12

Application of Advanced Data Analytics to Improve Haul Trucks Energy Efficiency in Surface Mines

Ali Soofastaei, Peter Knights, and Mehmet Kizil

CONTENTS

12.1 Introduction

Truck haulage is responsible for a majority of cost in a surface mining operation. Diesel fuel, which is costly and has a significant environmental footprint, is used as a source of energy for haul trucks in surface mines. Accordingly, improving truck energy efficiency would lead to a reduction in fuel consumption and therefore greenhouse gas emissions.

The determination of haul trucks fuel consumption is complex and requires multiple parameters including the mine, fleet, truck, speed, payload, operator inputs, fuel, climate, tire, and road conditions as inputs. Data analytics can be used to simulate the complex relationships between the input parameters affecting haul trucks fuel consumption. The aim of this chapter is to introduce an advanced data analytics model to improve the energy efficiency of haul trucks in surface mines.

The most important controllable parameters affecting fuel consumption are payload, truck speed, and total rolling resistance. From these parameters a comprehensive analytical framework can be developed to determine the opportunities for minimizing truck fuel consumption. The first stage of the analytical framework includes the development of the artificial neural network (ANN) model to determine the relationship between truck fuel consumption and payload, truck speed, and total resistance.

This model can be trained and tested using real data collected from some large surface mines in Australia, the United States, and Canada. A fitness function for the haul truck fuel consumption can be successfully generated. This fitness function is then used in the second stage of the analytical framework to develop a digital learning algorithm based on a novel multiobjective genetic algorithm (GA). The aim of this algorithm is to establish the optimum set points of the three controllable parameters to reduce the diesel fuel consumption, with these set points being specific to individual mines and fleet operations.

12.2 Context: Reducing Energy Consumption via Data Analytics

Energy efficiency has gradually become a more important consideration worldwide since the rise of the cost of fuel in the 1970s. The mining industry annually consumes trillions of British thermal units (BTUs) of energy in operations such as exploration, extraction, transportation, and processing. Mining operations use energy in a variety of ways, including excavation, materials handling, mineral processing, ventilation, and dewatering. It also uses significant quantities of power. The Mining industry consumed 520 petajoules (PJ) of energy in 2014–2015 or 9% of the national energy end use in Australia (Allison et al. 2016). Energy consumption in mining is rising at around 6% annually in Australia due to lower grade ores located deeper underground (EEO 2012), a trend seen in other developed countries (DOE 2012). As well as improving margins through efficiency savings, energy streamlining in the sector can also result in the reduction of millions of tons of gas emissions because the primary energy sources used in the mining industry are petroleum products: electricity, coal, and natural gas.

The potential for energy (and financial) savings has motivated the mining industry and governments to conduct research into the reduction of energy consumption. Consequently, a large number of research studies and industrial projects have been carried out in an attempt to do this in mining operations across the world (Soofastaei 2016). Current investments in the improvement of mining technologies and energy management systems have resulted in a significant reduction of energy consumption. Based on completed industrial projects, significant further opportunities exist within the mining industry to reduce energy consumption. The case study presented here—haulage equipment—is one of these potential areas.

Service trucks, front-end loaders, bulldozers, hydraulic excavators, rear dump trucks, and ancillary equipment, such as pick-up trucks and mobile maintenance equipment, are prominent examples of the diesel equipment and associated energy footprint of mining operations. In surface mines, the most commonly used means of mining and hauling of materials is by a truck and shovel operation. The trucking of overburden constitutes a major portion of energy consumption. However, as will be discussed, the rate of this energy consumption is the result of many different parameters (EEO 2010) which can be analyzed and altered to obtain optimal levels of performance.

Data analytics represents a very appropriate approach to pulling together these disparate data sources since it is the science of examining raw data to draw conclusions about that information. The main advantages of data analytics can be presented by cost reduction, faster and better decision-making, and finally new products and services (Soofastaei and Davis 2016). The uses of data analytics are many and can apply to areas that many might not have thought of before. One area that sees much potential in data analytics is the mining industry. For an industry that does trillions of dollars in business every year, data analytics should be considered a necessity not a luxury. Indeed, there are many phases of the mining process where data analytics can be put to use. The four main phases are the (1) extraction of ore, (2) materials handling, (3) ore comminution and separation, and (4) mineral processing. Of particular focus for some companies is efficiency improvements in the second phase, materials handling. Without data analytics at the heart of this phase, operators are more than likely to be subject to suboptimal functioning of their equipment, including in haulage vehicles and infrastructure.

As Figure 12.1 illustrates, use of data analytics in organizations cover two dimensions: time frame (the past, present, or future) and competitive advantage (value of insight generated). At the lowest level, analytics are routinely used to produce reports and alerts. These are simple, retrospective processing and reporting tools, such as pie graphs, top-ten histograms and trending plots, typically addressing variations of the basic question of "what happened and why?" Increasingly, sophisticated analytical tools, capable of working at or near real-time and providing rapid insights for process improvement, can show the user "what *just* happened" and assist them in understanding "why" as well as the next best action to take. However, at the top of the pictured comparative advantage scale, there are predictive models and optimization tools, aimed at evaluating "what will happen" and identifying the best available responses.

