ORIGINAL PAPER

INFLUENCE OF MODEL PARAMETER UNCERTAINTIES ON FORECASTED SUBSIDENCE

Wojciech GRUSZCZYŃSKI *, Zygmunt NIEDOJADŁO and Dawid MROCHEŃ

AGH University of Science and Technology, Faculty of Mining Surveying and Environmental Engineering, Al. Mickiewicza 30, 30-059 Cracow, Poland

*Corresponding author’s e-mail: wgrusz@agh.edu.pl

ABSTRACT

Surface deformation due to underground exploitation affects the safety of overlying structures. Forecasting can predict risks to surface structures and facilitates actions designed to improve their resilience and reduce the potential impact of mining activities. However, forecasting accuracy is limited. Therefore, in practice, model parameters are determined within a certain margin to ensure that critical values of deformation indicators for surface objects are not exceeded. For economic reasons, it is important to minimize these margins while also ensuring that safety is maintained. One important factor influencing forecasting accuracy is the uncertainty in deformation model parameters used for calculations. Therefore, it is critical to adopt an appropriate methodology for determining and addressing the uncertainties in deformation model parameters used in forecasting.

This study presents methods for estimating the Knothe's model parameters needed to forecast surface deformation caused by underground mining and defining the uncertainties in those forecasts. Depending on the parameter uncertainties, one of two methods for propagation is proposed: the Monte Carlo method or the law of propagation of uncertainty. Using this approach, it is possible to account for uncertainty and reduce forecast margins. A case study of hard coal mining in the Upper Silesian Coal Basin region of Poland is presented.

1. INTRODUCTION

Surface deformation resulting from underground exploitation is a direct hazard to overlying structures. Basing on forecasts of deformation indicators, one can predict the at-risk surface structures. Such predictions can inform and facilitate modifications to increase the resilience of overlying structures, optimize the distances between mining sites and structures, and/or determine whether to use filling materials (Kratzsch, 1983; Kwiatek, 2007; Peng, 1992; Whittaker and Reddish, 1989).

Forecasting accuracy is limited. Therefore, in practice, indicators are determined within certain margins to ensure that critical deformation values do not exceed the safety thresholds for surface structures. For economic reasons, it is extremely important to minimize these margins while also ensuring that safety is maintained (Ostrowski, 2006).

All forecasts in the article were conducted with use of the Knothe's model (Section 3.1) used in many countries and in particular in Poland. One important factor affecting forecasting accuracy in general and particularly the Knothe's model is the uncertainty in the deformation model parameters used in calculations (Hejmanowski and Malinowska, 2016). This uncertainty depends directly on the method used to determine the parameters. For forecasting in mine areas with a dearth of sufficient geodetic observations, simple assumptions are made based on historical average values for model parameters (Kowalski, 2005). These assumptions can result in large discrepancies between the forecasts and true deformation. Therefore, it is critical to adopt an appropriate methodology for determining and addressing the uncertainty in deformation model parameters used in forecasting (Hejmanowski and Malinowska, 2009). This study focuses on determining the uncertainties in Knothe's deformation model parameters, propagating these uncertainties, and evaluating their effects on forecast results. The results are illustrated with a case study on hard coal mining in the Upper Silesian Coal Basin region of Poland. The purpose of the paper is to show how source of the Knothe's model parameters affects subsidence prediction uncertainty. The paper aims to answer a questions:

- How much uncertainty there is in subsidence prediction based on Knothe's model parameters from regional observations from USCB in general or from region of given mining plant?
- How measurements carried out during earlier longwall operation affect uncertainty of estimates of Knothe's
Table 1 Exploitation characteristics in the eastern Upper Silesian Coal Basin, Poland.

<table>
<thead>
<tr>
<th>Longwall</th>
<th>Mining depth [m]</th>
<th>Operating height [m]</th>
<th>Width [m]</th>
<th>Length [m]</th>
<th>Mining start date</th>
<th>Mining end date</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>640</td>
<td>4.7</td>
<td>210/250</td>
<td>690</td>
<td>2012/04</td>
<td>2013/01</td>
</tr>
<tr>
<td>A2</td>
<td>595</td>
<td>4.7</td>
<td>250</td>
<td>320</td>
<td>2013/04</td>
<td>2013/09</td>
</tr>
<tr>
<td>B</td>
<td>630</td>
<td>4.6</td>
<td>250</td>
<td>610</td>
<td>2014/06</td>
<td>2015/04</td>
</tr>
<tr>
<td>C1</td>
<td>600</td>
<td>4.5</td>
<td>180</td>
<td>515</td>
<td>2015/01</td>
<td>2016/06</td>
</tr>
<tr>
<td>C2</td>
<td>595</td>
<td>4.3</td>
<td>250</td>
<td>120</td>
<td>2016/07</td>
<td>2016/09</td>
</tr>
</tbody>
</table>

Fig. 1 Sketch of the mining area with completed exploitation and locations of observation lines in the eastern Upper Silesian Coal Basin, Poland.

model parameters and what is even more important uncertainty of subsidence prediction?

This article deals with forecasting of subsidence at the final (later) stage of excavation based on earlier observations (i.e., predicting the development of deformation based on information known at the time of forecasting). This article is not concerned with modeling, which only illustrates the distribution of deformation for optimal parameter values. Most past studies (Cui et al., 2013; Hu et al., 2011; Huayang et al., 2002; Quinta, 2012; Lian et al., 2011; Liao, 1993; Nicieza et al., 2005; Polanin, 2015; Weifeng and Xiaohong, 2013) deal with modeling the phenomenon (i.e., determining how the phenomenon developed when the optimum parameter values have already been determined based on observations made after the end of exploitation). Although the Monte Carlo method has been used in past studies (Kwinta, 2010; Hejmanowski and Malinowska, 2016; Niedojadło, 2008; Niedojadło and Gruszczynski, 2015), none of these studies analyzed subsidence, and no attempt was made to verify the results of calculations using measurement data.

