# Energy usage optimisation of heavy haul freight trains

A. Bogaers* and N. Botha

*Modelling and Digital Science, CSIR*

abogaers@csir.co.za

**ABSTRACT**

The aim of this paper is to outline a preliminary investigation into an energy optimisation model with the aim of eventually being incorporated into a real-time driver assist program. A significant portion of the South African economy is sustained by a large and extensive mining industry. Integral to sustaining the mining economy is the transport of mined raw material via freight rail over large distances, typically from the mines to central distribution or processing centres. Due to the heavy tonnage and long distances an enormous amount of energy is required. Using classical mechanics, an energy usage model for typical freight trains based on available tractive power and typical rolling stock resistances (curvature, friction, gravity etc) is considered. The aim is to find the optimal velocity profile over the full distance of the intended track for which the total locomotive energy usage is reduced. We will illustrate the robustness of the proposed model to predict optimal operational velocity profiles for a number of interesting scenarios, many supported by similarly related analytical studies.

**KEYWORDS**

Energy, Optimisation, Freight rail

**INTRODUCTION**

The large amount of energy used in the operation of heavy haul freight trains have led to much research into energy efficient operation [1, 2, 3, 4, 5, 6]. Some of the earliest literature on optimal train control dates back to the 1960's when Ishikawa [3] determined the optimal control for a train traveling along a flat straight track. It was found that by maximising the Hamiltonian of the control problem (assuming that the control variables follow linear relationships with simple bounds) it can analytically be shown that the optimal energy control for a train on a flat track starts with a period of maximum power (maximum acceleration) followed by a period of partial power (cruising), a period of no power or braking (coasting) and finally a period of maximum braking, as illustrated in Figure 1(a). As the trip time is reduced, the cruising phase will decrease until it disappears, leaving only the full power phase, the coasting phase followed by maximum braking, as illustrated in Figure 1(b).
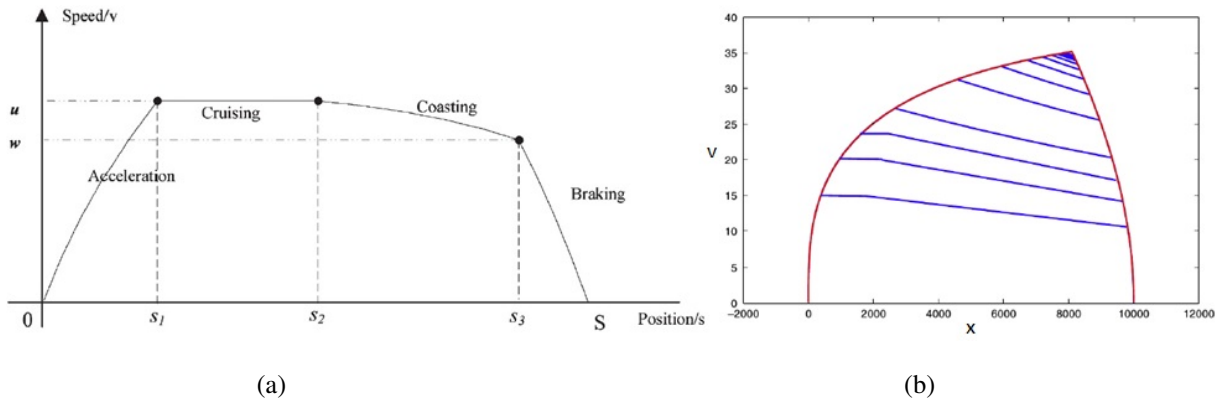
(a)



(b)

**Figure 1.** (a) An illustration of the various phases, or switching points between the different driving phases for optimal control on a flat track. (b) The analytical profile for various control phases on a flat track at different times for optimal train control, images taken from Su et. al. [7] and Howlett et. al. [2] respectively.

