Write up related to Non- stationary Version of Continuous Density Hidden Markov Model for PD Pattern Recognition

- **Methodology for Hidden Markov Model Formulation**

  - **Choice of HMM Topology**

Fundamentally, two types of models are utilized: 1. Ergodic and 2. Non-ergodic.

- Ergodic type (also called the fully connected HMM) is a model wherein every state can be reached (in possibly a single step or at least in a finite number of steps) from every other state of the model.
- **Non-ergodic type (left-right model)** is a model in which **as time increases the state index MAY ALSO increase** (the states proceed from left to right)
- Some of the major research studies carried out by researchers such as Bartnikas et al, Satish et al [1], [2] have found that invariably 4 (or) 5 state (hidden state) representation is a sufficient description to be considered as appropriate for pattern recognition studies.

  - **Basis for Choice of Hidden State for Left- Right HMM utilized in MATLAB Coding**

- PD patterns, as a first approximation is considered comprising two zones namely "PD zone" and "background".
- **Overall, typically in the patterns considered for the studies, the states are apparently found to be 4 (or) 5 states** [1], [2]. Figure 1 shows a case as an example for better understanding. Figure 2 depicts the HMM topology utilized in this study.

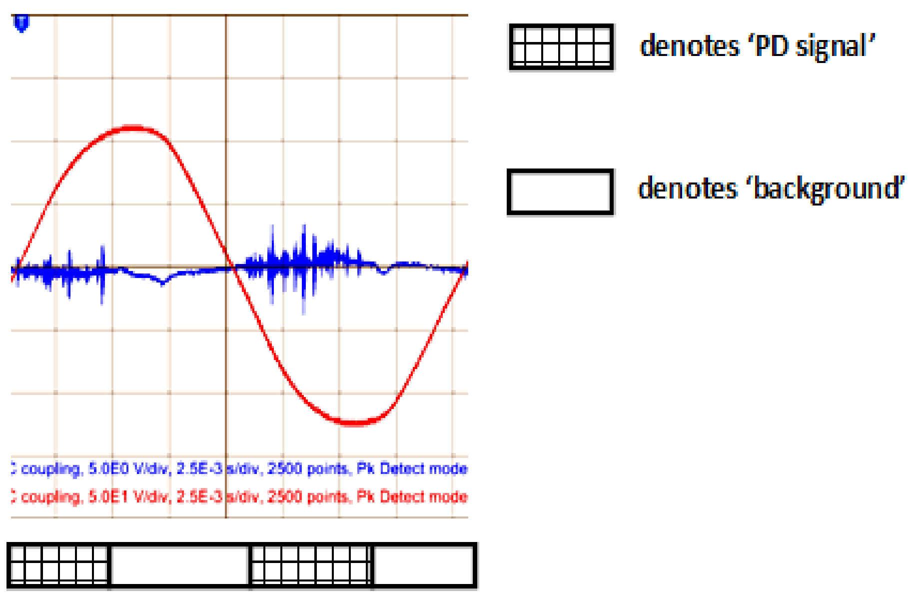

Figure 1: Schematic for Hidden State Labels for obtaining Observation Sequence Matrix

