JAAS

Cite this: *J. Anal. At. Spectrom.*, 2012, **27**, 432

www.rsc.org/jaas

PAPER

Use of artificial neural networks in measuring characteristics of shielded plutonium for arms control

Matt J. Aitkenhead,*[a] Mark Owen[b] and David M. Chambers[b]

Received 4th August 2011, Accepted 4th January 2012
DOI: 10.1039/c2ja10230g

Artificial neutral networks were developed for use as a potential 'information barrier' technology in the verification of arms control treaty accountable items. They were used to identify and measure specific attributes from γ-ray spectra. These attributes included the presence or absence of plutonium, the plutonium Pu-239/Pu-240 isotopic ratio or ^{239}Pu content and the material age. A set of over 400 training spectra were generated using a spectral simulation software package and various methods for the selection of input data were tested. An input data set which discounted low energy regions susceptible to shielding effects was found to be most effective. Once trained, the network correctly identified the presence or absence of plutonium from real γ-ray spectra. Accurate results were also achieved for estimating the content of ^{239}Pu. In simulated test spectra a root mean squared error (RMSE) of less than 0.1 was found when using the optimum number of inputs. The network was also able to distinguish between spectra from plutonium samples of different ages. Further work is planned to investigate the estimation of a confidence level for whether a specific threshold of ^{239}Pu content is exceeded. An improved training set is anticipated to improve accuracy in determining the material age, which was not achieved accurately.

1. Introduction

The potential for a future treaty in the disarmament of nuclear weapons has prompted the development of various forms of verification technologies[1–3] as well as assurance protocols.[4,5] Although specific requirements vary, such technologies share a common goal in providing a trusted system that allows verification at defined stages of the disarmament process. Previous mock scenarios between states have highlighted the necessity for any disarmament process to present complete transparency to both the host and inspecting parties. This raises many important issues with regards to procedures and technology implementation.

Utilising γ-ray spectrometry has significant potential in warhead verification due to its ability to disclose information regarding isotopic content and material age.[6–9] Both ^{239}Pu and ^{235}U have distinct γ-ray signatures; however the latter displays prominent peaks only at low energies presenting problems in obtaining useful spectra when shielding is present. On the other hand, plutonium isotopes generate peaks of higher energy which are detectable even when shielding such as that of typical warhead containers is present. Therefore, the work presented here focuses on obtaining γ-ray spectrometry of plutonium isotopes, particularly ^{239}Pu, ^{240}Pu and ^{241}Pu.

1.1. Information barriers

Although the spectral information of such isotopes allows information to be extracted relating to the isotopic content and material age, such measurements also pose significant issues with the disclosure of sensitive or proliferative information. Information regarding isotopic content, for example, allows identification of the source of the radioactive material (*i.e.* which reactor it came from).[10] This presents a problem as any verification measurements to be taken are likely to be carried out by an inspecting party external to the host state.

An "information barrier" has been seen for some time as a potential solution to this problem.[11] The idea represents a physical barrier which allows information through that verifies the presence of a nuclear warhead or components of one, whilst preventing the disclosure of any sensitive information. There have been various interpretations of what an information barrier may involve and require.[12,13] The two main requisites are:

- That the host can be confident that sensitive information will not be disclosed to the inspecting party;
- That the inspecting party can be confident that the inspection system will represent an accurate and reproducible result.

The work presented here proposes the application of an artificial neural network (ANN) as an information barrier. The advantage of using an ANN is its ability to be trained to produce

[a]*The James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH, Scotland, UK*
[b]*Threat Reduction Division, Atomic Weapons Establishment, E-mail: matt.aitkenhead@hutton.ac.uk; Tel: +44 (01224) 395257*

a specific output in response to an input dataset. In this instance its function is primarily to (A) recognise the presence of plutonium based on the emission spectrum. The network is also required to identify (B) whether the plutonium present exceeds a specific isotopic ratio and (C) the age of the sample, a factor that influences its isotopic composition. The first two of these requirements involve presenting only a positive or negative Boolean output and thereby do not disclose any additional information which may be deemed classified. Determining the age of the sample, while providing additional information that may be deemed sensitive and subject to classification can be considered an 'optional extra'. A further advantage of using a neural network for the process of isotopic analysis is that the γ-ray spectra generated for training the network can be completely unclassified. Therefore once trained, the network itself contains no classified information and also requires no further alteration; so it can be freely distributed to both hosting and inspecting parties for their own testing and verification of the technology. In this way the ANN can become a trusted and transparent piece of equipment to parties on both sides of the verification process.

While it is true that highly reliable gamma spectral codes currently exist, for example MGA,[14–17] it would be far from simple to convert these codes for Arms Control monitoring. The issue of how such complex codes could be authenticated by two independent parties (host and inspector) is currently an unsolved problem. Indeed, any code which relies on significant computing power to run would be necessity be used on host supplied computing equipment, leaving any inspecting party with considerable issue in trusting any output gained. Assessing alternative routes for gamma analysis is therefore highly desirable for Arms Control purposes. The use of neural networks provides the possibility of future implementation in hardware rather than software for the analysis of gamma spectra, potentially significantly reducing the complexity of monitoring equipment. This potentially would be a big step forward towards trustable measurement technologies for Arms Control purposes. Clearly much work still remains to achieve this, but this initial demonstration of the concept is an important step in this direction.

