Jour. Indian Association of Sedimentologists, Vol. 36, No. 1 & 2 (2019), pp.186-204
ISSN Online 2582 – 2020
ISSN Print 0970 – 3268

Statistical Approach to Study the Lithostratigraphic Sequence in the Proterozoic Kolhans

Rohini Das
Department of Geology and Geophysics, IITKharagpur. PIN: 721301
romiyadas@gmail.com

Abstract: Lithofacies succession in the Proterozoic Kolhan Group has been studied statistically using modified Cross-Association Analysis, Markov chain model and Entropy function. The lithofacies analysis based on the field descriptions and their vertical packaging has been done for assessing the sediment depositional framework and the environment of deposition. Six lithofacies arranged in two genetic sequences have been recognized within the succession. The result of Markov chain and cross-association analysis indicates that the deposition of the lithofacies is in NonMarkovian and non cyclic process and represents asymmetric fining- upward. The chi-square test has been done to test randomness in hypotheses for lithofacies transition at confidence level of 95%. The entropy analysis has been done to evaluate the randomness of occurrence of lithofacies in a succession. Two types of entropies are related to every state: one is relevant to the Markov matrix expressing the upward transitions (entropy after deposition), and the other, relevant to the matrix expressing the downward transitions (entropy before deposition). The total energy regime calculated from the entropy analysis showing maximum randomness suggests that changing pattern in deposition has been a result of rapid to steady flow. This resulted a change in the depositional pattern from deltaic to lacustrine bypassing that finally generated non-cyclicity in the sequence.

Keywords: - *cyclicity, assymetricity, fan-delta, braided-ephemeral*

Introduction

The Kolhan Group is preserved as linear belt extending for 80-100 km with an average width of 10-12 km revealing deposition of Kolhan sediments in narrow and elongated troughs. The Kolhan Group lying unconformably above the Singhbhum granite is bounded by the Jagannathpur lavas on the southeast & south and the Iron Ore Group on the west. The western contact of the basin is faulted against the Iron Ore Group. The Kolhan Group of sediments are found in four detached sub-basins - Chaibasa-Noamundi basin, Chamakpur-

Keonjhargarh basin, Mankarchua basin and Sarapalli-Kamakhyanagar basin (Saha, 1994). Interpreting the depositional environment of the Proterozoic Kolhan sequence was difficult because of (a) the absence of zoo-fossils,(b) the absence of land vegetation, which has profound influence on precipitation, run-off, and sediment yield and (c) scarcity of exposures. Vertical variations of lithofacies within a given sequence play an important role in the recognition of depositional environment and their lateral dispersal (Walker, 1963). In order to determine the depositional architecture and its regional variations, a check of the results obtained so far (Tewari and Singh, 2009) by mathematical means seems desirable. Markov chain is one of the statistical methods that can be used to study the probability of occurrence and repetition of different rock units during deposition (Hota, 2000).

Study Area

The Kolhan Group in general (Fig.1) displays low (5˚-10˚) westerly dip. It unconformably overlies the Singhbhum granite to the east and with a faulted contact on the Iron Ore Group of rocks to the west (Saha, 1994). A pyrophyllitic shale layer (10 m thick) is locally present in between the Singhbhum granite and the Kolhans (Saha, 1994). The Chaibasa-Noamundi basin extends from Chaibasa (Long. 85^0 48’E: Lat. 22^0 33’N) in the north to Noamundi (Long. 85^0 28′ E: Lat. 22^0 09′N) in the south (length: 60-80 km; width: 8-10 km). The Chamakpur - Keonjhargarh (Long. 85°20’-85°35’ E ; Lat. 21°35’-22°10’ N) on the other hand covers an area approximately 375 km^2

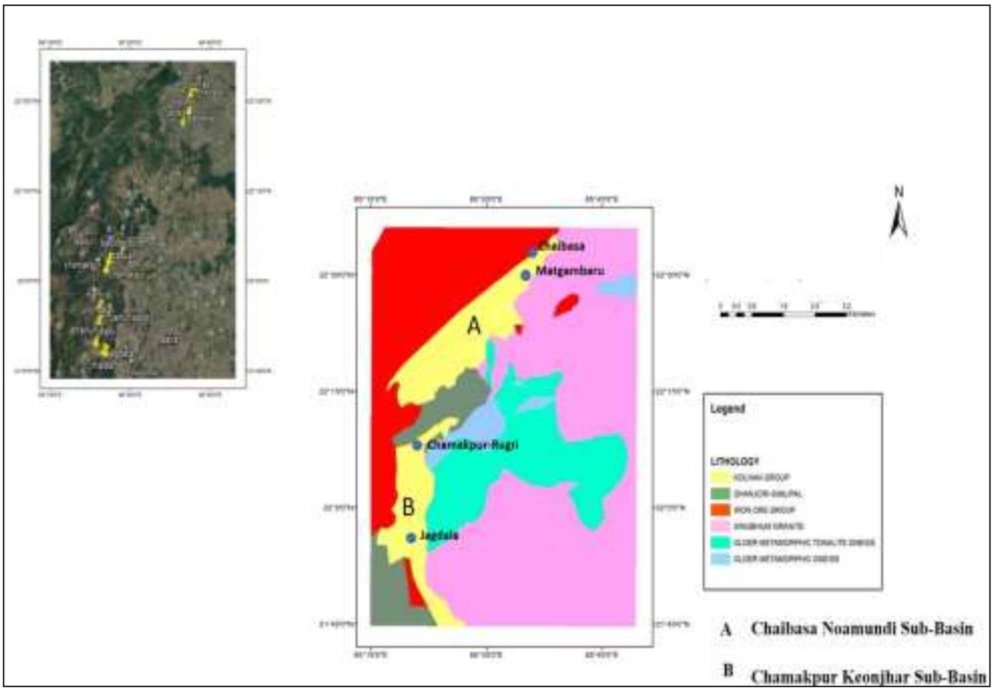

Fig. 1The geological map of Kolhan basin showing the two sub-basins (After Saha, 1994) with location index from Google earth.

(length : 50-55 km ; width : 6-8 km).