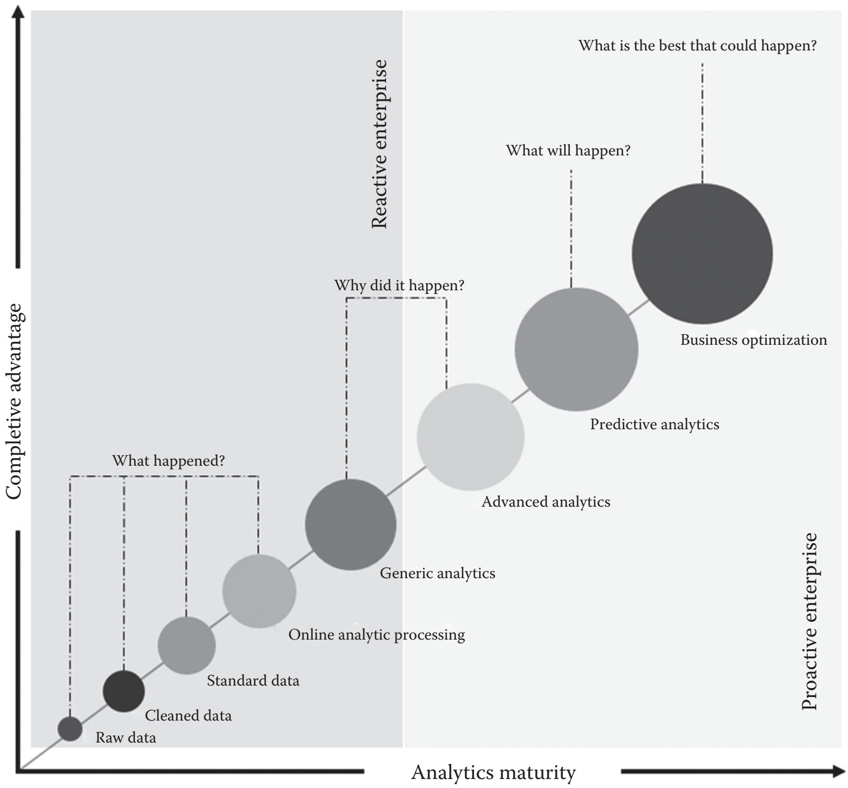

FIGURE 12.1
Competitive advantages of data analytics in organizations. (From Davenport, T. H. et al., *Analytics at Work: Smarter Decisions, Better Results*, Harvard Business Press, New York, 2010.)

12.3 Modeling Haul Trucks' Fuel Consumption

In this chapter, the effects of the three most important and effective parameters on fuel consumption of haul trucks are examined. These parameters are payload (P), truck speed (S), and total rolling resistance (TR). On a real mine site, the correlation between fuel consumption and the three parameters is complex. We use two artificial intelligence methods to create an advanced data analytic model to estimate and reduce haul truck fuel consumption in surface mines. The model can estimate the energy consumption of haul trucks in surface mines using an artificial neural network (ANN) and can also find the optimum values of P, S, and TR using a GA. We analyze each of these in turn and then present the results of our modeling.

12.3.1 Artificial Neural Network

ANNs, also known as neural networks (NNs), simulated neural networks (SNNs), or parallel distributed processing (PDP), are the representation of methods that the brain uses for learning (Hammond 2012). ANNs are a series of mathematical models that imitate a few of the known characteristics of natural nerve systems and sketch on the analogies of adaptive natural learning (Rodriguez et al. 2013). The key component of a ANN paradigm is the unusual structure of the data processing system. ANNs are utilized in various computer applications to solve complex problems. They are fault-tolerant and straightforward models that do not require information to identify the related factors and do not need the mathematical description of the phenomena involved in the process (Beigmoradi et al. 2014).

The main part of a NN structure is a "node." Biological nodes sum the signals received from numerous sources in different ways and then carry out a nonlinear action on the results to create the outputs. NNs typically have an input layer, one or more hidden layers, and an output layer (Figure 12.2).

Each input is multiplied by its connected weight, and in the simplest state, these quantities and biases are combined. They then pass through the activation functions to create the output (Equations 12.1 through 12.3).