In general, minimising the total energy for a trip within a given time frame can be fairly complex. The terrain typically consists of a number of inclines, descents and curves. Furthermore, for safety reasons, either in regions of sharp corners, or steep track gradients, speed limits are imposed. To further complicate matters, locomotive technology does not produce constant tractive power or braking force; both traction and braking efficiencies are typically functions of velocity, illustrated in Figure 2 for an E14 locomotive.
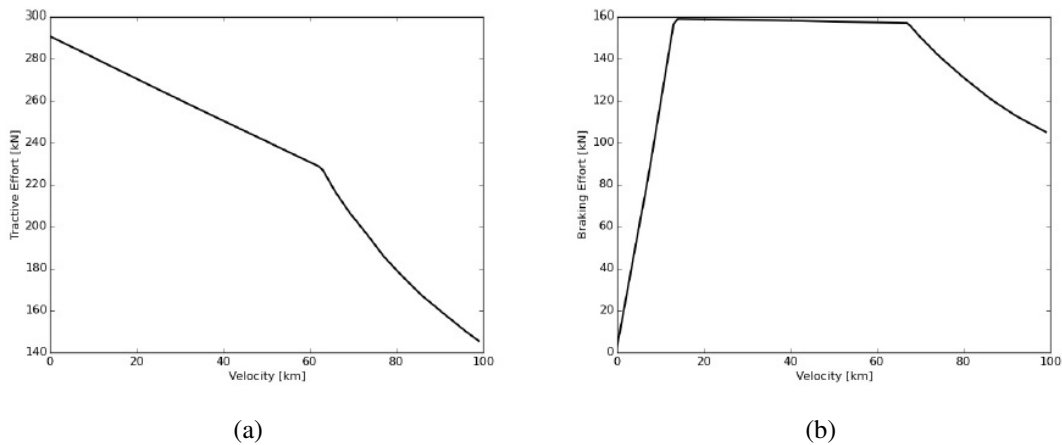


(a)



(b)

**Figure 2.** Maximum available (a) tractive and (b) braking effort for an E14 locomotive as a function of velocity, images produced from Lombard [8].

To enable application of optimal train control in practice, the analytical studies have been extended to include speed limits, variable traction efficiencies and variable gradients [2]. These analytical studies are limited to lumped mass models. While evidence tends to suggest that by minimising energy one naturally improves the overall train coupling forces, to date there has been no formal attempt to optimise both in train coupling forces and energy usage. In an attempt to improve the cost effectiveness and meet increasing scheduling constraints heavy haul operators are continuously increasing the length of trains. This leads to higher inter-train coupling forces, increased wear and an increased risk of derailment. Longer trains further pose the possibility of a single train simultaneously experiencing a multitude of different track conditions. For example, the lead portion of a

2

train may be traveling downhill while the trailing end of the train is still traveling uphill, see figure 3.

Minimising energy usage, while simultaneously including coupling forces, implies that analytical methods are no longer strictly suitable. To adopt such a multi-parameter optimisation problem would require the use of non-linear optimisation methods. Some early attempts have focused on using evolutionary optimisation methods. Consider for example the work presented by McClanachan et al. [9] who used genetic algorithms (GA) to optimise for the optimal speed profile. The study focused on a preliminary investigation illustrating that GA can provide speed profiles closely matching those predicted by analytical methods. The results appear sufficiently promising to warrant further investigation by including coupling forces.

GA algorithms are general purpose directed random optimisation algorithms which employ no special knowledge of the problem, to determine optimal solutions. They typically require several tens of thousands of function evaluations, similar to other evolutionary methods and related methods, such as ant colony optimisation, particle swarm optimisation or cuckoo search; more information on these methods can be found in a review by Wahab et. al. [10]. In turn these evolutionary methods can be extremely expensive (orders of several hundreds to thousands of CPU hours depending on the computational complexity of the function evaluations, and the profile of the design space). Due to these computational constraints it is currently infeasible to converge to machine precision, often resulting in highly sporadic, unrealistic results.