Methodology

Fieldworks were carried out to describe and characterize the lithounits of the Kolhan basin from Chaibasa to Chamkpur. At each exposure, the different lithounits were studied and identified on the basis of their bed geometries, gross lithologies, and sedimentary structures. The textural and the structural aspects of the lithounits observed in the outcrops were then clubbed into six lithofacies for better representation. The identity of each lithofacies was based on the presence of a set of primary textures and structures (Selley, 1970, 1976). Six lithofacies identified are:

(a) Granular lag facies (GLA),
(b) Granular sandstone facies (GSD),
(c) Sheet sandstone facies (SSD),
(d) Plane laminated sandstone facies (PLSD),
(e) Rippled sandstone facies (RSD), and
(f) Thin laminated siltstone-sandstone facies (TLSD)

Software package used for the statistical analyses and graphical representations was STATISTICA 8.0. Vertical sedimentary logs are prepared using software Sedlog. Only the representative sections required for the statistical analysis are shown. Programming of algorithm for Markov chain analysis and entropy analysis is carried out in Matlab and R.

Cross Association Analysis

Cross association detects the similarity between correlated geological vertical sections. Cross association is used to compare several geological vertical sections which are arbitrarily selected from different localities. Tracing the beds from section to section finally leads to what is known as lithological correlation. Lithological correlation meets with difficulties due to: 1) Lateral variation in bed thickness, 2) Lithology and 3) Missing of strata by erosion, lack of fossils, and tilting of strata The present paper confirms the validity of the geostatistical cross association method in facilitating the correlation between different geological sections which have great variation in litholog, thickness and obscurity in a single sedimentary succession.

Structuring Data for Cross Association Analysis

Lithounits	GLA	GSD	RSD	SSD	TLSD	PLSD
Gangabasha	1	2	1	0	0	4
Behind IT college	0	1	1	0	2	2
Rajanbasha	0	1	1	0	1	1
Gumuagara	0	4	0	0	2	4
Arjunbasha	2	1	1	2	3	4
Tunglai	0	3	2	0	0	3
Gutuhatu	0	1	1	0	2	0
Bingtopang	0	0	1	2	4	3
Bistampur	1	2	0	1	2	2
Diliamarcha	0	3	4	0	0	0
Matgamburu	1	1	2	0	0	1
Rajanka	1	2	2	0	0	3

Name of the Location	Numerical value of lithofacies present												
Gangabasha	6	3	5	3	5	3	4	3					
Behind IT college	5	2	3	4	3	2							
Rajanbasha	5	3	2	4									
Gumuagara	2	3	5	3	5	3	5	2	5	3			
Arjunbasha	2	3	1	4	1	5	2	6	3	6	3	2	3
Tunglai	3	5	3	5	4	5	4	3					
Gutuhatu	5	2	4	2									
Bingtopang	3	2	3	2	4	1	2	3	2	1			
Bistampur	2	5	2	5	3	6	3	1					
Diliamarcha	4	5	4	5	4	5	4						
Matgamburu	6	4	3	4	5								
Rajanka	6	5	3	4	3	5	3	4					

Table 2: Summary Statistics of different litho-columns of the study area

Analytical Procedure

To assess the degree of similarity between two sequences (sections), the nominal values in a given sequence are moved stepwise past the nominal values of a second sequence.

Δ = Number of comparisons (the length of the overlapped segments) and

t= the number of matches

The **Cross Association Index (CAI**) = the ratio of the number of matches to the length of the two overlapping segments.

Assuming that the number of matches at position *i*, then **CAI** is **CAI (I) = t/Δ.** The results of cross association analysis are described in Table.4.

X= Number of observations in the **k^th^** state of the chain

Y=Number of observations in the **k^th^** state of the chain.

XY=Product of **X** and **Y**.

M=Match position

P=Probability of match at any position of comparison

For any two random sequence of the same composition.

1- P= Probability of mismatch at any position of comparison for any two random sequence of the same composition...

E= Expected number of matches =**PΔ**

E'=Expected number of mismatches =**Δ-E**

O=Observed number of matches = **t**

O' = Observed number of mismatches =**Δ -O** Let the values of **1st** sequence be **a_1, a2, a3, AK**.

then **$\sum a_i$=n** where ai denotes the total number of occurrences of state **i**.

similarly, for **2nd** sequence of values are denoted by **b_1, b2, b3,......., b_k** then **$\sum b_i$=m** where bi denotes the total number of occurrences of state **i**. Using some counting method, the total number of possible ways of filling any match position is **m*n,** with 2 different values of **k**. While the same is **$\sum a_i b_i$** for identical values of **k**.

Thus the probability of match position of comparison for any two random sequence of the same composition is **$P = \sum a_i b_i /mn$**. The chi square test used as an approximation test.

χ 2 ={ (O-E) 2 /E }+ {(O'-E')2/E' }………equation (1)

Yates correction may be applied in statistics especially when the expected number of matches is small. This correction calls for subtraction of 0.05 from the absolute difference of expected and observed number of matches for a given significant level of 5%. the corresponding critical value is 3.84.. Thus, the equation (1) becomes,

χ 2 ={ (O-E-0.05) 2 /E }+ {(O'-E'-0.05)2/E' } When **Gangabasha** section is moved by **Behind ITI college** section one position at a time and is compared at each match position, it is observed that the best match position is at **10th** with **3** matches and **Δ=3/5**. When the last state of Gangabasha section (plane laminated sandstone) matches with 5th state of Behind ITI college section (plane laminated sandstone).

Null hypothesis,

H_0: Two sequence are not similar

H_1: H_0 is not true

The p-value is a probability that measures the evidence against the null hypothesis. Lower probabilities provide stronger evidence against the null hypothesis. It is used to determine whether to reject or fail to reject the null hypothesis, which states that the variables are independent. The p value calculated is 0.007 for r (degrees of freedom) varying from 0 to 5 from the chi-swuare table which is less than 0.05(5% significant level) and strongly support that H_0 is not true.

