



DANIEL JANC\*, RAFAŁ WIŚNIEWSKI\*\*, STANISŁAW STRYCZEK\*\*

## Core recovery criterion – a new optimization method of geological information quality in the geological conditions of the Grodziec Syncline Area

### Introduction

Geologic-prospecting drilling is currently oriented to acquire the most abundant geologic, engineering and reservoir data. Despite the development of drilling and drilling geophysics' methods, rock samples, i.e. cores with intact structure and collected directly from the rock mass, are the most reliable source of information.

Although attempts to improve the techniques and technologies were made, the results are not fully satisfactory due to the amount and the quality of the geological information. The coring process was optimized with the following standard drilling criteria: minimum cost, maximum rate of penetration, maximum footage and minimum time of drilling (Wójtowicz 1975). The authors proposed optimization by focusing on maximum geological information quality and increasing the efficiency of drilling works in the Grodziec Syncline, thereby gaining better quality and quantitative information about the drilled layer. The low core recoveries have a direct influence on the amount and the dependability of the geological information possible to be obtained.

---

\* Ph.D. Eng., Exalo Drilling SA, Piła; PGNiG Warszawa, Poland

\*\* Professor, AGH University of Science and Technology, the Faculty of Drilling, Oil and Gas, Krakow, Poland; e-mail: wisniows@agh.edu.pl; strydzek@agh.edu.pl

## 1. Core drilling in the Grodziec Syncline area

The analyzed drilling works were conducted in South West Poland within the Grodziec Syncline, which is a subunit of the North-Sudetic Basin (Pietrzyński et al. 2007). They were targeted at a detailed recognition of copper ore deposits in the C1 category, and increasing the recognized resources beyond the most extreme balance wells (Kwaśny et al. (2007). The scope of the geologic works covered two stages within which 15 (9 + 6) wells were performed from the surface. The field and laboratory tests followed to establish reservoir, hydrogeological, geologic-engineering and gas parameters of the deposit.

The proposed optimization of the geological information quality was based on data from 9 wells performed in 2011–2012. The project was realized with the use of rotary rigs. The coring was performed with standard single- or double-section core barrels (inner tube immobile with respect to the core) with an adjustable safety joint (Janc et al. 2011). When selecting core bits, attention was paid to the lithology of the drilled rocks and the protection of the core. Low invasion PDC tools of a steel body and four or six blades were applied. Owing to the considerable lithological diversity, with the predominance of hard and compact strata, only a standard ring core catcher with surface weld overlay was used. The problems encountered during the realization of works were connected with the low recovery of cores, lost circulations of mud (50 to 956 m<sup>2</sup> per well) and shorter drilling trips due to the jamming of the core in the core barrel.

Prior to optimizing, a representative set of data had to be prepared and verified as on this basis the optimized model could be construed according the assumed criterion. Therefore, the first decision step was working out a method for data selection and verification.

In the analyzed problem, it was crucial to correctly determine and characterize intervals where core recovery problems (zones of low recovery) occur and confirm their homogeneity and properties which influence the coring process.

The zones of lower core recovery were selected in intervals, which were homogeneous in their geomechanical and lithological parameters, as well as other properties influencing the coring process, where the core recovery was under 100%. They were also found in the same sequence of the lithological profile. Four zones of lower core recovery were discovered on the basis of core recovery analyses (daily drilling reports, final well reports) and lithological profiles from 9 wells in the Grodziec Syncline:

- ◆ Zone A) Upper Zechstein – upper silt;
- ◆ Zone B) Upper Zechstein – carbonate strata between shale and silt packages;
- ◆ Zone C) Upper Zechstein – lower silt;
- ◆ Zone D) Middle Zechstein – carbonate beds.

The smallest recovery was observed in zone B, where in 60% of cases the core was below the boundary value of 80%. The total recovery from this zone was 57.4%. On the whole, 25% of drilling trip in all analyzed zones did not meet the contractual requirements.

## 2. Verification of zones of lower core recovery with statistical classification methods

The geologic-technical analysis allowed for zones of similar lithological type, geologic conditions and the same sequence in the stratigraphic profile to be determined. At this stage it was not possible to clearly state whether the analyzed interval in the neighboring well had the same geomechanical properties influencing the core drilling process or not. Therefore, to ensure that the determined zones in particular wells were homogeneous, they were verified with statistical classification methods.

These methods are commonly applied in economics and medicine for predicting whether a given observation belongs to a predefined class. These are supervised learning methods, i.e. first the classifier is taught on a teaching set, and on this basis a given observation is ascribed one of the classes. The analysis of homogeneity of the drilled zones was conducted with the use of 5 methods presented below. They differ in principle, i.e. the way in which the classification rule is built and which are the most common in use (Kornach and Ćwik 2005; Krzyśko et al. 2008; Stanimirov et al. 2012; Józwik 2013).

### 2.1. Quadratic discriminant analysis (QDA) method

This method belongs to Gaussian classifiers. It employs the estimation of a relation between the density and the assumed parametric model for density, i.e. the assumed known form of density, except for the parameters enclosed in it. The classifier takes the following form:

$$d_B(x) = \arg \max_k \delta_k(x) \quad (1)$$

where:

$$\delta_k(x) = -\frac{1}{2} \ln |\Sigma_k| - \frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) + \ln \pi_k \quad (2)$$

- ↪  $\Sigma_k$  – covariance;
- $\mu_k$  – probabilistic measure of vector  $X$ ;
- $\pi_k$  – probability *a priori*.

