Fuzzy mathematical model for the analysis of geomagnetic field data

M. Sridharan

Magnetic Observatory, Indian Institute of Geomagnetism, University Campus, Pondicherry-605 014, India

(Received March 2, 2008; Revised February 14, 2009; Accepted June 30, 2009; Online published November 30, 2009)

The Indian network of magnetometers provides an opportunity to examine the pattern of geomagnetic field variations during magnetic storms. In this study, fuzzy transitive closure analysis, which is a powerful technique for pattern recognition, has been employed. The pattern of variation differs at the nonequatorial, equatorial and the observatory situated nearer to the geomagnetic S_q focus. The results of the analysis are compared with those of classical cluster analysis. The comparison confirms the validity of applying this model for the analysis of geomagnetic storms. The superiority of fuzzy concepts over the conventional method and the analytical techniques are presented here.

Key words: Fuzzy logic, transitive closure, geomagnetic storm, geomagnetic S_q, EEJ current.

1. Introduction

Dynamic processes on the Sun deliver plasma of charged particles, principally protons, electrons, and the associated fields, to the Earth's environment, causing geomagnetic disturbances at the Earth's surface that have been named geomagnetic storms. The solar effect on the magnetic perturbation that occurs during geomagnetic storms has been studied by many scientists (e.g., Sugiura, 1953; Rastogi *et al*., 1964; Srivastava *et al*., 1999). India has a unique set of geomagnetic observatories spanning the magnetic equator and the S_q focus in a region of the world where the geomagnetic and geographic meridian planes are least separated. A number of geomagnetic storms have been analyzed in terms of their (1) solar flare effects and sudden commencement (Rastogi *et al*., 1997), (2) the H amplitude of sudden commencement at Sabhawala (Jain, 1978), (3) spectral characteristics (Chandra *et al*., 1995), (4) effects on the ionosphere total electron content and evidence of electro-jet control (Jain, 1978), (5) multi-dimensional scaling (Sridharan and Ramasamy, 2002), (6) extreme behavior (magnetic storm of 1–2 September 1859; Tsurutani *et al*., 2002), and (6) fractal behavior (29–31 October 2003; Sridharan and Gururajan, 2006). Although no two storms are identical, most storms have certain features in common (Parkinson, 1983). Valdivia *et al*. (1996) constructed a nonlinear predictive model for the prediction of magnetic storms based on the D_{st} data alone (not utilizing the data of the interplanetary magnetic field) by applying the time delay embedding technique and phase space reconstructions method. Although the magnetic storm onset cannot be predicted in the absence of the solar wind input, this model can quantitatively describe the subsequent evolution of a storm that can reach a threshold value (-50 nT). The methodology is that 50 hours of sample predictions with iteration and updated every hour can form a sequence of iterated prediction. Time delay and phase space reconstruction methods have been applied to study the fractal behavior of the geomagnetic storm on 29–31 October 2003 (Sridharan and Gururajan, 2006). Valdivia *et al*. (1999) studied the spatiotemporal activity of magnetic storms. The database contains the horizontal component of the magnetic field perturbation at six midlatitude magnetometer stations during the first 6 months of 1979. Time delays adjusted for a spatiotemporal system have been applied to reveal the dynamical properties of the ring current, including the coupling to the solar parameters, and to predict the evolution of the longitudinal profile of magnetic storms. Vassiliadis *et al*. (1999) analyzed the four stages of a magnetic storm, namely, the initial growth, the main and recovery phases, and the storm-substorm relationships. Non-linear, second-order autoregressive moving average models are fit for the data covering the interval of January 1 to June 30, 1979, at a 5-min resolution.

In this paper, variations in geomagnetic storm time ranges for the sensitive horizontal (H) component of the earth's magnetic field recorded at the geomagnetic observatories in India are analyzed by applying a fuzzy transitive closure algorithm. The data of the recently constructed observatories at Pondicherry and Tirunelveli are considered for this analysis.

2. Superiority of Fuzzy-Based Clustering

The purpose of clustering is to distill natural groupings of data from a large data set, thereby producing concise representation of a system's behavior. The fuzzy clustering technique is complementary to earlier classical techniques. It incorporates an iterative approach to data analysis that allows the design of complex systems using a higher level of abstraction, including subjective concepts such as "good", "better", and "little bit satisfied", among others, that can be mapped into exact numeric ranges. Fuzzy techniques enable a researcher to lump together those data points that populate some multidimensional space into a specific number of different clusters. For a given set X of data, the problem of clustering in X is to find several cluster centers that can properly characterize relevant classes of X. In classical cluster analysis, these classes are required to form a partition of X, such that the degree of association is strong for data within blocks of the partition and weak for data in different classes. This requirement is beyond the reach of many applications when the data are chaotic (such as geomagnetic storm time data) in nature, and so it is necessary to replace it with a weaker requirement of a fuzzy partition. The distance of an input sample to the centre of the cluster to which the input sample belongs is used as a criterion to measure the cluster compactness. The compactness and separation validity are more accurate and can produce a better segmentation result. Unknown parameters between one class of data and the other classes can be estimated by a fuzzy model in a naturally consistent way.

Fuzzy logic uses three simple steps:

1. Fuzzification—conversion of numeric data in a real-world domain to fuzzy numbers in the fuzzy domain.
2. Aggregation—computation of fuzzy numbers (all between 0.0 and 1.0) in the fuzzy domain.
3. Defuzzification—conversion of the obtained fuzzy numbers back to the numeric data in the real-world domain.

Advantages of using fuzzy concepts:

1. No need to have prior knowledge about the relationships of data.
2. Modeling of complex, non-linear problems.
3. Handling of any kind of information (numeric, linguistic, logical, etc.).
4. Management of imprecise, partial, vague, or imperfect information.
5. Fast computation using fuzzy number operations.
6. Natural language processing/programming capability.

