These occurrences can easily lead to weak source signals, for instance a source at depth or a low activity surface source, not being identified.

Finally, the short acquisition times (typically 1 s) demanded by the high density of measurements during a survey, when combined with the low energy efficiency of handheld detectors, results in spectra containing a large stochastic noise element (Du et al., 2010). This renders conventional spectral analysis methods founded on determination of peak area unworkable (Alamaniotis et al., 2013).

These complications mean that source-background separation is never going to be seamless and a trade-off between detection rate and false alarm rate is encountered (Ely et al., 2006). To safeguard against the operator losing confidence in the detection system, a typical false alarm rate of 1 in 200 ($\alpha = 0.005$) should be attained (Kock et al., 2012). Subsequently to attain this false alarm rate, straightforward deterministic algorithms such as the Gross Counting algorithm are still heavily relied upon and can produce poor detection rates (Runkle, 2006).

1.3. Machine learning

Here we hypothesise whether *it is more appropriate to track general changes in spectral shape* by dividing raw spectra into more appropriate energy bins (Fagan et al., 2012). However, this procedure is complicated by the fact source and background populations may overlap considerably within the new high dimensional feature space (depending on the number of energy bins used) and possibly exhibit non-linear class boundaries. This scenario can be unfeasible to separate using conventional Newtonian mathematics (Smola et al., 1998). Instead, supervised Machine Leaning (ML) algorithms can be employed to map the underlying relationship between explanatory and response variables (Galushkin, 2007). For this reason, ML has been employed previously to develop alarm thresholds for gamma-ray data (Kangas et al., 2008; Sharma et al., 2012; Varley et al., 2015; Wei et al., 2010).

Customarily ML attempts to address a statistical problem by learning the underlying structure of a sample of data provided to it during a training phase. Once training is complete, the performance of the model fit can be assessed using independent cross-validation dataset (Dragović et al., 2006). There are a number of possible ML methods available (Ao et al., 2010). However, preliminary investigations demonstrated neural networks (NNs) and support vector machines (SVMs) to be the most encouraging for this application.

1.4. Neural networks

The structure of a NN is analogous of the brain in that it is made up of processing units called neurons connected by synaptic weights (Olmos et al., 1992). Neurons are separated into three individual layers: input, hidden and output. Spectral data can be fed into the input layer, where it is passed through weighted synapses to the hidden layer where a non-linear function is used to map the problem to the output layer. Global convergence of the problem is attempted by minimising the error between training outputs and actual outputs through an iterative procedure of updating the weights between neurons (Gurney, 2003). To avoid getting trapped in local minima during this process training algorithms such as "resilient backpropogation" are used (Riedmiller and Braun, 1993).

1.5. Support vector machines

SVMs approach the problem differently to NNs, instead making use of kernel functions to enlarge the feature space ensuring dot products of the support vectors can be easily computed (Smola et al., 1998). This allows the introduction of maximal-separating hyperplanes providing a means of separating complex populations without using vast quantities of computer memory (James et al., 2013). This technique is referred to as the "kernel trick". There are a number of commonly applied kernels to perform this task including: polynomial, sigmoid, radial basis function and spline (Sangeetha and Kalpana, 2010).

The aim of this study was to assess the performance of a number of different detector-algorithm combinations (henceforth referred to as detector configurations) by spiking background spectra with representative Monte Carlo source spectra providing a means of establishing detection rate and false alarm rate.

2. Materials and methods

2.1. Field site

In the early 1990s ^{226}Ra was discovered at Dalgety Bay, Fife, Scotland (−3.3505 °E, 56.0349 °N) (Fig. 2) attributed to actions once carried out at its historical airfield, notably during wartime periods (Patton et al., 2013). A housing estate, sailing club and public footpath now exist in close proximity to known contaminated areas and erosion events have redistributed large quantities of contaminated material onto the public beach. As a result of the dynamic nature of the beach, *hot particles* are regularly brought to the surface or relocated laterally (Dale et al., 2013). This has prompted initially intense and subsequently large-scale routine monitoring efforts to reduce the risk to the public. The beach however presents a challenging environment in which to monitor, since there are considerable variations in background and density gradients, alongside large sections that contain relatively benign homogeneously distributed contamination (Fig. 2). Tyler et al. (2013) present an interesting study into the physical and chemical formation and the resultant risks associated with a number of *hot particles* found at Dalgety Bay.

2.2. Background data acquisition

Background spectra can change considerably from one site to next and over time. Therefore, background data was acquired from Dalgety beach using a NaI:Tl and LaBr:Ce (both 71 × 71 mm) attached to a wheeled mounted frame, one behind the other, to ensure the same ground was being covered by each detector. The detectors were mounted at a height of 0.1 m. 1024 channel spectra were acquired every second using Ortec's Maestro software alongside GPS coordinates and integrated for real time mapping and assessment using software developed at Stirling University. A walking speed of 0.5 m s^{-1} and transect spacing of 0.5 m were maintained during the survey. A 1 second lag was introduced to ensure that comparable background spectra were used to test algorithms. A total of 35,000 "background" spectra were recorded for each detector over the course of 3 days.

2.3. Simulating the radiation field of a "hot" particle

Monte Carlo calibration spectra were preferred over analytical-derived calibration spectra given that laboratory calibration can introduce large systematic uncertainties (Hendriks et al., 2002). The software package Monte Carlo N-Particle 5 (MCNP5) was used to produce full-spectral responses (Briesmeister, 1993). To validate the modelling method, Monte Carlo spectral responses were compared to experimental spectra taken from concrete calibration pads (Minty et al., 1997). Above 150 keV, good agreement was found between calibration pad and Monte Carlo spectra, therefore only energies above 150 keV were used. The background dataset was known to exhibit a large range of shielding conditions brought about by changes in geological matrix composition and density. Attempting to model and then correlate this variation would have been unfeasible, therefore a standard geological

