Case-based reasoning system for prediction of fuel consumption by haulage trucks in open-pit mines

Gomolemo Tadubana1, Boyce Sigweni2, Raymond Suglo3
1,2Department of Mining and Geological Engineering, Botswana
2Department of Electrical, Computer and Telecommunications Engineering
Botswana International University of Science and Technology, Botswana

ABSTRACT
The shovel-truck system is commonly used in open-pit mining operations. Truck haulage cost constitutes about 26% of open-pit mining costs as the trucks are mostly powered by diesel whose cost is escalating annually. Therefore, reducing fuel consumption could lead to a significant decrease in overall mining costs. Various methods have been proposed to improve fuel efficiency in open-pit mines. Case-based reasoning (CBR) can be used to estimate fuel consumption by haulage trucks. In this work, CBR methods namely case-based reasoning using forward sequential selection (CBR-FSS), traditional CBR, and Naïve techniques were used to predict fuel consumption by trucks operating at Orapa Mine. The results show that the CBR method can be used to predict fuel consumption by trucks in open-pit mines; the predicted values of fuel consumption using the CBR-FSS technique gave much lower absolute residual values, higher standardised accuracy values, and effect sizes than those of other prediction techniques on all the datasets used. The system will enable mine planners to know the fuel consumed per trip and allow them to take mitigation measures on trucks with high fuel consumption.

This is an open access article under the CC BY-SA license.

1. INTRODUCTION
Mining is the process of obtaining useful minerals from the earth’s crust at a profit [1]. Open-pit mining operations generally employ conventional methods to mine ore or waste [2]. The profitability of any mine is determined by the efficient management of the unit operations [3]. Research shows that truck haulage costs account for 26% of the production costs, and together with shovels, they account for 40% of the total mining cost [4-6]. Diesel operated dump trucks are commonly used to haul materials in most surface mining operations [7]. The mobile materials handling fleet often account for a sizeable amount of both capital and operational costs. Fuel consumption is one of the primary operating costs associated with shovel-truck operations, with fuel cost constituting a significant part of the materials handling cost [8] and accounts for about 30% of total energy costs in surface mines. During idling, there is no production by the truck but the engine continues to run. Hence it is necessary to control fuel consumption by haulage trucks to reduce mining costs. The need for cost-saving has motivated mine operators and governments to conduct several research studies on how to reduce energy consumption in the mining industry [9].

Journal homepage: http://ijece.iaescore.com
A paper using case based reasoning (CBR) shows that good driving style reduced fuel consumption by 12% for light driving vehicles in urban areas [10]. However, CBR cannot be used in predictions when the available data is misleading or irrelevant [11]. Hence, methods that lessen the impact of irrelevant or misleading features are required to ensure that reliable and accurate predictions are made using CBR for the CBR frameworks to benefit from improvements of the feature sets.

The use of the CBR method could enable mine operators to predict the fuel consumed by each truck per trip and the total amount of fuel consumed per shift or per day by the truck fleet. This will assist in making logistical arrangements for fuel supplied to the mine and to determine when the fuel consumption of any of the trucks is getting too high for remedial action. This study investigated the factors that affect fuel consumption by haul trucks in open pit mines and developed an algorithm for predicting the consumption of fuel by haul trucks. The CBR algorithm that was developed was used to predict fuel consumption of randomly selected haulage trucks in an operating open pit mine.

2. BACKGROUND RESEARCH

Case based reasoning is the process of effectively dealing with a new problem based on the solution of the past problems [12, 13]. After getting a prediction each time, a new problem has been solved, the new solution is retained and made available for future identical problems [14]. For accurate predictions to take place the cases must be similar and close to the new problem [15, 16].

