

Research Article

GEO-STATISTICAL ANALYSIS OF THE BARAKAR CYCLOTHEMS (EARLY PERMIAN): A CASE STUDY FROM THE SUBSURFACE LOGS IN SINGRAULI GONDWANA SUB-BASIN OF CENTRAL INDIA

Z. A. Khan¹ and *Ram Chandra Tewari²

¹*Directorate of Geology and Mining, Khanij Bhavan, Lucknow-226001, India*

²*Department of Geology, Sri J. N. P. G. College, Lucknow-226002, India*

**Author for Correspondence*

ABSTRACT

Geo-statistical models such as quasi-independence Markov chain analysis, Entropy analysis. Linear regression, Principal component analysis and Factor analysis are used to define, analyze and interpret coal bearing Barakar cyclothems from Singrauli Gondwana sub basin of central India. The fining upward cycles are Type B symmetrical cycles that varying in thicknesses from 4-5m to several tens of meters. Lateral migration of stream channels, channel aggradations and differential subsidence of the basin floor in response to variable sediment supply are the most likely processes for the origin of the cyclothems. Significant interrelationships between stratigraphic and lithologic variables are indicated due to positive and definite correlation between total thickness of strata, total thickness and number of sandstone, total thickness and number of shale beds, and total thickness and number of coal seams. The statistical results suggest an optimum balance between the rate of deposition and the rate of subsidence throughout the deposition of Barakar cyclothems. This may indicate, in turn, that the essential components of the depositional framework did not change materially through time, and the formation of coal was in various sub environments of low lying abandoned flood plains of sinuous streams, interchannel areas, protected lakes and distal crevasse splays.

Key Words: *Geostatistics, Gondwana, Permian, Barakar Formation, Singrauli Sub-basin*

INTRODUCTION

Sedimentologists, in sharp contrast to researchers in other geology disciplines, have recently become interested to search for simple relationships between the lithological variables within a basin. Indeed, the complexity of the sedimentological processes greatly affects lithological variables. Till date, most work on the relationships between lithological variables within particular areas has been based on the visual comparison (Casshyap and Tewari, 1984; Tewari, 2004 and Khan, 2007). This approach has the undoubted advantage that it is comprehensive and has proved to be an excellent method of detecting qualitative relationships. By and large, the lithofacies and paleoflow analysis were used to reveal the broad, regional sedimentation patterns of lithological variables in coal bearing succession around the world (Khan and Casshyap, 1982; Casshyap and Tewari, 1984; Tewari, 2005; Tewari and Singh, 2008 and Tewari *et al.*, 2012). It cannot readily adapt to express quantitative relationships between lithological variables and the knowledge of which is vital to simulate depositional processes by computer (Krumbien, 1968; Harbaugh and Bonham-Carter, 1970; Davis, 2002 and Anderson, 2003). In addition, the quantitative information provides a basis for future computer simulation studies, designed to tackle effectively the underlying causes of cyclical sedimentation within a framework of relevant geological knowledge (Casshyap *et al.*, 1988; Khan and Tewari, 1991, 2007; Tewari, 2008 and Tewari *et al.*, 2009). Following Khan (1978) who used product moment correlation coefficient and linear regression lines to demonstrate approximately linear relationship between the total thickness of strata, the number of cycles and average thicknesses of coal cycles in Barakar coal measures in the East Bokaro coalfield, Casshyap *et al.*, (1988) used similar method to investigate the relationships between seven lithological variables from the similar succession in different Gondwana sub-basins of eastern India. It was realized, however, that the calculation of correlation coefficients and linear regression lines did not constitute a complete analysis

Research Article

of all the interacting variables and the more sophisticated techniques of principal component analysis and factor analysis were therefore applied to the data (Tewari, 2008 and Khan and Tewari, 2010). These techniques produce results which are closely in access with the results of the earlier regression analysis.

The depositional environment and facies changes, paleoflow characters and paleohydrology of the Barakar sequence of the eastern-central India Gondwana basins have been investigated in considerable detail (Casshyap and Tewari, 1984 and Tewari, 2005). In addition, the Early Permian Barakar sediments are also studied using statistical models such as Markov chain and Entropy function (Khan and Casshyap, 1981; 2007; Casshyap and Khan, 1982; Casshyap and Tewari, 1984; Tewari and Casshyap, 1983; Tewari *et al.*, 2009 and Hota and Maejima, 2004) linear regression and correlation coefficients (Casshyap *et al.*, 1988 and Khan and Tewari, 1991) Cluster analysis (Tewari, 1997 and Khan, 2007), factor analysis (Khan and Tewari, 2010) and Principal component analysis (Tewari, 2008). However, no attempt has been made to apply these models simultaneous on a single data set (sub basin) so as to provide an integrated quantitative model. The present case study may thus be considered as a contribution towards our researches on mathematical modeling of Early Permian Gondwana coal measures of Peninsular India using the combination of geo-statistical models in a given area. The results not only serve as an independent check on the results obtained earlier but also provide a suitable quantitative model of coal forming swamps in fluvial system. This reasonably simple half graben coal Gondwana basin was selected for detailed case study with the view that the results might prove to be applicable to cyclically deposited sequences in other basins of India and abroad.

Stratigraphic Summary

The intracratonic Gondwana basins of Peninsular India occurring as graben and half-graben interpreted as rift valley by early workers (Ghosh and Mitra, 1972). However, subsequent workers (Casshyap, 1979; Casshyap and Tewari, 1984; Veevers and Tewari, 1995; Tewari, 2005 and Tewari and Maejima, 2010) believed that the original Gondwana basins of Peninsular India grew in size progressively with sedimentation and that they were down faulted later (Mesozoic), followed by prolonged erosion that resulted in the present disposition of the outliers. The Singrauli Gondwana sub-basin constitutes the northern most part of the Son –Mahanadi Gondwana basin is, likewise, a half graben with the southern contact erosional and the northern contact faulted against the Archean basement (Fig.1).

Table 1: Stratigraphy and sedimentary characters of Gondwana rocks of Singrauli sub basin, Central India

Age	Formation	Lithology and Sedimentary characters
Late Triassic	Mahadeva	Coarse grained multistory and sheet like sandstone bodies, occasionally ferruginous. Lens like thin shale beds. Conglomerate and pebble beds in lower part.
Early Triassic	Panchet	Medium to coarse grained, white, greenish white, pink micaceous sandstone. Greenish brown and red siltstone and shale. Occasional pebble beds in lower part.
Late Permian	Raniganj	Fining upward cycles of coarse to medium and fine grained sandstone, arenaceous and carbonaceous shale, and coal.
Middle Permian	Barren Measures	Very coarse to coarse and medium grained ferruginous sandstone, and red and green shale.
Early Permian	Barakar	Fining upward cycles of medium to coarse grained, channel to sheet like sandstone, carbonaceous shale and coal.
Permo-Carboniferous	Talchir	
-----	U n c o n f o	r m i t y -----
Pre-Cambrian		Phyllites, schists, quartzite and gneisses.

Research Article

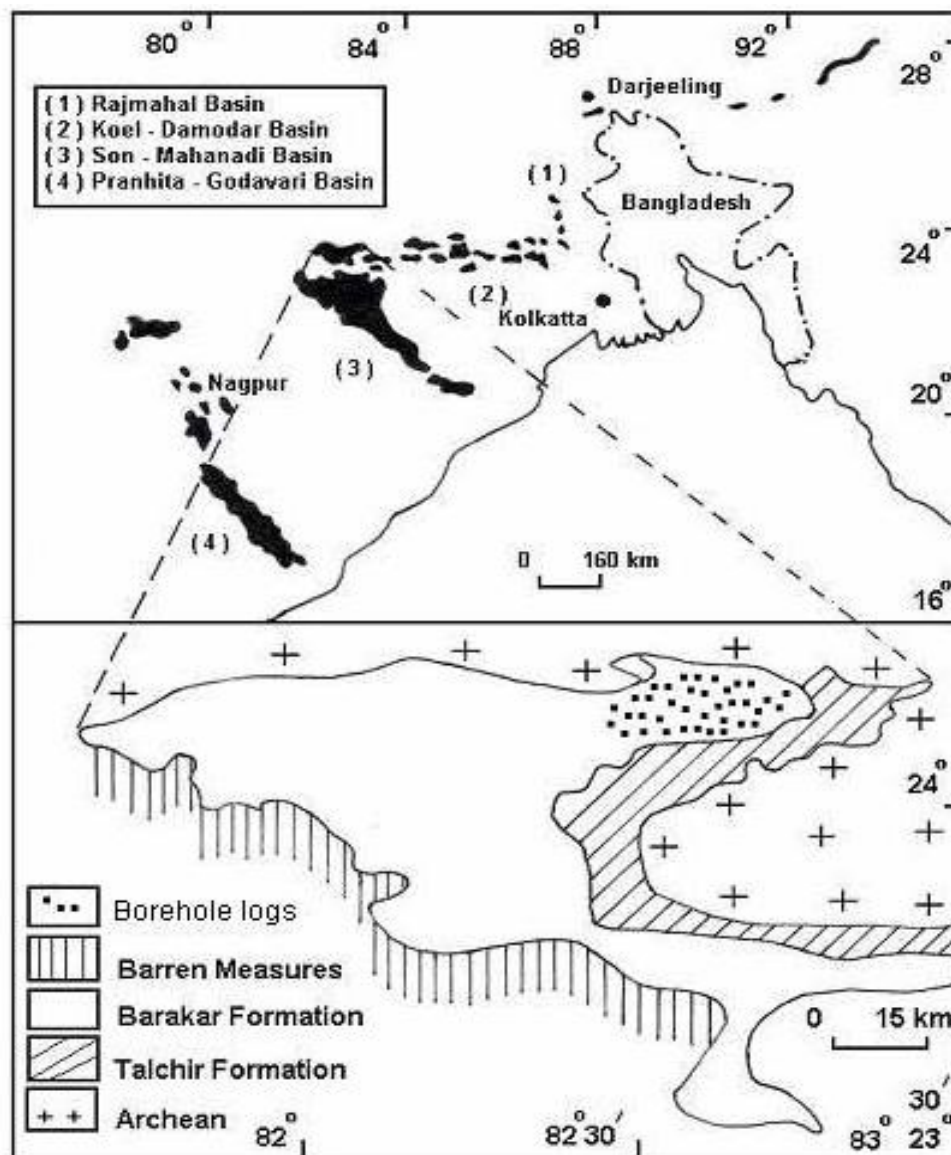


Figure 1: Map showing distribution of Gondwana basins of India and geological map of Singrauli sub-basin