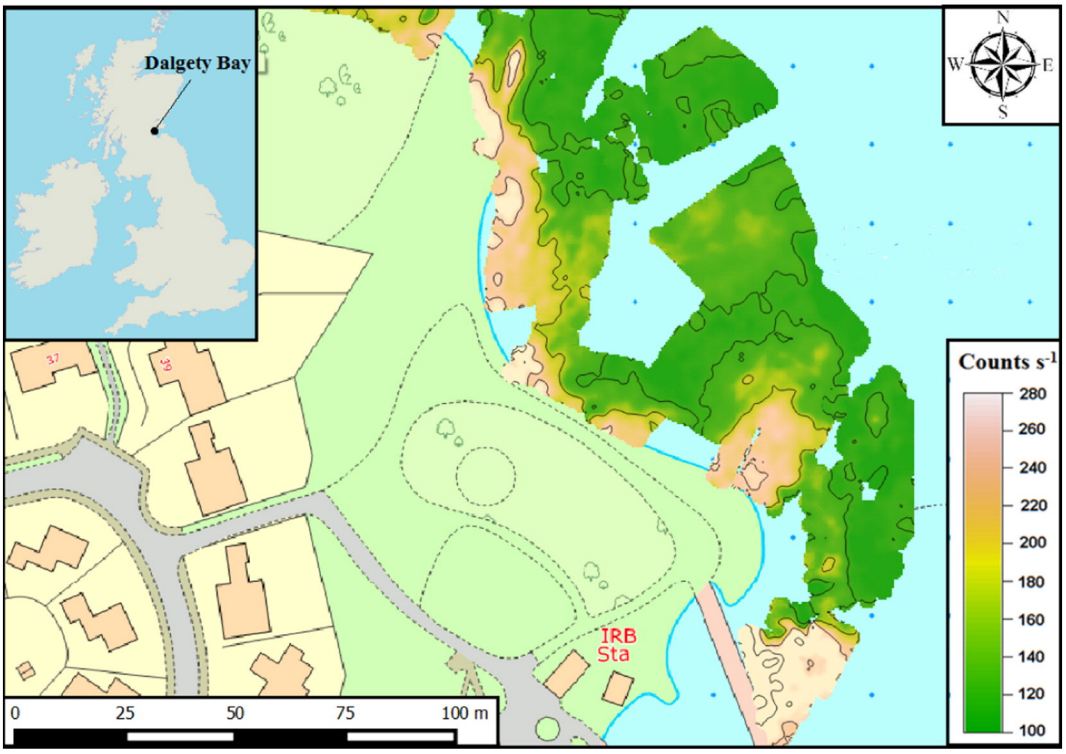

Fig. 2. Location of Dalgety Bay and a heat map demonstrating the variation in total count rate (counts s^{-1}) taken using a 71 × 71 mm NaI:Tl detector.

matrix (wet sand; 17% water) and density (1.5 g cm^{-3}) were used (Table 1).

To recreate realistic monitoring conditions a transect spacing of 0.5 m and walking speed of 0.5 m s^{-1} were assumed. In MCNP5 geometry, this involved integrating the detector's response 0.5 m through the radiation field of a point source. Table 1 describes the geometric range for the integration. A random sample was taken from the four dimensional array, possessing the dimensions x, y, depth and activity. To ensure an unbiased routine the random number seed in MCNP5 was changed between individual runs (Moreira et al., 2010). To reproduce the characteristic resolution of each detector statistical broadening was introduced after individual runs. Secular equilibrium was assumed and physical data was obtained from the National Nuclear Data Center (2013). Source spectra were then injected into the different detector "background" datasets at the same point to ensure consistency. A total of 20,000 spikes (with varying depth, activity and offset) were introduced and the spike rate was kept below 1% of the background dataset to prevent substantial overlap. To accurately assess SEPA mandate detection rate, a separate dataset generated spiked with 6000 20 kBq sources at 0.1 m burial depth.

Table 1
MCNP5 model parameter inputs.

Parameter	Inputs
Geological matrix	Wet sand 17% water
Soil density	1.5 g cm^{-3}
Detector height	0.1 m
[a]x offset (direction of travel)	0–0.25 m
[a]y offset (perpendicular to travel)	0–0.5 m
[a]Depth	0–0.9 m
[b]Activity	1–100,000 kBq

[a] 50 mm discrete increments were applied between geometric constraints.
[b] Systematic sampling used between activity constraints.

3. Pre-processing and algorithm execution

3.1. Gross Counting algorithm

Gross Counting (GC) algorithms are used extensively on handheld detectors, as they are relatively easy to setup (Ely et al., 2004; Jarman et al., 2008). GC is currently the only algorithm to be employed at Dalgety Bay (Dale et al., 2013). The method treats the entire spectrum as a single bin, using a rolling average filter of the previous few measurements to estimate the background (N) (Eq. (1)).

$$T = N + K\sqrt{N} \tag{1}$$

K (sigma multiplier) defines the number of standard deviations ($\sqrt{N}$) above N the alarm threshold should be set (Fig. 3). If a substantial increase in signal is received the alarm should sound as T is breached. The parameters K and the number of lags used on the moving average were optimised on part dataset.

3.2. Spectral comparison ratios

Dividing the raw spectrum into broad energy bins, instead of using raw data, alleviates some counting noise and reduces input dimensionality, although some loss of energy distinction is inevitable (Pfund et al., 2010) (Fig. 4). Selecting optimal energy bins is a contentious subject and novel algorithms have been implemented to optimise their number, placement and size (Wei et al., 2011). In this study though, two systematic binning systems were applied as input for ML. The first, Resolution Bins (RB), focused on the deterioration of resolution with increasing energy (Runkle, 2006) (Fig. 4). This produced 30 energy bins for LaBr:Ce and 18 for NaI:Tl. Whilst systematic, this approach cannot account for the fact that many of the bins, particularly at higher energy, may contain zero counts for a single measurement. A second method, Regions Of Interest Bins (ROIB), focussed around full energy peaks and scattering regions was implemented producing 13 energy bins for each detector.