Function  $d_B(x)$  is a quadratic function of argument  $x$  (classification function of group  $K$ ). In this method, classes are divided by second order hypersurfaces.

## 2.2. Linear discriminant analysis (LDA) method

The LDA method is a simplified version of the QDA method. The classifier in this method has the following form:

$$d_B(x) = \arg \max_k \delta_k(x) \quad (3)$$

where:

$$\delta_k(x) = -\frac{1}{2} \ln |\Sigma_k| - \frac{1}{2} (x - \mu_k)' \Sigma_k^{-1} (x - \mu_k) + \ln \pi_k \quad (4)$$

In this method, the classes are divided with hypersurfaces.

## 2.3. Naive Bayes classifier

This is a static classifier based on the Bayes theorem. Its principle is based on estimating the density of each property in each class followed by the analysis of the product of such density for each class. Variables of the vector of observation are assumed to be independent of one another. The density of probability equals  $p(x|K) = f_K(x)$ , whereas  $\pi_K$  is an absolute *a priori* probability that the observation belongs to class  $K$  ( $K = 1, 2, \dots, g$ ). Function  $f_k$  is a density of normal distribution. Then, through the Bayesian law, a posteriori probability that the observed (vector) value  $x$  belongs to class  $K$  equals:

$$p(K|x) = \frac{\pi_K p(x|K)}{\sum_{r=1}^g \pi_r p(x|r)} \quad (5)$$

The Bayesian rule has its general form:

$$d_B(x) = \arg \max_K \pi_K f_K(x) \quad (6)$$

The evaluation of function  $f_k$  can be significantly simplified on the assumption that all the constituents of the observed random vector  $X$  are independent. Then:

$$f_k(x_1, x_2, \dots, x_p) = \prod_{j=1}^p f_{kj}(x_j) \quad (7)$$

This reduces the problem to the estimation  $p$  of single-dimensional functions of density for each of the classes  $K$ . The observed element is ascribed to class  $K$ , for which the function  $f_x$  has the highest values.

#### 2.4. K-Nearest Neighbor Method (KNN)

This classification method is based on a non-parametric estimation of probability distributions. For an observation (test sample) to be classified,  $K$  nearest neighbors (observations) is determined from the learning set. Then they are included in the class, to which most of the determined neighbors belong. *A posteriori* estimator of probability that observation  $x$  belongs to class  $k$  is calculated as participation of observations from this class among its  $K$  nearest neighbors:

$$\hat{p}(k | x) = \frac{1}{K} \sum_{i=1}^n I(\rho(x, x_i) \leq \rho(x, x^{(K)})) I(y_i = k), \quad (8)$$

$$k = 1, \dots, L$$

↪  $x^{(K)}$  is  $K$ -th point of a teaching sample (as far as distance from  $x$  goes),  $\rho$  is a distance (the so-called measure of dissimilarity of objects). The classifier in the KNN method has the following form:

$$\hat{d}_{KNN}(x) = \arg \max_k \hat{p}(k | x) \quad (9)$$

#### 2.5. Decision trees method

This family of statistical methods for data mining makes use of diagrams called 'trees'. The construction of binary trees lies in a sequential complementary division of subsets of space of sample  $\chi$  into two subsets, starting from the entire set  $\chi$ . In each step, the division is aimed at obtaining the most possible homogeneous parts. The subsets which do not undergo further division, are final subsets which are ascribed a class label.

A 'tree' consists of a 'root' and 'branches' and the successive 'nodes' ( $t$ ). A certain condition imposed on a given observation is checked in each 'node', and on this basis one of the 'branches' is selected from where we proceed to a specific 'node' below. 'Leaves', where we can read the class, to which a given observation can be ascribed are located underneath. The classifier construed on the basis of a classification tree  $T$  has the following form:

$$\hat{d}_T(x) = \sum_{t \in \tilde{T}} ind(t) | (x \in t) \quad (10)$$

The measure of the classifier's quality is the coefficient of accuracy describing the percentage of test samples correctly qualified by the model. The fastness is also another measure of the classifier's usability. The classifiers were validated with two methods: cross validation and one leave out methods (Janc 2014). The homogeneity of zones was verified on the basis of the relation: momentary rate of penetration – weight on bit – rotation, characteristic of a given zone (class). The classification was performed on 5 series of data constituting individual technological cases defined on the basis of the type of the applied tool and wear code according to IADC (Hughes Christensen 1996). For the sake of eliminating distortions resulting from the order of magnitude of the involved data, rotations were recalculated for [1/s], weight on bit on [kG · 10<sup>3</sup>] and the penetration rate to [m/h]. The following technological situations were analyzed:

- ◆ S10 – tool 1, wear 0;
- ◆ S11 – tool 1, wear 1;
- ◆ S12 – tool 1, wear 2;
- ◆ S13 – tool 1, wear 3;
- ◆ A21 – tool 2, wear 1;

The classification was performed in a development environment R 3.0.2 (Rybiński 2012) with the use of packages *MASS*, *klaR*, *party*, *grid*, *zoo*, *sandwich*, *strucchange*, *stats4* and *modeltools*. The list of fit coefficients of accuracy for a classifier is presented in Table 1.

