Application of Markov chain and entropy analysis to lithologic succession – an example from the early Permian Barakar Formation, Bellampalli coalfield, Andhra Pradesh, India

Ram Chandra Tewari[1,*], D P Singh[2,**] and Z A Khan[3,†]

[1]*Department of Geology, Sri J.N.P.G. College, Lucknow 226 001, India.*
[2]*SMEC India Pvt. Ltd., 5th Floor, Tower C, DLF Building 8, Cyber City, Phase II, Gurgaon 122 002, Haryana, India.*
[3]*Directorate of Geology & Mining, Khanij Bhawan, Lucknow 226 001, India.*
[*]*e-mail: ram_tewari@yahoo.com*
[**]*e-mail: dp.singh@smecindia.com*
[†]*e-mail: zahid26aug@yahoo.co.in*

A statistical approach by a modified Markov process model and entropy function is used to prove that the early Permian Barakar Formation of the Bellampalli coalfield developed distinct cyclicities during deposition. From results, the transition path of lithological states typical for the Bellampalli basin is as: coarse to medium-grained sandstone → interbedded fine-grained sandstone/shale → shale → coal and again shale. The majority of cycles are symmetrical but asymmetrical cycles are present as well. The chi-square stationarity test implies that these cycles are stationary in space and time. The cycles are interpreted in terms of in-channel, point bar and overbank facies association in a fluvial system. The randomness in the occurrence of facies within a cycle is evaluated in terms of entropy, which can be calculated from the Markov matrices. Two types of entropies are calculated for every facies state; entropy after deposition E(post) and entropy before deposition E(pre), which together form entropy set; the entropy for the whole system is also calculated. These values are plotted and compared with Hattori's idealized plots, which indicate that the sequence is essentially a symmetrical cycle (type-B of Hattroi).

The symmetrical cyclical deposition of early Permian Barakar Formation is explained by the lateral migration of stream channels in response to varying discharge and rate of deposition across the alluvial plain. In addition, the fining upward cycles in the upper part enclosing thick beds of fine clastics, as well as coal may represent differential subsidence of depositional basin.

1. Introduction

Vertical variations of lithofacies within a given sequence play an important role in the recognition of depositional environment and their lateral dispersal. It is of particular importance in the context of the widely accepted law of the correlation of facies as proposed by Walther (1893) and elaborated by Middleton (1973), Reading (1991) and recently by Sengupta (2006). According to this law, only those facies can be superimposed which can be observed beside each other at a given time. Importantly, only a gradual transition from one facies to another implies that the two facies represent environments that once were adjacent laterally.

Observations of outcrop and borehole sections of the coal bearing formation indicate a systematic repetition of fining upward sequences. Although,

Keywords. Barakar coal measures; entropy functions; Markov chain; cyclicity.

J. Earth Syst. Sci. **118**, No. 5, October 2009, pp. 583–596

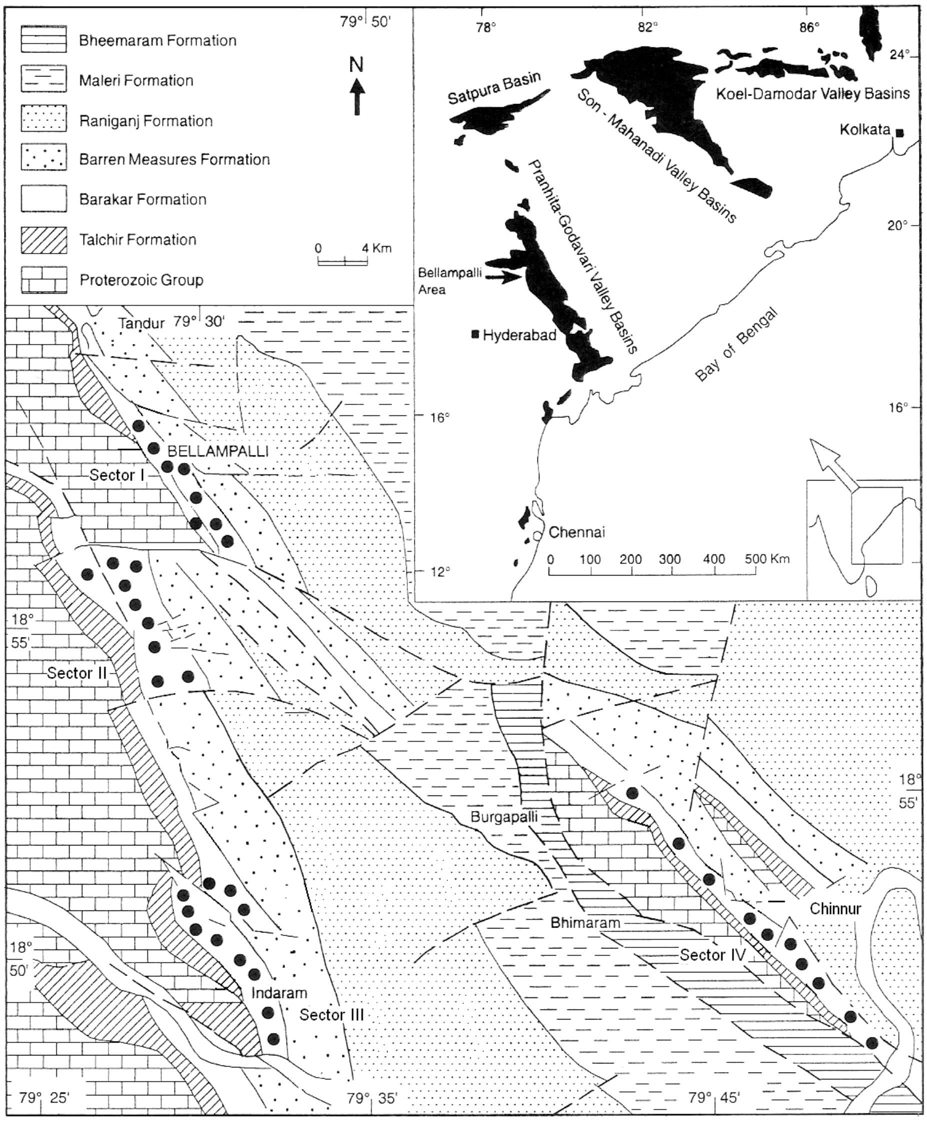

Figure 1. Geological map of Bellampalli coalfield (modified after Raja Rao 1982), showing location of borehole logs.

individual cycles are present, the scarcity of comprehensive exposures due to weathering makes it difficult to determine regional distribution of cyclicity. Hence, information obtained from 38 boreholes from Bellampalli coalfield, Andhra Pradesh penetrating the entire Barakar section providing a precise record of lithologic transitions has been utilized for various statistical analyses. In order to determine the depositional architecture and its regional variations, a check of the results obtained so far (Tewari and Singh 2008) by mathematical means seemed desirable. The vast amount of data obtained through counting of lithologic transitions of the borehole logs justify the application of Markov chain and entropy functions.

The objectives of the present study are:

- to deduce lithologic transitions in vertical sequences through space and time;
- to analyze the significance of succession and sediment cyclicity; and
- to compare cyclic characters of early Permian sediments with those of eastern and central India Gondwana basins.