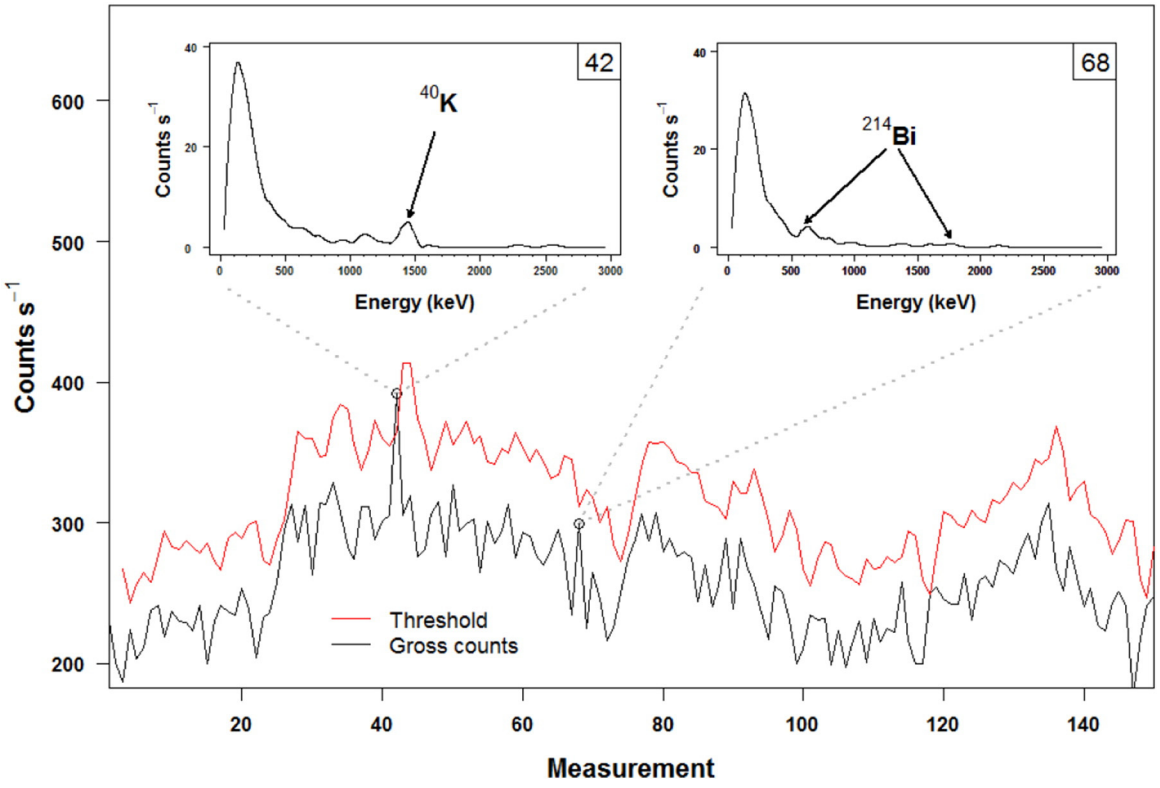

Fig. 3. Gross Counting obstacles: a ^{40}K dominated spectrum (42) triggers the alarm as the threshold is broken, but the ^{226}Ra spike possessing characteristic ^{214}Bi peaks (68) is missed as it does not reach the threshold.

Note, energy bins are narrower for LaBr:Ce given its superior energy resolution, although significantly more background counts were encountered in each bin caused by intrinsic contamination.

Spectral Comparison Ratios (SCR) outlined by Trost and Iwatschenko (2002), were used to transform energy bins into the time series to measure how closely the observed spectrum matched that of the previous background measurement (Eq. (2)).

$$SCR_i = N_1^C - \frac{N_1^B}{N_i^B} N_i^C \tag{2}$$

where N_i^c is bin i, and N_1^c is the first bin, of the current measured spectrum. N_i^B and N_1^B are a respective moving averages of bin i and the first bin of the previous spectra. This produces a transformed spectrum where the first channel is 0 and all channels contained within a background spectrum should be close to zero as small temporal variations are expected to occur (Du et al., 2010). Source contributions are anticipated to introduce larger deviations in spectral comparison ratios, though these are expected to occur across the spectral energy range, be non-linear in nature and contain a large counting noise element. Therefore, the effectiveness of conventional source-background separation approaches often relying on Gaussian distribution statistics tends to be too simplistic (Runkle, 2006).

3.3. Machine learning implementation

To train and assess the performance of detector configurations two datasets were formed: a training dataset and a cross-validation dataset. Typically 3000–5000 samples were used to train ML. For NNs, 30% of the training set was set aside as a test set to track the progress of training. The remaining data (~30,000) was used as the cross-validation dataset. Data was mean centred and scaled to the variance prior to training. Importantly, to produce the desired False Alarm Rate (FAR), approximately 5–7 times more background samples were included during training. It was discovered that by adding two more inputs (alongside spectral comparison ratios), one containing a moving average of gross counts and the other the total counts, an overall improvement in detection for both detectors could be realised. In our previous work this has also been found to be a benefit (Varley et al., 2015). This is not surprising given that areas of high background and homogeneous contamination will exhibit larger systematic variation, which ML could take into account by considering the total count rate and its localised variation. Another study employed a similar approach only alongside a deterministic gross alarm instead of providing it as an input for ML (Kangas et al., 2008).

3.4. Neural Network optimisation

To optimise the number of hidden neurons for NNs an extensive grid search was conducted (Medhat, 2012). As anticipated, NNs implemented on RB needed more hidden neurons compared to ROIB due to the greater number of inputs (Dragovic et al., 2005). That said 15–20 hidden neurons provided the lowest Relative Mean Squared Error (RMSE) for all NNs. RMSE on training and test sets for all NNs tended to converge after a few hundred learning epochs. The R package "RSNNS" was used (Bergmeir and Benítez, 2012).

3.5. Support Vector Machine optimisation

A radial basis function was found to provide the best degree of separation for SVMs in this application. Each radial basis function had two parameters that were optimised by an extensive grid search for each individual problem. The so-called cost parameter (C) providing the tolerance of the number of support vectors to include either side of the hyperplane presented no clear pattern between detector systems (600–2000) (Hornik et al., 2006). However, the gamma parameter specifying the width of the radial basis function varied between the