Further, the demonstration of network training using computer simulated data to high material identification success has potential applications in other radiological detection areas. Widening this training to other gamma detection systems could lead to improved detection success for portal monitoring systems for nuclear security and safeguards purposes, cutting down current high false positive detection rates. Such a system would therefore have significant commercial applications.

1.2. Artificial neural networks

An ANN is an assembly of interconnected processing elements, known as units or nodes. The processing ability comes from the strength or weights of the connections between nodes which are adapted during a learning phase in order to achieve the desired result for a set of training data.[18,19] Their structure is loosely based on the biological neural system of neurons and synapses. A standard network consists of an input layer, one or more hidden layers and an output layer as shown in Fig. 1.

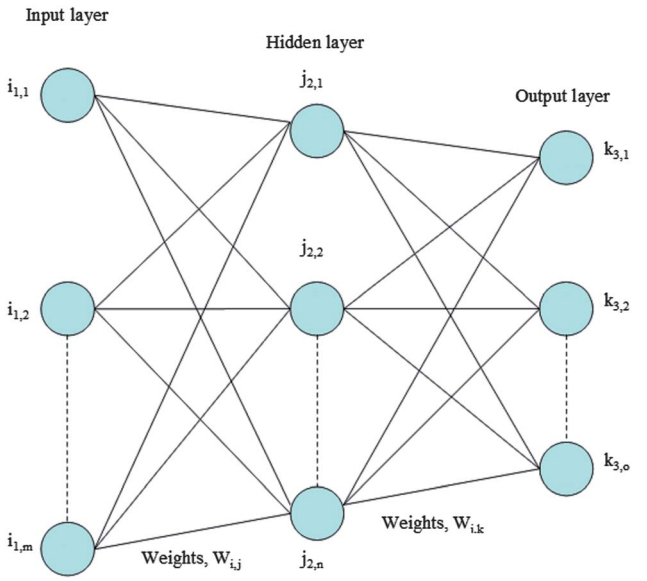

Fig. 1 A simple neural network, showing the basic architecture.

ANNs have previously been developed for spectral classification typically for the identification of a radioactive or nuclear source.[20–22] Reported results have indicated their capability in this field. However previously, these networks have been developed using small numbers of input variables thereby limiting the information extracted from the spectrum. While using a smaller input dataset allows the network to be trained more rapidly, it also removes potentially useful information from the emission spectrum dataset, and could lower the accuracy of the system's response. Objectives of this work include looking at the effectiveness of using larger data sets (*i.e.* some or all of the available emission spectrum data) in order to measure and identify isotopic ratios and material age from complex spectra such as are emitted by plutonium.

Many neural network architectures exist, and there are various techniques employed for training an ANN, with the two most popular being backpropagation (BPNN)[23] and Radial Basis Function Networks (RBFN). It has been observed that the performance of the RBFN fails to match that of the BPNN when the network complexity and the amount of data available are the constraining factors. However, when a simpler training procedure and reduced computational times are required, RBFN is the preferred choice.[24] In the present study where training times were not an issue and the network is of high complexity, the most appropriate for this application was identified as the backpropagation algorithm, using the training approach described by[19] from earlier work by.[23]

Backpropagation neural network training has been shown to be extremely useful for identification and characterisation of many kinds of sources based on their spectral signatures, including stars,[25] plasma chamber leaks[26] and EEGs used to detect sleep patterns in humans.[27] In comparison to some other neural network training methods, backpropagation is relatively easy to implement. As mentioned above, it does often take a larger number of training steps to produce comparable performance, but is more robust following training and with large input data sets. While it has been around for a long time as

a method (approximately 25 years), it continues to be useful and relevant.

In the backpropagation method, the network uses the difference (or error) between the actual output and the desired output from each output node. The actual output node activation is achieved by activating the layers of the network in sequence, with the output signal from a node being in direct relation to the summed input signals from all nodes connected to it in the previous layer. The error at each output node is reduced by adjustment of the weights preceding that node according to simple mathematical rules. This process is then propagated back through the network. The error is gradually reduced by iteration of this process over a number of steps, with the training being halted at either some predefined error threshold, or when no perceivable improvement is being made. A training rate is also required which governs the magnitude of the changes made to the weights, thereby controlling how fast the network is trained. Further information about the number of nodes within each neural network layer, and additional design factors, are given in later sections.

1.3. Spectral regions of interest

Two regions of the plutonium γ-ray spectrum were identified to be of particular interest for the network input, based on information given in.[28] The isotopic ratio of ^{239}Pu to ^{240}Pu can be measured in the relatively high energy region of 635–665 keV. A distinguishing feature of the spectrum produced by plutonium is a symmetrical triplet of peaks centred at 640 keV. Fig. 2 shows how this triplet varies from low ^{239}Pu content (ratio 1 : 1 of ^{239}Pu to ^{240}Pu and ^{241}Pu) up to high ^{239}Pu content (ratio 1 : 0) due to the ^{240}Pu peak at 643 keV.