Satellites (e.g., InSAR, DInSAR, and TimeSAR) (Chaussard et al., 2014; Herrera et al., 2007; Grzovic and Ghulam, 2015; Perski et al., 2009; Qu et al., 2014; Zhang et al., 2015) and aerial imagery (LiDAR) (Ao et al., 2016; Khan et al., 2014) are often utilized to make observations of surface deformation. Unmanned aerial vehicles (UAVs) (Gruszczynski et al., 2017) have also been employed to conduct studies on surface deformation. However, in Poland and many other countries, conventional measurements from points stabilized on observation lines remain the most common technique for characterizing land surface changes. On the example of such observations, the methodology of determining the uncertainty of model parameters and deformation indices was illustrated.

2. RESEARCH SITE DATA

Analyses were carried out for influences of mining operations conducted on the northern slope of the primary basin in the eastern Upper Silesian Coal Basin (USCB). The deposits are characterized by beds with strata deposited monoclinically in individual tectonic blocks and separated by larger fault-type dislocations.

The main notable feature is the presence of carbon deposits in layers composed primarily of sandstone layers of relatively high stiffness. This relates to the characteristic manner in which the effects of mining are revealed during mining on subsequent longwalls and seams. Initiating exploitation where it has not yet been conducted results in significant movement of the terrain surface only after a significant exposure of the layer ceiling. Mining the first longwall in the initial stage of mining produces only slight movements of the terrain surface. Only a significant length of longwall mining or initial mining of another neighboring longwall results in a subsidence of the surface close to the expected value.

2.1. MINING OPERATIONS

Mining was conducted by collapsing the roof of five longwalls (Fig. 1). Longwalls A1, B, and A2 form a complex of excavated deposit space situated under the observation lines. These longwalls were mined from the second quarter of 2012 to the end of the first quarter of 2015.

Longwall C1 was excavated west of the initial north section of the N–S line. Its exploitation began at the end of 2015 and was completed in the second quarter of 2016; it has been extending to the west, gradually moving away from the observation line. The most recent mining in the area occurred in the C2
longwall; mining in this longwall began in the third quarter of 2016 directly at the north end of the N–S line. The development of this longwall parallel to longwall C1 (from the north) extends to the west. The mining of this longwall had not been completed as of September 2016. The characteristics of each longwall are presented in Table 1.

2.2. OBSERVATIONS

In the immediate area of mining operations height measurements were performed at points along two observation lines located above the completed mining operation, as shown in Figure 1. The north-to-south (N–S) observation line has 72 points that are spaced almost uniformly every 22.9–28.3 m. The total length of the line is approximately 1780 m. The west-to-east (W–E) observation line consists of 83 points spaced at 12.3 to 30.2 m. The total length of the line is just over 2055 m.

The subsidences observed on points from these lines were used to determine Knothe model parameters and their uncertainties in Section 4.3.

The measurements were performed between March 2009 and September 2016. This period covers the mining operation time of four of the described longwalls (Table 1); the analyzed observation lines are located within the range of influence of these longwalls. The mining of the C2 longwall was only initiated at the end of the measurement period. Measurements along the N–S line made in mid-September 2016 cover the initial development stage of this longwall. Consequently, the analyzed results for this longwall do not show the effect of final mining on the surface. This caveat only applies to the initial north portion of the N–S line, up to a maximum distance of 400 m.

The measurement periods are summarized in Table 2. These measurements allow the subsidence of Series A (A1+A2), Series B (A1+A2+B) and Series C (A1+A2+B+C1+C2) to be determined, as shown in Figure 1. Only the total subsidence (i.e., the subsidence calculated against the baseline measured in March 2009) was analyzed in this study. The measurement periods were selected so that the influence of time on the subsidence caused by the mining of subsequent longwalls could be neglected.

Table 2 Dates at which the measurement series were performed in the eastern Upper Silesian Coal Basin, Poland.

<table>
<thead>
<tr>
<th>Series</th>
<th>Time of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2009/03</td>
</tr>
<tr>
<td>A</td>
<td>2014/04</td>
</tr>
<tr>
<td>B</td>
<td>2015/04</td>
</tr>
<tr>
<td>C</td>
<td>2016/09</td>
</tr>
</tbody>
</table>

3. CALCULATION METHODS

3.1. KNOTHE’S MODEL

Many methods have been developed to predict surface deformation (Kratzsch, 1983; Peng, 1992; Whittaker and Reddish, 1989). Generally, these methods can be divided into two groups: numerical and empirical methods. Numerical methods digitize the medium as a continuous material, including the finite element method (Tajduś, 2009; Uluń et al., 2013) and finite difference method (Alejano et al., 1999; Nengxiong et al., 2013; Shahriar et al., 2009) or a non-continuous material, such as the discrete element method (Barbato et al., 2016; Salmi et al., 2017). Another noteworthy method is cellular automata theory (Lian et al., 2011; Saavedra-Rosas and Drage, 2014; Sikora, 2016).

The second group includes the profile function (Asadi et al., 2004; Diez and Álvarez, 2000; Torano et al., 2000; Waddington and Kay, 1995), graphical method (NCB, 1975), application of artificial neural networks (Ambrożič and Türk, 2003; Gruszczynski, 2007; Weifeng and Xiaohong, 2013) and influence function (Álvarez-Fernandez et al., 2005; Cui et al., 2013, 2000; Ghabraie et al., 2017; Huayan et al., 2002; Liao, 1993; Nicieza et al., 2005; Ren et al., 1987, 2014; Sheorey et al., 2000; Singh and Singh, 1998). Based on large collections of observations, the relationships between measurement results and geometric conditions from a mining operation are determined. Generally, geologic conditions are not directly taken into account.

In the research all calculations were made using the Knothe model (Knothe, 1957), which belongs to influence function method group. Knothe’s model and its modified versions (Ruhrkohle methods or SDPS; Ehhradt and Sauer, 1961; Karmis et al., 1987) have been widely used to predict surface deformation in many countries, including Australia, Czech Republic, Germany, USA, Poland and China (Byrnes, 2003; Doležalová et al., 2009; Ehhradt and Sauer, 1961; Karmis et al., 1990; Knothe, 1957; Luo and Cheng, 2009). This method is based on the assumption of a normal distribution of influence. For the spatial problem, the surface of the influence is given by Equation (1):

$$G(x, y) = \frac{a \cdot g}{r^2} \exp \left[ -\frac{\pi (x^2 + y^2)}{r^2} \right], \quad (1)$$

where:

- $a$ is the exploitation coefficient,
- $g$ is the operating height,
- $r$ is the radius of main influences, and, $x, y$ are the coordinates of the influenced point.