The Gondwana stratigraphy of the area represents about 2350 m thick sequence of Permian-Triassic Gondwana rocks comprising Talchir, Barakar, Barren Measures, Raniganj, Panchet and Mahadeva Formations in ascending order. Table 1 summarises the Gondwana stratigraphy and sedimentary characters of various litho-units of Singrauli Gondwana sub-basin. The glacial Talchir formation (80m) which lie unconformable on the Archean basement represents the basal sedimentary formation of the Gondwana sequence. The overlying Early Permian Barakar formation occurs extensively in the sub-basin and is about 600m in eastern and central Son Valley, but is relatively thin in the western Son Valley (300m). It shows a variable relationship with underlying formations, lying gradationally above the Karharbari or Talchir and overlaps them to rest directly on the Archean basement. The above overlapping relationship suggests aerial expansion of basinal area during Barakar sedimentation following the termination of glacial episode as elsewhere in other Gondwana basins of Peninsular India (Tewari, 2005 and Tewari and Maejima, 2010).

Research Article






Facies	Description	Interpretation
	Shaly coal and / or bituminous coal	Backswamp
	Interbedded lenses of shale and fine grained sandstone. Common are parallel to ripple cross laminations.	Vertical accretion deposits from overbank flood/ and or swales.
	Coarse to medium grained feldspathic sandstone : cosets of planar and trough cross-beds ubiquitous	Middle and upper part of point bar / braid bar.
	Gritty to coarse feldspathic to lithic sandstone showing large scale planar cross-beds; bedding at intervals horizontal.	Largely channel deposits. Braid bar sands of braided rivers.
	Pebbly to gritty lithic sandstone with thin interbeds of conglomerate. Large trough cross-bedding common.	Channel bar, locally channel lag of braided rivers.

Figure 2: Diagrammatic representation of lithofacies and sedimentary characters of Barakar Formation, Singrauli sub-basin. Depositional environments are listed alongside

The bulk of the Barakar formation consists of pink, grey to white coarse to very coarse grained sandstone, fine grained sandstone, siltstone, shale and coal, or, more often is overlain directly by any one or more than one of the units listed above (Fig. 2). Interbedded with or scattered in the coarse grained sandstone, sporadically, are fine, medium to coarse pebbles and cobbles. Generally, speaking the fine clastic, like siltstone and shale, constitute a small proportion of the Barakar cyclothems and, where present rarely

Research Article

exceed couple of meter thickness and are seldom extensive laterally. The overlying Barren Measures occurring in the southern part of the area are uniformly composed of Interbedded coarse to medium grained channel to sheet like sandstone and grey to red and micaceous shale. The succeeding Raniganj Formation of Late Permian confirms the return of coal and composed of fining upward cycles. Indeed, the thickest known coal seam of about 134 m in thickness occurs in the Raniganj sediments of this area. The Gondwana rocks strike east-west with low dips ($< 10^\circ$) directed towards north, though steep dips up to 25° are recorded close to northern faulted boundary.

Basic Data

The data used in the study is simple cored borehole successions of lithological members coded into a limited number of states for the Markov chain and Entropy analysis. 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. The lithological data were coded into four states, namely coarse-medium grained sandstone (SS), arenaceous shale /argillaceous shale (SH), carbonaceous shale (CS) and coal (C). All four states are well represented in each of the thirty seven boreholes. For the quantitative interrelationships between lithological variables, following nine variables were extracted from the logs of thirty seven cored boreholes. The nine lithological variables and the symbols used to designate them are as:

- The total thickness of strata (B1)
- The total thickness of sandstone members (B2)
- The total thickness of shale (B3)
- The total thickness of coal seams of any thickness (B4)
- The number of sandstone members (B5)
- The number of shale (B6)
- The number of coal seams of any thickness (B7)
- The sand/shale ratio (B8)
- The clastic ratio (B9)

Following the examples of (Duff and Walton, 1962; Casshyap, 1975 and Tewari and Casshyap, 1983) an arbitrary definition of the term coal cycle was adopted. In the present study the top of coal cycle was placed at the top of coal, or if coal was absent, the top of shaly coal. To constitute a separate coal cycle the coal or shaly coal must be separated from the next coal or shaly coal horizon in the sequence by at least 30cm of clastic sediment.

MATERIALS AND METHODS

Markov Chain Analysis

The Markov chain proposed by (Vistelius, 1949) is used to identify and evaluate stratigraphic trends, which can often be obscured by non-cyclic elements. The objective of a Markov chain analysis is essentially to take away the randomness of each component within a population. Once this is done, the remaining transitions are then assumed to be due to non-random processes which may be geologically interpreted. The Markov chain analysis searches for the best probable transition from one to another facies through time. The transitions in a sequence of lithologies can be summarized in a matrix of one step transition. A one step Markov process is a stochastic process in which the facies of the system at time t_n is influenced by or dependent on the facies of the system at time $t_{(n-1)}$, but not the previous history that led to the facies at time $t_{(n-1)}$. When sedimentary facies are used, observations within the same facies results in auto correlation of thicker facies within themselves, thus over shadowing any Markovian tendency present between different units (Krumbien, 1968) and the transitions between different facies (regardless of their stratigraphic thickness) were considered. The resulting transition matrix has the property that main diagonal frequencies are zero. Subsequent examination of this commonly used method revealed its susceptibility to inappropriate conclusions. (Schwarzacher, 1975 and Power and Sterling, 1982) indicated that the major obstacle to rigorous analysis was the presence of previously defined zeros

Research Article

in the transition count matrix. They stated that matrices containing predefined zeros cannot result from a simple independent random process, and believed subsequent statistical tests are meaningless.

Quasi – independence models, a statistically rigorous group of techniques that are applicable to the evaluation of incomplete matrices are described by (Power and Sterling, 1982). The quasi-independence model can be used to generate the expected cell values from the independent trails matrix and takes account of restriction upon independence within sequential data thus providing somewhat less random conditions and this procedure has been adapted in current study. As discussed above, each interval regardless of its thickness forms a single step in the chain. The discrete steps are then used to construct a transition count (or tally) matrix (Table 2). Each cell displays the number of observed transitions from the facies of row i (t_{n-1}) to the overlying facies of column j (t_n). These upward transitions for the Barakar coal measures sequences were recorded from 37 different boreholes, was pooled in 488 transitions and structured in 4x4 transition count matrix (n_{ij}) (Table 2a).

Following matrices are computed from the transition count matrix:-

1). Observed transition probability matrix (p_{ij}) which gives the actual probabilities of the given transition occurring in given sequence was calculated as

$$p_{ij} = n_{ij} / n_{++} \\ = \text{cell value} / \text{row sum}$$

2). Expected frequency transition estimates with quasi-independence (E_{ij}) given by $E_{ij} = a_i b_j$ ($i \neq j$) derived by using an iterative procedure till a_i and b_i attain an arbitrary constant [29, p.916]. Below we describe the estimation of the expected transition frequencies under quasi-independence.

Let $E(n_{ij})$ denotes the expected value of the number of transitions from state i to state j in a particular geologic section where n facies are possible. Then the discrete Markov process is said to possess model “quasi-independence”, when each $E(n_{ij})$ is given by

$$E(n_{ij}) = a_i b_j \quad i \neq j \quad i, j = 0, 1, 2, 3, \dots, m$$

Where a_i and b_j denotes frequency of individuals in the i^{th} row and j^{th} column, respectively. Estimating the parameter, a_i and b_j , $i, j = 1, 2, 3, \dots, m$ require an iterative solution as follows:

First Iteration:

$$a_i^1 = n_{i+} / (m-1), \quad i = 1, 2, 3, \dots, m$$

$$b_j^1 = n_{+j} / \sum a_i^1$$

Similarly I^{th} Iteration:

$$a_i^{(I)} = n_{i+} / \sum b_j^{(I-1)}, \quad i = 1, 2, 3, \dots, m$$

$$b_j^{(I)} = n_{+j} / \sum a_i^{(I)}, \quad j = 1, 2, 3, \dots, m$$

Iteration is continued until some specified accuracy is obtained. [29, p.916] found a convergence criterion of 1% more than adequate and is retained in present study. That is, iteration is continued until

$$\{a_i^{(I)} - a_i^{(I-1)}\} < 0.01 a_i^{(I)} \quad \text{for } i = 1, 2, 3, \dots, m.$$

And

$$\{b_j^{(I)} - b_j^{(I-1)}\} < 0.01 b_j^{(I)} \quad \text{for } j = 1, 2, 3, \dots, m.$$

Let A_i and B_j denotes the final value of $a_i^{(I)}$ and $b_j^{(I)}$ then the estimated expected frequencies under quasi-independence are given by $E_{ij} = A_i B_j$, for $i \neq j$. The calculated values are shown in Table 2C.

