



Data Article

Data on methane concentration collected by underground coal mine sensors

Michał Kozielski*, Marek Sikora, Łukasz Wróbel

Department of Computer Networks and Systems, Silesian University of Technology, Akademicka 16, Gliwice 44-100, Poland

ARTICLE INFO

Article history:

Received 7 May 2021

Revised 29 September 2021

Accepted 4 October 2021

Available online 11 October 2021

Keywords:

Sensory measurements

Coal mine

Methane concentration

Prediction

ABSTRACT

Coal mining requires working in hazardous conditions. Miners in an underground coal mine can face several threats, such as, e.g. methane explosions. To provide protection for people working underground, systems for active monitoring of production processes are typically used. One of their fundamental applications is screening dangerous gas concentrations (methane in particular) to prevent spontaneous explosions. Such a system is the source of the data set containing raw data collected at an underground coal mine. The data is collected from 28 different sensors placed at various locations around the coal mine. All the attributes except one are numeric, and the examples collected form a time series. This data set can be used in a variety of analytical tasks, including classification, regression, time series and stream data analysis.

© 2021 The Author(s). Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

* Corresponding author.

E-mail address: michal.kozielski@polsl.pl (M. Kozielski).

Specifications Table

Subject	Applied Machine Learning, Mining Engineering
Specific subject area	Coal mine, Methane concentration, Sensory data, Time series
Type of data	Table
How data were acquired	Measurements were acquired by several sensors: methane meter, anemometer, barometer, humidity sensor, temperature sensor, pressure sensor, current meter, velocity meter, driving direction indicator. Data from sensors measuring environmental parameters were collected by the SMP-NT system. Data from sensors measuring the parameters of the longwall shearer were collected by the MAK5 DBC system. The data was integrated by the THOR dispatching system and its data base was the final source of this data set.
Data format	Raw Integrated Synchronised
Parameters for data collection	The data was recorded by two groups of sensors - the first describes the climatic conditions in a selected part of the mine (methane meter, anemometer, barometer, humidity sensor, temperature sensor, pressure sensor), the second describes the activity of a longwall shearer (current meter, velocity meter, driving direction indicator). The collected data was synchronised and shows values every 1 s.
Description of data collection	The data was collected at an underground coal mine, and it consists of two sets of characteristics that were combined. The first one is formed of typical data collected and visualised by a monitoring system that controls the environmental parameters underground. The latter set consists of the values characterising the cutter loader operation. The set of sensors is typical and representative for monitoring underground workings. The values of each type of sensors were collected by a dedicated system installed in the mine. The merged data set was downloaded from the dispatching system.
Data source location	Institution: a coal mine in the Upper Silesian coal basin (co-ordinates 50.066, 18.438) City/Town/Region: Silesian Voivodeship Country: Poland
Data accessibility	Sikora, Marek; Wróbel, Łukasz (2021), "Methane", Mendeley Data, V1, https://doi.org/10.17632/yd7vw4c5mk.1
Related research article	D. Ślęzak, M. Grzegorowski, A. Janusz, M. Kozielski, S.H. Nguyen, M. Sikora, S. Stawicki, Ł. Wróbel, A framework for learning and embedding multi-sensor forecasting models into a decision support system: A case study of methane concentration in coal mines. Information Sciences, 451 (2018) 112-133. https://doi.org/10.1016/j.ins.2018.04.026

Value of the Data

- This data set is useful because it consists of real-life industrial sensory data. The data can be used to generate models predicting the increased concentration of methane in the coal mine, which is a threat to miners and causes an emergency stoppage of production. Therefore, the data is important because it relates to critical events on which human life may depend and additionally, it relates to predictive maintenance of production task.
- This data has a form of multidimensional time series, therefore, it may be of benefit to data scientists and researchers in general who are interested in time series analysis. After processing the data so that the labels indicate that the accepted safety threshold for methane concentration was exceeded, the data may also be of interest to data scientists dealing with unbalanced data. Moreover, the dataset may be of interest to researchers in the field of Internet of Things and data stream analysis, as it contains frequent measurements from a network of distributed sensors (the network is wired - intrinsically safe). Finally, researchers involved in mining, safety engineering and production maintenance can benefit from this dataset as predicting methane concentration can help prevent emergency electrical disconnection in a mining area.