Relatively low values of fit coefficients in the LDA method for S12 (61.7%) and S11 (76.86%) stem from the way in which the space has been divided into classes. Particular classes distinguish hypersurfaces, whereas the spatial distribution of measurement points indicates results obtained by the QDA method (S12 – 88.29% and S11 – 85.11%), where the space is divided with second order hypersurfaces. The curvature of surfaces separating classes more accurately determines the boundaries and encompassed most of the data. The best fit for the classifier is obtained with the KNN method (90.86–100%), which analyzes neighboring measuring points. Equally good results are obtained with the Naive Bayes classifier (82.98–94.50%) based on the estimation of density of properties.

The level of correctly classified elements is very high and in most cases exceeds 80%. The analyzed data comes from industry and is burdened with a number of disturbances and errors; the coring process itself is complex and depends on a number of factors. For this reason this value can be assumed as satisfactory. The main factors influencing the height of the acceptability threshold are: jamming of the core (change of the character of drilling); accuracy of weight indicator, calibration, technical condition; kinematic characteristic of the rig, fluency with which parameters change; human factor, experience of the crew, way in which the rig is handled, data registered by the mud logging service accuracy of sensors used for control and measurement purposes.

Table 1. Fit coefficients of accuracy for a classifier.

Tabela 1. Zestawienie współczynników dopasowania klasyfikatora

Classification method	S10	S11	S12	S13	A21
<b>LDA</b>	<b>89.29</b>	<b>76.86</b>	<b>61.70</b>	<b>92.66</b>	<b>81.43</b>
Qualitative evaluation of LDA – cross validation method	86.89	76.66	60.03	92.19	79.25
Qualitative evaluation of LDA – one leave out method	89.29	76.85	60.70	92.66	81.43
<b>QDA</b>	<b>89.29</b>	<b>88.29</b>	<b>85.11</b>	<b>93.58</b>	<b>89.64</b>
Qualitative evaluation of QDA – cross validation method	81.53	87.42	80.32	91.08	86.85
Qualitative evaluation of QDA – one leave out method	89.29	88.29	85.11	93.58	89.64
<b>Naive Bayes classifier (NB)</b>	<b>91.07</b>	<b>88.29</b>	<b>82.98</b>	<b>94.50</b>	<b>86.07</b>
Qualitative evaluation of NB – cross validation method	91.07	88.03	80.95	92.17	87.01
Qualitative evaluation of NB – one leave out method	87.25	86.62	77.81	91.67	83.67
<b>K-Nearest Neighbor (KNN)</b>	<b>100.00</b>	<b>90.86</b>	<b>90.96</b>	<b>94.50</b>	<b>90.36</b>
Qualitative evaluation of KNN – cross validation method	87.82	75.21	69.84	91.61	75.83
Qualitative evaluation of KNN – one leave out method	92.86	75.29	70.08	97.22	69.89
<b>Classification tree</b>	<b>91.07</b>	<b>91.71</b>	<b>73.94</b>	<b>94.50</b>	<b>90.00</b>
Qualitative evaluation of trees – cross validation method	87.50	73.50	60.32	91.67	86.02
Qualitative evaluation of trees – one leave out method	86.68	81.63	67.00	93.44	82.75

As the percentage of elements correctly qualified by all tested classifiers is considerably high, one may venture a conclusion that these zones are homogeneous. The values of the fit coefficient of accuracy and qualitative evaluation confirm the individual character of the determined zones and their homogeneity of geomechanical properties having influence on the coring process. From among the analyzed classification methods, the best results were obtained for the QDA and Naive Bayer classifier methods. This choice was based on the qualitative evaluation of the classifier.

### 3. Criterion of maximum geological information quality

Obtaining maximum geologic information has priority when realizing works in the Grodziec Syncline. The new criterion of maximum geological information quality worked out at the Faculty of Drilling, Oil and Gas at AGH-UST meets both the technical and economic criteria (Janc 2014). This results from the way in which core recoveries above 80% are additionally appreciated and granted. Technically, this is a quality criterion (Pająk and Wieczorowski 1982). It tends to maximize the geologic information, which directly depends

on the amount of core material to be analyzed. The more numerous the samples and the larger their scope, the higher the quality of analyses and better recognition (and documentation) of the deposit. As far as the contractor is concerned, this criterion is directly connected with the economic result of the undertaking. According to the contract, these are ‘turnkey’ wells, where works are paid on the basis of the footage. In the case of core recovery under the boundary value of 80% the contractor is supposed to cover the entire or part of the cost of such works. For this reason, the maximization of core recovery is favorable both for the operator who obtains fuller geologic information, and the operator, who achieves better economic effect of the conducted works.

Optimized core recovery for given geologic conditions is a function of many parameters:

$$U = f(T, Z, L, P, n, Q, C, M, \rho, \dots, \omega) \quad [\%] \quad (11)$$

- $T$  – type of drilling core bit, [-];
- $Z$  – wear of the tool, [IADC];
- $L$  – length of the trip, [m];
- $P$  – weight on core bit, [kG];
- $n$  – rotations, [1/min];
- $Q$  – rate of pumped volume of mud, [l/min];
- $C$  – pressure of pumped mud, [atm];
- $M$  – torque, [kGm];
- $\rho$  – mud density, [g/cm<sup>3</sup>];
- $\omega$  – unmeasurable factors and disturbances.

