Service Regularity Loss in High-Frequency Feeder Bus Lines: Causes and Self-Driven Remedies

Jorge González ‡, René Doursat † and Arnaud Banos †

Abstract. An agent-based model of a single bus line is designed to assess quality of service in terms of regularity and occupancy. Agents represent buses moving on a linear network of stops, where passengers (summarized by counts) get on and off. Three performance-loss factors are considered: distances between stops, drivers’ behavior, and influx of users along the line. We show that service regularity is less impacted by section lengths and bus speeds than by demand. Two solutions are tried: increasing the number of buses through a higher departure frequency, and improving the exchange of passengers with all-door boarding instead of the usual one-door scheme. These solutions are self-driven since no central supervision is required. We defined a measure of regularity based on the distribution of “headways” (time intervals) between buses. According to our model, higher frequencies do not improve regularity, whereas all-door boarding does for large numbers of incoming passengers. We describe a case study based on real data collected from a “feeder” service of the Parisian regional train network. It is highly unbalanced, as it serves opposite commuting directions during the morning and evening peaks, making it an especially interesting example.

Keywords. Public transportation network, bus, reliability, delay, agent-based modeling, boarding.

1 Introduction

Every transportation system in the world relies on buses due to their numerous advantages. A bus network does not need a huge infrastructure investment because it can use existing streets and roads [4, 7], making it more economical than other public transportation systems. It is resilient, in the sense that if a bus encounters a problem, other buses can overtake it or be rerouted as needed [1, 6]. At relatively low operating costs, it is also a profitable enterprise [4, 8]. On the other hand, buses are often perceived by the public as uncomfortable and unreliable because of highly variable travel and waiting times, which depend on the traffic conditions over the shared road network [1, 2].

Nowadays, priority at crossroads and exclusive lanes are recognized to greatly improve the efficiency of bus transportation systems, in particular commercial speed [6, 5]. Yet, even high-standard systems suffer a loss of regularity as consecutive buses inevitably start “clumping” together after a while, something that can be observed in every city around the world. There can be long empty periods followed by the arrival of two or more buses at the same time. Not many explanatory models of this phenomenon have been proposed, most of the research work focusing on the supervision of bus operation, the fast detection of perturbations, and optimal strategies to restore service after a disturbance [3].

We explore here two self-regulatory approaches toward improving the quality of service of a bus line. The first approach increases supply through the number of buses and frequency of departure. The second approach facilitates demand by implementing all-door boarding instead of one-door boarding. It allows passengers to board the bus through any door by installing fare collectors at every entrance or at every bus stop. Our study is based on the agent-based modeling and simulation (ABMS) of a single bus line, applied to real data collected from Nr. 91.06C (Massy-Palaiseau; Gare RER – Saclay; Christ) in the greater Paris area. This line is a “feeder” service providing the main connection between the Plateau de Saclay, an academic and industrial hub, and the regional suburban train network (RER). During the morning and evening rush hours, it has a high frequency of one bus every five minutes, as the great majority of commuters work at the facilities located in Saclay. This particularity is also the reason for a highly asymmetrical schedule, with a morning peak in the outbound direction, and an evening peak heading inbound. Moreover, as the region is also experiencing a recent development boom, customer demand has increased rapidly in the last few years and this growth is expected to continue. For all these reasons, this case study is an interesting object of application, both from a scientific and a practical perspective.

2 Methodology

We propose an agent-based model whose main agents are the buses running on one line. Individual passengers are
not differentiated, but grouped by destination. To build
the network, bus stops are connected by edges, called
“sections”, forming a 1D pathway from one terminal to
the other. Bus stops and sections are also considered
agents with attributes and indexed accordingly. On each
section, between two consecutive stops, a bus moves at its
own constant speed. We assume that a bus driver has a
preferred speed and each section has both a recommended
and a maximum speed. The preferred speed of a bus is the
speed that its driver tends to. The recommended speed
of a section is the speed that buses tend to over that
section, and may reflect geographical constraints, traffic
flow, or some degree of central regulation. These origins
are not taken into account, however, only the effect of
this quantity is factored in by assigning it equally to all
buses. In sum, the actual speed of a bus depends on the
recommended and maximum speeds of a section and the
preferred speed of the driver. When a bus arrives at a
station, it stops if passengers want to board or exit the
bus, and stays there as long as needed. Then, it moves
to over the next section and so on, until the last stop.