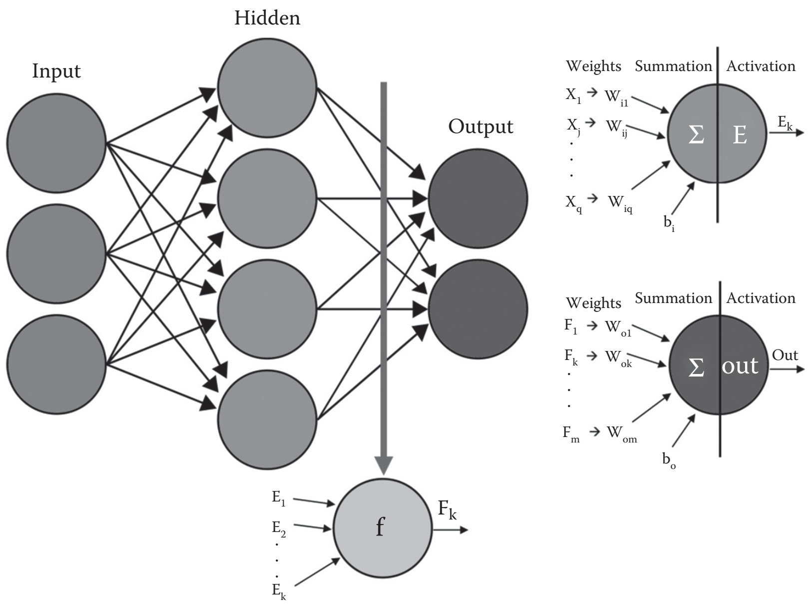

FIGURE 12.2
Artificial neural network structure. (From Soofastaei, A., *Development of an Advanced Data Analytics Model to Improve the Energy Efficiency of Haul Trucks in Surface Mines*, PhD thesis, The University of Queensland, School of Mechanical and Mining Engineering, Brisbane, Australia, 2016.)

$$E_k = \sum_{j=1}^{q} (w_{i,j,k} x_j + b_{i,k}) \qquad k = 1, 2, \ldots, m \tag{12.1}$$

where:
- x is the normalized input variable
- w is the weight of that variable
- i is the input
- b is the bias
- q is the number of input variables
- k and m are the counter and number of NN nodes, respectively, in the hidden layer

In general, the activation functions consist of linear and nonlinear equations. The coefficients associated with the hidden layer are grouped into matrices $w_{i,j,k}$ and $b_{i,k}$. Equation 12.2 can be used as the activation function between the hidden and the output layers (in this equation, f is the transfer function).

$$F_k = f(E_k) \tag{12.2}$$

The output layer computes the weighted sum of the signals provided by the hidden layer, and the associated coefficients are grouped into matrices $W_{o,k}$ and b_o. Using the matrix notation, the network output can be given by Equation 12.3.

$$\text{Out} = \left(\sum_{k=1}^{m} w_{o,k} F_k \right) + b_o \tag{12.3}$$

Network training is the most important part of NN modeling and is carried out using two methods: controllable and uncontrollable training. The most common training algorithm is that of back-propagation. A training algorithm is defined as a procedure that consists of adjusting the coefficients (weights and biases) of a network to minimize the error function between the estimated network outputs and the real outputs.

12.3.2 Optimization of Effective Parameters on Haul Truck Fuel Consumption

Optimization is a part of computational science that represents a very effective way to find the best measurable solution for problems. To solve a given problem, it is important to consider two components: (1) search area and (2) objective function. In the search area, all the possibilities of the solution are considered. The objective function is a mathematical function that associates each point in the search area to a real value, applicable to evaluate all the members of the search area.

Traditional optimization methods are characterized by the stiffness of their mathematical models, making their application limited in representing "real-life" dynamic and complex situations (Selvakumar et al. 2013). Introducing optimization techniques based on artificial intelligence, underpinned by heuristic rulings, have reduced the problem of stiffness. Heuristic rules can be defined as reasonable rules derived from experience and observations of behavior tendencies within a system of analysis.

Using analogies with nature, some heuristic algorithms were proposed during the 1950s by trying to simulate biological phenomena in engineering. Accordingly, these algorithms were termed natural optimization methods. One of the best advantages of using the mentioned algorithms is their random characteristic. Due to their innate flexibility, they have been found to be appropriate to solve all types of problems in engineering (Singh and Rossi 2013; Soleimani et al. 2013; Soofastaei et al. 2016). Rapid advances in computing during the 1980s made the use of these complex algorithms for optimization of functions and processes more practicable when traditional methods were not successful in this area. During the 1990s some new heuristic methods were created by the previously completed algorithms, such as swarm algorithms, simulated annealing, ant colony optimization, and the method used in this study, GAs.

GAs were proposed by Holland (1975) as an abstraction of biological evolution, drawing on ideas from natural evolution and genetics for the design and implementation of robust adaptive systems (Sivanandam and Deepa 2008). Use of the new generation of GAs is comparatively novel in optimization methods. They do not use any derivative information and, therefore, have good chances of escape from local minimums. As a result, their application in practical engineering problems can bring more optimal, or at least more satisfactory, solutions than those obtained by other traditional mathematical methods (Whitley et al. 1990).

GAs are analogous with the evolutionary aspects of natural genetics. From randomly selected "individuals" in any search area, the fitness of the solutions, which is the result of the variable that is to be optimized, is determined subsequently from the "fitness function." The individual that generates the best fitness within the population has the highest chance to return in the next generation with the opportunity to reproduce by crossover with another individual, thus producing decedents with both characteristics. If a GA is developed correctly, the population (a group of possible solutions) will converge to an optimal solution for the proposed problem (Xing and Qu 2013). The processes that have more contribution to the evolution are the crossover, based on the selection and reproduction and the mutation.

