Journal of Engineering and Technology for Industrial Applications

ITEGAM-JETIA

ISSN ONLINE: 2447-0228

Manaus, v.7 n.29, p. 13-24. May/Jun, 2021

DOI: https://doi.org/10.5935/jetia.v7i29.755

RESEARCH ARTICLE OPEN ACCESS

MINE SUBSIDENCE PREDICTION USING GENE EXPRESSION PROGRAMMING BASED ON MULTIVARIABLE SYMBOLIC REGRESSION

Hadi Rasouli[1], Kourosh Shahriar*[2] and Sayyed Hasan Madani[3]

[1, 2, 3] Department of Mining and Metallurgy Engineering, Amirkabir University of Technology, Tehran, Iran.

[1] http://orcid.org/0000-0002-1422-6225, [2] http://orcid.org/0000-0002-8561-6984, [3] http://orcid.org/0000-0002-6447-3646

Email: hadi.rasouli@aut.ac.ir, *k.shahriar@aut.ac.ir, hmadani@aut.ac.ir

ARTICLE INFO

Article History
Received: May 27th, 2021
Accepted: June 18th, 2021
Published: June 30th, 2021

Keywords:
Longwall mining,
Maximum subsidence,
Empirical model,
Gene expression programming.

ABSTRACT

Accurate prediction of surface subsidence becomes a significant challenge for active industrial companies in coal mining fields due to the importance of the economic impacts of longwall mining-induced subsidence. This article explores a new variant of genetic programming, namely gene expression programming (GEP). The GEP-based method is utilized to present a new mathematical formula for subsidence prediction in longwall coal mining. The derived model includes both geometrical and geological variables. The data set consists of field measurements obtained through 37 longwall panels of Ulan coal mine, NSW, Australia. The GEP-based model concluded satisfactory subsidence prediction outcomes compared to other empirical methods such as NCB, DMR, ACARP, and IPM. The predictive ability of the GEP-based models, which captures the complex nonlinear effects of the critical factors on the magnitude of subsidence, resulted in a statistically significant improvement in predictive capacity compared to the aforementioned empirical methods. The sensitivity analysis results indicated that Panel width and cover thickness with 31% and 23% were the most influential parameters in the proposed model. Also, the extracted seam thickness, thick layer location, and thick layer thickness had 19%, 16%, and 11% impact on the GEP proposed model, respectively.

I. INTRODUCTION

Longwall coal mining is the most common underground extraction method globally because of its relatively low cost, high safety, and efficiency in mining seams from depths. Longwall coal mining subsidence can affect groundwater resource and causes changes in permeability, porosity, and groundwater levels. Longwall-induced subsidence also has impacts on surface waters and associated ecosystems. Subsidence impacts can be divided into impacts on ecology, hydrology, geomorphology, and topography. Due to the importance of the mentioned effects, accurate prediction of surface subsidence due to longwall coal mining becomes a significant challenge in mining engineering. This importance includes environmental, economic, and social aspects.

Surface subsidence prediction methods, in general, include four categories: Empirical, numerical, hybrid, and physical methods. Empirical methods are based on the back analysis of field measurements and are the most common subsidence prediction methods. Some examples of these methods are the National coal Board (NCB) method, Australian Department Mineral Resource (DMR) method, Australian Coal Research Program (ACARP), influence profile method and Incremental Profile Method (IPM) [1-5].

Numerical methods use various mathematical functions to study ground movements in and around the longwall panels. Some examples of these methods include: cutting cantilever beam, key strata, and Voussoir beam structure methods [6-8].

Hybrid methods involve various mixtures of back-analysis of field observation data and using numerical and intelligent techniques. The fourth category is physical methods, which provide visual means but have little predictive value. In this study, Symbolic regression (SR)-based Gene Expression Programming (GEP) is used for subsidence prediction in longwall coal mines [9-10]. SR is a cluster of regression analysis methods to find and build the best model for accuracy and simplicity. This method requires many input and output data points to find an accurate regression. Similar to the genetic algorithm, this method includes both simple, linear chromosomes of fixed length. Like the parse trees of Genetic Programming (GP), GEP uses ramified structures of different sizes and shapes. The final output of the GEP model is a series of linear chromosomes of fixed length. In GEP, the genotype and phenotype are separated, and the model can benefit from all the evolutionary advantages. The main objective of this study is to develop a new mathematical model for predicting the maximum subsidence of longwall panels using the GEP method.

Among the various geometrical and geological factors that influence mine subsidence, five parameters include: panel width (W), extracted seam thickness (T), Cover thickness (H), the thickness of overburden thick layer (t), and the location of the massive unit above workings (Y) are used in the proposed model. Figure 1 shows the schematic of the longwall mining method [11].

The following sections of the paper are organized as follows: Section 2 gives brief literature surveys on the subsidence prediction methods. The most applicable empirical methods in Australia are discussed in more detail. In section 3, the method of research is explained. Discussion and results are provided in section 4. Finally, the conclusion is given in section 5.

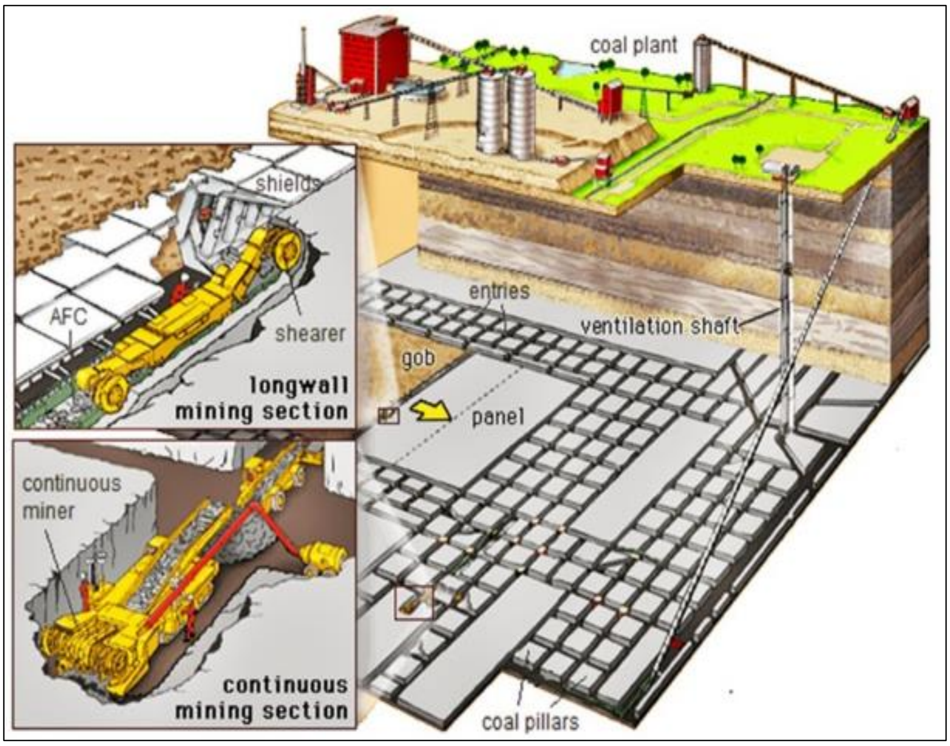


Figure 1: Schematic of longwall coal mining.
Source: [11].

II. LITURATURE SURVEY

The accuracy of mine subsidence prediction methods should never be taken for granted. The magnitude of subsidence depends on the input parameters of the specific site conditions. Empirical methods are based on the actual field measurements. These approaches predict subsidence based on parameter relationships developed from field monitoring and experience [1-4, 12-14]. The most widely used methods for predicting longwall mining-induced subsidence in Australia are described in detail.