Age determination of plutonium samples is typically done using the analysis of ^{241}Am content exploiting the decay of ^{241}Pu to ^{241}Am and ^{237}U. The energy region of 330–350 keV is particularly useful for these measurements as it contains peaks at 332.4 keV and 335.4 keV generated by excited states of ^{237}Np which are populated from ^{241}Am decay and ^{237}U. Fig. 3 shows how these peaks vary as the plutonium age is increased. It should be noted that these peaks are generally unresolved from ^{239}Pu peaks which occur at 332.8 keV and 336.1 keV. The contribution from ^{239}Pu

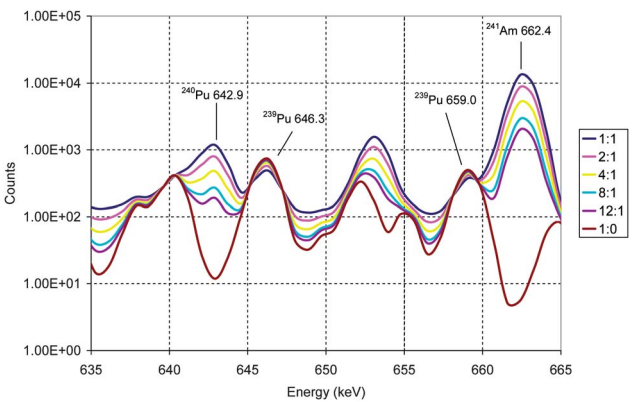

Fig. 2 Simulated γ-ray spectrum of 5 years old plutonium for different ratios of ^{239}Pu to ^{240}Pu and ^{241}Pu, with a logarithmic scale of counts (intensity) *versus* wavelength energy.

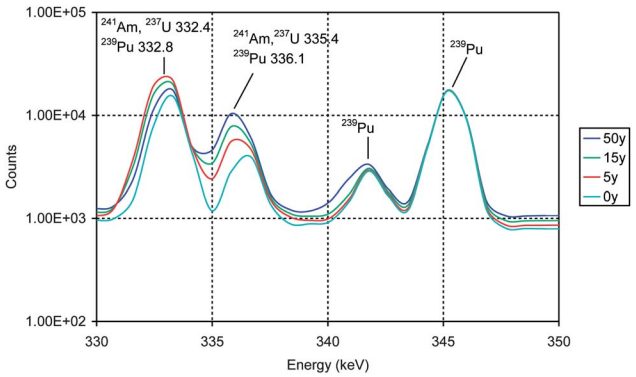

Fig. 3 Simulated γ-ray spectrum of plutonium (92% ^{239}Pu) for various ages displaying changes in ratio of ^{241}Pu to ^{241}Am and ^{237}U, with a logarithmic scale of counts (intensity) *versus* wavelength energy.

can however be analysed using the well resolved ^{239}Pu peaks at 341.5 keV and 345.0 keV.

This analysis is normally carried out using a combination of visual peak analysis (to identify the presence or absence of signature peaks) and measurement of the comparative height of different peaks to give an indication of isotopic ratio. Clear mathematical relationships between peak height and isotopic ratio are difficult to produce, as individual peaks are strongly influenced by sample shape, size, sampling time, shielding material and the isotopic mix that is present. Both energy regions discussed above are of use in determining the presence or absence of plutonium, but it is difficult to draw any more information from the spectrum than this using traditional approaches. The use of neural networks in producing more accurate isotopic ratio information is based on the ability of neural networks to detect patterns in large, noisy and inconsistent datasets.[29,30]

In addition, peaks from the isotopes of ^{239}Pu, ^{240}Pu and ^{241}Pu are also found outside of these identified regions, meaning that there might be additional identification/measurement accuracy to be found through the inclusion of other portions of the emission spectra. In order to develop this idea, an investigation was carried out to determine whether varying the size of the spectral region presented to the neural network affects the accuracy of the output results.

2. Methods

2.1. Generating training data

A deterministic 1-dimensional modelling software package developed in-house at AWE was used for spectral simulation to generate a set of training data for the network. This allowed a wide range of controlled spectra to be generated in a reduced time scale and for a reduced cost compared to that of obtaining the measurements experimentally. The model consisted of a 500g sphere of plutonium of defined isotopic ratios, surrounded by a defined uniform layer of shielding material. The shielding sphere was surrounded by a vacuum. A 40% high purity germanium detector was modelled at 1 metre from the centre of the sphere and took measurements for 1000 s intervals. A matrix was generated to define the most relevant model parameters and how they would be varied; consisting of:

• Isotopic ratio of ^{239}Pu to ^{240}Pu and ^{241}Pu; 1 : 1, 2 : 1, 4 : 1, 8 : 1 and 12 : 1;

• Material age; 0, 5, 15 and 50 years;

• Shielding material; plastic, aluminium, steel, lead and depleted uranium;

• Shielding material thickness; 3 to 4 values depending on the material to give realistic values.

A similar model was used to generate non-plutonium spectra, typically of common radioisotopes found in the background or used in the commercial or medical industries. This data was used to train the network to recognise the absence of plutonium and consisted of spectra generated from radioisotopes including: ^{241}Am, ^{133}Ba, ^{109}Cd, ^{252}Cf, ^{57}Co, ^{60}Co, ^{137}Cs, ^{152}Eu, ^{131}I, ^{192}Ir, ^{40}K, ^{237}Np, ^{226}Ra, ^{99}Tc, ^{232}Th, ^{233}U, ^{235}U, ^{238}U and ^{88}Y. Overall, the training set consisted of a total of 420 samples including both plutonium and non-plutonium spectra.