2. General geology and nature of data

The Bellampalli coalfield occupies the northern part of Pranhita–Godavari Gondwana basin (figure 1, inset). Raja Rao (1982) briefly summarized the geological details of this area. Recently, Singh and Tewari (2007) and Tewari and Singh (2008) have carried out detailed paleochannel and paleocurrent analysis of the Gondwana sediments of this area but a detailed lithofacies analysis and their relationships is still lacking. The lower Gondwana rocks of the area comprise in ascending

Table 1. *Gondwana stratigraphy of Bellampalli coalfield, Godavari valley basin, Andhra Pradesh (modified after Raja Rao 1982).*

Age	Formation	Lithology
Middle to Upper Triassic	Bheemaram	Coarse-grained sandstone with clay galls and few clay intercalations.
Lower to Middle Triassic	Maleri	Soft red mudstone with calcareous bands of sandstone.
Upper Permian	Raniganj	Coarse-grained ferruginous sandstone with clay clasts and pebbles, cherty siltstone and pebble beds.
Middle Permian	Barren Measures	Medium to coarse-grained greenish grey to grey white feldspathic sandstone with subordinate variegated clay and micaceous siltstones.
Lower Permian	Barakar	Coarse-grained white sandstones with lenses of conglomerates, subordinate shale/clay and coal seams.
Permo-Carboniferous	Talchir	Fine grained sandstone, splintery green shale/clay, khakhi coloured clay, pebble beds and diamictite.
- - - - - - - - - - - - - - - -	Unconformity	- -
Upper Proterozoic	Sullavai Group	Medium to coarse-grained white to brick red sandstone, quartzitic at places and mottled shale.
- - - - - - - - - - - - - - - -	Unconformity	- -
Lower Proterozoic	Pakhal Group	Greyish white to buff quartzites, grey shales, phyllites, dolomites and marbles.
- - - - - - - - - - - - - - - -	Unconformity	- -
Archeans		Granites, banded gneisses, biotite gneisses, hornblende gneisses, quartz-magnetite-schist, biotite schist, pegmatite veins.

order, a basal sequence of diamictite, pebble bed and other glacigene deposits. These are succeeded conformably by the main coal-bearing Barakar Formation, which is about 300 m in thickness, consisting of coarse-grained sandstone, argillaceous and arenaceous shale and coal (table 1). The basal sandstone of this unit forms an extensive cover of current bedded, braided stream deposits. Commonly, they overlap the glacigene sediments, indicating rapid expansion of the basin area, as described elsewhere (Tewari and Singh 2008). In view of the non-availability of good vertical sections in the area, the vertical lithofacies relationship of the Barakar Formation could not be established. Since the Barakar strata have very little outcrop in the area studied, the present investigation is based entirely on the data derived from borehole logs. A total of 38 borehole logs, from localities scattered throughout the area (figure 1) is used. The drilled thickness of the strata in these boreholes varies from 50–250 m.

Of particular importance is the noting transition of one lithology to another in a stratigraphic section. 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 four lithofacies, which are distinctly marked in each borehole log as well as in outcrop sections, are used in this paper. To analyze cyclic characters through space and time, the lithofacies transitions are analyzed separately in each borehole log, and by pooling the data for four sectors as well as for the entire area. The four facies are:

- **Facies A:** Coarse to medium-grained sandstone.
- **Facies B:** Interbedded fine-grained sandstone and shale.
- **Facies C:** Carbonaceous and argillaceous shale.
- **Facies D:** Coal.

3. Analytical procedure

The concept of cycles of sedimentation implies that the initial state or lithology determine to some extent the subsequent state or lithology. This led Vistelius (1949) to propose the use of Markov chain as an analytical tool in the study of vertical lithofacies relationship in stratigraphic sequences. However, the approach is useful in that it can often point out subtle relationships in the stratigraphic succession that would not otherwise be noticed. The literature on Markov chain and entropy analyses is now rapidly growing. Test cases have been published by Khan and Casshyap (1982); Tewari and Casshyap (1983); Mack and James (1986); Khan (1997); Sharma *et al* (2001); Hota and Maejima (2004); Khan and Tewari (2007) and others.

The method described in this section is based mainly on those of Gingerich (1969); Power and

Easterling (1982) and Davis (2002). In addition, the nature of cyclic order of a sequence is studied using entropy concept following Hattori (1976).

3.1 *Embedded Markov chain analysis*

The modified Markov process model after Power and Easterling (1982) used in this study incorporates following successive steps:

- Structuring of one step embedded tally count matrix (fij), where i, j corresponds to row and column number. It will be noticed that where $i = j$, zeros are present 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.
- Estimated expected frequency after Goodman's model (1968) of quasi-independence given by $Eij = \widehat{aibj}$ (where $i \neq j$) derived by using an iterative procedure till ai and bj attain an arbitrary constant (Power and Easterling 1982, p. 916). The computational procedure is as follows:

Let $E(nij)$ denote the expected value of the number of transitions from state i to state j. Then the model proposed by Goodman, termed 'quasi-independence', is:

$$Enij = aibj, \quad i \neq j,$$
$$= 0, \quad i = j.$$

Estimating the parameters, ai and bj, i, $j = 1, 2, 3, \ldots, m$, requires an iterative scheme which is given as follows:

First iteration:

$$ai^{(I)} = n_{i+}/(m-1), \quad i = 1, 2, 3, \ldots, m,$$

$$bj^{(I)} = n_{+j} \Big/ \sum_{i \neq j} ai, \quad j = 1, 2, 3, \ldots, m,$$

Ith iteration:

$$ai^{(I)} = n_{i+} \Big/ \sum_{j \neq i} bj^{I-1}, \quad i = 1, 2, 3, \ldots, m,$$

$$bj^{(I)} = n_{+j} \Big/ \sum_{i \neq j} ai^{I}, \quad j = 1, 2, 3, \ldots, m,$$

where n_{i+} and n_{+j} are the row i and column j totals, respectively.

Iteration may be continued until some specified accuracy (1% in the present case) is obtained. That is, iteration is continued until

$$ai^{(I)} - ai^{(I-1)} < 0.01ai^{(I)}, \quad \text{for } i = 1, 2, \ldots, m$$

and

$$bj^{(I)} - bj^{(I-1)} < 0.01bj^{(I)}, \quad \text{for } j = 1, 2, \ldots, m.$$

Let $\hat{a}i$ and $\hat{b}j$ denotes the final values of $ai^{(I)}$ and $bj^{(I)}$, then the estimated expected frequencies of quasi-independence are given by $Eij = \widehat{aibj}$, for $i \neq j$.

- Using the values of fij and Eij in the expression $\sum_{i=1}^{n} \sum_{j=1}^{n} (fij - Eij)^2/Eij$ for χ^2 yields a statistics which is distributed as a chi-squared variable with $(n-1)^2 - n$ degrees of freedom. The larger the χ^2 value for a given value of n, the stronger the evidence in favour of the Markovian model of lithologic transition, i.e., for the presence of cyclicity.
- Construction of normalized difference matrix, here symbolized as Zij. This provides a framework for identifying large difference (+values) between observed (fij) and expected transition frequencies (Eij).

The entire computation was performed on desk calculator.

3.2 *Structure of transition count matrix*

Lithologic transitions, to test the presence or absence of Markov property or lack of it are investigated at sector and coalfield levels separately using data from the available 38 borehole logs. Figure 2 illustrates portions of the lithological sequence from a borehole log of the study area. The transition 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 (see step 1, section 3.1). Tally count matrix based on borehole logs numbering 8 to 9 for each sector is structured. Subsequently, data for all 38-borehole logs are added and a bulk matrix is structured at coalfield level. Data for the bulk tally matrix at sector level and coalfield level are computed separately following the procedure stated above.