- This data set can be used in a variety of analytical tasks. It can be used in classification task, where the warning methane concentrations are predicted. It can be used in regression task, where future value of methane concentration is predicted. Finally, it can be used to verify various methods of processing time series and stream data.
- This data set is interesting because it shows a fusion of data from which information fusion can be made – sensors are of different types, they collect underground environmental data, methane concentration and additionally production intensity characteristics.

1. Data Description

The data set contains raw data collected in an underground coal mine between March 2, 2014 and June 16, 2014 [1]. It consists of 9,199,930 data examples, where each example consists of a time stamp, which consists of 6 attributes (year, month, day, hour, minute, second) and measurements collected from 28 different sensors which are given at one second intervals. The data set is contained in a CSV file and is accompanied by a file containing a description of the attributes and a file containing a plot showing the minimum, maximum and average values of measurements in a 1 h sliding window.

The sensors recording the measurements are placed at various locations around the coal mine. The map of the coal mine area of interest is presented in Fig. 1.

The sensors, their type, kind of action they trigger and the warning and alarm thresholds are listed in Tables 1–3. The sensors correspond to the attributes of the data set. The characteristics of the numerical attributes are presented in Table 4. The values of the only non-numeric F_SIDE attribute are explained in Table 3.

The data set does not contain any missing values. The data is real-life and therefore it can contain outliers. Moreover, some values are outside the allowed measuring ranges indicated in Tables 1–3.

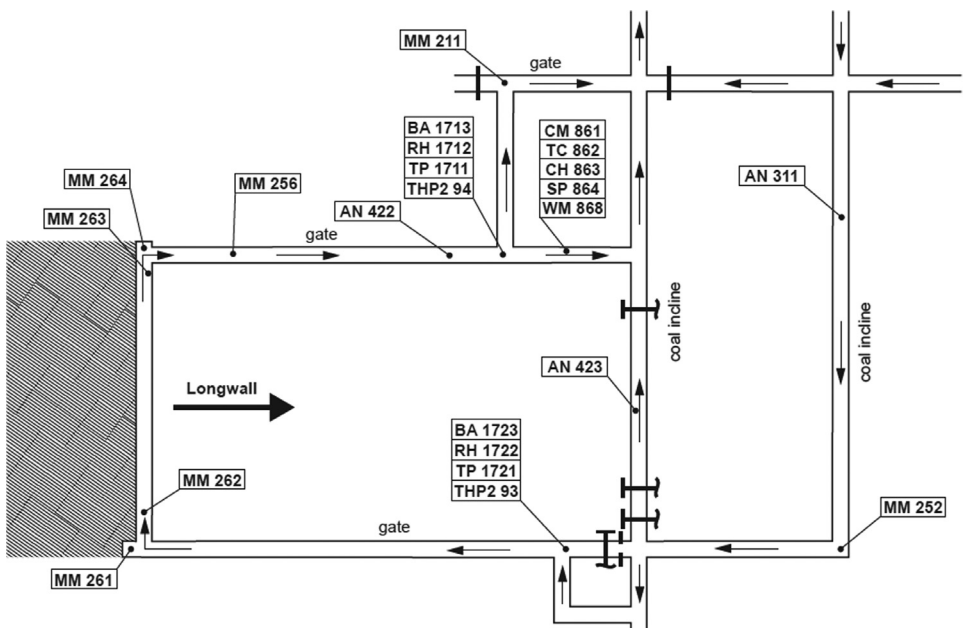


Fig. 1. Coal mine area map containing sensor locations.

Table 1
Characteristics of sensors collecting data-part 1.

AN311	anemometer (distant) [m/s] sensor type: anemometer [−5, 5] kind: alarming
AN422	anemometer [m/s] sensor type: anemometer [−5, 5] kind: switching off
AN423	anemometer [m/s] sensor type: anemometer [−5, 5] kind: switching off
TP1721	temperature [°C] sensor type: temperature THP (three-component sensor THP2/93) kind: registering
RH1722	humidity [%RH] sensor type: humidity THP (three-component sensor THP2/93) kind: registering
BA1723	barometer [hPa] sensor type: barometer THP (three-component sensor THP2/93) kind: registering
TP1711	temperature [°C] sensor type: temperature THP (three-component sensor THP2/94) kind: registering
RH1712	humidity [%RH] sensor type: humidity THP (three-component sensor THP2/94) kind: registering
BA1713	barometer [hPa] sensor type: barometer THP (three-component sensor THP2/94) kind: registering

Table 2
Characteristics of sensors collecting data-part 2.