If fuzzy logic is applied for the data analysis, then the data should be converted into fuzzy numbers (between 0 and 1). To estimate subjectively the resemblance between pairs of data points, we adopt the convention of arranging data for numerical classification in the form of matrix. Each entry $[ij]$ in such a matrix is the score of the proximity relation (subjective similarity) between data points i and j. It should be noted that numerical values in the proximity matrix are only quantitatively descriptive numbers whose significance cannot be evaluated by conventional statistical techniques and thus are determined subjectively. The proximity relation is not necessarily transitive. We must utilize the theory of inexact matrices in order to formulate a transitive closure structure that will enable us to separate the data set into mutually exclusive clusters which are, in essence, equivalent in class. Many different methods are used for clustering, and their selection depends primarily on the specific application. In all other clustering methods, proximity relations defined by arbitrary similarity measures are not necessarily similarity relations. In our method, cluster variables are classified based on the Minkowski formula; other methods are obvious and direct. This method requires accuracy and precession. An α-cut or α-level set, $\alpha\varepsilon(0, 1]$, of a fuzzy relation R is defined as the crisp binary relation R_α. As α runs through $(0, 1]$, α-cuts of R form a nested sequence of crisp relations such that whenever $\alpha_1 \geq \alpha_2$, R_{α_1} is contained in R_{α_2}; that is, $R_{\alpha_1} \leq R_{\alpha_2}$. The scalar '$\alpha$' is an indication of the validity of each clustering in the interval [0, 1].

Here, we report, for the first time, on the application of this technique for the available data of geomagnetic storm time ranges. The result of the analysis is compared with that using the non-fuzzy method (Johnson and Wichern, 2006).

3. Significance of Fuzzy Transitive Closure Algorithm in the Cluster Analysis

Clustering techniques have been applied to a wide variety of research problems. Hierarchical (local, graph, theoretic) clustering methods can be divided into agglomerative and divisive methods. In this paper, the agglomerative algorithm is applied because of its connection to fuzzy relational methods that produce hierarchical clusters. The widely used technique is a c-means clustering, which has certain drawbacks. Fuzzy c-means clustering has been applied for various atmospheric and geophysical studies (e.g., Dekkers *et al.*, 1994; Kruiver *et al.*, 1999). This method requires the specification of the desired number of clusters, which is a disadvantage whenever the clustering problem does not specify any desired number of clusters. Even if the population is known to consist of k groups, the sampling method may be such that the data from the rarest group do not appear in the sample. Forcing the data into k groups would lead to nonsensical clusters. In such problems, the number of clusters should reflect, in a natural way, the structure of the given data. Methods based on fuzzy transitive closure work in this way.

4. Data

Simultaneous data of the geomagnetic storm time ranges of the sensitive horizontal (H) component for a network of geomagnetic observatories situated at Alibag, Hyderabad, Kodaikanal, Nagpur, Pondicherry, Sabhawala, and Tirunelveli available for the period between 2001 and 2003 are considered for this analysis. The Tirunelveli and Kodaikanal observatories situated nearer to the dip equator are influenced by the daytime equatorial electrojet current system. Sabhawala observatory is situated nearer to the focus of the S_q currents in the northern hemisphere. At middle and low latitudes on the Earth, one storm each year is larger than 250 gammas, and about ten storms per year are over 50 gammas and display a similar general appearance for the H component field (Campbell, 1997). The range obtained from the maximum and minimum values recorded in the magnetograms of the observatories during the geomagnetic storms are considered for this analysis. The locations of the observatories are provided in Table 1, and its corresponding map is provided in Fig. 1. The data used for this study in Table 2 are taken from the data book published by the Indian Institute of Geomagnetism. Graphical representation of data is presented in Fig. 2.

Table 1. Location of observatories (geographic and dipole coordinates).

Serial number	Station	Geographic	Dipole
1	Sabhawala (SAB)	Latitude 30°22′N	21.2°N
		Longitude 77°48′E	151.9°
2	Nagpur (NGP)	Latitude 21°09′N	11.96°N
		Longitude 79°05′E	152.1°
3	Alibag (ABG)	Latitude 18°37′N	10.02°N
		Longitude 72°52′E	145.9°
4	Hyderabad (HYB)	Latitude 17°25′N	8.3°N
		Longitude 78°33′E	151.3°
5	Pondicherry (PON)	Latitude 11°55′N	2.7°N
		Longitude 79°55′E	152.1°
6	Kodaikanal (KOD)	Latitude 10°14′N	1.23°N
		Longitude 77°28′E	149.6°
7	Tirunelveli (TIR)	Latitude 8°42′N	0.33°S
		Longitude 77°48′E	149.76°

Table 2. Geomagnetic storm time ranges for the horizontal H component (in nT).

Serial number	Date	ABG-H	HYB-H	KOD-H	NGP-H	PON-H	SAB-H	TIR-H
1	31.01.2001	149	146	163	158	157	128	194
2	19.03.2001	271	273	275	266	279	271	291
3	22.03.2001	146	142	211	145	170	138	267
4	27.03.2001	202	223	271	208	206	131	280
5	08.04.2001	177	177	194	173	234	228	232
6	11.04.2001	332	338	323	347	337	382	323
7	18.04.2001	118	117	189	115	147	160	240
8	28.04.2001	222	226	246	237	240	194	293
9	27.05.2001	142	145	133	154	134	128	144
10	18.06.2001	158	172	225	153	198	125	265
11	17.08.2001	215	207	219	223	213	220	230
12	25.09.2001	185	203	184	191	185	255	207
13	11.10.2001	136	138	181	145	167	153	215
14	21.10.2001	284	295	270	297	286	347	267
15	28.10.2001	219	238	292	213	272	224	315
16	29.12.2001	188	167	228	184	200	170	276
17	17.02.2002	108	113	222	110	139	51	246
18	28.02.2002	176	193	251	193	219	100	211
19	18.03.2002	119	121	133	135	123	158	187
20	23.03.2002	181	176	243	177	180	173	225
21	19.04.2002	183	193	222	200	203	164	272
22	23.04.2002	179	191	234	187	216	144	292
23	11.05.2002	157	158	241	163	165	189	178
24	23.05.2002	255	220	147	235	244	65	239
25	01.08.2002	138	152	186	200	163	155	228
26	18.08.2002	81	98	221	102	113	110	236
27	07.09.2002	162	76	275	166	169	133	252
28	30.09.2002	147	150	308	150	146	134	174
29	11.11.2002	91	90	127	93	102	65	145
30	26.11.2002	83	95	92	81	95	114	94
31	20.03.2003	161	158	158	169	169	126	223
32	29.05.2003	169	180	207	182	171	241	253
33	17.08.2003	254	261	251	258	265	258	248
34	15.11.2003	122	122	119	127	123	122	150