Assuming that the exploitation surface is a semiplane (practically, a field with a minimum size of $2r \times 2r$), the subsidence reaches a maximum value of $s_{max}$ as in Equation (2):
$s_{\text{max}} = a \cdot g.$ \hspace{1cm} (2)

The parameter determining the dispersion of influences $r$ (Fig. 2) can be calculated from the relationship in Equation (3):

$$r = \frac{H}{\tan \beta},$$ \hspace{1cm} (3)

where:

- $H$ is the exploitation depth, and
- $\beta$ is the angle of main influences.

There are two methods for determining the values of Knothe's model parameters. Average values from a mining region or mine are assumed when data from earlier surveys are lacking (Kowalski, 2005). Alternatively, surveys along observation lines are used (Kwinta, 2012b). The least-squares method can be used to determine optimal parameter values (Białek, 2003; Kwinta, 2011) to provide a fit between the theoretical and true subsidence.

The meaning of the parameters in the model can be classically conceptualized over a large mining field, as shown in Figure 2:

1. The exploitation coefficient $a$ characterizes the mechanism for closing post-mining voids for a given mining method (e.g., collapse of roof mining, backfill mining, room and pillar mining). It describes the ratio of the maximum subsidence $s_{\text{max}}$ to the average mining thickness $g$ and is a dimensionless value that ranges from 0 to 1:

$$a = \frac{s_{\text{max}}}{g}.$$ \hspace{1cm} (4)

2. The tangent of the angle of main influences, $\tan \beta$, is a parameter characterizing the rock mass and its geo-mechanical properties. Its value is determined based on the profile of the edge of the subsidence basin, where the maximum slope $T_{\text{max}}$ occurs. $T_{\text{max}}$ is the inclination of the profile at the inflection point, as described by Equation (5):

$$T_{\text{max}} = \frac{s_{\text{max}}}{r},$$ \hspace{1cm} (5)

Taking into account Equations (5) and (3), we obtain:

$$\tan \beta = \frac{H \cdot T_{\text{max}}}{s_{\text{max}}}.$$ \hspace{1cm} (6)

3. The operating rim $p$ is the result of analyzing the subsidence basins while addressing the systematic divergences between theoretical subsidence and results of geodetic measurements. Knothe's model assumes the occurrence of a subsidence of $0.5 s_{\text{max}}$ (i.e., the inflection point of the subsidence basin profile) over the rim of the exploitation (Fig. 2). The displacement of the inflection point by a certain value $p$ in the direction of the exploited space is observed. In forecasting calculations, the operating rim $p$ can be defined as a function of the depth of exploitation (Kowalski, 2007):

$$p = k \cdot H,$$ \hspace{1cm} (7)

where $k$ is the factor of the operating rim parameter.

Introducing the operating rim reduces the dimensions of the effective mining field, which is assumed for calculation, by parallel moving the mining edges in the working direction by $p$.

The values of the parameters $a$, $\tan \beta$, and $p$ are constant over time for a mining operation at a given site. However, with multiple exploitation events, these parameters can change due to the changes in the geo-mechanical properties of the rocks.
or the possibility of residual effects from earlier exploitation.

The Knothe’s model itself doesn’t take into account the prior mining history of the site, which is somewhat limiting. The mining history can be incorporated into model by gradual change in model parameters for subsequent predictions. One of the solutions to this problem can be an introduction of additional parameter to the Knothe model accountable for influence of activation of the old oobs on deformation process (Białek, 2003).

In addition, the effect of rock-mass dehydration is taken into account in the model; rock-mass dehydration is assumed to increase subsidence by a constant value for a given mining operation. The subsidence after removing the effects of dehydration was analyzed.

### 3.2. METHOD FOR ESTIMATING MODEL PARAMETER VALUES

When estimating Knothe’s model parameters, currently, the most frequently used criterion minimizes the sum of the squared differences between observable and modeled subsidence values. Many different algorithms are used to effectively minimize this error function. Some algorithms do not use the gradient of this function, whereas others require the gradient to be determined. From a practical point of view, these methods differ from each other in terms of stability and speed.

In this study, a Gauss-Newton method based on the solution to the supernumerary system of equations linearized using the Taylor series was used in determination of Knothe’s model parameters and their uncertainties based on current geodetic observation (section 3.3). The unknowns in such a system of equations are the changes in model parameters, and individual equations correspond to the observation of subsidence at one point, as in Equation (8):

$$\Delta \alpha, \Delta p, \text{and } \Delta \tan \beta \text{ are the changes in a parameter value in a given algorithm step,}$$

$$\frac{\partial s}{\partial a}, \frac{\partial s}{\partial p}, \text{and } \frac{\partial s}{\partial \tan \beta} \text{ are the partial derivatives of modeled subsidence values for model parameters of the } i^{th} \text{ point,}$$

$$s_i \text{ is the modeled subsidence of the } i^{th} \text{ point, and}$$

$$s_i^0 \text{ is the observed subsidence of the } i^{th} \text{ point.}$$

The first forward difference was used to calculate the derivative values:

$$f'(x_i) = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} + O(x_{i+1} - x_i),$$

where:

$f(x_i)$ is the derivative value,

$f(x_i)$ and $f(x_{i+1})$ are function values appropriate for argument values of $x_i$ and $x_{i+1}$, respectively, and

$O(x_{i+1} - x_i)$ is the error order for the numerically determined derivative value.

Due to the linearization of equations, the final solution is obtained iteratively, and the corresponding function values (subsidence) and derivatives are determined for the values of actual parameters in the given iteration. If the distances between the optimal values of the unknowns and those determined in a given step are sufficiently small, the solution errors due to equation linearization become negligible. In this iteration, the parameter-setting process is concluded. The advantage of using this method is the potential for a relatively simple estimation of solution uncertainty (i.e., model parameter values). The variance–covariance matrix of unknowns (parameters) can be determined according to Equation (13):

$$\sum_x = \sigma_x^2 (A^T A)^{-1},$$

where:

$\sigma_x^2$ is the variance of the random factor of subsidence computed from Equation (14) as

$$\sigma_x^2 = \frac{\sum_{i=1}^{n} (s_i - s_i^0)^2}{n - u},$$

where:

$n$ is the number of observations (equations), and

$u$ is the number of unknowns (parameters).

