Estimating energy savings from a train driving advice system

A. J. Fisher[1] A. V. Metcalfe[2] P. J. Pudney[3]

(Received 4 March 2022; revised 4 July 2022)

Abstract

TTG Energymiser is an in-cab system that provides real-time driving advice to train drivers with the aim of reducing energy use subject to meeting the train schedule. A survey of the efficacy of Energymiser has been undertaken, to provide evidence for marketing claims. Results from 23 different trials are analysed, where 16 of the trials were on passenger routes and 7 were on freight routes. Each trial consists of many trips, with Energymiser activated for around half, and yields an estimate of the change in energy use when Energymiser is used. A Bayesian hierarchical model is fitted to the 16 estimates from passenger routes and provides an estimate of the mean saving and the standard deviation of individual trials about the mean. The mean saving is 7.2% and the standard deviation of individual trials is estimated as 3.3%. The corresponding mean and standard deviation for freight routes are 8.4% and 5.8%, respectively.

DOI:10.21914/anziamj.v63.17209, Published 2022-07-17, as part of the Proceedings of the 15th Biennial Engineering Mathematics and Applications Conference. ISSN 1445-8810.

Contents

1 Introduction

TTG Energymiser is an in-cab system that provides real-time driving advice to train drivers to help them stay on time and reduce energy use. It was developed by researchers at the University of South Australia and Sydney company TTG Transportation Technology, and is used by railways around the world.

The calculation of the optimal driving strategy takes into account traction and braking characteristics of the train, rolling resistance and aerodynamic drag, track gradients and curves, and scheduled arrival times at stops and intermediate locations.

The principles used to determine the optimal control strategy are described in a pair of papers by Albrecht et al. [2, 1], where Pontryagin's principle is used to show that an optimal control strategy has only five driving modes: maximum power, cruising at constant speed, coasting, regenerative braking on a decline at constant speed, and braking. Albrecht et al. then used analysis of an adjoint variable to determine the optimal sequence of controls and precise switching points between control modes.

Figure 1 shows an optimal journey profile created with the Energymiser software for an example train journey between Belp and Bern in Switzerland.

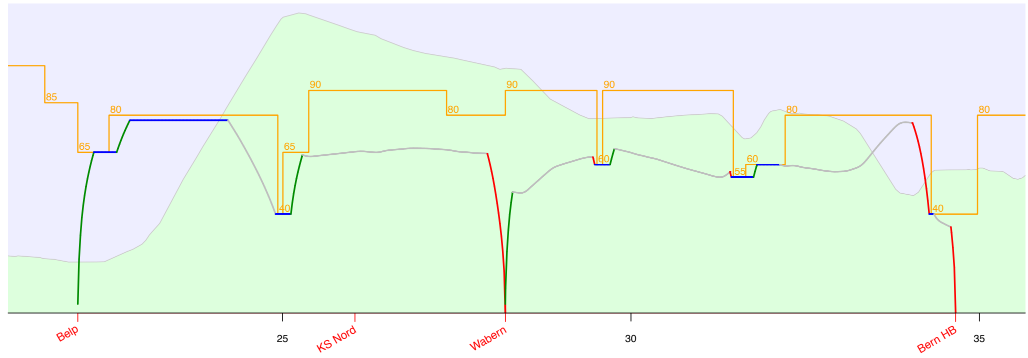

Figure 1: Example speed profile for an optimal journey. The horizontal axis is distance. The orange curve indicates the route speed limit in km/h. The bold curve shows the ideal speed profile and control modes: power (green), hold (blue), coast (grey) and brake (red). The green shading indicates the elevation profile.

Energymiser uses GPS to determine the current location and speed of the train, and calculates an optimal driving strategy to the next timing point in real time.

Rail operators run short trials to evaluate potential savings. They run trips without advice then trips with advice, and compare the energy use. There can be large variations in energy use for the same trip, but the mean energy use with advice is generally less than the mean energy use without advice, and there is less variation with advice.

Figure 2 shows energy use from a trial on an urban railway in Europe, with a trip length of 56 km and 16 stops.

How can we predict savings from a small number of trips, and estimate the savings over many trials? In Section 2 we describe a method for estimating energy savings for a single trial with a small number of trips. Section 3 describes a meta-analysis of results from 23 different trials.