National Coal Board in the UK proposed a subsidence prediction method during the 1960s [1]. NCB method results are based on the UK geology and do not predict subsidence magnitude accurately for other countries. During the 80s and 90s, the NCB method had been used in Australia. DMR, ACARP, and IPM methods have now replaced them. The NCB method gave good predictions when used in British mining conditions, but it provided much higher values than measured data in Australia. The difference in calculated subsidence magnitude is because the rock mechanics, geological conditions, and overlying strata of the extracted coal seam in British coalfields are different from those in Australia. The strata rocks in Britain are generally less strong and competent. Therefore, for a given seam thickness, the calculated maximum subsidence by the NCB method is greater than it would typically be for the Australian mining conditions. Figure 2 shows the NCB curves for subsidence prediction in different w/h ratios for caving and solid stowing cases.

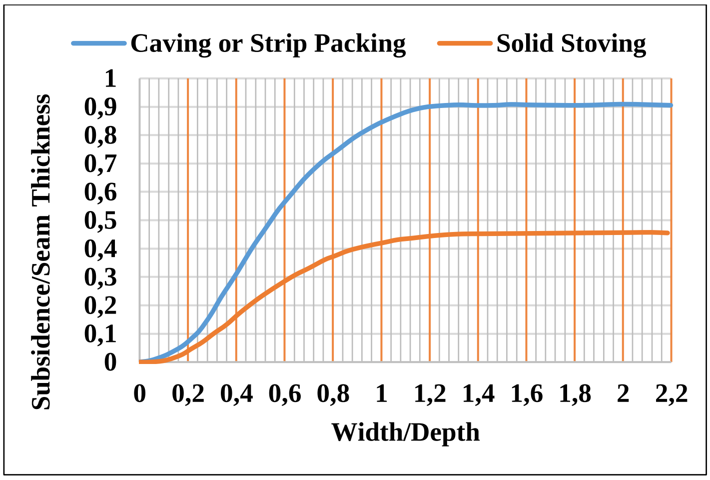


Figure 2: NCB subsidence prediction method.
Source: [1].

As a result of extensive research, NEW South Wales Department of Mineral Resources (DMR) introduced a new subsidence prediction model for Australian conditions [2]. DMR model is a modified form of the NCB method for subsidence prediction. The graphical charts of the DMR proposed model were presented in three handbooks for major coalfields of NSW. This method is also applicable for greenfield sites and where a worst-case scenario prediction is required. The data inputs to the DMR method are limited to the panel geometric variables include panel width, cover depth, and seam thickness. Like other empirical models, the DMR model is only suitable for subsidence prediction when geometrical variables are within the ranges that the model has been developed. In model development, The database from over fifty Newcastle and southern were used to model construction. As Figure 3 shows, this model is applicable when the panel (W/H) ratio is 0.2 to 2.0 and covers depth ranges between 70 m and 350 m. In recent years, the DMR method has mostly been superseded by the incremental profile method.

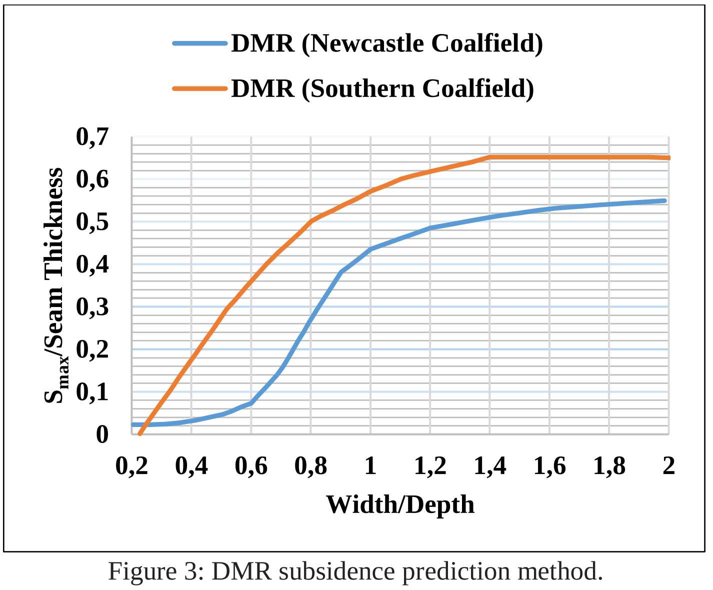


Figure 3: DMR subsidence prediction method.
Source: [2].

In 2003, the Australian Coal Research Program introduced a new model for subsidence prediction in Newcastle Coalfields [3]. ACARP model was developed based on the LAMODEL program and provided a reliable subsidence prediction model using both geometrical and geological information of longwall panels (See Figure 6).

The main focus was on the behavior of massive sandstone and conglomerate strata above the extracted coal seam. The massive geological units are classified into high, moderate, and low SRP. In the next step, according to obtained SRP factor and the thickness of the massive unit (Figure 4), maximum subsidence magnitude can be calculated from Figure 7. Upper and Lower bound prediction lines of this method for depths between 50m to 150 m is presented in Figure 7. For others depth diagrams are presented in [3]. Geometrically, the subsidence above a series of longwalls is strongly influenced by the panel width, the cover thickness, and the extraction height. Regarding geology, massive strata units above longwall panels result in reduced subsidence compared to longwall panels with similar geometry but thinner strata units.

$$S_{max} = \sqrt{12(1-\nu 2)/t)\ (\gamma H/E)\ (W2/4)} \tag{1}$$

Where are:

ν – Poisson ratio (dimensionless),

t – Overburden thickness (m),

γ – Unit weight (N/m^3),

H – Cover thickness (m),

E –Young Modulus (N/m^2),

W– Panel width (m).

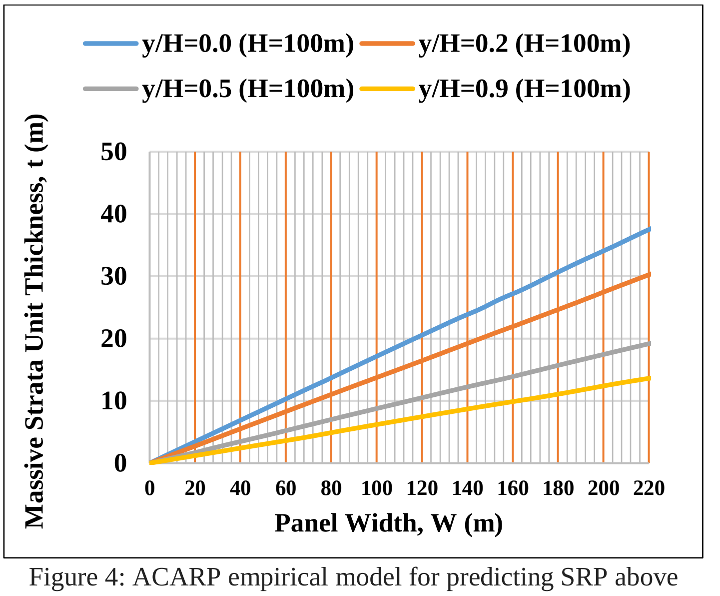


Figure 4: ACARP empirical model for predicting SRP above panels with cover thicknesses between 50 and 150m.
Source: [3].

IPM method was proposed to predict subsidence in the Newcastle Coalfield [4]. To predict subsidence with IPM requires panel width (W) and the thickness of cover (H). Figure 5 shows the details of subsidence prediction by the IPM method.

