Bayesian inference for high-speed train dynamics and speed optimization under uncertainty for energy saving

J. Nespoulous [1,2], **C. Soize** [1], **C. Funfschilling** [2], **G. Perrin** [3]
[1] Université Gustave Eiffel, MSME UMR 8208,
5 Bd Descartes, 77454 Marne-La-Vallée, France

[2] SNCF, DTIPG,
1-3 Av François Mitterrand, 93574 Saint-Denis, France

[3] Université Gustave Eiffel, COSYS,
5 Bd Descartes, 77454 Marne-La-Vallée, France
e-mail: **julien.nespoulous@free.fr**

Abstract
The train is a complex nonlinear system, whose dynamic behavior is difficult to predict accurately because of its environmental sensitivity. Indeed, in spite of a relative fine modeling of the vehicle and its rolling environment (track and wind), the slightest uncontrolled disturbance can modify the dynamic comportment of the train. For this reason, uncertainty must be considered in the physical models. The industrial objective of this work is twofold. Firstly, the construction of a longitudinal dynamic model for high-speed trains able to take into account the fluctuations inherent to the system. Secondly, the optimization under uncertainty of the driver's command with the objective of reducing the energy consumed by the train, under a set of punctuality and physical nonlinear constraints (speed limitation, final speed, and final position constraints).

1 Introduction

Controlling energy consumption has turned to be an important challenge of the 21^{st} century and particularly in the railway world, since the transport sector constitutes one of the largest consumers. For this reason, the railway companies pay close attention to their energy consumption and seek to reduce it. Recently, this objective has become even more crucial because of the growing demand, stemming from the increase of the trains' frequency, as well as their speed. To achieve this reduction, three levers can be activated: modify the rolling environment, the vehicle characteristics, or the speed profile. The present work focuses on the optimization of the latter. As a result, the optimal speed profile can be used as a guide to help drivers or it could be implemented in autonomous trains.

It is difficult to perfectly predict trains' behavior because of its environmental sensitivity. Thus, particular attention must be paid to define the system entries (track and wind characteristics), in order to construct the train longitudinal dynamic model [1] from a Lagrangian multi-body approach and the energy consumption model. These models should consider the uncertainty [2] as all trains do not have the same dynamic response to a specific entry.

When the models are satisfactory, that is to say, when they are sufficiently precise, we can try to minimize the energy consumed under a set of constraints. These constraints aim to assure the security, the punctuality of the journey, and the comfort of the passengers. Several methods can be used to solve this optimization problem as dynamic programming [3], evolutionary algorithms [4], or pseudo-spectral method [5].

The objective of this paper is to present a method, which includes uncertainty in high-speed train models. These models were then used to optimize the driver's command, to minimize the energy consumed by the train under uncertainty. Section 2 presents the construction of the dynamic and energy consumption models. Afterwards, the optimization problem is defined, simplified, and solved in Section 3.

2 Bayesian inference for high-speed train modeling

2.1 High-speed train dynamic and energy consumption models

The first step is to find the impact that the driver has on the dynamic system and its energy consumption.

The train dynamics highly depends on the environment. For example, a strong headwind will impact the resistant force and therefore the energy consumed during the journey. Consequently, the dynamic and energy consumption models have to consider the effect of the rolling environment $\mathcal{E}$. It is composed of both the track characteristics (declivity and curvature) and the wind description (amplitude and direction).

The action of the driver on the system is called the driver's command u. It is an abstract representation of the traction/braking manipulator. In other words, it corresponds to the choice of a normalized traction or braking longitudinal force that the driver decides to inject in the system. It is a deterministic time-dependent variable that takes any values inside the interval $[-1, 1]$, 1 (respectively -1) representing the maximum traction torque (respectively braking torque) available for a given speed.

After a sensitivity analysis on the model parameters, we have identified 9 of them, which impact greatly the system and can be considered uncertain. Effectively, they may vary depending on the weather or the train under consideration. The uncertain parameters are summarized below:

- Total mass of the train M_T.
- Davis parameters A, B, and C.
- Auxiliary power P_a.
- Efficiency parameters A_η, B_η, C_η, and D_η.

We write $\boldsymbol{X}$ the random vector containing the random variables associated with the uncertain parameters.

For a given rolling environment $\mathcal{E}$, a known deterministic driver's command u and a specific uncertain vector $\boldsymbol{X}$, we can solve the whole dynamic problem. The train dynamics can be modeled by a rigid body approach, from which we derive the Lagrangian equations. These equations are written for the whole train and are projected on the longitudinal to the track axis:

$$\boxed{M_T k^r \ddot{Y}(t) = \sum_\alpha F^\alpha\left(u(t), \boldsymbol{X}, \mathcal{E}\right)} \tag{1}$$

with k^r a factor including the wheels' rotation, $\ddot{Y}$ the train longitudinal acceleration. F^α regroups traction and braking forces (adapted to the TGV Dasye high-speed train), the *Davis* or resistant force, a corrective force applied in curve, and the weight.