For the sake of geologic conditions, the function was determined individually for each zone.

#### 4. Function of core recovery

The identification, aimed at determining a mathematical model which would describe a given process, is based on determining parameters, which have a significant influence on their interactions. This was done with the use of the correlation analysis. Unfortunately, the analyses did not reveal any direct relation between the particular parameters of the processes and the core recovery in all technological cases. Therefore, the idea of controlling was based on the detection of basic factors influencing core recovery, i.e.: mechanical wear of the core due to jamming, erosional activity of the drilling mud and washing of the core, unstable operation of tools downhole (mechanical wear of the core due to vibrations).

Prior to building the model, general assumptions regarding the form, functionality, limitations and way of updating were made, i.e.:



1. The model is based on the detection of phenomena responsible for core wearing in the drilling process.
2. Managing the process through the control of influence of major parameters of drilling on indices which directly depend on it.
3. Functionality and simplicity of design – easy applicability and upgrading in industrial conditions while coring.
4. Universal nature.
5. Model under construction – with the model will be developed the growing number of data.
6. Limitation of the model results from technological and technical causes.

Taking the major factor influencing the core recovery into account, the influence of major parameters of drilling on the directly related indices was analyzed in 7 models (Table 2).

Table 2. Core recovery models

Tabela 2. Modele funkcji uzysku rdzenia

No	Model	Function	Description
1.	Model A	$U = f\left(\frac{V}{P}\right)$	Momentary rate of penetration to weight on bit ratio as an index of core jamming and its wear
2.	Model B	$U = f\left(\frac{V}{P \cdot n}\right)$	Ratio of momentary rate of penetration and product of mechanical parameters of drilling as an index of core jamming and its wear
3.	Model C	$U = f\left(\frac{M}{P}\right)$	Torque to weight on bit ratio as an index of core jamming and its wear
4.	Model D	$U = f\left(\frac{M \cdot P}{V}\right)$	Ratio of momentary rate of penetration and product of torque and weight on bit as an index of core jamming
5.	Model E	$U = f(P \cdot n)$	Influence of mechanical parameters of drilling on recovery
6.	Model F	$U = f\left(\frac{Q}{C}\right)$	Ratio of pumped volume of mud and pumping pressure of mud pump as an index of erosional wear of the core
7.	Model G	$U = f\left(\frac{Q}{C} \cdot \frac{V}{P}\right)$	Product of models A and F as mechanical and erosional indices of core wearing

The form of the model was determined with the use of linear and nonlinear analysis methods. The selection criterion deciding about the form of the core recovery function accounts for regression coefficient  $R^2$ , and also a more universal aspect, i.e. high adaptability of the model to actual data in all the technological cases. Among the analyzed models,

the best fit was noted for the logarithmic model A. The function of core recovery takes the following general form:

$$y = a \ln(x) + b \quad (12)$$

where the value of function  $y$  (dependent variable – recovery  $U$ ) is limited by an interval  $\langle 0;100 \rangle$ , whereas the interval of arguments  $x$  ( $V/P$  ratio) depends on technical and technological limitations of the process. The core recovery function in particular technological cases is presented in Table 3.

Table 3. Equations of core recovery models (model A)

Tabela 3. Współczynnik modeli uzysku rdzenia (model A)

PT	Regression equation	R <sup>2</sup>
A11	$y = 42.089 \ln(x) + 52.044$	0.7217
A21	$y = 108.40 \ln(x) + 78.758$	0.7381
B11	$y = 83.565 \ln(x) + 50.577$	0.8318
B12	$y = 80.572 \ln(x) + 89.978$	0.8671
B21	$y = 48.858 \ln(x) + 81.297$	0.7623
C11	$y = 24.537 \ln(x) + 88.859$	0.5701
D11	$y = -23.30 \ln(x) + 100.1$	0.5966
D13	$y = -27.18 \ln(x) + 66.658$	0.7569
	<b>Mean R<sup>2</sup></b>	<b>0.7306</b>

The obtained values of the regression coefficient are on an acceptable level and prove that the model has a sufficient fit to the real data. Attention should be paid to the fact that the stochastic model was created on the basis of a binary experiment. The analyzed data is averaged industrial data, usually at a narrow range of variability. The occurring noises and disturbances result from the accuracy of indications and the way in which the data was recorded by measuring devices and the subjective evaluation of wear of the tool by the crew. The assumption that the rock medium is homogeneous, although in reality its homogeneity is not on an acceptable level is equally important. The technical and trip characteristic of the rig sub-assemblies imposes certain limitations regarding the range of controllable parameters, fluency of regulation and stability in time. The human factor, experience of the crew and organization of work should also be taken into consideration.

The optimization of the coring process depends on finding an extreme of the function in view of certain limitations. This also requires defining some detailed assumptions for the model functions. In the analyzed case, the following assumptions were made:

1. Macroscopically homogeneous rocks.
2. Constant parameters of drilling.
3. Weight on core bit is evenly distributed on all non-deformable blades.
4. The tool does not wear out during the trip (short sections, strong tool).
5. Technical and technological limitations.