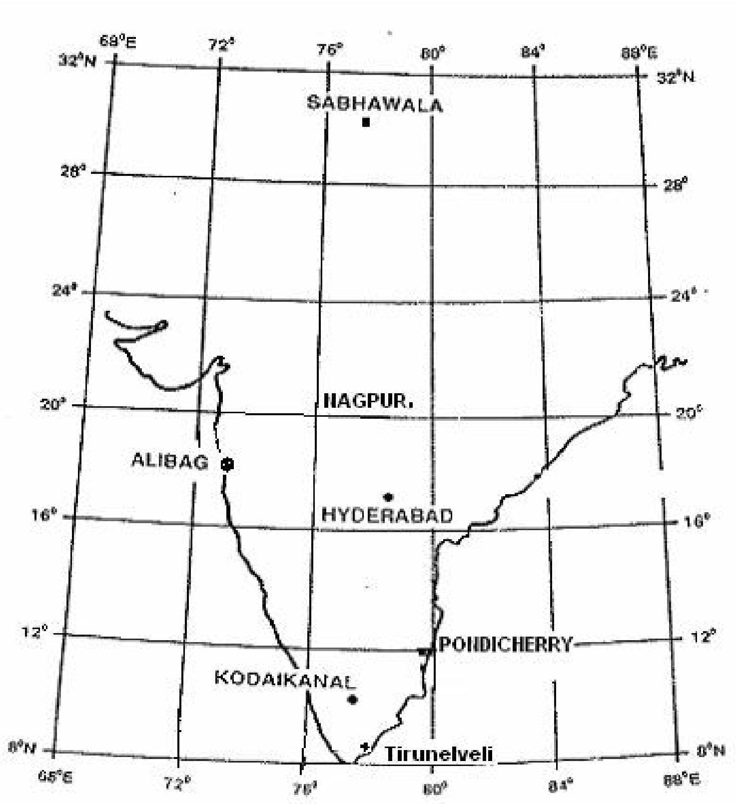

Fig. 1. Location map of the observatories.

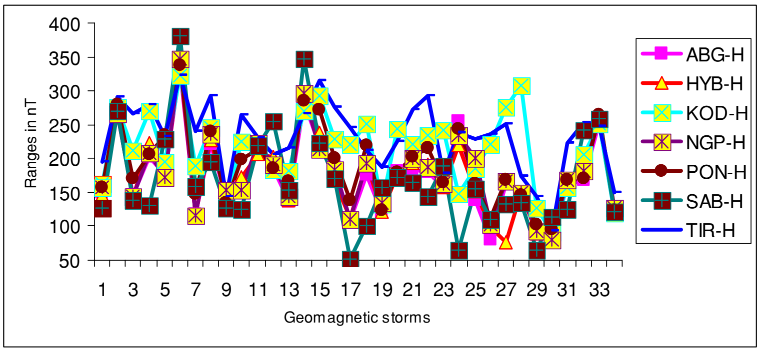

Fig. 2. Geomagnetic storm time ranges for horizontal H component in nT.

5. Classical Clustering Method

5.1 Euclidean distance matrix (*R*)

The geomagnetic storm time ranges described in Table 2 are considered for the derivation of the Euclidean distance matrix by the following formula:

$$\text{Distance } = \left[(x_{11} - x_{12})^2 + (x_{21} - x_{22})^2 + \ldots.. + (x_{n1} - x_{n2})^2\right]^{1/2}$$

The stations are abbreviated as follows: ABG, Alibag; HYB, Hyderabad; KOD, Kodaikanal; NGP, Nagpur; PON, Pondicherry; SAB, Sabhawala; TIR, Tirunelveli. The matrix R is constructed using the data given in Table 2, and the Euclidean distance is determined between one place and other places using the formula described above.

For example, the distance between ABG and HYB $= [(149-146)^2+(271-273)^2+(146-142)^2+\ldots+(254-261)^2+(122-122)^2]^{1/2} = 109.86$. The distance between ABG and KOD $= [(149-163)^2+(271-275)^2+(146-211)^2+\ldots(254-251)^2+(122-119)^2]^{1/2} = 371.87$.

All of the distances are calculated in a similar manner, and the matrix R is formed as follows:

$R =$

	ABG	HYB	KOD	NGP	PON	SAB	TIR
ABG	0	109.86	371.87	84.469	132.97	289.02	431
HYB	109.86	0	379.31	117.32	148.31	275.73	435.85
KOD	371.87	379.31	0	347.89	302.89	461.38	262.46
NGP	84.469	117.32	347.89	0	130.92	282.01	401.09
PON	132.97	148.31	302.89	130.92	0	319.18	334.08
SAB	289.02	275.73	461.38	282.01	319.18	0	536.07
TIR	431	435.85	262.46	401.09	334.08	536.07	0

5.2 Algorithm for hierarchical clustering (non-fuzzy method)

Agglomerative hierarchical methods start with individual objects, and initially there are as many clusters as objects. The following steps are performed in the agglomerative hierarchical clustering algorithm for grouping N objects:

1. Start with N clusters, each containing a single entity and an $N \times N$ symmetric matrix of distances $R = \{d_{ik}\}$.
2. Search the distance matrix for the nearest pair of clusters. Let the distance between "most similar" clusters U and V be d_{UV}.
3. Merge clusters U and V. Label the newly formed cluster (UV). Update the entries in the distance matrix by (1) deleting the rows and columns corresponding to clusters U and V and (2) adding a row and column giving the distances between cluster (UV) and the remaining clusters.
4. Repeat Steps 2 and 3, a total of $N - 1$ times. (All objects will be in a single cluster after the algorithm terminates). Record the identity of clusters that are merged and the levels (distances or similarities) at which the mergers take place.