The energy consumed by a train varies for different reasons. Firstly, the traction chain converts electric energy to mechanical energy with an unknown efficiency. Secondly, a part of the electric power collected by the pantograph is not transmitted to the traction chain but is used to assure the passenger comfort or the security of the journey. It is called the auxiliary power P_a. The energy consumed is written:

$$\boxed{F^E(u, \boldsymbol{X}, \mathcal{E}) = \int_{t_s}^{t_f} P^E\left(u(t), \boldsymbol{X}, \mathcal{E}\right)\, dt\,,} \tag{2}$$

where P^E is the electric power, t_s, and t_f are the initial and arrival times.

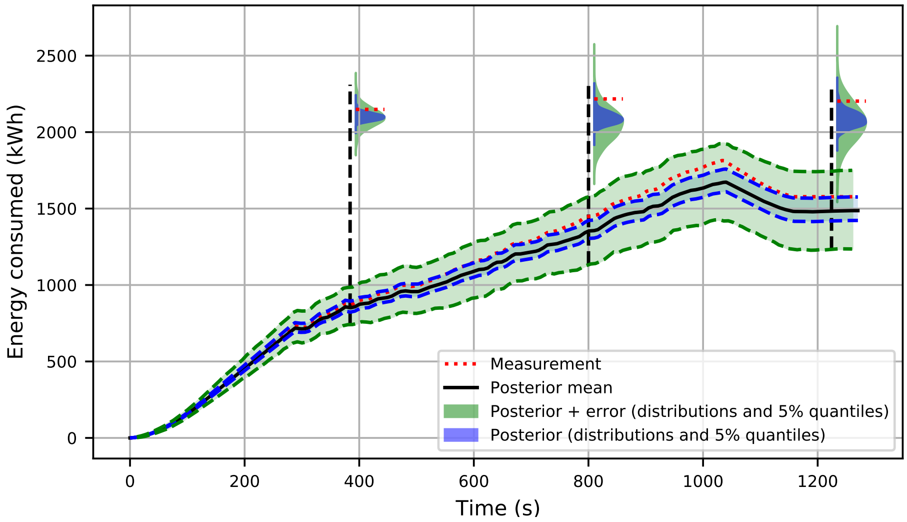

Figure 1: Simulated and measured energy consumed by the train in function of the position calculated from realizations of the uncertain parameters drawn in the posterior distributions.

2.2 Bayesian inference method to quantify the parameter uncertainty

All trains do not behave the same way and the consideration of the system fluctuations appears to be necessary. The uncertainty is proposed to be comprised in the model parameters, defined probabilistic. These models are calibrated with a Bayesian inference from energy measurements carried out on commercial trains.

The first step of the Bayesian calibration is to construct the prior distributions for each parameter. To do so, we need to exploit as much information as we have. Thus, we regroup all the available information on each uncertain parameter. Based on this information, the maximum entropy principle [6] yields specific prior laws, whose hyperparameters are identified analytically.

The next step consists in the definition of a likelihood function. It aims to compare the energy measured and simulated at specific instants. In practice, we consider the measurements to be deterministic, and the error comes from the model. We decompose two kinds of error. ϵ^F is applied to the longitudinal force, and ϵ^P is applied to the electric power. These two random variables are supposed to be normal, centered, with a variance estimated from the maximization of the likelihood function. As the likelihood function may take very small values, we prefer to use its logarithmic form.

Finally, we apply the Metropolis-within-Gibbs algorithm to the TGV Dasye train on the LGV Rhin-Rhone track. This algorithm observes the uncertain parameters one-by-one, and it is useful when they do not have the same influence on the system. By using a classic Metropolis-Hastings algorithm, we may slow down the convergence of the algorithm for some parameters.

Once the algorithm has converged, the posterior distributions would give an accurate description of the system. Drawing realizations of these random variables is equivalent for choosing specific parameter values. Figure 1 shows the energy consumed by the train depending on the position. The blue envelope is the 95% quantile interval from realizations of $\boldsymbol{X}$ and the green envelope is the quantile interval for realizations of both $\boldsymbol{X}$, ϵ^F, and ϵ^P. The mean value is the solid line, and the energy measurements are plotted in red. The distributions of the green and the blue envelopes are plotted at three specific instants.

