Clustering of Self Powered Neutron Detectors: Combining Prompt and Slow Dynamics

Rihab A. Razak, Mani Bhushan, Madhu N. Belur, Akhilanand P. Tiwari, Mahendra G. Kelkar, and Mahitosh Pramanik

***Abstract*—The focus of this work is on clustering Self Powered Neutron Detectors (SPNDs) with different dynamic characteristics into smaller groups, with each group containing highly correlated SPNDs. In order to cluster the SPNDs correctly, we propose novel ways to compensate for the effect of different dynamic response characteristics. In particular, two types of SPNDs: (i) cobalt, which give a prompt response, and (ii) vanadium, which give a delayed response, are considered. We propose and compare three compensation methods to cluster both types of SPNDs *together* using their measurement data: (i) pure delay applied to cobalt SPND data, (ii) cobalt SPND data 'slowed down' to match the vanadium SPND dynamics by passing it through vanadium SPND transfer function, and (iii) vanadium SPND data 'speeded up' to match the cobalt SPND dynamics by passing it through the inverse of vanadium SPND transfer function. Based on extensive simulations, it is found that slowing down cobalt SPND measurements to match the vanadium SPND dynamics yields the best results. This method is then used to obtain clusters from data obtained from a nuclear reactor in India for both vanadium and cobalt SPNDs and the resulting clusters appear reasonable.**

***Index Terms*—Clustering, correlation, dynamics matching, self powered neutron detectors.**

I. Nomenclature

Reactor/SPND Related Variables and Their Units

$G_v(s)$ vanadium SPND transfer function, $\mathrm{A.cm^2.s}$

$I_{60}(t)$ current in cobalt SPND due to β-decay of ^{60}Co, A

$I_{61}(t)$ current in cobalt SPND due to β-decay of ^{61}Co, A

$I_d(t)$ delayed current for cobalt SPND, A

$I_{p,tot}(t)$ total prompt current produced by cobalt SPND, A

$I_{p,n}(t)$ current due to the γ-rays produced internally, A

$I_{p,\gamma}(t)$ current generated by cobalt SPND due to the external γ flux, A

$I_v(t)$ output current of vanadium SPND, A

K_{60} sensitivity constant for $N_{60} = 6.3147 \times 10^{-30}$ $\mathrm{A.cm^3}$

K_{61} sensitivity constant for $N_{61} = 1.7668 \times 10^{-24}$ $\mathrm{A.cm^3}$

K_{gv} sensitivity constant corresponding to the delayed current component of vanadium SPND, $\mathrm{A.cm^3.s}$

K_{pv} sensitivity constant corresponding to the prompt current component of vanadium SPND, $\mathrm{A.cm^3.s}$

$N_{51}(t)$ concentration of ^{51}V, $\mathrm{atoms.cm^{-3}}$

$N_{52}(t)$ concentration of ^{52}V, $\mathrm{atoms.cm^{-3}}$

$N_{59}(t)$ concentration of ^{59}Co, $\mathrm{atoms.cm^{-3}}$

$N_{60}(t)$ concentration of ^{60}Co, $\mathrm{atoms.cm^{-3}}$

$N_{61}(t)$ concentration of ^{61}Co, $\mathrm{atoms.cm^{-3}}$

λ_{52} decay constant of ^{52}V $= 3.6 \times 10^{-3}$ $\mathrm{s^{-1}}$

λ_{60} decay constant of ^{60}Co $= 416.9 \times 10^{-12}$ $\mathrm{s^{-1}}$

λ_{61} decay constant of ^{61}Co $= 116.7 \times 10^{-6}$ $\mathrm{s^{-1}}$

ϕ neutron flux, $\mathrm{n.cm^{-2}.s^{-1}}$

S_c sensitivity constant corresponding to flux for cobalt SPND $= 0.813 \times 10^{-20}$ $\mathrm{A.cm^2.s.n^{-1}}$

σ_{51} microscopic absorption cross section of ^{51}V for neutron capture $= 4.9 \times 10^{-24}$ $\mathrm{cm^2}$

σ_{59} microscopic absorption cross section of ^{59}Co for neutron capture $= 37 \times 10^{-24}$ $\mathrm{cm^2}$

σ_{60} microscopic absorption cross section of ^{60}Co for neutron capture $= 2 \times 10^{-24}$ $\mathrm{cm^2}$

τ_d parameter in the denominator of vanadium SPND transfer function $G_v(s)$ (also the time-constant of vanadium SPND response) $= 315$ s

τ_p parameter in the numerator of vanadium SPND transfer function $G_v(s) = 23.68$ s

T_1 delay corresponding to the time constant of the vanadium SPND transfer function $= 315$ s

T_2 delay corresponding to minimization of the integral of squared error between vanadium SPND transfer function output and delayed flux input $= 217.3$ s

T_3 delay corresponding to maximization of the correlation between vanadium SPND transfer function output and delayed flux input $= 365$ s

Clustering Related Variables

A_t output of SPND A at time t

B_t output of SPND B at time t

$\bar{A}$ average value of output of SPND $A = \frac{1}{N}\sum_{t=1}^{N} A_t$

$\bar{B}$ average value of output of SPND $B = \frac{1}{N}\sum_{t=1}^{N} B_t$

Manuscript received November 04, 2013; revised March 28, 2014; accepted October 27, 2014. Date of publication November 20, 2014; date of current version December 11, 2014. This work was supported by the Board of Research in Nuclear Sciences (BRNS), India.

R. A. Razak and M. N. Belur are with the Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India.

M. Bhushan is with the Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India, on leave at the Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: mbhushan@iitb.ac.in).

A. P. Tiwari is with the Reactor Control Division, Bhabha Atomic Research Centre, Mumbai 400085, India, and also with the Homi Bhabha National Institute, Mumbai 400085, India.