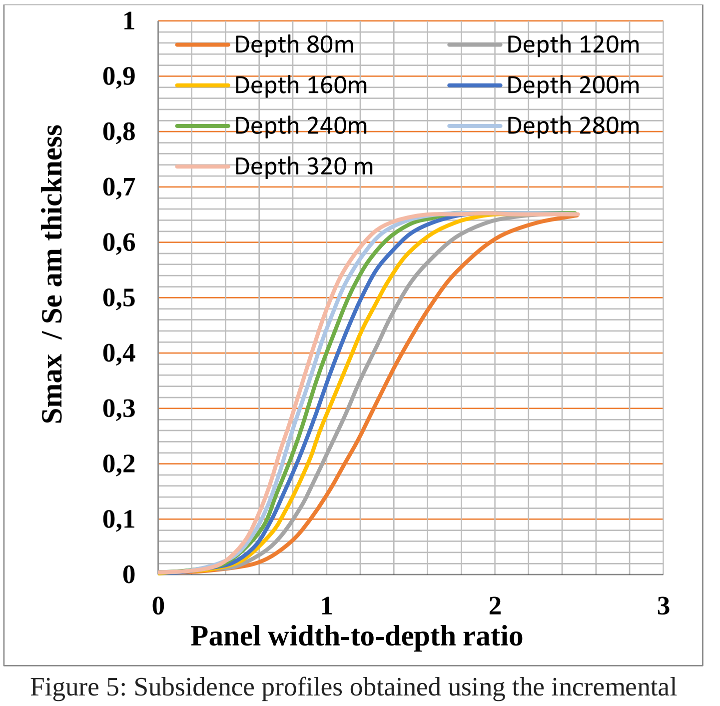


Figure 5: Subsidence profiles obtained using the incremental profile method.
Source: [4].

Empirical methods require extensive field measurement data to develop relationships applied to the mine subsidence prediction. However, when there are few case histories, an alternative approach is needed.

Numerical modeling methods simulate the geological and geotechnical conditions of the mining site to predict the

impact of various mining scenarios. Numerical modeling methods are often used when the mining method, strata conditions, and coal seam thickness differ from situations used in previously presented empirical methods. A conceptual model of the geotechnical conditions and the proposed mining methods must be developed before using numerical modeling. Computer modeling to simulate subsidence has been undertaken with various levels of success over the past years [15-17].

Developments in machine learning fields have created several new computer-aided data mining and hybrid approaches applicable for prediction problems. Artificial Neural Networks (ANN) have extensively been used to develop the nonlinear relationships between input parameters in mining and other geotechnical engineering systems [18-21]. A genetic algorithm (GA) is a robust stochastic approach for predicting various civil and mining problems. In contrast with ANNs and GA, the application of GP and its different variants, such as GEP in mining engineering, is entirely new and original. Various studies have also shown that GP, Linear Genetic programming (LGP), Multi-Expression Programming (MEP), and Gene Expression Programming (GEP) have advantages over ANNs in dealing with prediction problems [22-23].

The GEP and GP-based methods have useful applications where other standard modeling methods are complicated or detailed information for model construction don't exist [24-32].

III. METHODOLOGY

Genetic programming (GP) is one of the machine learning techniques that searches a program space. Gene Expression Programming (GEP) is an advanced form of GP technique. Ferreira presented gene expression programming [9, 10]. Traditional regression analyses, in some cases, have significant uncertainties. The regression analysis has a considerable impediment relating to complex processes. Besides that, the application of regression methods in model construction is the normality assumption of residuals. The capability of classical regression methods is also limited for the formulation of complexity, idealization of material behavior, and excessive empirical Parameters. The Gene ecpresion programming approach overcomes the constraints of various subsidence prediction methods that were previously presented. Contrary to many other soft computing tools, GEP provides prediction equations that can readily be used for subsidence prediction in longwall coal mines. The constitutive models derived using these methods can be incorporated into the different models. They may also be used to quickly check on other empirical models such as NCB, DMR, ACARP, and IPM. Similar to the genetic algorithm, this method includes both the simple, linear chromosomes of fixed length.

Like the parse trees of GP uses ramified structures of different sizes and shapes. The final output of the GEP model is a series of linear chromosomes of fixed length. In GEP, the genotype and phenotype are separated, and the model can benefit from all the evolutionary advantages. Ferreira book includes a basic algorithm of GEP and its implementation details. GEP method consists of five essential components: function set, terminal set, fitness function, control parameters, and termination points. GEP uses a fixed length of character strings to represent solutions to the problems, expressed as parse trees of different sizes and shapes. These trees are called GEP expression trees (ETs). One advantage of the GEP technique is that genetic diversity is highly simplified as genetic operators work at the chromosome level. Another advantage of GEP is its unique, multi-genic nature which allows the evolution of more complex programs composed of several subprograms. The fundamental of the GEP is schematically represented in Figure 8. The algorithm uses the following steps until a termination condition is achieved:

(1) Randomly generation of the fixed-length chromosome of each individual for the initial population;

(2) Chromosome expression as ET and fitness evaluation of individuals;

(3) Selection of the best individuals according to the fitness function;

(4) Repeating the previous stages to define several generations or until an acceptable solution is found. Standard fitness functions that are used in GEP model evaluation are as follows:

III.1 NUMBER OF HITS

When the precision is chosen for the evaluated models, the number of hits fitness function favors other fitness functions in evaluating the goodness of constructed model. The error can be either absolute or relative. The fitness F (ij) of an individual program i for fitness case j is evaluated by Equation 2 [9]:

$$\text{If } E_{(\mathrm{ij})} \leq p, \quad \text{then } F_{(\mathrm{ij})} = 1; \quad \text{else } f_{(\mathrm{ij})} = 0 \tag{2}$$

Where are:

E (ij) – is the error of an individual program i for fitness case j,

p –is the precision,

F (ij) – is the fitness of an individual program i for case j.

Equation 3 and Equation 4 show the error of an individual program in the cases in which errors are absolute and relatives, respectively [8]:

$$E_{(ij)} = \left|P_{(ij)} - T_j\right| \tag{3}$$

$$E_{(ij)} = \left|\frac{P_{(ij)} - T_j}{T_j} \cdot 100\right| \tag{4}$$

Where are:

$P_{(ij)}$ – Predicted value by the program i for the case j,

T_j– Target value for the case j.

$f_{\max} = n$, can occur when n is the number of fitness cases.

III.2 MEAN SQUARE ERROR (MSE)

The mean squared error expresses how close a regression line is to a set of points. It does this by taking the distances from the points to the regression line (these distances are the errors) and squaring them. The squaring is necessary to remove any negative signs. E (ij) is the error of an individual program i for fitness case j. Equation 5 and Equation 6 show the error of an individual program in instances in which errors are absolute and relatives:

$$E_i = \frac{1}{n}\sum_{j=1}^{n}(P_{(ij)} - T_j)^2 \quad (5)$$

Where are:
n – Number of fitness cases,
E_i – Absolute error for the program i,
$P_{(ij)}$ – Predicted value by the program i for the case j,
T_j– Target value for the case j.

$$E_i = \frac{1}{n}\sum_{j=1}^{n}\left(\frac{P_{(ij)} - T_j}{T_j}\right)^2 \quad (6)$$

Where are:
E_i– is the relative error for the program i.
Other parameters are previously stated in Equation 5.
Thus, for a perfect fitness, $P_{(ij)} = T_j$, and $E_{ij} = 0$.
To evaluate the fitness f_i of the program i, the following.

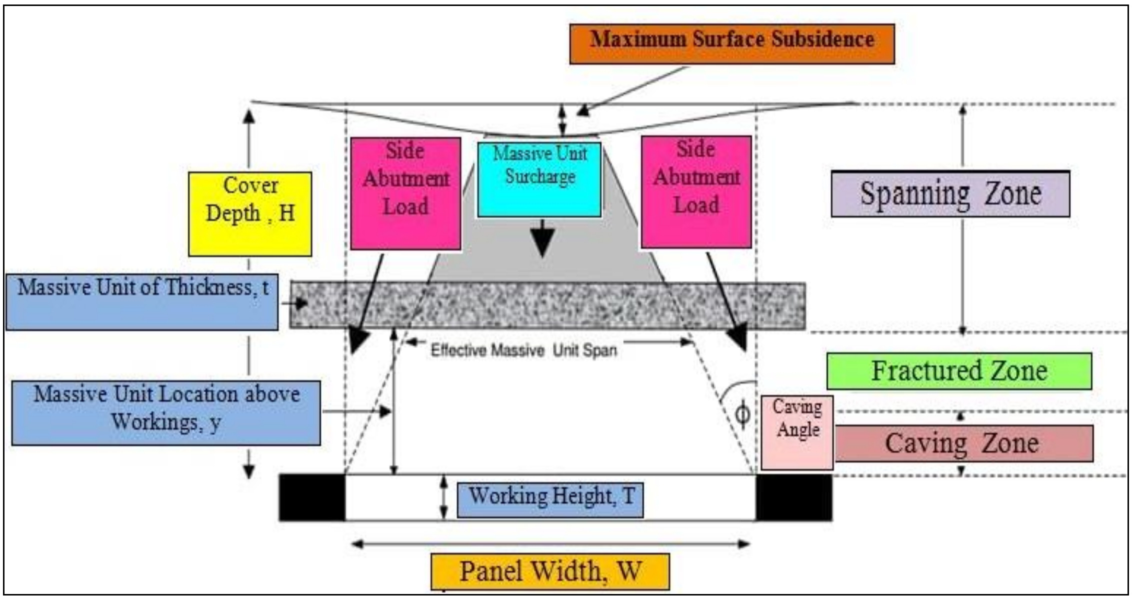


Figure 6: Key parameters of ACARP subsidence prediction method.
Source [3].