3). Transition probability estimates with quasi-independence (P_{ij}). The observed transition probabilities contain both random and deterministic component. To evaluate the latter, the estimated expected frequencies, E_{ij} , obtained under quasi-independence are divided by the row sums, and corresponding transition probability estimates (P_{ij}) are obtained. The sum of transition probabilities from a particular facies to the other facies will be equal to 1. The calculated values are given in Table 2D.

4). Probability difference matrix (d_{ij}). This matrix is obtained by subtracting the observed transition probability matrix (p_{ij}) from the expected transition probability matrix (P_{ij}). It has both positive and negative entries (Table 2E), where positive entries are interpreted as being dominant transition.

Research Article

5). Normalized difference matrix (Nd_{ij}). Whether or not a difference represents a “significant” departure from quasi-independence depends upon the size of the probabilities being estimated and on the amount of data involved in the estimates. Table 2E, one cannot tell which difference are “signal” and which are “noise”. To make this problem easier, (Turk, 1979 and Power and Sterling, 1982) propose a normalized difference matrix as an aid in interpreting large differences between observed transition frequencies and transition probabilities estimated with a model of quasi-independence. Mathematically it can be expressed as

$$Nd_{ij} = f_{ij} - E_{ij} / \sqrt{E_{ij}}$$

Where f_{ij} = observed number of transition from i^{th} facies to j^{th} facies, E_{ij} = expected number of transition probability from i^{th} to j^{th} facies.

Table 2: Quasi independence Markov matrices and chi-square of lithological state in Barakar cyclothems, Singrauli Gondwana sub- basin

Table 2A: Transition Count Matrix (f_{ij})

	Sandstone	Shale	Carb. shale	Coal
Sandstone	0	92	21	10
Shale	48	0	71	23
Carb. shale	42	25	0	61
Coal	34	21	40	0

Table 2B: Transition Probability Matrix (p_{ij})

	Sandstone	Shale	Carb. shale	Coal
Sandstone	0	0.748	0.170	0.081
Shale	0.338	0	0.500	0.161
Carb. shale	0.328	0.195	0	0.476
Coal	0.357	0.221	0.421	0

Table 2C: Expected Frequency Transition Matrix (E_{ij})

	Sandstone	Shale	Carb. shale	Coal
Sandstone	0	50.17	44.51	26.32
Shale	50.38	0	55.98	35.62
Carb shale	44.59	55.86	0	31.54
Coal	28.25	35.38	31.38	0

Table 2D: Expected Transition Probability Matrix under Quasi-independence (P_{ij})

	Sandstone	Shale	Carb.shale	Coal
Sandstone	0	0.414	0.367	0.218
Shale	0.354	0	0.394	0.250
Carb. shale	0.337	0.423	0	0.239
Coal	0.298	0.370	0.331	0

Table 2E: Probabilities Difference Matrix (d_{ij})

	Sandstone	Shale	Carb. shale	Coal
Sandstone	0	+0.334	-0.197	-0.137
Shale	-0.016	0	+0.106	-0.089
Carb. shale	-0.009	-0.228	0	+0.237
Coal	-0.059	-0.149	+0.090	0

Research Article

Table 2F: Normalized Difference Matrix (Nd_{ij})

	Sandstone	Shale	Carb. shale	Coal
Sandstone	0	+5.91	-3.52	-3.44
Shale	-0.33	0	+2.01	-2.1
Carb. shale	-0.39	-3.59	0	+5.88
Coal	+0.71	-2.42	+1.54	0

Test of Significance: The null hypothesis of a random model for the whole sequence is tested again using a test statistics based upon the concept of quasi-independence as

$$\chi^2 = \sum_{i=1}^n \sum_{j=1}^n (f_{ij} - E_{ij})^2 / E_{ij}$$

Where f_{ij} = observed number of transition from i to j , E_{ij} = expected number of transition from i to j under the assumption of quasi-independence. Under the hypothesis of independence χ^2 is approximately distributed as a chi-squared variable with $(m-1)^2 - m$ degree of freedom (because m cells were omitted). Thus, the observed value of χ^2 can be compared to tables of the chi-squared distribution to assess the conformance of the data to the model of statistical independence. The larger the χ^2 value, for a given value of m , the stronger the evidence is against the hypothesis of independence.

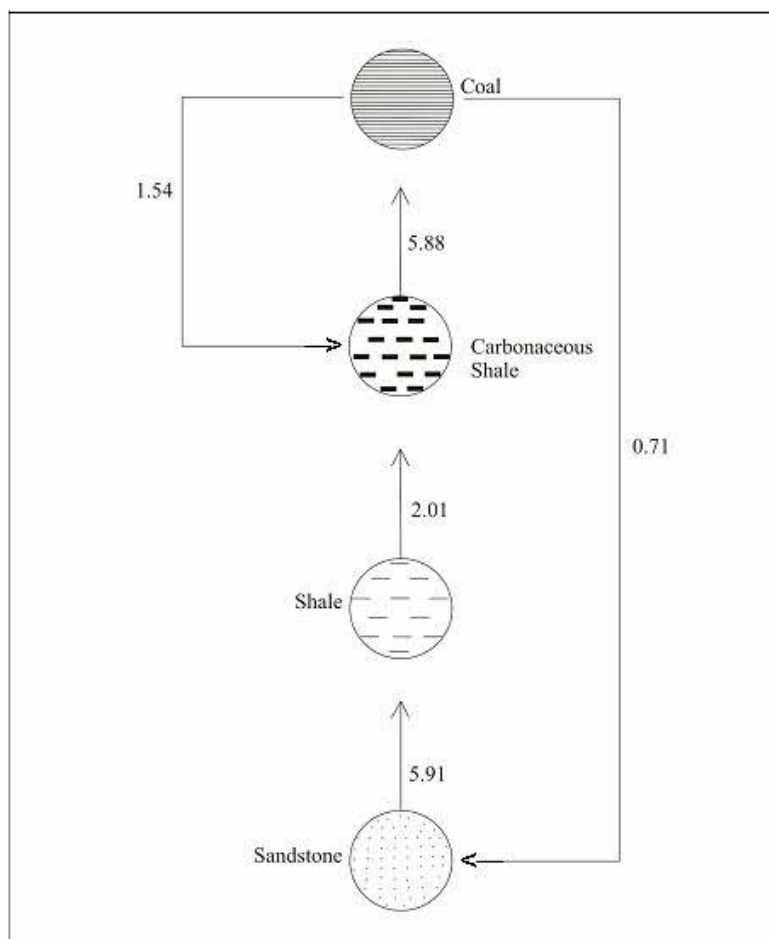


Figure 3: Markov diagram based on positive values of normalized difference matrix (Nd_{ij}) showing upward transition between different lithologic states

Research Article

Table 2 records the bulk tally matrices as well as calculated values of expected transition frequency (E_{ij}), difference matrix (d_{ij}) and normalized difference matrix (Nd_{ij}). Chi-square statistics ($\chi^2 = 117.49$) also listed, for which the tabulated value at given degree of freedom and at 99.5% confidence limit are high enough to justify the presence of Markov property. Table 2 which gives probability difference (d_{ij}) and normalized difference (Nd_{ij}) to the data set shows that the primary contributor to the large (χ^2) is the excessive number of transitions from lithofacies SS to lithofacies SH and lithofacies CS and lithofacies C. With only five degrees of freedom among the twelve d_{ij} and Nd_{ij} , there is considerable dependency among them, so thus four cells also contribute to the several fairly large negative differences in difference matrices (Table 2E and 2F). Figure 3 shows Markov transition diagram based on positive values of normalized difference matrix (Nd_{ij})

Entropy Analysis

Hattori (1976) introduced the concept of “Entropy” to the Markov probability matrix to analyze the extent and ordering of the transitions in measured sections. It was further applied by many sedimentologists to analyze and interpret cyclic successions (Khan and Casshyap, 1981; Hota and Maejima, 2004; Khan and Tewari, 2007 and Tewari *et al.*, 2009). (Hattori, 1976) recognized two types of entropies- one is entropy after deposition (i.e. post depositional) and refers to leaving a particular j^{th} state from any other state and designated as $E_i^{(post)}$, while the other which refers to entering a particular state j^{th} state from any other state is entropy before deposition (i.e. pre depositional) and designated as $E_i^{(pre)}$. These two types of entropies pertain to every state; one is relevant to the Markov matrix expressing the upward transition and the other relevant to matrix expressing the downward transitions.

Entropy after deposition (i.e. across the row) with respect to state i can be calculated as

$$E_i^{(post)} = - \sum_{j=1}^n p_{ij} \log_2 p_{ij}$$

Where $E_i^{(post)}$ is entropy after deposition with respect to state i; n is the number of lithologic state; and p_{ij} is relative frequency that state j follows i.