2.1 Model

The parameters of the model are the following: there are
M sections and M + 1 bus stops. The path-like graph
representing the bus line is denoted by G = (V, A), where
V is the set of bus stops and A the set of sections. The
elements of V are indexed by i = 0, 1, . . . , M, corresponding
to their position along the line, with the first stop at
i = 0 and the last stop at i = M. Let da be the length of
section a ∈ A, and vcom a and vrec a the recommended and
maximum speeds on this section. The first two values are
uniformly drawn in intervals [da, da+1] and [vcom a, vrec a],
respectively. We denote by p i,j the rate of passenger influx
going from origin i to destination j, with (i, j) ∈ V 2, by
P i = ∑j∈V p i,j the total influx rate at stop i, and by
P = ∑i∈V P i the global influx rate. The p’s are drawn
from a power-law distribution in [P−, P+], with exponent α.
Let D i,j the fraction of influx at stop i heading for
j, so that 0 ≤ D i,j ≤ 1 and ∑j D i,j = 1. Another index,
b ∈ B = {1, 2, . . . , N} denotes the b-th bus to leave from
stop 0: its departure time is T b, and its capacity C b. The
scheduled headway between consecutive buses is H b, and
its standard deviation σH. The preferred speed of bus
driver is denoted by v b pref , and is drawn from a normal
distribution of mean µ pref and variance σ pref 2. Finally, ∆t
represents the time-step of our discrete model.

Next, we introduce the variables of the dynamics (omitting
the time-dependency notation “... (t)” for greater
readability). Let r i,j be the actual number of passengers
waiting at stop i who want to go to j, o b j the number
of passengers on bus b whose destination is also j, and
a b j = ∑i′∈V, i′<i o b j the current occupancy of the bus.
Additionally, vcom b denotes the commercial speed, and vb the
actual speed of bus b. Finally, t′(b) represents the arrival
time of bus b at stop i.

2.2 Initialization

Given M, we build the graph G = (V, A) of our model as
follows. Each bus stop i ∈ V is assigned a total passenger
influx rate P i ∼ ΠL(α, P−, P+), which is a power-law
distribution ΠL(p) ∼ p −α restricted to [P−, P+], and per-
destination influx rates p i,j = P i D i,j for all j > i. When
using real data, we may set one or more of these values
to actual measurements. Each bus b ∈ B leaves stop
0 according to a randomly perturbed timetable with a
departure at T b ∼ N(bH0, σH 2 ), where N(µ, σ 2 ) is the
normal distribution of mean µ and variance σ 2, and is
also given a preferred speed v b pref and capacity C b . There
are no users at the beginning of the simulation: r i,j = 0
and o b j = 0 for all i, j ∈ V, b ∈ B.

2.3 Microscopic rules

At each time step, variables are updated according to the
following rules (Fig. 1):

1. “Waiting” rule: passengers come to bus stops at con-
   stant rates that depend on their itinerary, thus we
   write: r i,j ← π p i,j ∆t + r i,j for all j > i, where
   π(p) is the Poisson law of mean λ.

2. “Exiting” rule: when bus b arrives at stop i, the o b j
   passengers inside who wanted this destination exit
   the bus, therefore: o b j ← 0.

3. “Boarding” rule: at this point, we use the following
   temporal variables: f = ∑j∈V r i,j for the number
   of passengers waiting at stop i, g = o b j for the
   passengers who just left the bus, c = C b − o b j + g for
   the
remaining capacity on the bus, and \( n \) for the passengers who are boarding the bus. Two cases arise:

- “Undercapacity boarding” rule: if \( f \leq c \), then all \( n = f \) passengers are able to take the bus, hence: \( s_i' = s_i + r^{i,j} \), and \( r^{i,j} = 0 \) for all \( j > i \) (Fig. 1b).
- “Overcapacity boarding” rule: if \( f > c \), only \( n = c \) passengers can take the bus and \( n' = f - c \) must stay out. In this case, we draw an \( n \)-combination of destinations from the set of \( f \) passengers: let \( s^{i,j} \), with \( 0 \leq s^{i,j} \leq r^{i,j} \) and \( \sum_{j \in V} s^{i,j} = n \), denote the number of passengers with destination \( j \) who boarded the bus. Then we apply: \( s_{i}' = s_{i} + s^{i,j} \) and \( r^{i,j} = r^{i,j} - s^{j,i} \) for all \( j > i \). Finally: \( a_b \leftarrow C_b \).

