

# Geostatistical Simulation under Orthogonal Transformed Indicator Model

Türkan CENGİZ<sup>1</sup> and Erhan A. TERCAN<sup>2</sup>

<sup>1</sup> General Directorate of Mineral Research and Exploration, Ankara, Turkey

<sup>2</sup> Department of Mining and Engineering, Hacettepe University, Ankara, Turkey

**ABSTRACT:** Conditional cumulative distribution function (ccdf) plays an important role in geostatistical estimation and sequential simulation. A variety of method for estimating conditional distribution functions is suggested. These are classified as parametric and nonparametric. This study is concerned with nonparametric approach, in particular, orthogonal transformed indicator method. Orthogonal transformed indicator method (OTIM) is a compromise between the two extremes of indicator cokriging and indicator kriging. It requires less estimation and modelling over indicator cokriging and uses more information over indicator kriging. The idea behind this approach is to transform the indicator functions into a set of spatially orthogonal functions (factors) and to use the autokrigability property of these functions. This paper includes an application of OTIM to estimation and simulation of the thickness data of the upper lignite seam of the Kalburçayırı field, Kangal Basin, Sivas, Turkey. The Cholesky-spectral algorithm is used as orthogonalization algorithm.

**KEY WORDS:** spatial orthogonalization, conditional distribution, stochastic simulation.

## Introduction

In mining and geological analysis the assessment of uncertainty about an unknown value at an unsampled location is one of the most important problems. The unknown values could be estimated by linear geostatistical techniques (variogram, kriging). However, linear geostatistics contains many problems due to its data independent character. In these techniques the estimated values are smoothed, showing less variability in respect to real one. This has an important effect in all phases of a mining project, including feasibility, mine planning and production scheduling. One way around this problem is to use conditional distribution functions that completely solve data-independence problem. A variety of method for estimating conditional distribution functions is suggested. These are classified as parametric and nonparametric. This study is concerned with nonparametric approach, especially orthogonal transformed indicator method of this approach.

The conditional distribution functions and their nonparametric estimation are described in Goovaerts (1997), Tercan and Kaynak (1999), and Tercan (1999). Orthogonal transformed indicator kriging (OTIM) is a compromise between the two extremes of indicator cokriging and indicator kriging. It requires less estimation and modeling over indicator cokriging and uses more information over indicator kriging. The idea behind this approach is to transform the indicator functions into a set of spatially orthogonal functions (factors) and to use the autokrigability property of these functions (Tercan 1999). Orthogonalization of indicator function relies on principally the decomposition of the indicator variogram matrices as a matrix product. Depending on the type of decomposition, varying degrees of the spatial orthogonality among the factors are produced. Tercan (1999) considers three decomposition algorithms: Spectral (SPEC), Symmetric (SYMM) and Cholesky-Spectral (CHSP) decomposition. The results of his studies indicate that the estimation algorithm based on the

CHSP decomposition performs better than the other decomposition algorithms.

This paper introduces estimating conditional distributions based on orthogonal transformed indicator method and applies it to estimation and simulation of the upper lignite seam thickness of the Kalburçayırı field, Kangal Basin, Sivas, Turkey. The Cholesky-spectral algorithm is used as orthogonalization algorithm due to its superior performance among others. This study is mainly based on the doctoral work of the author (Cengiz 2003).

## Estimation of Conditional Distribution Based on Cholesky-Spectral Decomposition

Indicator variable  $I(\mathbf{x}; z_k)$  is obtained by coding the  $Z(\mathbf{x})$  random variable as 0 and 1.

$$I(\mathbf{x}; z_k) = \begin{cases} 1, & Z(\mathbf{x}) \leq z_k \\ 0, & Z(\mathbf{x}) > z_k \end{cases} \quad (2.1)$$

The sequence of indicator variables which obtained by defining it for more than one cut-off value  $[z_k, k = 1, \dots, K]$  is defined as indicator vector.

$$I(\mathbf{x}; z) = [I(\mathbf{x}; z_1) \dots I(\mathbf{x}; z_K)] \quad (2.2)$$

Conditional distribution function is equal to the expected value ( $E[x]$ ) of indicator variable (Eq. 2.3).

$$E[I(\mathbf{x}; z_k)|Z_n] = 1.F(\mathbf{x}; z_k|Z_n) + 0.[1 - F(\mathbf{x}; z_k|Z_n)] = F(\mathbf{x}; z_k|Z_n) \quad (2.3)$$