M. G. Kelkar and M. Pramanik are with the Nuclear Power Corporation of India Ltd., Mumbai 400094, India.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNS.2014.2366931

$\mathrm{corr}(A, B)$	correlation coefficient between SPNDs A and B
c_j	centroid of the jth cluster, i.e. a time series with value at time t being the average of the values at time t of the time series of the individual SPNDs in the jth cluster
c_{ji}	ith SPND in jth cluster
$\bar{c}$	centroid of all the SPNDs
$d(A, B)$	(correlation-based) distance between SPNDs A and B defined as $1 - \lvert\mathrm{corr}(A, B)\rvert$
E	sum (across all clusters) of the distances of each SPND from its cluster centroid
k	total number of clusters
$\min_{\mathrm{corr}}$	the minimum of the magnitude of all pair-wise correlations between the SPNDs belonging to the same cluster
n_j	number of SPNDs in the jth cluster
N	number of available observations for an SPND
s_b	sum of *inter*-cluster distance
s_w	sum of *intra*-cluster distance (also equals E)
sep_i	index indicating separability of clusters = $\frac{s_w}{s_b}$

II. Introduction

Self powered neutron detectors (SPNDs) are a class of widely used sensors for measuring neutron flux in a reactor in real-time. SPNDs operate based on the principle of activation by neutron interaction ([1], [2], [3]). An SPND consists of an emitter material that undergoes radioactive decay upon absorption of neutrons leading to the production of electrons. The output of an SPND is a current signal, usually of the order of micro/nano amperes. Some SPNDs provide a prompt response to the variation in neutron flux while others provide a delayed response. The response characteristics depend on the exact process of current generation in the emitter material. In this paper, we consider two types of SPNDs - cobalt and vanadium, which are commonly used in Indian Pressurized Heavy Water Reactors (PHWR). Cobalt SPNDs are of prompt type and thus provide an instantaneous estimate of the flux value. On the other hand, vanadium SPNDs are of delayed type. In spite of this, vanadium SPNDs have several advantages over cobalt SPNDs, such as low burn-up rate, small gamma response and negligible background signal [4].

A typical nuclear reactor can have hundreds of such SPNDs providing vital information for reactor monitoring and control purposes. It may thus be useful to cluster these SPNDs into smaller groups, each containing highly correlated detectors. Once such groups are obtained, they could be used for several purposes. As an illustration, data driven techniques such as Principal Component Analysis (PCA) could be used to develop models relating measurements of SPNDs in a particular group. These models could then be used to identify a faulty SPND and estimate its corresponding true flux in real-time, based on the output signals of the other SPNDs in that group, thereby allowing the reactor operations to continue [5].

It is essential that while obtaining these groups, SPNDs with differing dynamics are considered simultaneously so that information from all the SPNDs can be effectively utilized. Developing techniques for clustering SPNDs with different dynamics is the aim of our current work. Here, by 'dynamics' we mean the speed of the response. In particular, we consider cobalt and vanadium SPNDs which have different dynamics. We propose three data processing techniques to ensure that cobalt and vanadium SPNDs have similar time characteristics so that clusters with highly correlated SPNDs can be obtained. These three techniques are: (i) introducing a pure delay in cobalt SPND data to match vanadium SPND dynamics, (ii) appropriately 'slowing down' cobalt SPND data to match vanadium SPND dynamics, and (iii) 'speeding up' vanadium SPND data to match cobalt SPND dynamics. These procedures are elaborated in Section V in this paper. Based on extensive simulations comparing these dynamics matching methods, it is shown that the second method provides the most meaningful results which is then used to obtain SPND clusters for the data available from a PHWR operating in India. More specifically, we had data corresponding to 42 cobalt and 102 vanadium SPNDs. These SPNDs were clustered after appropriately slowing down the cobalt SPND outputs. Several of the cobalt SPNDS known to be in a particular spatial zone in the reactor were indeed identified as belonging to the same cluster thereby validating our overall approach.

Using correlation based analysis for clustering of SPNDs with different dynamics has not received any attention in the literature. While [6] used correlation between SPNDs as a criterion for validation of SPND signals, they did not perform any clustering of these SPNDs.

The rest of the paper is organized as follows: in Section III, the response characteristics of cobalt and vanadium SPNDs are discussed based on their first principles models. The widely used k-means clustering algorithm, which is also used in our current work, is briefly discussed in Section IV. The various dynamics matching techniques used in this work are discussed in Section V. These methods are compared based on simulations in Section VI and the method which works best is identified. The application of the proposed method to the SPND data available from an Indian PHWR is given in Section VII and conclusions are drawn in Section VIII.

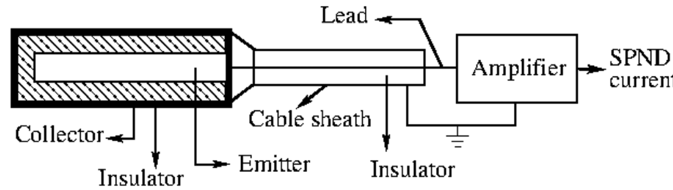

Fig. 1. Basic structure of SPNDs [4].

III. SPND Models

A representative SPND is shown in Fig. 1 [4]. It consists of four main parts- emitter, insulator, sheath (collector) and mineral insulated coaxial cable ([1], [3], [4]).

The materials that can be used for the emitter include rhodium, vanadium, platinum, silver and cobalt [7]. Typically, Magnesium Oxide (MgO) or Aluminium Oxide (Al_2O_3) are used as insulators while Inconel is used as the sheath [4]. The insulator has high resistance of the order of 10^9 to 10^{12} ohms. The SPNDs in the Indian PHWR are typically less than 5 mm in diameter and 20 to 30 cm in length. The insulator in the detector portion is same as that in the cable. It is to be noted that the analysis presented in our work is specific neither to SPND sizes nor to the insulation material, but depends on their prompt/delayed characteristics only.

Detailed information on the working principle of SPNDs can be found in literature ([1], [2], [3], [4], [8]). For completeness, a brief summary is presented here. The emitter undergoes neutron capture to produce radioactive nuclei which undergo β decay to produce more stable nuclei. Electrons are generated during these processes. The electrons which cross the insulation and reach the collector form a current whose path is completed through a cable. The strength of the current is very small (of the order of micro/nano amperes) and it needs to be amplified before further processing. The current generation in SPNDs is due to the following three phenomena:

(a) Neutron capture by emitter atoms directly leads to the generation of current-carrying electrons. When an emitter atom undergoes neutron capture, γ-rays are produced which generate current-carrying electrons due to photo-electric/Compton effect. Electrons are generated immediately after neutron capture and hence this part of the current signal is prompt ([1], [2]).
(b) When emitter atoms undergo neutron capture, radioactive isotopes are produced. These undergo β decay during which high energy β-rays (high energy electrons) are emitted. These β-rays contribute to the current signal of the SPND. Also, the β decay may produce γ-rays which again generate electrons as mentioned above. Since the β decay occurs with a certain half-life, this part of the current represents a delayed response [8].
(c) External γ flux in the reactor also generates current (background noise) in the SPND signal [1].