The function of core recovery  $U$  has one local extreme at point  $[100; W_{100}]$ . The index of jamming is defined as:

$$W = \frac{V}{P} \left[ \frac{m}{\text{kG} \cdot 10^3 \cdot h} \right] \quad (13)$$

By inserting equation (13) to (12) we obtain:

$$U = a \ln(W) + b \quad (14)$$

After a simple transformation of the dependence above (14), the value of the jamming index equals:

$$W = e^{\frac{U-b}{a}} \quad (15)$$

Then the optimum value of argument  $W$  for 100% recovery is:

$$W_{opt} = W_{100} = e^{\frac{100-b}{a}} \quad (16)$$

- ↪  $U$  – core recovery, [%];
- $a, b$  – coefficients of the model;
- $W$  – jamming coefficient, [ $\text{m}/\text{kG} \cdot 10^3 \cdot \text{h}$ ];

The jamming coefficient depends on weight on core bit ( $P$ ) exerted by the tool and on the difference of the entry friction of the core to the core barrel, on the assumption that the drilled formation was homogeneous. The jamming of the core (entry friction) in the bit and core barrel is not directly measurable in the course of drilling. The idea of controlling and optimizing the process of coring according to the criterion of maximum core recovery lies in controlling the core wear by investigating the influence of weight on bit on the momentary rate of penetration. If no jamming is observed, the entire weight is transferred on the core bit and the acquired momentary rate of penetration is on a specific level, characteristic of a given technological case. After jamming, part of the weight is transferred onto the rock material which is destroyed and washed out, and the momentary rate of penetration decreases despite the fact that a constant weight is registered on the surface. In such a situation,

apart from the criterion of maximum geological information quality, it is also the criterion of the maximum momentary rate of penetration which is indirectly accounted for, and which has a significant influence on the financial effect of the conducted works. When applying optimum weight on core bit  $P_{dop}$ , the theoretical momentary rate of penetration for 100% core recovery in the model is:

$$V_{opt100\%} = W_{100} \cdot P_{opt} \text{ [m/h]} \quad (17)$$

$$P_{opt} = P_{dop} \text{ [kG]} \quad (18)$$

A list of theoretical values of optimum momentary rate of penetration at admissible weight on bit is presented in Table 4.

Table 4. List of optimum parameters of coring according to the maximum core recovery criterion - model A

Tabela 4. Zestawienie wartości optymalnych parametrów procesu rdzeniowania według kryterium maksymalnego uzysku rdzenia – model A

TS	Function of model	$W_{opt}$ [m/(kG·10 <sup>3</sup> ·h)]	$P_{opt}$ [kG·10 <sup>3</sup> ]	$W_{opt}$ [m/h]	$W_{80}$ [m/(kG·10 <sup>3</sup> ·h)]	$W_{100}$ [m/(kG·10 <sup>3</sup> ·h)]
A11	$U = 42.089 \ln(W) + 52.044$	3.12	4	12.48	1.94	3.12
A21	$U = 108.40 \ln(W) + 78.758$	1.22	4	4.88	1.01	1.22
B11	$U = 83.565 \ln(W) + 50.577$	1.81	4	7.24	1.42	1.81
B12	$U = 80.572 \ln(W) + 89.978$	1.13	4	4.52	0.88	1.13
B21	$U = 48.858 \ln(W) + 81.297$	1.47	4	5.88	0.97	1.47
C11	$U = 24.537 \ln(W) + 88.859$	1.57	4	6.28	0.7	1.57
D11	$U = -23.30 \ln(W) + 100.1$	1.00	4	4.00	2.37	1.00
D13	$U = -27.18 \ln(W) + 66.658$	0.29	4	1.16	0.61	0.29

Theoretical values of optimum parameters of coring are mean values. While drilling in industrial conditions their momentary values will vary in a certain interval of values. This results from the technical characteristic of the device, control and assumed homogeneity of the rock mass. Taking the contract conditions into consideration, the optimum value of the core recovery equals to 80–100%. Then the optimum value of the jamming coefficient belongs to an interval  $\langle W_{80}; W_{100} \rangle$ . When the jamming factor is below the boundary values  $W_{80}$ , the core recovery will be lower than agreed upon in the contract.

## 5. Statistical validation of the model

The core recovery models were validated with the use of statistical tests of the mean jamming index employing Student's  $t$ -distribution. It depends on analyzing the mean value  $W_{80-100\%}$  calculated on the basis of a depth database and comparing it with the theoretical index of jamming, determined from the recovery mean values in the interval  $\langle 80; 100 \rangle$  determined with a mathematical model.

The analysis of the statistical tests of the mean jamming index was made for two sets, i.e.  $W_{80-100\%} \in \langle 80; 100 \rangle$  and  $W_{0-80\%} \in \langle 0; 80 \rangle$ . Sets defined in such a manner: zero hypothesis (they are equal) and an alternative hypothesis (they are different) were checked. The alternative hypothesis was assumed in all cases after the test was completed. The list of optimum intervals determined on the basis of the significance test for two expected values is presented in Table 5.

