A Statistical Correlation of Sulfur Content in Coal 5 and the Overlying Gray Shale

S. N. Ghose
Amax Coal Company, Indianapolis, Indiana 46225

Introduction

The marketability of coal today largely depends on sulfur content - ash and BTU being comparable. Low sulfur coal, which is not abundant in the Midwest, is in great demand and to an extent, may significantly affect the future of the coal industry in that region. Competition for low sulfur coal has influenced the midwestern coal industry from exploration to operation.

The association of relatively low sulfur coal with overlying thick, freshwater gray shale is a familiar fact in the geology of the Illinois Basin. This interesting phenomenon is mentioned in the literature by various authors such as Hopkins, Wier, Friedman, Eggart and others. Hopkins (1968) described this association in detail and the function of the gray shale in preventing exterior sulfur contamination of the coal. In this paper, an effort is made to establish an empirical relation between these two variables: gray shale thickness and total sulfur content. Based on these variables a predictive scale is constructed to approximate the total sulfur content per given gray shale thickness for quick appraisal. This correlation is based on the data of Coal 5 which is a major commercial seam in the Illinois Basin. Coal 5 shows significant uniformity in quality, thickness and roof lithology. Also, unlike Coal 6, Coal 5 does not show any persistent parting (blue band). Therefore, the depositional environment of Coal 5 is probably less complicated and hence is easier to correlate and predict the total sulfur content. However, the method presented here also can be applied to other coal seams for predicting total sulfur content under similar geological conditions.

Coal 5 in the Illinois Basin normally shows two distinctive roof lithologies: extensive marine black shale and local fresh water gray shale. Although several geological conditions can lead to the deposition of a fresh water gray shale, the gray shale encountered in Coal 5, in many parts of the basin, is that of an active channel origin. Periodic flooding of the coal swamp near the channel deposited the fine sediments on Coal 5 near the channel as shown in Figure 1 and Figure 2.

Procedure

In this study ninety analyses of Coal 5 core samples and the corresponding thickness of overlying gray shale are used as the data base for defining the attitude of the curve shown in Figure 3. The characteristics of these two variables are correlated and categorized at 5 feet intervals. Each 5 feet of increment (unit block) provides one mean position on the curve. Presumably, the smaller the unit block, the greater expression of curve-detail which in turn adds to the accuracy of the method. As shown in Figure 2, the degree of scatter is considerable, and the overlapping of data is unavoidable if point to point correlation is desired. The scatter of data points represents variations in the depositional system. Therefore, to obtain a meaningful interpretation within the limits of practicality, 5 feet unit blocks were chosen. To provide a measure of reliability the basic statistical parameters of error distribution are estimated for each incremental 5 feet (Table 1 and Table 2). The data represented in each block are free from systematic error