- **Typical Left- Right HMM utilized in MATLAB Coding**

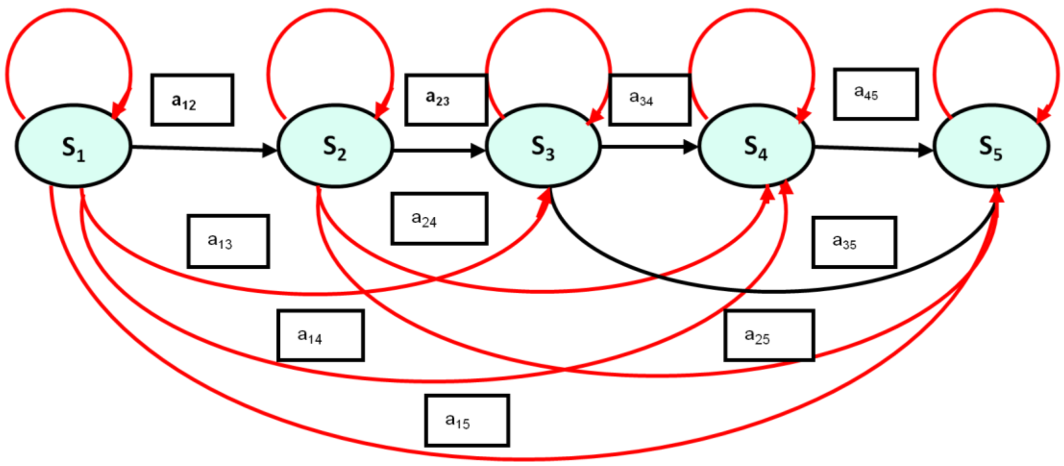

Figure 2: Non-ergodic (Left- Right) HMM Topology utilized for PD Pattern Discrimination

- Model denoted by $\lambda = (A, B, \Pi)$
  where 'A' is State Transition matrix
  'B' is Observation Probability Density
  'Π' is Initial State Matrix

- **Methodology for Formulation of Clustering Algorithm for PD Signature Patterns related to Internal and Surface Discharge**

- **Data Extraction Stage**

- Data obtained for Internal and Surface discharge patterns from PDBase® software is extracted in the form of phi-q-n signature patterns.
- This dataset is further preprocessed into phase windows with each phase window having data related to $\phi - q - n$ for every 10°. Hence, a total of 36 sets of 3 tuple vectors are obtained for ever sinusoidal time sequence.
- Coding is developed for extracting this $\phi - q - n$ (10°) datasets for several PD training patterns. In this case a total of 20 sets (each) of internal and surface discharge patterns are extracted as $\phi - q - n$ (10°) pattern sequence.
- Each sample of PD pulse patterns will have a 36 window sequence in each row. For example the sequence will be of the form:
  $\phi_1 - q_1 - n_1$; $\phi_2 - q_2 - n_2$; $\phi_2 - q_2 - n_2$; . . . . . . ; $\phi_{36} - q_{36} - n_{36}$

- **Clustering Stage**

- Both labeled (Learning Vector Quantization) and unlabelled (K-means) clustering algorithms have been taken up for clustering the preprocessed data obtained as $\phi - q - n$ signature patterns.

- In this preliminary study, a total of 12 clustering sets were obtained from a total of 20 sets provided for training.
- Hence, the output of the clustering would result in a matrix of size 12 x 108 (12 rows of clustered data with 108 columns i.e. 36 windows comprising 3 features of ϕ –q – n for each of the 36 windows). Figure 3 shows the methodology adopted during clustering.

1	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
2	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
3	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
4	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
5	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
6	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
7	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
8	31	0	0	41	0	0	51	0	0	67	0.215625	1	71	0	0	81	0	0	91	0
9	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
10	31	0	0	41	0	0	56	0.290625	1	61	0	0	71	0	0	81	0	0	91	0
11	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
12	31	0	0	41	0	0	60	0.1875	1	61	0	0	71	0	0	81	0	0	91	0
13	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
14	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
15	31	0	0	41	0	0	51	0	0	61	0	0	71	0	0	81	0	0	91	0
16	31	0	0	41	0	0	51	0	0	67	0.525	1	75	0.1125	3	81	0	0	91	0
17	31	0	0	45	0.087197	1	51	0	0	61	0	0	71	0	0	85	0.384375	1	91	0
18	31	0	0	41	0	0	51	0	0	68	0.225	1	77	0.103125	1	81	0	0	91	0
19	31	0	0	41	0	0	51	0	0	61	0	0	78	0.478125	1	81	0	0	91	0
20	31	0	0	41	0	0	51	0	0	62	0.665625	1	71	0	0	84	0.13125	1	91	0

Figure 3: Snapshot of the Clustering Methodology utilized during PD Pattern Discrimination

- The cluster numbers allotted by the algorithm pertaining each of the 20 datasets presented to the algorithm for training and clustering is shown in Figure 4. The clusters automatically allotted the 12 cluster as follows:
  - Cluster Label 1: Training set no. 1,9
  - Cluster Label 2: Training set no. 3
  - Cluster Label 3: Training set no. 5, 12
  - Cluster Label 4: Training set no. 6
  - Cluster Label 5: Training set no. 8
  - Cluster Label 6: Training set no. 10
  - Cluster Label 7: Training set no. 2, 7, 11, 14, 17
  - Cluster Label 8: Training set no. 13
  - Cluster Label 9: Training set no. 15
  - Cluster Label 10: Training set no. 16
  - Cluster Label 11: Training set no. 18, 19
  - Clustering Label 12: Training set no. 20