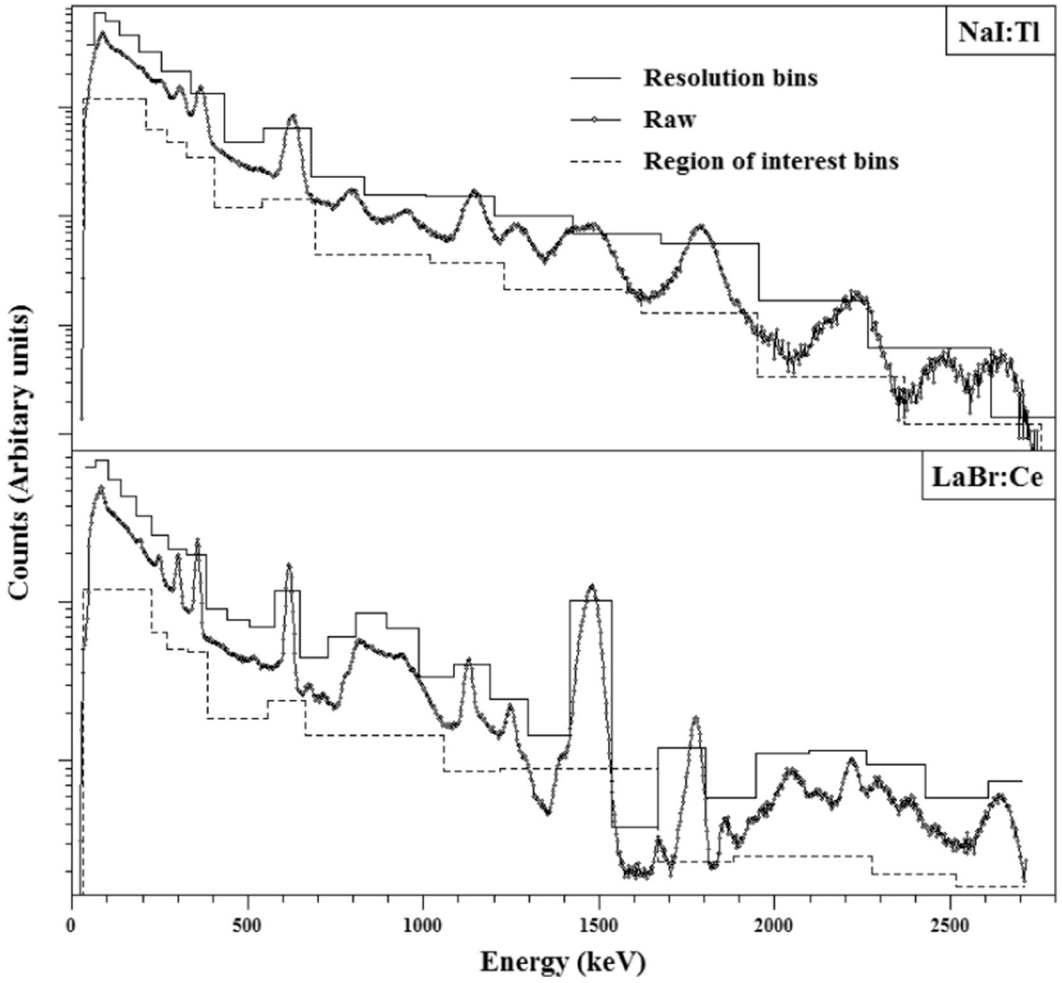

Fig. 4. Binning systems for NaI:Tl and LaBr:Ce detectors. Note the magnitude of bins has been altered for visual clarity.

binning systems, RB (<3) and ROIB (~300). The R package "e1071" was used (Dimitriadou et al., 2008).

3.6. Quality assessment

Quality assessment of individual detector configurations was based on the Overall Detection Rate (ODR) (Eq. (3)), SEPA's 20 kBq Mandate Detection Rate (MDR) (Eq. (4)) and the False Alarm Rate (FAR) (Eq. (5)) of an independent cross-validation dataset. Once ML architectures were obtained, 10 resamples (Section 3.4) were taken to obtain a mean and standard deviation to ensure that final values were not a chance representation.

$$\mathrm{ODR} = \frac{\text{Number of sources detected}}{\text{Total number of injected sources}} \tag{3}$$

$$\mathrm{MDR} = \frac{\text{Number of mandate sources detected}}{\text{Total number of injected mandate sources}} \tag{4}$$

$$\mathrm{FAR} = \frac{\text{Number of background alarms}}{\text{Total number of background}} \tag{5}$$

ODR is a general measure of detectability and does not provide information into the minimum detectable activity (MDA) as a function of depth. For example, certain detection systems might be better at identifying sources at the surface but less effectively at depth or vice versa. Subsequently, MDA ($\alpha = 0.95$) was fitted through maximum likelihood using a binary logistic regression (Crawley, 2012).

4. Results and discussion

4.1. Optimisation of Gross Count algorithm

Initially, K values and the number of lags used in the moving average were optimised to generate a GC baseline to compare ML to. Exponential moving average formulations, $2/(n+1)$ and $(1/n)$ for LaBr:Ce and NaI:Tl respectively, were found to yield the highest ODR and MDR for the require FAR ($\alpha > 0.005$). ODR and MDR behaved very similarly with varying lag and K values for both detectors, subsequently only MDR is discussed at this stage (Fig. 5). Final ODR for the optimised alarms can be found in Table 2.

NaI:Tl produced the highest MDR (0.8112), at a lag of 3, either side of this lag a decrease in detectability was witnessed. K values appeared to mirror MDR for both detectors suggestive that, as more noise was added through a change in lag, K attempted to negate its influence. The best detectability was witnessed with the smallest K values (NaI:Tl = 3.9 and LaBr:Ce = 6.3). The higher K value for LaBr:Ce and the greater lag (5), infers that the algorithm struggled to cope with the additional noise introduced by internal contamination (Fig. 1). This significantly lowered its MDR (0.5352). Overall, the GC method could not reach the MDR set by SEPA for either detector.

4.2. Performance of machine learning algorithms

The performance of all algorithms is summarised for different binning and detector configurations (Table 2). The combination of a NaI:Tl, an NN and ROIB provided the best ODR of 0.6927 ± 0.008

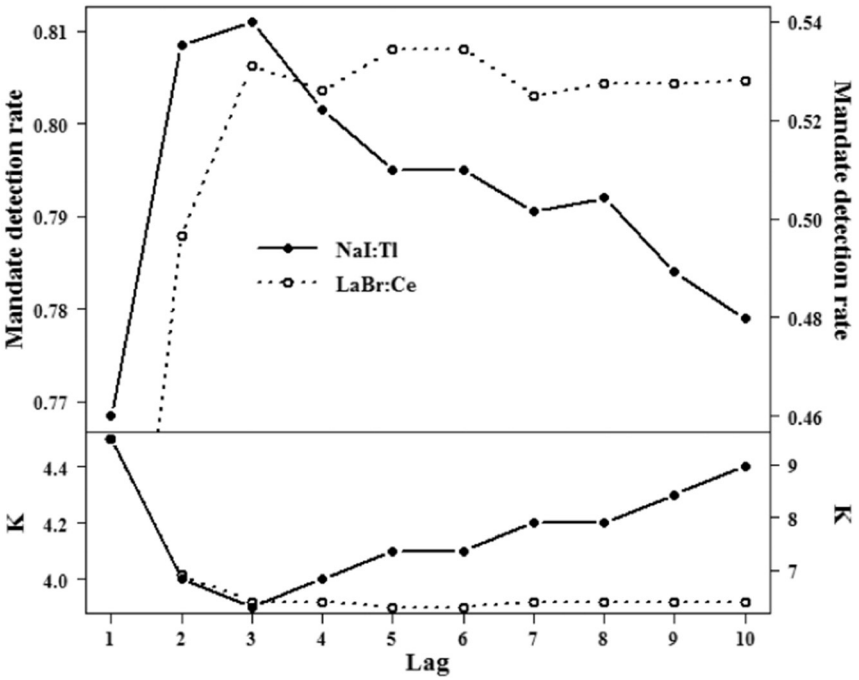

Fig. 5. Mandate detection rates (top plot) and K values (bottom plot) as a function lag for NaI:Tl (empty circles and broken line) and LaBr:Ce (full circles and unbroken line) detectors. NaI:Tl is on the primary axis and LaBr:Ce on the secondary.