GAs have been applied to a diverse range of scientific, engineering, and economic problems due to their potential as optimization techniques for complex functions (Singh and Rossi 2013; Stanković et al. 2013; Tian et al. 2013). There are four significant advantages when using GAs to optimize problems (Yousefi et al. 2013). First GAs do not have many mathematical requirements regarding optimization problems. Second, they can handle many types of objective functions and constraints (i.e., linear or nonlinear) defined in discrete, continuous, or mixed search spaces. Third, the periodicity of evolution operators makes them very efficient at performing global searches (in probability). And finally, they provide us with great flexibility to hybridize with domain-dependent heuristics to allow an efficient implementation for a problem.

It is also important to analyze the influence of certain parameters on the behavior and the performance of the GA, to establish their relationship with the problem necessities and the available resources. The influence of each parameter on algorithm performance depends on the context of the challenge being treated. Thus, determining an optimized group of values to these parameters will depend on a good deal of experimentation and testing. There are a few main parameters in the GA method. Details of these five core parameters are illustrated in Figure 12.3 and tabulated in Table 12.1.

The primary genetic parameters are the size of the population that affects the global performance and the efficiency of the GA, the mutation rate that avoids that a given position remains stationary in value, or that the search becomes essentially random.

12.3.3 The Developed Model

An innovative combined model was introduced to improve the three key effective parameters on the energy consumption of haul trucks. Taking the facets of

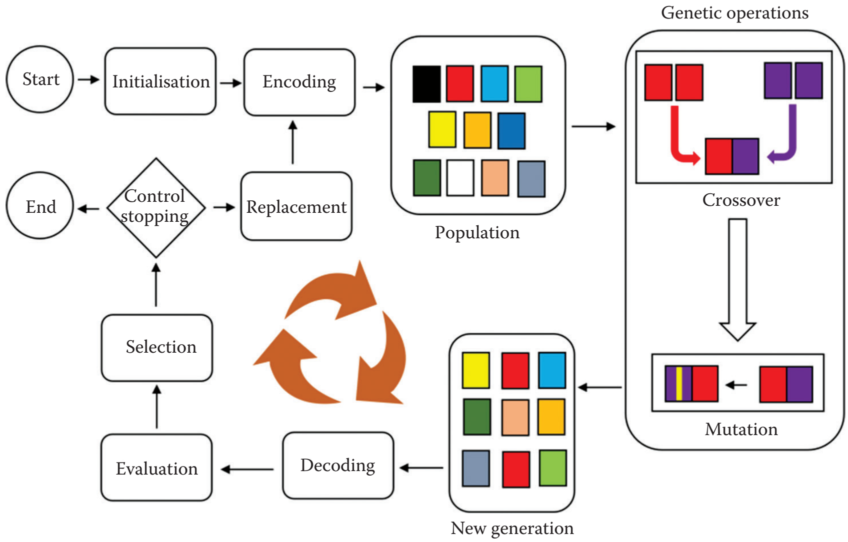

FIGURE 12.3
A simple structure of the genetic algorithm. (From Soofastaei, A., *Development of an Advanced Data Analytics Model to Improve the Energy Efficiency of Haul Trucks in Surface Mines*, PhD thesis, The University of Queensland, School of Mechanical and Mining Engineering, Brisbane, Australia, 2016.)

TABLE 12.1

Genetic Algorithm Parameters

GA Parameter	Details
Fitness function	The primary function for optimization.
Individuals	An individual is any parameter to apply the fitness function. The value of the fitness function for an individual is its score.
Populations and generations	A population is an array of individuals. At each iteration, the GA performs a series of computations on the current population to produce a new population. Each successive population is called a new generation.
Fitness value	The fitness value of an individual is the value of the fitness function calculated for that individual.
Parents and children	To create the next generation, the GA selects certain individuals in the current population, called parents, and uses them to create individuals in the next generation, called children.

the GA approach, in this model P, S, and TR are the individuals and the main function for optimization of the fitness function is fuel consumption. The fitness function was created by an ANN model. This function is a correlation between the fuel consumption of the haul truck, P, S, and TR. After the first step, the completed function goes to the GA phase of the computer code as an input. The developed code starts all GA processes under stopping criteria defined by the model (MSE and R^2). Finally, the improved P, S, and TR will be presented by the model. These optimized parameters can be used to minimize the fuel consumption of haul trucks (Figures 12.4 and 12.5).