We can observe that the confidence intervals, proposed for quantifying the model uncertainty, are appropriate as they include all the measurements. The variance of the confidence interval increases with time as the error accumulates. Moreover, the measurements are better described by the posterior distributions in comparison with the priors. It signifies that posterior distributions have a great capacity to represent the case study. The energy consumed decreases at the end of the journey as the train is braking and recovers a part of the energy. The distributions plotted at three instants show how the simulated energy is dispersed around the mean value and where the measurements lie.

3 Robust optimization under uncertainty

Once the models are sufficiently accurate, we use them to find the optimal driver's command that minimizes the energy consumed. The optimal solution has to respect a set of constraints $\mathcal{C}$ assuring the security, the punctuality, and the comfort of the passengers. As we have decided to define the parameters with random variables, the optimization method has to handle the uncertainty.

3.1 Deterministic optimization of specific configurations

The first idea is to find a family $(u_k^*)_k$ that minimizes the energy consumed in specific configurations $\boldsymbol{x_k}$ randomly drawn in $\boldsymbol{X}$. The optimization problem can be written:

$$\boxed{u_k^*(\mathcal{E}, \boldsymbol{x_k}) = \arg\min_{u \in \mathcal{C}} F^E(u, \boldsymbol{x_k}, \mathcal{E})} \tag{3}$$

With this method, all the quantities are subject to the initial choice of $\boldsymbol{x_k}$, and the optimization problem remains deterministic. To consider the constraints, we choose to penalize the cost function with an Augmented Lagrangian. Another issue is the high dimension of the problem. Indeed, optimizing a continuous function is very costly. Consequently, we propose to reduce the dimension of the problem by approaching the driver's command to a constant piecewise function.

To solve this problem, we use an iterative algorithm called CMA-ES [7], which stands for Covariance Matrix Adaptation - Evolution Strategy. At each iteration, the algorithm draws a population of few points, evaluates the cost function and its penalization (from Augmented Lagrangian). From this knowledge, it calculates the covariance matrix to estimate the best directions to explore in order to orientate the draws of the next iteration. After around two thousand iterations, the calculation converges to an optimal solution that verifies the constraints and reduces the energy consumed of around 28%.

The family of solutions obtained with this method are optimal trajectories, that should be used in specific configurations $\boldsymbol{x_k}$. But the driver cannot know which solution is the best for his journey. And a specific trajectory can be optimal in a configuration $\boldsymbol{x_k}$ but it can be rather bad in another situation $\boldsymbol{x_j}$. For this reason, we propose a second strategy.

3.2 Optimization of the driver's command under uncertainty

To help drivers reducing the energy consumption, we should give them a nominal deterministic command to follow. But each command is associated with a particular configuration. Therefore, the second strategy looks for an optimal trajectory that is easy to transform in order to fit each configuration. As a result, the optimization problem changes slightly. Effectively, the optimal command u^* is defined deterministic in an admissible domain $\mathcal{D}$ and should be easily adaptable to a given configuration $\boldsymbol{x_k}$ with a new transformed command $T(u, \boldsymbol{x_k})$. This transformation corresponds to a slight increase (or decrease) of the command to fit the simulated final position with the real position of the train station. In other words, the initial command is modified, so that the new solution directly verifies the security, punctuality, and comfort constraints.

Finally, we minimize the expected value of the energy consumed over $\boldsymbol{X}$ by approaching it numerically. By removing the difficulty of respecting the constraints, which are highly dependent on $\boldsymbol{X}$, and by approaching the mean numerically, we deflect the issue of optimizing under uncertainty. Finally, the optimization problem can be written:

$$\boxed{u^* = \arg\min_{u \in \mathcal{D}} \frac{1}{K} \sum_{k=1}^{K} F^E\left(T(u, \boldsymbol{x_k}), \boldsymbol{x_k}, \mathcal{E}\right).} \tag{4}$$