The test revealed that both the mean values typical of a given set and their intervals on the significance level of 0.05 differ. The comparison of previously determined optimum intervals of index  $W$  from  $\langle 80; 100 \rangle$  (Table 4) and mean value intervals from statistical tests (Table 5) confirms the correctness and usability of the optimization method. In all technological cases mean values are in the optimum interval or beyond its right side limit determined on the basis of the core recovery mathematical model. Moreover, the intervals overlap (Fig. 1) in both methods. This confirms the correctness of the optimization.

Table 5. Mean values of jamming index - model A

Tabela 5. Wartości średnie wskaźnika klinowania – model A

TS	Population	Mean >79.5%	Interval		Population	Mean <79.5%	Interval		var. test	t. test	Scope of recovery changes
A11	164	3.18	2.83	3.52	67	1.34	1.06	1.63	1.59e-08	2.48e-14	1-100%
A21	34	1.20	0.93	1.48	34	0.47	0.58	0.69	1.96e-06	5.61e-05	30-100%
B11	38	1.51	1.31	1.71	35	0.92	0.78	1.06	0.01587	3.71E-06	0-100%
B12	36	0.97	0.84	1.11	48	0.50	0.43	0.57	0.001574	2.36E-08	20-100%
B21	36	1.40	1.20	1.60	43	0.81	0.68	0.94	0.03257	1.97E-06	6-100%
C11	125	1.66	1.38	1.94	19	0.78	0.55	1.01	1.11E-06	1.59E-06	40-100%
D11	153	1.57	1.35	1.80	23	17.88	2.20	33.6	< 2.2e-16	0.02111	10-100%
D13	44	0.44	0.38	0.50	44	3.13	1.61	4.66	< 2.2e16	0.000461	43-100%

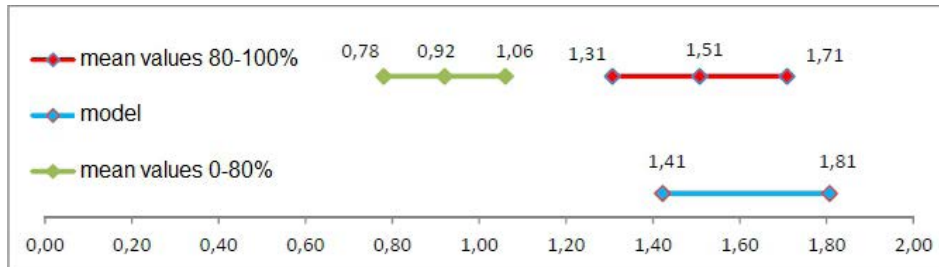


Fig. 1. Graphical comparison of optimization methods (B11)

Rys. 1. Graficzne zestawienie metod optymalizacji (B11)

## 6. Field test

An industrial test was performed to confirm the efficiency of the elaborated optimization of the coring method for the maximum core recovery criterion. A total of 9 drilling trips were analyzed in all lower core recovery zones. The applied tool was of type 1 and wear 1 according to the IADC code. The depth ranges, mean parameters of the process (time base) and obtained results are presented in Table 6.

Table 6. List of drilling parameters and core recovery in particular zones

Tabela 6. Zestawienie parametrów wiercenia i uzysków rdzenia w poszczególnych strefach

TS	No. of trip	Depth [m b.s.]		Length of trip [m]	Recovery [%]	V [m/h]	P [kG·10 <sup>3</sup> ]	W [m/(kG·10 <sup>3</sup> ·h)]	Recovery in model	
		beginning	end						model [%]	difference [%]
A11	1	1 089.0	1 091.3	2.3	83	4.92	2.48	1.98	80.9	2.1
B11	2	1 091.3	1 095.4	4.1	0	1.86	2.82	0.66	15.8	-15.8
B11	3	1 104.0	1 106.0	2.0	85	3.47	2.42	1.43	80.7	4.3
B11	4	1 106.0	1 108.0	2.0	100	3.31	2.05	1.61	90.6	9.4
C11	5	1 115.4	1 120.0	4.6	100	4.96	1.99	2.49	100	0
C11	6	1 120.8	1 123.0	2.2	100	5.26	2.46	2.14	100	0
C11	7	1 133.0	1 138.5	5.5	100	3.57	2.5	1.43	97.6	2.4
D11	8	1 175.4	1 184.0	8.6	100	3.33	3.59	0.93	100	0
D11	9	1 184.0	1 192.6	8.6	100	2.91	3.31	0.88	100	0

Statistical classification methods, i.e. the QDA method and the Naive Bayes classifier were used for identifying and evaluating homogeneity of the drilled rocks. The teaching set consisted of data from 9 previous wells (S11), whereas the test set consisted of data from the newly drilled well. The test based on ascribing measuring points classes (zones), and then determining the percentage of correctly classified samples. The results of the classification are presented in Table 7.