Similarly Entropy before deposition (i.e. along the column of downward matrix) has been expressed with respect to state i as

$$E_i^{(pre)} = - \sum_{j=1}^n q_{ji} \log_2 q_{ji}$$

Where $E_i^{(pre)}$ is entropy before deposition which respect to state i; q_{ji} is relative frequency that state j precedes state i. $E_i^{(post)}$ and $E_i^{(pre)}$ serve as indications of the variety of transitions immediately after and before the occurrence of state i, respectively. Hattori (1976) listed interpretations for a set of relations between $E_i^{(pre)}$ and $E_i^{(post)}$ including the significance of state i with respect to succeeding or preceding states in a sequence, and also the influence of state i on its successor ($E_i^{(pre)} > E_i^{(post)}$) or dependency of its precursors ($E_i^{(pre)} < E_i^{(post)}$). (Hattori (1976) showed that by plotting $E_i^{(pre)}$ versus $E_i^{(post)}$ for each lithological state, one can make some interpretation to the style of cyclicity and the way in which the cycles are truncated. He drew a number of diagrams of the distribution of $E_i^{(pre)}$ verses $E_i^{(post)}$ for idealized, truncated, symmetrical and symmetrical successions. Apart from entropies with respect to individual sets, (Hattori, 1976) introduced entropy for the whole sedimentation system as

$$E_{(system)} = - \sum_{i=1}^n \sum_{j=1}^n r_{ij} \log r_{ij}$$

where $r_{ij} = f_{ij}/n_{++}$, f_{ij} is entries in the tally matrix;

$$n_{++} = \sum_{i=1}^n \sum_{j=1}^n f_{ij}$$

The $E_{(system)}$ can take a value between $-\log 1/n$ and $-\log 1/(n-1)$

Research Article

Regression Analysis

To investigate the relationships between lithological variables, the correlation coefficient of every possible pair of variables were calculated using Statistics 2010, using data from all 37 boreholes. The programmed used computes means, standard deviation, and correlation coefficient. The correlation coefficients, which are listed in Table 4, were tested for significance at the 5%, 1% and 0.1% (Fisher and Yates, 1963).

Table 3: Markov matrices equated to Entropies for the individual lithologic state of the Barakar cyclothems, Singrauli Gondwana sub- basin

Transition Count Matrix (f_{ij})

	Sandstone	Shale	Carb. shale	Coal
Sandstone	0	92	21	10
Shale	48	0	71	23
Carb. shale	42	29	0	61
Coal	34	21	40	0

Upward Transition Matrix (p_{ij})

	Sandstone	Shale	Carb.shale	Coal
Sandstone	0	0.748	0.170	0.081
Shale	0.338	0	0.500	0.161
Carb shale	0.328	0.195	0	0.476
Coal	0.357	0.221	0.421	0

Downward Transition matrix (q_{ij})

	Sandstone	Shale	Carb. shale	Coal
Sandstone	0	0.648	0.159	0.106
Shale	0.388	0	0.537	0.244
Carb Shale	0.389	0.205	0	0.649
Coal	0.274	0.148	0.318	0

Independent Trail matrix (r_{ij})

	Sandstone	Shale	Carb. shale	Coal
Sandstone	0	0.189	0.042	0.020
Shale	0.097	0	0.144	0.047
Carb. shale	0.085	0.058	0	0.124
Coal	0.069	0.042	0.081	0

Table 4: Calculated values of Entropy set for the Barakar cyclothem of the Singrauli Gondwana sub- basin

Lithological state	$E^{(pre)}$	$E^{(post)}$	Relation
Sandstone	1.17	1.04	$E^{(pre)} > E^{(post)}$
Shale	1.45	1.51	“
Carb. shale	1.40	1.42	“
Coal	1.44	1.54	$E^{(pre)} < E^{(post)}$

$$E_{(max)} = 1.585 \quad E_{(system)} = 3.014$$

Research Article

Table 5A: Linear Regression equations for pairs of variables with coefficient of correlation

Linear Regression Line	Correlation coefficient
$b = 0.85a - 23.33 (\pm 26.4)**$	0.99
$c = 0.09a + 4.81 (\pm 23.2)$	0.67
$d = 0.06a + 18.53 (\pm 12.5)$	0.30*
$c = 0.04b + 13.11 (\pm 24.8)$	0.58
$d = 0.05b + 22.65 (\pm 13.6)$	0.21 not significant
$e = 0.03b + 3.60 (\pm 3.71)$	0.78
$d = 0.13c + 26.91 (\pm 14.8)$	-0.17 not significant
$f = 0.62c + 1.67 (\pm 3.81)$	0.77
$g = 0.20d + 2.75 (\pm 7.90)$	0.32*
$f = 1.82e - 1.12 (\pm 12.85)$	0.42
$g = 0.67e + 3056 (\pm 7.70)$	0.30*
$g = 0.42 f + 3.33 (\pm 5.12)$	0.67

a = total thickness of strata

e = number of sandstone beds

b = total thickness of sandstone

f = number of shale beds

c = total thickness of coal seams

g = number of coal seams

**confident limits are given in brackets.

- *correlation coefficient significant at 95%
- correlation coefficient significant at 99%

Table 5B: Equations of linear lines showing statistical relationship of the number of coal cycles to total thickness of strata and to average thickness of coal cycles

Dependent variable	Independent variable	Correlation coefficient	Confidence level	Linear equation
No. of coal cycles (y)	Total thickness (x)	0.85	99%	$y = 0.06x + 1.70$
Average coal cycles thickness (z)	Total thickness (x)	0.25	95%	$Z = 0.08x + 16.15$
Average cycle thickness (z)	Number of coal cycles (y)	-0.17	80%	$z = -1.82y + 17.85$

The data were then used to calculate equations of linear regression lines and 95% fiducial limits of all pairs of variables that had coefficient of correlation which were significant at the 5% level or less, The equations and 95% fiducial limits are listed in Table 5 and graphs of the pairs of lithological variables are shown in Figure 4. It must be stressed that a high coefficient of correlation between any pair of variables does not necessarily imply a causative correlation and may be due to both being closely related to a third variable (Read and Dean, 1967 and Khan and Tewari, 2007)

Factor Analysis

Factor analysis, is a statistical technique designed to explain complex relations among variables in terms of a few factors, which themselves represent simpler relations among fewer variables. The occurrence of each of lithological variable is assumed to be completely and linearly determined by p independent factors or sedimentological processes. Different portion of the total variance or the loading of a variable can then be assigned to different factors and may be expressed as

$$X_j = \sum_{r=1}^p C_{jr} fr$$

Where fr ($r=1,2,3,\dots,p$) represents r th common underlying factors, x_j is the j th observed variable, p is the specified number of factors and C_{jr} indicates the factor loading of variable x_j on factor fr . The theoretical unknown factors can thus be expressed in terms of distinct groups of lithological variables which, when

Research Article

correlated with the observed features of sedimentation of the area of investigation, provide significant insight into the nature of causal factors.

A fundamental purpose of the factor analysis lies in the reduction of dimensionality. It is generally found that only the first few factors would show recognizable pattern of the data matrix, the remainder representing largely the random effect or noise. This calls for the selections of a meaningful and useful minimum number of factors ($p < m$) which will account for most of the variances in the data set and therefore convey the same information. Various criteria for the selection of suitable factors prior to analysis were suggested by different workers (Harris, 1985; Morrison, 1990; Davis, 2002). Some recommended retaining all those which have eigenvalues > 1 and others extracted only those factors which lie above a distinct break in the descending eigenvalues. In the present study, the numbers of factors were limited to two on the basis of cluster analysis (Khan, 2007) and secondly no geological advantages seem to be gained by using more than two factors (Khan and Tewari, 2010).

Principal Component Analysis (PCA)

In principal component analysis no hypotheses need to make about the variables (Lawley and Maxwell, 1971). It is a relatively straight forward method of “breaking down” a correlation (or covariance) matrix into a set of orthogonal components or axes, equal in number to the number of variables concerned. These correspond to the eigenvalues (latent roots) and eigenvectors (latent vectors) of the matrix, the eigenvalues being extracted in descending order of magnitudes and the eigenvectors being mutually orthogonal. Thus, if each variable is considered as a vector in multidimensional space with the number of dimensions in such space equal to the number of samples, these variables can be referred to the eigenvectors used as an orthogonal system of reference axes. The loading of each variable-vector on to an eigenvector constitute a column of the principal components matrix and the sum of square of these loadings is equal to the corresponding eigenvalues. Each eigenvalue indicates the amount of the total variance of the observed variables which can be related to the appropriate eigenvector.

Although the first few components may extract a high proportion of the total variance, all components are generally required to reproduce the correlations between observed variable exactly. It is usual however to concentrate attention upon the first few components (McCommon, 1966; Cooley and Lohnes, 1971 and Davis, 2002).

RESULTS AND DISCUSSION

Interpretation from Markov Chain Analysis

Figure 3 shows Markov transition diagram based on positive values of probability difference matrix (d_{ij}) and normalized difference matrix (Nd_{ij}). Highest positive values of d_{ij} and Nd_{ij} matrices (Table 2E and 2F) link lithologic states distinctly resulting in a strong transition path for lithologic sequence that can be derived as:

Sandstone (SS) → Shale (SH) → carbonaceous shale (SH) → Coal (C) → Sandstone (SS)

Often, in the literature, the positive differences in tables such as Table 2E and Table 2F are interpreted as being the “dominant” transitions. As states above figure are drawn with arrows connecting the lithologies for which there are positive differences and these are regarded as being “fully developed” or “ideal cycles”. The implicit assumption underlying this procedure is that the expected matrix represents “noise”; subtracting the expected matrix from observed matrix implies we can filter the observations from randomness or noise. The remainder should then be expected to represent signal.

This transition path is typical of the coal bearing Barakar Formation and displays a progressive fining upward of particle size from coarse grained sandstone through siltstone/shale, carbonaceous shale to coal. The lithological transitions as here deduced are by and large comparable with a few exceptions as also are closely similar to the cyclical sequences of other Late Paleozoic coal measures around the world (Read, 1968; Johnson and Cook, 1971; Casshyap, 1975 and Casshyap *et al.*, 1987) including the Permian coal measures of other Lower Gondwana basins of India (Casshyap and Khan, 1982; Casshyap and Tewari, 1984; Khan and Tewari, 2007; Tewari and Casshyap, 1983 and Tewari *et al.*, 2009). Each fining upward

Research Article

cycle comprises a lower coarse grade member of dominantly sandstone (channel bar) and an upper fine grade member of siltstone/ shale (overbank) topped by shaly coal or coal (coal swamp). The coal bearing cycles as deduced in this study when subjected to Entropy analysis reveals Type-B symmetrical cycles with varying thicknesses from couple of meters to several tens of meters.