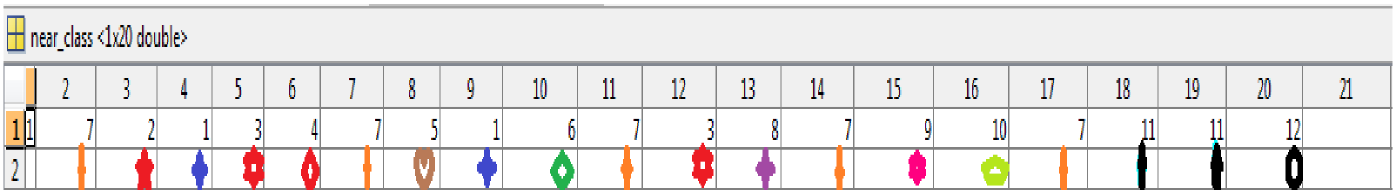

Figure 4: Snapshot of the random allotment of clusters and their respective numbers

- **Implementation of non- stationary version of CDHMM for PD Signature Patterns Analysis related to Internal and Surface Discharge**
  - HMMs are classified based on the process of obtaining the density estimates of the observations as Discrete Density Hidden Markov Model (DDHMM) and Continuous Density Hidden Markov Model (CDHMM).

- **Discrete Density Hidden Markov Model (DDHMM):**
  - A HMM comprises 'N' states wherein a unique type of physical significance may be attributed to it. In a Markov process, a new state is given by $s_t \in \{1,2,...,N\}$ in steps of $\mathrm{t} = 1,2,...,\mathrm{T}$.
  - The initial state distribution matrix is labeled 'Π' wherein $\prod = \{\pi_k\}, \pi_k = P_r(s_1 = k)$.
  - Computation of the state transition probability matrix 'A' involves calculation of $A = \{a_{km}\}, a_{km} = P_r(s_{t+1} = m \mid s_t = k)$ for $\mathrm{i} = 1,2,..., \mathrm{N}$.
  - The resulting state sequence is denoted by $S = \{s_t, t = 1,2,...T\}$ and the probability obtained is given by $P_r(S \mid A, \prod) = \pi_{s1} \prod_{t=2}^{T} a_{s_{t-1}s_t}$.
  - The emission or output vector is described by a sequence $O = \{o_1, o_2, ..., o_T\}$ where $o_t \in \Re^M$ with M dimension for the 'O' observation sequences.
  - The output vector is produced according to the probability distribution called the observation probability distribution 'B' given by $B = \{b_m(o_t)\}$ where $b_m(o_t) = P(o_k = k \mid s_t = m)$.

- **Continuous Density Hidden Markov Model (CDHMM):**

  - If the emission probability distributions are described as a Gaussian mixture model with a conditional mean vector μ_m and a covariance matrix V_m obtained for density related to 'm^{th}' state, the density is described as $N(\mu_m, V_m)$.

- **Difference between Stationary and Non- stationary Version of HMM:**

  - CDHMM described above, models the transition probabilities from't' to 't +1' and the transition is dependent on the state at 't' only. Though this may provide classification/ recognition capability yet the dynamic behaviour of the PD patterns may not always provide satisfactory results.
  - To enhance the dynamic behaviour of the model a non-stationary component is added into the model which aims at capturing the time-dependence in the state transition probabilities [3], [4].

- The state transition matrix is now generated to have T-1 number of matrices and is modified to A_t where $t = 1,2,\ldots,T-1$.

- **Training Phase of Non-stationary CDHMM for PD Pattern Discrimination**
  - This phase involves obtaining the state optimized likelihood function (using the maximum likelihood algorithm) pertaining to the parameters A, Π, B related to a class (category) of PD source.
  - Maximization of the state optimized likelihood for each training sequence of a set of observations is obtained by utilizing the forward Viterbi algorithm. Figure 5 shows the sequence for the implementation of non-stationary CDHMM for implementation.