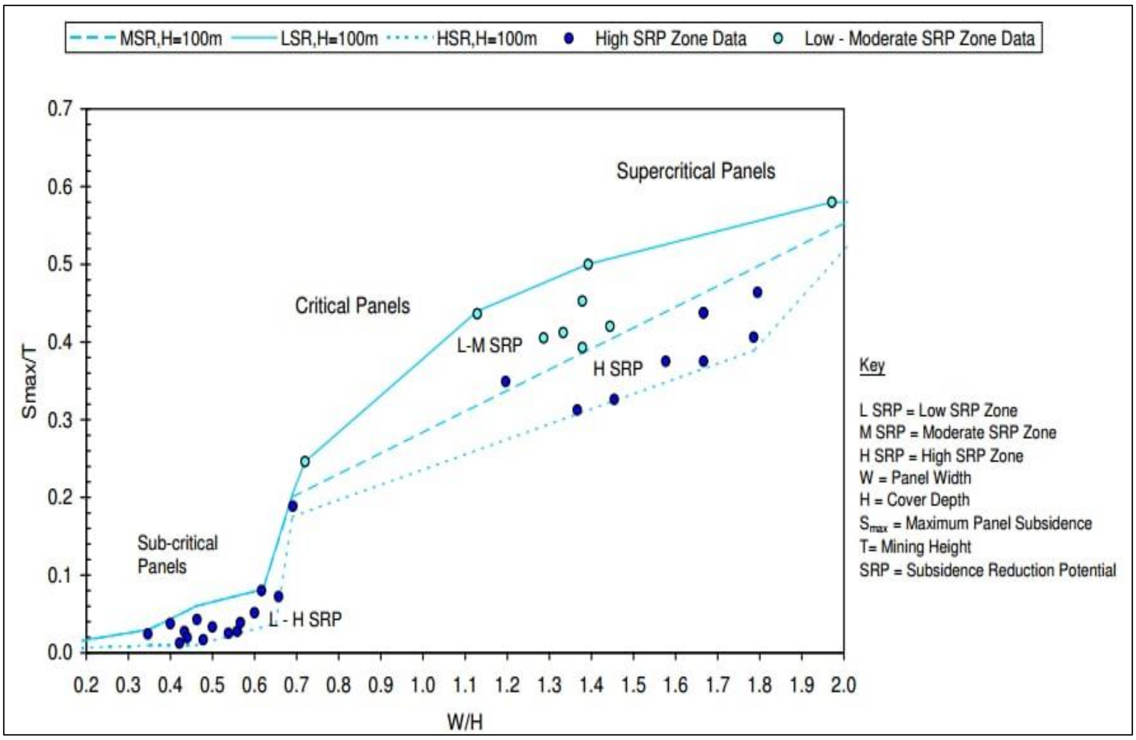


Figure 7: Empirical model for predicting subsidence above panels (cover thicknesses 50 m to 150 m and low to high SRP zones).
Source [3].

Equation can be used [9]:

$$f_i = 1000 \cdot \frac{1}{1 + E_i} \quad (7)$$

Where are:
f_i–Total fitness of the program i,
E_i– Total error of the program i.
f_i ranging from 0 to 1000, with 1000 corresponding to the ideal conditions.

III.3 R-SQUARE

R-square is the square of the Pearson product-moment correlation coefficient, which can be calculated as Equation 8 [9]:

$$R_i = \frac{n\sum_{j=1}^{n}(T_j P_{(ij)}) - (\sum_{j=1}^{n} T_j)(\sum_{j=1}^{n} P_{(ij)})}{\sqrt{\left[n\sum_{j=1}^{n} T_j^2 - (\sum_{j=1}^{n} T_i)^2\right]\left[n\sum_{j=1}^{n} P_{(ij)}^2 - (\sum_{j=1}^{n} P_{(ij)})^2\right]}} \quad (8)$$

Where are:
R_i– Pearson correlation coefficient.
Other parameters are stated in Equation 5.

The fitness of an individual program is a function of the correlation coefficient and is defined by the Equation 9 [9]:

$$f_i = 1000 \cdot R_i^2 \quad (9)$$

f_i ranges from 0 to 1000, with 1000 corresponding to the ideal fitness.
d. Precision and Selection Range.

The fitness f_i of program i is expressed by Equation 10 and Equation 11 for absolute and relative errors [9]:

$$f_i = \sum_{j=1}^{n}(R - |P_{(ij)} - T_j|) \quad (10)$$

$$f_i = \sum_{j=1}^{n}(R - |\frac{P_{(ij)} - T_j}{T_j} \cdot 100|) \quad (11)$$

Where are:
R – selection range
$P_{(ij)}$, and T_j are the parameters previously defined.
Thus, for a perfect fit, $P_{(ij)} = T_j$ for all cases *and* $f_i = R$.

In the second step, the terminals and functions are determined. The third step involves selecting chromosomal structures such as head and tail size and the number of genes. and finally, in the fourth step, the linking function is defined, and genetic operators are determined. Detailed descriptions of stages 2 to 4 are mentioned in [9].

IV. RESULTS AND DISCUSSION

Data from Ulan mines have been used in the modeling process [33-38]. Ulan coal mines are located in New South Wales, Australia, and are very similar in geological and geotechnical characteristics. The collected empirical data include 37 longwall panels and 119 measured subsidence from them. Most of the model data, about 79% of them were related to old Ulan mine. This section describes Ulan coal mines characteristics. Ulan Coal Mine Complex (UCMC) is situated in the central west of New South Wales. It is located near the village of Ulan, approximately 38 kilometers north-northeast of Mudgee and 19 kilometers northeast of Gulgong. Coal mining started in the Ulan area in the 1920s, consisting of Old Ulan, No.3 underground mines, Ulan West areas, and open-cut mining. Figure 9 shows the locality plan of the complex [38].

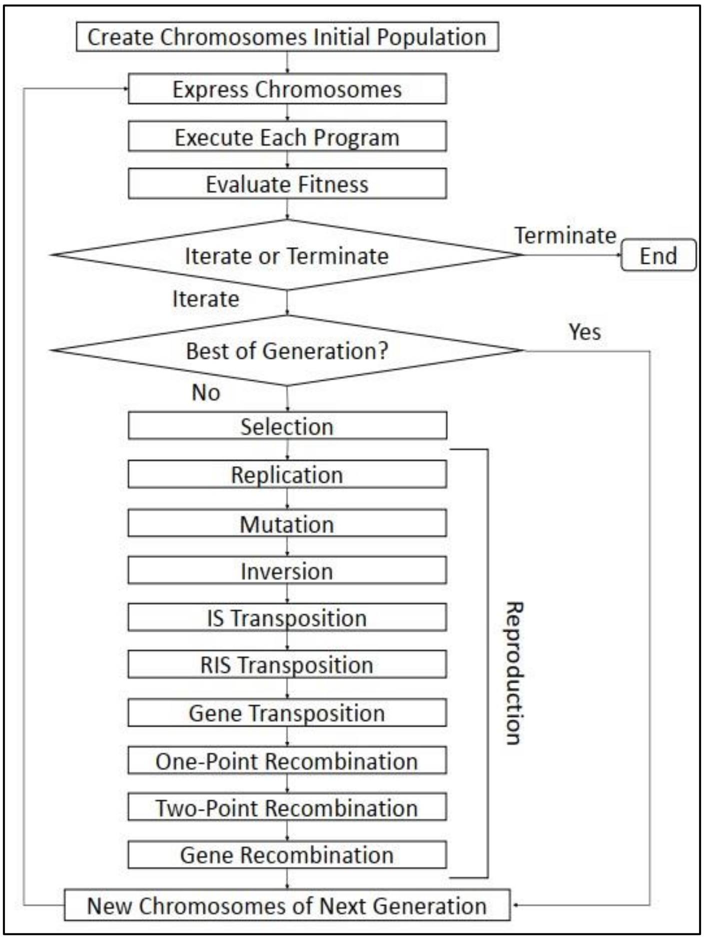


Figure 8: The flowchart of a gene expression algorithm.
Source: [9].