3.3 *Entropy analysis*

The starting point in entropy analysis is (a) transition tally count matrix (fij). This is a 2-D

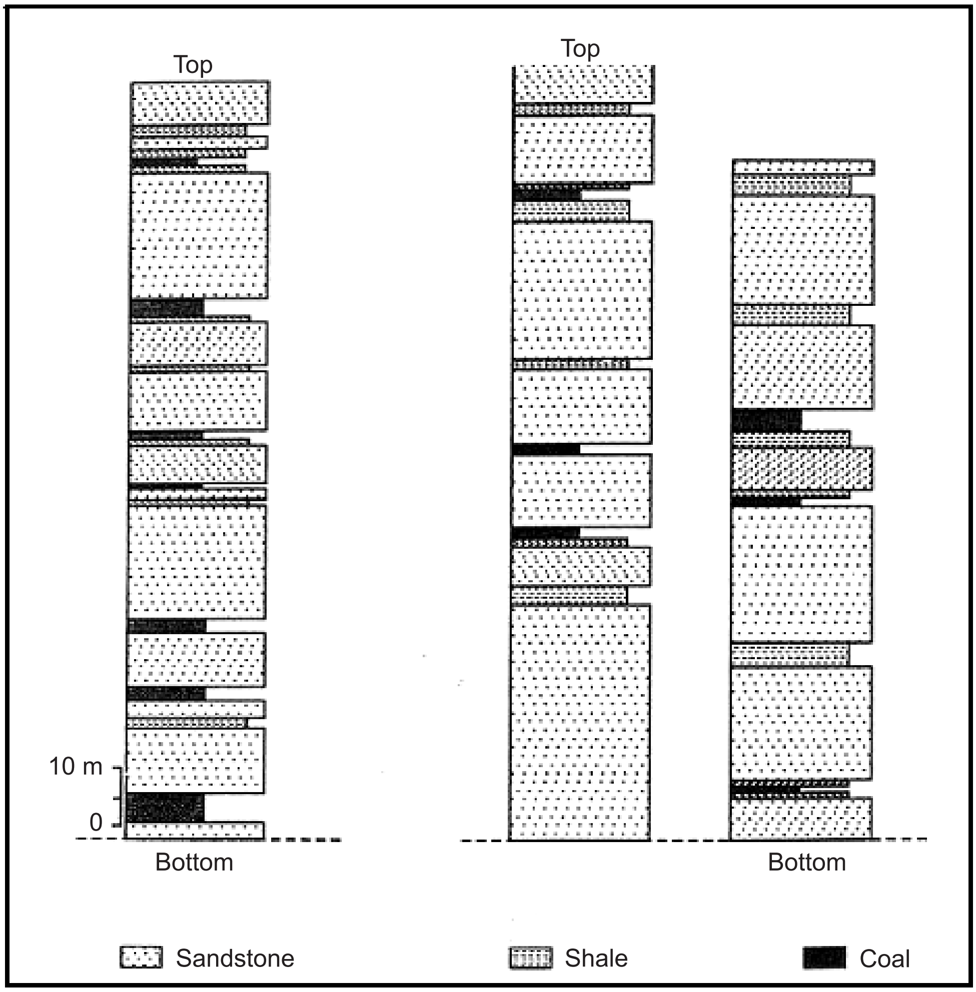

Figure 2. Subsurface Barakar stratigraphy of Bellampalli coalfield reproduced from borehole logs.

array, which tabulates the number of times that frequency of all possible vertical lithologic transitions occurring in a given stratigraphic succession, (b) from the transition count matrix (fij) two probability matrices may be derived. The first is an upward probability matrix composed of pij, which gives the actual probabilities of the transition occurring in the given section and is calculated as fij/n_{i+}, where n_{i+} is the row i total. In the pij matrix, the row total sums to unity. From this matrix E(post) (i.e., entropy after deposition) for each lithological state has been calculated using relationship $E(\text{post}) = -\sum_{j=1}^{n} \sum pij \log_2 pij$. The second matrix, containing element qji which represents the probability of the given transition being preceded by any other transition and given by fij/n_{+j}, where n_{+j} is the column total. The column totals in the qji matrix sums to unity. E(pre) (i.e., entropy before deposition) can be calculated using relationship $E(\text{pre}) = -\sum_{j=1}^{n} \sum qji \log_2 qji$. E(pre) and E(post) indicate the variety of lithological transitions which immediately lead into, and follow from, that state and they serve as a useful supplement to the transition probability information. By plotting E(post) against E(pre) for each lithology one can make some interpretation of the styles of cyclicity and the way in which cycles are truncated. Hattori (1976) drew a number of diagrams of the distribution of E(post) *versus* E(pre) for idealized, truncated, symmetrical and asymmetrical lithologic succession. Hattori's (1976) normalized entropies were also calculated as $R(\text{pre}) = E(\text{pre})/E(\text{max})$ and $R(\text{post}) = E(\text{post})/E(\text{max})$; where $E(\text{max}) = -\log_2 1/(n-1)$, which denotes the maximum entropy possible in a system where n states operate. This concept allows comparisons between states in different system, regardless of the number of state variables selected in each system. Apart from entropies with respect to individual sets, the entropy of the whole sedimentation unit can be calculated as $E(\text{system}) = -\sum_{i=1}^{n} \sum_{j=1}^{n} rij \log_2 rij$, $rij = fij/n\,{+}{+}$: where fij are entries in the tally matrix, $n\,{+}{+}$ is number of states $= fij$, which can be used for deciphering

Table 2. *Transition count, expected cell value, normalized difference and chi-square matrices of Barakar Formation (sector I), Bellampalli coalfield.*

Transition count matrix (fij)

	A	B	C	D
A	00	08	113	35
B	08	00	03	01
C	78	03	00	65
D	59	00	79	00

Expected cell value matrix (Eij)

	A	B	C	D
A	00	4.47	79.39	71.56
B	4.49	00	3.93	3.55
C	79.09	3.91	00	62.49
D	71.65	3.54	62.81	00

Normalized difference matrix (Zij)

	A	B	C	D
A	00	+1.59	+3.79	−4.34
B	+1.63	00	−0.48	−1.36
C	−0.16	−0.51	00	+0.28
D	−1.51	−1.90	+2.06	00

Chi-square matrix

	A	B	C	D
A	00	2.79	14.23	18.68
B	2.74	00	0.22	1.83
C	0.02	0.21	00	0.10
D	2.23	3.54	4.17	00

χ^2	Degree of freedom	Limiting value at 99.5%
50.76	5	22

Table 3. *Transition count, expected cell value, normalized difference and chi-square matrices of Barakar Formation (sector II), Bellampalli coalfield.*

Transition count matrix (fij)

	A	B	C	D
A	00	16	94	78
B	06	00	07	07
C	74	04	00	101
D	102	00	01	85

Expected cell value matrix (Eij)

	A	B	C	D
A	00	6.89	86.32	94.54
B	6.68	00	6.28	6.88
C	86.41	6.3	00	86.41
D	94.54	6.89	86.32	00

Normalized difference matrix (Zij)

	A	B	C	D
A	00	+3.48	+0.83	−1.70
B	−0.34	00	+0.29	+0.05
C	−0.133	−0.92	00	+1.57
D	+0.77	−2.25	−0.14	00

Chi-square matrix

	A	B	C	D
A	00	12.05	0.68	2.89
B	0.11	00	0.08	0.002
C	1.78	0.84	00	2.46
D	0.59	5.04	0.02	00

χ^2	Degree of freedom	Limiting value at 99.5%
26.54	5	22

Table 4. *Transition count, expected cell value, normalized difference and chi-square matrices of Barakar Formation (sector III), Bellampalli coalfield.*

Transition count matrix (fij)

	A	B	C	D
A	00	08	23	15
B	06	00	06	02
C	18	03	00	46
D	22	03	38	00

Expected cell value matrix (Eij)