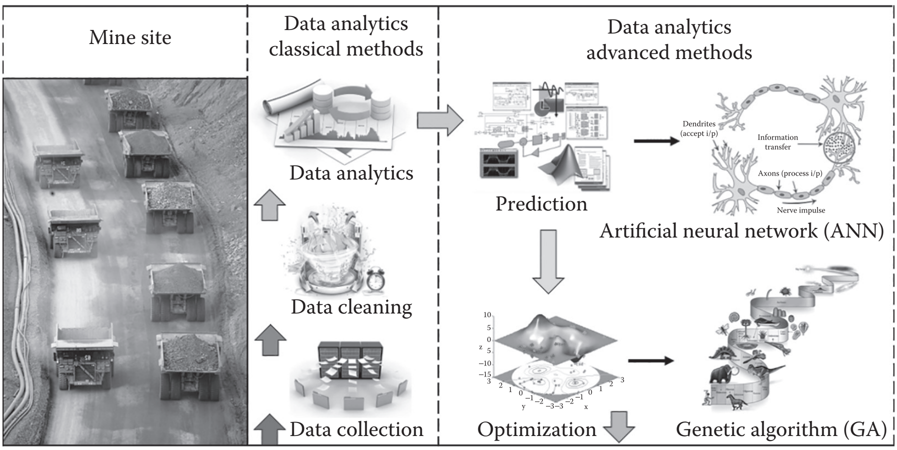

FIGURE 12.4

A schematic of the developed idea to create a combined artificial intelligence model. (From Soofastaei, A., *Development of an Advanced Data Analytics Model to Improve the Energy Efficiency of Haul Trucks in Surface Mines*, PhD thesis, The University of Queensland, School of Mechanical and Mining Engineering, Brisbane, Australia, 2016.)

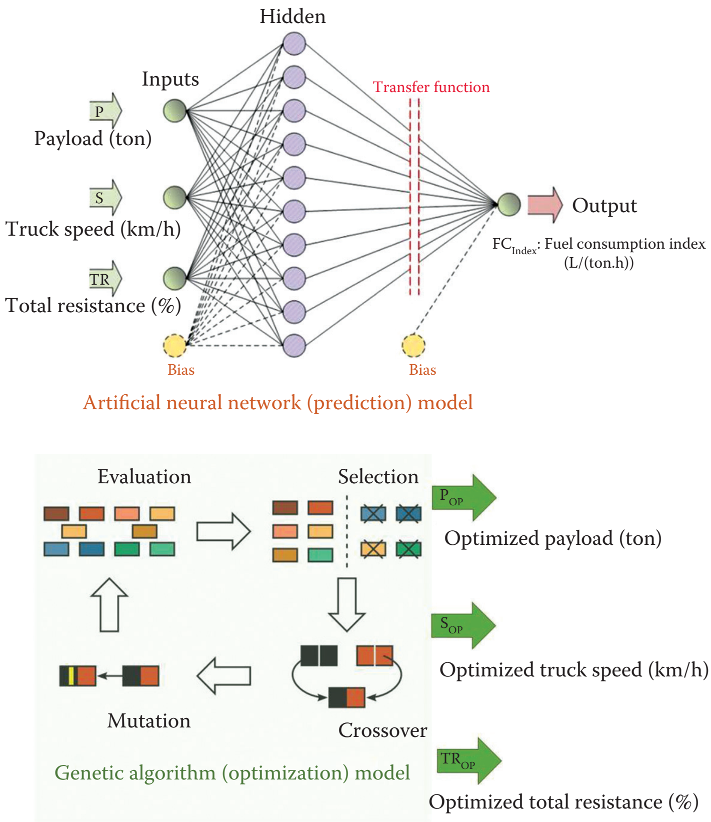

FIGURE 12.5
Details of developed model. (From Soofastaei, A., *Development of an Advanced Data Analytics Model to Improve the Energy Efficiency of Haul Trucks in Surface Mines*, PhD thesis, The University of Queensland, School of Mechanical and Mining Engineering, Brisbane, Australia, 2016.)

12.4 Results

The indicated artificial intelligence model that was developed was then tested against real data taken from some types of popular trucks in four big surface mines in the United States, Canada, and Australia. Some information about these mines and trucks is presented in Table 12.2 (Figures 12.6 through 12.9).

To test the developed networks and validate the developed model, 1,000,000 independent samples collected from four mines were used. As our figures illustrate, the results show good agreement between the actual and estimated values of fuel consumption. Figure 12.10 presents sample values

TABLE 12.2

Case Studies

Case Study	Location	Mine Type	Mine Details	Investigated Truck
Mine 1	Queensland, Australia	Surface coal mine	The mine has coal reserves amounting to 877 million tons of coking coal, one of the largest coal reserves in Asia and the world. It has an annual production capacity of 13 million tons of coal.	CAT 793D
Mine 2	Arizona, United States	Surface copper mine	The mine represents one of the largest copper reserves in the United States and the world, having estimated reserves of 3.2 billion tons of ore grading 0.16% copper.	CAT 777D
Mine 3	Arizona, United States	Surface copper mine	The deposit had estimated reserves (in 2017) of 907 million tons of ore grading 0.26% copper and 0.03% molybdenum.	CAT 775G
Mine 4	Ontario, Canada	Surface gold mine	This mine produced 235,000 ounces of gold in 2016, at the cost of sales of $795 per ounce, and all-in sustaining costs of $839 per ounce. The mine's proven mineral reserves as of December 2016, were 1.6 million ounces of gold.	CAT 785D

for the estimated (using the ANN) and the independent (tested) fuel consumption to highlight the insignificance of the values of the absolute errors in the analysis for the four mines that were studied.