Figure 9: locality of the Ulan coal mines complex.
Source: [38].

The coal seams in the region range in thickness from 0.4 to 1o meters, and the Ulan coal seam is the thickest among them. Except for the Ulan, other seams are uneconomical due to high ash content. Massive Triassic sandstone, siltstone, and Narrabeen conglomerate overlie the Permian Coal Measures. Figure 10 shows Ulan coal mine longwalls and their positions [33].

The research aims to develop a new mathematical model for predicting the maximum subsidence of longwall panels using the GEP method. The GEP model uses the five influencing input parameters as Equation 12:

$$\mathrm{Smax} = \mathrm{F}\,(\mathrm{W}, \mathrm{T}, \mathrm{H}, \mathrm{t}, \mathrm{Y}) \qquad (12)$$

Where are:
W– Panel width (m),
T– Extracted seam thickness (m),
H – Cover thickness (m),
t – Thickness of overburden thick layer (m),
Y– Massive unit location above workings (m).

The fitness function for model development is the mean square error (MSE). The mean squared error Ei of an individual program i is evaluated using Equation 5 and Equation 6. For assessing the fitness fi of individual program i, Equation 6 is used. Various parameters used in the GEP model are shown in Table 1.

Table 1: Experiment parameters of the GEP model.

Parameters	Values
Fitness Function	Equation 6
Population Size	30
Number of Generation	1000
Head Length	5
Number of Genes	3
Chromosome Length	33
Function Set	+, -, /, *, tan, inv
Terminal Set	c0, …, c3
Link Function	+
Mutation Rate	0.004
Inversion Rate	0.01
IS Transposition Rate	0.01
RIS Transposition Rate	0.01
One-Point Recombination	0.3
Two-Point Recombination	0.3
Gene Recombination	0.1
Random Numbers	[-100,100]

Source: Authors, (2021).

Ulan longwall panels have all been carefully scheduled according to the timetable provided in Table 2.

Table 2: Extraction timetable of Ulan longwall panels.

Long wall	Start	Finish	Longwall	Start	Finish
LW1	07-12-86	30-11-87	W4	01-09-13	02-05-14
LW2	20-12-87	15-10-88	LW28	03-07-14	05-03-15
LW3	30-11-88	31-08-89	W5	06-05-15	01-03-16
LW4	05-12-89	15-09-90	LW29	02-05-16	01-11-16
LW5	15-10-90	04-01-92	W6	02-01-17	01-10-17
A	15-05-92	30-08-92	LW30	02-12-17	04-04-18
B	05-10-92	28-02-93	W7	05-06-18	07-04-19
LW6	15-03-93	30-07-93	LW31	08-06-19	01-02-20
LW7	07-10-93	30-05-94	W8	09-04-20	07-11-20
LW8	15-06-94	15-02-95	LW32	08-01-21	09-07-21
LW9	22-03-95	26-10-95	W9	09-09-21	10-03-22
LW10	05-12-95	23-08-96	W10	11-05-22	08-10-22
LW11	25-10-96	26-11-97	LW33	09-12-22	08-02-23
LW12	23-10-97	01-07-98	W11	11-04-23	11-05-23
LW13	29-07-98	21-04-99	Ulan West		
LW14	21-07-99	01-04-00	UW1	02-01-12	01-01-14
LW15	31-05-00	22-02-01	UW2	01-03-14	01-03-15
LW16	20-03-01	08-10-01	UW3	01-05-15	01-05-16
LW17	06-11-01	21-07-02	UW4	01-07-16	01-09-17
LW18a	26-07-02	23-02-03	UW5	01-11-17	01-11-18
LW19	11-04-03	03-11-03	UW6	10-01-19	12-05-20
LW20a	10-12-03	10-12-03	UW7	13-07-20	10-10-21
LW21	27-10-04	27-10-04	UW8	10-12-21	12-06-23
LW22	23-09-05	23-09-05	UW9	11-08-23	10-01-25
LW23	25-10-06	11-09-07	UW10	12-03-25	12-04-26
LW24	05-11-07	20-03-08	UW11	12-06-26	10-08-28
W1	26-05-08	12-02-09	North 1		
LW25	01-04-09	01-11-09	LW-C	01-04-14	08-09-14
W2	01-01-10	01-11-10	LW-D	01-04-11	08-09-11
LW26	01-11-11	01-0911	LW-E	01-04-12	08-09-12
W3	01-11-11	03-08-12	LW-F	01-04-13	08-09-13
LW27	01-11-12	01-07-13	LW-G	01-04-15	08-09-15

Source: Authors, (2021).

For example, Table 3 shows a part of the measured subsidence data of old Ulan coal mine panels.

Table 3: Measured subsidence of old Ulan coal mine panels.

LW Panel	Measured Subsidence (m)	LW Panel	Measured Subsidence (m)
A	1.2	11C	1.4
B	0.93	11X	1.4
1	1.5	12D	1.3
5	1.0	13D	1.3
6	0.13	14D	1.1
7	1.0	15D	0.96
8	1.0	16D	1.1
9	1.2	17D	1.2
10	1.3	18E	1.1
11	1.4	19E	1.2

Source: [39].

Table 4 expresses the Performance results of the proposed GEP model in its different stages. Table 6 suggests the final GEP model equations and sub-trees in details.

Table 4: Proposed GEP model performance in different stages.

Model	Experiment vs Prediction		
Stage	R	MAD	RMSE
Learning	0.904	0.191	0.202
Test	0.892	0.184	0.197
Validation	0.881	0.181	0.192

Source: Authors, (2021).

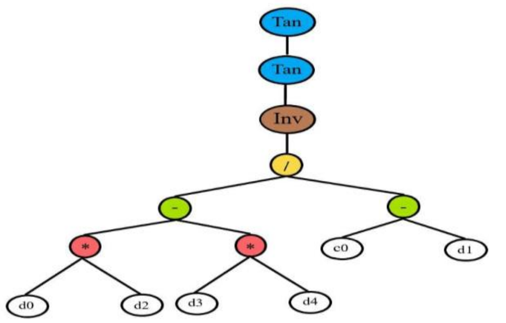
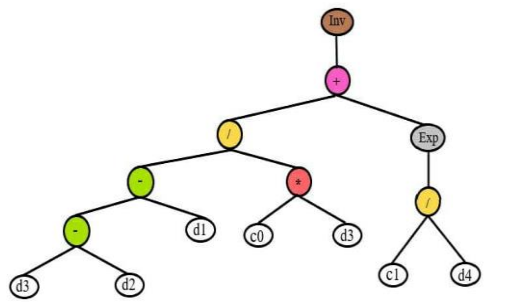
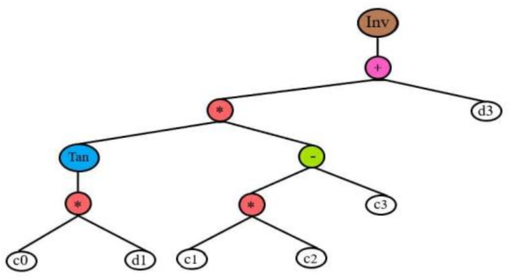
Table 5 and Figure 10 shows the importance of each parameter in the proposed GEP model. Panel width parameter with 31%, cover thickness with 23%, extracted seam thickness with 19%, thick layer location with 16%, and thick layer thickness with 11% impact the proposed model.