The generated training spectra provided a huge range in total counts, primarily due to the effects of different shielding. In order to allow the neural network to learn effectively, each input value was required to be in the range [0, 1]. Therefore some kind of normalisation of the data was required. In order to ensure that the normalisation procedure carried out was the best possible, a range of different approaches was tested. The most effective method was found to use a moving average over a window of a certain size centred on each data point, with the value each point divided by this moving average. All datasets were also normalised to fully occupy the range [0, 1], by dividing each value after conversion by the maximum value in the converted spectrum. This improved the network training and accuracy and ensured greater consistency between datasets. Tests were carried out with a range of moving average window widths. An optimal width was found to be within the range [60, 200], with no significant variation across this range. A window width of 100 was therefore selected as being as close to the best value as could be determined. Throughout these optimisation tests, all 420 training datasets were used.

2.2. Generating test data

A small dataset of spectra was generated for testing using the trained neural network models. Different parameter values were used in order to test the network on datasets that were not the same as any of the training data. The parameter values that were used did however fall within the parameter value ranges used for training the neural network models. While it is possible that the network could be tested in situations where certain parameters, such as shielding thickness or material age, are outwith the range used for training, it is not reasonable to expect the network to achieve the same levels of accuracy as it can for parameters lying within its training experience. The total number of modelled test spectra was 152.

A number of experimentally-generated spectra were also identified for network testing. These included a set of spectra generated from a plutonium sphere with various types of shielding, a number of PIDIE (Plutonium Isotopic Detonation Intercomparison Exercise) spectra (standardised plutonium spectra developed by[31] and a set of non-plutonium spectra consisting of some common radioisotopes, giving a total of 24 spectra. In order to present these spectra as an input to the network, it was necessary to reformat and normalise them into the same configuration as the modelled spectra. A small piece of software was written to enable this reformatting; providing the experimental spectra with the same number of input points at the same calibration as the modelled spectra. It should be noted that such treatment of the experimental measurements along with the data normalisation discussed previously would be required in any potential future systems, due to the range of measuring devices that are available and their different output formats.

2.3. Neural network development

2.3.1. Structure & training algorithm. The chosen neural network structure consisted of an input layer, two hidden layers and an output layer. Each layer was fully connected to the following layer, with connection weightings being randomized in the range [−0.25, 0.25] prior to training. The size of the input layer was varied in order to identify the most effective energy range to be used, with the first node in the layer corresponding to the lowest energy band for that spectral 'window', and the last node corresponding to the highest energy band in that window. In one training case, two windows were used, at different energy ranges. Each energy band was approximately 0.85 keV wide. Three sets of windows of the spectra were chosen for training, to identify optimal performance: 300–400 keV and 600–700 keV; 260–850 keV; 5–1277 keV. These windows were chosen to include known peaks from the plutonium spectrum, or in the case of the third option to include the entire spectral range.

In each case, the number of nodes in the hidden layer equalled double the number of input nodes or 100 nodes, whichever was less. The output layer consisted of three nodes relating to:

(a) Pu presence/absence;

(b) Percentage ^{239}Pu content;

(c) Age determination.

Regardless of size, each network was trained using the back-propagation neural network training algorithm. All possible combinations of training step count (1000, 2000, 5000, 10000, 20000 and 50000 steps trialled), training rate (values of 0.001, 0.002, 0.005, 0.01, 0.025, 0.05, 0.1) and momentum rate (values of 0.1, 0.3, 1, 2, 5) were investigated. The final training was carried out for 10000 steps, with a training rate of 0.025 and a momentum rate of 1. It was found that using these parameter values, the network was not overfitted to the training data but that accuracy in comparing against test data was optimised. For further information on the implementation of the back-propagation training method, the reader is referred to[23] or to.[19]

2.3.2. Using simulated annealing to identify optimal network inputs. As the number of data points used for network input could potentially be very high, resulting in a network that required a great amount of computational time to train and use, the possibility of reducing the number of inputs was explored. This would speed up the network operation and potentially improve its accuracy. For this to be achieved the points which provide the greatest information content within the overall dataset would have to be found. A technique of finding these data points known as simulated annealing was used. This technique essentially 'evolves' the locations of a fixed number of sample sites from across the overall dataset, *i.e.* sample site

locations are adjusted by small random values until a situation of least error between the actual and predicted values is found. By using this strategy and varying the number of sample points chosen, it is possible to find near-optimal locations for each sample point and determine the accuracy of a neural network trained with data from only these points.

This technique was used for a network trained to identify the presence of plutonium and its isotopic ratio. The number of input sample points was varied between 10, 20, 50 and 100 with two network outputs. In each case, the number of nodes in each of the hidden layers equalled the number of nodes in the input layer, which equalled the sample count. The simulated annealing was carried out using the following algorithm over a total of 10000 annealing steps:

1. Initialise the sample locations randomly across the full spectral range (5–1277 keV), and set the initial sample 'jump' size to 50.

2. Measure the performance of the network trained with the initialised sample positions. Call this performance Q_1, normalised using the range of possible values to the range [0, 1] (where low values indicate lower error rates).

3. Randomly reposition the sample positions up to the current jump size away from their current positions (higher or lower).