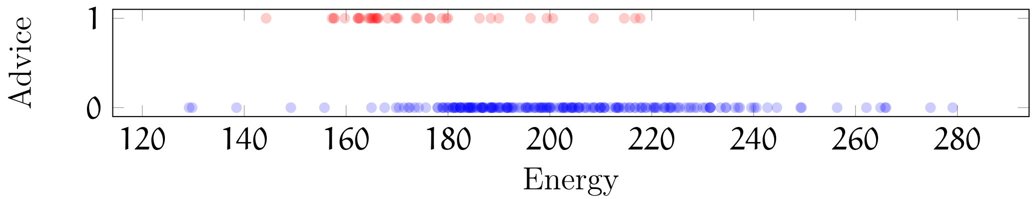

Figure 2: Energy use from from 244 trips over the same route. There were 39 trips with advice (red) and 205 trips without advice (blue). Each dot represents the energy use for a trip.

2 Estimating savings

Within each trial, we model energy use for a trip j as

$$E_j = \left(1 - \theta a_j + \sum_k \beta_k x_{kj}\right) E_0 + \epsilon_j$$

where $j = 1, \ldots, n$ is the trip number, and n is the number of trips in the trial; E_0 is the mean energy use for all trips without advice and with covariates set at zero; θ is the saving due to advice; a_j is an indicator that is 1 when advice is given, and 0 otherwise; x_{kj} are values taken by covariates; β_k are coefficients for the covariates; ϵ_j is random error, independently distributed with zero mean and a constant variance.

Each trial trip provides a_j and E_j. The primary aim is to estimate θ and its standard error, but this entails estimating E_0 and β_k as well. The covariates considered were x_1, coded as 0 or 1 to indicate direction along the route, and x_2 for the mass of the train on freight trains.

The estimation was performed using non-linear least squares.

Figure 3 shows the probability of exceeding a saving of θ, calculated using n total trips from the data shown in Figure 2. Half the trips had advice. The top graph is the result for $n = 20$ and the bottom graph is the result for $n = 60$. Each graph shows the range of possible savings indicated by the data. The pair of graphs shows how the certainty improves with the number of trips.

3 Meta-analysis

Different trials could be different operators, different routes or different timetables. Let θ_i be the underlying mean saving for trial i, and let $\hat{\theta}_i$ be the estimated saving for trial i. What is the overall mean saving Θ that Energymiser can provide, and how much will individual trials vary about this mean? A Bayesian hierarchical model has the form

$$\theta_i \sim \mathsf{N}\left(\Theta, \tau^2\right), \quad \hat{\theta}_i \sim \mathsf{N}\left(\theta_i, \sigma_i^2\right),$$

together with prior distributions for Θ and τ [e.g., 3]. The main objective is to estimate the mean saving Θ, and the standard deviation of individual trials about the mean τ. We know $\hat{\theta}_i$ and their associated standard errors σ_i. We ignore the uncertainty in the estimation of σ_i in order to facilitate implementation in the R package `bayesmeta` [4], but this has negligible effect on the results. The assumed priors for Θ and τ are typically normal and half-Cauchy, respectively,

$$\Theta \sim \mathsf{N}(\mu, \sigma_0^2) \quad \text{and} \quad \tau \sim \mathcal{HC}(x_0, s),$$

where $\mathcal{HC}$ is the half-Cauchy distribution, and μ, σ_0^2, x_0 and s are hyper-parameters.

The analysis follows from factorisation of the multivariate distribution of unknowns conditional on the data

$$f(\boldsymbol{\theta}, \Theta, \tau \mid \boldsymbol{y}) \propto f(\tau \mid \boldsymbol{y}) f(\Theta \mid \tau, \boldsymbol{y}) f(\boldsymbol{\theta} \mid \Theta, \tau, \boldsymbol{y}), \tag{1}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_I)$ and $\boldsymbol{y} = (y_1, \ldots, y_I)$ where I is the number of trials. The second and third terms on the right of equation (1) follow from inverse