Table 5: Sensivity analysis results of the proposed GEP model.

Parameter	Unit	Symbol	Importance in the GEP model (%)
Panel width (W)	m	d0	31%
Cover thickness (H)	m	d1	23%
Extracted seam thickness (T)	m	d2	19%
Thickness of overburden thick layer (t)	m	d4	11%
Massive unit location above workings (Y)	m	d3	16%

Source: Authors, (2021).

Table 6: Sub-equations and sub-trees of the proposed GEP model results.

Equations	Expression Trees (Ets)
$$\mathrm{Tan}\left[\tan\left(\frac{c9-d1}{(d0*d2)-(d3*d4)}\right)\right] \quad (13)$$ Where are: d0 – is the Panel width (m), d1 – is the cover thickness (m), d2 – is the thickness of extracted seam (m), d3 – is the thickness of overburden thick layer (m), d4 – is the massive unit location above workings (m), c0 = 11.2418.	
$$\left[(d3-d2)-\frac{d1}{c6*d3}\right]+e^{\frac{c5}{d4}} \quad (14)$$ Where are: d1 – is the cover thickness (m), d2 – is the thickness of extracted seam (m), d3 – is the thickness of overburden thick layer (m), d4 – is the massive unit location above workings (m), c0 = -1.9457 , c1 = -8.8704.	
$$\mathrm{Tan}(c0*d1)]*[(c2*c4)-c1]+d3] \quad (16)$$ Where are: d1 – is the cover thickness (m), d3 – is the thickness of overburden thick layer (m), c0 = 144.932 , c1= 1.6201, c2 = 9.0104 , c3 = -3.7042.	

Source: Authors, (2021).

Table 7: Summary of the most applicable empirical methods for prediction of mine subsidence in Australia.

Method	Use in Australia	Main advantages	Main disadvantages
NCB	Less commonly used today and is used more as a rule of thumb to compare with other methods	Fast, low cost, requires fewer input parameters	The accuracy of the method when used in Australia is low
DMR	Widely used before 90s. It is used as an auxiliary method along with other subsidence prediction methods	Fast, low cost, requires fewer input parameters	Doesn't take into account the presence of the overlying thick strata
ACARP	Relatively common method	Takes into account the presence of the overlying thick strata	Its application limited to Newcastle, New South Wales
IPM	The most common subsidence prediction method in Australia	It can be implemented quickly and cheaply relative to numerical modeling methods	It requires more input data than NCB and DMR methods

Source: Authors, (2021).

Table 8: Subsidence prediction results in the most applicable empirical methods for prediction of mine subsidence in Australia.

Mine	W	H	T	t	Y	NCB	DMR	ACARP	IPM	GEP	Measured Subsidence
Ulan West	261	75	3.2	10	15	2.88	1.68	1.72	1.68	1.72	1.91
Ulan West	261	75	3.2	15	20	2.88	1.68	1.72	1.68	1.91	1.82
Ulan West	315	140	3.2	20	65	2.88	1.68	1.22	1.42	1.48	1.36
Ulan West	315	215	3.2	35	30	2.88	1.6	1.47	1.42	1.13	0.97
Old Ulan	216	160	2.54	15	13	2.27	1.19	1.17	1.29	1.55	1.37
Old Ulan	216	162	2.52	18	19	2.27	1.19	1.23	1.27	1.49	1.36
Old Ulan	216	167	2.54	20	23	2.29	1.17	1.27	1.25	1.44	1.35
Old Ulan	198	169	2.54	30	55	2.23	1.14	1.07	1.14	0.84	0.64

Source: Authors, (2021).

For validation purposes, the performance of the proposed GEP model and other maximum subsidence prediction methods reviewed in this research are compared with measured subsidence. The DMR, ACARP, and IPM methods had been proposed based on the database from NSW. Ulan coal mine complex has the same geological and geotechnical similarities with those of mentioned methods. Although the NCB method is proposed for UK mines, it has been used as a rule of thumb in comparing results of other empirical methods.

Table 7 provides a summary of the main empirical and proposed GEP models. Validation results of the proposed GEP model vs. other methods are provided in Table 8.

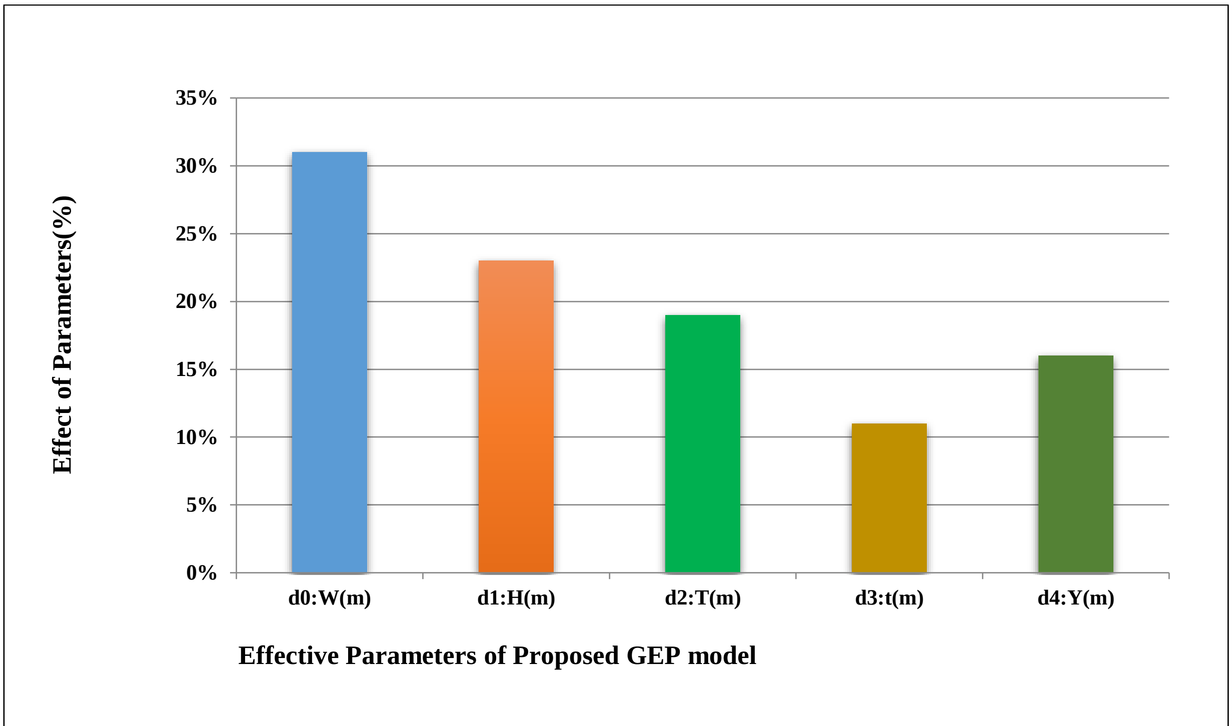


Figure 10: Sensitivity analysis results of the proposed GEP model.
Source: Authors, (2021).

