

Response to reviewer 1's feedback –

- Title: Holistic approach promotes failure prevention of smart mining machines based on Bayesian Networks.
- Manuscript ID: machines- 2560637

Dear reviewer,

we would like to thank the reviewer again for the effort and the excellent and valuable feedback we received. It was very helpful and improved the quality of the paper. We have taken these comments into consideration and the article has been updated according to the feedback to our best interpretation. In the following answers, the comment from the reviewer is shown in **black** and our responses in **blue**. We hope that the response to the feedback is satisfactory. We are willing to consider any further suggestion that reviewers might have on the revised manuscript.

Best Regards

Feedback

1.1 Carefully check and improve not only the examples given by the reviewers, but also the entire manuscript.

The manuscript has undergone a revision, with notable enhancements like for instance the provided rewritten abstract below. Other examples: spaces have been removed, slang and American English has been changed to English and figure sequence corrected and equation editor in Word was applied. Finally, acronyms exchanged to capital letters. For instance, the complete term 'artificial intelligence' is spelled out the first time it is mentioned, while subsequently, it is referred to as 'AI'. All changes have been marked with yellow colour in the provided manuscript.

Abstract: In the forthcoming era of fully autonomous mining, spanning from drilling operations to port logistics, novel approaches will be essential to pre-empt hazardous situations in the absence of human intervention. The progression towards complete autonomy in mining operations must be meticulous approaches and uncompromised security. By ensuring a secure transition, the mining industry can navigate the transformative shift towards autonomy while upholding the highest standards of safety and operational reliability. Experiments involving autonomous pathways for mining machinery, utilizing AI for route optimization, demonstrate a capacity to achieve higher speeds compared to manually operated. This translates to enhanced productivity, subsequently fostering increased production capacity to meet the rising demand for metals. Nonetheless, it has been observed an accelerated wear on crucial elements like tires, brakes, and bearings on mining machines. Autonomous mining processes will require smarter machines without humans that guide and support actions prior to a hazardous situation occurring. This paper will delve into a comprehensive perspective on safety of autonomous mining machines by using Bayesian Networks (BN) detecting possible hazards fires. The BN is tuned with a combination of empirical field data and laboratory data. Various faults have been recognized and their correlation with the measurements has been established.

1.2 Pay attention to the newly added content, such as "... it is essential to possess the capability to not only detect errors but also to comprehend their underlying causes ...".

Thanks for the feedback, the sentence has been rewritten. We will review other long sentences as well and have been marked with yellow colour in the manuscript.

It is essential to have the capability to detect underlying causes as well as necessary preventive measures. From these, accurate predictions can be established.

1.3 Figure 6 and Figure 9: Axis titles and their units should be added directly to the figure.

Figure 6: why is only the secondary axis is explained?

To enhance reader comprehension, we have replaced Figure 6. The previous figure may not have been entirely clear in depicting the relationship between machine movement (primary y-axis) and its corresponding CO₂ ppm values (secondary y-axis). The new figure only shows ppm values for the same hour but different years.