Match position 1 : 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 0 and CAI(1) = 0,*

Match position 2: 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 0 and CAI(2) = 0,*

Match position 3: 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 1 and CAI(3) = 1/3.*

Match position 4: 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 0 and CAI(4) = O*

Match position 5 : 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 2 and CAI(5) =2/5,*

Match position 6: 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 0and CAI(6) = 0,*

Match position 7: 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 2and CAI(7) = 2/6=1/3.*

Match position 8: 6 3 5 3 5 3 4 3 5 2 3 4

3 2 In this case *t = 1 and CAI(8) = 1/6*
Match position 9: 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 3 and CAI(9) = 3/5,*
Match position 10: 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 1 and CAI(10) = 1/4,*
Match position 11: 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 1 and CAI(11) = 1/3.*
Match position 12: 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 0 and CAI(12) = 0.*
Match position 13: 6 3 5 3 5 3 4 3 5 2 3 4 3 2 In this case *t = 0 and CAI(12) = 0.*

m*n= **∑Xi∑Yi** =6*8 =48
∑XiYi =2+1+8=11

Thus, **p= (∑$X_i Y_i$ /m*n)** =11/48=.2291
E=Δp=.2291*5=1.145 **{asΔ= 5 from match num 10}**
E'=Δ-E=5-1.1872=3.855 And **O=t**=3
O'=Δ-O=5-3=2
χ 2Y ={ (O-E-0.5) 2 /E }+ {(O'-E'-0.5)2/E' } ={(3-1.145-0.5)2/1.145}+{(2-3.855-0.5)2/3.855}
= 1.5+1.44 =**2.9** < **3.84** hence null hypothesis can't be rejected. **Correlation is not significant at 5%** significant level**.**

Markov Chain Analysis

In the field it is observed that there is gross lithological asymmetricity present between various lithofacies. There is marked difference in the sandstone and shale thickness, with shale thickness very high as compared to sandstone. It is difficult to prove in the field the presence of time independent depositional relation, if any, between the two facies as there is absence of unconformity. To prove cyclic arrangement in the lithofacies in the study area, the Markov property was applied (Gingerich,1969;Miall, 1973; Powers and Easterling,1982). Cyclic sedimentation defines cyclic or rhythmic sedimentation as a series of lithologic elements repeated through time. Alternatively, two types of observable cyclicity may be noteworthy: one in which there exists an order of sequence only, and another in which there is a certain order of repetition along the vertical scale of the sedimentary succession. In this study it is considered to ignore thickness altogether. Markov process, named after the Russian mathematician Andrey Markov, is a stochastic process that satisfies the Markovian property. It can be used to model a random system that changes states according to a transition rule that only depends on the current state. In a first-order Markov process, a lithologic unit or a facies state, F_j observed at a point n depends upon the facies state F_i observed at point (n-1). In other words, the geologic situation at point (n-1) governs the event that will happen at n. The transition probability of a facies being in the state F_j at n given that the facies is in state F_i at (n-1) is denoted by

$P_{ij}(n\text{-}1, n)$, i.e. for discrete-time Markov chains,

$$Pij(n\text{-}1, n) = P[(Fn_j)|(F(n\text{-}1)_i)]$$

Structuring Data for Markov Chain Analysis

The data used in the study is vertical sedimentary log successions of lithological members coded into a limited number of states for the Markov chain and Entropy analysis (Fig. 2). No account has been taken of the thickness of each member and no multistory lithologies are recognized. Thus, it is not considered possible for a given lithological state to pass upward into the same lithological state. In the present study only discrete lithofacies transitions regardless of individual bed thickness are counted, therefore, focus is on the evolution of the depositional process. In order to prevent transition tendencies from being too diffused throughout the count matrix, only six lithofacies, which are distinctly marked in each sedimentary log as well as in outcrop sections, are used in this study. To analyse cyclic characters through space and time, the lithofacies transitions are analyzed together in all sedimentary logs, and by pooling the data for four sectors as well as for the entire area. Seventeen lithological sections (Fig. 2) were considered for studying the vertical and areal distributions of the lithofacies within the Kolhan basin.

Analytical procedure

Frequency count matrix (F): Frequency count matrix is calculated from the vertical sequence profile of sedimentary logs shown in Fig. 2. Since markov chain is used which has memory less property i.e. the geologic situation at point (n-1) governs the event that will happen at n. That's why all seventeen sedimentary logs can be used to calculate matrix F without loss of information. Subsequently, data for all logs are added and a bulk matrix is structured at Basin level. Number of transition from facies i to j is represented in row i and column j of matrix F (Table 4a), which signifies number of times state j followed immediately after state i in the sedimentary logs. The frequency count matrix is structured into embedded Markov chain considering only transition lithologies and not their thickness as stated elsewhere. Since a transition is supposed to occur only when it results in a different lithology, the diagonal elements are all zeros in the resulting tally matrix.

The modified Markov process model used in this study incorporates structuring of one step embedded tally count matrix (F_{ij}), where i, j corresponds to row and column number. It will be noticed that where i = j, zeros are present