	A	B	C	D
A	00	3.06	22.99	19.87
B	3.06	00	5.86	5.07
C	22.92	5.85	00	37.98
D	19.86	5.07	30.06	00

Normalized difference matrix (Zij)

	A	B	C	D
A	00	+2.81	+0.002	−1.09
B	+1.68	00	+0.06	−1.36
C	−1.04	−1.18	00	+1.30
D	−0.48	−0.92	−0.009	00

Chi-square matrix

	A	B	C	D
A	00	7.98	00	1,19
B	2.82	00	0.003	1.86
C	1.06	1.39	00	1.69
D	0.23	0.85	00	00

χ^2	Degree of freedom	Limiting value at 99.5%
29.07	5	22

Table 5. *Transition count, expected cell value, normalized difference and chi-square matrices of Barakar Formation (sector IV), Bellampalli coalfield.*

Transition count matrix (fij)

	A	B	C	D
A	00	04	35	14
B	02	00	02	00
C	30	00	00	32
D	20	00	27	00

Expected cell value matrix (Eij)

	A	B	C	D
A	00	1.15	32.84	18.86
B	1.21	00	1.74	0.99
C	32.73	1.66	00	27.03
D	19.05	0.96	27.39	00

Normalized difference matrix (Zij)

	A	B	C	D
A	00	+2.66	+0.38	−1.12
B	+0.72	00	+0.20	−1.00
C	−0.48	−1.28	00	+0.96
D	−0.22	−0.98	−0.07	00

Chi-square matrix

	A	B	C	D
A	00	7.06	0.14	1.25
B	0.52	00	0.04	0.99
C	0.23	1.66	00	0.91
D	0.05	0.96	0.006	00

χ^2	Degree of freedom	Limiting value at 99.5%
33.82	5	22

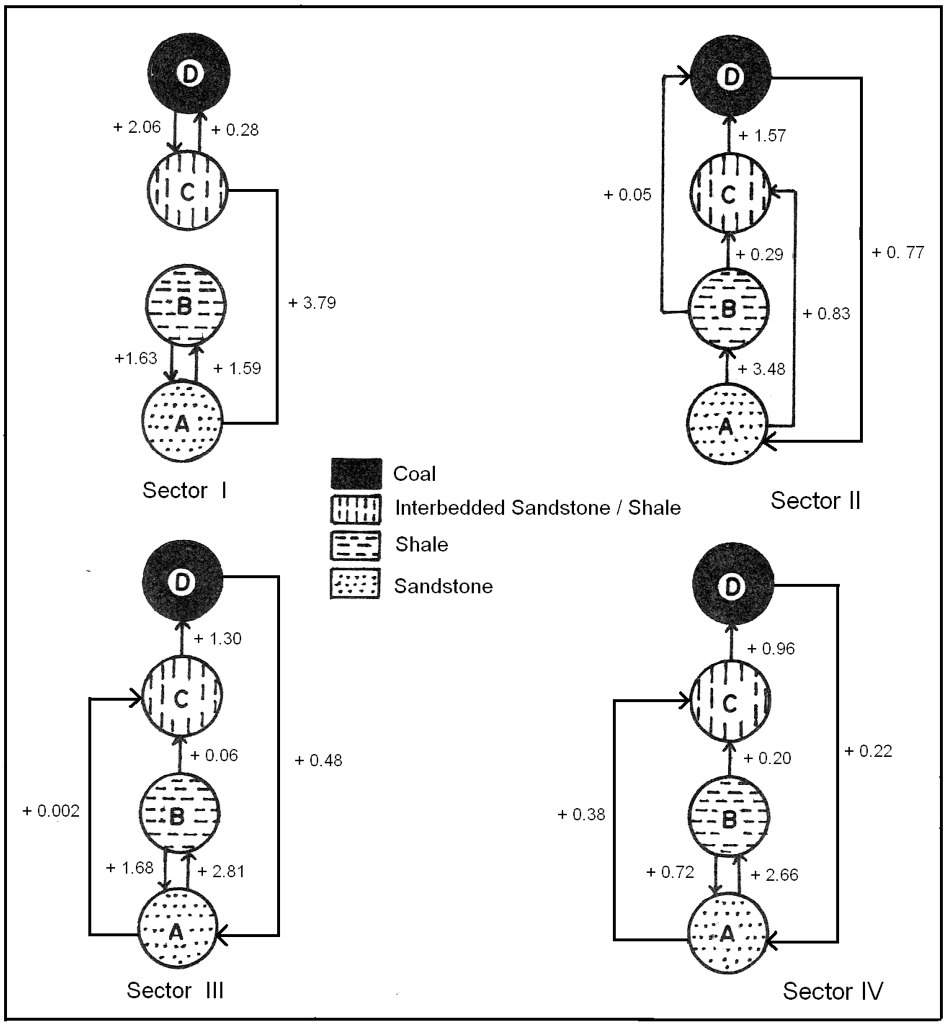

Figure 3. Lithofacies relationship diagram showing upward transitions (based on positive values in Zij matrix) in the four sectors of the Barakar Formation.

the overall depositional environment of cyclical units. The E(system) can take a value between $-\log_2 1/n$ and $-\log_2 1/n(n-1)$, where n is the number of states.

Data for the tally count matrix at sector level and coalfield level are used to compute separately E(pre), E(post) and E(system) following the procedure outlined above.

4. Quantitative results

4.1 *Lithologic transition at sector level*

The bulk transition counts matrices (fij), expected cell value matrices (Eij), normalized difference matrices (Zij), and chi-square matrices separately for four sectors of the Bellampalli coalfield are listed in tables 2, 3, 4, and 5. Indeed, there is a strong tendency of Markovian property or cyclicity in the Barakar strata at sector level at given degrees of freedom and 99.5% confidence level. In figure 3, A, B, C and D shows Markov transition diagram based on positive values of normalized difference matrix (Zij). Highest positive values of Zij matrix link lithologic states distinctly resulting in a strong transition path for lithologic sequence that can be derived as follows:

Sector I: Coarse to medium-grained sandstone (facies A) → interbedded fine-grained sandstone and shale (facies B) → shale (facies C) → coal (facies D) → shale (facies C).

Sectors II, III, IV: Coarse to medium-grained sandstone (facies A) → interbedded fine-grained sandstone and shale (facies B) → shale (facies C) → coal (facies D) → coarse to medium-grained sandstone (facies A).

Thus transition path is typical of the coal bearing Barakar strata and by and large display a progressive fining of particle size from coarse-grained sandstone through fine-grained sandstone/shale

Table 6. *Transition tally count, upward transition, downward transition, independent trial matrices and entropy values for the Barakar Formation (sector I), Bellampalli coalfield.*

Transition count matrix (f_{ij})

	A	B	C	D
A	00	08	113	35
B	08	00	03	01
C	78	03	00	65
D	59	00	79	00

Upward transition matrix (p_{ij})

	A	B	C	D
A	00	0.05	0.72	0.22
B	0.67	00	0.25	0.08
C	0.53	0.02	00	0.45
D	0.43	00	0.57	00

Downward transition matrix (q_{ij})

	A	B	C	D
A	00	0.73	0.58	0.35
B	0.06	00	0.02	0.009
C	0.54	0.27	00	0.64
D	0.41	00	0.41	00

Independent trial matrix (r_{ij})

	A	B	C	D
A	00	0.02	0.25	0.08
B	0.02	00	0.007	0.002
C	0.17	0.007	00	0.14
D	0.13	00	0.17	00

	E(pre)	E(post)	R(pre)	R(post)
A	1.318	1.038	0.832	0.655
B	0.417	1.175	0.263	0.743
C	1.402	1.117	0.885	0.705
D	1.055	0.986	0.665	0.622