So, the conditional distribution functions  $F(\mathbf{x})$  can be obtained by estimation (in equations denoted by “\*”) of the expected values of indicator variables (Eq. 2.4).

$$F(\mathbf{x}; z_k|Z_n) = I(\mathbf{x}; z_k)^* \quad (2.4)$$

The conditional distribution functions are obtained by estimation of indicator vectors. When OTIM is used for estimation of conditional distribution functions, the indicator vector (Eq. 2.2)

is transformed into factors  $\mathbf{Y}(\mathbf{x})$  or random functions, that shows orthogonality at each distance.

$$\mathbf{Y}(\mathbf{x}) = \mathbf{I}(\mathbf{x}; \mathbf{z})\mathbf{P} \quad (2.5)$$

Where;  $\mathbf{P}$ , denotes a  $K \times K$  full ranked matrix that linearly transform the indicator vector into factors. Factors are used as kriging estimators.

$$\mathbf{Y}^*(\mathbf{x}) = \left[ \sum_{a=1}^N \lambda_a(\mathbf{x}_a) Y_1(\mathbf{x}_a), \dots, \sum_{a=1}^N \lambda_K(\mathbf{x}_a) Y_K(\mathbf{x}_a) \right] \quad (2.6)$$

Here,  $\lambda_k(\mathbf{x}_a)$ ,  $k=1, \dots, K$ , denote the kriging weights. Once the vector of the factor estimators,  $\mathbf{Y}^*(\mathbf{x})$  is obtained an inverse transformation will provide an estimate of the conditional distribution function vector:

$$\mathbf{I}^*(\mathbf{x}; \mathbf{z}) = \mathbf{Y}^*(\mathbf{x}) \mathbf{P}^{-1}, \quad (2.7)$$

where the superscript  $-1$  denotes inverse.

In this study the matrix  $\mathbf{P}$ , is calculated by using Cholesky-Spectral (CHSP) decomposition. The CHSP algorithm uses both the Cholesky and spectral decompositions and decomposes the two indicator variogram matrices  $\Gamma_I(\mathbf{h}_1)$  and  $\Gamma_I(\mathbf{h}_2)$  for  $\mathbf{h}_2 > \mathbf{h}_1$ . The indicator variogram function matrix is a  $K \times K$  matrix that contains the indicator direct variograms along its major diagonal and the indicator cross variograms off that diagonal (Eq. 2.8):

$$\Gamma_I(\mathbf{h}) = \begin{bmatrix} \gamma_I[\mathbf{h}; (z_1, z_1)] & \dots & \gamma_I[\mathbf{h}; (z_1, z_K)] \\ \vdots & \ddots & \vdots \\ \gamma_I[\mathbf{h}; (z_K, z_1)] & \dots & \gamma_I[\mathbf{h}; (z_K, z_K)] \end{bmatrix} \quad (2.8)$$

The CHSP decomposes these matrices such that

$$\begin{aligned} \Gamma_I(\mathbf{h}_2) &= \mathbf{X}\mathbf{X}^T \\ \Gamma_I(\mathbf{h}_1) &= \mathbf{D}\mathbf{X}\mathbf{X}^T \quad \mathbf{X} = \mathbf{G}\mathbf{S} \end{aligned} \quad (2.9)$$

where the matrix  $\Gamma_I(\mathbf{h}_2)$  is positive definite,  $\mathbf{G}$  is the lower triangular matrix from the Cholesky decomposition of  $\Gamma_I(\mathbf{h}_1)$  (Eq. 2.10):

$$\Gamma_I(\mathbf{h}_1) = \mathbf{G}\mathbf{G}^T, \quad (2.10)$$

$\mathbf{S}$  and  $\mathbf{D}$  are orthogonal and diagonal matrices from the spectral decomposition of  $\mathbf{C} = \mathbf{G}^{-1}\Gamma_I(\mathbf{h}_2)(\mathbf{G}^{-1})^T$ . In this equation the superscript “T” is denote the transpose of that matrix. By using these equations the factors and factor variograms are calculated as Equation (2.11):