The primary idea of CBR is that when a new prediction of fuel consumed per trip is required, a similarity measure will be used to select the most similar past trips to predict the amount of fuel to be consumed. The performance of CBRs in relation to other prediction methods has been shown to be encouraging [17]. The level of success of CBRs depend upon attributes of the dataset. Discontinuities existing in fundamental connections between dependent and independent variables tend to make the CBR technique more effective [18]. Unfortunately, CBR method is often exposed to the likelihood of misleading or irrelevant factors in the prediction exercise [11]. Therefore, all irrelevant or misleading features must be removed to reduce their effects in the predictions. This is done by using feature subset selection (FSS) which determines the most favorable feature subsets for an accurate prediction [19]. In FSS the factors that carry significant information regarding the output are considered as relevant features. Thus, in assigning weights, it is best to allocate higher weight values to factors that carry significant information about the output. Also, all CBR methods must assign weights to factors. For instance, a CBR technique that uses all features basically assigns non-zero weights to all factors. FSS techniques assign 0 or 1 weights to features indicating exclusion or inclusion respectively [12]. FSS techniques can be categorised into two main groups using filter and wrapper methods. The filter method reduces the number of factors before training [20]. Thus, it is less computationally complex but less accurate [10]. Wrapper method combines with indicators to limit fitting errors [21]. Hence, wrapper method has high fitting accuracy, high complexity and a low generalisation of chosen factors to other conditions.

2.1. Dataset creation

In this paper, a dataset is formed entailing dependent variables (DV) and independent variables (IV) as required by the CBR technique. DV is the variable that must be estimated using a lot of independent variables or parameters. Fuel consumption in light duty vehicles has been successfully evaluated using CBR techniques [10]. In most researches, DV and IV of datasets are chosen utilising domain experts and the choice of the best parameters do not necessarily lead to optimal results. Therefore, to derive the right independent variables, factors that affect fuel consumption by haulage trucks are used.

2.2. Independent variables

From the literature, 19 parameters where identified to influence fuel consumption by haulage trucks. Interviews, site visits and data was collected from Komatsu and Barloworld who are the service providers for haulage trucks at Debswana Orapa Mine. Table 1 (see in Appendix) summaries the key 19 parameters identified by various researchers as determinants of fuel consumption by trucks in open pit mines. These were employed to predict fuel consumption by haulage trucks in this study.

3. RESEARCH METHODS

3.1. Repeated measure design

Repeated measure design is taking measurements on the same subject over time and under different conditions. It is characterised by having more than one measurement of at least a given variable for each subject [34]. Repeated design measures were used in this research because they reduce the variance of
estimates of fuel consumed per trip, fewer trips need to be trained to complete the analysis enabling many analyses to be done in a shorter period and it permits monitoring of how the weights of the features change after some time for both short- and long-term situations.

3.2. Creation of datasets

Domain experts have been used by most researchers to select dependent and independent variables for data sets. There are next to no hypothetical reflection to help select great variables [35]. To select the best variables, domain experts may not always be available. Therefore, this work presents variable selection systems which can be used by mine managers.

4. EVALUATION METHODS

The accuracy of various estimation techniques is used as the main criterion in analysing the value of CBR’s estimation of the fuel consumption by haulage trucks. Shepperd and MacDonnell method (SMM) was used to reduce the irregularities between validated research results and to provide a basis for understanding results with a specific emphasis on continuous prediction systems [36]. This work also followed the methodology generally employed when predicting the Design Reality Gap scores for telecentres [12].

The evaluation of an estimation system is based on standardised accuracy (SA), an unbiased statistic, calculation of effect sizes and testing the results of likelihood of the value relative to the baseline technique of random predictions (guessing). According to the SMM, an estimation system, $E_i$, is assumed. This is validated over a dataset $D$ using some accuracy statistic $S$ as per the validation scheme $V$ [37]. The SMM method can be utilised alongside other competing estimation systems. For example, given two estimation systems, $E_1$ and $E_2$ and an accuracy statistic $S$, one must answer basic questions such as: How is the performance against random guessing? What is significance testing? What is the effect of size?

4.1. Performance against random guessing

A baseline of random guessing is established to find out if the suggested method performs better than random guessing. It is expected that any theoretically good system should execute better than random guessing over time. If that is not the case, then it is assumed that the indicator is not utilising the target case features in any useful way.