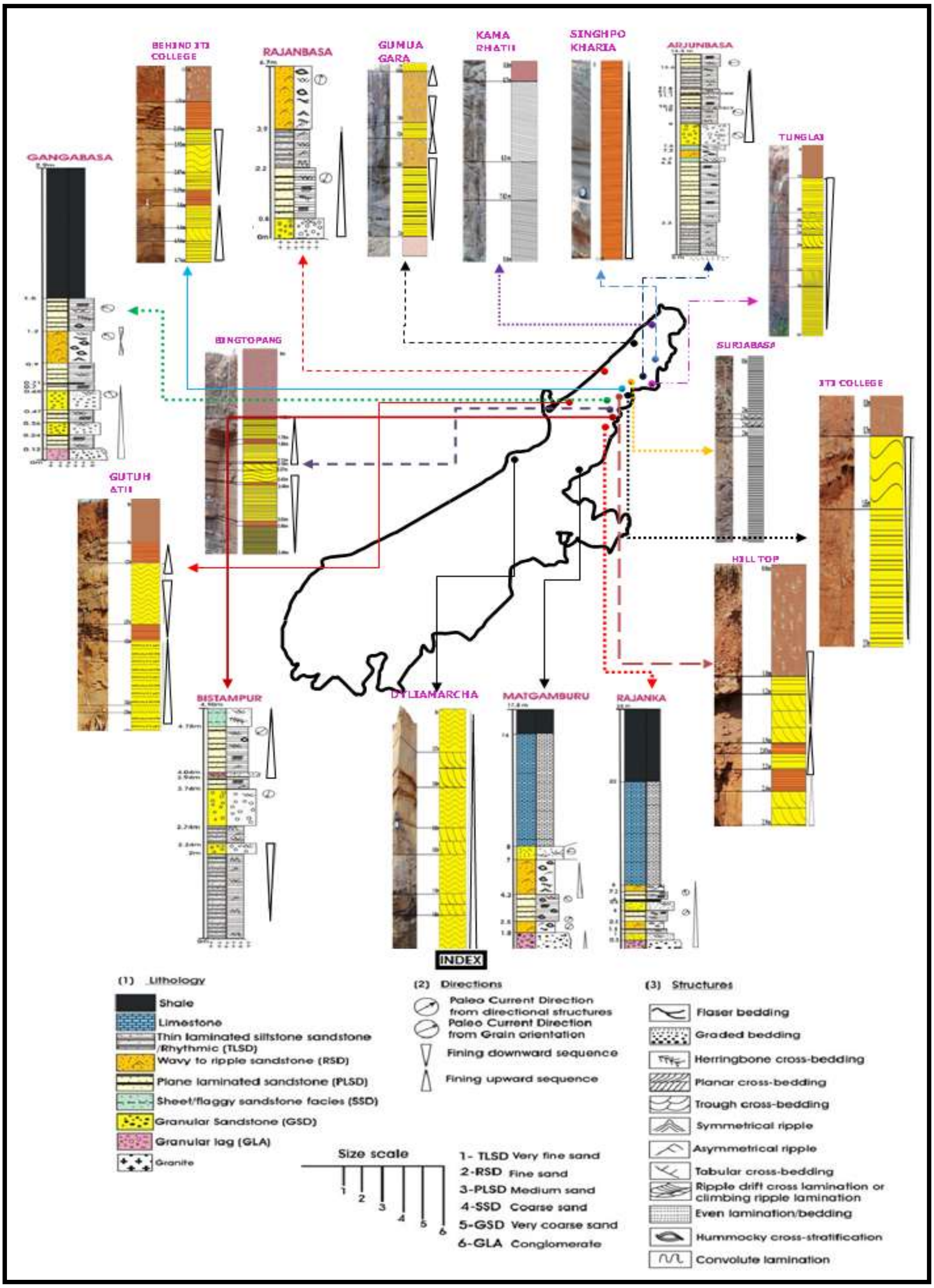

Fig. 2 Location of Lithologs showing the vertical distribution of the lithofacies in the study area

in the matrix, i.e., probability of moving from one state to another state has only been recorded where the lithofacies shows an abrupt change in character, regardless of the thickness of the individual bed.

Transition Frequency Matrix (F)

As mentioned above, a first order embedded chain matrix is structured by counting transition from one facies to another, and the resulting frequency matrix will contain zeros along the principal diagonal (F_{ij} =0)(Table 4a). This is a two dimensional array which records the frequency of the vertical transitions that occur between the different lithofacies in a given stratigraphic succession. The lower bed / facies of each transition couplet are given by the row numbers of the matrix, and the upper bed / facies by the column numbers. Each lithofacies is coded with column numbers or capital letters. The transition count matrix is expressed as F_{ij}, where i = row number and j = column number. When i=j, the transition between same lithofacies are denoted by zeros. In other words, transitions are only recorded when the lithofacies shows abrupt changes in character in spite of the thickness of the individual lithofacies.

Upward Transition Probability Matrix (P)

The upward transition probability matrix pertains to the upward ordering of lithologies in a succession and is calculated in the following manner: $P_{ij} = F_{ij} / SR_i$ Where, SR_i is the corresponding row total (Table 4b). The transition probability matrix represents the actual probabilities of the transition between one lithofacies to another in a vertical section. This array is obtained by taking the number of transitions of one facies to another and dividing by the total number of transitions involving the first facies.

Downward Transition Probability Matrix (Q)

Similar to the upward transition probability of lithologies a downward transition probability (Q matrix) can also be determined by dividing each element of the transition frequency matrix by the corresponding column total, i.e., $Q_{ji} = F_{ij} / SC_j$ Where, SC_j is the column total (Table 4c).

Independent Random Matrix (R)

Assuming that the sequence of rock types was determined randomly an independent trials matrix can be prepared in the following manner: $R_{ij} = C_j / (T – C_i)$ Where, C_i is the column total of facies state F_i, C_j is the column total of facies state F_j, N is the total number of transitions in the system (Table 4d). The diagonal cells are filled with zeros. This matrix represents the probability of the given transition that occur in a random manner. If t=total number of lithofacies, n=rank of the matrix (the total number of rows and columns used), T= n, k=0 F_{ij}, where Fij =number of transition from facies A to facies A-F.

Difference Matrix (D)

A difference matrix (Table 4e) is calculated which highlights those transitions that have a probability of occurring greater than if the sequence were random. By linking positive values of the difference matrix, a preferred upward path of facies transitions can be constructed which can be interpreted in terms of depositional processes that led to this particular arrangement of facies (Miall, 1973). $D_{ij} = P_{ij} - R_{ij}$ A positive value in difference matrix indicates that a particular transition occurs more frequently and a negative value indicates that it occurs less frequently. In difference matrix the values in each rows of the matrix sum to zero. If the values are close to zero, a vertical succession with little or no memory indicates independent nature of deposition of facies in a basin.

Expected Frequency Matrix (E)

It is necessary to construct an expected frequency matrix, since a statistician‟s rule of thumb states that chi – square tests should only be applied when the minimum expected frequency in any cell not exceeds 5. The matrix of expected values is given by $E_{ij} = R_{ij} * SR_i$ (Table 4f).

Test of Significance

Non-parametric chi-square ($\chi 2$) test has been applied to ascertain whether the given sequence has a Markovian memory or no memory. To test null hypothesis, chi-square ($\chi 2$) values are calculated for vertical successions (Table 4g).

$$\chi^2 = \sum_{i=0}^{n} \sum_{j=0}^{n} (Fij - Eij)^2/Eij$$

F_{ij} = transition count matrix or observed frequency of elements in the transition count matrix; Eij = Expected frequency matrix

ν = degree of freedom given by the number of non-zero entries in the $[r_{ij}]$ matrix minus the rank of the matrix=n2–2n, where n denotes rank of the matrix If the computed values of chi-square exceed the limiting values at the 5% significance level suggests the Markovity and cyclic arrangement of facies states.