All processes in the developed model, then, certainly work based on the present dataset collected from four large surface mines. The results of using developed model for the selected real-life mines are given in Tables 12.3 through 12.6. They, therefore, could presumably be replicated using the same method for other surface mines.

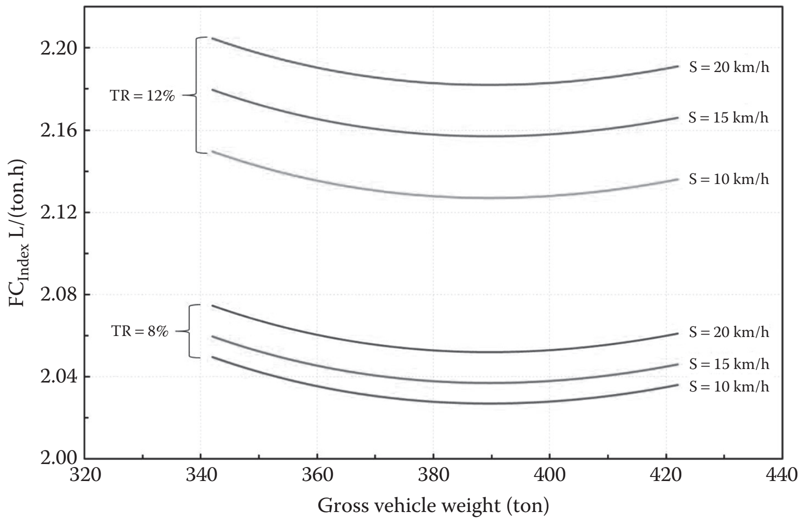

FIGURE 12.6
Correlation between Gross Vehicle Weight, S, TR, and FC_{Index} based on the developed ANN model for CAT 793D. All data were collected from a surface coal mine located in Central Queensland, Australia (Mine 1).

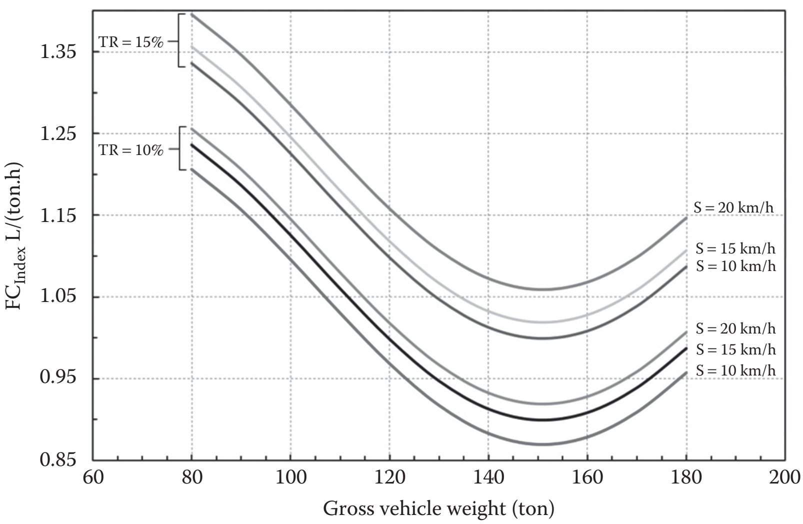

FIGURE 12.7
Correlation between GVW, S, TR, and FC_{Index} based on the developed ANN model for CAT 777D. All data were collected from a surface copper mine located in Arizona, United States (Mine 2).

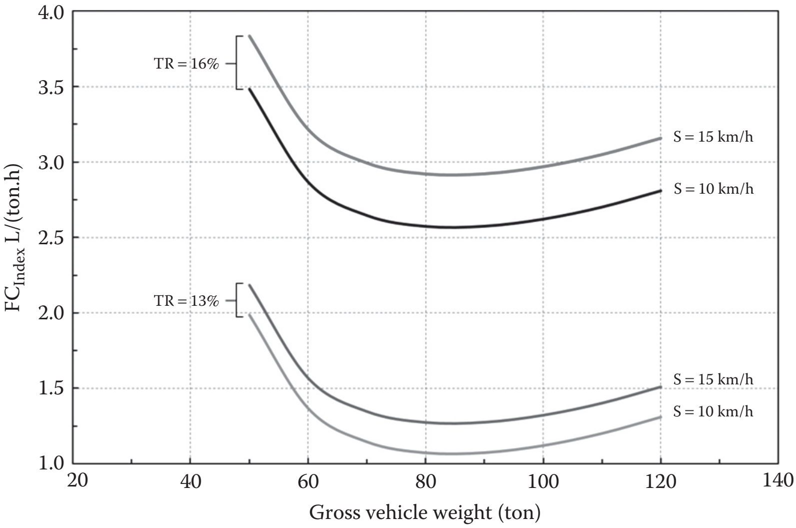

FIGURE 12.8
Correlation between GVW, S, TR, and FC_{Index} based on the developed ANN model for CAT 775G. All data were collected from a surface copper mine located in Arizona, United States (Mine 3).