SPNDs can be classified as prompt SPNDs or delayed SPNDs depending on whether the major component of the signal is prompt or delayed ([3], [8]). Cobalt SPNDs are prompt whereas vanadium SPNDs are delayed. We now describe the dynamic models of cobalt and vanadium SPNDs.

A. Cobalt SPND Model

A brief description of cobalt SPNDs is presented here. More details can be obtained from [1]. The isotope ^{59}Co is the most common and naturally occurring form of cobalt. This isotope undergoes neutron capture to produce ^{60*}Co, which emits γ-rays to produce the ^{60}Co isotope. Next, ^{60}Co undergoes β-decay with a half-life of 5.26 years to produce ^{60}Ni. The isotope ^{60}Co may also undergo neutron capture yielding ^{61}Co, which finally decays into ^{61}Ni with a half-life of 99 minutes.

The mathematical model of cobalt SPND is given by the following set of equations [1]:

$$\begin{aligned}
\frac{d}{dt}N_{59}(t) &= -N_{59}(t)\sigma_{59}\phi(t),\\
\frac{d}{dt}N_{60}(t) &= N_{59}(t)\sigma_{59}\phi(t) - N_{60}(t)(\lambda_{60} + \sigma_{60}\phi(t)),\\
\frac{d}{dt}N_{61}(t) &= N_{60}(t)\sigma_{60}\phi(t) - \lambda_{61}N_{61}(t).
\end{aligned} \tag{1}$$

The output current of the cobalt SPND is the sum of the prompt and delayed components of current, and is given by

$$I(t) = I_{p,tot}(t) + I_d(t), \tag{2}$$

Where

$$\begin{aligned}
I_{p,tot}(t) &= I_{p,\gamma}(t) + I_{p,n}(t),\\
I_{p,n}(t) &= K_{60}N_{60}(t) + K_{61}N_{61}(t) + S_c\phi(t),
\end{aligned}$$

and

$$I_d(t) = I_{60}(t) + I_{61}(t).$$

In the case of a cobalt SPND, the delayed component is a small fraction of the total current and will be ignored in our work when using the cobalt SPND model to generate simulated data. The delayed current component is negligible for a newly installed SPND but increases with time because of the build-up of ^{60}Co and ^{61}Co. Most of the prompt signal is produced by the internally generated γ-rays. In addition, there is also a contribution from externally produced γ flux. These may come from fission in the reactor or neutron capture in other parts of the reactor. While generating our simulation data, the prompt component due to the external γ flux is not considered, since it is likely to be small.

B. Vanadium SPND Model

A first principles model for vanadium SPND has been developed in [4], [9] using the information available in [10] and is described here. The isotope ^{51}V undergoes neutron capture to produce ^{52}V which in turn undergoes β decay. The equations are given below,

$$\begin{aligned}
\frac{d}{dt}N_{51}(t) &= -N_{51}(t)\sigma_{51}\phi(t),\\
\frac{d}{dt}N_{52}(t) &= N_{51}(t)\sigma_{51}\phi(t) - \lambda_{52}N_{52}(t),\\
I_v(t) &= K_{pv}\sigma_{51}N_{51}(t)\phi(t) + K_{gv}\lambda_{52}N_{52}(t).
\end{aligned} \tag{3}$$

Assuming the burn-up of ^{51}V to be negligible [4], i.e. assuming its concentration to be constant, leads to a linear model. We thus get the vanadium SPND transfer function from $\phi(t)$ to $I_v(t)$ as

$$G_v(s) = K\frac{\tau_p s + 1}{\tau_d s + 1}, \tag{4}$$

where $\tau_p = \frac{K_{pv}}{\lambda_{52}(K_{pv}+K_{gv})}$, $\tau_d = 1/\lambda_{52}$ and $K = \sigma_{51}N_{51}(K_{pv} + K_{gv})$. The value of τ_p will vary depending upon the SPND characteristics and reactor operation. In our work, we obtain both τ_p and τ_d from the PHWR trip data for an individual vanadium SPND. The values were found to be $\tau_p = 23.68$ s and $\tau_d = 315$ s respectively [11]. The value of the gain K is obtained as 1.415×10^{-20} A.cm^2.s.n^{-1} based on experiments performed during SPND calibration in the Indian PHWR. However, the value is not important in our work since we will work with correlation, which is insensitive to the value of the gain K. Note that the ability to partially compensate for burn-up and build-up in a SPND is built into the associated amplifier shown in Fig. 1 by providing options for modifying the gain and specifying an offset.

IV. The k-Means Clustering Algorithm

In this section we briefly describe the k-means clustering algorithm and also discuss its implementation.

A. Brief Description

The *k-means* ([12], [13]) is a non-hierarchical clustering technique in which a set of points are partitioned into a specified number of clusters such that the sum of distances of each point from its cluster centroid is minimized. The distance between the points may be specified in a number of ways. In our work, we are going to work with time series of SPNDs. Thus a single point in clustering corresponds to a time-series. We require that those SPNDs which have high correlation (positive or negative) between each other belong to the same cluster. This is motivated by a possible end use of the resulting clusters, namely to develop linear models between SPNDs in a given cluster [5]. Since high correlation between a pair of SPNDs implies the possibility of a strong linear relationship between them, we consider two SPNDs to be close together if the absolute value of the correlation coefficient between them is high. Thus, for the purpose of clustering, we define the distance between SPNDs A and B as,

$$d(A,B) = 1 - |\mathrm{corr}(A,B)|, \tag{5}$$

where

$$\mathrm{corr}(A,B) = \frac{\sum_{t=1}^{N}(A_t - \bar{A})(B_t - \bar{B})}{\sqrt{\sum_{t=1}^{N}(A_t - \bar{A})^2 \sum_{t=1}^{N}(B_t - \bar{B})^2}}.$$

We aim to cluster the SPNDs into k clusters where k is to be suitably chosen.