Table 7. List of classification results

Tabela 7. Zestawienie wyników klasyfikacji

QDA method							Naive Bayes classifier method						
Zone	Trip	A	B	C	D	[%]	Zone	Trip	A	B	C	D	[%]
A11	1	27	1	7	0	77.1	A11	1	30	2	0	0	93.8
B11	2	21	43	11	40	37.4	B11	2	26	21	52	52	13.9
B11	3	5	18	1	0	75.0	B11	3	3	17	4	0	70.8
B11	4	9	34	2	0	75.6	B11	4	3	40	2	0	88.9
C11	5	3	0	43	0	93.5	C11	5	3	1	42	0	91.3
C11	6	2	3	21	1	80.8	C11	6	2	2	19	3	73.1
C11	7	5	10	79	2	82.3	C11	7	10	0	85	1	88.5
D11	8	0	2	3	137	96.5	D11	8	0	7	20	115	81.0
D11	9	1	16	1	160	89.9	D11	9	0	43	1	134	75.3

Table 8. Core recovery models – updated after tests

Tabela 8. Uaktualnione na podstawie testu modele funkcji uzysku rdzenia

TS	Regression equation	R <sup>2</sup>
A11	$y = 42.089 \ln(x) + 52.044$	0.7217
<b>A11*</b>	<b><math>y = 42.082 \ln(x) + 52.173</math></b>	<b>0.7215</b>
B11	$y = 83.565 \ln(x) + 50.577$	0.8318
<b>B11*</b>	<b><math>y = 87.722 \ln(x) + 49.999</math></b>	<b>0.8697</b>
C11	$y = 24.537 \ln(x) + 88.859$	0.5701
<b>C11*</b>	<b><math>y = 22.115 \ln(x) + 88.073</math></b>	<b>0.5690</b>
D11	$y = -23.30 \ln(x) + 100.1$	0.5966
<b>D11*</b>	<b><math>y = -23.07 \ln(x) + 99.694</math></b>	<b>0.6063</b>

\* Accounting for the field test trips.

The next step is based on updating the core recovery functions. After introducing values from the tests trips, the determined core recovery models did not significantly vary from the original ones (Table 8).

The obtained core recoveries are equal or close to the theoretical value determined on the basis of the elaborated core recovery function. The amount of correctly classified elements on the level of 75–86.5% (QDA) and 70.8–93.8% (Naive Bayes classifier) proves that the drilled layers was homogeneous as far as factors affecting coring operations is concerned. The only exception is trip no. 2, during which this method was not applied. The obtained test results prove that this technology can be applied in industrial conditions.

## Conclusions

Lower core recovery in the analyzed area is mostly a result of the mechanical wearing of the core due to its jamming in the core barrel. The analysis of the influence of major parameters of the process did not reveal their direct influence on the core recovery in all the analyzed technological cases as jamming has a random and hardly predictable nature. Accordingly, the control of the process is based on detecting places of jamming by controlling the values of newly defined jamming coefficient, i.e. momentary rate of penetration to weight on core bit ratio.

The homogeneity of zones underwent statistical classification as far as their effect on the coring process is concerned. From among the 5 analyzed methods differing in their classification rule, the best results were obtained for QDA and the Naive Bayes classifier.

The optimization of the coring process for maximum geological information quality, based on a mathematical model of core recovery, depended on controlling the jamming coefficient. For practical and contractual reasons its value should be maintained at a level of 80–100% of core recovery. The jamming coefficient value is determined individually for each technological case. This results from the fact that both geological properties, type and technical state of the tool affects the rate of penetration.

The validation of the mathematical core recovery model with statistical tests for the mean value  $W_{80-100\%}$  proved the correctness of the proposed method. Mean values and intervals for index  $W$  obtained during the statistical tests correspond with the optimum intervals determined on the basis of the core recovery function. The test performed while drilling in the Grodziec Syncline confirmed the efficiency and applicability of the presented technology in industrial conditions.

Optimization of the drilling process in the Grodziec Syncline is an open issue. The mathematical models will be improved with the growing number of data. The optimization is fully justified in view of the planned works in that area.



## REFERENCES

- Christensen, H. 1996. *IADC dull grading system for rollers bits*, USA.
- Janc, D. 2014. *Optimization of the coring process in terms of the recovery factor in geological conditions of the KGHM Polska Miedź S.A. copper exploration concession area*. Ph.D. thesis, University of Science and Technology Cracow.
- Janc et al. 2011 – Janc, D., Żuromski, Z. and Zwierzchowski, W. 2011. Komplikacje wiertnicze podczas rdzeniowania oraz metody wpływające na zwiększenie uzysku rdzenia na przykładzie wierceń na obszarze koncesyjnym KGHM Polska Miedź S.A. *Wiadomości naftowe* 4(156) (in Polish).
- Jóźwik, A. 2013. *Nieparametryczne metody klasyfikacji obiektów*. Instytut Biocybernetyki i Inżynierii Biomedycznej PAN (in Polish).
- Kornach, J. and Ćwik, J. 2005. *Statystyczne systemy uczące się*. Warszawa: Wyd. Naukowo-Techniczne (in Polish).
- Krzyško et al. 2008 – Krzyško, M., Wołyński, W., Górecki, T. and Skorzybut, M. 2008. *Systemy uczące się*. Warszawa: Wyd. Naukowo-Techniczne (in Polish).
- Kwaśny et al. 2007 – Kwaśny, L., Bachowski, C. and Markiewicz, A. 2007. *Projekt prac geologicznych na poszukiwanie i rozpoznawanie złoże rud miedzi w obszarze synkliny grodzieckiej*. KGHM „CUPRUM” sp. z o.o., Wrocław (in Polish).
- Pająk, E. and Wiczerowski, K. 1982. *Podstawy optymalizacji operacji technologicznych w przykładach*. Warszawa–Poznań: PWN (in Polish).
- Piestrzyński et al. 2007 – Piestrzyński, A., Banaszak, A. and Zalewska-Kuczmierczyk, M. 2007. *KGHM Polska Miedź S.A. KGHM CUPRUM Sp. z o.o. na zlecenie KGHM Polska Miedź S.A., Lubin* (in Polish).
- Rybiński, M. 2012. *Krótkie wprowadzenie do R dla programistów z elementami statystyki opisowej*. WMiM, Uniwersytet Warszawski (in Polish).
- Stanimirova et al. 2012 – Stanimirova, I., Daszykowski, M. and Walczak, B. 2012. *Metody uczenia z nadzorem – kalibracja, dyskryminacja, klasyfikacja*. Zakład Chemometrii, Instytut Chemii, Uniwersytet Śląski, Katowice (in Polish).
- Wójtowicz, A. 1975. Problemy optymalizacji procesu wiercenia. *Skryty uczelniane AGH* nr 460. Kraków: Wyd. AGH (in Polish).