$$\begin{aligned} \mathbf{Y}(\mathbf{x}) &= \mathbf{I}(\mathbf{x}; \mathbf{z})(\mathbf{X}^T)^{-1} \\ \Gamma_Y(\mathbf{h}_2) &= \mathbf{X}^{-1}\Gamma_I(\mathbf{h}_2)(\mathbf{X}^T)^{-1} = \mathbf{I}_K \\ \Gamma_Y(\mathbf{h}_1) &= \mathbf{X}^{-1}\Gamma_I(\mathbf{h}_1)(\mathbf{X}^T)^{-1} = \mathbf{D} \end{aligned} \quad (2.11)$$

where,  $\mathbf{I}_K$  is the  $K \times K$  identity matrix. The orthogonality is guaranteed at two lag distances  $\mathbf{h}_1$  and  $\mathbf{h}_2$ . Note that because  $(\mathbf{X}^T)^{-1}$  is of full rank and  $\Gamma_I(\mathbf{h}_2)$  is positive definite, one can write:

$$\Gamma_I(\mathbf{h}_1)(\mathbf{X}^T)^{-1} = \Gamma_I(\mathbf{h}_2)(\mathbf{X}^T)^{-1} \mathbf{D} \quad (2.12)$$

and in fact, this is the matrix form of the generalized eigenvalue problem with the matrices  $\Gamma_I(\mathbf{h}_2)$  and  $\Gamma_I(\mathbf{h}_1)$ .

## Sequential Simulation

Consider the simulation of variable grade Z at N grid nodes  $\mathbf{x}_a$  conditional to the data set  $[z(\mathbf{x}_a), a = 1, \dots, n]$ . Sequential simulation (Gomez-Hernandez and Srivastava 1990) amounts to modelling the conditional distribution function then sampling it at each of the grid nodes visited along a random sequence.

When a nonparametric approach is considered an indicator-

based method is used. To ensure reproduction of the grade variogram model, each cdf is made conditional not only to the original n data but also to all values simulated at previously visited locations. Multiple realizations are obtained by repeating the entire sequential drawing process. Sequential simulation starts with the transform of an indicator vector into the spatially orthogonal factors (Tercan A.E. and Kaynak T. 2001).