The input to the k-means algorithm consists of the data based on which the clustering is to be performed and the number of clusters k. Let c_j for $j = 1, 2, \ldots, k$ be the centroid obtained for the j-th cluster consisting of n_j SPNDs, i.e. c_j is a time-series with the value at any time t being the average of the values of the SPNDs in the j^{th} cluster at time t. For each cluster, the sum of distances of the points from the centroid is given by

$$C_j = \sum_{i=1}^{n_j} d(c_{ji}, c_j). \tag{6}$$

The sum of distances of each SPND from its cluster centroid across all the clusters is then given by

$$E = \sum_{j=1}^{k} C_j. \tag{7}$$

The k-means algorithm minimizes E in an iterative manner. Given the initial cluster centroids, each SPND is assigned to one of the clusters such that the distance between the SPND and the cluster centroid is the least among all the clusters. The initial cluster centroids may be user-specified or chosen randomly. After all the SPNDs have been allocated to one of the clusters, the cluster centroids are recomputed and the assignment process is repeated with the new centroids. These iterations are continued till the convergence criterion is met, i.e. $\sum_{j=1}^{k} \|c_j^{(\ell+1)} - c_j^{(\ell)}\| < \epsilon$ where $c_j^{(\ell)}$ is the j-th centroid at the ℓ-th iteration and ϵ is chosen to be a small number.

B. Implementation of the k-Means Algorithm

For the work done in this paper we have used the k-means algorithm available in the Matlab *Statistics* toolbox [14], suitably modified so as to incorporate the distance described by (5). The k-means algorithm initially chooses a random set of cluster centroids. We thus rerun the algorithm fifty times and take the best set of clusters, i.e. the cluster set with the least value of sum of points-to-centroid distances defined by (7). The number of clusters k to be considered will also be unknown for real (plant) data. Hence, in Section VII we propose an approach for selecting a suitable number of clusters for the PHWR data. However, for comparison between various dynamics matching methods using simulated data, we use the true number of clusters k.

It should be noted that while we have used k-means algorithm in our work, any other clustering approach could also have been used. [15] presents a comprehensive review of various clustering methods available in literature. We have used k-means as it is one of the most popular algorithms for clustering.

V. Data Processing for Clustering

A. Clustering Procedure

The measurement data of cobalt and vanadium SPNDs may not be correlated even if they are subjected to the same neutron flux because the response of cobalt SPND is prompt while that of the vanadium SPND is delayed. Some form of dynamics matching is thus required so that both the cobalt and vanadium SPND data can be used together. The three methods which we propose for dynamics matching are as follows:

(a) Using cobalt SPND data delayed by time T (i.e. passed through a transfer function e^{-Ts}) for comparison with the original vanadium SPND data.
(b) Using the cobalt SPND data obtained after passing it through the vanadium SPND transfer function ($G_v(s)$, given by (4)) for comparison with the original vanadium SPND data.
(c) Comparing the original cobalt SPND data with the vanadium SPND data that has been passed through the inverse vanadium SPND transfer function ($G_v^{-1}(s)$).

A generic schematic showing these three methods of transforming SPND data is shown in Fig. 2. Table I lists the corresponding SPNDs I, II and the processing block for each of the three methods. These three methods are discussed in detail in the following sections. In each case, the unprocessed SPND I

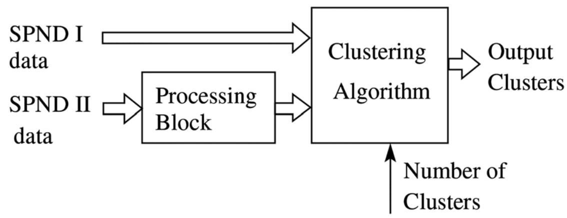

Fig. 2. Schematic of data processing framework for clustering.

TABLE I
DETAILS OF DATA PROCESSING METHODS

Method	SPND I	SPND II	Processing Block
Cobalt SPND signal delayed (pure delay)	Vanadium	Cobalt	e^{-Ts}
Cobalt SPND signal slowed (lag filter)	Vanadium	Cobalt	$G_v(s)$
Vanadium SPND signal speeded-up	Cobalt	Vanadium	$G_v^{-1}(s)$

data and the processed SPND II data is passed to the k-means clustering algorithm.

B. Pure Delay in Cobalt SPND Measurements

In this method, we use a pure delay to delay the cobalt SPND data by T samples i.e. pass the cobalt SPND measurements through a pure delay transfer function e^{-Ts}. The delayed cobalt SPND data is then clustered together with the original vanadium SPND data. The choice of the delay T is made in the following three ways:

- We choose $T = T_1 = 315$ s which is the value of τ_d in (4).
- We consider a unit step change $u(t)$ in input (i.e. neutron flux) to the vanadium SPND transfer function. We then seek the value of delay T_2 at which the error $e(t) = \int_0^\infty (r(t) - u(t - T_2))^2 dt$ is minimized, where $r(t)$ is the corresponding vanadium SPND transfer function output. This is analytically calculated to be $T_2 = 0.69\tau_d = 217.3$ s.
- We seek the value of delay T_3 which maximizes the correlation between $r(t)$ and $u(t-T_3)$. This value is numerically obtained to be $T_3 = 365$ s.

To summarize, we delay the cobalt SPND data in order to compensate for the slow response of the vanadium SPND data. Three delays are considered: $T_1 = 315$ s, $T_2 = 217.3$ s and $T_3 = 365$ s.

C. Cobalt SPND Measurements Passed Through Vanadium SPND Transfer Function

In this method, we pass the cobalt SPND data through the transfer function of vanadium SPND $G_v(s)$. This is used to transform the cobalt SPND data to similar time characteristics as that of vanadium SPND. Since $G_v(s)$ is a lag-compensator, this 'slows down' the cobalt SPND signal by attenuating the high frequency components in the cobalt SPND measurements.

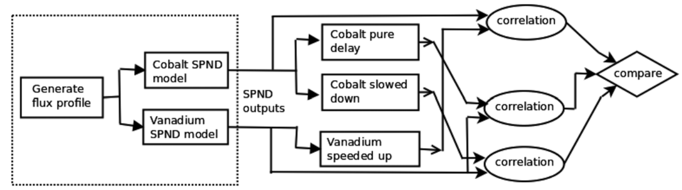

Fig. 3. Simulation process.

D. Vanadium SPND Measurements Passed Through its Inverse Transfer Function

In this method, we 'speed-up' the vanadium SPND data by passing its output through the inverse of its transfer function so that the delayed dynamics of vanadium SPNDs are compensated for. We then use the speeded up vanadium SPND data along with the cobalt SPND data for clustering. Passing vanadium SPND data through the inverse of vanadium SPND transfer function leads to amplification of high frequency components.

VI. COMPARISON OF THE DYNAMICS MATCHING METHODS

The proposed methods for matching the dynamics of vanadium and cobalt SPNDs are now compared based on simulations, that involve generating a flux profile and propagating it through the cobalt and vanadium SPND dynamic models so as to generate the corresponding SPND outputs. These outputs are processed according to the various dynamics matching methods listed in the last section and then compared.