The ability to estimate the uncertainty in the unknowns and the correlation between their values is important for assessing the forecast accuracy based on the parameter values. It should be stressed that this assessment is based on differences between the modeled and observed subsidence values. The
modeled values are the calculated subsidence values for the optimal Knothe's model parameter values based on observed subsidence values in a given exploitation state. Any discrepancy is treated as an additive noise of zero mean; that is, the errors are assumed to be completely random. This means that any systematic discrepancies between the model and observations will also be treated as random values, which will lead to erroneous evaluations of both the parameter values and their uncertainties. Therefore, it is important to include all the systematic factors in the model and to remove them from the differences between observed and modeled values. The uncertainty assessment for the parameters determined using this method is more accurate for smaller noise values. In addition, better accuracy is obtained for more linear relationships between parameter values and modeled function values (subsidence). Reliable estimates of the noise level are required to obtain high-quality evaluations. Meeting this requirement is directly related to the number of observations and thus to the number of supernumerary equations of the system to be solved.

3.3. METHODS OF PROPAGATION OF MODEL PARAMETER UNCERTAINTIES

There are many ways to propagate the uncertainty in model parameters. In this study, we use the Monte Carlo method (Metropolis and Ulam, 1949; Ulam, 1950) and law of propagation of uncertainty (JCGM, 2008). These two methods provide different possibilities and generate significantly different computing loads. Using both methods, the uncertainty in the model parameters and model accuracy are assumed. Model accuracy means that the model has parameter values with no systematic discrepancy between the determined and observed deformation indicators. Therefore, any discrepancy between the model and the observations is treated as noise with a zero mean and mutually non-correlated values.

Based on these assumptions, it is possible to estimate the uncertainty in the predicted deformation indicators using the aforementioned methods and to attribute a probabilistic interpretation to them.

3.3.1. MONTE CARLO METHOD

The Monte Carlo method generally requires considerably more computational effort than law of propagation of uncertainty (LPU) and provides a mechanism for generating arbitrary deformation indicator statistics. In addition, this method can easily operate with a large number of uncertainty parameters and strongly nonlinear functions.

One use of the Monte Carlo method is estimating the probability distribution (density function), where the arguments of the function (model) are randomly selected, and the function values are calculated based on sampled parameter values. The distributions from which these parameters are drawn should be predetermined based on previous observations. The simulation results can then be analyzed using statistical methods.

Model runs using the Monte Carlo method were performed for 10,000 iterations, which is considered sufficient for accurately determining the distribution of subsidence values at individual points and all related statistics.

3.3.2. LAW OF PROPAGATION OF UNCERTAINTY

The assessment of function uncertainty was also conducted using the law of propagation of uncertainty. Versions of this law differ based on the inclusion of the covariances of function argument values (independent variables) and the degree of function expansion in the Taylor series when approximating function values.

The law of propagation of uncertainty used in the study was based on the Taylor series expansion of a function using only first-degree terms and omitting higher-degree derivatives. In calculations, the covariances between the set parameter values were taken into account. In this case, Equation (15) describes the law as follows:

$$\sum_s = J \times \sum_x \times J^T,$$

where:

$$\sum_s$$ and $$\sum_x$$ are the variance–covariance matrices of predicted subsidence (or other indicators) and parameters, respectively, and $$J$$ is the Jacobian of the function.

In the specific case of propagation of uncertainties in Knothe's model parameters right hand side terms in Equation (15) take the following form:

$$J = \begin{pmatrix} \frac{\partial s_1}{\partial a} & \frac{\partial s_1}{\partial \tan \beta} & \frac{\partial s_1}{\partial p} \\ \vdots & \vdots & \vdots \\ \frac{\partial s_k}{\partial a} & \frac{\partial s_k}{\partial \tan \beta} & \frac{\partial s_k}{\partial p} \end{pmatrix},$$

$$\sum_n = \begin{pmatrix} \sigma^2_s & COV (a, \tan \beta) & COV (a, p) \\ COV (a, \tan \beta) & \sigma^2_{a,\beta} & COV (\tan \beta, p) \\ COV (a, p) & COV (\tan \beta, p) & \sigma^2_p \end{pmatrix}$$

where:

$$\sigma^2_s, \sigma^2_{a,\beta}, \sigma^2_p$$ are $$a$$, $$\tan \beta$$, and $$p$$ variances respectively, and

$$COV (a, \tan \beta), COV (a, p), COV (\tan \beta, p)$$ are covariances between parameters.

Here, only subsidence uncertainties (square roots of $$\sum_s$$ matrix diagonal elements) were estimated; the covariances between them were not calculated.

This form of the uncertainty propagation law can provide a relatively accurate determination of the uncertainty in deformation indicators only for
Table 3 Subsidence model parameters for the Upper Silesian Coal Basin region.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$a$</th>
<th>$\tan\beta$</th>
<th>$k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.80</td>
<td>1.92</td>
<td>0.10</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.12</td>
<td>0.28</td>
<td>0.05</td>
</tr>
</tbody>
</table>

relatively small uncertainties in model parameters. The advantage of this method is its low computational complexity; for each parameter and point, it is only necessary to determine the partial derivative value once, and there is no difficulty when considering covariances between model parameters.

The disadvantage of this approach is the increase in error in indicator uncertainty estimates with increasing parameter uncertainty for nonlinear functions.

Equation (15) requires knowing the values of the partial derivatives of the function with respect to its arguments (model parameters). For simple theoretical mining operation shapes, it is possible to analytically determine the exact derivative values. However, in practice (for realistic panel shapes), the analytical formulation of the necessary equations becomes too complex. In this case, the values of the partial derivatives can be determined numerically. In this study, the first forward difference is used for this purpose.