Distance Matrix $R = \{d_{ik}\} =$

	ABG 1	HYB 2	KOD 3	NGP 4	PON 5	SAB 6	TIR 7
ABG 1	0	109	371	84	132	289	431
HYB 2	109	0	379	117	148	275	435
KOD 3	371	379	0	347	302	461	262
NGP 4	84	117	347	0	130	282	401
PON 5	132	148	302	130	0	319	334
SAB 6	289	275	461	282	319	0	536
TIR 7	431	435	262	401	334	536	0

Min $(d_{ik}) = d_{41} = 84$.

Objects 4 and 1 (NGP and ABG, respectively) are merged to form the cluster (41). The distances between the cluster (41) and the remaining objects 2, 3, 5, 6 and 7 are needed to implement the next level of clustering. The nearest neighbor distances are

$$d_{(41)2} = \min\{d_{41}, d_{42}\} = \min\{84, 117\} = 84$$
$$d_{(41)3} = \min\{d_{43}, d_{13}\} = \min\{347, 371\} = 347$$
$$d_{(41)5} = \min\{d_{45}, d_{15}\} = \min\{130, 132\} = 130$$
$$d_{(41)6} = \min\{d_{46}, d_{16}\} = \min\{282, 289\} = 282$$
$$d_{(41)7} = \min\{d_{47}, d_{17}\} = \min\{401, 431\} = 401$$

By deleting the rows and columns of D corresponding to objects 4 and 1 and by adding a row and column for the cluster (41), we can obtain the following distance matrix.

	(41)	HYB 2	KOD 3		PON 5	SAB 6	TIR 7
(41)	0	84	347		130	282	401
HYB 2	84	0	379		148	275	435
KOD 3	347	379	0		302	461	262
PON 5	130	148	302		0	319	334
SAB 6	282	275	461		319	0	536
TIR 7	401	435	262		334	536	0

The smallest distance between a pair of clusters is now $d_{(41)2} = 84$, and object 2(HYB) can be merged with cluster (41) (NGP and ABG) to obtain the next cluster (241) (NGP, ABG, and HYB). The distances between the cluster (241) and the remaining objects 3, 5, 6 and 7 are needed to implement the next level of clustering. The nearest neighbor distances are

$$d_{(241)3} = \min\{d_{(41)3}, d_{23}\} = \min\{347, 379\} = 347$$
$$d_{(241)5} = \min\{d_{(41)5}, d_{25}\} = \min\{130, 148\} = 130$$
$$d_{(241)6} = \min\{d_{(41)6}, d_{26}\} = \min\{282, 275\} = 275$$
$$d_{(241)7} = \min\{d_{(41)7}, d_{27}\} = \min\{401, 435\} = 401$$

The following matrix is obtained by deleting the rows and column of D corresponding to (41) and 2 and adding a row and column for the cluster (241).

	(241)		KOD 3		PON 5	SAB 6	TIR 7
(241)	0		347		130	275	401
KOD 3	347		0		302	461	262
PON 5	130		302		0	319	334
SAB 6	275		461		319	0	536
TIR 7	401		262		334	536	0

The smallest distance between the pair of cluster is now $d_{(241)5} = 130$. One can merge the object, 5(PON) with cluster (241) to get the next cluster (5241).

$$d_{(5241)3} = \min\{d_{(241)3}, d_{53}\} = \min\{347, 302\} = 302$$
$$d_{(5241)6} = \min\{d_{(241)6}, d_{56}\} = \min\{275, 319\} = 275$$
$$d_{(5241)7} = \min\{d_{(241)7}, d_{57}\} = \min\{401, 334\} = 334$$

The distance matrix for the next level of clustering is

	(5241)		KOD 3			SAB 6	TIR 7
(5241)	0		302			275	334
KOD 3	302		0			461	262
SAB 6	275		461			0	536
TIR 7	334		262			536	0

The smallest distance between pair of clusters is now (37) (KOD, TIR). This cluster does not have any link with the previous (5241) cluster and is a separate cluster. Therefore, two clusters, namely {5241} and {37}, are obtained.

Deleting the rows and columns of 3 and 7 and adding a cluster (37), we get

	(5241)	(37)				SAB 6	
(5241)	0					275	
(37)		0					
SAB 6	275					0	

One can see three clusters {5241}, {37}, and {6}, representing {NGP, ABG, HYD, PON}, {KOD, TIR}, and {SAB} clusters, respectively.

6. Fuzzy Transitive Closure Method

6.1 Equivalence relations

The equivalence relation (ER) is very important in pattern recognition (Bezdek *et al.*, 1986) because an equivalence relation on a set of objects defines a set of equivalence classes. An equivalence relation is defined on a set which is reflexive, symmetric and transitive.

Let the set of objects be $O = \{o_1, o_2, ..., o_n\}$. A crisp (either 0 or 1) binary relation R in O is a crisp sub-set $R \subset O \times O$. The R can be described by a membership function $\rho : O \times O \to \{0, 1\}$. The n^2 numbers $\{\rho(o_i, o_j)\}$, which characterize the membership of (o_i, o_j) in the relation R, is defined as $\rho(o_i, o_j) = 1 \Leftrightarrow o_i R o_j$ (if o_i is related to o_j, the membership value is '1'; if membership value is '1', then o_i is related to o_j). The $n \times n$ relational matrix is $R(\rho : O) = [r_{ij} = \rho(o_i, o_j)]_{n \times n}$. The terms reflexive, symmetric, and transitive relationships are explained as follows:

(1) Reflexive relationship: Any square binary relation R is reflexive if $r_{ii} = \rho(o_i, o_j) = 1 \,\forall\, o_i \in O$. Reflexivity means that every element is fully related to itself.
(2) Symmetric relationship: Any square binary relation R is symmetric if $r_{jk} = r_{kj} \,\forall\, j \neq k$, that is if $R = R^T$ (transpose of the matrix R). This means that whenever o_j is related to o_k at any level, o_k is related to o_j at the same level.
(3) Transitive relationship: A crisp relation $R = [r_{jk}]_{n \times n}$ is transitive if $r_{jk} = 1$ whenever $r_{ji} = 1$ and $r_{ik} = 1$ for some i.