Entropy Analysis

Hattori (1976) applied the concept of entropy to sedimentary successions possessing Markov property to determine the degree of random occurrence of lithologies in the succession. Methods of calculation of entropy as suggested by Hattori (1976) have been largely followed in the present study. Hattori (1976) recognized two types of entropies with respect to each lithological state; one is post–depositional entropy corresponding to matrix P and the other, pre-depositional entropy, corresponding to matrix Q. Hattori (1976) defined post-depositional entropy with respect to lithological state i as

$$E_i^{(post)} = -\sum_{j=0}^{n} P_{ij} * \log(P_{ij}) \quad \text{....eq1}$$

If E_i (post)= 0.0, state i is always succeeded by state j in the sequence. If E_i (post)>0, state i is likely to be overlain by different states. Hattori (1976) defined pre – depositional entropy with respect to state i as

$$E_i^{(pre)} = -\sum_{j=0}^{n} Q_{ij} * \log(Q_{ij}) \quad \text{...eq2}$$

Large E_i (pre) signifies that i occur independent of the preceding state. E_i (post) and E_i (pre) together

Form an entropy set for state i, and serves as indicators of the variety of lithological transitions immediately after and before the occurrence of i, respectively. Hattori (1976) used the interrelationships of E_i (post) and E_i (pre) to classify various cyclic patterns into asymmetric, symmetric and random cycles. The values of E_i (pre) and E_i (pre)calculated by equations (1) and (2) increases with the number of lithological states recognized. To eliminate this influence, Hattori (1976) normalized the entropies by the following equation:

$$R = E/E_{max}, \text{ where, } E_{max} = -\log 2 \frac{1}{n-1}$$

Where R is the normalized entropy, E is either post–depositional entropy or pre–depositional entropy, and Emax is the maximum entropy possible in a system where n state variable operates Table6.

RESULTS

The results of cross association analysis displays associatograms showing the position of maximum match between the widely separated stratal associations of litho sections (Fig. 3; Table 3).

Gangabasha Vs Behind ITI College

The maximum match at 10th position with 3 matches and 2 mismatches for 5 set of comparisons. The probability of matches is .30 and the expected number of matches and mismatches are 1.15 and 3.86 respectively. The computed value of χ^2 (2.90) does not exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the acceptance of null hypothesis (H_0) and rejection of the alternative hypothesis (H_1) at 5% level of significance.

Behind ITI college Vs Rajanbasha

The maximum match at 7th position with 2 matches and 1 mismatches for 3 set of comparisons. The probability of matches is .25 and the expected number of matches and mismatches are .75 and 2.25 respectively. The computed value of $\chi 2$ (2.11) does not exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the acceptance of null hypothesis (H_0) and rejection of the alternative hypothesis (H_1) at 5% level of significance.

Rajanbasha Vs Gumuagara

The maximum match at 2nd position with 2 matches and 0 mismatches for 2 set of comparisons. The probability of matches is .25 and the expected number of matches and mismatches are .50 and 1.50 respectively. The computed value of χ^2 (4.60) exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the rejection of null hypothesis (H_0) and acceptance of the alternative hypothesis (H_1) at 5% level of significance.

Gumuagara Vs Arjunbasha

The maximum match at 2nd position with 2 matches and 0 mismatches for 2 set of comparisons. The probability of matches is .20 and the expected number of matches and mismatches are .4 and 1.6 respectively. The computed value of $\chi 2$ (5.75) exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the rejection of null hypothesis (H_0) and acceptance of the alternative hypothesis (H_1) at 5% level of significance.

Arjunbasha Vs Tunglai

The maximum match at 20th position with 1 matches and 0 mismatches for 5 set of comparisons. The probability of matches is .16 and the expected number of matches and mismatches are .15 and .85 respectively. The computed value of $\chi 2$ (2.94) does not exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the acceptance of null hypothesis (H_0) and rejection of the alternative hypothesis (H_1) at 5% level of significance.

Tunglai Vs Gutuhatu

The maximum match at 9th position with 1 matches and 2 mismatches for 3 set of comparisons. The probability of matches is .16 and the expected number of matches and mismatches are .47 and 2.5 respectively. The computed value of $\chi 2$ (.42) does not exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the acceptance of null hypothesis (H_0) and rejection of the alternative hypothesis (H_1) at 5% level of significance.

Gutuhatu Vs Bingtopang

The maximum match at 5th position with 3 matches and 1 mismatch for 4 set of comparisons. The probability of matches is .25 and the expected number of matches and mismatches are 1 and 3 respectively. The computed value of $\chi 2$ (4.3) exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the rejection of null hypothesis (H_0) and acceptance of the alternative hypothesis (H_1) at 5% level of significance.

Bingtopang Vs Bistampur

The maximum match at 2nd position with 1 matches and 1 mismatches for 2 set of comparisons. The probability of matches is .2 and the expected number of matches and mismatches are .4 and 1.6 respectively. The computed value of $\chi 2$ (.75) does not exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the acceptance of null hypothesis (H_0) and rejection of the alternative hypothesis (H_1) at 5% level of significance.

Bistampur Vs Dyliamircha

The maximum match at 5th position with 2 matches and 3 mismatches for 5 set of comparisons. The probability of matches is .14 and the expected number of matches and mismatches are .71 and 4.29 respectively. The computed value of $\chi 2$ (1.61) exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the rejection of null hypothesis (H_0) and acceptance of the alternative hypothesis (H_1) at 5% level of significance.

Dyliamircha Vs Matgamburu

The maximum match at 2nd position with 2 matches and 0 mismatches for 2 set of comparisons. The probability of matches is .31 and the expected number of matches and mismatches are .63 and 1.37 respectively. The computed