Genetic interpretation of sedimentary members comprising fining upward cycles is recorded in Figure 2. As stated earlier, the sandstone assemblage comprising the coarse grade member of the Barakar cyclothems commonly fine upward from the base and exhibit large scale cross bedding, horizontal bedding as features that characterized by sands of many river channel bars (Reading, 1996; Miall, 2000 and Bogg, 2005). These bedding types are dominant ones in the side-and point bar sand of many modern streams. Side- or point bars result from the lateral accretion of stream bed load on the sideward migration of meandering channels. That the Barakar Paleocurrent system consisted essentially of anabranching and (or) migration system of streams (meandering) has been suggested on the basis of Paleocurrent analysis (Casshyap and Tewari, 1984 and Tewari, 2005). Extensive development of the coarse grade member in the Barakar cyclothems of Singrauli Gondwana sub-basin may indicate that the growing channel bar deposits migrated sideward as the meandering river channel anabranches through the slowly subsiding flood plains. The upper portion of cycle is dominated by transition probability siltstone/ shale (SH) \rightarrow carbonaceous shale (CS) ($Nd_{ij} = +2.01$). This is suggestive of continuous fining up sequence and supported by typical sedimentary structures, such as ripples, low angle cross bedding, consistent foreset azimuths. This may be interpreted as a possible levee deposition which interfingered frequently with back swamp sub environment. The carbonaceous shale, in turn, consistently show a strong preference to be overlain by coal seams as shown by high positive Nd_{ij} value ($Nd_{ij} = +5.88$) as well as positive value in difference matrix ($d_{ij} = +0.237$) which led to stable coal forming conditions. These episodes of undisturbed peat accumulation over a considerable time span was repeated several times (6 coal seams of workable thickness, ranging in thickness from 2-23 m occur). This relationship further suggests that peat swamp were developed in abandoned channel as well as in distal flood plains.

Interpretation from Entropy Analysis

Upward transitional matrix (p_{ij}), downward transition matrix (q_{ij}) and independent trial matrix (r_{ij}) as defined above were calculated from transition tally matrix (f_{ij}) (Table 3). Using these matrices, entropies $E^{(pre)}$, $E^{(post)}$ and $E^{(system)}$ were calculated for individual states and for the whole sedimentation unit as shown in Table 4. Computed entropy values of individual lithologic state are sub equal to equal implying that the deposition of these lithologies was not a random event for the Barakar cyclothems in the Singrauli sub basin strengthen the result obtained by quasi independence Markov analysis. For coarse to medium sandstone and shale, $E^{(pre)} > E^{(post)}$ implies that sandstone units possibly occur after different lithologic states as recognized in available borehole logs, and may be followed by them. In geological terms, this relation indicates that channel sub environment accumulating coarse to medium grained sandstone developed widely and repeatedly, as recognized in the field by several workers (Casshyap, 1979; Casshyap and Khan, 1982; Casshyap and Tewari, 1984; Tewari, 2005; Tewari and Casshyap, 1983; Tewari and Singh, 2008; Tewari *et al.*, 2009; Hota and Maejima, 2004 and Maejima *et al.*, 2008). The sandstone lithofacies thereby have exerted a considerable influence upon succeeding lithological states. By contrast, the remaining lithological states (shale, carbonaceous shale, coal) indicate $E^{(post)} > E^{(pre)}$, this relation in respect to these lithologies indicates that the dependency of each lithological state on its precursor is stronger than the influence of the lithological state on the successor, which in geological terms suggests that the deposition of each of these lithologies was strongly influenced by the preceding state. Calculated values of $E^{(pre)}$ and $E^{(post)}$ for each lithological state were graphically plotted (Fig. 4). Entropy sets pertaining to sandstone; shale and carbonaceous shale fall linearly along or close to a diagonal line. The plot of entropy set for coal falls slightly outside, indicating that it was most random event. Geologically this indicates that lithofacies C, a deposit of swamp areas is not a regular feature of the alluvial setting in which Barakar cyclothems were deposited. It developed occasionally, as well as locally; but wherever it appears it shows a symmetrical relationship.

Research Article

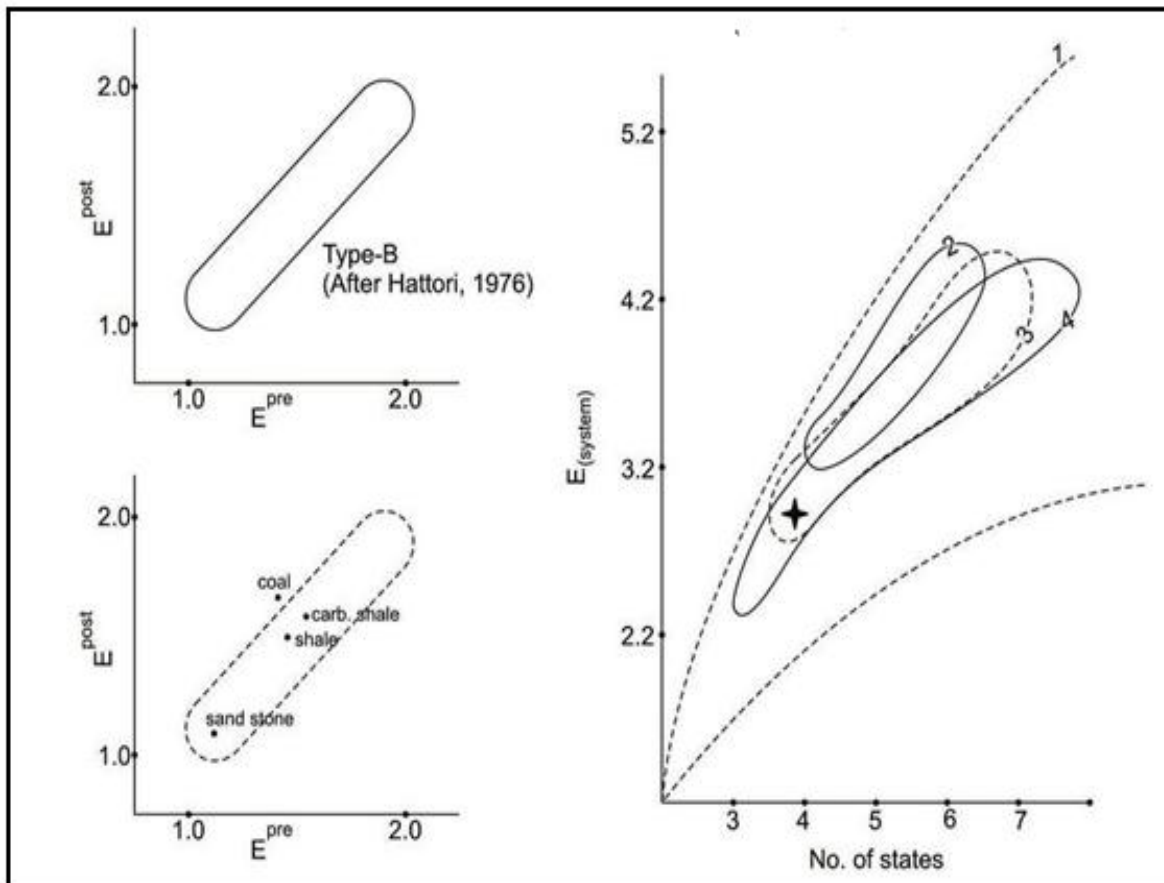


Figure 4: (A) Entropy sets derived from Barakar cyclothem showing Type B (symmetrical) pattern. For comparison the Type B pattern after Hattori (1976) is given at top. (B) Relationships between Entropy and depositional environments of lithologic sequences (after Hattori, 1976). * Entropy for Barakar cyclothem, Singrauli sub-basin

This pattern of entropy sets can be compared with the Type-B cyclical pattern of (Hattori, 1976, Fig. 1) which signifies symmetrical cycles. This inference compares well with a conclusion deduced independently from the quasi independence model suggesting a symmetrical pattern. Indeed, this cyclical pattern for the given Gondwana sub- basin is similar to that reported from other areas (Tewari *et al.*, 2009). Computed value of $E_{(system)}$ lie well within the zone delineated by (Hattori, 1976) as a “fluvial-alluvial environment” (Fig. 4B), thus confirming the dominance of fluvial environment.

Interpretation from Linear Line Equations

In order to illustrate the interrelationships between the listed variables, regression lines were plotted for each separately. The results strongly suggest close relationship between the total thickness of strata on one hand and total thickness and number of constituent lithologies (Figure 5). The varying slope (dy/dx) of regression lines implies that the rate of increase, though regular was not the same for each lithologic type, and was too low for coal seams, perhaps due to greater compaction of vegetal debris and plant subsequent to their formation and subsidence under a thick cover of overlying clastic sediments. This interpretation is supported by the Markov chain analysis described elsewhere.