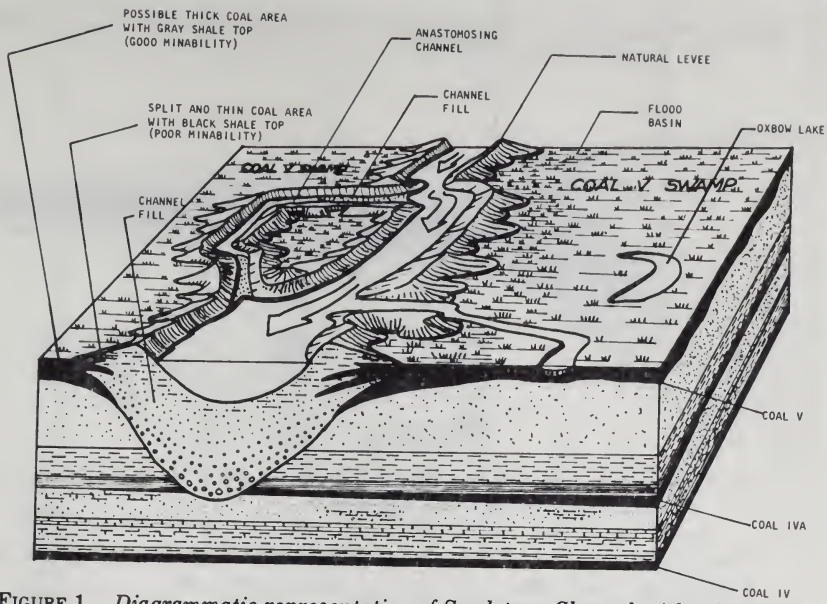

FIGURE 1. *Diagrammatic representation of Sandstone Channel with relation to Coal V Swamp Environment*

because $\sum_{i=1}^{n} vi = 0$.

Sample means from each 5 feet interval block were fitted by computer. The best fit is a polynomial regression of degree two. The observed and estimated sulfur percentages are given below.

Gray Shale Thickness	Observed (Percent)	Estimated (Percent)
5	3.9860	3.6465
10	3.9600	3.1096
15	1.8575	2.6288
20	1.5360	2.2042
25	1.5162	1.8358
30	1.3850	1.5235
35	1.2750	1.2674
40	1.1175	1.0674
45	1.3375	.92350
50	0.7100	.83579
55	1.1175	.80422
60	1.3855	.82880
65	.9340	.90953
75	.70667	1.2394

To make this correlation easily usable and to avoid overlapping errors further adjustment and refinement is necessary. This refinement is achieved by expanding the incremental gray shale unit or condensing the predictive scale. A 4 step scale is proposed here, which is more compatible with the margin of error and allows the steps to be more easily identified.

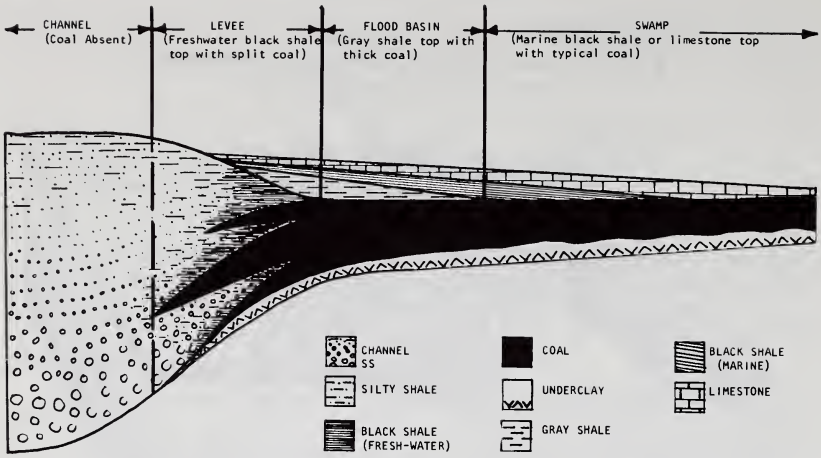

FIGURE 2. *Schematic cross-section relating environments of deposition to coal and roof lithologies*

Black Shale	3.5% to 6% or more sulfur
2 to 10 feet of gray shale	2% to 3.5% sulfur
10 to 25 feet of gray shale	1.5% to 2% sulfur
25 to 80 feet of gray shale	0.5% to 1.5% sulfur