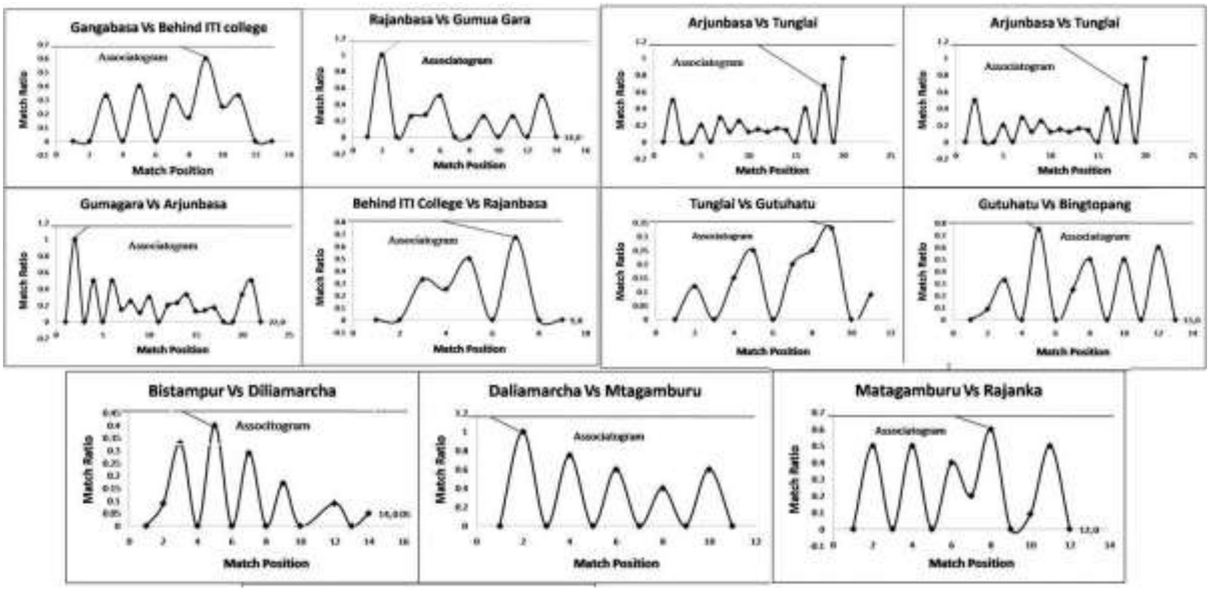

Fig.3Associatogram showing the position of maximum match between the widely separated strata associations.

Value of χ2 (3.75) exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the rejection of null hypothesis (H_0) and acceptance of the alternative hypothesis (H_1) at 5% level of significance.

Matgamburu Vs Rajanka

The maximum match at 8th position with 3 matches and 2 mismatches for 5 set of comparisons. The probability of matches is .23 and the expected number of matches and mismatches are 1.13 and 3.88 respectively. The computed value of χ^2 (2.65) does not exceed the critical value (3.84) for 1 degree of freedom at 5% significant level. This leads to the acceptance of null hypothesis (H_0) and rejection of the alternative hypothesis (H_1) at 5% level of significance.

Comparision Between logs	lithofacies	X	Y	∑X∑Y	M	P	1-P	O	O'	E	E'	χ^2Y	Conclusion
Gangabasha Vs Behind IT college	GLA	1	0	0	10	.3	.7	3	2	1.15	3.86	2.9	Correlation is not significant
	GSD	2	1	2									
	SSD	1	1	1									
	PLSD	0	0	0									
	RSD	0	2	0									
	TLSD	4	2	8									
	TOTAL	8	6	11									
Behind IT college Vs Rajanbasha	GLA	0	0	0	7	.25	.75	2	1	.75	2.25	2.11	Correlation is not significant
	GSD	1	1	1									
	SSD	1	1	1									
	PLSD	0	0	0									
	RSD	2	1	2									
	TLSD	2	1	2									
	TOTAL	6	4	6									

Comparision Between logs	lithofacies	X	Y	∑X∑Y	M	P	1-P	O	O'	E	E'	χ^2Y	Conclusion
Rajanbasha Vs Gumagara	GLA	0	0	0	2	.25	.75	2	0	.50	1.50	4.8	Correlation is significant
	GSD	1	4	4									
	SSD	1	0	0									
	PLSD	0	0	0									
	RSD	1	2	2									
	TLSD	1	4	4									
	TOTAL	4	10	10									
Gumagara Vs Arjunbasa	GLA	0	2	0	2	.2	.8	2	0	.4	1.6	5.75	Correlation is significant
	GSD	4	1	4									
	SSD	0	1	0									
	PLSD	0	2	0									
	RSD	2	3	6									
	TLSD	4	4	16									
	TOTAL	10	13	26									
Arjunbasha Vs Tunglai	GLA	2	0	0	20	.16	.84	1	0	.35	.85	2.94	Correlation not significant
	GSD	1	3	3									
	SSD	1	2	2									
	PLSD	2	0	0									
	RSD	3	0	0									
	TLSD	4	3	12									
	TOTAL	13	8	17									
Tunglai Vs Gutuhatu	GLA	0	0	0	9	.16	.84	1	2	.47	2.5	.42	Correlation not significant
	GSD	3	1	3									
	SSD	2	1	2									
	PLSD	0	0	0									
	RSD	0	2	0									
	TLSD	3	0	0									
	TOTAL	8	4	5									

Comparision Between logs	lithofacies	X	Y	$\sum X_i \sum Y_i$	M	P	1-P	O		O'	E	E'	$\chi^2 Y$	Conclusion
Gutuhatu Vs Bingtopang	GLA	0	0	0	5	.25	.75	3	1	1	3		4.3	Correlation is significant
	GSD	1	0	0										
	SSD	1	1	1										
	PLSD	0	2	0										
	RSD	2	4	8										
	TLSD	0	3	0										
	TOTAL	4	10	9										
Bingtopang Vs Bistampur	GLA	0	1	0	2	.2	.8	1	1	.4	1.6		.75	Correlation is not significant
	GSD	0	2	0										
	SSD	1	0	0										
	PLSD	2	1	2										
	RSD	4	2	8										
	TLSD	3	2	6										
	TOTAL	10	8	16										
Bistampur Vs Dalianarcha	GLA	1	0	0	5	.14	.86	2	3	.71	4.29		1.61	Correlation is not significant
	GSD	2	3	6										
	SSD	0	4	4										
	PLSD	1	0	0										
	RSD	2	0	0										
	TLSD	2	0	0										
	TOTAL	8	7	10										

Comparision Between logs	lithofacies	X	Y	$\sum X_i \sum Y_i$	M	P	1-P	O		O'	E	E'	$\chi^2 Y$	Conclusion
Matgambura Vs Rajankas	GLA	1	1	1	8	.23	.77	3	2	1.13	3.88		2.65	Correlation is not significant
	GSD	1	2	2										
	SSD	2	2	4										
	PLSD	0	0	0										
	RSD	0	0	0										
	TLSD	1	3	3										
	TOTAL	5	8	10										
Dalianarcha Vs Matgambura	GLA	0	1	0	2	.31	.69	2	0	.63	1.37		3.75	Correlation not significant
	GSD	3	1	3										
	SSD	4	2	8										
	PLSD	0	0	0										
	RSD	0	0	0										
	TLSD	0	1	0										
	TOTAL	7	5	11										

Table 3: Results of Cross Association Analysis in the Kolhan Group

a) Transition Count Matrix (F)