(Table 2). What is more, it was the only detector configuration able to confidently attain SEPA's contractual mandate (MDR = 0.9531 ± 0.0072). However, ML produced a varying degree of performance depending on the detection configuration it was employed on. This can only be explained through the influence of individual detector constituents.

4.3. Detector influence

Between the two detectors, NaI:Tl consistently proved to be the most reliable for all ML (Table 2). For instance, the range of MDR for NaI:Tl (0.6092 ± 0.0224–0.9531 ± 0.0072) was consistently higher than LaBr:Ce (0.2028 ± 0.0505–0.5352). Notice, ML did not show any improvement over GC in MDR for LaBr:Ce (although some improvement is demonstrated in ODR). This indicates that LaBr:Ce, although it has higher energy resolution and efficiency, is not suitable for this application. Poor results were almost certainly attributed to the noise introduced, across the entire spectrum, by intrinsic counts over the short acquisition time. This noise would have disguised weaker source signals and made it more difficult for any algorithm to identify the occurrence of source signal. Even after certain bins corresponding to the most contaminated regions were removed no improvement in detection was found.

Conversely, NaI:Tl does not suffer from significant internal contamination permitting algorithms to isolate source signal with more confidence. From an economic perspective this finding is beneficial since NaI:Tl is approximately a fifth of the price of LaBr:Ce. That said, LaBr:Ce's superior energy resolution could play a more significant role in accurate depth and activity estimates useful for mapping purposes. What is more, post-processing noise reduction techniques could be applied in this scenario (Aage et al., 1999; Green et al., 1988). However, given the difference in performance between the detectors, further discussion will focus on NaI:Tl due to its superior *hot particle* detection performance.

4.4. Binning system influence

ROIB provided a consistent increase in detectability (MDR and ODR) over RB (Table 2). In fact RB provided no or little improvement over GC. This implies that by operating fewer bins and tracking changes in sensitive areas known to be associated with source contributions, rather than employing more of a formulated approach, source identification can be improved. The reason for this disparity could be down to the fact that many of the relatively narrow bins generated by the RB, for a single measurement, contained very few counts or even zeros: introducing excessive noise. To add to this complication, some spectral regions of importance were divided between two bins for instance the bin break at 2220 keV for NaI:Tl (Fig. 4). Some regions may also have been unnecessarily separated into multiple bins (i.e. below 200 keV) exhibiting unnecessary correlation. A combination of these elements resulted in the ML over-fitting to both the noise structure and benign changes, making it less sensitive to more subtle source contribution. The reduction in bins

Table 2
Statistics for all detector configurations.

Algorithm	NaI:Tl			LaBr:Ce		
	ODR	MDR	FAR	ODR	MDR	FAR
GC	0.6357	0.8112	0.004	0.5055	0.5352	0.0046
NN-*RB*	0.5982 ± 0.0069	0.7044 ± 0.0867	0.0024 ± 0.0008	0.5427 ± 0.0082	0.5277 ± 0.079	0.0034 ± 0.0017
NN-*ROIB*	0.6927 ± 0.0080	**0.9531 ± 0.0072**	0.00439 ± 0.001	0.5479 ± 0.0040	0.3642 ± 0.0504	0.0048 ± 0.0015
SVM-*RB*	0.5759 ± 0.015	0.6092 ± 0.0224	0.0049 ± 0.0020	0.4873 ± 0.0153	0.2028 ± 0.0505	0.0041 ± 0.0023
SVM-*ROIB*	0.6693 ± 0.0011	0.8855 ± 0.012	0.0039 ± 0.0017	0.5058 ± 0.0194	0.3858 ± 0.0653	0.0086 ± 0.0056

Bold values indicate SEPA's mandate detection rate attained (> 0.95).

(through ROIB) has alleviated some counting noise and simplified the fitting process. This allowed ML to be more robust to changes in overall spectral shape and ultimately leading to better source discrimination. As highlighted earlier, optimisation algorithms could aid in the decision of bins (Wei et al., 2011).

4.5. Machine learning algorithm comparison

In the majority of cases NNs provided better results than SVMs (Table 2), implying that the NNs were more efficient at dividing background-source populations. A possible reason for this difference may be that NNs were slightly better at defining non-linear boundaries within the feature space. However, finding support from the literature is difficult since this is the only study (to the best of the authors' knowledge) that compares the methods in a gamma spectroscopy setting. Studies in other scientific areas have found their performance to be similar (Byvatov et al., 2003; Zhang et al., 2008).

4.6. Advantages of machine learning

The advantage of using ML to interrogate spectral shape as opposed to total signal (GC) can be better understood by reviewing the MDR with total mean count rate (Fig. 6). At low total mean count rate, MDR was relatively high for all algorithms (>0.85). However, as total mean count rate increased, GC's MDR decreased to below 0.7 at the highest total mean count rates. Concurrently, ML tended to be significantly more stable across the total mean count range. At the limits of each population a large amount of uncertainty is observed due to the low sampling size.

One fundamental reason behind this is that GC attempts to separate source and background by assuming them to be two independent univariate Gaussian distributions where the standard deviation ($\alpha = 0.68$) is the square root of the mean (Knoll, 2010). To explain this, consider two background count rates of 100 ± 20 and 250 ± 32 and the introduction of a fixed source contribution of 35 ± 12 ($\alpha = 0.95$). For the first scenario, source and background distributions will not significantly overlap allowing the alarm to be triggered seeing as T is confidently breached (Eq. (1)). Yet in the second case, populations will significantly overlap ending up in the source being missed. Subsequently by using one bin, the signal to noise ratio is reduced relatively consistently with increasing count rate.

By dividing the spectrum into ROIB, more subtle spectral changes occurring across the entire spectrum are captured within spectral comparison ratios. Although the amount of counts has not increased an overall increase in signal to noise ratio is witnessed (Ely et al., 2006). This leads to source information being preserved more efficiently with increasing background count rate.