MM252	methane meter (distant) [%CH4] sensor type: methane meter MM-2PWk kind: switching off value of threshold A (alarm): 2.0%
MM261	value of threshold W (warning): 1.5% methane meter [%CH4] sensor type: methane meter MM-2PWk kind: switching off value of threshold A: 1.5%
MM262	value of threshold W: 1.0% methane meter [%CH4] sensor type: methane meter MM-2PWk kind: switching off value of threshold A: 1.0%
MM263	value of threshold W: 0.6% methane meter [%CH4] sensor type: methane meter MM-2PWk kind: switching off value of threshold A: 1.5%
MM264	value of threshold W: 1.0% methane meter [%CH4] sensor type: methane meter MM-2PWk kind: switching off value of threshold A: 1.5%
MM256	value of threshold W: 1.0% methane meter [%CH4] sensor type: methane meter MM-2PWk

(continued on next page)

Table 2 (continued)

MM211	kind: switching off value of threshold A: 1.5% value of threshold W: 1.0% methane meter [%CH4] sensor type: methane meter MM-2PWk
CM861	kind: switching off value of threshold A: 2.0% value of threshold W: 1.5% high concentration methane meter [%CH4] sensor type: methane meter [0, 100] kind: registering

Table 3
Characteristics of sensors collecting data-part 3.

CR863	sensor for pressure difference on the methane drainage flange [Pa] sensor type: pressure difference [0, 250] kind: registering
P_864	pressure inside the methane drainage pipeline [kPa] sensor type: pressure [0, 110] kind: registering
TC862	temperature inside the pipeline [°C] sensor type: temperature [10, 40] kind: registering
WM868	methane delivery calculated according to CM, CR, P, TC [m³/min] sensor type: methane delivery [0, 50] kind: registering
AMP1_IR	current of the left cutting head of the cutter loader [A]
AMP2_IR	current of the right cutting head of the cutter loader [A]
DMP3_IR	current of the left haulage in the cutter loader [A]
DMP4_IR	current of the right haulage in the cutter loader [A]
AMP5_IR	current of the hydraulic pump engine in the cutter loader [A]
F_SIDE	driving direction, 1=left, {0, 0.5}=right
V	cutter loader speed [Hz] Vmin=3Hz, Vmax=100Hz [Hz] values are transformed into [m/min] 100 Hz equal to about 20 m/min

Among the sensors collecting information about the environmental parameters underground, three methane sensors can be distinguished: MM263, MM264 and MM256. These sensors, as it is shown in Fig. 1, are located in the most exposed area of the monitored longwall, where the methane released from the longwall accumulates. Therefore, these sensors can indicate the highest concentration of methane released during mining activity, and thus they are of critical importance from the point of view of the explosion hazard.

The data set in a processed form was used in a data analysis competition [2]. In this case data was transformed into 51,700 time periods of 10 min each. Therefore, each example of this transformed data set consisted of 16,800 values, resulting from 600 measurements (10 min of measurements taken each second) collected by each of the 28 sensors. The time periods represented by each data example overlapped and were given in a chronological order.

The task of the competition was to predict a dangerous concentration of methane at a key location in the mine. The labels in the data indicated whether the warning threshold was reached in a period of the next three to six minutes, for three methane meters: MM263, MM264 and MM256. In particular, if a given example corresponded to a period between t-599 and t0, then the label for a methane meter MM in this row was set to warning if and only if $\max(\text{MM}(t181), \dots, \text{MM}(t360)) \geq 1.0$.

Table 4
Characteristics of numerical attributes.