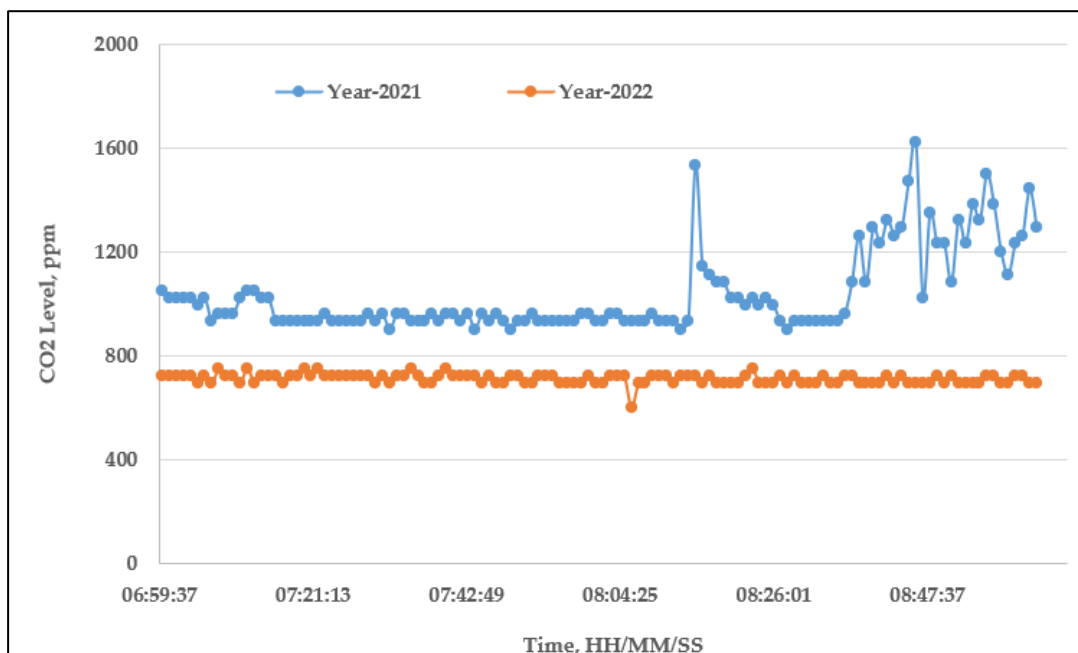


Figure 6. CO₂ level measured on a moving underground mining machine (diesel)

Figure 9: the timestep should have a unit such as second. "The values are scaled". What does it mean?.

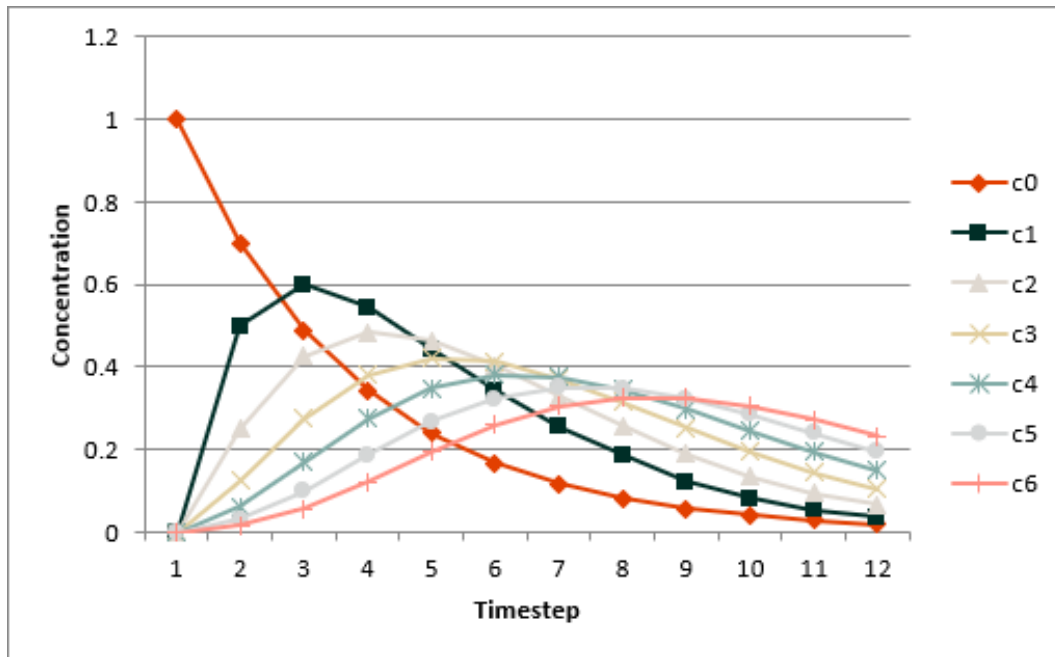


Figure 9. Concentration of CH_x (Y-axis) from the source C0 at the wall out into different positions in the tunnel. C0-C6 represents different positions from timestep =1 to 12 (X-axis). Not fixed dimensions.

To explain Figure 9, the following text have been added to the manuscript:

In Figure 9 general solution is being shown and not the absolute values. Timestep could be in seconds while concentration, C , could be ppm. Position 0 is the source, and the concentration varies depending on if overheating is a fact or not. In volume element 0, the concentration is c_0 . For instance, in timestep 2 the concentration in volume element 1 is 0.5. Equation 3 is showing how it can be simulated. The value in a certain position is depending on the effect of ventilation, vehicle movement and the behaviour of the source. This is addressed by using a fixed ratio between $\Delta t / \Delta x^2$. Assuming scaling up Δx from 1 dm to 3 dm the value of Δt is increase correspondingly.