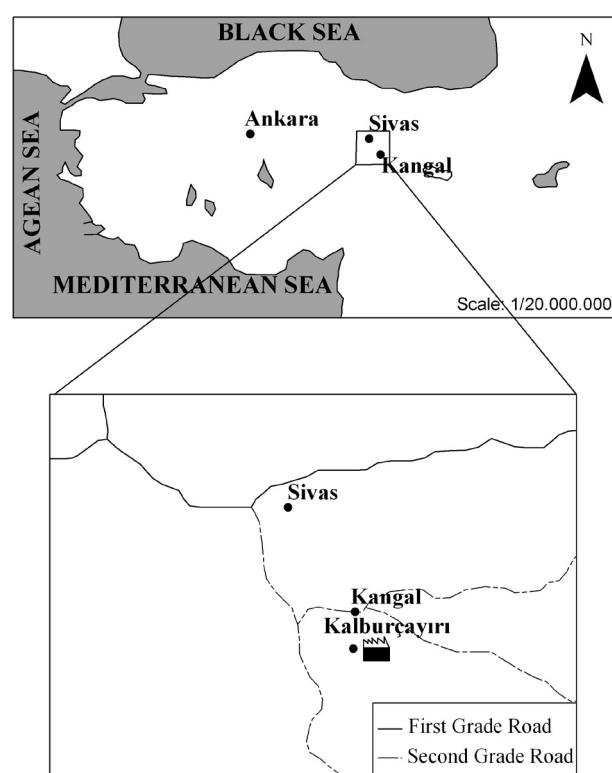
## Case Study

### Definition of Field

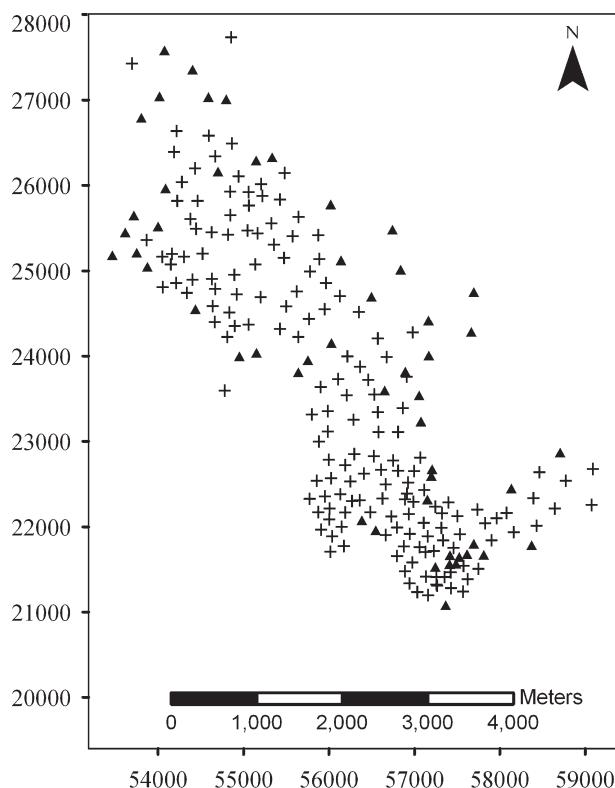
The orthogonal transformed indicator method is used for estimation and simulation of the thickness data of the upper lignite seam of the Kalburçayiri field, Kangal Basin, Sivas, Turkey (Figure 1.). The field includes two coal seams. The run-of-mine coal is directly fed into a power plant. Totally 222 drillings were performed in the field and about 170 of them made intersection with coal. Figure 2 shows the collar positions of drill holes. The stratigraphy of the study area is shown in the Figure 3.

The field was subjected to different geostatistical studies (Tercan 1996a, Tercan 1996b, Tercan 1998a, Tercan 1998b). In these studies, indicator kriging was used as an estimation method. The 170 thickness values were used. The average coal thickness is 7.05 m. The summary statistics and frequency distribution of thickness data are shown on Figure 4 and directional experimental variograms are shown in the Figure 5.

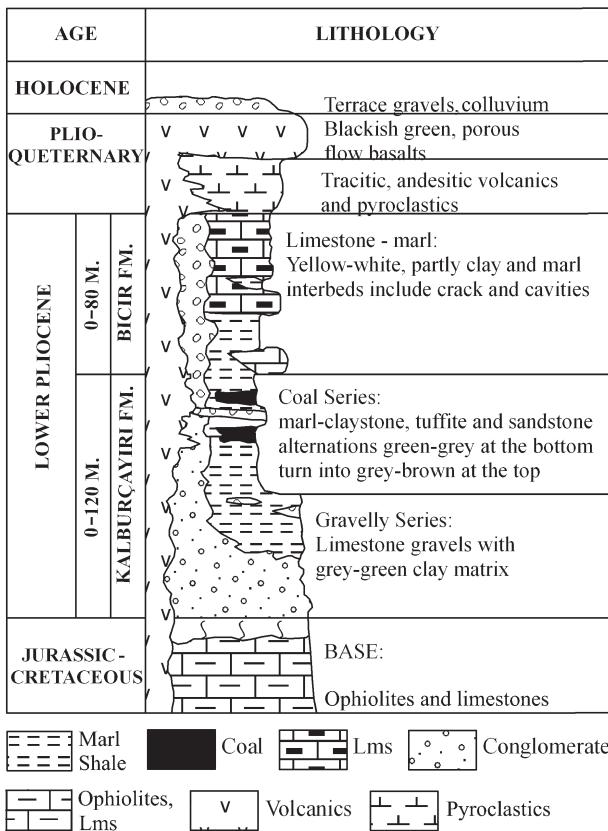
The directional variograms indicate the presence of anisotropy in the directions of N30W and N15E. The variogram



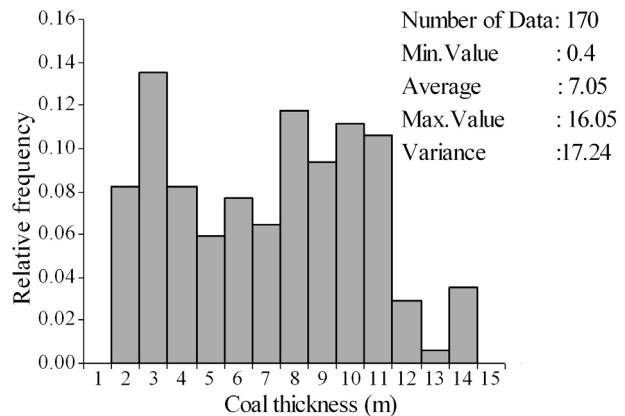
■ Fig. 1. Kalburçayiri field, Kangal Basin, Sivas, Turkey.