Research Article

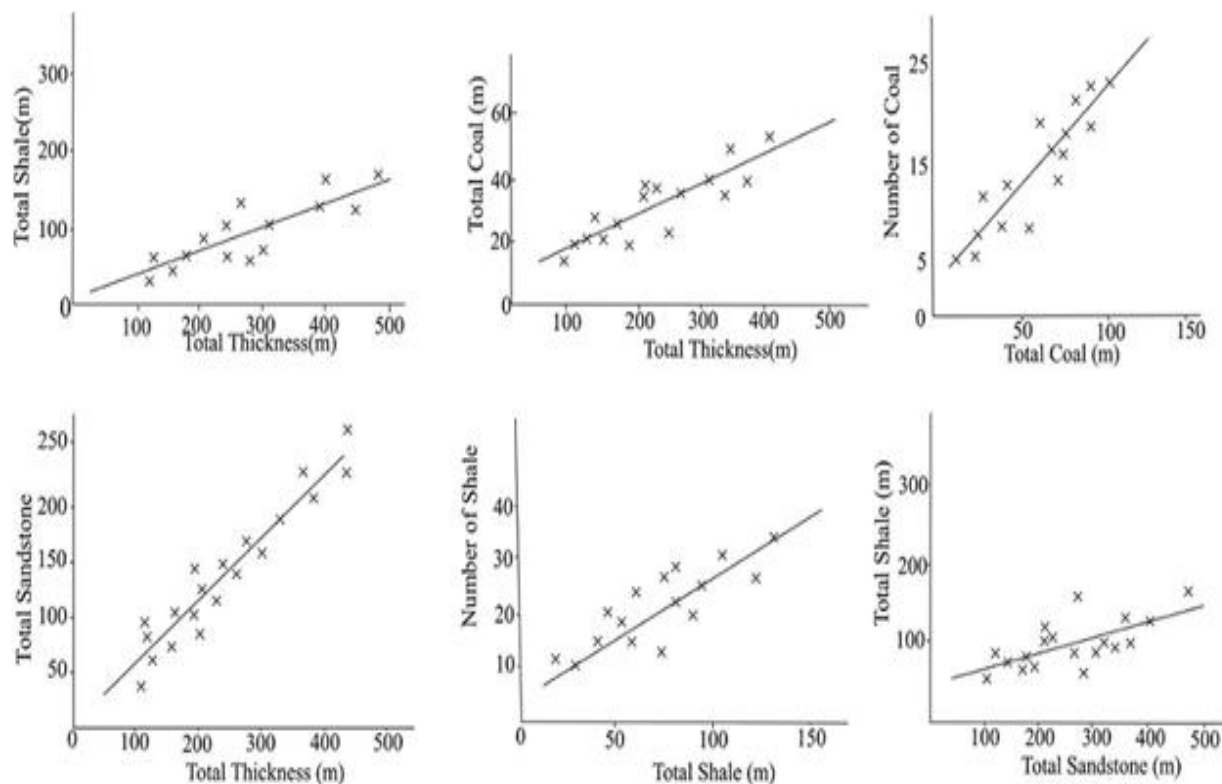


Figure 5: Linear regression lines between some critical lithologic variables of the Barakar cyclothems

The linear regression equation and the correlation coefficient between the number of cycles and the total thickness of strata demonstrate a tendency for number of cycles to be directly proportional to the total thickness of strata and thus to the net subsidence of the Singrauli Gondwana sub-basin. As the definition of a cycle given by (Duff and Walton, 1962) is followed, the number of cycles in a given section represents the number of horizons at which vegetation grew on the depositional surface at that point. Because it is unlikely that vegetation could colonize any surface that was consistently covered by more than a few centimeter of water or that peaty debris would be permanently preserved if it lay above the normal level of the water table. Commonly, these horizons probably represent a period of virtual standstill and rapid submergence brought the period of accumulation of vegetal material to close. During these periods local subsidence must have been sufficiently slow to be balanced by the accumulation of peat and any inorganic sediment that spilled over onto the coal swamp in times of flood. The appearance of clastic sediments above coal signifies a period when the depositional surface sank at rapid rate for vegetation to continue to live. This interpretation gets support from the results of Markov chain analysis where coal is shown by high probability of passage to sandstone (Fig. 3). It is difficult to explain the observed relationship between total thickness of strata and total number of cycles by the hypothesis of control of cyclical sedimentation by eustatic change in sea level (Duff and Walton, 1962) or by widespread diastrophic movement in the basin or in the source area of the clastic sediments, then the number of cycles would be expected to remain the same throughout the basin (Read and Dean, 1967 and Khan and Tewari, 1991). Since the Barakar cyclothems are clearly non-marine in character the theories of sea level changes are not applicable in the present case, however, in the opinion of the authors, the cycles of Barakar cyclothems were formed and essentially controlled by the interrelation of sedimentation and syntectonic (subsidence) rather than by factors external to the basin.

Research Article

Table 6: Correlation matrix of nine litho logical variables (computed after Standardization in respect to zero mean and unit deviation)

	A	B	C	D	E	F	G	H	I
A	1.000								
B	0.990	1.000							
C	0.671	0.578	1.000						
D	-0.125	-0.101	-0.171	1.000					
E	0.774	0.777	0.516	-0.056	1.000				
F	0.532	0.435	0.769	-0.195	0.423	1.000			
G	0.205	0.131	0.323	-0.105	0.031	0.670	1.000		
H	-0.242	-0.196	-0.495	0.475	-0.144	-0.418	-0.094	1.000	
I	0.197	0.259	-0.045	0.056	0.223	-0.246	-0.399	-0.086	1.000

A= Total thickness of strata

E= Total nos. of sandstone beds

B= Total thickness of sandstone bed

F= Total nos. of shale beds

C= Total thickness of shale bed

G= Total nos. of coal seams

D= Total thickness of coal seams

H= Ratio of sandstone/shale

I= Ratio of clastic sediments to coal

Correlation Coefficient: The basic data computed from 37 borehole logs is arranged in 37X9 matrixes where 9 refer to the number of lithological variables. The data was normalized following the procedure outlined by (Davis, 2002) and then correlation coefficients were computed for each pair of variables. The correlation coefficients between 16 pairs out of 36 pairs show fairly good positive correlation whereas ratio variables (sandstone/shale) ratio and (sandstone+shale)/coal ratio show less degree of correlation possibly in view of their dependency on other lithologic variables. Inspection of the top row of Table 6 reveals that, with the exception of the two “ratio variables (B8 and B9)” all the lithological variables tend towards linear relationship with the total thickness of strata (B1).

Interpretation from Principal Component Analysis

In the principal component analysis of data from all variables (B1 to B9)), only the first three eigenvalues prove to be greater than unity (Table 7), so that according to (Kaiser, 1948) criterion the first three eigenvectors should be used as reference axes. Such a system, which would account for 79.15% of the total variance of the nine observed variables, is illustrated graphically in Figure.

Table 7: Principal Ax's matrix derived from Table 6

Eigenvectors									
Variable	A	B	C	D	E	F	G	H	I
A	0.456	0.430	0.428	-0.137	0.384	0.396	0.197	-0.242	0.048
B	-0.214	-0.285	0.098	-0.216	-0.294	0.361	0.516	-0.178	-0.547
C	0.123	0.122	-0.043	0.624	0.133	0.116	0.306	0.615	-0.274
D	0.184	0.233	-0.374	-0.625	0.216	-0.213	-0.022	0.347	-0.407
E	0.025	0.058	-0.203	-0.238	-0.214	0.010	0.671	0.235	0.644
F	-0.360	-0.424	0.269	-0.246	0.499	0.303	-0.130	0.406	0.184
G	0.170	0.096	0.578	-0.191	-0.586	-0.036	-0.247	0.429	-0.019
H	0.064	0.120	-0.455	-0.003	-0.247	0.748	-0.376	0.071	0.080
I	-0.730	0.676	0.091	-0.006	-0.008	0.015	0.012	0.007	-0.023
Eigenvalues									
	3.955	1.857	1.310	0.693	0.503	0.302	0.271	0.104	0.001
Total variance (%) contributed by each eigenvalue									
	43.95	20.64	14.56	7.77	5.59	3.35	3.02	1.16	0.002

Research Article

Table 8: R-mode factor matrix derived from correlation matrix of nine litho logical variables

Factor									
Variable	F1	F2	F3	F4	F5	F6	F7	F8	F9
A	0.907	-0.292	0.142	0.153	0.018	-0.198	0.088	0.021	-0.025
B	0.855	-0.389	0.139	0.195	0.041	-0.233	0.050	0.039	0.023
C	0.853	0.134	-0.049	0.312	-0.144	0.148	0.301	-0.147	0.003
D	-0.273	-0.294	0.714	-0.520	-0.169	-0.135	-0.099	-0.002	-0.002
E	0.764	-0.402	0.153	0.180	-0.152	0.274	-0.305	-0.079	-0.003
F	0.787	0.493	0.132	-0.178	0.007	0.166	-0.018	0.242	0.005
G	0.392	0.703	0.350	-0.018	0.437	-0.071	-0.129	-0.121	0.004
H	-0.482	-0.244	0.704	0.289	0.166	0.223	0.233	0.023	0.002
I	0.096	-0.746	-0.313	-0.339	0.456	0.101	-0.009	0.026	-0.008

Table 9: First three factors extracted from Table 6 and communality (accounting for 79.2% of the total variance)

Variable	Factor -I	Factor-II	Factor-III	Communality
A	0.907	-0.292	0.142	0.964
B	0.855	-0.389	0.139	0.949
C	0.853	0.134	-0.049	0.864
D	-0.273	-0.294	0.714	0.819
E	0.764	-0.402	0.153	0.876
F	0.787	0.493.	0.132	0.938
G	0.392	0.703	0.350	0.878
H	-0.482	-0.244	0.704	0.887
I	-0.096	-0.746	-0.313	0.815
Eigen value	3.955	1.857	1.310	

This shows the projections of the nine variable-vectors, all of which are of unit length, on to the plane defined by the first three eigenvectors, all of which are also of unit length. In this three dimensional system the communalities for all variables are reasonably good but that for the total thickness of shale (B3) and number of sandstone members (B5) are lower than the others (Table 9).