In this paper we aim to pose an argument for gradient based optimisation methods. Gradient based methods are classified as non-linear optimisation methods, and utilise information pertaining to the gradients and gradient landscape of a cost function. They can therefore be considered as directed search algorithms and if the gradients associated with a cost function are available, typically require significantly fewer iterations to reach convergence than evolutionary methods.
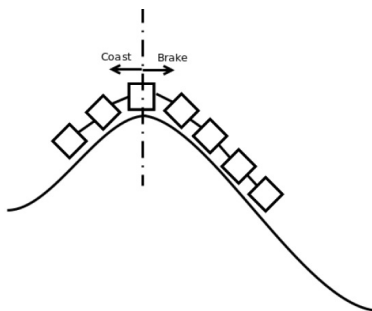


**Figure 3.** Long train on a hill experiencing different forces.

## PROBLEM FORMULATION

The general non-linear optimisation problem can be posed as [minimise $f(x)$], subject to a number of equalities and inequality constraints, namely $b(x) \geq 0$ and $c(x) = 0$.

The optimisation problem can be formulated as finding the sequence and magnitude of locomotive power which completes a given trip, over a given terrain, within the specified time limit, while minimising the overall energy used. We define energy as the integral of locomotive traction, $f(x)$

$$E(x) = \int f(x)dx, \tag{1}$$

where $x$ is the train location. The problem is subject to a number of constraints which can be summarised as

- Total travel time

3

- Speed limits

- Maximum available power (traction and braking), which varies with velocity

- Maximum rail-wheel adhesion (prevent slipping)

- Coupling forces.

The locomotive traction forces can be defined by

$$F_{\text{Loc}} + F_r + F_{cv} + m_{\text{tot}}g_a = m_{\text{tot}}a, \tag{2}$$

where $F_{\text{Loc}}$ is the force exerted by the locomotives, $m_{\text{tot}}$ represents the total mass of all the locomotives and wagons within the train [11]. $a$ represents the acceleration, $F_r$ the lumping of all the rolling resistances, $g_a$ the force due to gradient changes and $F_{cv}$ represents the resistances due to curvature. The rolling resistance term $F_r$ includes the resistance due to airflow friction losses and wheel rail frictional losses. Typically $F_r$ takes the following form

$$F_r = m(c_0 + c_1v + c_2v^2), \tag{3}$$

where the constants $c_{0,1,2}$ are typically empirically/experimentally determined based on the given locomotive and wagon properties as well as the track configuration. Slope and curvature are illustratively shown in Figure 4 and the associated losses/forces due to inclines or curvature can be approximated as $g_a = \sin(\theta)$ where $\theta$ is the angle of the slope and curve resistance as $F_{cv} = 0.004mD$ where $D = 0.5d_{wheelbase}/R$. R here represents the radius of curvature of the track curve.
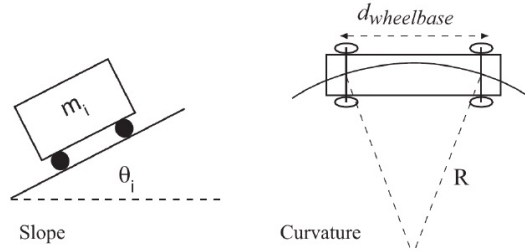


**Figure 4.** Simplistic illustration of losses due to gradients (slopes) and curvature [14].

Most gradient based methods require smooth, twice differentiable and continuous function definitions. Problem formulation is therefore of critical importance. Our design variables are therefore chosen to be velocity. While counter intuitive, this results in a relatively well behaved problem formulation. The track is discretised into a number of $n$ segments. At each discrete location $x_i$, we assume we have available the track height $h_i$, and for each segment between two respective positions we know the track radius of curvature $R_j$ . At $x_0$ and $x_n$ the starting and final velocities $v_0$ and $v_n$ are known. The unknown velocities $\{v_1 : v_{n-1}\}$ are therefore the velocities to be solved for, or the so called design variables, $\boldsymbol{dV}$ , which minimises the total energy usage. Therefore, given a series of design variables $\boldsymbol{dV} = \{dV_1, dV_2, ..., dV_{n-1}\}$ , our known velocities within a given iteration are given by