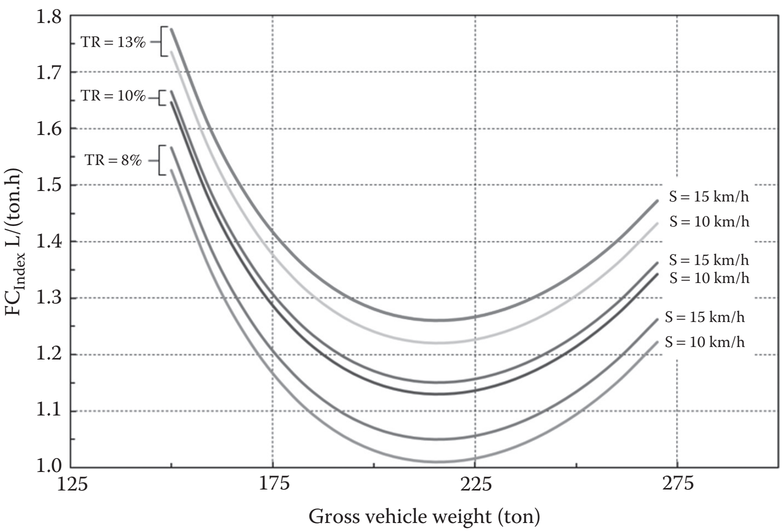

FIGURE 12.9
Correlation between GVW, S, TR, and FC_{Index} based on the developed ANN model for CAT 785D. All data were collected from a surface coal mine located in Ontario, Canada (Mine 4).

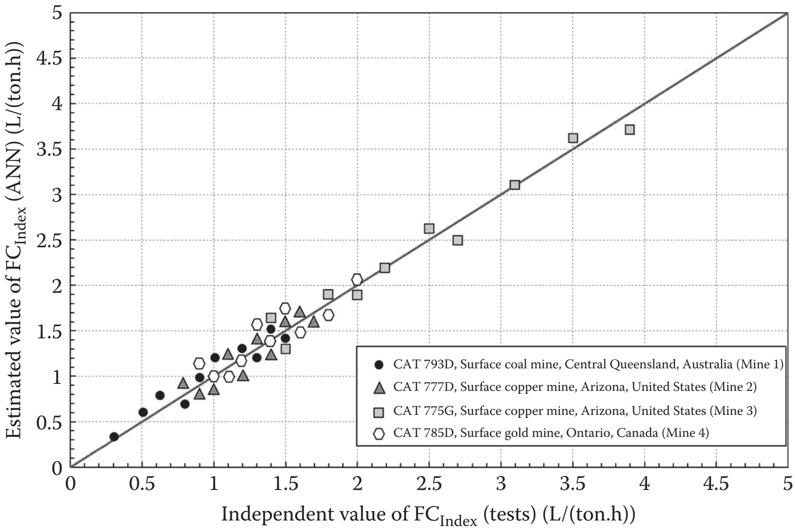

FIGURE 12.10
Sample values for the estimated (using the ANN) and the independent (tested) fuel consumption index.

TABLE 12.3

The Range of Normal Values and Optimized Range of Variables by GA Model to Minimize Fuel Consumption by Haul Trucks (Caterpillar 793D in Mine 1)

	Normal Values		Optimized Values	
Variables	**Minimum**	**Maximum**	**Minimum**	**Maximum**
Gross vehicle weight (ton)	340	430	380	400
Total resistance (%)	8	12	8	9
Truck speed (km/hr)	10	20	10	15

TABLE 12.4

The Range of Normal Values and Optimized Range of Variables by GA Model to Minimize Fuel Consumption by Haul Trucks (Caterpillar 777D in Mine 2)

	Normal Values		Optimized Values	
Variables	**Minimum**	**Maximum**	**Minimum**	**Maximum**
Gross vehicle weight (ton)	80	180	140	160
Total resistance (%)	10	15	10	11
Truck speed (km/hr)	10	20	10	12

TABLE 12.5

The Range of Normal Values and Optimized Range of Variables by GA Model to Minimize Fuel Consumption by Haul Trucks (Caterpillar 775G in Mine 3)

	Normal Values		Optimized Values	
Variables	**Minimum**	**Maximum**	**Minimum**	**Maximum**
Gross vehicle weight (ton)	50	120	70	90
Total resistance (%)	13	26	13	14
Truck speed (km/hr)	10	15	10	13

TABLE 12.6

The Range of Normal Values and Optimized Range of Variables by GA Model to Minimize Fuel Consumption by Haul Trucks (Caterpillar 785D in Mine 4)

	Normal Values		Optimized Values	
Variables	**Minimum**	**Maximum**	**Minimum**	**Maximum**
Gross vehicle weight (ton)	150	275	200	225
Total resistance (%)	8	13	8	9
Truck speed (km/hr)	10	15	10	12

12.5 Conclusions

The aim of this chapter was to formulate an advanced data analytics model capable of improving haul truck fuel consumption based on the relationship between P, S, and TR. From the available "real-life" datasets obtained from surface mining operations, this relationship is extremely complex to disect using traditional analysis. Therefore, an artificial intelligence method was adopted to create a reliable model to analyze the problem.