■ Fig. 2. The coal positions of the drillholes. +—drillhole made intersection with coal, ▲—drillhole without intersection with coal.



■ Fig. 3. The stratigraphy of the study area (modified from Sen 1999).



■ Fig. 4. The summary statistics and frequency distribution of thickness.

model is spherical<sup>1</sup> with nugget effect 5.0, partial sill 14.0 and range 1150 m in NW direction and 650 m in NE direction.

#### Estimation of conditional cumulative distribution functions

The first step in sequential simulation is estimation of conditional distribution functions. As there are no economic and technical restrictions, the nine cut off values corresponding to the nine deciles of the reference distribution are used: these are: 1.5, 2.4, 3.8, 5.8, 7.4, 8.6, 9.4, 11.0, 12.2. The conditional cumulative distribution functions were computed for nine cut off values using the coal thickness data of Kalburçayiri lignite field. The IK3D from GSLIB (Deutsch and Journel 1998) is modified for computations. The variance and means of distributions are computed using POSTIK from GSLIB (Deutsch and Journel 1998). The image map of real thickness data on Kalburçayiri coal field is shown on Figure 6.

#### The simulation of Kalburçayiri lignite field thickness values using OTIM

The simulations of Kalburçayiri lignite field thickness values are performed using OTIM. The simulations are made conditional to coal thickness value of the 170 drillings. Firstly the transformation matrices and factors are produced. SISIM given in Deutsch & Journel (1998) is modified in order to handle with OTIM. The CHSP technique is used as decomposition method in OTIM. The variogram parameters ( $C_0$ : nugget effect, C: sill and a: range value) of estimated factors are shown in the Table 1. The variograms are computed in the same direction with the real dataset anisotropy and modelled with spherical model.

<sup>1</sup> Spherical model is a model used in geostatistics and have a sill value. In this model:

$$\gamma(\mathbf{h}) = C_0 + Cx \left[ 1.5 \times \frac{h}{a} - 0.5 \times \left( \frac{h}{a} \right)^3 \right] ; h \leq a$$

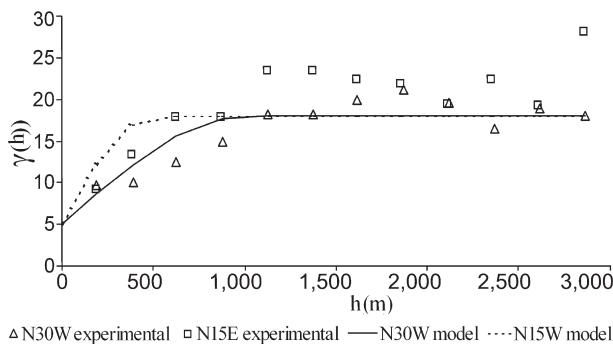
$$\gamma(\mathbf{h}) = C_0 + C ; h > a$$

$$\gamma(\mathbf{h}) = 0 ; h = 0$$

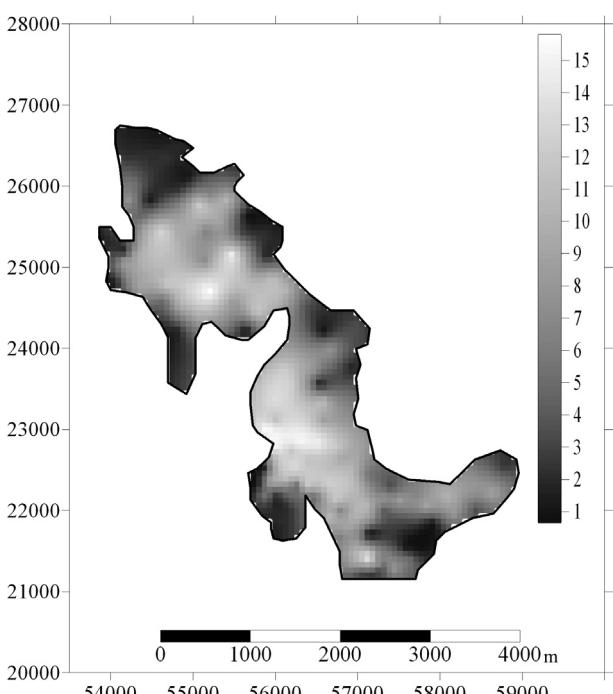
OTIM simulations are realized conditional to real data. The experimental variograms in NW30 and NE15 directions of simulated thickness and real thickness data are shown in Figure 7