Another factor is that larger benign systematic fluctuations tend to occur with higher total count rate as some areas will be cluttered with discrete background sources (i.e. masonry or rocks). GC tends to false alarm more regularly in such areas (Fig. 3). To circumvent this occurrence a larger value of K has to be set, thus lowering the MDR in higher count rate areas (Eq. (1)). A dynamic function fitted to K allowing it to take into account general systematic changes with varying mean count rate could arguably be employed. This could decrease FAR slightly. ML methods (NN and SVM), because they take into account changes in shape as well as overall magnitude, are better able to cope with benign fluctuations and ultimately be more sensitive to source contributions.

The improvement ML offers in comparison to GC is demonstrated by observing the MDA with depth (Fig. 7). The NN permitted the NaI:Tl to detect all activities on average 0.1 m deeper into the soil column compared to GC, for example 1 MBq could be detected down to a depth of 0.46 m using GC whilst the NN could distinguish down to a depth of 0.55 m. The SVM demonstrated less of an improvement over GC although still noticeable. The SVM could detect on average 60 mm further into the soil column.

The influence of detector is apparent. The optimised LaBr:Ce, operating an NN, could only just detect SEPA's mandate at the surface, whereas the optimised NaI:Tl could detect it down to a depth of 0.16 m. Notice that at greater depths this disparity diverges slightly, for instance LaBr:Ce could identify a 20 MBq *hot particle* to a depth of 0.7 m, however NaI:Tl (0.9 m) could detect the same activity 0.2 m deeper. This may have been caused by the majority of source photon being scattered to

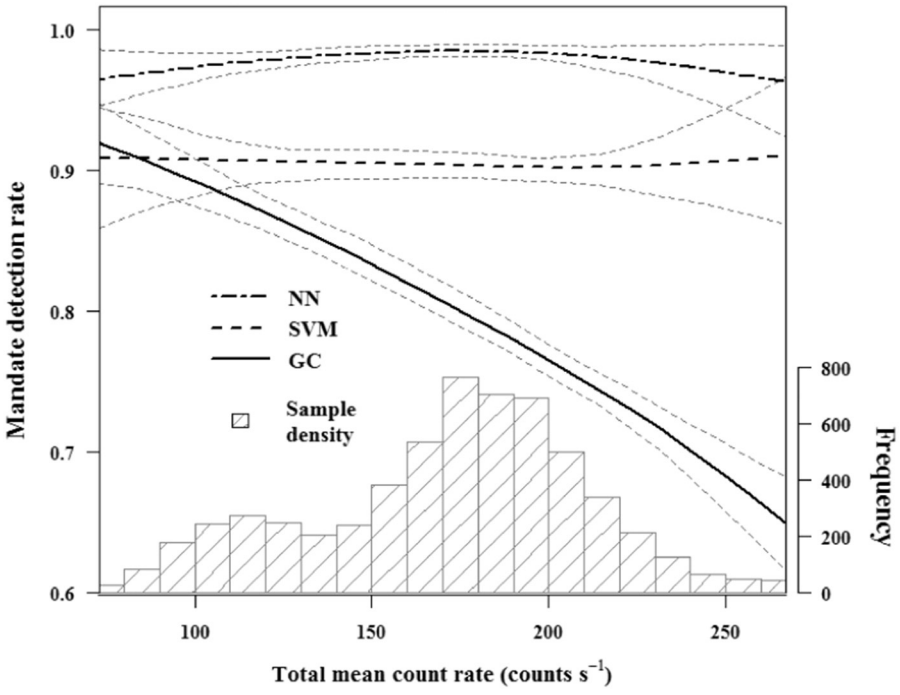

Fig. 6. General trends in mandate detection rate as a function of total mean count rate for optimised neural network, support vector machine and gross count algorithms for NaI:Tl. Machine learning algorithms used the region of interest bins. Sample density and confidence intervals ($\alpha = 0.95$) are included.

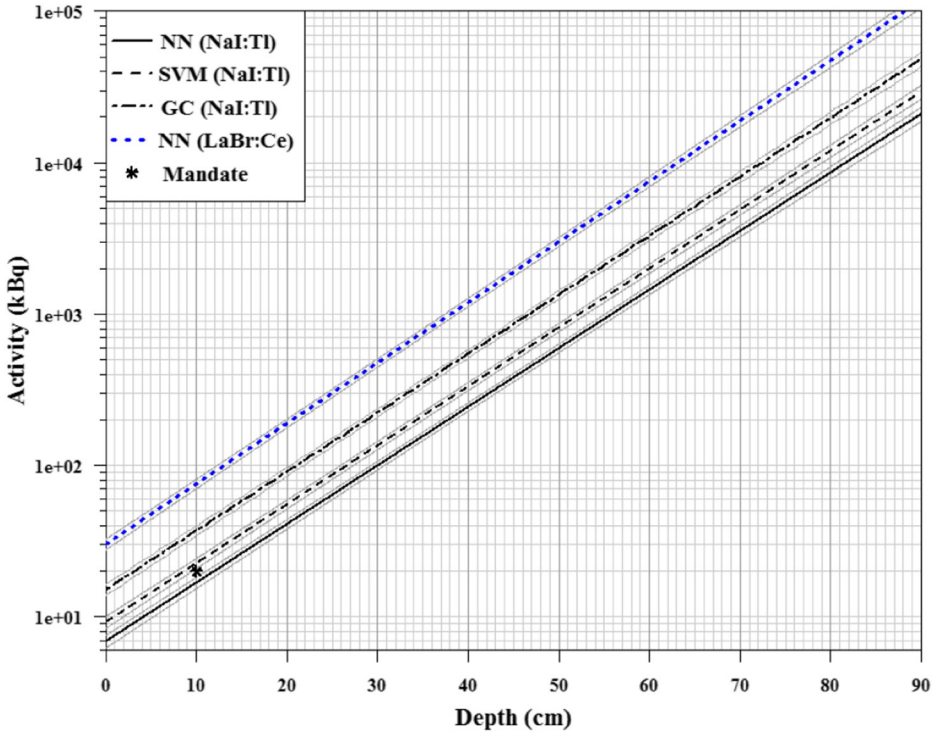

Fig. 7. Minimum detectable activity ($\alpha = 0.95$) functions fitted for selected detector configurations. 95% confidence intervals and SEPA's mandate are included. Region of interest bins was applied to all.

low energy into a region of substantial contamination for LaBr:Ce (<400 keV) (Fig. 1).