The first element of this method was to utilize an ANN model to establish a correlation between P, S, and TR with fuel consumption. The results of this correlation showed that fuel consumption has a nonlinear relationship with the investigated parameters. The ANN was then trained and tested using the collected real mine site dataset, with there being good agreement between the actual and estimated values of fuel consumption. Building upon this material, a GA model was developed for considering the optimization of effective parameters on fuel consumption in haulage trucks, which in turn could maximize the energy efficiency in haulage operations.

From these amalgamated models, the range of all studied effective parameters on fuel consumption of haul trucks were optimized, and the best values of P, S, and TR to minimize fuel consumption index (FC_{Index}) were highlighted. The developed model was applied to analyze data for four big coal, copper, and gold surface mines in the United States, Canada, and Australia.

References

Allison, B., A. Shamim, M. Caitlin, and P. Pha. 2016. *Australian Energy Update 2016*. Canberra, Australia: Office of the chief economist, Department of Industry Innovation and Science.

Beigmoradi, S., H. Hajabdollahi, and A. Ramezani. 2014. Multi-objective aeroacoustic optimisation of rear end in a simplified car model by using hybrid robust parameter design, artificial neural networks and genetic algorithm methods. *Computers and Fluids* 90:123–132.

Davenport, T. H., J. G. Harris, and R. Morison. 2010. *Analytics at Work: Smarter Decisions, Better Results*. Vol. 3. New York: Harvard Business Press, p. 93.

DOE. 2012. *Mining Industry Energy Bandwidth Study*. Washington, DC: Department of Energy, USA Government.

EEO. 2010. *Energy-Mass Balance: Mining*. Canberra, Australia: Australian Government, Department of Resources Energy and Tourism.

EEO. 2012. *First Opportunities in Depth: The Mining Industry*. Canberra, Australia: Australian Government, Department of Resources Energy and Tourism.

Hammond, A. 2012. Development artificial neural network model to study the influence of oxidation process and zinc electroplating on fatigue life of grey cast iron. *International Journal of Mechanical and Mechatronics Engineering* 12(5):128–136.

Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. Vol. 2. Washington, DC: University of Michigan Press, pp. 183–195.

Rodriguez, J. A., Y. El Hamzaoui, J. A. Hernandez, J. C. García, J. E. Flores, and A. L. Tejeda. 2013. The use of the artificial neural network (ANN) for modelling the useful life of the failure assessment in blades of steam turbines. *Engineering Failure Analysis* 35:562–575.

Selvakumar, S., K. P. Arulshri, and K. P. Padmanaban. 2013. Machining fixture layout optimisation using a genetic algorithm and artificial neural network. *International Journal of Manufacturing Research* 8(2):171–195.

Singh, A., and A. Rossi. 2013. A genetic algorithm based exact approach for lifetime maximisation of directional sensor networks. *Ad Hoc Networks* 11(3):1006–1021.

Sivanandam, S. N., and S. N. Deepa. 2008. *Introduction to Genetic Algorithms*. Vol. 3. New Delhi, India: Springer, p. 453.

Soleimani, H., M. Seyyed-Esfahani, and M. A. Shirazi. 2013. Designing and planning a multi-echelon multi-period multi-product closed-loop supply chain utilising genetic algorithm. *The International Journal of Advanced Manufacturing Technology* 68(1–4):917–931.

Soofastaei, A., S. M. Aminossadati, M. M. Arefi, and M. S. Kizil. 2016. Development of a multi-layer perceptron artificial neural network model to determine haul trucks energy consumption. *International Journal of Mining Science and Technology* 26(2):285–293.

Soofastaei, A., S. M. Aminossadati, M. S. Kizil, and P. Knights. 2016. Reducing fuel consumption of haul trucks in surface mines using genetic algorithm. *Applied Soft Computing.*

Soofastaei, A., and J. Davis. 2016. Advanced data analytic: A new competitive advantage to increase energy efficiency in surface mines. *Australian Resources and Investment* 1(1):68–69.

Stanković, L. J., V. Popović-Bugarin, and F. Radenović. 2013. Genetic algorithm for rigid body reconstruction after micro-Doppler removal in the radar imaging analysis. *Signal Processing* 93(7):1921–1932.

Tian, J., M. Gao, and Y. Li. 2013. Urban logistics demand forecasting based on regression support vector machine optimised by chaos genetic algorithm. *Advances in Information Sciences and Service Sciences* 5(7):471–478.

Whitley, D., T. Starkweather, and C. Bogart. 1990. Genetic algorithms and neural networks: Optimising connections and connectivity. *Parallel Computing* 14(3):347–361.

Xing, H., and R. Qu. 2013. A nondominated sorting genetic algorithm for bi-objective network coding based multicast routeing problems. *Information Sciences* 233(3):36–53.

Yousefi, T., A. Karami, E. Rezaei, and E. Ghanbari. 2013. Optimisation of the free convection from a vertical array of isothermal horizontal elliptic cylinders via a genetic algorithm. *Journal of Engineering Physics and Thermophysics* 86(2):424–430.