Cut-off Values	Nugget Effect $C_0$	Sill Value C	Range (m) $a_{\text{NW}30}$	Range (m) $a_{\text{NE}15}$
1.463	0.200	1.000	2200	700
2.37	0.250	1.000	1750	900
3.804	0.200	0.800	2000	900
5.82	0.500	0.300	1700	500
7.38	0.700	0.500	1200	650
8.644	0.750	0.250	1500	650
9.443	1.000	0.000	1.000	1.000
11.026	1.000	0.000	1.000	1.000
12.241	1.000	0.000	1.000	1.000

■ Tab. 1. Variogram parameters of factors.



■ Fig. 5. The directional experimental variograms of thickness data.



■ Fig. 6. The image map of mean type (E-type) estimation.

and Figure 8 Figure 9 shows the frequency distributions of simulated thickness and real thickness data.

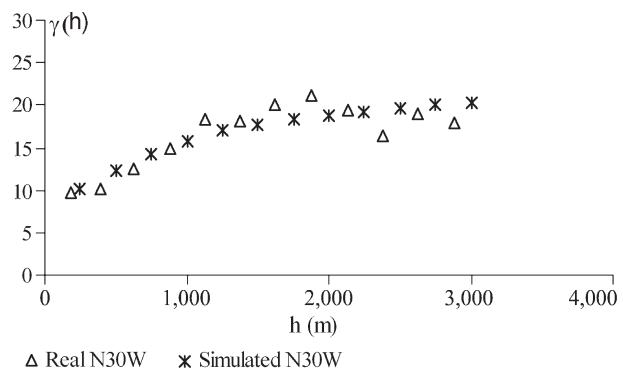
Figure 9 shows that the simulated values have the same frequency distribution as the real data. Despite to reproducing the spatial variability in N30W, there are some discrepancies (fluctuations) between real and simulated thickness variograms in N15E direction.

#### Usage of simulation values

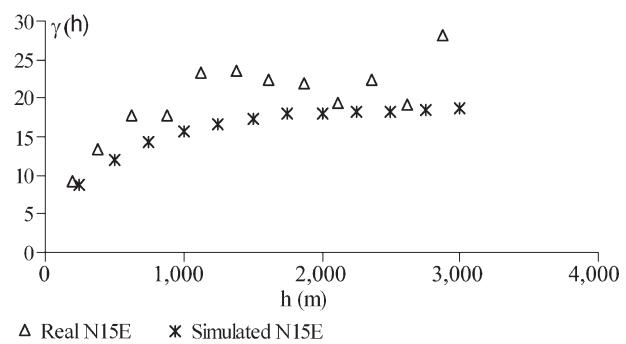
The purpose of the simulation is to make the corresponding data known at every point of the field. In this study, a grid field at 125 m intervals is generated. Figure 7 shows the spatial distribution of 658 conditionally simulated data on 125 m intervals at Kalburçayiri field.

## Conclusions

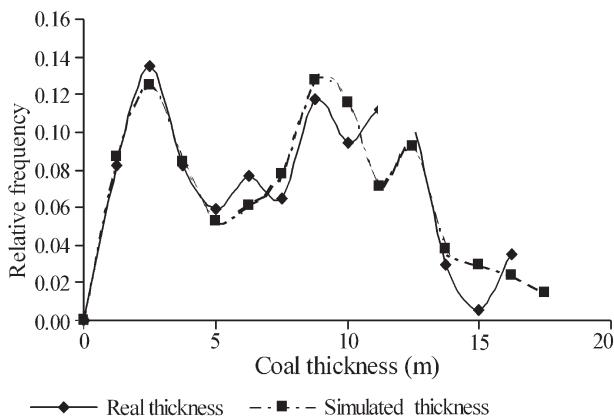
Simulation of a field is described as the generation of the numeric model of that field. Once the simulated values were generated, it can be used in many phases of mining such as mine evaluation, planning, reserve calculation, selection of mine equipment. The Orthogonal Transformed Indicator Method provides more reliable and robust results over other indicator methods in geostatistical simulation.