4. RESULTS

All the forecasts described in this section were made for the surface state after all mining of the longwalls shown in Figure 1 had finished (i.e., for longwall conditions categorized up to the C observation series in September 2016). The uncertainty of a subsidence forecasts based on mean parameters from the USCB region (section 4.1) and general mine parameters (section 4.2) was obtained by using the Monte Carlo method. The uncertainty of the subsidence forecast based on current geodetic observation (section 4.3) was obtained using law of propagation of uncertainty.

4.1. FORECASTING BASED ON REGIONAL PARAMETERS

When there is a lack of sufficient geodetic observations from the mine area of interest, forecasts are based on mean parameters for the USCB region. In this case, the uncertainty in the parameter values is very high, which clearly results in very high uncertainty in the values of the deformation indicators.

The mean parameter values of Knothe’s model and their uncertainties are shown in Table 3. These values are based on values from Kowalski (2007), who analyzed and determined Knothe’s model parameters independently for 29 subsidence basins from many different mines in the USCB region. These basins were formed in a number of deposits with different depths, relative and absolute dimensions, exploited panels, and geologic conditions. All analyzed basins were formed from the collapse of roof-mining hard coal deposits. An analysis of the determined parameter values indicates a lack of correlation between parameters, as well as the lack of correlation between the parameters and depth of exploitation. Shapiro-Wilk’s test showed that for each parameter, there is no basis for rejecting the normal-distribution hypothesis.

The uncertainty of forecast based on mean parameters from the USCB region was obtained using the Monte Carlo method. Estimates of operating rim parameter  $p$  was calculated according to Equation (7). Adopted for calculations mining depth of all longwalls (based on exploitation characteristics in Table 1) was equal to 600 m.

Figure 3 shows the observed and modeled subsidences based on mean parameters and the subsidence values for selected percentiles. The results are presented in percentiles instead of standard deviations because of the potential for strongly skewed non-normal distributions of subsidence values. In these cases, the percentile values are more illustrative and easier to interpret than the standard deviations. Percentile values were chosen so that 68% (16%–84%) and 95% (2.5%–97.5%) of cases are included within the intervals; these two intervals roughly correspond to the probabilities of finding the value of the random variable within the intervals  $\bar{x} \pm \sigma_x$  and  $\bar{x} \pm 2\sigma_x$  , respectively, for the normal distribution of the random variable. However, when using the Monte Carlo method, the percentile values can be selected almost arbitrarily provided that the values of percentiles in the tail of the variable distribution are determined less accurately than those close to the center of the distribution. Exemplary histograms and approximations of the distributions of modeled subsidence values at selected points are shown in Figure 4.

The width of the intervals between percentiles and discrepancy between the subsidence modeled for the average value and observed parameter value should be noted. The root mean square error (RMSE) of the subsidence for mean parameter values is  $\pm 436$ mm. More importantly, the widths between the percentile intervals clearly indicate a low confidence in the forecast. When analyzing the distributions shown in Figure 4, in the event of skewed distributions of random variables, the mean, median, and mode of the value do not overlap. Therefore, making any point estimate requires choosing the statistics it should reflect. The forecast should provide a mechanism for assessing the risk to the surface and
Fig. 3 Observed and predicted subsidence based on parameters from the USCB region along (a) the N–S and (b) the W–E observation lines shown in Figure 1.

Fig. 4 Subsidence histograms at selected points along the observation lines obtained for parameters from USCB region using the Monte Carlo method: points located about (a) 400 m from the beginning of the N–S line and (b) 700 m from the beginning of the N–S line.
Table 4 Subsidence model parameters for the eastern Upper Silesian Coal Basin mine.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>(a)</th>
<th>(\tan \beta)</th>
<th>(p) [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.49</td>
<td>1.81</td>
<td>52</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.10</td>
<td>0.28</td>
<td>12</td>
</tr>
</tbody>
</table>

Fig. 5 Observed and predicted subsidence based on parameters from the mine area, along (a) the N–S and (b) the W–E observation lines shown in Figure 1.

objects related to the influences of the underground operation. Typically, average values are used with some margin. The subsidence values for selected percentiles reflect this margin and facilitate estimating the probability of a given deformation indicator being exceeded. In practice, it is important that this margin is minimized while ensuring that the values of the deformation indicators are not underestimated. Geodetic observations of the terrain in the mine area or conducted directly over the considered exploitation are extremely helpful for forecasting and the associated calibration of model parameters.

4.2. FORECASTING BASED ON MINE PARAMETERS

A forecast using Knothe's model parameters derived from observations from the mine area provides narrower confidence intervals. However, observation analysis is required from at least a couple of the observation lines.

In this case, observations from 11 different excavations (and observation lines over longwall blocks) carried out before the excavation for which the subsidence forecast is made in other regions of the same mine were used. The number of observations is so small that the normality of the parameter distributions is only an assumption and has not been verified. Parameters from the analyzed mining locations were independently calculated based on the least-squares method of differences in subsidence determined between observations and modeled values. Individual parameters determined from one line can be strongly correlated, while the parameter triplets (\(a\), \(\tan \beta\), and \(p\)) determined on independent lines located in the mine are uncorrelated. Therefore, we can assume that there is no correlation between parameter values determined from different lines. An appropriate statistical evaluation of Knothe's model parameters estimated from observations in the mine area (other than from observation lines described in Section 2.2) is summarized in Table 4. The value of the operating rim and its uncertainty independent of the depth of exploitation was used because measurements from the mine area did not confirm any dependence.