### KRYTERIUM UZYSKU RDZENIA – NOWA METODA OPTIMALIZACJI JAKOŚCI INFORMACJI GEOLOGICZNEJ W WARUNKACH GEOLOGICZNYCH OBSZARU SYNKLINY GRODZIECKIEJ

#### Słowa kluczowe

wiercenie, rdzeniowanie, uzysk rdzenia, optymalizacja, klasyfikacja

#### Streszczenie

Coraz większe wymogi co do jakości i ilości informacji geologicznej wymuszają ciągle udoskonalanie technik i technologii ich pozyskiwania. W artykule tym autorzy przedstawili metodę optymalizacji procesu rdzeniowania dla nowo zdefiniowanego kryterium maksymalnego uzysku rdzenia, które jednocześnie jest kryterium maksymalnej jakości informacji geologicznej. Pozwala ona na poprawę efektywności prowadzonych prac bez konieczności inwestycji w nowy, specjalistyczny i drogi sprzęt. Do badania jednorodności przewierczanych stref zaadaptowane zostały powszechnie stosowane w medycynie i ekonomii statystyczne metody klasyfikacji. Spośród przeanalizowanych metod różniących się sposobem budowy reguły klasyfikacyjnej największą efektywnością wykazała

się metoda QDA (quadratic discriminant analysis) oraz naiwny klasyfikator bayesowski (Naive Bayes classifier). Zdefiniowane nowe kryterium maksymalnego uzysku rdzenia oraz opracowany matematyczny model uzysku rdzenia pozwolił na zwiększenie ilości, jak i jakości uzyskanego materiału do analiz, w warunkach geologicznych synkliny grodzieckiej. Sterowanie procesem polegało na detekcji zjawisk niszczenia rdzenia, tj. erozyjnego oddziaływania płuczki wiertniczej na rdzeń, mechanicznego niszczenia rdzenia na skutek klinowania się go w aparacie rdzeniowym, a także niestabilnej pracy narzędzia na dnie otworu. Analiza wykazała, że głównym czynnikiem wpływającym na zmniejszone uzyski rdzenia jest mechaniczne niszczenie na skutek klinowania. Do kontroli wartości uzysku rdzenia autorzy zdefiniowali i wykorzystali wskaźnik klinowania przedstawiający wpływ nacisku osiowego na chwilową prędkość wiercenia. Skuteczność opracowanej metodyki potwierdził test przeprowadzony w warunkach przemysłowych, podczas wiercenia kolejnego otworu na obszarze synkliny grodzieckiej.

Artykuł ten powstał na podstawie własnych doświadczeń i analiz przeprowadzonych przez autorów podczas realizacji prac na obszarze synkliny grodzieckiej w południowo-zachodniej części Polski.

#### CORE RECOVERY CRITERION – NEW OPTIMIZATION METHOD OF GEOLOGICAL INFORMATION QUALITY IN GEOLOGICAL CONDITIONS OF THE GRODZIEC SYNCLINE AREA

##### Keywords

drilling, coring, core recovery, optimization, classification

##### Abstract

The increasing requirements for the quality and amount of geological information induce the development of new technological solutions. The authors present the optimization of the coring method for a newly defined criterion of maximum geological information quality. It enables the efficiency of works to be improved without investing in new specialist and expensive equipment. The statistical classification methods commonly used in medicine and the economy were used for analyzing the homogeneity of the drilled rocks. The quadratic discriminant analysis (QDA) and the Naive Bayes classifier method turned out to be most efficient, therefore they were selected from among other methods differing in the way the classification rules are built. The newly defined criterion of maximum geological information quality and the mathematical model of maximum core recovery allowed the amount and the quality of recovered rock material in the geologic conditions of the Grodziec Syncline to be increased. The control of the process relied on detecting phenomena responsible for core damaging, i.e. erosiveness of drilling mud, mechanical damaging due to jamming in the core barrel and unstable operation of the core bit downhole. The analysis revealed that the main factor influencing the lower recovery of cores is mechanical damaging due to jamming in the core barrel. For the sake of controlling the core recovery, the authors defined and used the jamming index which defines the effect of weight on bit on momentary rate of penetration. The efficiency of this method was proved by a test performed in industrial conditions while drilling a successive wellbore in the Grodziec Syncline.

This paper was written on the basis of own experience and analyses performed by the authors while conducting works in the Grodziec Syncline in the South West Poland.