6.2 Fuzzy equivalence relations

A fuzzy equivalence relation (Klir and Yuvan, 1997) is defined on a set which is reflexive, symmetric, and maximum-minimum transitive. Similar to an ordinary equivalence relation, a fuzzy equivalence relation induces a partition in each of its α-cuts (Zadeh, 1965). As such, a meaningful fuzzy equivalence relation is defined on the transitive closure of the fuzzy compatibility relation (Anderberg, 1973). A fuzzy compatibility relation R on a set S consisting of n data items can be defined by an appropriate distance function (Klir and Folger, 2000) of the Minkowski class by the following formula: $R(x_i, x_k) = 1 - \delta \left[\sum (x_{ij} - x_{kj})^q\right]^{1/q}$ for all pairs $(x_i, x_k) \in S$ where q is a positive real number and δ is a constant that ensures that $R(x_i, x_k) \in [0, 1]$. The quantity δ is the inverse value of the largest distance in S.

6.3 Fuzzy transitive closure algorithm

Let R^1 be the square matrix of order k obtained from the given data matrix derived by the Minkowski class. The relational matrix $R^{(2)} = R^1 o R^1$ where an element of $R^1 o R^1$ is the maximum-minimum (x_{rj}, x_{js}) with j varying from 1 to k (k is the number of observatories which is '7' in this analysis), and x_{rs} is an element in the r^{th} row and s^{th} column of the matrix $R^{(2)}$.

Similarly,

$$R^{(4)} = R^{(2)} o R^{(2)}$$

$$\cdots\cdots\cdots\cdot$$

$$\cdots\cdots\cdot$$

$$R^{(2k)} = R^{(2k-1)}$$

This is continued until no new relationship is produced. Thus, the maximum-minimum transitive closure R is the relation $R^{(n-k)}$ which is denoted by R_τ. This relation R_τ induces a partition called α-cuts (Zadeh, 1965) in different intervals. S is the distance matrix R defined in 5.1. The operator 'o' is applied to define the maximum of the minimum values obtained from the corresponding rows and columns of a particular element in the matrix R^1, R^2, R^4, etc.

6.4 α-cut

An α-cut of a fuzzy set I is a crisp set I_α that contains all of the elements of the universal set X that have a membership grade in I that is greater than or equal to the specified value of α.

$$\text{i.e., } I_\alpha = \{x \in X / \mu_I(x) \geq \alpha\}$$

A fuzzy relation into crisp partitions with different values of α corresponds to α-cuts of I_α. The scalars (α) are often regarded as an indication of the validity of each hard clustering in the interval [0, 1] (Friedman and Kandel, 1999). 'I' is just the edge weight of the strongest adjacency link between each pair of nodes.

6.5 Data analysis

The distance matrix R described in the classical clustering method is applied to derive the relational matrix R^1. The inverse value of the largest distance between SAB and TIR = 1/536.07 is taken as the value 'δ'. The value of $q = 2$ for the Minkowski class. The element of the relational matrix R^1 is obtained by multiplying each element of the matrix R by δ and subtracting them from '1'; thereby, the data are converted into the fuzzy domain. The matrix obtained from the given data matrix derived by the Minkowski class is the relational matrix R^1.

Relational matrix R^1 =

	ABG	HYB	KOD	NGP	PON	SAB	TIR
ABG	1	0.795	0.3063	0.8424	0.7519	0.4609	0.196
HYB	0.795	1	0.2925	0.7812	0.7234	0.4857	0.1869
KOD	0.3063	0.2925	1	0.3511	0.435	0.1393	0.5104
NGP	0.8424	0.7812	0.3511	1	0.7556	0.474	0.252
PON	0.7519	0.7234	0.435	0.7556	1	0.4046	0.3768
SAB	0.4609	0.4857	0.1393	0.474	0.4046	1	0
TIR	0.196	0.1869	0.5104	0.252	0.3768	0	1

The relational matrix $R^2 = R^1 o R^1$ is obtained as follows:

Take the elements in the first row and first column in R^1:

First row:	1	0.795	0.3063	0.8424	0.7519	0.4609	0.196
First column:	1	0.795	0.3063	0.8424	0.7519	0.4609	0.196
Minimum:	1	0.795	0.3063	0.8424	0.7519	0.4609	0.196

Maximum of minimum: 1
The value '1' is defined as x_{11} in the matrix R^2.

Take the elements of first row and second column in R^1

First row:	1	0.795	0.3063	0.8424	0.7519	0.4609	0.196
Second column:	0.795	1	0.2925	0.7812	0.7234	0.4857	0.1869
Minimum:	0.795	0.795	0.2925	0.7812	0.7234	0.4609	0.1869

Maximum of minimum: 0.795
The value '0.795' is defined as x_{12} in the matrix R^2.

Similarly, all elements x_{ij} are found for the matrix R^2.