	A	B	C	D	E	F	SR_i	$T-SR_i$
A	0	1	1	3	0	1	6	33
B	3	0	1	1	1	1	7	32
C	0	1	0	0	0	1	2	37
D	1	2	0	0	1	1	5	34
E	0	2	0	0	0	0	2	37
F	2	1	0	1	0	0	4	35
SCj	6	7	2	5	2	4	Total=26	

b) Upward Transition Probability Matrix(P)

	A	B	C	D	E	F
A	0	0.35	0.1	0.25	0.1	0.2
B	0.316	0	0.105	0.263	0.105	0.21
C	0.25	0.292	0	0.208	0.083	0.167
D	0.286	0.333	0.095	0	0.095	0.19
E	0.25	0.292	0.083	0.208	0	0.166
F	0.273	0.318	0.091	0.227	0.091	0

c) Downward Transition Probability Matrix(Q)

	A	B	C	D	E	F
A	0	0.142	0.5	0.6	0	0.25
B	0.5	0	0.5	0.2	0.5	0.25
C	0	0.142	0	0	0	0.25
D	0.1667	0.285	0	0	0.5	0.25
E	0	0.285	0	0	0	0
F	0.333	.145	0	0.2	0	0

d) Independent Trails Probability Matrix(R)

	A	B	C	D	E	F
A	0	0.167	0.167	0.5	0	0.167
B	0.428	0	0.143	0.143	0.143	0.143
C	0	0.5	0	0	0	0.5
D	0.2	0.4	0	0	0.2	0.2
E	0	1	0	0	0	0
F	0.5	0.25	0	0.25	0	0

e) Difference Matrix (D)

	A	B	C	D	E	F
A	0	-0.183	0.67	0.25	-0.1	-0.033
B	0.113	0	0.37	-0.12	0.037	0.067
C	-0.25	0.208	0	-0.208	-0.083	0.334
D	-0.086	0.067	-0.095	0	0.105	0.009
E	-0.25	0.708	-0.083	0.208	0	-0.167
F	0.227	-0.68	-0.091	0.023	-0.091	0

f) Expected Frequency Matrix (E)

	A	B	C	D	E	F
A	0	2.1	0.6	1.5	0.6	1.2
B	2.212	0	0.735	1.841	0.735	1.47
C	0.5	0.584	0	0.416	0.166	0.334
D	1.43	1.665	0.475	0	0.475	0.95
E	0.5	0.584	0.166	0.416	0	0.332
F	1.092	1.272	0.364	0.908	0.364	0

g)Test Of Significance

Test of Equation	Computed value of	Limiting value at 0.5% significance level	Degree of freedom
Billingslay	14.343	30.14	19

Table 4: Matrices used to analyse transitions of lithofacies in the Kolhan Group

In the listed below tables (Table4a-f) A= GLA; B= GSD; C= SSD; D= PLSD; E= RSD; F= TLSD; SRi= Sum of ith row of the count matrix SC_j= Sum of jth column of the count matrix T= Total number of transition.

Discussion

Markov Chain and Cross Association Analysis

In the interpretation of significant facies transitions it is important to note that the calculated significant transitions

.

Represent the most probable facies transitions, but not their frequency in the studied sedimentary sequences. The real Frequencies of facies transitions are written down in the matrix of observed facies transitions (Table 4). Therefore, when interpreting sedimentary successions it is useful to consider both statistically significant and real facies transitions in order to better understand their significance and real occurrence in the studied sedimentary record. The highest values of <P> and the positive entries of <D> were analyzed to determine the cyclic processes.The computed values of chi-square is lower than the limiting values at the 5% significance level (Tables 4g) this means that the null hypothesis is false, suggesting the deposition of sediments is not by Markovian process and non-cyclic arrangement of facies states in Kolhan Group.The facies relationship diagram (Fig 4) is constructed from the difference matrix results (Table 4e) Relationship diagrams showing upward transition of facies states of Kolhan group in Fig 4.

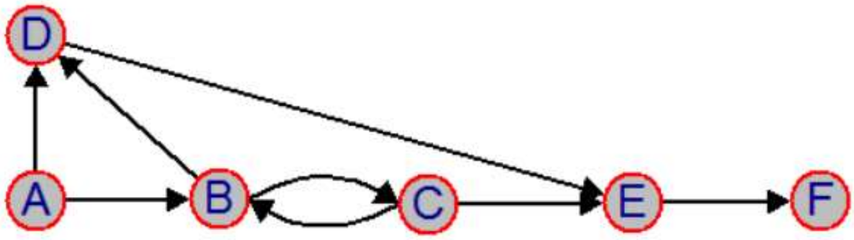

Fig. 4 Facies relationship diagrams showing upward transition of facies states of Kolhan group. A-GLA; B-GSD; C-SSD; D-PLSD; E-RDS; F-TLSD.

The preferred upward transition path for the lithofacies is GLA GSD SSD PLSD RSD TLSD.

The transition between GLA GSD, GSD SSD, SSD PLSD and RSD TLSD is non-Markovian and the lineage is non-repetitive in nature. The obvious aim of such approach was to detect and define cyclic relationships, if any. In the present case the cyclicity is absent or very weak. This information can greatly assist in environmental interpretation.

Entropy Analysis

Both E_{pre} and E_{post} are larger than 0.0 implies all six lithofacies (GLA, GSD, SSD, PLSD, RSD, and TLSD) overlies and also is overlain by more than one state (Hattori, 1976). E_{pre} and E_{post} are larger in number for GSD (Table 5), and it is deduced that the influx of pebbly sandstone into the basin was the most random event. For RSD and PLSD, $E_{pre} > E_{post}$. This relation indicates that rippled sandstones formed in shallow environment, though grainsize and hydrodynamic conditions may change the depth variations to some extent. So this type of findings need critical interpretation.

Large difference in E_{pre} and E_{post} and with $E_{pre} < E_{post}$ relationship in case of facies F indicates its strong dependence on its precursor which is visualised from the Markov metrics (Table 5 and Fig. 4). The depositional pattern in the TSLD facies is indicative of a low energy, suspension fall out during the waning phase of the sedimentation. In other words, these facies accumulated in environment located in the distal part of the basin in preference to other areas. The E (pre) and E (post) plots for coarse to medium-grained sandstone, interbedded fine grained sandstone/shale, shale fall far from the diagonal line (Fig 5). Energy regime related to the total entropy suggests that the shale in distal part of basin and is not of marine origin. The flow pattern overall changes from the deltaic environment to lacustrine environment. Fig 6 is well comparable with the type C cyclic pattern of Hattori, which signifies random lithologic series, as deduced independently by improved Markov process model. The cycles Kolhan basin belongs to the maximum entropy indicated by black dot (Fig.6) (Hattori, 1976).