Sensor	Min	Max	Mean	Std. Dev.	Median
AN311	-266	5	3.484	0.611	3.6
AN422	0	2.4	1.655	0.128	1.6
AN423	-2.4	5.3	1.498	0.33	1.4
TP1721	0	27.9	25.477	0.932	25.4
RH1722	0	71	49.283	6.143	48
BA1723	0	1131.7	1106.161	7.625	1105.9
TP1711	0	31.2	28.894	0.757	28.8
RH1712	0	86	68.687	7.268	69
BA1713	0	1130.9	1105.597	7.617	1105.3
MM252	-0.1	30	0.038	0.121	0
MM261	0	30	0.049	0.125	0
MM262	-0.2	30	0.051	0.136	0
MM263	-2	30	0.248	0.197	0.2
MM264	-2	40	0.327	0.206	0.3
MM256	0	30	0.43	0.204	0.4
MM211	-2	30	0.7	0.151	0.7
CM861	-0.2	67.7	32.92	21.395	43.7
CR863	-8	258	75.081	55.161	78
P_864	0	435.4	86.967	29.158	94.2
TC862	0	40.5	29.898	9.898	32.9
WM868	0	6.39	1.803	1.32	2.2
AMP1_IR	-255	988	5.854	24.413	0
AMP2_IR	-255	1009	5.741	24.25	0
DMP3_IR	-255	216	4.201	17.342	0
DMP4_IR	-255	198	3.97	17.313	0
AMP5_IR	-255	121	0.414	10.966	0
V	0	100	1.347	5.997	0

2. Experimental Design, Materials and Methods

The data set combines data from two sources. The first data source was the set of sensors collecting environmental measurements in an underground coal mine within the SMP-NT safety system¹ and transferred to the THOR dispatching system². SMP-NT is a safety and monitoring system for sites with a methane and coal dust explosion hazard. The main application of the system is monitoring of safety and production in underground mines. SMP-NT provides quasi-continuous communication with underground devices such as sensors and switch-off devices. The THOR system is a solution intended for monitoring industrial plants and technical facilities, e.g. mines, boiler houses or industrial enterprises where it is important to register and visualise environmental parameters data, and provide archiving and reporting that support analysis of the hazards occurring in the monitored facility.

The second data source was the cutter loader moving along the longwall in the coal mine. Its sensors collected the characteristics of the electricity consumed and the driving direction and velocity. The data from the sensors were collected by the MAK S DBC system³, which is a universal, distributed control system for mining machines and devices, and transferred to the THOR dispatching system.

The data was recorded by each of the sources at different frequencies. Therefore, the measurements were standardized so that they corresponded to successive values every one second. If several measurements were recorded within the same second, they were averaged. The missing values for cutter loader measurements (represented originally by NA) were replaced with 0 - meaning no work. The missing values for the remaining sensors were replaced with the last

¹ <https://sevitel.pl/product,3,SMP,SMP.html>.
² <https://sevitel.pl/product,25,THOR.html>.
³ <https://www.ibemag.pl>.

known value. These latter missing values resulted from the standardization of timestamps to every 1 s.

There are no missing values in the coal mine environmental data, as methane systems are safety ones and must operate constantly. If there is an interruption in transmission, the underground part of the system recording measurements in the mine should collect data and, after restoring the transmission, send it to the system operating on the ground for archiving.

Ethics Statement

The work did not involve any human or animal subjects, no data from social media platforms.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

CRediT Author Statement

Michał Kozielski: Writing – original draft, Writing – review & editing; **Marek Sikora:** Conceptualization, Data curation, Funding acquisition, Supervision, Writing – review & editing; **Łukasz Wróbel:** Data curation, Investigation, Methodology, Software, Validation, Writing – review & editing.

Acknowledgments

The data set was collected as part of research supported by the Polish National Centre for Research and Development (NCBiR) Grant PBS2/B9/20/2013 in the frame of the Applied Research Programme.

References

- [1] D. Ślęzak, M. Grzegorowski, A. Janusz, M. Kozielski, S.H. Nguyen, M. Sikora, S. Stawicki, Ł. Wróbel, A framework for learning and embedding multi-sensor forecasting models into a decision support system: a case study of methane concentration in coal mines, *Inf. Sci.* 451 (2018) 112–133, doi:[10.1016/j.ins.2018.04.026](https://doi.org/10.1016/j.ins.2018.04.026).
- [2] A. Janusz, M. Sikora, Ł. Wróbel, S. Stawicki, M. Grzegorowski, P. Wojtas, D. Ślęzak, Mining data from coal mines: IJCRS'15 data challenge, in: Y. Yao, Q. Hu, H. Yu, J.W. Grzymala-Busse (Eds.), *Rough sets, fuzzy sets, datamining, and granular computing*, Springer International Publishing, Cham, 2015, pp. 429–438.