4.2. Significance testing

Mean absolute residual (MAR) is used as the accuracy statistic, $S$, for continuous estimation systems. In contrast to mean magnitude of relative errors, MAR does not depend on proportions. Thus, it is unbiased. Unfortunately, with MAR comparisons across the datasets cannot be made because the residuals are not standardised and are difficult to interpret. Accordingly, a standardised accuracy measure (SA) has been presented [36] where accuracy is measured as the MAR relative to random guessing figure, $E_0$. Hence, the SA for estimation of $E_i$ is given by (1).

$$SA_{E_i} = 1 - \frac{\text{MAR}_{E_i}}{\text{MAR}_{E_0}} \times 100$$ (1)

where $\text{MAR}_{E_0}$ is the mean of a large number, normally 1000, runs of random guessing? SA is a ratio representing how better the estimation system, $E_i$, is than random guessing $E_0$. If the value of SA is near zero it would be discouraging, and a negative value would be regrettable.

4.3. Effect size

Effect size is a simple way of quantifying the difference between the methodologies. Research work has shown that the larger the effect size the stronger the relationship between the two methodologies. To assess the effect size, a standardised measure by Glass delta ($\Delta$) is used [36]. Glass delta is a measure of effect size which uses only the control standard deviations and it is only used when the standard deviations are significantly different between the techniques, this is given by (2).

$$\Delta = \frac{\text{MAR}_{E_i} - \text{MAR}_{E_0}}{\sigma_{E_0}}$$ (2)

where $E_0$ represents the sample standard deviation of the random guessing methodology. Glass delta standardises the contrast between the two estimation frameworks and afterwards it contextualises the distinction as far as measure of variety in the two measures of the accuracy statistic $S$. 

Case-based reasoning system for prediction of fuel consumption by... (Gomolemo Tadubana)
The standardised effect size is scale-free and considered as small (=0.2), medium (=0.5) and large i.e.=0.8 [37]. The \( \Delta \) has a unit of a standard deviation, so the effect is a decrease in the MAR of \( n \) amount of fuel consumed per trip.

5. EXPERIMENTS

Three CBR estimation techniques namely traditional CBR, case-based reasoning using forward sequential selection (CBR-FSS) and Naïve methods, were used in the prediction of fuel consumption by haulage trucks in an operating open pit mine in Botswana. These techniques have been utilised by other researchers in their work and they constitute a range of potential methodologies that can be used for case-based fuel consumption estimation [37, 38]. The Naïve method uses a sample mean to estimate a new fuel consumed per trip [36, 37] and CBR method uses all the parameters equally weighted to estimate the fuel consumed. The Naïve technique might be considered as a baseline technique and its application is from the 1990 s [39]. The case-based reasoning using forward sequential selection (CBR-FSS) method was preferred in this research because it has been proven to perform better that other techniques employed in their research [37]. CBR-FSS gives weights of 0 to irrelevant parameters and 1 to relevant parameters. This technique excludes all irrelevant features to fuel consumption prediction as compared to CBR.

5.1. Limitations of study

This study used data obtained from Debswana Orapa Mine only. Therefore, the results may not be applicable to all mines due to their different locations and operating conditions. Another limitation of the data used is that it does not include all parameters such as weather identified from the literature. Earlier studies have identified the parameters in Table 2 as those that influence fuel consumption by haulage trucks in open pit mines [22, 40-43].

| #  | Factor                  | Description/Motivation                                                                 | Source         |
|----|-------------------------|----------------------------------------------------------------------------------------|----------------|
| 1  | Extreme acceleration    | Refers to the rate of change of velocity over time. It is defined where fuel consumption increases rapidly while driving. An increase in fuel consumption occurs with acceleration increase or extreme acceleration. | [22, 40, 41]  |
| 2  | Extreme braking         | This is when the braking distance required to stop is shorter than expected.            | [44]           |
| 3  | Excess RPM              | It is the measure of how many times an engine turns in a minute. Fuel consumption is lower at low RPM due to friction. Increases levels of congestion lead to lower average speeds, longer travel times and increased delays at loading and dumping points. Hence flow of road traffic has a great effect on fuel consumption. | [45]           |
| 4  | Traffic conditions      |                                                                                       | [46, 47]       |
| 5  | Road geometry           | Road geometry has a great impact on overall energy consumption.                       | [48]           |
| 6  | Driver’s age            | Older drivers have been acknowledged as less aggressive when driving as compared to younger drivers | [49]           |
| 7  | Truck age               | Older vehicles have high shortfall in fuel consumption as compared to younger drivers. | [42]           |
| 8  | Driver’s experience     | A skilled driver can reduce the fuel consumption by more than 10% compared to an inexperienced driver, by reducing the need of braking | [43]           |
| 9  | Vehicle consuming accessories | These are heating, air-conditioning and other accessories that consume fuel.         | [45]           |