	E(Post)	E(Pre)	R(Post)	R(Pre)
A	1.793	1.643	0.772	0.707885
B	2.128	2.4643	0.9167	1.061741
C	1	0.9010	0.4306	0.388195
D	1.921	2.1474	0.8279	0.925205
E	0.009	0.5167	0.002	0.22262
F	1.5	0.6937	0.6423	0.250474

Table 5: Matrix used to analyse entropy value of lithofacies

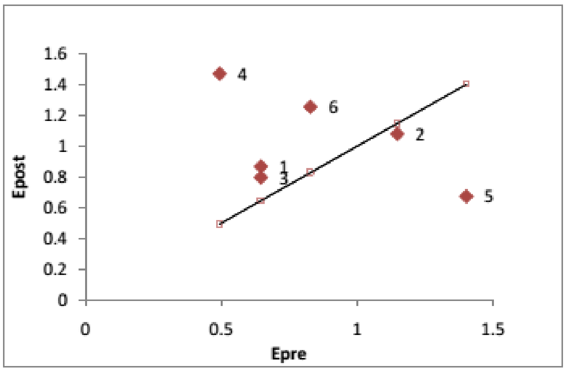

Fig. 5 Entropy set derived from Kolhan basin.1-GLA; 2-GSD; 3-SSD; 4-PLSD; 5-RDS; 6-TLSD.

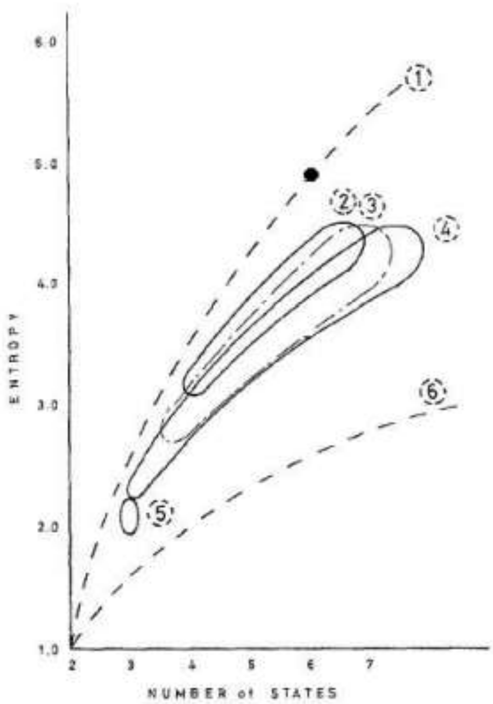

Fig. 6: Relationship between entropy and depositional environment of lithological sequences (after Hattori, 1976). 1-maximum entropy; 2-etropies for coal measure succession; 3-entropies for fluvial-alluvial successions; 4-entropies for neritic successions; 5-entropies for flysch sediments; 6-minimum entropy; Black dot indicate entropy of basin under study.

Conclusion

The application of the cross association analysis on the vertical sections shows that there is no significant correlation in between different lithofacies. The energy level of the deposition during the entire process shows a considerable fluctuation. Lack of correlation suggests lateral facies variation and existence of different environment at different places. The application of the first order Markov Chain analysis on the vertical, sections shows that there is a preferred fining upward transition path in the lithofacies. The operative geological processes were non-Markovian or independent in nature. The energy level during the entire process of sedimentation shows a considerable, fluctuation reflected by the Entropy analysis. Entropy analysis also proves type "C" cyclic pattern of Hattori, which signifies random lithologic series, as deduced independently by improved Markov process model. The cycles of Kolhan basin belongs to the maximum entropy. Asymmetric sequence can be well explained by sediment bypassing. The thinning upward sequences represent lacustrine deposits, while the thickening upward sequences represent point bar-sand flat deposits. Variation in layer thickness is suggestive of deposition by unsteady flow in a fluvial, regime within the channel. The flow was suddenly impeded, and as a result there was a quick fall in the energy of the solid-fluid system that resulted in rapid deposition. It appears that the GLA and the GSD facies represent the channel lag deposits of a braided river and the SSD, RSD, PLSD, and TLSD facies represent the portions of a fining upward sequence complex of a channel bar or possibly the longitudinal bartransverse bar-cross-channel bar complex in a fluvial environment. The result of this study described that the basin was non-cyclic, fining upward asymmetric sedimentary sequence of sandstone, shale with patches of carbonate.

Acknowledgement

The author is grateful to Department of Geology and Geophysics, IIT Kharagpur for giving the opportunity to carry out the research work.

References

Gingerich, P. (1969). Entropy in Markov chain and discrimination of cyclic pattern in lithologic successions. Journal of Sedimentary Research, 39(1), pp.330-497.

Hattori, I. (1976). Entropy in Markov chain and discrimination of cyclic pattern in lithologic successions. Mathematical. Geology, 8(4), pp.477-497.

Hota, R.N. (2000). Quantitative stratigraphic correlation-a case study from two borehole sections of Barakar Formation, Talchir coalfield, Jharkhand. Journal of Geological Society of India, 56, pp.517-524.

Miall, A.D. (1973). Markov chain analysis applied to an ancient alluvial plain succession. Sedimentology, 20, pp.347-364.

Powers, D. W. and Easterling, R. G. (1982). Improved methodology for using embedded Markov chains to describe cyclical sediments. Journal of Sedimentary Petrology, 52, pp.913-923.

Saha, A.K. (1994). Crustal evolution of Singhbhum North Orissa, Eastern India. Memoir Geological Society of India, 27, pp.341.

Selley, R.C. (1970). Ancient Sedimentary Environments and Their Sub-surface Diagnosis (2nd Edn). Routledge, Taylor and Francis, London, 320p.

Selley, R.C. (1976). Subsurface environmental analysis of North Sea sediments. Bulletin American Association of Petroleum Geology, 60, pp.184-195.

Tewari, R.C., Singh, D.P., Khan, Z.A. (2009). Application of Markov chain and Entropy analysis to lithologic succession – an example from early Permian Barakar Formation, Bellampalli coalfield, Andhra Pradesh, India. Journal Earth System Science, 18 (5) (2009), pp.583–596.

Walker, R.G. (1963). Distinctive types of ripple drift cross-lamination. Sedimentology, 2, pp.173-188.