E(max) = 1.585, E(system) = 2.787

Table 7. *Transition tally count, upward transition, downward transition, independent trial matrices and entropy values for the Barakar Formation (sector II), Bellampalli coalfield.*

Transition count matrix (f_{ij})

	A	B	C	D
A	00	16	94	78
B	06	00	07	07
C	74	04	00	101
D	102	01	186	186

Upward transition matrix (p_{ij})

	A	B	C	D
A	00	0.09	0.50	0.41
B	0.30	00	0.35	0.35
C	0.41	0.02	00	0.56
D	0.54	0.005	0.45	00

Downward transition matrix (q_{ij})

	A	B	C	D
A	00	0.76	0.51	0.42
B	0.033	00	0.04	0.40
C	0.41	0.19	00	0.54
D	0.56	0.05	0.46	00

Independent trial matrix (r_{ij})

	A	B	C	D
A	00	0.03	0.61	0.14
B	0.01	00	0.01	0.01
C	0.13	0.007	00	0.18
D	0.18	0.001	0.15	00

	E(pre)	E(post)	R(pre)	R(post)
A	1.322	1.339	0.834	0.845
B	0.390	1.578	0.246	0.996
C	1.463	1.108	0.923	0.699
D	1.200	1.037	0.757	0.654

E(max) = 1.585, E(system) = 2.916

to shale, then coal. The lithologic transitions are broadly alike in the sectors II, III and IV. However, in sector I the coal (facies D) has greater probability to be followed by shale (facies C) than coarse-grained sandstone (facies A). Markov diagram, which further indicates that each cycle, generally speaking is asymmetrical (ABCD–ABCD) beginning with erosional surface and coarse-grained sandstone. But symmetrical cycles ABCD–CBA also occurs locally in sector I. Individual cycles vary in thickness from a couple of meters to a few tens of meters.

The lithologic transitions deduced here closely resembles the cyclical sequences of other late Paleozoic coal measures (Read 1969; Casshyap 1975; Casshyap *et al* 1987) including the Permian coal measures of lower Gondwana of India (Casshyap and Tewari 1984; Tewari 1997; Hota and Maejima 2004; Khan and Tewari 2007).

4.2 *Entropy analysis at sector level*

The computed entropies E(pre) and E(post), normalized entropies R(pre) and R(post) of each lithologic state are subequal to equal (tables 6, 7, 8 and 9), implying that the deposition of these lithologies is not a random event. For coarse to medium-grained sandstone (facies A) $E(\text{pre}) \geq E(\text{post})$, implies with a high probability of this state passing up into interbedded fine grained sandstone and shale (facies B) but may occur after different lithologic state, as is also recognized in the stratigraphic sections (figure 2). By contrast, the remaining lithological states $E(\text{post}) \leq E(\text{pre})$ indicate that the deposition of each of these lithologies is strongly influenced by the preceding state. This relationship supports in a statistical way the otherwise geologically obvious conclusion that the deposition of these lithologies depends largely on specific environments.

The plot of the E(pre) and E(post) values for each lithological state is given in the figure 5 of Hattori's diagrams, which most closely (though not exactly) follows that expected for a symmetrical cyclic sequence (type-B category). Indeed, this cyclical pattern for the given Barakar Formation is similar to that reported from other areas based on the field study (Tewari 1997, 2005).

Table 8. *Transition count, upward transition, downward transition, independent trial matrices and entropy values for the Barakar Formation (sector III), Bellampalli coalfield.*

Transition count matrix (f_{ij})

	A	B	C	D
A	00	08	23	15
B	06	00	06	02
C	18	03	00	46
D	22	03	38	00

Upward transition matrix (p_{ij})

	A	B	C	D
A	00	0.17	0.50	0.32
B	0.43	00	0.43	0.14
C	0.27	0.04	00	0.69
D	0.34	0.05	0.60	00

Downward transition matrix (q_{ij})

	A	B	C	D
A	00	0.57	0.34	0.24
B	0.13	00	0.09	0.03
C	0.39	0.21	00	0.73
D	0.48	0.21	0.57	00

Independent trial matrix (r_{ij})

	A	B	C	D
A	00	0.04	0.12	0.08
B	0.03	00	0.03	0.01
C	0.09	0.02	00	0.24
D	0.12	0.02	0.20	00

	E(pre)	E(post)	R(pre)	R(post)
A	1.486	1.460	0.938	0.921
B	0.847	1.444	0.534	0.911
C	1.334	1.065	0.842	0.672
D	1.443	1.188	0.910	0.749

E(max) = 1.585, E(system) = 3.078

Table 9. *Transition count, upward transition, downward transition, independent trial matrices and entropy values for the Barakar Formation (sector IV), Bellampalli coalfield.*

Transition count matrix (f_{ij})

	A	B	C	D
A	00	04	35	14
B	02	00	02	00
C	30	00	00	32
D	20	00	27	00

Upward transition matrix (p_{ij})

	A	B	C	D
A	00	0.08	0.66	0.40
B	0.50	00	0.50	00
C	0.48	00	00	0.52
D	0.43	00	0.57	00

Downward transition matrix (q_{ij})

	A	B	C	D
A	00	01	0.55	0.30
B	0.04	00	0.03	00
C	0.58	00	00	0.70
D	0.38	00	0.42	00

Independent trial matrix (r_{ij})

	A	B	C	D
A	00	0.02	0.21	0.08
B	0.01	00	0.01	00
C	0.18	00	00	0.19
D	0.12	00	0.16	00

	E(pre)	E(post)	R(pre)	R(post)
A	0.996	1.216	0.628	0.767
B	0.338	1.000	0.213	0.631
C	0.816	0.999	0.515	0.630
D	1.056	0.986	0.666	0.622

E(max) = 1.585, E(system) = 2.701

4.3 *Lithologic transition at coalfield level*

Data for deducing Markov property in individual sectors are lumped together and processed at coalfield (basin) level (table 10). Chi-square value calculated by using the formula referred to earlier is significant at an appropriate degree of freedom at 99.5% level of confidence. Presence of Markov property is clearly indicated in the Barakar strata of Bellampalli coalfield. Markov transition diagram in figure 4 gives only those values of normalized difference matrix (Zij) for which the corresponding entries show positive differences. The positive value in Zij matrix links the lithologic states distinctly, and a strong preferred upward transition path for lithologic changes that can be derived is as:

Coarse to medium-grained sandstone (facies A) → interbedded fine-grained sandstone and shale (facies B) → shale (facies C) → coal (facies D) → shale (facies C).

Cyclicity deduced in the Barakar strata of Bellampalli area is symmetrical (figure 4), as is indicated by the presence of shale on top of coal; its environmental significance is deferred to a later part of the paper. The Barakar cycles as deduced here shows similarities with those of Son–Mahanadi and Koel–Damodar Gondwana basins of eastern-central India (Casshyap and Tewari 1984; Tewari 1997; Hota and Meijima 2004; Khan and Tewari 2007). Evidently, the results imply recurrence of the corresponding depositional environments both across the basin and through time, following Walther's (1893) law of facies.