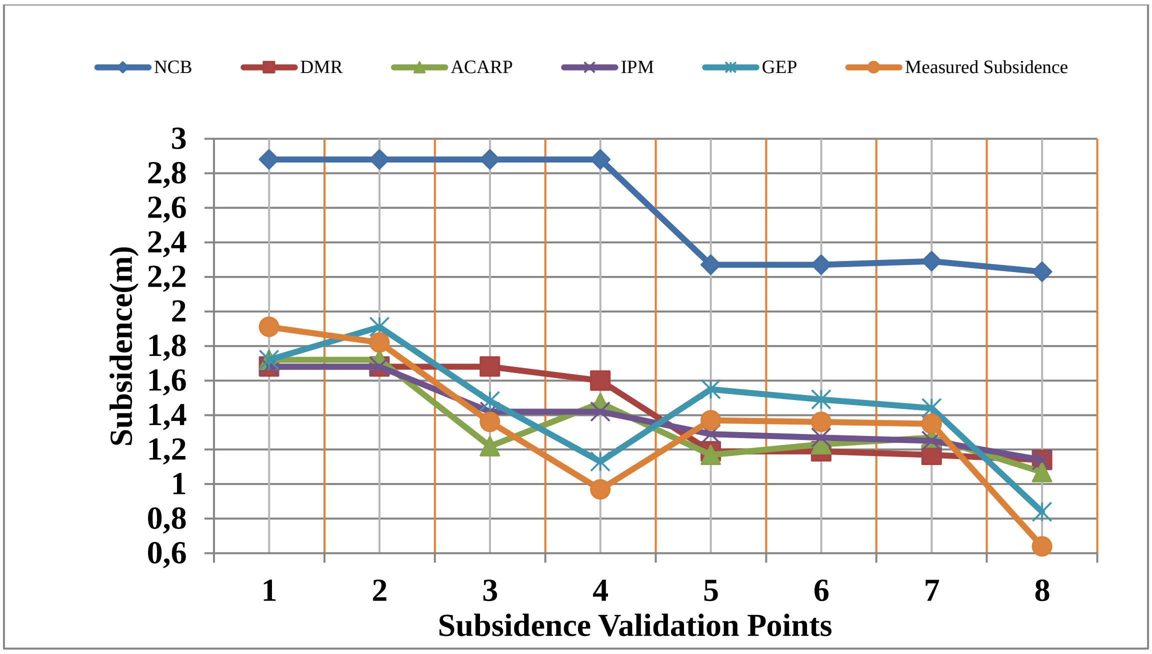


Figure 11: Prediction methods results vs. measured data.
Source: Authors, (2021).

As shown in Figure 11, the NCB method is based on the UK longwall mines. comparing the results of NCB method with the measured data is unsatisfactory. Compared to the previous methods presented with the measured data, the IPM has a minor error and higher accuracy. The IPM method has been proposed by Mine Subsidence Engineers Consultant into proprietary software and is available at a cost. Compared with IPM, the proposed GEP method has higher accuracy and requires less time and cost. Figure 12 and Table 9 suggest the statistics between measured subsidence and subsidence resulted from various prediction methods.

Table 9: Statistics results of the proposed GEP model vs. other methods.

Prediction method	Experiment vs Prediction			
	R	R^2	MAD	RMSE
NCB	0.46	0.21	1.26	1.28
DMR	0.72	0.52	0.26	0.34
ACARP	0.74	0.54	0.22	0.27
IPM	0.88	0.76	0.19	0.24
GEP	0.88	0.77	0.18	0.19

Source: Authors, (2021).

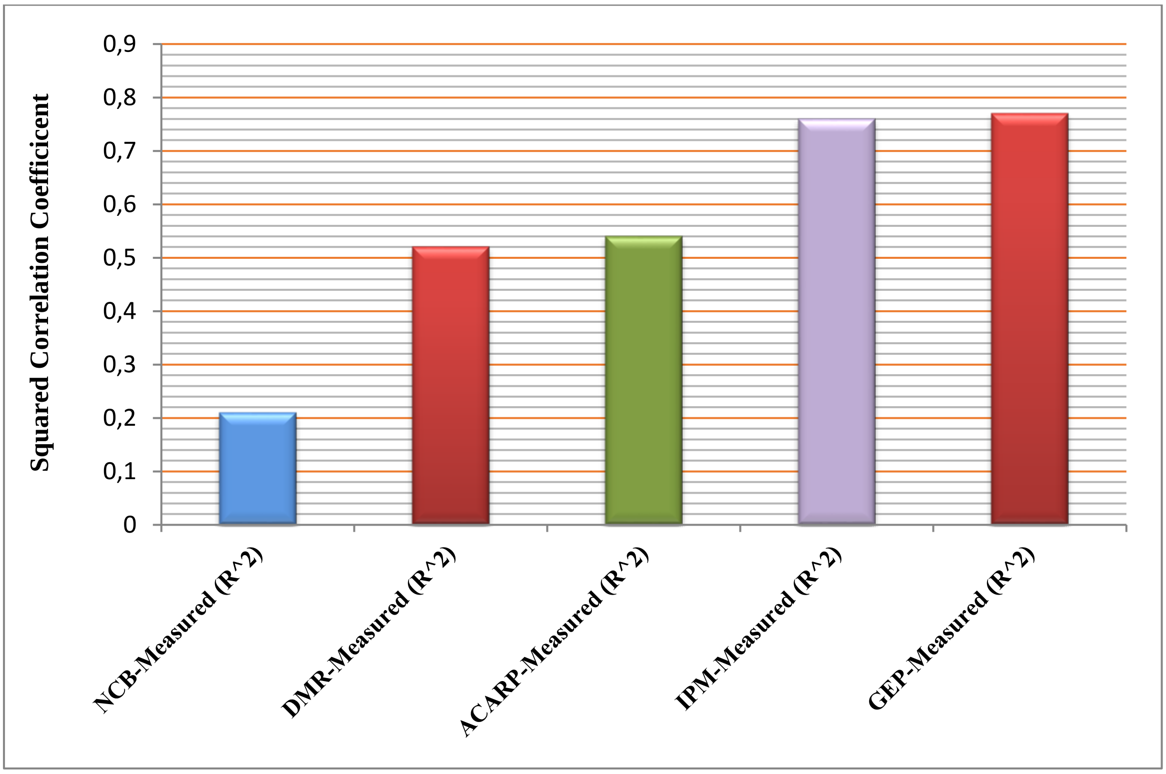


Figure 12: Squared correlation coefficients of measured subsidence and other prediction methods.
Source: Authors, (2021).

V. CONCLUSION

The present research introduced a new gene expression programming model based on multivariable symbolic regression for subsidence prediction in Ulan longwall mines. GEP can use linear, nonlinear functions and constant numbers without prior information about the final model. The proposed model was carried out through measured data as training inputs. Finally, a new mathematical formula for subsidence prediction was proposed. The maximum vertical subsidence was modeled in various affecting parameters (W, H, T, t, y). The proposed model was constructed using measured subsidence data from Ulan coal mines. By comparing the results of the proposed GEP model with NCB, DMR, ACARP, and IPM methods, it was observed that the GEP-based model was accurate enough and had the potential to be used in mines with similar geological and geotechnical conditions. The sensitivity analysis results indicated that Panel width with a 31% effect was the most influential parameter in the proposed model. Also, the impact of panel depth, extracted seam thickness, location, and thickness of the thick overburden layer were 23%, 19%, 16%, and 11%, respectively. The proposed model is expected to help predict subsidence where geological and geotechnical conditions are similar to that of the Ulan coal mines. Future research can include some new parameters individually or mixed to present new optimized subsidence prediction models.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Hadi Rasouli and Kourosh Shahriar.
Methodology: Hadi Rasouli, Kourosh Shahriar, and Sayyed Hasan Madani.
Investigation: Hadi Rasouli and Sayyed Hasan Madani.
Discussion of results: Hadi Rasouli and Kourosh Shahriar.
Writing – Original Draft: Hadi Rasouli and Sayyed Hasan Madani.
Writing – Review and Editing: Kourosh Shahriar.
Resources: Kourosh Shahriar and Sayyed Hasan Madani.
Supervision: Koroush Shahriar and Sayyed Hasan Madani.
Approval of the final text: Hadi Rasouli, Kourosh Shahriar, and Sayyed Hasan Madani.

VII. REFERENCES

[1] National Coal Board, "Subsidence engineers handbook", 2nd edition, Mining Department, London, UK, pp. 04-10, 1975.

[2] L. Holla, "Surface Subsidence Prediction in the Newcastle Coalfield", Department of Mineral Resources, NSW, Australia, pp. 13-23, 1987.