The relational matrix R^2 is given below:

$R^2 = R^1 o R^1 =$

	ABG	HYB	KOD	NGP	PON	SAB	TIR
ABG	1	0.795	0.435	0.8424	0.7519	0.4857	0.3768
HYB	0.795	1	0.435	0.795	0.7556	0.4857	0.3768
KOD	0.435	0.435	1	0.435	0.435	0.4046	0.5104
NGP	0.8424	0.795	0.435	1	0.7556	0.4857	0.3768
PON	0.7519	0.7556	0.435	0.7556	1	0.4857	0.435
SAB	0.4857	0.4857	0.4046	0.4857	0.4857	1	0.3768
TIR	0.3768	0.3768	0.5104	0.3768	0.435	0.3768	1

By successively applying the same procedures, corresponding elements are found for the matrix R^4 from R^2. The relational matrix R^4 is given below:

$R^4 = R^2 o R^2 =$

	ABG	HYB	KOD	NGP	PON	SAB	TIR
ABG	1	0.795	0.435	0.8424	0.7556	0.4857	0.435
HYB	0.795	1	0.435	0.795	0.7556	0.4857	0.435
KOD	0.435	0.435	1	0.435	0.435	0.4046	0.5104
NGP	0.8424	0.795	0.435	1	0.7556	0.4857	0.435
PON	0.7556	0.7556	0.435	0.7556	1	0.4857	0.435
SAB	0.4857	0.4857	0.4046	0.4857	0.4857	1	0.435
TIR	0.435	0.435	0.5104	0.435	0.435	0.435	1

Again, by the transitive closure algorithm $R^8 = R^4 o R^4$. The relational matrix R^8 is given below:

$R^8 = R^4 o R^4 =$

	ABG	HYB	KOD	NGP	PON	SAB	TIR
ABG	1	0.795	0.435	0.8424	0.7556	0.4857	0.435
HYB	0.795	1	0.435	0.795	0.7556	0.4857	0.435
KOD	0.435	0.435	1	0.435	0.435	0.4046	0.5104
NGP	0.8424	0.795	0.435	1	0.7556	0.4857	0.435
PON	0.7556	0.7556	0.435	0.7556	1	0.4857	0.435
SAB	0.4857	0.4857	0.4046	0.4857	0.4857	1	0.435
TIR	0.435	0.435	0.5104	0.435	0.435	0.435	1

When two successive matrices are identical, the stopping condition for the algorithm is achieved, leading to α-cuts in the interval [0, 1].

It is seen that $R^8 = R^4 o R^4 = R^4$. This means that the stopping condition for the transitive closure algorithm is achieved (Friedman and Kandel, 1999). The final matrix is defined as R_τ. As such, the transitive closure algorithm leads to cuts in the interval [0, 1] at 0.842, 0.795, 0.755, 0.510, 0.485, and 0.435 as shown below:

α-cuts:

0--... 0---... ---0---- 0---... 0---- .....0----- ...0---....

0 0.842 0.795 0.755 0.510 0.485 0.435 1

Thus, the following α-cuts are formed by R_τ

The brackets (,) define the open interval. The brackets [,] define the closed interval. The brackets {, } define the set.

$\alpha \in (0.842, 1]$: {(ABG), (HYB), (KOD), (NGP), (PON), (SAB), (TIR)}

$\alpha \in (0.795, 0.842]$: {(NGP − ABG), HYB, KOD, PON, SAB, TIR}]

$\alpha \in (0.755, 0.795]$: {(HYB − ABG), (NGP − HYB), (KOD, PON, SAB, TIR)}

: {(NGP − ABG − HYB), (KOD, PON, SAB, TIR)]

$\alpha \in (0.510, 0.755]$: {(ABG − PON), (HYB − PON), (NGP − PON), (KOD, SAB, TIR)]

: {(NGP − ABG − HYB − PON), (KOD, SAB, TIR)]

$\alpha \in (0.485, 0.510]$: {(KOD − TIR), (NGP − ABG −HYB − PON), (SAB)]

$\alpha \in (0.435, 0.485]$: {(ABG − KOD), (ABG − TIR), (HYB − KOD), (HYB − TIR), (KOD − NGP), (KOD − PON), (NGP − TIR), (PON − TIR), (SAB − KOD), (SAB − TIR)]

: [(NGP − ABG − HYB − PON), (KOD − TIR), SAB]

7. Dendogram

A dendogram is a graphical representation of the results of a hierarchical cluster analysis. The term “dendogram” is used in numerical taxonomy for any graphical drawing or diagram giving a tree-like description of a taxonomic system. The representation of a taxonomic system by a dendogram is particularly suitable in connection with a cluster analysis to investigate the structure of the corresponding operational taxonomic units; that is, entities or individuals considered to be the lowest ranking taxa within the system. This is a tree-like plot where each step of hierarchical clustering is represented as a fusion of two branches of tree into a single one. The branches represent clusters obtained on each step of the clustering.

8. Results

The α-cuts and the corresponding dendograms for the geomagnetic storm time ranges of the seven observatories are obtained using the fuzzy transitive closure algorithm. An examination of the dendograms with respect to α-cuts shows the following.

There are three clusters in the dendogram for the horizontal H component. The first cluster is formed by NGP, ABG, HYB, and PON; the second cluster is formed by KOD and TIR; SAB is isolated from the above two clusters. Rastogi and Patil (1992) explained the significant differences in the daily variation of the horizontal magnetic field H at an equatorial station during low and high sunspot years. The data taken for this analysis are between 2001 and 2003, which are relatively higher solar activity periods. During intense solar flare effects and storm sudden commencement events, the amplitude of H decreases progressively with increasing latitudes at the Indian chain of observatories (Rastogi *et al.*, 1997). The studies on multidimensional scaling (Sridharan and Ramasamy, 2002) show that the increase in latitudinal separation corresponds to an increase in Euclidean distances with respect to the geomagnetic storm time ranges at the Indian observatories.