2. As an application of Bayesian Network, a complete example based on a single hazard (engine overheating, cable overheating or other hazard) must be provided to demonstrate the entire procedure for developing the detection modules from identifying the variables and defining their states until testing the Model. In this example, the following questions should be answered based on detailed data analysis:

2.1 Why and how were the variables selected?

2.2 Why and how were the states of each variables determined?

2.3 How were the relevant probabilities determined and verified?

2.4 How was the model verified?

1000 Thanks for clarifying, for this reason additional text and an appendix C have been included in the manuscript.

A complete example for the engine overheating detection BN module has been provided in Appendix C. This demonstrates the entire procedure from identifying the variables to defining their states until training and testing the model. In this example, the following points have been addressed:

- Why and how the variables were selected

- Why and how the states of each variable were determined
- How the relevant probabilities were determined and verified
- How the model was verified

Appendix C

Example: Engine Overheating Detection Module

Step 1: Identify the Variables and Define Their States

Variables:

- Engine Temperature (absolute value) and CH_x deviation (from normal values)

States:

- Engine Temperature: OK (Below 90°C), Low (90°C - 105°C), Medium (105°C - 120°C), and High (Above 120°C)
- CH_x Deviation: OK (Below 0.1 ppm), Low (0.1 ppm - 0.7 ppm), Medium (0.7 ppm - 1.2 ppm), and High (Above 1.2 ppm)

Temperature and CH_x are parameters that can characterize engine overheating in the mining sector. Their states were defined based on recommendations and experiences of mining machine operators in relation to engine overheating states.

Step 2: Construct the Bayesian Network (BN)

Parent Node: Engine Overheating (States: OK, Low, Medium, High)

The states for the parent node (Engine Overheating) are defined based on the nature of the problem and its relevance in the maintenance decision making. In this case, "OK" represents normal operating conditions, while "Low," "Medium," and "High" represent increasing levels of engine overheating severity. The maintenance actions and readiness levels for engine overheating depend on the severity of the overheating condition. "Low" may involve minor maintenance and restricted operation, "Medium" requires immediate attention and potential vehicle shutdown, while "High" demands immediate action and may result in the vehicle being inoperable until the issue is fully resolved to ensure safety and prevent fire hazards.

Child Nodes:

- Engine Temperature (States: OK, Low, Medium, High)
- CH_x Deviation (States: OK, Low, Medium, High)

Step 3: Determine Prior Probabilities

Prior probabilities are determined based on historical data and expert knowledge. Overheating can occur on different levels, while severe overheating is rare over the engine's lifetime. For this example, the following prior probabilities are considered:

- $P(\text{Engine Overheating} = \text{OK}) = 0.99$
- $P(\text{Engine Overheating} = \text{Low}) = 0.025$

- $P(\text{Engine Overheating} = \text{Medium}) = 0.03$
- $P(\text{Engine Overheating} = \text{High}) = 0.001$

Step 4: Compute Conditional Probability Tables (CPTs)

The CPT of the Engine Overheating BN module is provided in Table 2, Section 2.4. The CPTs define the relationship between parent and child nodes. For instance, when the engine is under the 'OK' operating state, the probability of engine temperature to be in the 'OK' state is 0.99. The remaining 0.1 probability value is dedicated to sensor noise. A very small CH_x is expected at this operating condition of the engine. However, due to sensor errors, some CH_x observations are expected to fall in the 'M' zone.

Step 5: Infer the State of the System

To infer the state of engine overheating, we need new data on Engine Temperature and CH_x deviation. Assuming the following measurements:

- Engine Temperature: $110^\circ\text{C} = \text{'M' state}$
- CH_x Deviation: $0.8 \text{ ppm} = \text{'M' state}$

It is important to note that, for a given engine temperature within its 'Medium' state, it is more likely that the CH_x values fall within the same 'Medium' state range due to the relationship they have.

Bayes' theorem and inference algorithms too is used to calculate the posterior probabilities for each state of Engine Overheating.