$$v = \{v_0, dV_1, dV_2, ..., dV_n, v_n\}. \tag{4}$$

The cost function can be computed for each section of the discredited track $(\Delta x_j)$

$$\Delta x_j = x_{i+1} - x_i \text{ for } i = 1, 2, ..., n \qquad v_{average,j} = \frac{v_{i+1} + v_i}{2},$$

$$\Delta t_j = \frac{x_{i+1} - x_i}{v_{average,j}} \qquad\qquad a_j = \frac{v_{i+1} - v_i}{\Delta t_j},$$

$$\Delta h_j = \frac{h_{i+1} + h_i}{2} \qquad\qquad g_{a,j} = \sin(\arctan(\frac{\Delta h_j}{\Delta x_j}))9.81,$$

$$F_{r,j} = m(c_0 + c_1 v_{i+1} + c_2 v_{i+1}^2) \qquad F_{loc,j} = m_{total}(a_j + F_{r,j} + g_{a,j} + F_{cv,j}). \qquad (5)$$

Given a guess for velocities at each section of the track we can compute the locomotive traction and breaking forces $\boldsymbol{F}_{loc}$ needed to achieve the given velocities $\boldsymbol{dV}$.

$\boldsymbol{F}_{loc}$ would be positive if traction is applied and negative if braking is present. The convergence rates of gradient based methods reduce when handling highly skewed landscapes. We therefore normalise the forcing function to be between $\{0, 1\}$. The force function and braking function b is then defined as

$$f_i = \begin{cases} F_{loc,j}/P & \text{if} & F_{loc,j} > 0, \\ 0 & \text{if} & F_{loc,j} \leq 0, \end{cases} \qquad (6)$$

and

$$b_i = \begin{cases} F_{loc,j}/B & \text{if} & F_{loc,j} \leq 0, \\ 0 & \text{if} & F_{loc,j} < 0, \end{cases} \qquad (7)$$

where $P$ is the maximum available traction force to the locomotive and $B$ is the maximum available braking force. It is perhaps important to note that in typical trains all wagons also apply braking forces. For long trains when using pneumatic braking, the braking signal may take a long time to propagate along the length of the train, which is currently being omitted. The cost function to be optimised then becomes

$$\text{cost function} = \sum_{j=1}^{n-1} f_j \Delta x_j. \qquad (8)$$

Speed limits are handled as upper and lower bounds on the set of design variables $\boldsymbol{dV}$ and therefore can be handled by most optimisation algorithms.

Speed limits on tracks are typically imposed for safety reasons either in regions of sharp corners or other environmental factors which requires due diligence. In the absence of speed limits, it is possible to obtain estimates of maximum cornering speeds via simple summation of centrifugal forces and gravitational forces acting on a wagon running on a superelevated track with cant angle $\theta$, see figure 5 [12]. The critical, or maximum velocity for a given wagon can be found when the torque due to centrifugal force equals the torque due to the gravitational force, and is described as

$$v_{\max} = \sqrt{\frac{gRl \sin^2 \theta}{\cos \theta \left[h \cos \theta + l \sin \theta\right]}}. \qquad (9)$$