4. Measure the performance of the network trained using the new sample positions. Call this normalised performance Q_2.

5. Calculate the probability P of accepting the new sample positions as equal to:

$$P = e^{-(Q_2 - Q_1)/\Delta} \qquad \text{(eqn. 1)}$$

Where Δ equals 0.02 (determined through trial and error). If $Q_2 < Q_1$ (*i.e.* the performance has improved), then the new sample positions are accepted automatically and the value of Q_1 is replaced by that of Q_2. If $P < 1$ then a random number generator is used to determine whether or not the new sample positions are accepted, and Q_1 remains the same.

6. Reduce the sample 'jump' size by 1 every 200 annealing steps.

7. Repeat as from step 3.

The results for using this technique showed that as the number of samples decreases, the system performance worsens. It also shows that even for 100 samples, the RMSE is generally worse than when the entire range is used. Overall, the results show that even when a large number of inputs are used, the strategy of selecting a set of sample points using what should be the best available method degrades system performance. Therefore it was decided that the network would be provided with as much of the datasets for input as possible, with the only limiting factors being spectral effects of shielding at low energies or background peaks which would preferably be excluded. Determination of the dataset input size is thus described in the following section, with these constraints in mind.

3. Results

3.1. Optimising the window range

A key objective was to optimise the window range to be used for network input. Therefore three window ranges were selected for comparison:

(a) 300–400 keV and 600–700 keV;
(b) 260–850 keV;
(c) 5–1277 keV.

Size (a) was selected as this region provides the most valuable information from the spectrum regarding the isotopic content and the material age as discussed in Section 1.3. The two larger sizes were chosen due to conclusions of the network development stage, where it was found that in general, a larger dataset for input improved the network performance. Size (c) was chosen to capitalise on this result including a wide range of the spectral data whilst discounting the spectral region containing the prominent background peak at 1.46 MeV from potassium-40. Size (b) was chosen as an intermediate between the other two window sizes which also discounts the lower energy region of the spectrum due to its higher susceptibility to variations from shielding effects.

The root mean squared error (RMSE) was calculated for each output category and each window range for the training data set (420 spectra), the modelled data set (152 spectra) and the real data set (28 spectra). In addition, the r^2 values for correlation between the target and network values were also calculated, for each dataset involved. These results are displayed in Fig. 4. For presence/absence predictions, the p-values for training, model test and real test datasets were less than 0.01 for every window size, indicating very low probability that these results were obtained by chance. For Pu-239 content, p-values ranged from 0.023 (Training (c) to 0.056 (Training (a)), indicating low

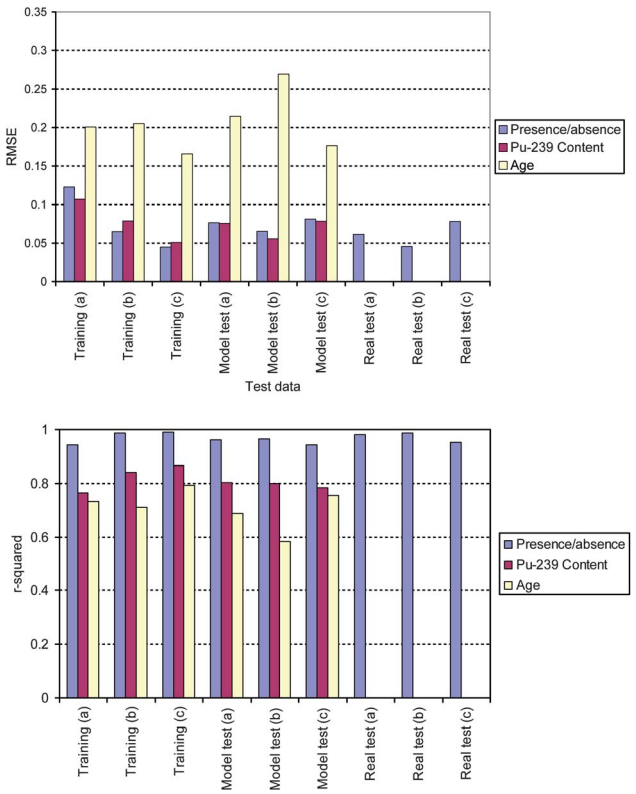

Fig. 4 RMSE and r-squared comparison for network which provides an estimate of ^{239}Pu content. Results for each network output are displayed for training data, modelled test data and real test data using the three window sizes (a), (b) and (c).

probability of the results being obtained by chance. For sample age, p-values tended to be higher, ranging from 0.050 (Training (c)) to 0.121 (Model test (b)), with all but Model test (b) giving values below 0.1. These values are moderately low, but not as powerful as for presence/absence or Pu-239 content.

For the real data set the RMSE could not be calculated for either ^{239}Pu content or age as neither of these attributes were known for the given spectra. While Fig. 4 gives the RMSE values for the presence/absence outputs, it would perhaps be more appropriate to express this result as a misclassification error rate. Setting the threshold for classification at 0.5, with values below this meaning absence and values above meaning presence, the network achieved a 100% accuracy rate for all datasets. The largest error value found for presence/absence was less than 0.2, indicating high certainty of classification in each case.