Geodetic observations from the mine area confirm a relatively atypical geologic structure and its influence on surface behavior. This is best illustrated by the difference between the average value of parameter \(a\), responsible for modeling the scale of the phenomenon and estimated for the whole USCB region (0.8), and that of the mine area (0.49). Figure 5 compares the observed subsidence with that modeled...
Table 5 Parameters estimated using observations from the analyzed lines for successive operating stages A, B, and C.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>0.74</td>
<td>0.60</td>
<td>0.61</td>
</tr>
<tr>
<td>$\tan \beta$</td>
<td>1.35</td>
<td>1.76</td>
<td>1.72</td>
</tr>
<tr>
<td>$p$ [m]</td>
<td>86</td>
<td>84</td>
<td>82</td>
</tr>
</tbody>
</table>

Table 6 Uncertainties and correlations between parameter values estimated for mining operation stage A (2014/04).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$a$</th>
<th>$\tan \beta$</th>
<th>$p$ [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>0.11</td>
<td>-0.46</td>
<td>0.98</td>
</tr>
<tr>
<td>$\tan \beta$</td>
<td>-0.46</td>
<td>0.04</td>
<td>-0.30</td>
</tr>
<tr>
<td>$p$ [m]</td>
<td>0.98</td>
<td>-0.30</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Table 7 Uncertainties and correlations between parameter values estimated for mining operation stage B (2015/04).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$a$</th>
<th>$\tan \beta$</th>
<th>$p$ [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>0.02</td>
<td>-0.90</td>
<td>0.96</td>
</tr>
<tr>
<td>$\tan \beta$</td>
<td>-0.90</td>
<td>0.04</td>
<td>-0.78</td>
</tr>
<tr>
<td>$p$ [m]</td>
<td>0.96</td>
<td>-0.78</td>
<td>2.4</td>
</tr>
</tbody>
</table>

Table 8 Uncertainties and correlations between parameter values estimated for mining operation stage C (2016/09).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$a$</th>
<th>$\tan \beta$</th>
<th>$p$ [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>0.02</td>
<td>-0.89</td>
<td>0.98</td>
</tr>
<tr>
<td>$\tan \beta$</td>
<td>-0.89</td>
<td>0.03</td>
<td>-0.82</td>
</tr>
<tr>
<td>$p$ [m]</td>
<td>0.98</td>
<td>-0.82</td>
<td>2.5</td>
</tr>
</tbody>
</table>

using the Monte Carlo method and estimates of mean parameter values and their standard deviations based on observations from the mine area.

Compared to the model based on USCB regional parameters (Fig. 3), the modeled subsidence values based on the average mine parameters approach the observed values, and the ranges between the percentiles clearly narrow. Therefore, using geodetic observations for forecasting is significant and reduces the necessary margin around the forecasted deformation indicator values. In many cases, this type of prepared forecast will be optimal given the available observations.

4.3. **FORECASTING BASED ON CURRENT GEODETIC OBSERVATION**

The measured data from the N–S and W–E observation lines (Section 2.2) were used to complete the short-term forecasts described in this section. The Knothe's model parameters and their uncertainties were determined based on observations from subsequent stages of operation according to the method described in Section 3.2.

The uncertainty in the determined deformation indicators was determined basing on parameter uncertainties determined using the law method described in Section 3.3.2. In this case, no distributions were modeled; only the uncertainty in the deformation indicator was estimated. The expanded forecast uncertainty for the deformation indicator was assumed to be three times the standard uncertainty. Assuming a normal distribution, these intervals correspond to approximately 68% and 99% of the probability that the observed values of the indicators will be within the intervals $\bar{X} \pm \sigma$ and $\bar{X} \pm 3\sigma$, respectively. The distribution of this "noise" around the predicted values will not show a normal distribution; however, for small values of uncertainty, the normal distribution approximation should not introduce significant errors.

Table 5 provides estimates of parameter values for subsequent operating stages. Tables 6, 7, and 8 show the estimated uncertainty values for the determined parameters (table diagonals) and Pearson's linear correlation coefficients (apart from diagonals).
Fig. 6  Observed and forecasted subsidence for stage C (2016/09) based on parameters calculated from the same observation line for operational stage A (2014/04) along the (a) N–S and (b) W–E observation lines shown in Figure 1.

Fig. 7  Observed and forecasted subsidence for stage C (2016/09) based on parameters calculated from the same observation line for operating stage B (2015/04) along the (a) N–S and (b) W–E observation lines shown in Figure 1.
Table 9 Predicted RMSE for subsidence in the mining area.

<table>
<thead>
<tr>
<th>Parameter sources</th>
<th>Predicted subsidence RMSE [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>USCB region (mean parameters)</td>
<td>436</td>
</tr>
<tr>
<td>Mine region (mean parameters)</td>
<td>88</td>
</tr>
<tr>
<td>Observations 2014/04 (stage A)</td>
<td>65</td>
</tr>
<tr>
<td>Observations 2015/04 (stage B)</td>
<td>40</td>
</tr>
<tr>
<td>Observations 2016/09 (stage C)</td>
<td>35</td>
</tr>
</tbody>
</table>

Figure 6 shows the predicted subsidence values for stage C based on parameters and their estimated uncertainties from stage A. Notably, there are wide ranges in predicted subsidence. These ranges arise directly from the early stage of the mining operation from which the parameters were estimated and the relatively high proportion of random factors with respect to the systematic (modeled) subsidence factors. During the operation stage used for forecasting, the maximum observed subsidence was −415 mm with an RMSE of ±29 mm.

Figure 7 shows the predicted subsidence for operation stage C based on the parameters and their uncertainties determined from operation stage B. There is a clear narrowing of the predicted subsidence intervals compared to the forecast based on parameters from operating stage A. This difference has two explanations. First, the parameters in stage B were closer to the conditions at the forecasted time. Second, there was a large increase in the value of the systematic subsidence factor compared to the random noise in the observed subsidence. The RMSE for fitting the model to observations at operating stage B was ±36 mm with a maximum observed subsidence of −1484 mm.

Table 9 shows the RMSE values for the subsidence forecasts calculated using the parameters determined at a given stage of development. Notably, there is a decrease in RMSE as the mining stage approaches the target condition (i.e., the time at which the forecast was made).

5. DISCUSSION

The results of this study show the significant effect of geodetic observations on forecasting accuracy. The proposed methods for estimating forecasted indicator values can provide reliable forecasts at any stage of exploitation. Decreasing the forecasting margin depends on the available observation data (i.e., resources). The forecast should be interval-based rather than point-based; for an interval-based forecast, certain confidence intervals for forecasted indicators are determined, including the uncertainty in estimated model parameter values. With such an approach, the forecast will be safe (i.e., it will be made with the necessary margin). This will ensure that the values observed by the forecaster at the chosen probability level do not exceed the forecasted deformation indicators.

When both estimating parameters and forecasting future values, it is extremely important to use a model that will not systematically deviate from observed deformation indicator values. Forecasting models assume an immutable set of parameters, both spatial and temporal, from which the indicators can be calculated as mining operations progress. In general, this assumption is fulfilled, at least to a good approximation. Possible deviations may result from unknown geologic differences in the area of operation. Such deviations may, to some extent, unpredictably distort the forecast based on the applied model.