$g$ represents the gravitational acceleration, $R$ the radius of curvature of the section of track, $l$ is half the gauge of the track and $h$ the height of the center of gravity of the wagon/locomotive above the rail.
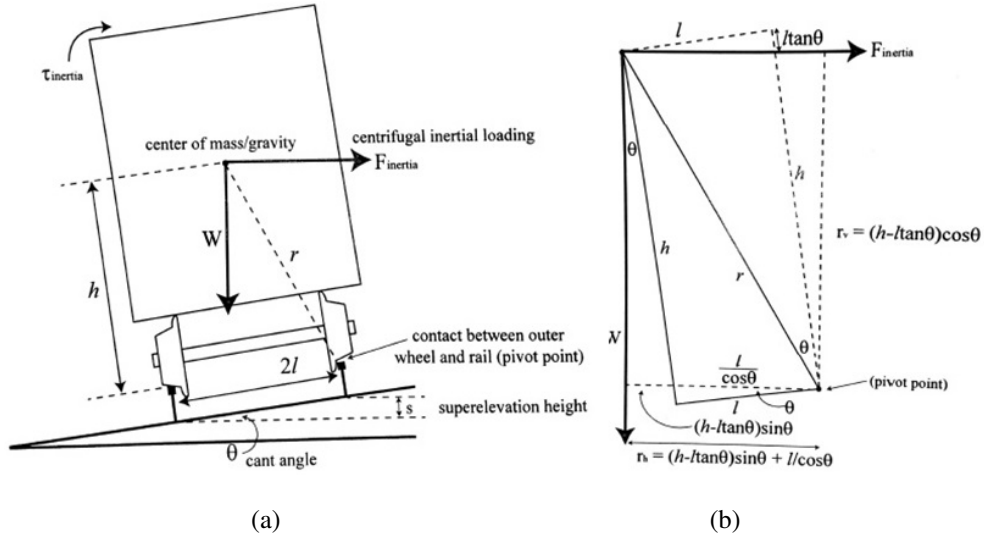
**Figure 5.** (a) Forces acting on a wagon on a super-elevated track used to compute the critical cornering speeds and (b) the schematic representation [12].

The remaining constraints on the system, namely the maximum allowable time constraint, maximum force and braking constraint would mathematically be posed as

$$
\begin{aligned}
&t_{\text{total}} \leq t_{\text{max}}, \\
&f_j \leq f_{\text{max}}(v), \qquad b_j \leq b_{\text{max}}(v), \qquad \text{for } j = 1, 2, ..., n-1, \\
&f_j > 0, \qquad\qquad b_j > 0.
\end{aligned}
\tag{10}
$$

$\mu$ represents the adhesion required to apply the prescribed $F_{\text{Loc}}$ without slipping. The adhesion is typically dependent on the track conditions where dry track conditions will have more adhesion available than wet tracks.

The traction and braking constraints are posed as functions of velocities (i.e. $f_{\text{max}}(v)$) since typical locomotive traction efficiencies are dependent on velocity and is not constant (as often assumed in literature). Alternative cost functions can be defined to optimise for minimum trip time, which becomes

$$
\text{minimise } T(x) = \sum_{j=1}^{n-1} t_j.
\tag{11}
$$

For a multi-object optimisation to minimise for a combination of trip time and energy usage the cost function will become

$$
\text{minimise } ET(x) = \omega \left( \int f(x)dx \right) + (1 - \omega) \sum_{j=1}^{n-1} t_j,
\tag{12}
$$

where $\omega$ is a weighting factor to give preference to either energy or time. Typically, because energy used and travel time length scales differ, it would be advisable to include an upfront scaling factor to provide some equality between energy and time. Optimising for minimum trip time and the multi-parameter energy and time optimisation are both potentially useful for trip scheduling.

Due to the large amount of constraints, we made use of sequential quadratic programming (SQP). SQP is useful for solving nonlinear optimisation problems, where the problem is transformed into a succession

of quadratic optimisation problems, with a quadratic objective and the linear constraint system consisting of equalities and inequalities. A detailed description of SQP can be found in the work by Bonnans et. al. [13].

## RESULTS

To highlight the ability of the proposed energy optimisation model on a number of scenarios are tested. It has previously been shown [3] that the optimal energy profile can be summarised (following Figure 1) as a period of acceleration at maximum power, followed by a period of partial power with no braking (cruising), a period of no power and no braking (coasting) and finally maximum braking. As the trip time is decreased the period of partial power disappears, thereby only having three phases of maximum power, coasting and finally maximum braking. And lastly, for shortest trip time there is only a maximum power phase, followed by maximum braking.