From Table 1 and Figs. 1 and 3, one can observe that the decrease in geomagnetic latitude from Nagpur to Pondicherry corresponds to the formation of the first cluster. This first cluster (NAG, ABG, HYB, and PON) clearly

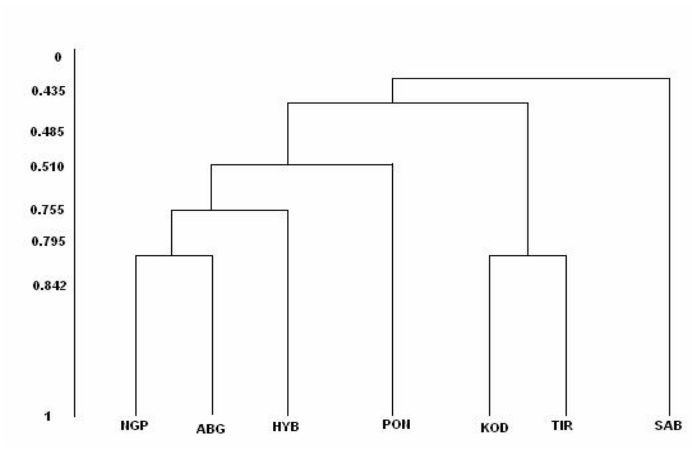

Fig. 3. Dendogram for geomagnetic storm time ranges.

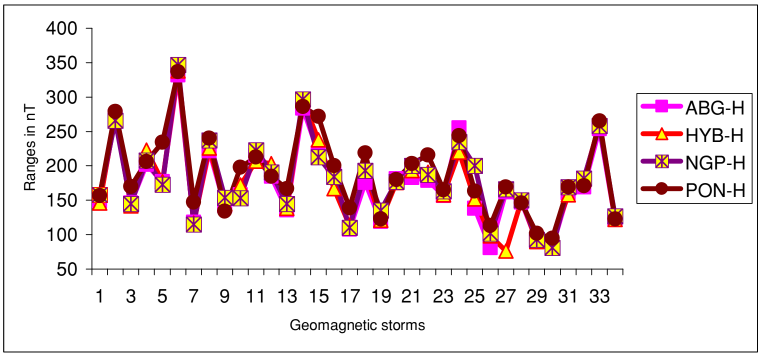

Fig. 4. Storm time ranges at non-equatorial regions.

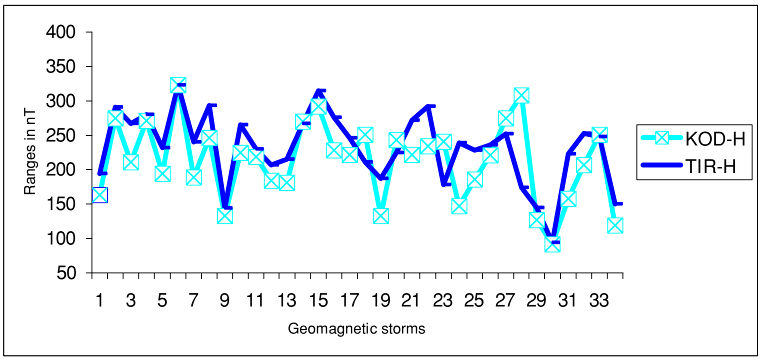

Fig. 5. Storm time ranges at equatorial region.

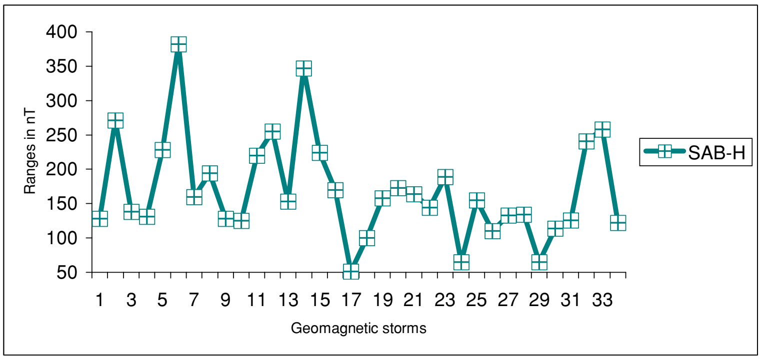

Fig. 6. Storm time ranges nearer to S_q focus.

shows the latitudinal dependence of geomagnetic variations on variations in geomagnetic storm time. Alex and Rao (1995) showed that the pattern of geomagnetic variations at equatorial latitudes differs significantly from that at higher latitudes. The abnormal variations and maintenance of larger spanning distances at the equatorial observatories have been explained by multidimensional scaling analysis (Sridharan and Ramasamy, 2002). The equatorial geomagnetic observatories forming a group in the dendogram are found in the second cluster formed by KOD and TIR. Jain and Sastri (1978) explained that neither ionospheric nor magnetospheric currents are directly responsible for the abnormal higher amplitude of variations at Sabhawala but, rather, this abnormal amplitude in variations is due to the sub surface electrical conductivity at the Himalayan foothills. Subsurface electrical conductivity may be the likely source, and this may be possibly following the trend of the Himalayas. This is seen in the dendogram (Fig. 3), where SAB is isolated from the other six places. Geomagnetic storm time variations at Sabhawala do not follow the pattern of latitudinal variations. Thus, the cluster validity at the equatorial and nonequatorial stations at the southern hemisphere and the isolated Sabhawala at the northern hemisphere is clearly seen in the dendogram described in Fig. 3. The graphs related to the dendogram are provided in Figs. 4, 5, and 6.

9. Conclusion

The fuzzy transitive closure algorithm is introduced in this paper as a means to identify the pattern of geomagnetic storm time ranges of certain observatories in the Indian region. Data obtained from the recently constructed observatories at Pondicherry and Tirunelveli are considered in this analysis. Pondicherry is a non-equatorial station and as such, its data follow the pattern of latitudinal dependence; Tirunelveli is nearer to the dip equator and hence affected by the daytime equatorial electrojet current system. Fuzzy logics are complementary technology in designing an intelligent data analysis approach. The utility of fuzzy concepts in pattern recognition is an emerging technique in various geophysical studies, and the fuzzy model presented here is a novel approach for the analysis of geomagnetic storms. The validity of the technique is confirmed by comparing the results with those obtained with classical cluster analysis. This analysis was performed using simultaneous data available only for a limited period of time. In general, for vast volumes of data in a complicated analysis, this technique yields accurate results. For researchers working in the field of geomagnetism, this fuzzy model will be important for furthering progress on the subject. The results reported here are expected to contribute to the application of this technique in future analyses exploring the hidden pattern(s) of various parameters in different fields of geomagnetism.