Confidence intervals for deformation parameters determined using the Monte Carlo method (i.e., based on parameters from the USCB region or mine area) account for variability in model parameter values. They are based on measurements from many mines or from different sites at a given mine; thus, the range in model parameter values includes their variability.

Determining parameters or making predictions for temporary states of deformation requires addressing the influence of time on land displacement values. Due to the correlation between determined parameters, the error in determining this parameter will influence the estimated values of all other parameters; it will also bias the forecast. For these reasons, it is advisable to estimate parameter values based on deformation indicator values after the ground has settled post-mining. Nonetheless, in this case, the problems described for the analyzed operation were not critical, and the uncertainties in the deformation indicators based on series B were correct.

A broader perspective for evaluating the estimated parameter values and uncertainty with which they were determined uses the distribution of RMSE changes (i.e., subsidence fitting), as presented in Figure 9. These graphs were generated according to Equation (18):

$$\Delta E = \sqrt{(E^2 - E_{\text{min}}^2)},$$

where:

- $\Delta E$ is the change in the RMSE value due to a change in the model parameter values relative to the optimum values for observations from a given operating stage,

- $E$ is the RMSE value for a modeled subsidence with respect to values determined based on observations from a given mining stage and model parameter values, and
INFLUENCE OF MODEL PARAMETER UNCERTAINTIES ON FORECASTED SUBSIDENCE

Fig. 8 Distribution of increasing RMSE values for fitting the model to observation data from individual mining operation stages A, B, and C.

$E_{\text{min}}$ is the minimum RMSE of the modeled subsidence (i.e., the RMSE value for optimal model parameter values based on observations for a given exploitation stage).

The contour lines in Figure 8 are shown in meters, with each successive contour representing a $\Delta E$ greater than the previous by 3 cm. Each of the distributions in Figure 9 was performed for two parameters assuming the optimum (constant) value of the third parameter based on observations.

The calculated RMSE values can be interpreted as follows. If the model is free from systematic errors, $E_{\text{min}}$ is the characteristic (statistic) of the random spread (noise) of the observed values relative to modeled values. In this case, $\Delta E$ can be interpreted as the impact of the error associated with the selected model parameters on the RMSE of the forecasted result.

The noise level ($E_{\text{min}}$) had similar values for all operating states: ±29 mm for stage A, ±36 mm for stage B, and ±35 mm for stage C. An examination of Figure 8 illustrates several important facts. First, "surfaces" within the contours of constant $\Delta E$ values for the assumed value (e.g., 0.03 cm) decrease as operations progress, which reflects the greater certainty in model parameters; the difference between status A and status B is particularly notable. Second, there are clear changes in parameter directions that cause large or small changes in $\Delta E$; these are specific main directions of the parameters. Finally, for a given operating state, there are many combinations of model parameters that result in almost identical subsidence changes at the observed site points based on the
subsidence RMSE values. However, the forecasts (i.e., predicted land subsidence as mining operations progress) are not identical based on the sets selected. The optimum estimated parameters at a given time period are not necessarily optimal for subsequent operating states. Therefore, this point-related estimate, which does not account for both the uncertainty in the estimated parameter values and correlations between them, can erroneously estimate the values of land surface point deformation indicators.

The comparison of the obtained results poses some difficulties as the majority of research works concentrate on the accuracy of the model (with the knowledge of optimal parameter values) and not on the accuracy of forecasting. The calculations presented herein were made based on information known at the time of preparing the forecast, in contrast to in many past studies (Cui et al., 2013; Hu et al., 2011; Huayang et al., 2002; Kwinta, 2012a; Lian et al., 2011; Liao, 1993; Nicieza et al., 2005; Polanin, 2015; Weifeng and Xiaohong, 2013), which assumed the knowledge of optimal model parameter values determined after the exploitation. The relative errors in the modeling of subsidence with optimal parameters in these works were on the order of 4%–8%, depending on the case considered. The problem with this measure of error is its dependence on the denominator value. Many observations suggest that final subsidence modeling error has a nearly constant value over one panel or block of panels. As a consequence, the relative error is much larger at the ends of observational lines (influence borders) compared to at their centers, where subsidence is relatively large.

Another frequently used measure of subsidence modeling or prediction error is scaled RMS error (MSEC, 2014a), which is defined as

$$\text{SRMS} = \frac{\text{RMS}}{s_{\text{ext}}^\text{pred}} \times 100\%$$

where:

- \(\text{SRMS}\) is scaled RMS error,
- \(\text{RMS}\) is subsidence RMS error, and
- \(s_{\text{ext}}^\text{pred}\) is the extreme forecasted subsidence absolute value.

Research conducted in the USCB region by several authors indicates that due to the irregularity of subsidence, the SRMS value in the case of longwall mining is not less than 1%–3% on average (Stoch, 2005; Kowalski, 2007). These studies were carried out without any association with a specific subsidence model and were based on the comparison of observed subsidence with regular curves fitted with several parameters. In this way, the spread of subsidence was estimated assuming the existence of a smooth average distribution of deformation indicator, for which the observed one is only one of the possible realizations of the deformation process. This scattering of subsidence is not included in the intervals of the subsidence values (percentiles, \(\bar{p} \pm \sigma_p\) and \(\bar{p} \pm 3\sigma_p\)) in Figures 3 and 5–7. The estimated value of this intrinsic noise (standard deviation) of subsidence for the analyzed case is on the order of 15–45 mm. However, it should be borne in mind that its inclusion does not linearly shift the value of the interval boundaries because this scattering square will be added to the squared \(\sigma_p\). For this reason, the significance of the subsidence intrinsic irregularity will be the higher the smaller is the forecast error related to the uncertainty of the model parameters.