Analysis

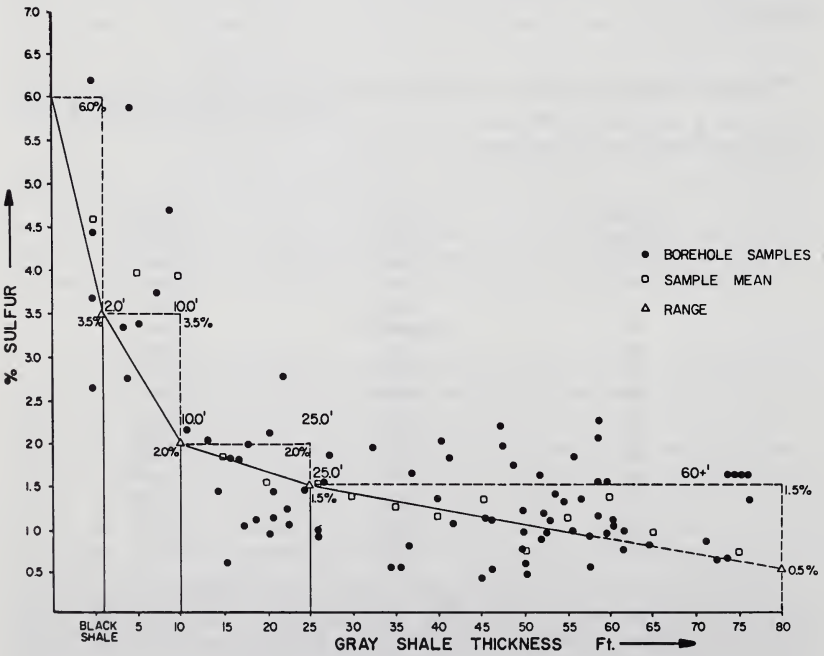

FIGURE 3. *Correlation of Total Sulfur Percentage & Gray Shale Thickness*

TABLE 1. *SUMMARY OF ESTIMATES*

Gray Shale	Black Shale	0 -	5 -	10 -	15 -	20 -	25 -	30 -	35 -	40 -	45 -	50 -	55 -	60 -	65 -	70 -75
(% of Sulfur) Mean = $\bar{x}$ (Sample)	4.60	3.986	3.960	1.858	1.536	1.516	1.385	1.275	1.118	1.338	0.7100	1.118	1.385	0.9340		0.70667
Σvi =	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0
Σv_i^2 =	9.178	5.720	0.990	0.329	0.695	2.776	0.622	0.966	0.693	1.563	0.117	0.848	2.728	0.087		0.309
Average Error =	1.84	1.273	0.527	0.268	0.695	0.471	0.385	0.695	0.398	0.390	0.150	2.120	0.432	0.123		0.0956
Maximum Error =	– 1.650	– 1.914	– 0.790	0.408	0.476	– 0.604	– 0.485	0.695	– 0.553	– 0.888	– 0.260	0.648	– 0.854	0.174		– 0.143
Median Error =	0.15	0.606	0.230	0.120	0.284	0.201	0.280	0.695	0.280	0.175	0.110	0.075	– 0.063	– 0.046		– 0.067
Probable Error (50%) =	1.02	0.954	0.475	0.224	0.281	0.629	0.404	0.663	0.324	0.487	0.133	0.234	0.352	0.099		0.083
Standard Deviation σ	1.514	1.414	0.703	0.331	0.417	0.629	0.599	0.982	0.480	0.722	0.198	0.348	0.522	0.148		0.124
Variance of the Sample $\sigma_{\bar{x}}$ = Mean	0.677	0.816	0.406	0.165	0.186	0.222	0.299	0.695	0.240	0.361	0.099	0.120	0.157	0.066		0.0717
Variance σ^2 =	1.836	1.90	0.330	0.824	0.139	0.347	0.156	0.483	0.173	0.391	0.029	0.105	0.248	0.017		0.010

TABLE 2. *FREQUENCY DISTRIBUTION*
Sulfur Percentages

	Total No. of Plots (Boreholes)	.5%	1%	1.5%	2%	2.5%	3%	3.5%	4%	4.5%	5%	5.5%	6%
BS	5					1		1		1			1
GS 5	3					1	1					1	
10	3					1	1					1	
15	4		1	1	2								
20	7	1	2	3						1			
25	8	1	5		1	1							
30	4	1	1	2									
35	2	1		1									
40	3	2		1									
45	5	1	2	1	1								
50	10	4	3	2	1								
55	8	3	4	1									
60	10	4	2	3	1								
65	5	3	2										
70													
75	3	3											
80	1		1										

NOTE: Frequency of data point occurences in each five feet block.