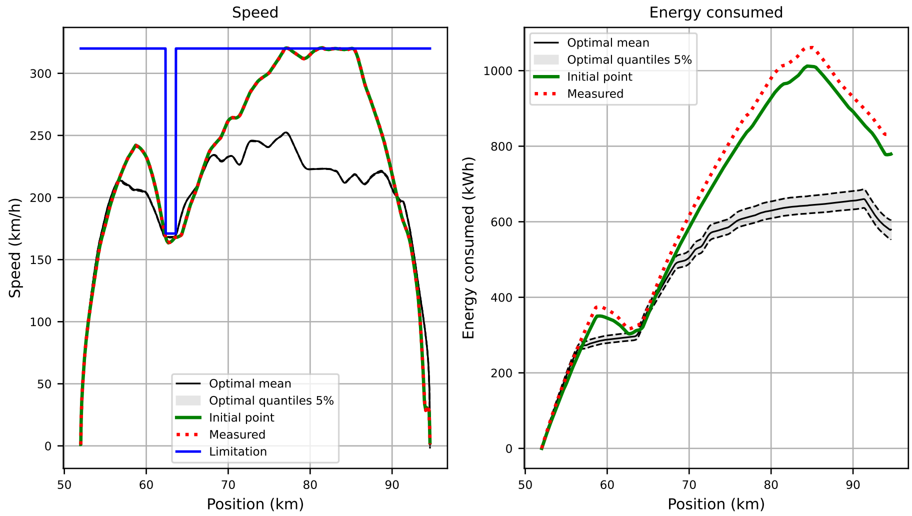

Figure 2: Measured and simulated speed trajectory (left figure) and energy consumed (right figure) calculated from the optimal driver's command and its transformations in function of the position.

To solve this problem, we use the evaluations of the cost function already calculated while achieving the previous method. In this way, we explore the admissible space highlighted before. A Principal Component Analysis is realized on these evaluations, weighted with their associated energy consumed, in order to extract the principal axes of research. These axes have the most impact on the solution regarding energy saving. The dimension of the optimization problem is reduced from the size of the command vector (200) to the number of principal axes selected (60 with 99% of the information conserved).

The optimization method applies the CMA-ES algorithm on the axes conserved after reduction. Consequently, keeping few directions limits the exploration but accelerates the convergence. The candidate vectors are calculated from the linear combination of the principal axes. They are slightly transformed to respect the constraints. The CMA-ES algorithm is applied to the coefficients of the linear combination and not the driver's command as before. The optimal trajectory and its transformations are shown on Figure 2.

The optimal speed profile respects the constraints. The variance of the speed is around $1km/h$ all along the journey. The transformations applied to u^* have an impact on the train dynamics. Nevertheless, the same speed trajectory is optimal for each realization of $\boldsymbol{X}$ and is obtained by adapting u^*. Even if the speed profiles are similar for different realizations of $\boldsymbol{X}$, the associated energy consumed varies a lot.

The energy is reduced by 26% compared with the initial trajectory. The optimal solution has approximately the same variability as the one calculated with the first optimization method (Section 3.1) and consumes barely the same amount of energy. However, the second method has a different objective and aims to find a trajectory easily transformed that covers all the configurations contained in $\boldsymbol{X}$.

4 Conclusion

In this paper, we have proposed a method, which aims to include the uncertainty in the dynamics and energy consumption models. This improves the quality of the models. However, it appears that the constraints of the optimization problem are very sensitive to the uncertainty. Consequently, we have proposed a new strategy, which aims to find an initial trajectory that can be easily transformed to fit with other possible configurations, still respecting the constraints.

In the near future, this algorithm may be used on different tracks and for various vehicles to help drivers minimizing the energy consumption. The solution can be applied easily because the algorithm directly outputs the command. Later, we can imagine using this method to find nominal trajectories for autonomous trains, in which the driver's feeling will disappear, and is going to be replaced by robust algorithms.

References

[1] Q. Wu, M. Spiryagin, and C. Cole, "Longitudinal train dynamics: an overview," *Vehicle System Dynamics*, vol. 54, no. 12, pp. 1688–1714, 2016.

[2] D. Lebel, C. Soize, C. Funfschilling, and G. Perrin, "Statistical inverse identification for nonlinear train dynamics using a surrogate model in a bayesian framework," *Journal of Sound and Vibration*, vol. 458, pp. 158–176, 2019.

[3] V. Calderaro, V. Galdi, G. Graber, A. Piccolo, and D. Cogliano, "An algorithm to optimize speed profiles of the metro vehicles for minimizing energy consumption," in *2014 International Symposium on Power Electronics, Electrical Drives, Automation and Motion*. IEEE, 2014, pp. 813–819.

[4] J. Nespoulous, C. Soize, C. Funfschilling, and G. Perrin, "Optimisation of train speed to limit energy consumption," *Vehicle System Dynamics*, pp. 1–18, 2021.

[5] P. Wang and R. M. Goverde, "Multi-train trajectory optimization for energy efficiency and delay recovery on single-track railway lines," *Transportation Research Part B: Methodological*, vol. 105, pp. 340–361, 2017.

[6] C. Soize, *Uncertainty quantification*. Springer, 2017.

[7] N. Hansen, "The cma evolution strategy: A tutorial," *arXiv preprint arXiv:1604.00772*, 2016.