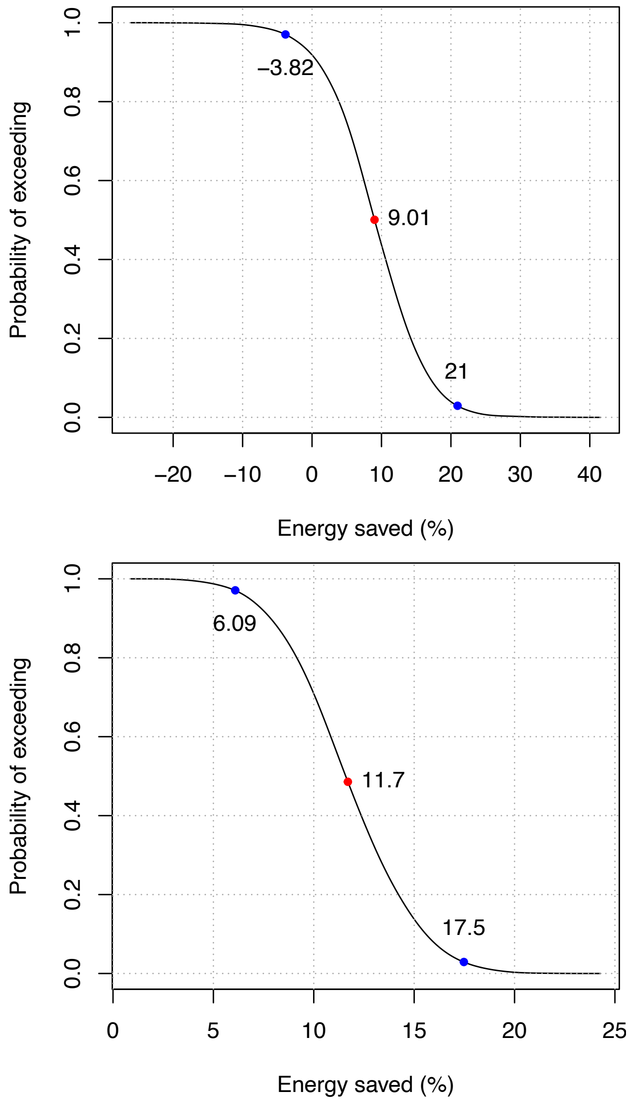

Figure 3: Predicted energy savings based on a comparison of n randomly chosen trips where half had advice, with $n = 20$ (top) and $n = 60$ (bottom). The red dots indicate the expected saving; the blue dots indicate the 95-percentile credible interval.

variance (precision) weighted averages. For the third term the independent components are

$$\theta_i \mid \Theta, \tau, y_i \sim N\left(\frac{y_i\phi_i + \Theta\phi_\tau}{\phi_i + \phi_\tau}, \frac{1}{\phi_i + \phi_\tau}\right),$$

where $\phi_i = 1/\sigma_i^2$ and $\phi_\tau = 1/\tau^2$. The second term is

$$\Theta \mid \tau, \mathbf{y} \sim N\left(\frac{\sum y_i\psi_i + \Theta_0\phi_0}{\sum \psi_i + \phi_0}, \frac{1}{\sum \psi_i + \phi_0}\right),$$

where $1/\psi_i = \sigma_i^2 + \tau^2$, and the prior distribution for Θ is $N(\Theta_0, 1/\phi_0)$. In the following it is convenient to introduce notation for the mean and precision of the distribution $\Theta \mid \tau, \mathbf{y} \sim N(\hat{\Theta}, \Psi^{-1})$ where $\hat{\Theta}$ is a vector of estimates of trial means and Ψ is a vector of reciprocals of variances of trial means.

The first term in equation (1) is obtained through the following nice argument. We start with

$$f(\tau \mid \mathbf{y}) = \frac{f(\tau, \mathbf{y})}{f(\mathbf{y})}.$$

Now the numerator and denominator are expressed as

$$f(\tau, \mathbf{y}) = \frac{f(\Theta, \tau, \mathbf{y})}{f(\Theta \mid \tau, \mathbf{y})}, \quad f(\mathbf{y}) = \frac{f(\Theta, \tau, \mathbf{y})}{f(\Theta, \tau \mid \mathbf{y})},$$

and so

$$f(\tau \mid \mathbf{y}) = \frac{f(\Theta, \tau \mid \mathbf{y})}{f(\Theta \mid \tau, \mathbf{y})} \propto \frac{f(\Theta, \tau) f(\mathbf{y} \mid \Theta, \tau)}{f(\Theta \mid \tau, \mathbf{y})}.$$

The left hand side of the above expression does not depend on Θ, so this relationship holds for any value for Θ. In particular we set Θ as $\hat{\Theta}$, and assuming that the distributions of Θ and τ are independent:

$$f(\tau \mid \mathbf{y}) \propto \frac{f(\tau) f(\mathbf{y} \mid \hat{\Theta}, \tau)}{f(\Theta \mid \tau, \mathbf{y})}.$$