The results shown in Fig. 4 demonstrate relatively small RMSE values for Pu presence/absence and ^{239}Pu content, with the RMSE values for age being significantly larger. This is assumed to be due to the training set consisting of an uneven spread of values with unrealistic values (such as 0 years) being included. Therefore, it is assumed that with improved variations of age in the training set, reduced RMSE values could potentially be achieved.

Fig. 4 also shows that the smallest RMSE for the training data occurs for the largest window size, (c). This is in agreement with the initial results found in the network development as discussed in Section 2.3.2 which found that the larger the dataset used the smaller the RMSE. Excluding the large RMSE produced for the age determination, the most accurate results for both the modelled and real test data sets were found using window size (b). This may correspond to the fact that window size (b) excludes the low energy region of the spectrum, which is most susceptible to shielding effects. Size (b) also includes additional information not provided by window size (a). Following this result, window size (b) is analysed for the remainder of this section.

3.2. Presence/absence determination

The most significant test performed on the trained network to look for the presence/absence of plutonium was that on the real experimental data. The network consistently gave an output in agreement with the desired value as shown in Fig. 5.

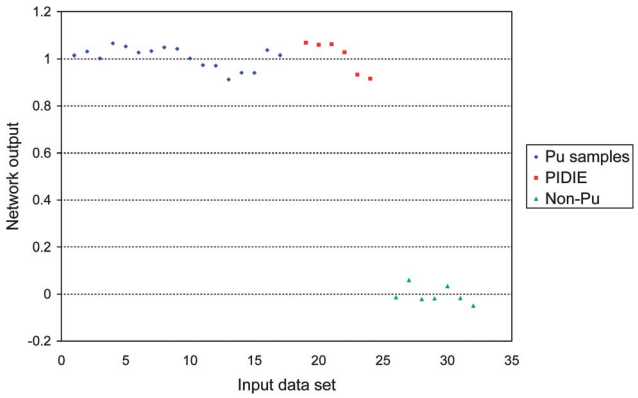

Fig. 5 Network outputs for the presence (1) or absence (0) of plutonium in real test spectra, using window range (b).

3.3. ^{239}Pu content

Due to the lack of variation regarding the plutonium-239 content of the real data set, the modelled data set was used to analyse this network output. The results are shown in Fig. 6, comparing the actual ^{239}Pu content with the network's estimated value. Although the overall RMSE for the plutonium-239 content is similar to that for the presence/absence of plutonium, Fig. 6 shows how this error can be much more significant as the various content ratios used can sometimes differ by less than the RMSE involved. This is reflected in the lower r-squared values for plutonium-239 content in Fig. 4. A more effective method may be to develop the network with a node output which gives a confidence level for whether a certain threshold content level is exceeded. For example this could be implemented by training the output node to give a value of 1 for a plutonium-239 content above a given threshold value and a value of 0 for a content below a given threshold value. Two threshold values could even be used, possibly separated by a difference of 0.1, thereby providing the full node range (from 0 to 1) for results which may appear in the most ambiguous region.

3.4. Age determination

As with the plutonium-239 content, the accuracy of the age determination was analysed for the modelled test data set as shown in Fig. 7. As can be seen, the only samples for which the age determination is relatively accurate (samples 25–28) are actually non-Pu. The large discrepancies between the network output and the desired output may be attributed to the irregular distribution of ages within the training data. Using an age of 0 years may also be unrealistic due to the complete exclusion of Am^{241}. The consequence of this is that the sample produces a spectrum similar to that when 100% Pu^{239} is used, as shown in Fig. 2. The large variation produced from either of these conditions may affect the efficiency and accuracy of training the network, therefore the inclusion of such data should be considered. However, it is also observable from Fig. 4 that the network has quite high r-squared values considering the high RMSE values; this indicates that adjustment of the network outputs

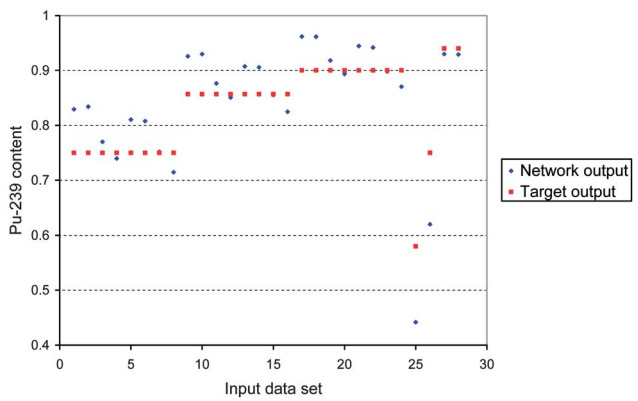

Fig. 6 Network outputs for ^{239}Pu content of modelled test spectra using window range (b).

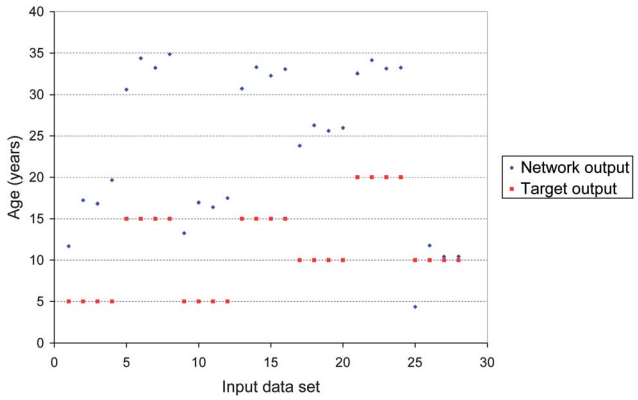

Fig. 7 Network outputs for material age of modelled test spectra using window range (b).

according to an inherent bias may improve the results. Examination of Fig. 7 shows that for the majority of the test datasets, the network outputs appear to be offset by approximately 15 years on average.