A comparative study of statistics to test the relation between the base and the predicted quantities and homogeneity of the population indicates moderate agreement.

The covariance shows that there is a positive correlation (both increasing)

$$\text{covariance} = \hat{\varrho}_{xy} = +0.042772$$

$$\underset{0.042772}{|\hat{\varrho}_{xy}|} < \underset{0.116227}{|\hat{\varrho}_x \hat{\varrho}_y|}$$

The covariance is smaller than the standard error of the two sets combined which indicates that these two sets are in good agreement.

The correlation coefficient (Pxy) between these two sets is 0.3680007 which is moderate. This suggests that other variables such as seam thickness and distance from the cut out affect the correlation.

An F test was given to check the equality of variances of these two sets. In other words hypotheses, Ho: $\hat{\varrho}_1^2 = \hat{\varrho}_2^2$ was tested against

$\hat{\varrho}_1^2 \neq \hat{\varrho}_2^2$ at the

significance level $\propto = .10$ ($1-\propto$ = confidence level). The degrees of freedom for both sets are taken as 8.

$$F = \frac{\hat{\varrho}_1^2}{\hat{\varrho}_2^2} = 0.3947540$$

$$\propto = .10,\ 1-\propto/2 = 0.95,\ \propto/2 = .05$$

$$F_{DF_1 DF_2}(1-\propto/2) = \frac{1}{F_{DF_2 DF_1}(\propto/2)}$$

From the table,

$$F_{8,8,0.95} = \frac{1}{F_{8,8,0.95}} = \frac{1}{3.44} = 0.2907$$

$0.39 > 0.29$ do not reject the hypothesis

$0.39 < 3.44$ do not reject the hypothesis

Conclusion

The method described in this paper, can be used not only in the regional type of study but also as an aid to select the most viable prospect among many for exploration. Also, in the developmental stages the method can be utilized to selectively increase the density of information. Further study of the homogeneity of the population, and analysis of data from different seams may provide a quantitative base from which environments of deposition can be assessed in order to locate good quality coal.

Acknowledgements

I am thankful to Dr. C. E. Wier for his permission to publish this paper. Mr. D. A. Darko and Mr. T. R. McCarthy painstakingly reviewed the manuscript and offered a number of useful comments for which I am grateful. I also wish to thank Dr. K. Badiozamani for his help in curve fitting by the computer.

Literature Cited

1. BEVINGTON, P. R. 1969. Data Reduction and Error Analysis for the Physical Sciences: McGraw-Hill Book Company, N.Y.
2. EGGART, D. L. and S. C. ADAMS. 1979. Distribution of Fluvial Channel Systems Contemporaneous with the Springfield Coal Member (Middle Pennsylvanian) in Southwestern Indiana: Ninth International Congress of Carboniferous Stratigraphy and Geology, University of Illinois, Urbana-Champaign.
3. FRIEDMAN, S. A. 1955. Split and Channel Sandstone Cut Out in Coal V in the Dresser Area, Vigo County, Indiana: Indiana Academy of Science, v. 65, P 165-168.
4. HOPKINS, M. E. 1968. Harrisburg (No. 5) Coal Reserves of Southern Illinois; Illinois Geol. Surv. Circ. 431.
5. PUGH, E. M. and G. H. WINSLOW. 1966. The Analysis of Physical Measurements: Addison-Wesley, Reading, Massachusetts.
6. WIER, C. E. 1952. Distribution, Structure and Mined Areas of Coal in Sullivan County, Indiana: Indiana Geological Survey Prelim. Coal Map 2. Wier, C. E., and Hutchison, H. C., 1977, Reduction of Sulfur in Indiana Coal by Washability Techniques: Department of Natural Resources, Geological Survey Bulletin 55.