To test whether our non-linear optimisation formulation can re-produce these optimal energy curves, we analyse a train with 1 locomotive towing 20 wagons loaded with 100 tonnes each. The locomotive has a constant power output of 400kN and constant braking tractive effort of 500 kN. The total trip is 10 km and is discretised into 100 segments. The optimisation algorithm reproduced the expected analytical results, shown in shown in Figure 6 for different trip times. These similar results verify that our model applied to train on a flat track compares well to previous authors work. Some of the constants and values used are given in Table 1. [14]
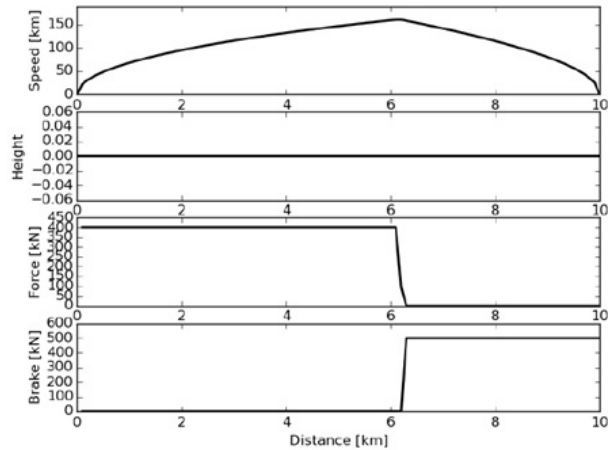
Table 1: Values for constants.

| Constant | Assigned value |
|----------|----------------|
| $c_0$ | $7.6558 \times 10^{-3}$ |
| $c_1$ | $1.08 \times 10^{-4}$ |
| $c_2$ | $1.4915 \times 10^{-5}$. |
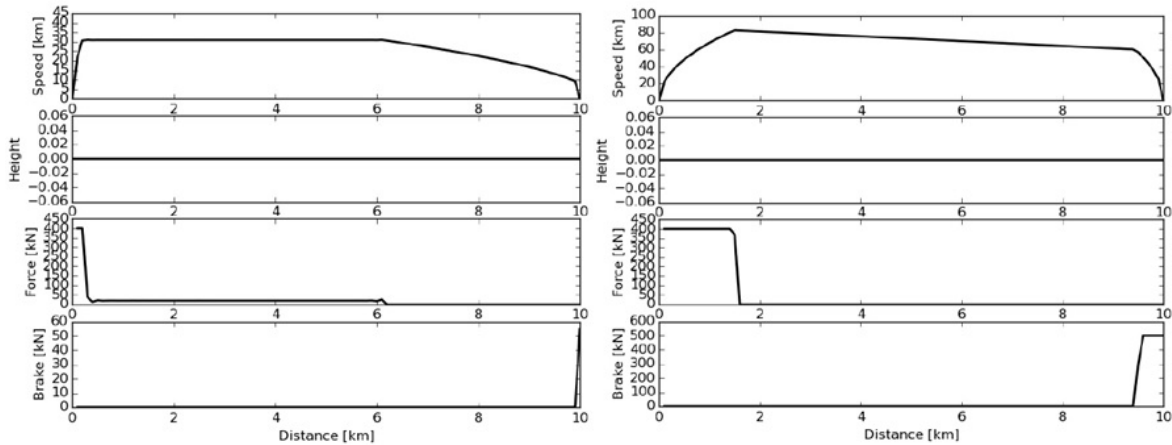


(a)



(b)

(c)

**Figure 6.** The optimised trip with a single locomotive pulling 2000 tonnes on a flat terrain for a time limit of (a) 25 min, (b) 10 min and (c) minimum trip time (7.3 min).

In Figures 6-8 the top subplot shows the train speed versus distance (the position on the track) with the track elevation versus distance in the second subplot. The third and fourth subplot shows the traction and braking force respectively, versus distance.