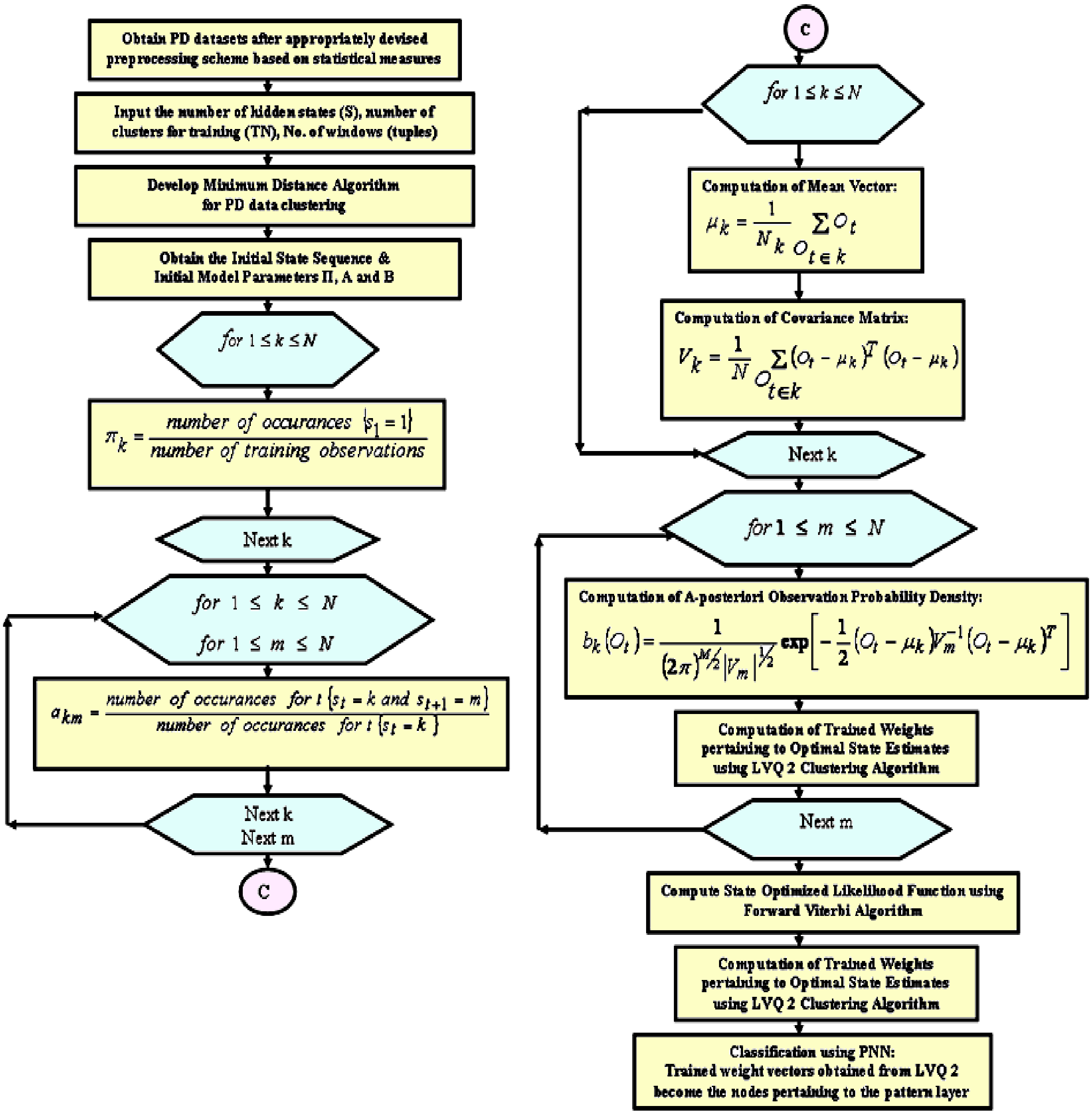

Figure 5: Flow Chart for Implementation of Non-stationary Version of CDHMM

- **Observations and Analysis**

- **Role of Hidden State Transition Labels during discrimination of Internal and Surface Discharge PD Signature Patterns**
    - It is evident from studies that for chosen hidden states (4) that subtle change are observed during the training phase of the non-stationary version of CDHMM considered for studies. Table 1 and Table 2 shows the comparison of the state transition labels for internal and surface discharges.
    - For a 20 training datasets considered in this study it was observed that the optimum likelihood estimates for the hidden states were obtained using the Viterbi Algorithm within 4 to 6 iterations.
    - Another vital aspect of this methodology is on the possibility to relate the physics of the discharge and the possible changes in the discharge patterns due to the effect of space charge (memory propagation effect during discharge). This factor is observed In the case of both internal and surface discharges as highlighted in Table 1 and Table 2.

Table 1: Optimized State Transition Labels for Internal Discharge Signature Patterns

Comparison and Analysis of the Role of State Transition Matrix due to Internal and Surface Discharge Patterns

A. State Transition Matrix for Internal Discharge Patterns (12 Clusters)

PHI-Q- N representation with Learning Vector Quantization (LVQ) Clustering: (comprising 36 tuple vectors); State Labels taken: 4

Cluster No.	Optimized State Transition Matrix Labels																																			
1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4
2	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4
3	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4
4	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4
5	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4
6	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4
7	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4
8	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4
9	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
10	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4
11	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4
12	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4

Table 2: Optimized State Transition Labels for Surface Discharge Signature Patterns

B. STATE TRANSITION MATRIX FOR SURFACE DISCHARGE PATTERNS (12 CLUSTERS)

PHI-Q- N representation with Learning Vector Quantization (LVQ) Clustering: (comprising 36 tuple vectors); State Labels taken: 4