6. EXPERIMENTAL RESULTS

All the statistical analysis contained in this research work is based on absolute residuals from the methodologies selected. The absolute residual results are summarised in Table 3 while Figure 1 shows the residual distribution of the results. The SMM method was used to evaluate the results from the various estimation techniques. Consequently, the results are analysed based on mean absolute residuals being 14.24, 20.24 and 24.33 for CBR-FSS, CBR and Naïve technique respectively as compared to 35.97 obtained from random guessing. The initial results on the standardised accuracy of the three methods compared to random guessing based on standardised accuracy (SA) and effect size (\( \Delta \)) are summarised in Table 4 [39]. The results in Table 4 shows that when the CBR-FSS, CBR and Naïve techniques are used to predict fuel consumption by haul trucks per trip in open pit mines, the predictions are 60.42%, 43.73% and 32.37% better than those based on random guessing respectively. The effect size relative to guessing by CBR-FSS can be said to be of medium effect size while that for the CBR and Naïve technique may be considered as small.
Table 3. Random sample (10%) of actual and predicted fuel consumed and absolute residuals

| Trip Number | Actual Fuel Consumed litres (x) | Predicted Fuel litres (y) | Absolute Residuals/Errors (|x - y|) |
|-------------|---------------------------------|---------------------------|----------------------------------|
| CBR-FSS     | CBR                             | Naïve                     | CBR - FSS                        | Naïve |
| 5           | 92.74                           | 92.07                     | 89.03                            | 85.17 | 0.68 | 3.71 | 7.58 |
| 20          | 87.06                           | 86.12                     | 99.38                            | 85.17 | 0.95 | 12.31 | 1.90 |
| 21          | 86.12                           | 86.59                     | 93.21                            | 85.17 | 0.47 | 7.09 | 0.95 |
| 29          | 21.77                           | 20.50                     | 39.74                            | 85.17 | 1.26 | 17.97 | 6.30 |
| 47          | 90.85                           | 90.69                     | 84.23                            | 85.17 | 0.16 | 6.63 | 5.68 |
| 49          | 88.96                           | 89.21                     | 99.33                            | 85.17 | 0.25 | 10.38 | 3.79 |
| 62          | 101.26                          | 101.23                    | 111.20                           | 85.17 | 0.03 | 9.94 | 16.09 |
| 70          | 86.12                           | 85.65                     | 88.96                            | 85.17 | 0.47 | 2.84 | 0.95 |
| 74          | 88.01                           | 88.64                     | 156.22                           | 85.17 | 0.63 | 68.21 | 2.85 |
| 107         | 29.34                           | 28.39                     | 64.16                            | 85.17 | 0.95 | 34.83 | 55.83 |
| 114         | 88.01                           | 88.48                     | 118.07                           | 85.17 | 0.47 | 30.05 | 2.85 |
| 120         | 90.85                           | 90.85                     | 99.79                            | 85.17 | 0.00 | 8.94 | 5.68 |
| 134         | 102.21                          | 103.15                    | 97.98                            | 85.17 | 0.95 | 4.22 | 17.04 |
| 138         | 34.07                           | 33.83                     | 83.48                            | 85.17 | 0.24 | 49.42 | 51.10 |
| 139         | 35.02                           | 35.02                     | 35.96                            | 85.17 | 0.00 | 0.95 | 50.15 |
| 142         | 35.02                           | 35.96                     | 142.66                           | 85.17 | 0.95 | 107.64 | 50.15 |

Table 4. Standardised accuracy and effect of sizes

| Technique | Criteria | SA (%) | ∆    |
|-----------|----------|--------|------|
| CBR-FSS   |          | 60.42  | 0.672|
| CBR       |          | 43.73  | 0.486|
| Naïve     |          | 32.37  | 0.360|