An important distinction is that real locomotives do not have constant tractive and braking efficiencies. For an E14 locomotive (using the maximum traction and braking curves as a function of velocity) we obtain a velocity and tractive effort profile, as shown in Figure 7, for minimum trip time, using the same track and number of wagons as used to generate the results in Figure 6. Following the expected profile of maximum power followed by maximum breaking, the optimisation algorithm successfully limits the maximum available tractive effort as a function of velocity. When comparing the results using the constant (Figure 6(c)) and velocity dependent (Figure 7) efficiencies there is a clear difference between the application of power and braking to get the same velocity profile.
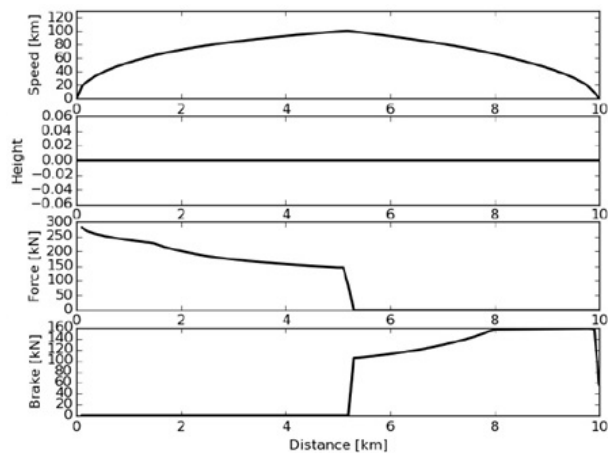


**Figure 7.** The optimised trip with a single locomotive pulling 2000 tonnes on a flat terrain for a E14 locomotive with variable maximum available traction and braking.

Finally a test track with steep gradients is considered, since typical tracks have numerous hills and descents. Following the work by Howlett et. al. [1, 2], it is expected that the optimal energy profile across a small hill would be to maintain cruising velocity by applying partial power to get over the hill. The gradient based optimisation can accurately predict this behavior as well. Consider here a 16km track with a 20m height change between 8 to 10km. This relates to a 1:100 steep banking gradient. Here we analyse a train with a locomotive with maximum constant power of 400kN, towing 2000 tonnes, where the optimal velocity profile for a 30 minute trip is shown in Figure 8(a). Notice how the cruising velocity is maintained by the application of partial power.

A really interesting problem occurs the moment an incline is too steep given the available power. In order to overcome the incline gradient requires application of power prior to the steep gradient, allowing momentum to carry the mass of the train over the incline. To illustrate the capability of the gradient based optimisation algorithm to cope with this complex situation we use the same track with the locomotive now towing 6000 tonnes. The results for this simulation is shown in Figure 8(b). The application of power prior to the hill is clear in Figure 8(b) as well as the application of more tractive power when compared to the tractive power application in Figure 8(a).
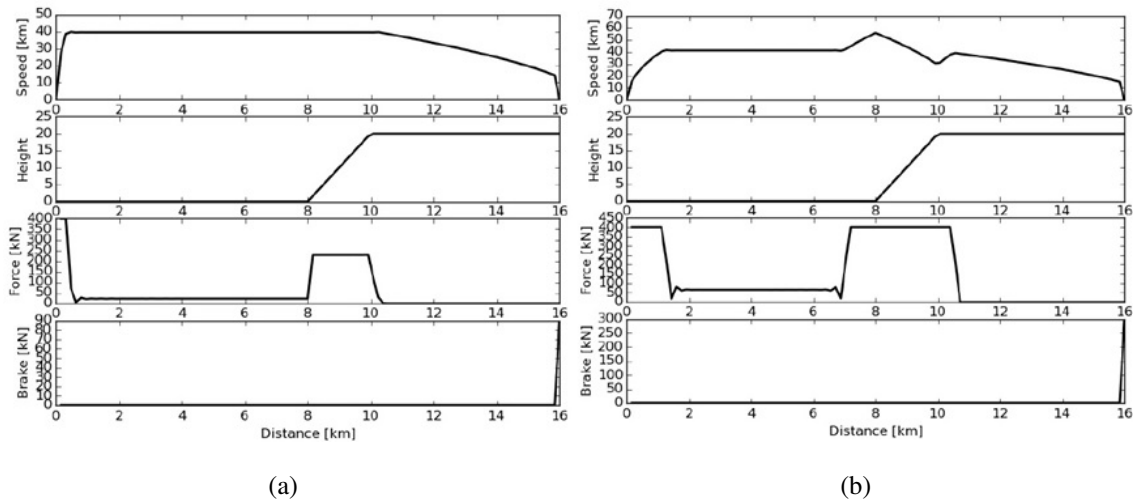


|     |     |
| :-: | :-: |
| (a) | (b) |

**Figure 8.** (a) Inclined track with train pulling 2000 ton and (b) pulling 6000 ton.

## CONCLUSION

In this paper we outline a preliminary investigation into a gradient based energy optimisation model with the aim of eventually being incorporated into a real-time driver assist program.