A. Pure Sinusoidal Flux Profile

Pure sinusoidal flux profiles of amplitude 10^{13} n.cm^{-2} s^{-1} and different frequencies were generated. The generated data consisted of 10,000 samples. The time periods of sinusoids were $T_n = 10000/2^n$ s, $n = 1, \ldots, 12$, thus giving 12 profiles. Each profile was passed through cobalt and vanadium SPND models as described in Section III to generate the corresponding output currents. Towards this end, the continuous time state space models were integrated using Runge-Kutta-4 method. The outputs (currents) were then processed according to the various dynamics matching methods proposed in Section V. The correlation between respective datasets is compared to find the method that gives a higher correlation. The process is shown in Fig. 3. For each method of dynamics matching, we obtain a plot of the correlation between the corresponding cobalt and vanadium SPND datasets versus the frequency of input flux. This was repeated for different process noise levels in the SPND models. The noisy states were generated by adding zero mean, white Gaussian noise in the differential equations of the SPND models given in Section III. The standard deviation of the noise in a state equation was taken as a fraction of the peak response of the corresponding noise free state trajectory obtained for a positive step of magnitude 10^{13} n.cm^{-2} s^{-1} in the input flux.

The resulting correlation plots are shown in Figs. 4–9 for the different techniques. In these figures, plots corresponding to different process noise levels are shown in different colors. The process noise level 1 represents the maximum process noise (standard deviation 10%) and noise level 7 represents no process noise. Noise levels from 1 to 6 represent decreasing process noise standard deviation by a factor of 10 at each step.

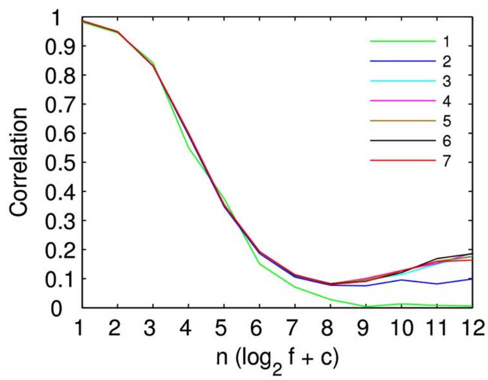

Fig. 4. Correlation versus n for various noise levels: raw cobalt SPND data and raw vanadium SPND data (i.e. without any dynamics matching).

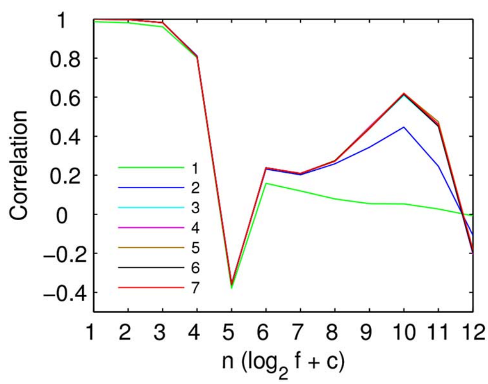

Fig. 5. Correlation versus n for various noise levels: cobalt SPND data delayed by T_1 time units and raw vanadium SPND data.

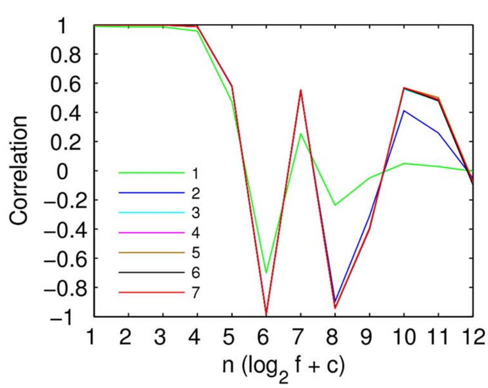

Fig. 6. Correlation versus n for various noise levels: cobalt SPND data delayed by T_2 time units and raw vanadium SPND data.

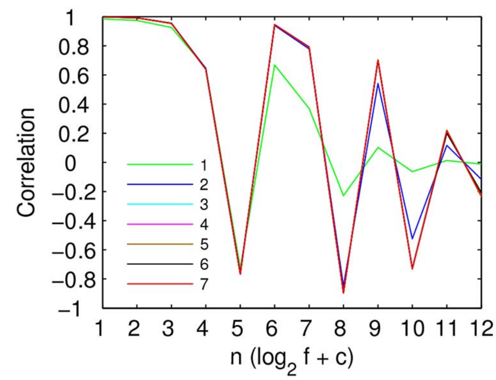

Fig. 7. Correlation versus n for various noise levels: cobalt SPND data delayed by T_3 time units and raw vanadium SPND data.

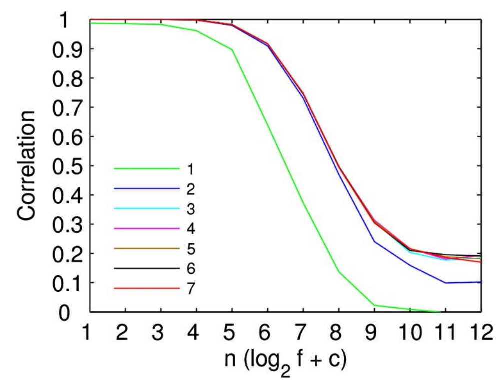

Fig. 8. Correlation versus n for various noise levels: slowed cobalt SPND data (passed through vanadium SPND transfer function) and raw vanadium SPND data.

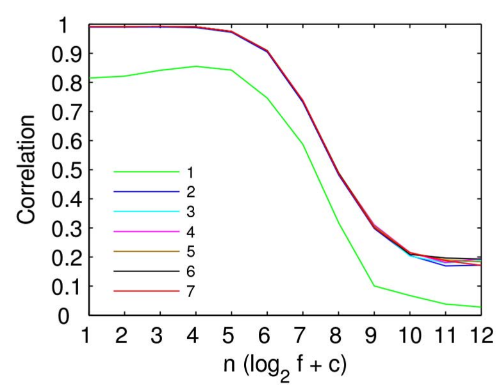

Fig. 9. Correlation versus n for various noise levels: raw cobalt SPND data and speeded-up vanadium SPND data (passed through the inverse of the vanadium SPND transfer function).