4.4 *Entropy analysis at coalfield level*

The bulk transition count matrix (fij) of Markov analysis given in table 10 is used to compute upward transition matrix (pij), downward transition matrix (qij), and independent trail matrix (rij). These matrices are subsequently used to calculate entropy before deposition E(pre), entropy after deposition E(post), normalized entropies R(pre) and R(post) and entropy for the system E(system) for the four lithofacies separately for each sector and for the entire Barakar Formation (table 11) using formula referred to above. The variations in pre- and post-depositional entropy values suggest variable degree of dependency of

Table 10. *Pooled transition count, expected cell values, normalized difference, and chi-square matrices of Barakar Formation in Bellampalli coalfield (based on 38 borehole logs).*

Transition count matrix (fij)				
	A	B	C	D
A	00	36	265	142
B	22	00	18	10
C	200	10	00	214
D	203	04	229	00
Expected cell value matrix (E_{ij})				
	A	B	C	D
A	00	16.46	221.50	209.52
B	16.58	00	17.40	16.46
C	221.50	17.28	00	219.93
D	204.92	15.99	215.09	00
Normalized difference matrix (Z_{ij})				
	A	B	C	D
A	00	+4.81	+2.92	−4.67
B	+1.33	00	+0.14	−1.59
C	−4.63	−1.75	00	+1.63
D	−0.13	−3.0	+0.95	00
Chi-square matrix				
	A	B	C	D
A	00	23.20	8.54	21.76
B	1.77	00	0.02	2.54
C	2.09	3.07	00	2.63
D	0.2	8.99	0.9	00
χ^2	Degree of freedom		Limiting value at 99.5%	
75.71	5		22	

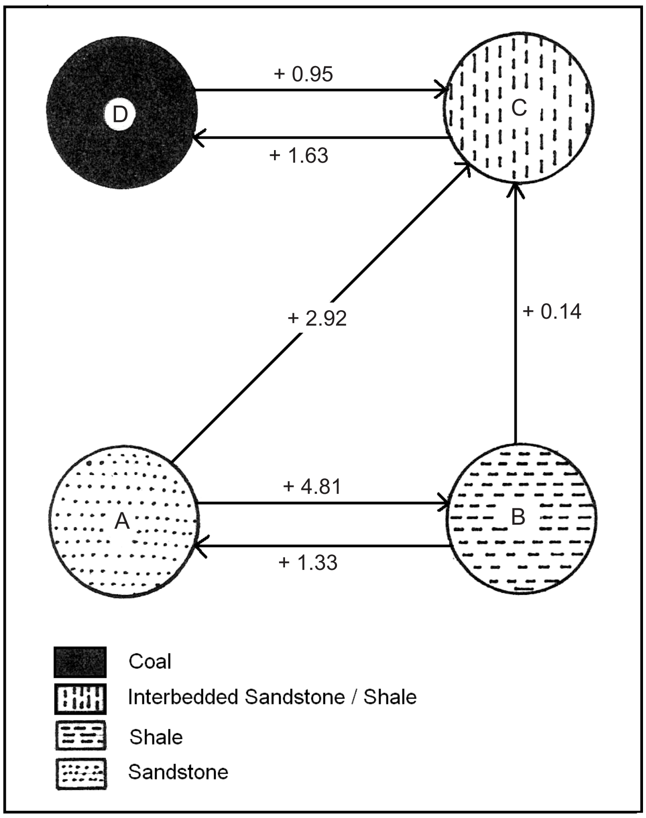

Figure 4. Lithofacies relationship diagram showing upward transitions (based on positive values in Zij matrix) for the bulk transition data of the Barakar Formation.

lithofacies on precursor and influence on successor during Barakar sedimentation. The computed values of E(pre) and E(post) and normalized entropies are almost sub-equal suggesting that the deposition of these lithofacies was not a random event (figure 5). However, if $E(\text{post}) > E(\text{pre})$, the deposition of given lithofacies is strongly influenced by the preceding lithofacies (Hattori 1976). The calculated value of E(system) falls well within the zone of 'Fluvial system' (figure 6) delineated by Hattroi (1976) corroborating the views expressed by early workers (see Khan and Tewari 2007, for references). Among the entropy sets those for facies A and facies C are located along the diagonal line; there is a fair probability that both lithologies tend to occur as a symmetrical cycle, that is, alternation in the succession. The entropy with respect to facies B (interbedded fine-grained sandstone and shale) deviates from the general distribution (figure 6). It is difficult to interpret clearly the cause responsible for the phenomenon, however, this may account for the allocyclic bahaviour of facies B (Beerbower 1964).

The E(pre) and E(post) plots for coarse to medium-grained sandstone, interbedded fine-grained sandstone/shale, shale and coal fall almost linearly or close to a diagonal line (figure 6), comparing well with the type 'B' cyclic pattern of Hattori, which signifies symmetrical cycle, as deduced independently by improved Markov process model.

5. Stationarity of cyclic sequence

In the analysis of Markov chain, it is assumed that Markov matrices are the result of a process that is stationary in time and space. The term 'stationary' implies that the transition probabilities are constant through time or space (Harbaugh and Bonham-Carter 1970, p. 122). Stationarity in a sequence can be verified by using following chi-square statistics after Anderson and Goodman (1957), which has been modified for computational convenience by Harbaugh and Bonham-Carter (1970, p. 124) as:

$$= 2\sum_{t}^{T}\sum_{ij}^{n} fij(t)\cdot \log_e pij(t)/qij$$

where $T = 1, 2, 3, \ldots, T$, giving the number of sequences tested against each other, $fij(t)$ and $pij(t)$ are the tally count and transition probability matrix values, respectively, for each sequence

Table 11. *Pooled transition count, upward transition, downward transition, and independent trials matrices, and entropy values for the Barakar Formation, Bellampalli coalfield.*

Transition count matrix (f_{ij})				
	A	B	C	D
A	00	36	265	142
B	02	00	18	10
C	200	10	00	44
D	203	04	229	00
Upward transition matrix (p_{ij})				
	A	B	C	D
A	00	0.08	0.60	0.32
B	0.44	00	0.36	0.20
C	0.44	0.02	00	0.54
D	0.47	0.009	0.53	00
Downward transition matrix (q_{ij})				
	A	B	C	D
A	00	0.72	0.52	0.36
B	0.05	00	0.04	0.03
C	0.47	0.20	00	0.62
D	0.48	0.08	0.45	00
Independent trial matrix (r_{ij})				
	A	B	C	D
A	00	0.03	0.19	0.10
B	0.02	00	0.01	0.007
C	0.14	0.007	00	0.18
D	0.15	0.003	0.17	00
	E(pre)	E(post)	R(pre)	R(post)
A	1.362	1.260	0.859	0.795
B	0.554	1.516	0.350	0.956
C	1.404	2.131	0.886	0.344
D	1.319	1.058	0.836	0.668
E(max) = 1.585, E(system) = 2.9319				

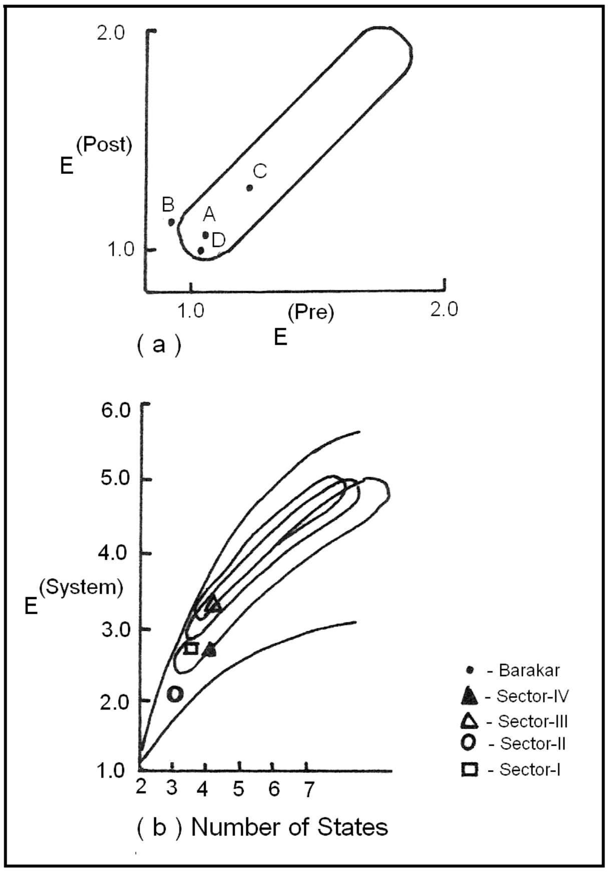

Figure 6. **(a)** Entropy set for entire Barakar Formation (based on pooled data from 38 borehole logs), and **(b)** relationship between entropy and depositional environment for the four sectors and entire Barakar Formation (environmental boundaries after Hattori 1976).