Figure 1. Boxplot showing residual distribution of the dataset

7. CONCLUSION

In this paper, it has been shown that the CBR-FSS technique can be used to predict fuel consumption per trip of haulage trucks in open pit mines without any expert knowledge. The results also show that when using CBR-FSS, CBR and the Naïve techniques to predict fuel consumption by haul trucks in open pit mines, the predictions are 60.42%, 43.73% and 32.37% respectively better than those based on random guessing. Furthermore, the standardised values show that predictions using CBR-FSS are 16.69% and 28.05% better than those of CBR and Naïve techniques. The effect size relative to guessing by CBR-FSS can be said to be of medium effect size while those of CBR and Naïve techniques may be considered as small.

It is concluded that using CBR prediction techniques to predict fuel consumption in open pit mines could be more cost-effective than using rough estimations to predict outputs or engaging experts to do extensive evaluations for the mine. It is acknowledged that predictions made in this work were based on datasets from one operating mine. Accordingly, the predictions made on fuel consumption by haulage trucks may only be applicable to the mine’s setting only and larger dataset features are needed to make the findings more applicable all mines irrespective of their location and operating conditions.

Case-based reasoning system for prediction of fuel consumption by... (Gomolemo Tadubana)
APPENDIX

Table 1. Feature description for haulage truck fuel consumption dataset [22]

| #  | Factor                     | Description/Motivation                                                                 | Source(s) |
|----|----------------------------|----------------------------------------------------------------------------------------|------------|
| 1  | Material                   | The material which is hauled by haul trucks                                          | [23]       |
| 2  | Weather conditions         | Ambient conditions refer to the external conditions such as wind, temperature and barometric pressure. These affect a vehicle’s fuel consumption as they influence the engine operation. They might also affect driver’s behavior as the driver must adjust his driving pattern accordingly.          | [24,25]   |
| 3  | Idle time                  | Refers to running a vehicle without moving the vehicle. When the duration of idling is longer than 10 s, an engine consumes more fuel compared to when it is restarted. | [26-28]   |
| 4  | Speed                      | The rate of fuel consumption increases with an increase in speed. Fuel consumption is increased by aggressive driving dramatically by up to 24%. | [23,29]   |
| 5  | Trip frequency             | Trip frequency of haulage trucks directly influences the production output in open pit mines. The trip frequency should be maximised to accomplish least fuel consumption per ton of materials hauled. | [30]       |
| 6  | Type of haulage truck      | Fuel consumption of vehicles vary primarily with vehicle type. Truck types used were CAT 789, KOMATSU 730 or KOMATSU 830 | Debswana Orapa Mine |
| 7  | Gradient                   | Rolling resistance of the haul trucks vary due to road conditions. A haul road that is dry and hard-packed keeps fuel costs and tyre wear to a minimum. | [31,32]   |
| 8  | Payload                    | Fuel consumption increases with an increase in the gross weight at which a truck operates. | [33]       |
| 9  | Trip distance              | One-way distance in meters                                                             | [23]       |
| 10 | Cycle time                 | For trucks is the sum of loading time, travel loaded, return (empty), dumping, spotting and wait times at excavator and dumping points. | [24,33]   |
| 11 | Loading time               | Measure of the time taken to load the haulage truck by the excavator.                  | [8,23]     |
| 12 | Spotting time              | Measure of the time taken during which the loading unit has the bucket in place to dump but awaits the truck to move into correct loading position. | [8,23]     |
| 13 | Dumping time               | Measure of the time taken for the truck to maneuver and dump its load. Measure of the time taken for the truck to maneuver and dump its load. | [8,23]     |
| 14 | Fixed time                 | Sum of load, spot and dump times.                                                      | [8,23]     |
| 15 | Travelling time            | A measure of the time taken to travel (loaded) and return (empty).                     | [8,23]     |
| 16 | Tonne-kilometre per hour (t-km/hr) | This represents Tyre Load x Average Speed                                              | Debswana Orapa Mine |
| 17 | Frame Torque (tonne-meter) | Measured as the frame time in either clockwise direction or anticlockwise               | Debswana Orapa Mine |
| 18 | Torque time                | Time taken from the start of the haul cycle to the peak torque                          | Debswana Orapa Mine |
| 19 | Sprung load                | Peak dynamic load calculations                                                            | Debswana Orapa Mine |

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Kwangu Sithelo, the Debswana Orapa Mine Senior Mining Engineer, for his assistance in this research.