That is

$$f(\tau \mid \mathbf{y}) \propto f(\tau) \frac{\prod_{i=1}^{I} \sqrt{\psi_i/(2\pi)} e^{-(y_i - \hat{\Theta})^2 \psi_i}}{\sqrt{\Psi} e^{(\Theta - \hat{\Theta})^2 \Psi}} .$$

Notice that when Θ is set to $\hat{\Theta}$ the exponential term $e^{(\Theta-\hat{\Theta})^2\Psi}$ in the denominator reduces to 1.

Implementing these calculations requires Monte–Carlo simulations. Each draw from the multivariate distribution begins with a draw from the scalar distribution $f(\tau \mid \mathbf{y})$ given in equation (1). The draw can be made by, for example, the Metropolis–Hastings algorithm or by a numerical inverse CDF procedure. Other draws are from normal distributions. The output is:

- a distribution for Θ, from which a mean and standard deviation is calculated;
- a distribution for τ, from which a mean and standard deviation is calculated;
- posterior distributions for θ_i, from which a mean and standard deviation is calculated as well as credibility intervals.

4 Results

We start by fitting the hierarchical distribution to all the trials. However, a substantial difference between freight and passenger services is that the latter typically make frequent stops according to a strict timetable, and this may impact energy saving. We therefore analyse these two categories independently.

Table 1 shows the estimates of expected energy savings from the 23 different trials, where $L_{0.95}$ and $H_{0.95}$ indicate the lower and upper bounds, respectively, of the 95% credible intervals. The credible intervals are set as the point estimate plus or minus two standard errors. These prior estimates of the θ_i and the corresponding standard errors are the inputs to the meta analysis.

Table 1: Percentage savings y_i and credible interval bounds $[L_{0.95}, H_{0.95}]$ from 23 trials.

Type	y_i	$L_{0.95}$	$H_{0.95}$
regional passenger	13.3	9.6	17.0
regional passenger	12.3	6.6	18.0
regional passenger	11.7	8.8	14.6
regional passenger	9.9	5.9	14.2
bulk freight	16.3	6.6	25.6
regional passenger	6.1	1.7	10.8
regional passenger	5.4	−3.2	14.2
regional passenger	5.0	1.4	8.9
regional passenger	6.7	2.7	10.6
regional passenger	2.8	0.5	5.1
regional passenger	5.9	2.9	9.1
regional passenger	5.3	2.0	8.6
regional passenger	7.9	3.2	12.5
regional passenger	3.5	0.5	6.6
regional passenger	7.9	3.7	12.0
freight	14.4	6.7	22.0
freight	7.4	2.4	12.0
intercity passenger	9.2	4.9	13.5
regional passenger	2.8	−10.6	15.4
freight	0.1	−7.5	7.3
freight	28.1	18.3	37.7
freight	7.7	−0.6	15.6
freight	9.2	−10.3	27.5

Table 2: Posterior distributions of percentage energy saving.

trial: number	parameter	median	mean	$L_{0.95}$	$H_{0.95}$
All: 23	Θ	8.15	8.18	6.3	10.2
	τ	3.65	3.76	2.0	5.7
Passenger: 16	Θ	7.41	7.42	5.6	9.2
	τ	2.83	2.92	1.5	4.5
Freight: 7	Θ	11.03	11.04	5.5	16.6
	τ	7.11	7.60	2.2	14.1

Tests of Energymiser during development suggested an overall mean saving of 10%, and the developers thought there was a 2/3 probability that it would be between 5% and 10%. So, the prior distribution for Θ is set as normal with a mean of 10 and a standard deviation of 5. The developers also estimated that there was a 1/2 probability that the standard deviation of trials about the mean would be less than 5%. So, the median parameter of the half-Cauchy prior for τ is set at 5. The results are summarised in Table 2, including the shortest 95% credible intervals. A forest plot showing the posterior estimates of θ_i is shown in Figure 4.

For passenger routes, the estimate of Θ is taken as the median of its posterior distribution which is 7.4%. The posterior distribution is near symmetric and the mean and median are close. The estimate of τ is taken as the median of its posterior distribution and is 2.8%. For freight routes the estimates of Θ and τ are 11.0% and 7.1%, respectively.