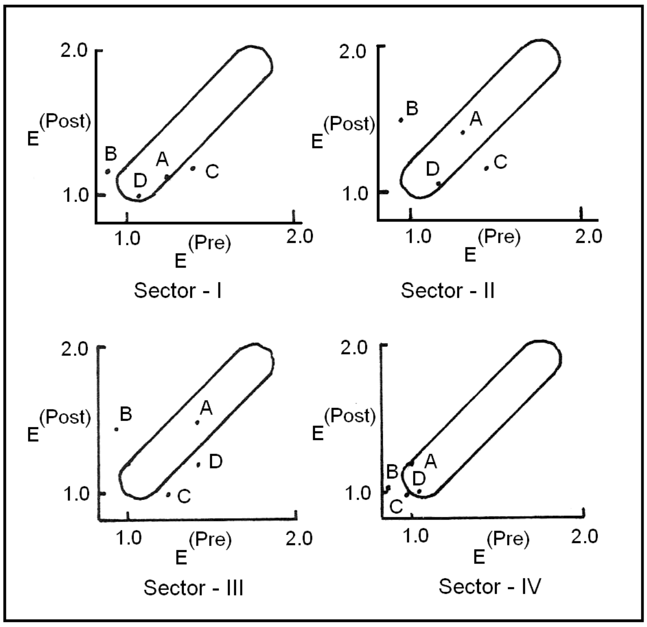

Figure 5. Entropy sets for the lithofacies of the Barakar Formation in four sectors. A: Sandstone, B: Interbedded sandstone-shale, C: Shale, and D: Coal.

and qij equals to the bulk transition probability matrix values calculated for data from both sequences. The number of degrees of freedom equals to $(T-1)n(n-1)$, where n equals to total number of lithologic states (4, in the present case).

Stationarity of the Markov process is tested at sector level on grouping the entire borehole logs in each sector. Applying the equation mentioned above, chi-square statistics of stationarity are computed and recorded in table 12. If the null hypothesis (i.e., $pij(t) = qij$) is computed then the 'calculated value must be less than the tabulated value at some preselected level of significance for the total number of degrees of freedom' (Harbaugh and Bonham-Carter 1970, p. 125). The highest calculated value of chi-square statistics is 21.18, which at 12 degrees of freedom is below the limiting value of 26.22 at 95% significance level; emphasizing that at each sector the sequence of lithologic transitions has been stationary through space (table 12).

The results suggest that the nature of the cyclic sequence is stationary at sector level and strongly indicates that the corresponding set of

Table 12. *Chi-square stationarity statistics within Barakar coal-bearing cycles of Bellampalli coalfield.*

Sector	χ^2	Limiting values 99%	95%	Markov process
I and II	4.826	26.22	21.03	Stationary
I and III	6.969	26.22	21.03	Stationary
I and IV	9.942	26.22	21.03	Stationary
II and III	12.390	26.22	21.03	Stationary
II and IV	21.186	26.22	21.03	Stationary
III and IV	12.126	26.22	21.03	Stationary

sub-environments in the depositional basin conforms to a definite pattern.

6. Sedimentological interpretation

The application of improved Markov process model and entropy functions evidently indicate symmetrical cyclic pattern in the Barakar Formation of Bellampalli coalfield. The coarse to medium-grained sandstone (facies A) forms the basal unit of each fining upward cycle. It is dominant lithofacies, which shares about 65% of the Barakar Formation by volume. The available outcrop sections suggest that it generally occurs as channel to sheet-like and multistory bodies. The channel- and sheet-like sandstone bodies showing erosional base and almost flat-top in addition to cosets of planar and trough cross-beds; these sandstone bodies also exhibit horizontal beds and scour-and-fill structure. The sandstone lithofacies along with erosional structures and cross-beds resemble channel deposition (point bar/braid bar) in close conformity with the Barakar Formation in eastern and central India (Casshyap and Tewari 1984; Tewari 1998). Multistory sandstone bodies in fluvial environments are probably formed when the rate of migration within the aggrading channel belt is large enough to cause superposition of channel bars, before the channel belt is abandoned (Gordon and Bridge 1987; Bridge and Mackey 1993), On the other hand, Bridge and Leeder (1979) have suggested that low subsidence rates than sedimentation give rise to sheet sandstone bodies. The interbedded fine-grained sandstone/shale (facies B) and shale (facies C), preferentially overlie coarse to medium-grained sandstone throughout the area. These lithofacies respectively constitutes about 12% and 13% of the given Barakar sequence. They occur as thin sheet-like (1.5–2.5 m) or lens like bodies along with the sandstone bodies. The fine-grained sandstone is ripple cross-laminated whereas shale is laminated. The interbedded assemblages of the fine clastics is attributed to deposition by vertical/lateral accretion on top of channel bars during a low water stage or as overbank/levee facies during periods of overflow. Following the flood stage, laccustrine conditions of stagnant water may have developed in the low lying areas beyond channel and overbank sub-environments, resulting the deposition of shale/carbonaceous shale or in coal forming sub-environments where swamp or marshy conditions develop. Similar peat forming environments may have developed in areas of abandoned channels. Indeed, shale records a strong upward transition to coal ($Zij = +0.95$). However, upward linkage of coal (facies D), which represents the top unit of Barakar cyclical units has more preference for shale (facies C) than the coarse to medium-grained sandstone (facies A), resulting in the symmetrical fining cycles, implies a gradual encroachment of coal swamp by adjacent back swamp and levee sub-environments as a consequence of slow and gradual lateral shift of channel course across the alluvial plain of meandering streams. Coal accounts about 5% of the Barakar sequence and contains six workable coal seams varying in thickness from 5 to 15 m (Raja Rao 1982). In addition, there are several coal beds of 1 m or less in thickness and are generally laterally discontinuous and lens like, may represent their formation in interchannel and/or distal flood plains or frequent shifting of channels should have prevented development of thick peat swamps to produce only thin impersistent coal seams (Tewari 1997). The fining upward cycles deduced here are similar to those, which characterize meandering and/or braided streams environment. Similar coal-bearing Barakar cycles of eastern and central India Gondwana basins have been interpreted in terms of lateral migration of stream channel across the alluvial plains (Casshyap and Tewari 1984; Hota and Maejima 2004; Khan and Tewari 2007). However, the coal-bearing Barakar cycles enclosing relatively thick shale and coal beds exhibiting sharp vertical relationship between lithofacies, particularly in the middle and upper parts of Barakar sequence of eastern and central India have been recently attributed to repeated slow and rapid subsidence of the depositional basin (Maejima *et al* 2008).

7. Conclusions

The present study is designed to quantify the nature of cyclicity observed with in the Barakar Formation of the Bellampalli coalfield in Andhra Pradesh, India, using modified Markov chain analysis and entropy functions. The cyclical sequence shows a fining upward character and is commonly symmetrical represented by coarse to medium-grained sandstone → interbedded fine sandstone/shale → shale → coal → shale, similar to Hattori's type 'B' pattern. This order of lithologic transition is closely comparable with that suggested for the early Permian Gondwana coal measures of India (Khan and Tewari 2007).