The x-axis represents the value n which is a monotonically increasing function of frequency f ($n = \log_2 f + c$, with $c = \log_2 10000$). For the sake of comparison with dynamics matching techniques, the correlations between the raw (unprocessed) cobalt and vanadium SPND measurements are also shown in Fig. 4.

It is expected that as the process noise level is decreased, the curve shifts to the right implying high correlation for a larger range of frequencies. The curve to the extreme right in each Fig. corresponds to the case of zero process noise. It can be seen from the plots that the correlation is higher for larger ranges of frequencies in the case of slowing down cobalt SPND data (Fig. 8) and speeding up vanadium SPND data (Fig. 9). Consider the value of n where the correlation drops to 0.9 for the zero process noise level curve. This value can be observed to be 3 for raw data case (Fig. 4), 3.5 for the case with cobalt SPND signal delayed by time T_1 (Fig. 5), 4.5 for the case with cobalt SPND signal delayed by time T_2 (Fig. 6), 3.5 for the case with cobalt SPND signal delayed by time T_3 (Fig. 7), 6 for the case with cobalt SPND signal slowed down (Fig. 8) and 6 for the case with vanadium SPND signal speeded up (Fig. 9). Comparing the responses shown in Figs. 8 and 9, it is evident that the cobalt SPND signal when slowed down gives higher correlation overall at various noise levels and thus seems to perform better. For the case of pure delay, it can be seen that the curves show random variations and even show negative correlations for some frequencies. This is because after a certain frequency, the

TABLE II
CORRELATIONS FOR VARIOUS METHODS: FLUX PROFILE CONSISTS OF SUM OF TWO SINUSOIDS WITH TIME-PERIODS AS MENTIONED IN THE RESPECTIVE CASES

	Case 1: 1250 s, 19.5 s	**Case 2: 10^4 s, 4.9 s**	**Case 3: 5,000 s, 9.76 s**	**Case 4: 10^4 s, 5000 s**
Raw data	0.4270	0.6988	0.6771	0.9699
Delay T_1	-0.5669	0.6684	0.5499	0.8695
Delay T_2	-0.2989	0.6788	0.5484	0.9065
Delay T_3	-0.6347	0.6648	0.4644	0.8492
Cobalt SPND signal slowed	0.9892	0.9971	0.9969	0.9999
Vanadium SPND signal speeded	0.7208	0.7179	0.8776	0.9955

delay values (T_1, T_2 or T_3) become larger than the time period of the input sinusoid and this changes the sign of the correlation.

B. Mixed Sinusoidal Flux Profile

In this case, we added two sinusoidal profiles, one of which is of high frequency and the other of low frequency, from the sinusoidal profiles listed earlier. The flux profiles thus obtained were passed through the noise free cobalt and vanadium SPND models. The resulting correlations for the various cases are listed in Table II from which it can be seen that slowing the cobalt SPND data leads to highest correlations.

In order to further test the performance of various dynamic matching methods, a sinusoidal flux profile consisting of 10 sinusoids equally spaced in the range 0 to 0.5 Hertz was given as input and the correlations were found to be as: 0.3148 for the raw data case, 0.2951 for the delay T_1 case, 0.2958 for the delay T_2 case, 0.2899 for the delay T_3 case, 0.9753 for the cobalt SPND slowed down case and 0.3326 for the vanadium SPND speeded up case. Once again, slowing the cobalt SPND data leads to much better results than other data processing methods.

C. Clustering Multiple SPNDs

Until now, we considered correlations amongst the processed outputs of a cobalt and a vanadium SPND, each of which is exposed to the same single input flux profile. Since we intend to identify the best data processing method that can be used to extract clusters from a group of cobalt and vanadium SPNDs using data obtained from the PHWR, we now perform the following simulation experiments that involve clustering of multiple SPNDs.

We generate data for 24 SPNDs such that the SPNDs belong to four predefined clusters. Using this data, clustering is performed according to the various dynamics matching methods and compared to find the methods leading to correct identification of the clusters on using k-means clustering algorithm. The data is generated as follows: we generate four different flux profiles uncorrelated to each other. Each of these flux profiles are used to generate six sets of SPND data - two for cobalt and four for vanadium SPNDs. White, zero mean Gaussian process noise with standard deviation of 1% as mentioned in Section VI-A is also added to the SPND models during data generation. Since the four flux profiles are uncorrelated to each other, the SPND data generated from a single profile forms a single cluster. Thus, each true cluster consists of two cobalt and four vanadium SPNDs. Different types of flux profiles were generated as discussed below.

1) *Each flux profile was a random Gaussian sequence (mean 0 and standard deviation 10^{13}) uncorrelated to each other.*
2) *Same as case 1 but with additional random noise added to the fluxes input to the individual SPNDs within a given cluster*: The input flux profile for a given SPND was thus an addition of two components: (i) a random Gaussian sequence (mean 0 and standard deviation 10^{13}) which was the same for each SPND in a given cluster, and (ii) a random Gaussian sequence (mean 0, standard deviation 10^{12}) which was different for each SPND.
3) *Each flux profile consisted of a sinusoidal component and a random Gaussian component along with random noise:* In other words, the flux profiles were same as in case 2 but with additional sinusoidal component added to them. The sinusoidal component consisted of five sinusoids which were of different frequencies for the different flux profiles. Two of the four flux profiles consisted of low frequency sinusoids in the range 0.15-0.20 and 0.25-0.30 respectively in terms of the normalized frequency, which refers to the frequency as a percentage of $F_s/2$ where F_s is the sampling frequency (1 Hertz in our case). The other two profiles consisted of high frequency sinusoids in the range 0.75-0.80 and 0.85-0.90 respectively in terms of normalized frequency. The sinusoidal component was the same for all SPNDs within a given cluster.

In each of the above cases, clustering was performed after processing SPND signals using the techniques discussed in previous sections. It was found that correct clusters were retrieved when the data was transformed by either slowing the cobalt SPND measurements or speeding-up the vanadium SPND responses. For the cases when pure delays were introduced in the cobalt SPND measurements, the clusters were sometimes not correctly identified. In order to further compare the slowing-down cobalt and speeding-up vanadium SPND data processing methods, we additionally calculate,

$$\mathrm{sep}_i := \frac{\text{sum of intra-cluster distance, } s_w}{\text{sum of inter-cluster distance, } s_b} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} d(c_{ji}, c_j)}{\sum_{j=1}^{k} n_j d(c_j, \bar{c})}, \tag{8}$$

with $d(c_{ji}, c_j)$ being the distance of the i^{th} SPND in cluster j from the centroid of cluster j and $d(c_j, \bar{c})$ representing the distance of the centroid of the j^{th} cluster with the overall centroid. Thus, in general a large s_b means that the cluster centres are far apart while a smaller value of s_w indicates that the individual clusters are quite compact. However, neither s_w nor s_b can give a complete picture of the obtained clusters. Ideally, one would like each individual cluster to be compact, i.e. s_w to be small, while being well separated from each other, i.e. s_b to be high. Thus, their ratio sep_i can be used as a measure of separation of the resulting clusters. In particular, a lower value of sep_i can be associated with better clustering performance.