Improvements on the results of Fig. 7 could potentially be achieved by better defining the age variation in the training spread or using a more realistic and even spread of ages. It is also possible that a separate network dedicated to age alone could improve prediction accuracy. However, the results obtained here show that age determination using this neural network approach is not sufficiently accurate to be usable. Therefore, it cannot be considered to have been accomplished.

4. Conclusions

We have demonstrated a method for using neural networks to identify and characterise radioactive material, in this case plutonium-239, on the basis of gamma-ray emission spectra taken through shielding. Such a neural network could be implemented easily within a laptop or PDA attached to a gamma-ray detector, and used in the field for arms control verification. The network itself would not contain any information of a sensitive nature, and could be used transparently and rapidly to verify claims regarding potentially threatening material.

The network performance has been consistent in identifying the presence or absence of plutonium in a set of real γ-ray spectra different to that of the simulated data used for training. In addition, the network was tested on its ability to identify the ^{239}Pu content from a range of simulated γ-ray spectra, also different to that of the training data. Although the RMSE of the output values were of similar magnitude to those for the absence/presence of plutonium, the effect was much greater as the difference in ^{239}Pu content of sample data was comparable to the RMSE. To improve the accuracy in these results, a second method will be trialled for the network development. Here the output will offer a confidence level relating to whether a certain threshold ^{239}Pu content level is exceeded or not, rather than calculating a precise value.

Such a method could be more useful as it does not output an estimate for the ratio which if accurate could be deemed classified when employed in a real application. However a comparison of the two methods is recommended to provide a more comprehensive exploration of the network's capability. If the first method was in fact found to be more effective then additional code could be added to compare the ratio estimate with a given threshold and give a positive/negative output relating to the presence/absence of that ratio.

As an extension of analysing the network's capability, its ability to identify plutonium age from γ-ray spectra was also investigated. The error was much larger for this output and it was found that the network was not able to distinguish between spectra of plutonium of different ages. By imposing an offset to the network output the error could be reduced significantly, but this is not an approach that could be relied upon in practice as in some cases it would result in a negative value for age. Further improvement in determining plutonium age might be achieved by improved selection of training data and including spectra from plutonium of a more even distribution of ages.

Ideally once an optimum neural network is developed using software, a hardware implementation would be much more appropriate for the desired application. However the accuracy of identifying ^{239}Pu content requires significant improvement before such development could be investigated as the confidence levels would be have to be much higher for an operational verification technology. Future work in this area should focus on improvement of the method using more advanced neural network designs and training algorithms, and increasing the quantity of the training data.

References

1 B. McGinnis, A. Raffo-Caiado, M. Whitaker, S. Hayes, O. M. Guidicini, 2004. Ten years of technical cooperation between the DOE and ABACC. *Proceedings of the 7th International Conference on Facility Operations: Safeguards Interface*, pp. 417–430.
2 S.-P. Liu, F.-H. Hao, J. Gong, H.-L. Wu, W. Kang, G.-C. Hu, Y.-C. Xiang and J.-H. Zhang, γ-ray integrated prototype system for the measurement of plutonium component attributes, *Wuli Xuebaol Acta Physica Sinica*, 2007, **56**(8), 4406–4411.
3 S.-P. Liu, J. Gong, F.-H. Hao and G.-C. Hu, Template identification technology of nuclear warheads and components, *Chin. Phys. B*, 2008, **17**(2), 363–369.
4 Kinji Koyama, What the verification regime under a fissile material cutoff treaty could be like: A preliminary view. JNMM, *Journal of the Institute of Nuclear Materials Management*, 1999, **27**(2), 48–52.
5 J. Loading and T. Rauf, IAEA & NPT: The verification challenge, *IAEA Bulletin*, 2005, **46**(2), 20–25.
6 R. K. Walter, R. L. James, F. Leon, 1997. *The Application of High-Resolution Gamma-Ray Spectrometry (HRGS) To Nuclear Safeguards, Nonproliferation and Arms Control Activities* (New York: Brookhaven National Laboratory).
7 A. Peurrung, R. Arthur, R. Elovich, B. Geelhood, R. Kouzes, S. Pratt, R. Scheele and R. Sell, The 871 keV gamma ray from 17O and the identification of plutonium oxide, *Nucl. Instrum. Methods Phys. Res., Sect. A*, 2001, **474**(3), 285–293.
8 A. Bernstein, B. A. Brunett, N. R. Hilton, J. C. Lund and J. M. Van Scyoc, The "radiation continuity checker": An instrument for monitoring nuclear disarmament treaty compliance, *IEEE Trans. Nucl. Sci.*, 2002, **49 II**(3), 864–869.
9 X.-J. Zhang, J. Wang, S.-L. Wang and Y.-Q. Fan, Intercomparison and evaluation of several software for gamma-ray spectra, *Hedianzixue Yu Tance Jishu/Nuclear Electronics and Detection Technology*, 2007, **27**(5), 825–829.
10 K. J. Moody, I. D. Hutcheon, P. M. Grant, 2005. *Nuclear forensic analysis*, 512pp. CRC Press.
11 W.-W. Zhao, Information barrier in nuclear disarmament: methods and techniques, *Hedianzixue Yu Tance Jishu/Nuclear Electronics and Detection Technology*, 2005, **25**(6), 866–873.