The Markov chain and entropy pattern substantiated proposed fluvial meandering depositional model as derived independently from paleocurrent analysis (Singh and Tewari 2008). As widely believed (Casshyap and Tewari 1984; Khan and Tewari 2007) meandering river channels of moderate to high sinuosity in response to varying discharge and rate of deposition should account for cyclical deposition in the study area. Rapid and frequent lateral shift of channel course, a common phenomenon in modern river basin may favourably explain the development of symmetrical fining upward cycles. However, those cycles which enclose thick coal beds exhibiting sharp relationship between lithofacies possibly developed due to the differential subsidence of the depositional basin.

Acknowledgements

We are thankful to Sri Bassava Chari, the then General Manager (Exploration), Singareni Collieries Company Ltd., Kothagudem for his kind permission to collect borehole records for this study. Our sincere appreciation is due to Sri Jagan Mohan, Dr Uday Bhasker and staff of the S.C.C.L. for many fruitful discussions. We also thank Mr. S K Agarwal D.G.M, U P for his cooperation in making data files about mathematical operations. We appreciate the constructive comments of the manuscript by the reviewers of the Journal of Earth System Sciences.

References

Anderson T W and Goodman L A 1957 Statistical inference about Markov chain; *Ann. Math. Statist.* **28** 89–110.

Beebower J R 1964 Cyclothems and cyclic depositional mechanism in alluvial plain sedimentation; *Bull. Kans. Univ. Geol. Survey* **169** 35–42.

Bridge J S and Leeder M R 1979 A simulation model of alluvial stratigraphy; *Sedimentology* **26** 617–644.

Bridge J S and Mackey S D 1993 A revised alluvial stratigraphic model; In: Alluvial Sedimentology (eds) Marzo M and Pingdafabregas C, *Spec. Publ. Intern. Assoc. of Sedimentologists* **17** 319–336.

Casshyap S M 1975 Cyclic characteristics of coal bearing sediment in the Bochumer Formation (Westphal A-2) Ruhrgebiet, Germany; *Sedimentology* **22** 237–255.

Casshyap S M and Tewari R C 1984 Fluvial models of the Lower Permian Gondwana coal measures of Son–Mahanadi and Koel–Damodar basins, India; In: Coal and Coal bearing Sequences (eds) Rahmani R A and Flores R M, *Spec. Publ. Intern. Assoc. of Sedimentologists* **7** 121–147.

Casshyap S M, Kreuser T and Wopfner H 1987 Analysis of cyclical sedimentation in the Lower Permian Machuchuma coalfield (SW Tanzania); *Geologische Rundschau* **76** 869–883.

Davis J C 2002 *Statistics and Data Analysis in Geology*; (Singapore: John Wiley & Sons) 638pp.

Gingerich P D 1969 Markov analysis of cyclical alluvial sediments; *J. Sedim. Petrol.* **39** 330–332.

Goodman L A 1968 The analysis of cross-stratified data: Independence, quasi-independence and interactions in contingency tables with or without missing entries; *J. Anner. Statistist. Assoc.* **63** 1091–1131.

Gordon E A and Bridge J S 1987 Evolution of Catskill (upper Devonian) river system; Intra and extrabasinal controls; *J. Sedim. Petrol.* **57** 234–249.

Harbaugh J W and Bonham-Carter G 1970 Computer Simulation in Geology (New York: Wiley Inter. Science) 575 pp.

Hattori I 1976 Entropy in Markov chains and discrimination of cyclic patterns in lithologic successions; *Math. Geol.* **8** 477–497.

Hota R N and Maejima W 2004 Comparative study of cyclicity of lithofacies in Lower Gondwana formations of Talchir basin, Orissa, India: A statistical analysis of subsurface logs; *Gondwana Res.* **7** 353–362.

Khan Z A 1997 Quasi-independence and entropy analysis of a late orogenic Middle Siwalik molasses sequence in Kuluchaur area, Pauri Garhwal, India; *Geoinformatics* **7** 135–145.

Khan Z A and Casshyap S M 1982 Entropy in Markov chain of Late Paleozoic cyclical coal measures of East Bokaro basin; *Math. Geol.* **13** 153–162.

Khan Z A and Tewari R C 2007 Quantitative model of early Permian coal bearing cycles from Son–Mahanadi and Koel–Damodar basins of eastern-central India; *Gondwana Geol. Magazine, Spec. Publ.* **9** 115–125.

Mack G S and James W C 1986 Cyclic sedimentation in the mixed zone (Lower Permian), Southwestern Mexico; *J. Sedim. Petrol.* **56** 635–647.

Maejima W, Tewari R C and Hota R N 2008 On the origin of Barakar coal bearing cycles in the Gondwana basins of peninsular India; *J. Geosci. Osaka City University, Japan* **51** 21–26.

Middleton GV 1973 Johannes Walther's Law of the correlation of facies; *Bull. Geol. Soc. America* **84** 979–988.

Power D W and Easterling G R 1982 Improved methodology for using embedded Markov chains to describe cyclical sediments; *J. Sedim. Petrol.* **53** 913–923.

Raja Rao C S 1982 Coalfields of India: Coal resources of Tamil Nadu, Andhra Pradesh, Orissa and Maharashtra; *Bull. Geol. Surv. India, Series A* **45** 103 p.

Read W A 1969 Analysis and simulation of Namurian sediments in central Scotland using a Markov process model; *Mathematical Geol.* **1** 199–219.

Reading H G 1991 Sedimentary Environments and Facies; 2nd edn. (London: Blackwell) 641 p.

Sengupta S M 2006 Introduction to Sedimentology; 2nd edn. (New Delhi: Oxford and IBH Publishing Co.) 314 pp.

Sharma S, Sharma M and Singh I B 2001 Facies characteristics and cyclicity in lower Siwalik sediments, Jammu area: A new perspective; *Geol. Mag.* **138** 455–470.

Singh D P and Tewari R C 2007 Paleochannel patterns of early and middle Permian Gondwana streams of Bellampalli–Chinnur area of Andhra Pradesh; *Gondwana Geol. Magazine* **22(2)** 73–78.

Tewari R C 1997 Numerical classification of coal bearing cycles of early Permian Barakar coal measures of eastern central India Gondwana basins using Q-mode cluster analysis; *J. Geol. Soc. India* **50** 593–600.

Tewari R C 1998 Channel sandstone bodies in fluvial Permian–Triassic Gondwana succession of peninsular India; *J. Geol. Soc. India* **51** 747–754.

Tewari R C 2005 Tectono-stratigraphic-sedimentary events in Gondwana succession of peninsular India; *J. Geol. Soc. India* **65** 636–638.

Tewari R C and Casshyap S M 1983 Cyclicity in early Permian Gondwana coal measures: An example from Giridih and Saharjuri basins of Bihar; *Sedim. Geol.* **35** 297–312.

Tewari R C and Singh D P 2008 Permian Gondwana paleocurrents in Bellampalli coalbelt of Godvari valley basin, Andhra Pradesh and paleogeographic implications; *J. Geol. Soc. India* **71** 266–270.

Vistelius A B 1949 On the question of the mechanism of formation of strata; *Dokl. Acad. Nauk. SSSR* **65** 191–194.

Walther J 1893 Einleitung in de Geologie als historische Wissenschaft-Jena; V. 1055.

MS received 16 June 2008; revised 14 February 2009; accepted 20 April 2009