Figure 6 A, B and C show the projections of each of the variable vectors on to the plane defined by every possible pair of the first three eigenvectors. Figure 6A which show the plane defined by the first three eigenvectors demonstrate the clustering of the four variable vectors namely the total thickness of sandstone (B2), total thickness of shale beds (B3), total thickness of coal seams (B4) and the number of sandstone beds (B5). The vector for the total thickness of strata (B1) and the number of shale beds (B6) are somewhat more loosely associated with this group, nevertheless, the communalities is fairly high, so that these variables are particularly well represented in the system. The vector s for the sand/shale ratio (B8) and clastic ratio (B9) lies remote from all the other variable vectors. Figure 6 B in which second and third eigenvectors are used show similar grouping of vectors as does the corresponding diagram (Fig. 6A). The relationship between these lithological variables remain similar to those already noted but the relative closeness of the vector for the vectors for the total thickness of strata (B1) and number of shale beds (B6). Of the two vectors representing “ratio” variable, which for the clastic sediment to coal (B9) lies close from all the other variable vectors around the first eigenvector. Figure 6C, which shows the projections of the variable vectors on to the plane defined by the second and third Eigen vectors does not show as clear a grouping of vectors as does the Figure 6A and 6B. The clustering of the vector variables about the first eigenvector (Fig. 6) demonstrate how these lithological variables tends towards a linear relationship with each other and how all tend to be related to the total thickness of strata i.e. net subsidence of the basin strengthen the views expressed by (Casshyap *et al.*, 1988; Khan and Tewari, 2007

Research Article

and Tewari, 2008). In geological terms the clustering of vectors around the first eigenvector strongly suggest a close relationship between the total thicknesses of strata on the one hand and total thickness of sandstone beds, the total thickness of shale beds, the total thickness of coal seams of any thickness and number of sandstone and shale beds. This leads to a striking conclusion that a given increase in the total thickness of strata or net subsidence of the basin tends to result in an increase in the number of constituents lithologies and their thicknesses.

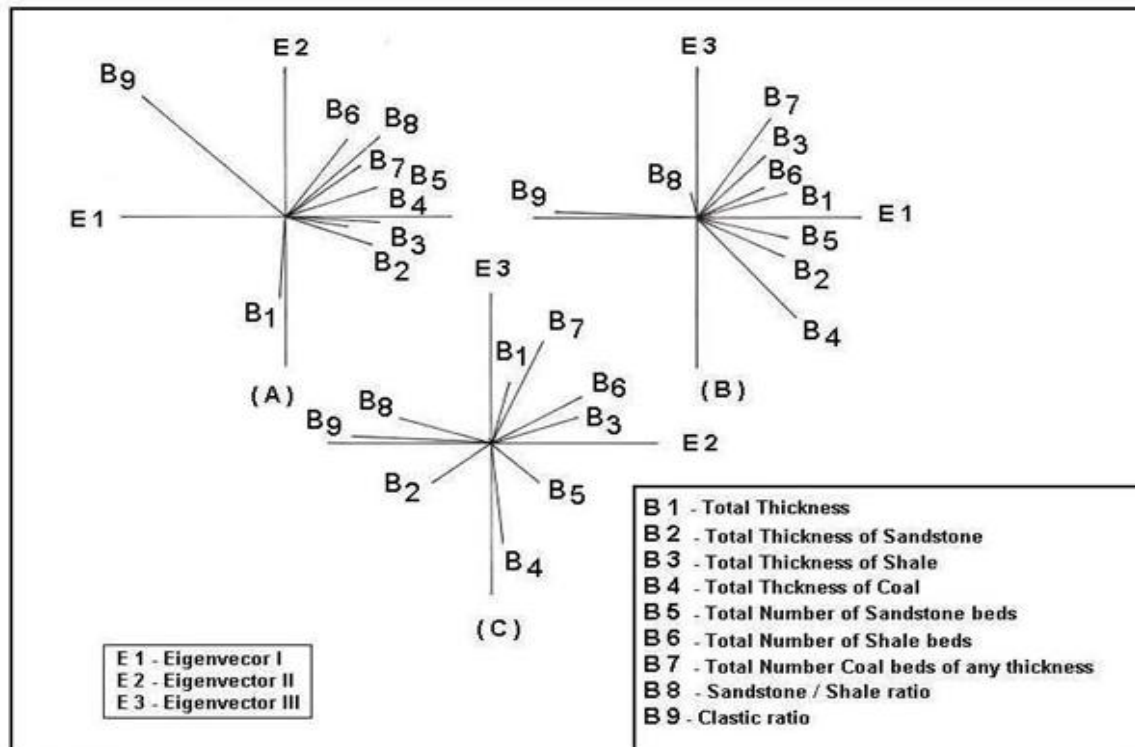


Figure 6: Projection of lithologic variable vectors of unit length on the planes defined by the First and Second (A), first and Third (B) and Second and Third (C) eigenvector in Principal Component Analysis

This may indicate, in turn, that the essential components of the depositional framework did not change materially even though their relative importance was not always the same as reflected in the compositional variation displayed by rocks of the Barakar coal measures (Aslam *et al.*, 1991). Poor relation or lack of it in the total thickness of coal in the Singrauli Gondwana sub-basin may imply inconsistent correlation between the factors, which determine thickness of strata, and those, which control coal formation. The depositional changes from levees to back swamp and/or coal swamp environment possibly explain this tendency. The fine clastics (shale/siltstone) represent over bank deposits in adjoining flood plains that choked the development of peat accumulation. Alternatively, it may be due to the progressive change in paleochannel morphology and paleohydrodynamic through time from bed load streams to mixed load streams (Casshyap and Tewari, 1984; Tewari, 2005 and Khan, 2013).

Interpretation from Factor Analysis

Data matrix consists of nine lithological variables were reduced to three factors (Table 9), and these together explain 79.15% of the total variance. The first factor (F-I) has the highest eigenvalue of 3.955 and it also explain 43.95% of the total variance, whereas, the second factor (F-II) having the eigenvalue of 1.857, explains 20.64% of the total variance. The eigenvalue of the third factor (F-III) is 1.310 and it

Research Article

explains 14.56% of the variance. Among the three factors, the third factor does not have the loading of either the total thickness of strata nor is number of constituent lithologies. This third factor has no role for discussion following recommendations of several workers.

The Factor-I, the total thickness of strata (0.907), the total thickness of sandstone beds (0.855), the total thickness of shale (0.853), the number of sandstone beds (0.764) and number of shale beds (0.788) have been significantly loaded (Table 9). The highest positive loading in respective lithologies indicates that the contribution of the lithological variables increases with the increasing loading in a dimension. In another words that an increase in total thickness i.e. net subsidence is due to increase in the thicknesses and the number of constituent lithologies of the Barakar coal measures in Singrauli sub-basin. This geological interpretation get support from linear regression line as well as from the principal component analysis as discussed earlier. Since the total thickness (net subsidence) shows, a negative correlation with the total thickness of coal seams may imply that factors that controlled the formation of coal and those, which determine net subsidence, are inconsistent. The coal forming environment whether a lake or marsh (swamp), is apparently not a normal feature of alluvial flood plains (Strahler, 1963) and probably neither was this so during the deposition of Barakar coal measures in Singrauli Gondwana sub-basin. The inverse relationship as described above is reported, especially from Late Paleozoic coal-bearing deposits of India (Tewari, 2008) and also from the Carboniferous coal measures of different parts of world (Johnson and Cook, 1973 and Casshyap, 1975). To sum up, principal component analysis and factor analysis, as applied in present study, yield similar results, which confirm those of the regression line described elsewhere in (Khan and Tewari, 1991 and 2007 and Tewari, 2008).

Conclusion

Improved Markov chain analysis was applied to 37 data sets derived from the Barakar coal measures successions which were laid down in fluvial environment, all show a definite tendency toward a preferred path of upward lithological transitions as follows:

Sandstone (SS) → Shale (SH) → carbonaceous shale (CS) → Coal (C) → Sandstone (SS)

Entropy analysis has shown that these coal measures cyclic units probably were essentially of Type-B symmetrical cyclic pattern

The number of coal bearing cycles of the Barakar cyclothems seems to be intimately related to the thickness of the strata, indicating that the local processes of sedimentation and subsidence have been important in the development of cyclothems. Most of the Barakar cyclothems seem to be explained in terms of sedimentation variations in area undergoing differential subsidence. The varying slope of regression lines implies that the rate of increase though regular was not the same for each constituent lithologies and was too low for coal, perhaps due to greater compaction of vegetal debris and peat swamp subsequent to their formation and subsidence under a thick cover of overlying clastic sediments. This geological interpretation is supported by the Markov chain results discussed above.

The quantitative results of principal component analysis and factor analysis confirm those of linear regression lines. In both types of multivariate analyses the mutual interrelationships between lithological variables can be discerned in system with two orthogonal reference axes. The results strongly suggest a close interrelationship between the total thickness of strata on one hand and total thicknesses and number of constituent lithologies on the other hand. This leads to a striking conclusion that a given increase in the total thickness of strata or net subsidence of the basin tends to result in an increase in the number of lithologic members and their thicknesses. However, a similarity of pattern of regression lines together with high factor loading and corresponding high communalities on constituent lithologies suggest that a balance was maintained throughout the deposition of Barakar cyclothems between the rate of deposition and the rate of subsidence. The multivariate principal component and factor analysis corroborating quasi-independence Markov analysis, entropy analysis and regression lines. The study provides quantitative data base in identifying potential depositional sites of peat formation in the Early Permian fluvial coal measures succession. We hope that the present case study may stimulate other workers to investigate

Research Article

other fluvial successions, so that more evidence can be accumulated in an attempt to solve the problem of sedimentary and tectonic processes in the formation of cyclically deposited successions.