Acknowledgments. The author is very grateful to the referees for their critical evaluation and their suggestions for the improvement of the paper. The author is also very grateful to Prof. Archana Battacharyya, the Director, Indian Institute of Geomagnetism; Prof. S. Alex, Area Chairman, Observatory and Data Analysis, and Prof. B. M. Pathan, Head of Observatories, for their kind support in carrying out this research.

References

Alex, S. and D. R. K. Rao, Magnetospheric contribution to low latitude Sq variations, *J. Geomag. Geoelectr.*, **47**, 599–606, 1995.
Anderberg, M. R., *Cluster Analysis for Applications*, Academic Press, New York, 1973.
Bezdek, J. C., G. Biswas, and L.Y. Huang, Transitive closures of fuzzy thesauri for information and retrieval systems, *Int. J. Man-Machine Studies*, **25**, 343–356, 1986.
Campbell, W. H., *Introduction to Geomagnetic Fields*, Cambridge University Press, Cambridge, 1997.
Chandra, H., G. D. Vyas, B. M. Pathan, and D. R. K. Rao, Spectral characteristics of magnetic storm-induced F-region scintillations extending into daytime, *J. Atmos. Terr. Phys.*, **57**, 1273–1283, 1995.
Dekkers, M. J., C. G. Langereis, S. P. Vriend, P. J. M. van Santvoort, and G. J. de Lange, Fuzzy c-means cluster analysis of early diagenetic effects on natural remanent magnetization acquisition in a 1.1 Myr piston core from the Central Mediterranean, *Phys. Earth Planet. Inter.*, **85**, 155–171, 1994.
Friedman, M. and A. Kandel, *Introduction to Pattern Recognition-Statistical, Structural, Neural and Fuzzy Logic Approaches*, Imperial College Press, London, 1999.
Jain, A. R., Geomagnetic storm effects on ionospheric total electron content in Indian zone, *Indian J. Radio Space Phys.*, **7**, 111–118, 1978.
Jain, A. R. and N. S. Sastri, The H amplitude of Sudden Commencements of Magnetic Storms at Sabhawala (Dehra Dun), *J. Geophys.*, **45**, 101–106, 1978.
Johnson, R. A. and D. W. Wichern, *Applied Multivariate Statistical Analysis*, Pearson Education, Inc. Pearson Prentice Hall, NJ, USA, 2006.
Klir, G. J. and T. A. Folger, *Fuzzy Sets, Uncertainty and Information*, Prentice hall of India Private Ltd., New Delhi, 2000.
Klir, G. J. and B. Yuvan, *Fuzzy Sets and Fuzzy Logic, Theory and Applications*, Prentice Hall of India Private Ltd., New Delhi, 1997.
Kruiver, P. P., Y. S. Kok, M. J. Dekkers, C. G. Langereis, and C. Laj, A pseudo-Thellier relative palaeo intensity record, and rock magnetic and geochemical parameters in relation to climate during the last 276 kys in the Azores region, *Geophys. J. Int.*, **136**(3), 757–770, 1999.
Parkinson, W. D., *Introduction to Geomagnetism*, Scottish Academic Press, Edinburgh and London, 1983.
Rastogi, R. G. and A. R. Patil, On certain aspects of daily variation of geomagnetic field at low latitudes, *J. Geomag. Geoelectr.*, **44**, 495–503, 1992.
Rastogi, R. G., N. D. Kaushika, and N. B.Trivedi, Some relations between the sudden commencements at the equatorial electrojet, *J. Atmos. Terr. Phys.*, **26**, 771–776, 1964.
Rastogi, R. G., D. R. K. Rao, S. Alex, B. M. Pathan, and T. S. Sastry, An intense SFE and SSC event in geomagnetic H, Y and Z fields at the Indian chain of observatories, *Ann. Geophys.*, **15**, 1301–1308, 1997.
Sridharan, M. and A. M. S. Ramasamy, Multidimensional scaling technique for analysis magnetic storms at Indian observatories, *Proc. Ind. Acad. Sci.*, **111**, 459–465, 2002.
Sridharan, M. and N. Gururajan, Fractal behavior of geomagnetic storm on 29-31 October 2003, *Int. J. Mat. Sci.*, **54**, 303–311, 2006.
Srivastava, B. J., H. Abbas, D. R. K. Rao, and B. M. Pathan, Extended main phase of some sudden commencement great geomagnetic storms with double SSCs, *J. Atmos. Terr. Phys.*, **61**, 993–1000, 1999.
Sugiura, M., The solar diurnal variation in the amplitude of sudden commencements of magnetic storms at the geomagnetic equator, *J. Geophys. Res.*, **58**, 558–559, 1953.
Tsurutani, B. T., W. D. Gonzalez, G. S. Lakhina, and S. Alex, The extreme magnetic storm of 1-2 September 1859, *J. Geophys. Res.*, **108**, 2002.
Valdivia, J. A., A. S. Sharma, and K. Papadopoulos, Prediction of magnetic storms by nonlinear methods, *Geophys. Res. Lett.*, **23**(21), 2899–2902, 1996.
Valdivia, J. A., D. Vassiliadis, and A. Klimas, Spatiotemporal activity of magnetic storms, *J. Geophys. Res.*, **104**, 12239–12250, 1999.
Vassiliadis, D., A. J. Klimas, J. A. Valdivia, and D. N. Baker, The Dst geomagnetic response as a function of storm phase and amplitude and the solar wind electric field, *J. Geophys. Res.*, **104**, 24957–24976, 1999.
Zadeh, L. A., Fuzzy sets, *Inform. Control*, **8**, 338–353, 1965.

M. Sridharan (e-mail: mopondy@bsnl.in)