12 D. MacArthur, R. Whiteson, J. W. Wolford, 1999. Functional Description of an Information Barrier to Protect Classified Information. *Proceedings of the 40th Annual Meeting of the Institute of Nuclear Materials Management, 25–29 July 1999*, Phoenix, Arizona.

13 The Joint United States DOE-DOD Information Barrier Working Group, 1999. The Functional Requirements and Design Basis for Information Barriers. *Pacific Northwest National Laboratory Report, PNNL-13285*.

14 R. Gunnink., 1990. MGA: A gamma-ray spectrum analysis for determining plutonium isotopic abundances. *Lawrence Livermore National Laboratory Report, vol. 1, UCRL-LR-103220*, Livermore, California.

15 R. Gunnink, W. D. Ruhter, P. Miller, J. Goerten, M. Swinhoe, H. Wagner, J. Verplancke, M. Bickel, S. Abousahl, 1994, MGAU: a new analysis code for measuring U-235 enrichments in arbitrary samples. In: *IAEA Symposium on International Safeguards, March 8–14, 1994*, Vienna, Austria.

16 T. E. Sampson, G. W. Nelson, T. A. Kelley, 1989. FRAM: a versatile code for analyzing the isotopic composition of plutonium from γ-ray pulse height spectra. *Los Alamos National Laboratory Report LA-11720-MS*.

17 G. Brunetti, F. V. Frazzoli, D. Parise, T. Dragnev, R. Abedin-Zadeh, M. Aparo, S. Guardini and A. Ravazzani, An application of gamma/X-ray/weighing methods for the complete in field assay of nuclear materials, *Nucl. Instrum. Methods Phys. Res., Sect. B*, 2004, **213**, 414–417.

18 K. Gurney., 1997. *An Introduction to Neural Networks*. UCL Press.

19 M. J. Aitkenhead, A. J. S. McDonald, J. J. Dawson, G. Couper, R. P. Smart, M. Billett, D. Hope and S. Palmer, A novel method for training neural networks for time-series prediction in environmental systems, *Ecol. Modell.*, 2003, **162**(1–2), 87–95.

20 P. E. Keller and R. T. Kouzes, Gamma spectral analysis *via* neural networks, *Nuclear Science Symposium and Medical Imaging Conference*, 1995, **1**, 341–345.

21 E. Yoshida, K. Shizuma, S. Endo and T. Oka, Application of neural networks for the analysis of gamma-ray spectra measured with a Ge spectrometer, *Nucl. Instrum. Methods Phys. Res., Sect. A*, 2002, **484**, 557–563.

22 S. Clarke, R. Gao, Y. Xu, L. H. Tsoukalas and T. J. Downar, Artificial networks for nuclear source characterisation, *Transactions of the American Nuclear Society*, 2005, **93**, 327–328.

23 D. E. Rumelhart, G. E. Hinton and R. J. Williams, Learning representations by back-propagating errors, *Nature*, 1986, **323**(6088), 533–536.

24 S. Garg, S. Pal and D. Chakraborty, Evaluation of the performance of backpropagation and radial basis function neural networks in predicting the drill flank wear, *Neural Comput. Appl.*, 2007, **16**(4), 407–417.

25 D. Ordonez, C. Dafonte, M. Manteiga and B. Arcay, Parameter extraction from RVS stellar spectra by means of artificial neural networks and spectral density analysis, *Hybrid Artificial Intelligence Systems*, 2008, **5271**, 212–219.

26 B. Kim and S. Kwon, Wavelet-coupled backpropagation neural network as a chamber leak detector of plasma processing equipment, *Expert Syst. Appl.*, 2011, **38**(5), 6275–6280.

27 R. K. Sinha, Artificial neural network and wavelet based automated detection of sleep spindles, REM sleep and wake states, *Journal of Medical Systems*, 2008, **32**(4), 291–299.

28 D. Archer, T. Gosnell, J. Luke, L. Nakae, 2001. Pu300, Pu600, and Pu900 Systems. *Arms Control and Non-proliferation technology*, pp. 32–34, Office of Non-proliferation Research and Engineering, US Department of Energy, Spring.

29 M. J. Aitkenhead, M. J. Mustard and A. J. S. McDonald, Using neural networks to predict spatial structure in ecological systems, *Ecol. Modell.*, 2004, **179**(3), 393–403.

30 T. Kavzoglu, Increasing the accuracy of neural network classification using refined training data, *Environ. Modell. Software*, 2009, **24**, 850–858.

31 J. Morel, B. Chauvenet, ESARDA NDA-WG Members, 1992. Intercomparison of plutonium isotopic composition measurements by X- and gamma-ray spectrometry, results from the PIDIE exercise. *Rapport CEA R-5582*, CEA Saclay, France.