Cluster No.	Optimized State Transition Matrix Labels																																			
1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
2	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
3	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
4	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
5	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
6	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
7	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
8	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
9	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
10	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
11	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
12	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4

- **Role of Optimized Observation Probability Density Estimates (B Matrix) during discrimination of Internal and Surface Discharge PD Signature Patterns**

  - It is evident from Figure 6 and Figure 7 that perceptible difference in the optimized probability density estimates is made out during discrimination of internal and surface discharge patterns.
  - Further, this aspect is reiterated from Table 1 and Table 2 on the effect of space charge on the transition on state labels as highlighted in "red" for a few specific clusters.
  - The shift in the distribution is ascribed to the variations in the state labels.
  - It is observed that for internal discharges the first 7 windows comprise state label '1', the next 10 windows relate to state label '2', while the next 13 windows is attributed to label '3' and the balance 11 windows relate to label '4'.
  - In the case of surface discharges the first 6- 8 windows relate to state label '1', next 7-8 labels is ascribed to label '2', subsequent 13 windows for label '3' and the remaining 10 windows attributed to label '4'.
  - This clearly indicates the difference in signature patterns of internal and surface discharges and the role played by memory propagation effect during dynamics of signature variations within each source of PD.
  - Figure 8 shows the distribution of probability estimates for datasets obtained for ϕ –q patterns. These patterns are obtained with a phase window of 360 (each window comprises discharge magnitude for 1° window) totally to 720 dataset for each sample of a row.

- It is evident that the probability density estimates are more clearly described and more perceptible in the case of internal discharge than the plot indicated in Figure 6.