4.7. Future prospects for machine learning in routine monitoring

It has been shown that through the use of ML a marked improvement in detection rate could be obtained by a simple change in software as opposed to expensive hardware changes. However, there are considerations that must be taken into account before employing the technique.

NNs, and to some extent SVMs, are powerful when making predictions within the confines of the training dataset as demonstrated through this work. Yet when presented with data outside of this convex hull they are very poor extrapolators (Dragović et al., 2006; Haley and Soloway, 1992). One of the main concerns in this application is the ability of NN to manage changes in background populations brought about by variations on spatial and temporal scales. The spectral comparison ratios transformation to some extent negates this influence (2).

However, it was found that by using two subsets of our data from different parts of the time series and geographical positions, one to train and test and one to cross validate, the detection system performance was significantly reduced. This infers that background variations at Dalgety Bay are complex and cannot be taken fully into account by using spectral comparison ratios. Hence for optimum performance, the background population must be well characterised for all individual areas. This eventually is very much attainable in routine monitoring given that large amounts of data, typically hundreds of thousands of spectra, can be generated over a relatively short period of time. This should allow ML to work within the limits of the known population and ultimately be more sensitive to radiation fields from *hot particles*. Nonetheless, the inter-site capability of the approach may be limited for real-time detection.

Another important issue to be addressed is the representativeness of source calibration data. This study adopted a simplified model on the grounds that correlating materials between the Monte Carlo model and background data is unfeasible. However, acquiring data from routine monitoring may provide a means of supplying additional source data to update the model.

5. Conclusions

The identification of ^{226}Ra *hot particles* through real-time analysis of gamma spectroscopy data can be problematic particularly in the case of a weak source signal and variable background. It has been demonstrated that ML can significantly improve detection limits in this situation by focussing on changes in spectral shape compared to conventional total count rate algorithms. In this context, the intrinsic contamination of LaBr:Ce resulted in poor detection rates compared to NaI:Tl for the detection of ^{226}Ra hot particles. These findings represent a relatively inexpensive development in routine monitoring. The approach outlined in this study could arguably be applied to other sites and other radioisotopes; notably ^{137}Cs.

Acknowledgements

This work was funded by NERC (project — NE/I018956/1) and Nuvia Limited (project — 46007/008). We thank Corynne McGuire and Mark Toner for assisting in field data acquisition and Nuvia Limited for providing the lanthanum bromide detector. We would also like to thank Pete Burgess and an anonymous reviewer for their useful suggestions.

References

Aage, H.K., Korsbech, U., Bargholz, K., Hovgaard, J., 1999. A new technique for processing airborne gamma ray spectrometry data for mapping low level contaminations. Appl. Radiat. Isot. 51, 651–662.

Alamaniotis, M., Heifetz, A., Raptis, A.C., Tsoukalas, L.H., 2013. Fuzzy-logic radioisotope identifier for gamma spectroscopy in source search. IEEE Trans. Nucl. Sci. 60, 3014–3024.

Ao, S., Rieger, B.B., Amouzegar, M.A., 2010. Machine Learning and Systems Engineering. Springer.

Baker, A., Toque, C., 2005. A review of the potential for radium from luminising activities to migrate in the environment. J. Radiol. Prot. 25, 127.

Ball, T., Cameron, D., Colman, T., Roberts, P., 1991. Behaviour of radon in the geological environment: a review. Q. J. Eng. Geol. Hydrogeol. 24, 169–182.