TABLE III
VALUES OF s_w, s_b AND sep_i WITH VARIOUS PROCESSING METHODS

	s_w	s_b	sep_i
Raw data (no processing)	6.5789	4.1938	1.5687
Cobalt SPND slowed down	0.4194	0.5367	0.7814
Vanadium SPND speeded up	12.5931	11.5363	1.0916

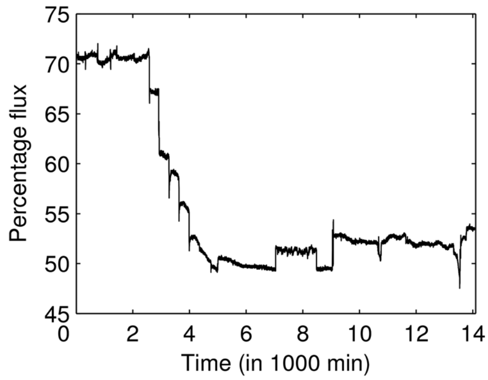

Fig. 10. Data for cobalt-18 SPND (an illustrative cobalt SPND).

The values of s_w, s_b and sep_i for the two methods as well as for the unprocessed raw data are shown in Table. III for the case 3 above, i.e., flux profile consisting of sinusoidal components and random Gaussian component along with random noise. It can be seen that the value of sep_i is lowest for the clusters obtained when cobalt SPND output was slowed down. This suggests that this method would be able to identify the clusters better as compared to the method of speeding up the vanadium SPND output. Based on the above comprehensive simulations, slowing down cobalt SPND output is identified as the method to be used for dynamics matching for clustering the SPND data available from the PHWR.

VII. Application to Data From Indian PHWR

We now perform clustering on SPND data available from the Indian PHWR that consists of a cylindrical core with 42 cobalt and 102 vanadium SPNDs. The reactor core is divided into fourteen control zones with each zone containing three cobalt SPNDs. Data from the SPNDs sampled at an interval of one minute are available for a period of 10 days. The 144 SPNDs are labeled using serial numbers from 1 to 144. The numbers 1 to 42 represent cobalt SPNDs and 43 to 144 represent the vanadium SPNDs. Data for 15 of the vanadium SPNDs were temporarily unavailable for the duration under consideration and therefore those SPNDs have been neglected in the following work.

Plots of the raw data of one of the cobalt SPNDs and the corresponding transformed data obtaining by passing the raw data through the vanadium SPND transfer function are shown in Figs. 10 and 11 respectively. It is seen from these plots that as expected, some attenuation of high frequency components has occurred due to this transformation. In order to facilitate a better comparison, a plot of the data of a randomly chosen vanadium SPND is also shown in Fig. 12. In this figure, the regions appearing shaded correspond to very high frequency oscillations in the detector measurements and are not actually shaded. As

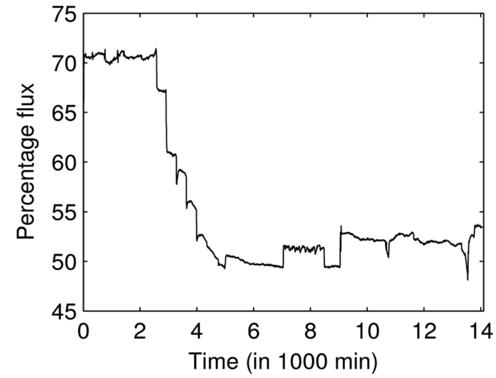

Fig. 11. Cobalt-18 SPND data passed through vanadium SPND transfer function.

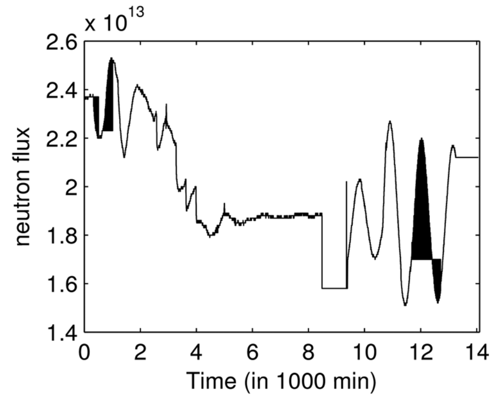

Fig. 12. Data for vanadium-46 SPND (an illustrative vanadium SPND).

mentioned in Section VI, we will slow down the cobalt SPND data for the purpose of clustering. It can be noted from Figs. 10 and 12 that the cobalt and vanadium SPND measurements from the plant are available in different units (have different scaling). However, this is not important for us, as we are using correlation based distance for clustering.

Choosing number of clusters: We considered three criteria for deciding the number of clusters k. Towards this end, we performed clustering for various values of k from 2 to 50. In each case, we computed the following three quantities:

1) For each cluster, $1 - \min_{\mathrm{corr}}$ was calculated. This quantity gives the maximum pairwise spread within each cluster. Then, the average value of $1 - \min_{\mathrm{corr}}$ is computed over all the clusters.
2) The sum of the point-to-centroid distances within each cluster is calculated and is summed over all the clusters, i.e. E (7) is computed.
3) The ratio sep_i (8) is computed.

The above quantities are plotted with respect to the number of clusters in Fig. 13 for the cobalt SPND slowed down method. As expected, these quantities generally decrease with increasing number of clusters. It is seen that there is no appreciable decrease in any of the three quantities on increasing the number of clusters beyond 15. Hence we choose the number of clusters $k = 15$.