It was established that for a flat terrain the optimal control strategy starts with maximum acceleration at maximum power, followed by a period of cruising, a period of coasting and finally maximum braking. The cruising period disappears with decreasing total trip time, so that there is only three phases: maximum power, coasting and maximum braking. For the shortest trip time there is only a maximum power phase, followed by maximum braking. When the terrain has steep gradients the application of additional power is necessary, while for very steep gradients the application of additional power may be necessary before the gradient. The numerical model mirrors the behaviour from literature for each scenario, which verify that our model applied to a train on a flat track as well as steep gradients, correlate well with the work of previous authors.

In this study the train has been described mathematically as a lump mass. This lump mass model has laid the groundwork for further investigation into integration of the coupling forces between the different locomotives

and wagons by using a longitudinal dynamics model.

## REFERENCES

1    Howlett, P, "Optimal strategies for the control of a train", Automatica, Vol. 32. No. 4, pp. 519-532, 1996.

2    Howlett, P.G. et. al., "Local energy minimization in optimal train control", Automatica, Vol. 45. No. 11, pp. 2692-2698, 2009.

3    Ichikawa, K., "Application of optimization theory for bounded state variable problems to the operation of train", Bulletin of JSME, Vol. 11. No. 47, pp. 857-865, 1968.

4    Jiaxin, C. and Howlett, P., "A note on the calculation of optimal strategies for the minimization of fuel consumption in the control of trains", Automatic Control, IEEE Transactions on, Vol. 38.(11), pp. 1730-1734, 1993.

5    Khmelnitsky, E., "On an optimal control problem of train operation", Automatic Control, IEEE Transactions on, Vol. 45. No. 7, pp. 1257-1266, 2000.

6    Liu, R.R. and Golovitcher, I.M., "Energy-efficient operation of rail vehicles", Transportation Research Part A: Policy and Practice, Vol. 37. No. 10, pp 917-932, 2003.

7    Su, S. et. al., "A Subway Train Timetable Optimization Approach Based on Energy-Efficient Operation Strategy", IEEE Transactions on Intelligent Transportation Systems, Vol. 14. No. 2, 2013.

8    Lombard,P.J., "Lecture notes in train service planning. Introductory course on multi-disciplinary concepts in railway engineering", University of Pretoria, July 2015.

9    McClanachan, M. and Cole, C., "Optimising the operation of a heavy haul train", In CORE 2012: Global Perspectives; Conference on railway engineering, 10-12 September 2012, Brisbane, Australia, page 569. Engineers Australia, 2012.

10   Wahab, A.b. et. al., "A Comprehensive Review of Swarm Optimization Algorithms", PLoS One, Vol. 10. No. 5, e0122827.

11   Cole, C, "Chapter 9 Longitudinal Train Dynamics", Editor Iwnicki, S., Handbook of Railway Vehicle Dynamics (pp. 239-277), CRC Press, 2006.

12   Moody, J.C., "Critical Speed Analysis of Railcars and Wheelsets on Curved and Straight tracks", Bates Collage, Honors Thesis, Paper 107, 2014.

13   Bonnans, J.F., et al., "Numerical optimization: Theoretical and practical aspects", Springer Science & Business Media, 2006.

14   Chou, M. and Xia, X., "Optimal cruise control of heavy-haul trains equipped with electronically controlled pneumatic brake systems", Control Engineering Practice, Vol. 15. No. 5., pp. 511-519, 2007.