- $P(\text{Engine Overheating} = \text{OK} | \text{Engine Temperature} = \text{M}, \text{CH}_x \text{ Deviation} = \text{M})$
- $P(\text{Engine Overheating} = \text{L} | \text{Engine Temperature} = \text{M}, \text{CH}_x \text{ Deviation} = \text{M})$
- $P(\text{Engine Overheating} = \text{M} | \text{Engine Temperature} = \text{M}, \text{CH}_x \text{ Deviation} = \text{M})$
- $P(\text{Engine Overheating} = \text{H} | \text{Engine Temperature} = \text{M}, \text{CH}_x \text{ Deviation} = \text{M})$
- $P(\text{Engine Overheating} = \text{OK} | \text{M}, \text{M}) \propto P(\text{Engine Overheating} = \text{OK}) * P(\text{Engine Temperature} = \text{M} | \text{OK}) * P(\text{CH}_x \text{ Deviation} = \text{M} | \text{OK}) \propto 0.99 * 0.004 * 0.05 = 0.000198$
- $P(\text{Engine Overheating} = \text{L} | \text{M}, \text{M}) \propto P(\text{Engine Overheating} = \text{L}) * P(\text{Engine Temperature} = \text{M} | \text{L}) * P(\text{CH}_x \text{ Deviation} = \text{M} | \text{L}) \propto 0.005 * 0.02 * 0.95 = 0.0000475$
- $P(\text{Engine Overheating} = \text{M} | \text{M}, \text{M}) \propto P(\text{Engine Overheating} = \text{M}) * P(\text{Engine Temperature} = \text{M} | \text{M}) * P(\text{CH}_x \text{ Deviation} = \text{M} | \text{M}) \propto 0.004 * 0.97 * 0.95 = 0.01846$
- $P(\text{Engine Overheating} = \text{H} | \text{M}, \text{M}) \propto P(\text{Engine Overheating} = \text{H}) * P(\text{Engine Temperature} = \text{M} | \text{H}) * P(\text{CH}_x \text{ Deviation} = \text{M} | \text{H}) \propto 0.001 * 0.01 * 0.99 = 0.000000099$

Calculate the normalisation constant (Z):

Normalisation constant (Z) = Sum of all unnormalised probabilities

$$Z = 0.000198 + 0.0000475 + 0.01846 + 0.000000099 \approx 0.018705599$$

Now the probabilities are normalised:

- $P(\text{Engine Overheating} = \text{OK} | \text{M}, \text{M}) \approx 0.000198 / 0.018705599 \approx 0.010597$

- $P(\text{Engine Overheating} = L | M, M) \approx 0.0000475 / 0.018705599 \approx 0.002540$
- $P(\text{Engine Overheating} = M | M, M) \approx 0.01846 / 0.018705599 \approx 0.986062$
- $P(\text{Engine Overheating} = H | M, M) \approx 0.000000099 / 0.018705599 \approx 0.000005$

The state with the highest probability is "Engine Overheating = Medium" with a probability of approximately 0.986. Based on the observed data (Engine Temperature = Medium, CH_x = Medium), and the estimated posterior probabilities, the decision by the BN is that the severity of engine overheating is "Medium."

The sensitivity of the analysis to the prior probability values has also been examined based on the following three cases:

Case 1:

- $P(\text{Engine Overheating} = \text{OK}) = 0.98$
- $P(\text{Engine Overheating} = L) = 0.01$
- $P(\text{Engine Overheating} = M) = 0.008$
- $P(\text{Engine Overheating} = H) = 0.002$

Case 2:

- $P(\text{Engine Overheating} = \text{OK}) = 0.97$
- $P(\text{Engine Overheating} = L) = 0.015$
- $P(\text{Engine Overheating} = M) = 0.012$
- $P(\text{Engine Overheating} = H) = 0.003$

Case 3:

- $P(\text{Engine Overheating} = \text{OK}) = 0.95$
- $P(\text{Engine Overheating} = L) = 0.025$
- $P(\text{Engine Overheating} = M) = 0.02$
- $P(\text{Engine Overheating} = H) = 0.005$

The estimated probabilities through the inferencing indicated that "Engine Overheating = Medium" is the most likely state.

3.0 Line 355: "...the BN models will be tested on a new dataset from..." Have the BN models tested?

We are not sure if we have interpreted the question correctly, however if it is about validating the performance of the BN model of a new dataset, this is described in step 5. If it is concerning the practical implementation of BN model, this will be part of our future research work.

To support the readers, we redefined the sentence:

Implementation of the BN model: The BN model will be tested by using new dataset from machines in operation to evaluate its performance in detecting faults. This phase is ongoing and will demand several months of testing to retrieve relevant data and make statistical analyses. Furthermore, fine tuning of the BN model parameters will be carried out. From the first experiments on a single machine, initial data was retrieved. However, the entire fleet is now in focus.

Thank you once again!

Best regards

Authors