5 Conclusion

Based on 16 passenger trials and seven freight trials:

- passenger trains have an average saving of 7.4%, and it is estimated that around two thirds of trials have a saving between 4.6% and 10.2%;
- freight trains have a slightly higher average saving of 11.0%, and it

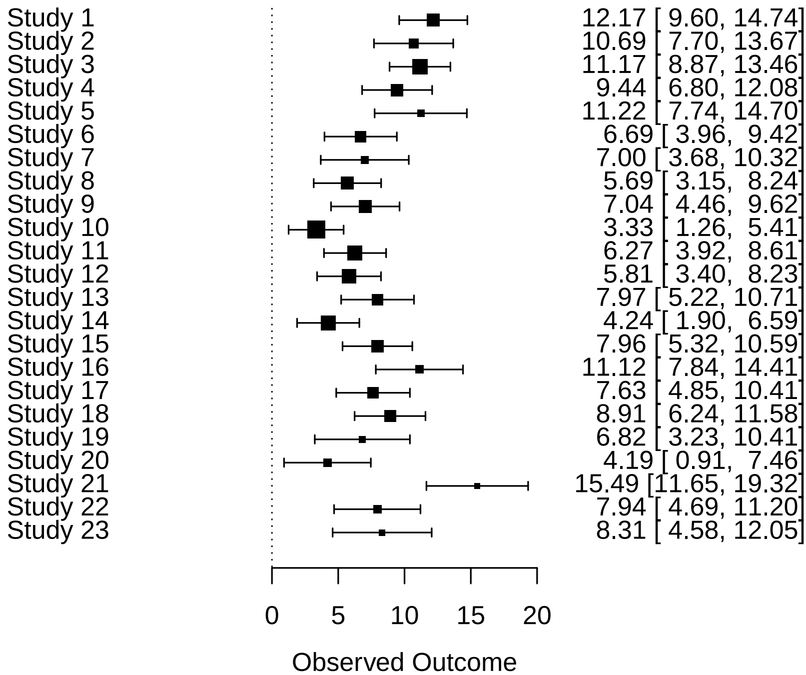

Figure 4: Posterior trial (study) means.

is estimated that around two thirds of trials have a saving between 3.9% and 18.1%.

So, train operators can expect to save around 8% on fuel costs if they adopt Energymiser. Since energy costs are a substantial proportion of a rail operator's outgoing payments, and the cost of Energymiser is relatively low, this would be a substantial saving. Given sufficient trials, it would be feasible to explain, in part at least, the deviations of θ_i about Θ in terms of variables such as train operator, train driving crew, and energy source using multiple regression. However, this was not an aim of this research.

Acknowledgements

This research was partly supported by funding from the Australian Government through the Australian Research Council's Linkage Project funding scheme (project LP110100136), and by TTG Transportation Technology.

References

[1] A. Albrecht, P. Howlett, P. Pudney, X. Vu, and P. Zhou. "The key principles of optimal train control—Part 2: Existence of an optimal strategy, the local energy minimization principle, uniqueness, computational techniques". In: *Transport. Res. B: Method.* 94 (2016), pp. 509–538. DOI: 10.1016/j.trb.2015.07.024. (Cit. on p. C111).

[2] A. Albrecht, P. Howlett, P. Pudney, X. Vu, and P. Zhou. "The key principles of optimal train control—Part 1: Formulation of the model, strategies of optimal type, evolutionary lines, location of optimal switching points". In: *Transport. Res. B: Method.* 94 (2016), pp. 482–508. DOI: 10.1016/j.trb.2015.07.023. (Cit. on p. C111).

[3] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC Press, 2012. DOI: 10.1201/b16018 (cit. on p. C114).

[4] C. Röver. "Bayesian random-effects meta-analysis using the bayesmeta R package." In: *J. Stat. Software* 93.6 (2020), pp. 1–51. DOI: 10.18637/jss.v093.i06 (cit. on p. C114).

Author addresses

1. **A. J. Fisher**, School of Mathematical Sciences, University of Adelaide, Australia.
2. **A. V. Metcalfe**, School of Mathematical Sciences, University of Adelaide, Australia.

orcid:0000-0002-7680-3577

3. **P. J. Pudney**, CIAM Industrial and Applied Mathematics, University of South Australia, Australia.
orcid:0000-0001-5596-5473