Clustering Results: The k-means algorithm in Matlab version R2008b is then applied to obtain 15 clusters. The implementation took 55.98 seconds (for 50 reruns with different initial-cluster choices) when executed on a 4 GB RAM Linux ma-

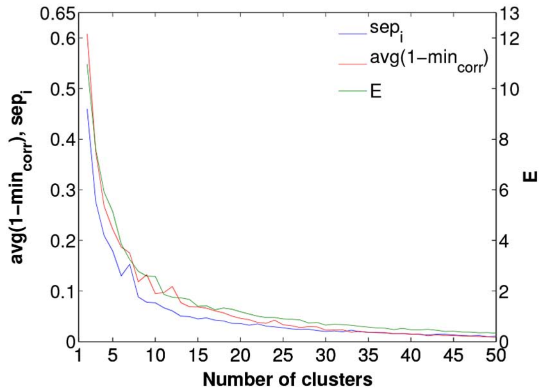

Fig. 13. Variation of avg(1-$\min_{\rm corr}$), E, and sep_i with number of clusters.

TABLE IV
CLUSTERS OBTAINED FOR THE INDIAN PHWR

Cluster No.	SPNDs
1	1, 3, 8, 10, 15, 17, 22, 24, 29, 31, 36, 38
2	2, 9, 16, 23, 30, 37
3	4, 11, 18, 25, 32, 39, 49, 139
4	5, 12, 19, 26, 33, 40
5	6, 7 ,13, 14, 20, 21, 27, 28, 34, 35, 41, 42
6	43, 48, 55, 59, 60, 65, 80, 81, 89, 104, 105, 115, 138
7	44, 45, 46, 47, 57, 67, 68, 75, 76, 85, 90, 91, 116, 117, 118, 137
8	50, 53, 54, 121, 140, 143, 144
9	51, 52, 73, 74, 99, 100, 101, 109, 110, 120, 123, 124, 141, 142
10	56, 58, 66, 92, 93, 107
11	61, 62, 71, 82, 83, 84, 106
12	63, 64, 77, 78, 79, 86, 87, 88, 113, 114
13	69, 94, 95, 108, 119
14	72
15	96, 97, 98, 102, 111, 122

chine with Fedora FC 15 as the operating system. An additional 4.33 seconds were spent on the preprocessing step i.e. the dynamics matching step. The resulting clusters had sep_i value of 0.6043 and are listed in Table IV. It can be seen that while most of the clusters are pure clusters consisting of only cobalt (SPND numbers 1-42) or vanadium (SPND numbers 43-144) SPNDs, one of the clusters (cluster number 3) is a mixed one containing six cobalt and two vanadium SPNDs. It is also noted that while several cobalt sensors located in the same spatial zone get clustered together, the process of combining SPNDs of different dynamics reveals high correlation between SPNDs across multiple zones as well.

VIII. CONCLUSION

In this paper, we simultaneously cluster cobalt and vanadium SPNDs in a nuclear reactor into smaller groups. Towards this end, we explore three methods for dynamics matching to ensure that the cobalt and vanadium SPNDs have same time characteristics before clustering is performed. These methods are: delaying cobalt SPND data, slowing cobalt SPND data and speeding up vanadium SPND data. It was found by extensive simulations that the method of slowing the cobalt SPND data resulted in the best clustering performance. This is reasonable since this method involves a lag-compensator (the vanadium SPND transfer function), which attenuates high-frequency components including noise. This method was then used to cluster SPNDs in an Indian pressurized heavy water reactor. The resulting clusters can be used for building data-driven models amongst the SPNDs. These models can then be used for a variety of tasks such as data reconciliation and fault detection and diagnosis of SPNDs [5].

REFERENCES

[1] J. A. Sovka, “Response of cobalt neutron flux detectors,” Atomic Energy of Canada Limited, Chalk River, ON, Canada, Tech. Rep, 1969, AECL-3368.

[2] M. N. Agu and H. Petitcolas, “γ-compensated self-powered detectors for reactor instrumentation and control,” *Meas. Science Technol.*, vol. 1, no. 10, pp. 1047–1051, Oct. 1990.

[3] S. O. Yusuf and D. K. Wehe, “Analog and digital compensation techniques for delayed self-powered neutron detectors,” *Nucl. Sci. Eng.*, vol. 106, no. 4, pp. 399–408, Dec. 1990.

[4] A. K. Mishra, S. R. Shimjith, T. U. Bhatt, and A. P. Tiwari, “Dynamic compensation of vanadium self powered neutron detectors for use in reactor control,” *IEEE Trans. Nucl. Sci.*, vol. 60, no. 1, pp. 310–318, Feb. 2013.

[5] R. A. Razak, M. Bhushan, M. N. Belur, A. P. Tiwari, M. G. Kelkar, and M. Pramanik, “Data reconciliation and gross error analysis of self powered neutron detectors: Comparison of PCA and IPCA based models,” *Int. J. Adv. Eng. Sci. Appl. Math.*, vol. 4, no. 1–2, pp. 91–115, Mar.–Jun. 2012.

[6] F. Adorjan and T. Morita, “A correlation-based signal validation method for fixed in-core detectors,” *Nucl. Technol.*, vol. 118, no. 3, pp. 264–275, Jun. 1997.

[7] R. B. Shields, “A platinum in-core flux detector,” *IEEE Trans. Nucl. Sci.*, vol. NS-20, no. 1, pp. 603–608, Feb. 1973.

[8] L. A. Banda and B. I. Nappi, “Dynamic compensation of rhodium self-powered neutron detectors,” *IEEE Trans. Nucl. Sci.*, vol. NS-23, no. 1, pp. 311–316, Feb. 1976.

[9] S. Krishnan, “Fault diagnosis of self powered neutron detectors: Model based observers for input estimation,” M.Tech. thesis, Dept . of Elect. Eng., Indian Inst. of Technol. Bombay, Mumbai, India, 2011.

[10] C. M. Lederer, J. M. Hollander, and I. Perlman, *Table of Isotopes*. New York, NY, USA: Wiley, 1967.

[11] R. A. Razak, “Data based techniques for fault diagnosis of self powered neutron detectors,” M.Tech. thesis, Dept of Elect. Eng., Indian Inst. of Technol. Bombay, Mumbai, India, 2012.

[12] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*, ser. Advanced Reference Series. Englewood Cliffs, NJ, USA: Prentice-Hall, 1988.

[13] H. Spath, *Cluster Analysis Algorithms for Data Reduction and Classification of Objects*, ser. Computers and their Applications. Chichester, U.K.: E. Horwood, 1980.

[14] “Mathworks,” Statistics toolbox documentation: Mathworks [Online]. Available: www.mathworks.com/help/toolbox/stats, 2008

[15] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: A review,” *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.