[3] MAN-001/1, "Summary of the Modified ACARP 2003 Empirical Model", Ditton Geotechnical Services, NSW, Australia, https://majorprojects.planningportal.nsw.gov.au, pp. 1-152, 2013.

[4] MSEC 309, "The Prediction of Subsidence Parameters of Proposed Austar Longwalls A6 to A17", Mine Subsidence Engineering Consultants, NSW, Australia, https://majorprojects.planningportal.nsw.gov.au, pp. 21-27, 2008.

[5] M. Hood, R.T. Evy, and L.R. Riddle, "Empirical methods of subsidence prediction, a case study from Illinois", International Journal of Rock Mechanics, 20 (4), pp. 153-170, 1983.

[6] C. He, J. Xu, "Subsidence prediction of overburden strata and surface based on the Voussoir beam structure", Advance in Civil Engineering, 1, pp. 1-14, 2018.

[7] Z. Li, J. Xu, "The effects of the rotational speed of voussoir beam structures formed by key strata on the ground pressure of stopes", International Journal of Rock Mechanics, 108, pp. 67-79,2018.

[8] C. Liu, "Voussoir beam model for lower strong roof strata movement in longwall mining", J. Rock Mech. Geotech. Eng. 9 (6), pp. 1171–1176, 2018.

[9] C. Ferreira, "Gene expression programming, Mathematical modelling by an artificial intelligence", vol.21, Springer, United Kingdom, pp. 1-492, 2006.

[10] C. Ferreira, "Gene expression programming: a new adaptive algorithm for solving problems", Complex Systems, 13(2), pp. 87–129, 2001.

[11] https://www.britannica.com/technology/coal-mining, Last accessed 03/20/2017.

[12] J. Yue, W. Cheng, and L. Fan, "The Study on mathematical model of urban land subsidence based on statistical analysis", International Conference on Management and Service Science, Wuhan, China, pp. 20–22, September 2009.

[13] G. Yu, W. Mi, D. Wang, L. Gao, and S. Lu, "Research on the relationship between the surface dynamic subsidence and overburden separated strata of coal mine and its model", Procedia Eng,, 191, pp. 196–205,2017.

[14] B. Wang, J. Xu, and D. Xuan, "Time function model of dynamic surface subsidence assessment of grout-injected overburden of a coal mine", Int. J. Rock Mech. Min. Sci, 104, pp. 1–8,2018.

[15] H. Li, J. Zha, G. Guo, "A new dynamic prediction method for surface subsidence based on numerical model parameter sensitivity", J. Clean Prod, 233, pp. 1418–1424, 2019.

[16] P. Tzampoglou, C. Loupasakis, " Numerical simulation of the factors causing land subsidence due to overexploitation of the aquifer in the Amyntaio open coal mine", Greece, Hydro Research, 1, pp. 8 –24, 2019.

[17] A. Zingano, A. Weiss, " Subsidence over room and pillar retreat mining in a low coal seam", Int. J. Min. Sci. Technol, 29, pp. 51–57, 2019.

[18] Y. Cha, W. Choi, and O. Buyukozturk, "Deep learning-based crack damage detection using convolutional neural networks", Comput. Aided Civ. Infrastruct. Eng, 32, pp. 361–378, 2017.

[19] M.E. Yetkin, M.K. Ozfirat, "Selection of thick coal seam mining method using analytic hierarchy process", JETIA, vol. 5, no. 20, pp. 06-11, Dec. 2019.

[20] M. Juszczyk, A. Le'sniak, "Modelling construction site cost index based on neural network ensembles", Symmetry, 11, pp. 411, 2019.

[21] V. Singh, S. Bano, A.K. Yadav, and S. Ahmad, "Feasibility of artificial neural network in civil engineering", Int. J. Trend Sci. Res. Dev, 3, pp. 724–728, 2019.

[22] B.S. Narendra, P.V. Sivapullaiah, S. Suresh, and S.N. Omkar, "Prediction of unconfined compressive strength of soft grounds using computational intelligence techniques: a comparative study", Computers and Geotechnics, Vol. 33, pp. 196-208, 2006.

[23] C. Kayadelen, O. Naydın, M. Fener, A. Demir, and A. Ozvan,"Modeling of the angle of shearing resistance of soils using soft computing systems", Expert Systems with Applications, Vol. 36, pp. 18- 26, 2009.

[24] M.N. Amar, "Prediction of hydrate formation temperature using gene expression programming", Journal of Natural Gas Science and Engineering, 87, pp. 24-36, 2021.

[25] I. Azim, J. Yang, M.F. Iqbal, M.F., Z. Mahmood, F. Wang, and Q.F. Liu, "Prediction model for compressive arch action capacity of RC frame structures under column removal scenario using gene expression programming", Structures, 25, pp. 212-228, 2020.

[26] A. Fathi, M. Mazari, and M. Saghafi, "Multivariate global sensitivity analysis of rocking responses of shallow foundations under controlled rocking", Eighth International Conference on case histories in geotechnical engineering, Geo- Congress, GSP 307, pp. 490-498, 2019.

[27] T. Hong, K. Jong, and Ch. Koo, "An optimized gene expression programming model for forecasting the national CO2 emissions in 2030 using the metaheuristic algorithms", Applied Energy, 228, pp. 208-220, 2018.

[28] A.I. Lawal, S. Kwon, O.S. Hammed, and M.A. Idris, "Blast-induced ground vibration prediction in granite quarries: An application of gene expression

programming, ANFIS, and sine cosine algorithm optimized ANN", International Journal of Mining Science and Technology, 31, pp. 265-277, 2021.

[29] H. Majidfar, B. Jahangiri, W.G. Buttlar, and A.H. Alavi, "New machine learning-based prediction models for fracture energy of asphalt mixtures", Measurement 135, pp. 438–451, 2019.

[30] S.M. Mousavi, P. Aminian, A.H. Gandomi, A.H. Alavi, and H. Bolandi, "A new predictive model for compressive strength of HPC using gene expression programming", Advances in Engineering Software, 45, pp. 105-114, 2012.

[31] S. Moyano, O. Reyes, H. Fardun, and S. Ventura, "Performing multi-target regression via gene expression programming-based ensemble models", Neurocomputing, 432, pp. 275-287, 2021.

[32] Y.Z. Murad, R. Hunifat, and W. Al-Bodour, (2020): Interior Reinforced Concrete Beam-to-Column Joints Subjected to Cyclic Loading: Shear Strength Prediction using Gene Expression Programming. Case Studies in Construction Materials, 13, pp. 1-9, 2020.

[33] ULA 3367, "Subsidence assessment of Ulan coal mine", Strata Control Technology, NSW, Australia, https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal/reporting-documents, pp. 1-119, 2009.

[34] ULA 3370, "Subsidence assessment, North 1 underground mining area", Strata Control Technology, NSW, Australia, https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal/reporting-documents, pp. 3-19, 2009.

[35] ULA, "Analysis of subsidence results from longwall W2, comparison with predictions", Strata Control Technology, NSW, Australia, https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal/reporting-documents, pp. 1-92, 2009.

[36] ULA 4510, "2015 annual review of subsidence monitoring at Ulan west mine", Strata Control Technology, NSW, Australia, https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal/reporting-documents, pp. 1-33, 2009.

[37] ULN SD PLN 0103, "Ulan West Extraction Plan for longwalls LW1 to LW4", Ulan Coal GLENCORE, NSW, Australia, https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal/reporting-documents, pp. 11-29, 2015.

[38] https://www.glencore.com.au/operations-and-projects/coal/current-operations/ulan-coal, NSW, Australia, Last accessed 07/14/2019.

[39] WHT/1, Mine subsidence impact assessment for the proposed longwall panels LWs 1 to 14, No.4 underground area, Moolarben coal project, NSW, Australia, http://www.moolarbencoal.com.au/icms-docs/259060-response-to-submissions-appendices-8-9.pdf, pp. 30-47, September 2006.