ACKNOWLEDGEMENT

We thank the Director, Geology and Mining UP (ZAK) and the management of the Sri J N P G College, Lucknow (RCT) for providing working and library facilities. Our sincere appreciation is due to Mr. S K Agarwal who processed the data at the Geology & Mining Computer lab.

REFERENCES

- Anderson TW (2003).** An introduction to multivariate statistical analysis. 3rd edition John Wiley and Sons 721.
- Aslam M. Arora M and Tewari RC (1991).** Heavy mineral suite in the Barakar sandstone of Moher-subbasin in Singrauli coalfield, central India. *Journal Geological Society of India* **38** 66-75.
- Bogg, Jr S (2005).** Principles of Sedimentology and Stratigraphy. 2nd edition Printis and Hall, UK 688.
- Casshyap SM (1975).** Cyclic characteristics of coal bearing sediments in Bochumer Formation (Westphel A2) Ruhrgebeit, Germany. *Sedimentology* **22** 236-256.
- Casshyap SM (1979).** Patterns of sedimentation in Gondwana basins, Fourth International Gondwana Symposium, Calcutta **2** 525-551.
- Casshyap SM and Khan ZA (1982).** Discrete Markov analysis of Permian coal measures of East Bokaro basin, Bihar, India. *Indian Journal of Earth Science* **9**.
- Casshyap SM Kreuser T and Wopfner H (1987).** Analysis of cyclical sedimentation in the lower Permian Machuchuma coalfield (SW Tanzania). *Geological Rundschau* **76** 869-883.
- Casshyap SM and Tewari RC (1984).** Fluvial models of the lower Permian Gondwana coal measures of the Koel-Damodar and Son-Mahanadi basins, India, In 'Sedimentology of Coal and Coal bearing sequences'. Rahmani RA and Flores RM (edition), *Special Publication of International Association of Sedimentologists* **7** 121-147.
- Casshyap SM Tewari RC and Khan ZA (1988).** Interrelationship of stratigraphic and lithological variables in Permian fluvial Gondwana coal measures of eastern India. *Proceedings of Symposium on Quantitative Stratigraphic Correlation*, IIT, Kharagpur 111-125.
- Cooley WW and Lohnes PR (1971).** Multivariate Data Analysis. John Wiley and Sons New York 164.
- Davis JC (2002).** Statistics and Data analysis in Geology. 3rd Edition John Wiley and Sons New York 639.
- Duff PMcLD Walton EK (1962).** Statistical basis for cyclothems: a quantitative study of the sedimentary succession in the East Pennine Coalfield. *Sedimentology* **1** 225-235.
- Fisher RA and Yates F (1963).** Statistical Tables for Biological, Agricultural and Medical Research 6th edition Oliver and Boyd, Edinburg 646.
- Ghosh PK and Mitra ND (1972).** A review of the recent progress in the studies of Gondwana of India. *Second Gondwana Symposium South Africa* 29-47.
- Harbough JH and Bonham-Carter G (1970).** Computer simulation in Geology John Wiley and Sons New York 575.
- Harris RJ (1985).** Primer of Multivariate statistics. 2nd Edition Academic Press, Florida 576.
- Hattori I (1976).** Entropy in Markov chain and discrimination of cyclic patterns in lithologic succession. *Mathematical Geology* **8** 477-497.
- Hota RN 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 Research* **7** 353-362.
- Johnson KR and Cook AC (1873).** Cyclic characteristics of sediments in the Moon Island Beach Subgroup, New Castle Coal measures, New South Wales. *Mathematical Geology* **5** 91-110.

Research Article

- Kaiser HF (1968).** The varimax criterion for analytical rotation of factor analysis. *Psychometrika* **23** 187-200.
- Khan ZA (1978).** Lithofacies, sedimentation trends and paleoflow characters of Karharbari and Barakar strata in East Bokaro Coalfield, Bihar. *PhD Thesis Aligarh Muslim University* 297.
- Khan ZA (2007).** Q mode analysis of cyclically deposited Barakar coal measures of Gondwana basins, eastern and central India. 12th Conference, International Association of Mathematical Geology, Beijing, China **702-707**.
- Khan ZA (2013).** Paleochannel metamorphosis in Late Paleozoic Gondwana Rivers of Peninsular India. In Earth Resources and Environment (edition) Venkatachalapathy R *Research Publishing Services, Singapore* 245-261.
- Khan ZA and Casshyap SM (1981).** Entropy in Markov chain analysis of Late Paleozoic cyclical coal measures of East Bokaro basin, Bihar. *Mathematical Geology* **13** 153-162.
- Khan ZA and Casshyap SM (1982).** Sedimentological synthesis of Permian fluvial sediments of East Bokaro basin, Bihar. *Sedimentary Geology* **33** 111-128.
- Khan ZA and Tewari RC (1991).** Net subsidence and number of cycles: their interrelationship in different Permian Gondwana basins of Peninsular India. *Sedimentary Geology* **73** 161-173.
- Khan ZA Tewari RC (2007).** Quantitative model of early Permian coal bearing Barakar cycles from Son-Mahanadi and Koel-Damodar Gondwana coalfields of eastern-central India. *Gondwana Geological Magazine* **9** 115-125.
- Khan ZA and Tewari RC (2010).** Multivariate analysis of coal bearing Barakar cyclothem (late Paleozoic) in Singrauli sub-basin, central India. *Gondwana Geological Magazine* **12** 133-140.
- Krumbein WC (1968).** Fortran IV computer programme for simulation of transgression and regression with continuous time Markov models. *Contributions of Geological Survey Kansas* **26** 38.
- Lawley DN and Maxwell AE (1971).** Factor analysis as a statistical method. 2nd Edition *Bullerworth and Company London* 153.
- Maejima W Tewari RC and Hota RN (2008).** On the origin of Barakar coal bearing cycles in Gondwana basins of Peninsular India. *Journal of Geosciences, Osaka City University, Japan* **51** 21-26.
- McCommon RB (1966).** Principal component analysis and its application to large scale correlation studies. *Journal of Geology* **74** 721-733.
- Miall AD (2000).** Principles of Sedimentary Basin Analysis. *Springer Verlag, New York* 582.
- Morrison DF (1990)** Multivariate Statistical Methods. 3rd edition *McGraw Hill New York* 495.
- Power DW and Easterling GR (1982).** Improved methodology for using embedded Markov chains to describe cyclical sediments. *Journal of Sedimentary Petrology* **53** 913-923.
- Read WA (1969).** Analysis and simulation of Namurian sediments in central Scotland using a Markov process model. *Mathematical Geology* **1** 199-219.
- Read WA and Dean JM (1967).** Cycles and subsidence: Their relationship in different sedimentary and tectonic environments in the Scottish Carboniferous. *Sedimentology* **23** 107-120.
- Tewari RC (2008).** Net subsidence and evolution of coal swamps in Early Permian coal measures of eastern Indian Gondwana basins using Principal component analysis. *Journal of Geosciences* **51** 27-34.
- Schwarzacher W (1975).** Sedimentation models and Quantitative Stratigraphy. *Elsevier Publishing Company, New York* 396.
- Strahler AN (1963).** The Earth Science. Harper and Row, New York 681.
- Tewari RC (1997).** Numerical classification of coal bearing cycles of Early Permian Barakar coal measures of eastern-central India Gondwana basins using Q-mode cluster analysis. *Journal of Geological Society of India* **50** 593-599.
- Tewari RC (2004).** Sedimentological synthesis of Gondwana formations of Peninsular India and its bearing on stratigraphy. *Indian Journal of Petroleum Geology* **13** 75-85.
- Tewari RC (2005).** Tectono-stratigraphic-sedimentary events in Gondwana succession of Peninsular India. *Journal of Geological Society of India* **65**(5) 636-638.

Research Article

Tewari RC (2008). Net subsidence and evolution of coal swamps in Early Permian coal measures of eastern Indian Gondwana basins using Principal component analysis. *Journal of Geosciences, Osaka City University, Japan* **51** 27-34.

Tewari RC and Maejima W (2010). Origin of Gondwana basins of Peninsular India. *Journal of Geosciences* **53**(5) 43-49.

Tewari RC and Casshyap SM (1983). Cyclicity in Early Permian fluvial Gondwana coal measures: an example from Giridih and Saharjuri basins, Bihar, India. *Sedimentary Geology* **35** 297-312.

Tewari RC, Hota RN and Maejima W (2012). Fluvial architecture of Early Permian Barakar rocks of Korba Gondwana basin, eastern-central India. *Journal Asian Earth Science* **52** 43-52.

Tewari RC Singh DP (2008). Permian Gondwana paleocurrents in Bellampalli coal belt of Godavari Valley basin, Andhra Pradesh and paleogeographic implications. *Journal of Geological Society of India* **71** 266-270.

Tewari RC, Singh DP and Khan ZA (2009). Application of Markov chain and entropy analysis to lithologic succession- an example from the early Permian Barakar formation, Bellampalli coalfield, Andhra Pradesh, India. *Journal of Earth System Science* **118** 583-596.

Turk G (1979). Transition analysis of structural sequences: Discussion. *Geological Society of America Bulletin* **90** 989-991.

Veevers JJ and Tewari RC (1995). Gondwana master basin of Peninsular India between Tethys and the interior of the Gondwanaland province of Pangea. *Geological Society of America Bulletin Memoir* **187** 72.

Vistelius AH (1949). The mechanism of formation of sedimentary beds. *Doklady Academy Nauk SSSR* **65** 191-194.