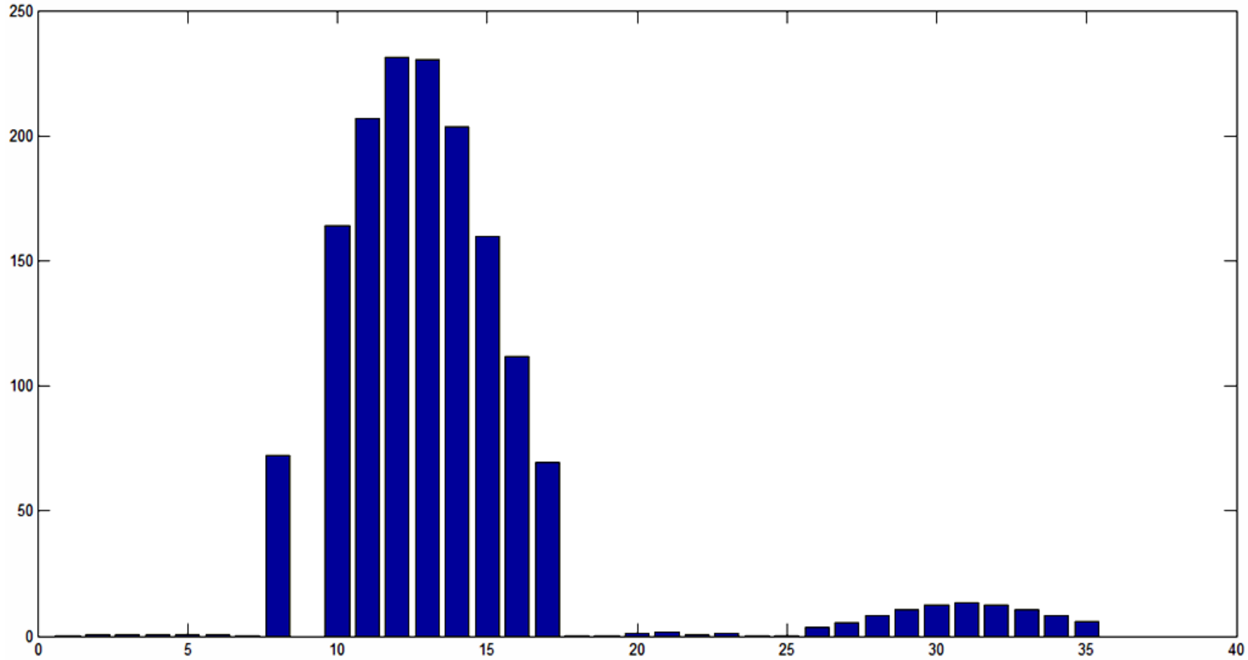

Figure 6: Histogram indicating variations in Probability Density Estimates for Internal Discharges ϕ –q – n (10°)