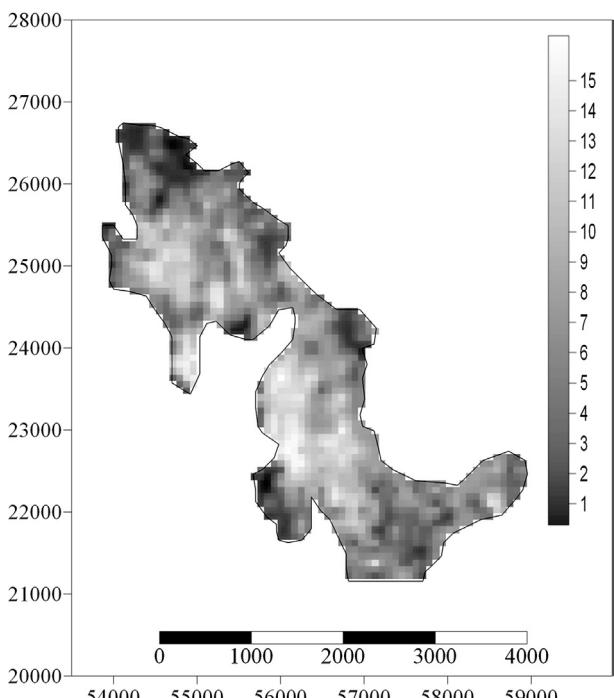
■ Fig. 7. The experimental variograms in NW30 directions of simulated thickness and real thickness data.



■ Fig. 8. The experimental variograms in NE15 directions of simulated thickness and real thickness data.



**Fig. 9.** Frequency distributions of simulated thickness and real thickness data.



**Fig. 10.** Image map of simulated data on Kalburçayiri coal field.

## References

- CENGİZ T., 2003. Reserves estimation and simulation by orthogonal indicator simulation. Ph.D. Thesis, Hacettepe University, Ankara, Turkey (Unpublished, in Turkish, with English Abstr.).
- DEUTSCH C.V. and JOURNAL A.G., 1998. GSLIB: Geostatistical software library and user's guide. Oxford University Press, New York.
- GOMEZ-HERNANDEZ J. and SRIVASTAVA R.M., 1990. ISIM3D: An ANSI-C three dimensional multiple indicator conditional simulation program. Computers & Geosciences, 16, 395-440.
- GOOVAERTS P., 1997. Geostatistics for natural resources evaluation. Oxford University Press, New York.
- SEN O., 1999. Evaluation of Kalburçayiri (Sivas) Coal Deposit reserve using geometric/geostatistical methods. MSc. Thesis, Hacettepe University, Ankara, Turkey (Unpublished, in Turkish, with English Abstr.).
- TERCAN A.E., 1996a. Evaluation of border uncertainties in mine deposits using indicator kriging and an application at Sivas-Kangal-Kalburçayiri Mine Deposit. Madencilik, 35, 4, 3-12 (in Turkish).
- TERCAN A.E., 1996b. Determining optimum drilling location using geostatistics at Sivas, Kangal Coal Deposit. 10<sup>th</sup> Coal Congr. of Turkey, pp. 1-6 (in Turkish).
- TERCAN A.E., 1998a. Assessment of boundary uncertainty in a coal deposit using probability kriging. Technical Note, IMM, Mining Industry, Section A, 107, 51-54.
- TERCAN A.E., 1998b. Estimation of coal quality parameters using disjunctive kriging. 5th Intern. Symp. on Environmental Issues and Waste Management in Energy and Mineral Production. Ankara, Turkey, pp. 353-356.
- TERCAN A.E., 1999. Importance of orthogonalization algorithm in modelling conditional distributions by orthogonal transformed indicator methods. Mathematical Geology, 31, 155-173.
- TERCAN A.E. and KAYNAK T., 1999. Conditional distributions functions and their place in reserve estimation. 16<sup>th</sup> Mining Congr. of Turkey, Ankara, Turkey, pp. 237-244 (in Turkish, with English Abstr.).
- TERCAN A.E. and KAYNAK T., 2001. Anisotropy problem in geostatistical simulation by orthogonal transformed indicator methods. Int. 17<sup>th</sup> Mining Congr. of Turkey, Ankara, Turkey, pp. 603-608.