REFERENCES

[1] D. Espinoza, M. Goycoolea, E. Moreno, and A. Newman, “Minelib: a library of open pit mining problems,” *Annals of Operations Research*, vol. 206, no. 1, pp. 93–114, 2013.
[2] H. L. Hartman and J. M. Muttansky, “Introductory mining engineering,” *John Wiley & Sons*, 2002.
[3] P. P. V. Coronado, “Optimization of the haulage cycle model for open pit mining using a discrete-event simulator and a context-based alert system,” Theses, The University of Arizona, 2014.
[4] T. Bishop, “Trucks,” *Surface Mining*, pp. 553–583, 1968.
[5] S. Michaelson, “Open pit mining past, present, and future,” *Open Pit Mine Planning and Design* (eds. JT Crawford and WA Hustrulid), pp. 5–16, 1979.
[6] T. Sullivan, “New technology and economies of scale in shovel-truck sizing,” in *Proc. of Second Inter-national Symposium on Mine Planning and Equipment Selection*, Calgary, Canada, 1990, pp. 3–6.
[7] E. Siami-Irdeemoosa and S. R. Dindarloo, “Prediction of fuel consumption of mining dump trucks: A neural networks approach,” *Applied Energy*, vol. 151, pp. 77–84, 2015.
[8] N. Dumakor, V. Temeng, and K. Bansah, “Optimising shovel-truck fuel consumption using stochastic simulation,” *Ghana Mining Journal*, vol. 17, no. 2, pp. 39–49, 2017.
[9] A. Soofastaei, “Development of an advanced data analytics model to improve the energy efficiency of haul trucks in surface mines,” Theses, The University of Queensland, 2016.
[10] J. Rios-Torres, J. Liu, and A. Khattak, “Fuel consumption for various driving styles in conventional and hybrid electric vehicles: Integrating driving cycle predictions with fuel consumption optimization,” *International Journal of Sustainable Transportation*, pp. 1–15, 2018.
Case-based reasoning system for prediction of fuel consumption by... (Gomolemo Tadubana)
K. Brundell-Freij and E. Ericsson, “Influence of street characteristics, driver category and car performance on urban driving patterns,” Transportation Research Part D: Transport and Environment, vol. 10, no. 3, pp. 213–229, 2005.

I. M. Berry, “The effects of driving style and vehicle performance on the real-world fuel consumption of US light-duty vehicles,” PhD dissertation, Massachusetts Institute of Technology, 2010.

M. Ivarsson, “Fuel optimal powertrain control for heavy trucks utilizing look ahead,” Ph.D. dissertation, Linköping University Electronic Press, 2009.

P. A. Rimmo and L. Hakamies-Blomqvist, “Older drivers’ aberrant driving behaviour, impaired activity, and health as reasons for self-imposed driving limitations,” Transportation Research Part F: Traffic Psychology and Behaviour, vol. 5, no. 1, pp. 47–62, 2002.

J. C. Ferreira, J. de Almeida, and A. R. da Silva, “The impact of driving styles on fuel consumption: A data-warehouse-and-data-mining-based discovery process,” IEEE Trans. Intelligent Transportation Systems, vol. 16, no. 5, pp. 2653–2662, 2015.

B. Jereb, S. Kumpersˇcak, and T. Bratina, “The impact of traffic flow on fuel consumption increase in the urban environment,” FME Transactions, vol. 46, no. 2, pp. 278–284, 2018.

A. Y. Bigazzi and K. J. Clifton, “Modeling the effects of congestion on fuel economy for advanced power train vehicles,” Transportation Planning and Technology, vol. 38, no. 2, pp. 149–161, 2015.

B. Luin and S. Petelin, “Impact of road geometry on vehicle energy consumption,” Transport Problems, vol. 12, 2017.

R. D. Mather, “Age and driving behaviour: contributions from human factors,” Journal of Scientific Psychology, pp. 24–31, 2007.