Bergmeir, C., Benítez, J.M., 2012. Package 'RSNNS'.
Briesmeister, J.F., 1993. MCNP—A General Monte Carlo N-Particle Transport Code. LA-12625.
Byvatov, E., Fechner, U., Sadowski, J., Schneider, G., 2003. Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. J. Chem. Inf. Comput. Sci. 43, 1882–1889.
Crawley, M.J., 2012. The R Book. John Wiley & Sons.
Dale, P., Robertson, I., Toner, M., 2008. Radioactive particles in dose assessments. J. Environ. Radioact. 99, 1589–1595.
Dale, P., Gemmill, J., Milne, J., 2013. Dalgety Bay Radioactive Contaminated Land Risk Assessment. Scottish Environmental Protection Agency, SEPA Corporate Office, Erskine Court, Castle Business Park, Stirling.
De Groot, A., Van der Graaf, E., De Meijer, R., Maučec, M., 2009. Sensitivity of in-situ γ-ray spectra to soil density and water content. Nucl. Instrum. Methods Phys. Res., Sect. A 600, 519–523.
DECC, 2012. Radioactive Contaminated Land Statutory Guidance Environmental Protection Act 1990: Part 11a. Contaminated Land. Department of Energy and Climate Change, Area 3C, 3 Whitehall Place, Londonm, SW1A 2AW.
Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., Weingessel, A., 2008. Misc Functions of the Department of Statistics (e1071), TU Wien (R package:1.5-24).
Dragovic, S., Onjia, A., Stankovic, S., Anicin, I., Bacic, G., 2005. Artificial neural network modelling of uncertainty in gamma-ray spectrometry. Nucl. Instrum. Methods Phys. Res., Sect. A 540, 455–463.
Dragović, S., Onjia, A., Bačić, G., 2006. Simplex optimization of artificial neural networks for the prediction of minimum detectable activity in gamma-ray spectrometry. Nucl. Instrum. Methods Phys. Res., Sect. A 564, 308–314.
Du, Q., Wei, W., May, D., Younan, N.H., 2010. Noise-adjusted principal component analysis for buried radioactive target detection and classification. IEEE Trans. Nucl. Sci. 57, 3760–3767.
Ely, J.H., Kouzes, R.T., Geelhood, B.D., Schweppe, J.E., Warner, R.A., 2004. Discrimination of naturally occurring radioactive material in plastic scintillator material. IEEE Trans. Nucl. Sci. 51, 1672–1676.
Ely, J., Kouzes, R., Schweppe, J., Siciliano, E., Strachan, D., Weier, D., 2006. The use of energy windowing to discriminate SNM from NORM in radiation portal monitors. Nucl. Instrum. Methods Phys. Res., Sect. A 560, 373–387.
Fagan, D.K., Robinson, S.M., Runkle, R.C., 2012. Statistical methods applied to gamma-ray spectroscopy algorithms in nuclear security missions. Appl. Radiat. Isot. 70, 2428–2439.
Galushkin, A.I., 2007. Neural Networks Theory. Springer.
Green, A.A., Berman, M., Switzer, P., Craig, M.D., 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Trans. Geosci. Remote Sens. 26, 65–74.
Gurney, K., 2003. An Introduction to Neural Networks. CRC Press.
Guss, P., Reed, M., Yuan, D., Cutler, M., Contreras, C., Beller, D., 2010. Comparison of CeBr(3) with LaBr(3):Ce, LaCl(3):Ce, and NaI:Tl detectors. Hard X-Ray, Gamma-Ray, and Neutron Detector Physics Xii 7805, p. 78050L.
Haley, P.J., Soloway, D., 1992. Extrapolation Limitations of Multilayer Feedforward, Neural Networks.
Harvie, D.I., 1999. The radium century. Endeavour 23, 100–105.
Hendriks, P.H.G.M., Maučec, M., de Meijer, R.J., 2002. MCNP modelling of scintillation-detector γ-ray spectra from natural radionuclides. Appl. Radiat. Isot. 57, 449–457.
Hornik, K., Meyer, D., Karatzoglou, A., 2006. Support vector machines in R. J. Stat. Softw. 15, 1–28.
IAEA, 2003. Guidelines for Radioelement Mapping Using Gamma Ray Spectrometry Data. International Atomic Agency, Austria.
ICRP, 2007. ICRP Publication 103. Ann. ICRP 37 (2).
Iltis, A., Mayhugh, M.R., Menge, P., Rozsa, C.M., Selles, O., Solovyev, V., 2006. Lanthanum halide scintillators: properties and applications. Nucl. Instrum. Methods Phys. Res., Sect. A 563, 359–363.
James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning. Springer.
Jarman, K.D., Runkle, R.C., Anderson, K.K., Pfund, D.M., 2008. A comparison of simple algorithms for gamma-ray spectrometers in radioactive source search applications. Appl. Radiat. Isot. 66, 362–371.
Kangas, L.J., Keller, P.E., Siciliano, E.R., Kouzes, R.T., Ely, J.H., 2008. The use of artificial neural networks in PVT-based radiation portal monitors. Nucl. Instrum. Methods Phys. Res., Sect. A 587, 398–412.
Knoll, G.F., 2010. Radiation Detection and Measurement. John Wiley & Sons.
Kock, P., Lanke, J., Samuelsson, C., 2012. A real-time statistical alarm method for mobile gamma spectrometry—combining counts of pulses with spectral distribution of pulses. Nucl. Instrum. Methods Phys. Res., Sect. A 681, 55–60.
Medhat, M.E., 2012. Artificial intelligence methods applied for quantitative analysis of natural radioactive sources. Ann. Nucl. Energy 45, 73–79.
Menge, P.R., Gautier, G., Iltis, A., Rozsa, C., Solovyev, V., 2007. Performance of large lanthanum bromide scintillators. Nucl. Instrum. Methods Phys. Res., Sect. A 579, 6–10.
Minty, B., Luyendyk, A., Brodie, R., 1997. Calibration and data processing for airborne gamma-ray spectrometry. AGSO J. Aust. Geol. Geophys. 17, 51–62.
Moreira, M.C.F., Conti, C.C., Schirru, R., 2010. A new NaI(Tl) four-detector layout for field contamination assessment using artificial neural networks and the Monte Carlo method for system calibration. Nucl. Instrum. Methods Phys. Res., Sect. A 621, 302–309.
National Nuclear Data Center, 2013. Nuclear Datasheets. 2013.
Olmos, P., Diaz, J., Perez, J., Garcia-Belmonte, G., Gomez, P., Rodellar, V., 1992. Application of neural network techniques in gamma spectroscopy. Nucl. Instrum. Methods Phys. Res., Sect. A 312, 167–173.
Patton, N., Gemmill, J., Milne, J., 2013. Appropriate Person Report. Scottish Environmental Protection Agency, SEPA Corporate Office, Erskine Court, Castle Business Park, Stirling.
Pfund, D.M., Jarman, K.D., Milbrath, B.D., Kiff, S.D., Sidor, D.E., 2010. Low count anomaly detection at large standoff distances. IEEE Trans. Nucl. Sci. 57, 309–316.
Riedmiller, M., Braun, H., 1993. A direct adaptive method for faster backpropagation learning: the RPROP algorithm. Anonymous Neural Networks, 1993, IEEE International Conference onIEEE.
Runkle, R., 2006. Analysis of spectroscopic radiation portal monitor data using principal components analysis. IEEE Trans. Nucl. Sci. 53, 1418–1423.
Sangeetha, R., Kalpana, B., 2010. A comparative study and choice of an appropriate kernel for support vector machines. Anonymous Information and Communication Technologies. Springer.
Sharma, S., Bellinger, C., Japkowicz, N., Berg, R., Ungar, K., 2012. Anomaly detection in gamma ray spectra: a machine learning perspective. Anonymous Computational Intelligence for Security and Defence Applications (CISDA), 2012 IEEE Symposium onIEEE.
Smola, A.J., Schölkopf, B., Müller, K., 1998. The connection between regularization operators and support vector kernels. Neural Netw. 11, 637–649.
Trost, N., Iwatschenko, M., 2002. Method and Device for Detecting Man-Made Radiation.
Tyler, A.N., 2008. In situ and airborne gamma-ray spectrometry. Radioactivity in the Environment. Elsevier (407 pp.).
Tyler, A., Dale, P., Copplestone, D., Bradley, S., Ewen, H., McGuire, C., et al., 2013. The radium legacy: contaminated land and the committed effective dose from the ingestion of radium contaminated materials. Environ. Int. 59, 449–455.
Varley, A., Tyler, A., Smith, L., Dale, P., 2015. Development of a neural network approach to characterise ^{226}Ra contamination at legacy sites using gamma-ray spectra taken from boreholes. J. Environ. Radioact. 140, 130–140.
Wei, W., Du, Q., Younan, N.H., 2010. Particle swarm optimization based spectral transformation for radioactive material detection and classification. Anonymous Computational Intelligence for Measurement Systems and Applications (CIMSA), 2010 IEEE International Conference onIEEE.
Wei, W., Du, Q., Younan, N.H., 2011. Parallel optimization-based spectral transformation for detection and classification of buried radioactive materials. Anonymous Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011. IEEEIEEE.
Zhang, Y., Cong, Q., Xie, Y., Zhao, B., 2008. Quantitative analysis of routine chemical constituents in tobacco by near-infrared spectroscopy and support vector machine. Spectrochim. Acta A Mol. Biomol. Spectrosc. 71, 1408–1413.