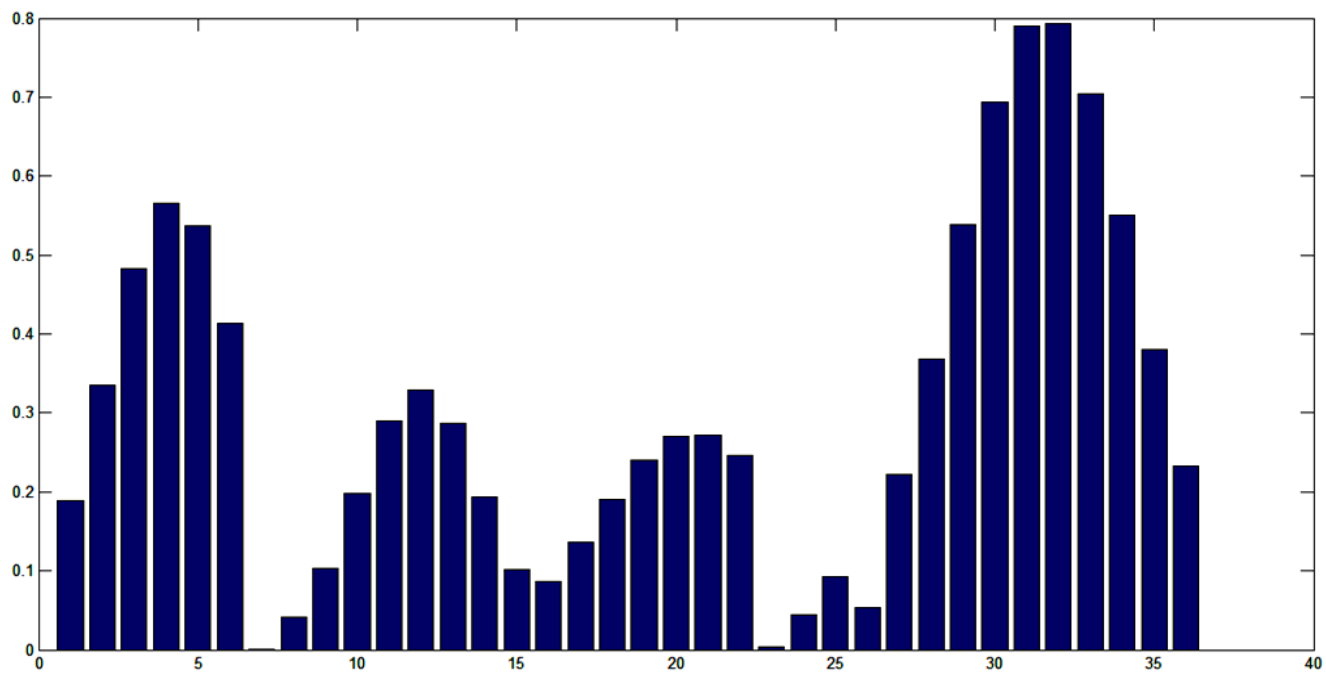

Figure 7: Histogram indicating variations in Probability Density Estimates for Surface Discharges ϕ –q – n (10°)

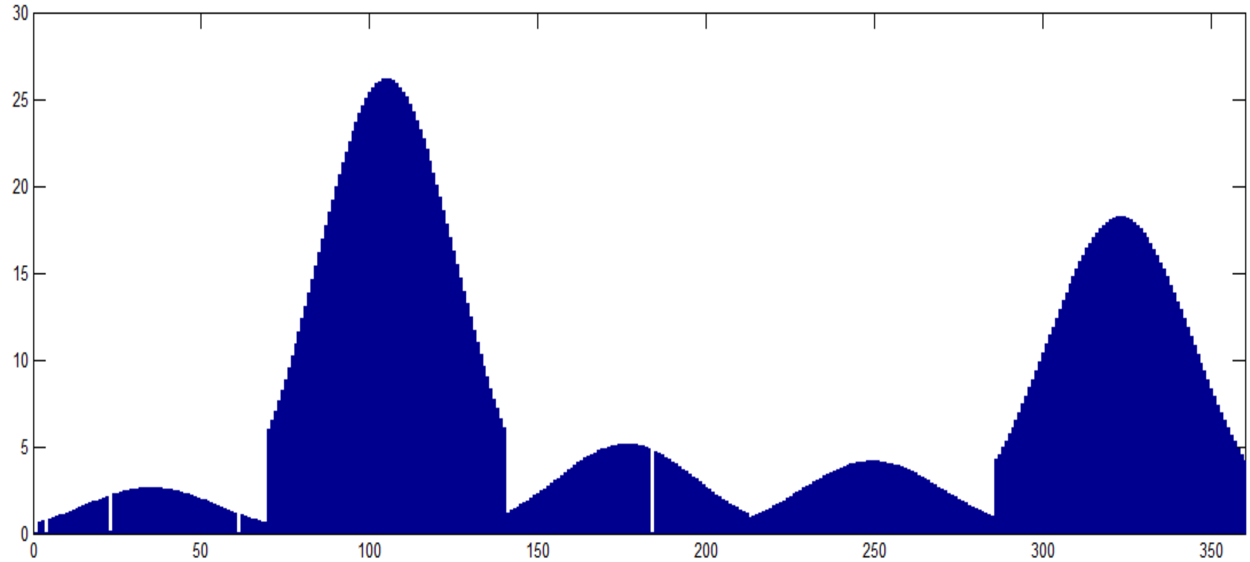

Figure 8: Histogram indicating variations in Probability Density Estimates for Internal Discharges $\phi - q$ (360 phase windows i.e. 1° per phase window)

Major References

[1] L. Satish and B. I. Gururaj, “Use of Hidden Markov Models for Partial Discharge Pattern Classification”, *IEEE Transactions on Electrical Insulation*, Vol. 28, No. 2, pp. 172- 181, April 1993.

[2] T. K. Abdel-Galil, Y. G. Hegazy, M. M. A. Salama and R. Bartnikas, “Partial Discharge Pulse Pattern Recognition Using Hidden Markov Models’, *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 11, No. 4, pp. 715- 723, August 2004.

[3] Yaman Barlas and Korhan Kanar, “A Dynamic Pattern-oriented Test for Model Validation”, *Proceedings of 4th Systems Science European Congress*, Valencia, Spain, pp. 269-286, 1999.

[4] Bongkee Sin and H. Jin Kim, “Non-stationary Hidden Markov Model”, *Signal Processing*, 46, pp. 31-46, 1995.