A large study on Australian coal fields in Bowen Basin and in NSW regions (MSEC, 2014b) covering data from few hundreds of observational lines showed that the prediction errors are much larger than the modeling errors cited above. An assessment of prediction errors in that study indicated that \(\text{SRMS}\) error for the incremental profile method (IPM) is in the range of about 15% to 30%, depending on the mining area. At the same time, the study indicated a 10%–15% probability of the predicted maximum subsidence being exceeded. For comparison, the \(\text{SRMS}\) values for the lines analyzed in this study based on the average values of parameters are presented in Table 10. These values show only the accuracy of forecasts based on average parameter values without accounting for the margin of safety associated with the uncertainty in their designation. Accounting for the safety margin will generally increase the forecast error; however, it will reduce the probability of exceeding the values forecasted based on the observed values.

<table>
<thead>
<tr>
<th>Parameter sources</th>
<th>(\text{SRMS}) [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>USCB region (mean parameters)</td>
<td>17.3</td>
</tr>
<tr>
<td>Mine region (mean parameters)</td>
<td>5.7</td>
</tr>
<tr>
<td>Observations 2014/04 (stage A)</td>
<td>4.7</td>
</tr>
<tr>
<td>Observations 2015/04 (stage B)</td>
<td>2.7</td>
</tr>
<tr>
<td>Observations 2016/09 (stage C)</td>
<td>2.3</td>
</tr>
</tbody>
</table>

In another accuracy study of forecasts made for a large number (>30) of exploitations in the USCB area (Kowalski, 2015), the following indicator was used to determine the forecast error:
\[ \Theta_s = \frac{s_{pred}^{\text{pred}} - s_{obs}^{\text{obs}}}{s_{obs}^{\text{obs}}} \times 100 \% \], 

where: 
\( \Theta_s \) is the error indicator of the extreme subsidence forecast, and 
\( s_{obs}^{\text{obs}} \) is the extreme observed subsidence absolute value.

The average absolute value of the \( \Theta_s \) index (average error) obtained in these studies is about 15\%, and the standard deviation of this index is \( \pm 17.5 \% \). The average value of this index calculated over the whole dataset used in the study is close to zero (less than 1\%). Interestingly, in an analysis of the NSW region in Australia (MSEC, 2009–2017) focused on the accuracy of forecasts using IPM, the obtained value of this indicator was similar (an average error of approximately 16\%). The indicator \( \Theta_s \) compares only extreme values of subsidence; it does not verify their location.

In the analyzed case, the application of mean USCB parameters results in a discrepancy in the extreme subsidence of 60\%; in contrast, the use of mine-region average parameters gives a \( \Theta_s \) value close to 0\%. The analysis of the data collected in Table 10 and Figure 5 indicates that such high compliance for the use of mine-region average parameters is accidental and does not represent the whole of the forecast results. The value of this indicator using parameters determined from only the A-series measurement results is 11\%; the values determined from only the B- and C-series measurements are approximately 3.5\%. The obtained values of \( \Theta_s \) are similar in both works.

The high value of the \( \Theta_s \) obtained using the average parameters for the USCB also coincides with the observations from Kowalski’s paper (2015), who found a \( \Theta_s \) error indicator of the forecast of up to 50\% using the mean parameters for the USCB. These results confirm that using the average parameters of a large region (e.g., the entire USCB) for forecasting is a last resort and carries a significant risk of large errors.

The presented solution is a special case that can be adapted to most prognostic methods. As a result, a simple tool for determining the uncertainty of prepared mining deformation forecasts is obtained, which is particularly important for urbanized areas (e.g., USCB) and for assessing the safety of sensitive objects (e.g., monuments, engineering structures, and technical infrastructure), where even a slight exceedance of permissible deformation values may lead to structural damage. As rightly pointed out in the MSEC study (2014b), the probability of exceeding the forecasted values is a measure of the conservativeness of the forecast and should not be analyzed separately from its accuracy (e.g., measured by SRMS error). In practice, a balance in this trade-off between the accuracy and conservativeness of the forecast should be struck while taking into account both the applicable legal regulations and the resistance of objects located in the rock mass and on the surface.

6. CONCLUSIONS

The following conclusions were drawn based on the research results:

1. Model parameters determined based on observations of surface points from a specific period of mining operations, especially in the early stages of development, are characterized by high uncertainty.

2. Parameter uncertainty decreases as mining operations progress, which is associated with an increase in the value of the systematic subsidence factor relative to the value of the random factor.

3. Making predictions without accounting for the uncertainty in the assumed model parameters and correlations between them can lead to significant discrepancies between the actual and modeled deformation indicators. Furthermore, accounting for these uncertainties can reduce the risk of overestimating or underestimating the deformation indicator values, which may both be of significant economic importance.

4. Using the Monte Carlo method provides a mechanism for making predictions, even when there is high uncertainty in the model parameters.

5. The disadvantage of the Monte Carlo method is its relatively high computational effort, which depends on the required accuracy of the calculated deformation indicator statistics.

6. For deformation model parameters with relatively small uncertainties, good results are achieved by applying the law of propagation of uncertainty; the advantage of this method is the relatively low computational effort.

ACKNOWLEDGMENTS

This work was supported by the statutory research WGGiS’ AGH University of Science and Technology from the Department of the Protection of Mining Areas, Geoinformatics and Mining Surveying grant number 11.11.150.195 and by the Ministry of Science and Higher Education under grant number 15.11.150.390.

REFERENCES


DOI: 10.1016/S0148-9062(99)00022-4


DOI: 10.1007/s10596-009-9134-1


Khan, S.D., Huang, Z. and Karacay, A.: 2014, Study of ground subsidence in northwest Harris county using GPS, LiDAR, and InSAR techniques. Natural Hazards, 73, 1143–1173. DOI: 10.1007/s11069-014-1067-x


Polamin P.: 2015, Application of two parameter groups of the Knothe-Budryk theory in subsidence prediction.
Journal of Sustainable Mining, 14, 67–75. DOI: 10.1016/j.jsm.2015.08.010
Sikora, P.: 2016, Simulation of rock mass horizontal displacements with usage of cellular automata theory. Archives of Mining Sciences, 61, 4, 749–763. DOI: 10.1515/ams-2016-0051
Zhang, J.Z., Huang, H.J. and Bi, H.: 2015, Land subsidence in the modern Yellow River Delta based on InSAR time series analysis. Natural Hazards, 75, 3, 2385–2397. DOI: 10.1007/s11069-014-1434-7