4. “Stopping” rule: bus \( b \) stays at stop \( i \) for a duration \( \tau(t_{in}(f), t_{out}(g)) \), where \( \tau(t, t') \) can represent either a maximum, if the entrance and exit doors are separated: \( \tau = t_o + \max\{t_{in}(f), t_{out}(g)\} \), or a sum, if inward and outward passenger flows are mixed: \( \tau = t_0 + t_{in}(f) + t_{out}(g) \). In both cases there is a fixed stopping overhead \( t_0 \), including the time to open the doors. The arguments, \( t_{in} \), and \( t_{out} \) represent the time taken by passengers who are boarding and exiting the bus, respectively. Generally, they are nonnegative logistic functions but we adopt here a simple linear scheme: \( t_{in}(f) = \rho_{in} f \) and \( t_{out}(g) = \rho_{out} g \), where the \( \rho \)'s are time-per-person rates.

5. “Lingerling” rule: when bus \( b \) is about to leave stop \( i \) after a time \( t \), and if it is still under capacity, the simulation verifies if other passengers have arrived in the meantime: if \( f = \sum_{j \in V} s^{i,j} > 0 \) at this instant, then the bus stays longer to allow the additional passengers to board, otherwise it leaves. This process may occur several times as long as new passengers are showing up. However, if the boarding rate is less than the average time between two incoming passengers, \( \rho_{in} < 1/P_i^b \), the likelihood of repeating this process decreases with successive iterations.

6. “Moving” rule: when bus \( b \) leaves a stop, its speed is set to \( v_b = (v_{rec} + v_{pref})/2 \) on the whole section \( a \in A \) that lies ahead. From there, at each iteration, if \( \Delta x = v_b \Delta t \) is greater than the remaining distance to the next stop, the bus arrives at that stop; otherwise, its location is simply increased by \( \Delta x \) (Fig. 1a).

### 2.4 Output values

Let \( B^i \) denote the set of buses \( l = 1, \ldots, |B| \) sorted by increasing arrival time \( t'(b) \) at stop \( i \). For all \( b \in B^i \) with \( l < |B| \), we have by construction: \( t'(b) \leq t'(b_{l+1}) \), therefore we can calculate the headways between buses observed at stop \( i \) as follows: \( h'(b_i') = t'(b_{i+1}) - t'(b_i) \). This allows us to define a bus group \( K = \{k, k+1, \ldots, k'\} \) for a given time threshold \( \theta > 0 \) such that, for all \( l \in K \setminus \{k'\} \), we get \( h'(b_i) \leq \theta \), but \( h'(b_{k+1}) > \theta \) and \( h'(b_{k'}) > \theta \).

At every time step, the following quantities are measured: occupancies \( o_{b_i} \), headways \( h'(b) \), bus group sizes \( |K| \) and commercial speeds \( v_{com}^0 \). The latter are the ratio of the total length of the line to the total time spent by each bus on this line. We also define the average headway at stop \( i: \mu_b^i = (h'(b))_{b \in B} \), its standard deviation \( \sigma_b^i = \sqrt{(h'(b)^2)_{b \in B} - (\mu_b^i)^2} \), and its maximum \( h'_+ = \max_{b \in B}(h'(b)) \).

### 3 Main Results

Several parameters have been introduced in Section 2.1. In order to analyze their impact on the dynamics and the outcome of the model, we conduct different tests in which only one of them is modified at a time. With the goal of applying this methodology to practical situations, we use real data from the bus line 91.06C described above. Most of the parameters can be set or estimated on the basis of the available data. Since this line is mainly used by commuters coming from the RER train system in the morning and returning there in the evening, buses generally start full and empty themselves in the morning, whereas they start empty and fill up in the evening. This important asymmetry, typical of feeder lines, is taken into account in our work. We choose two datasets to test our hypotheses: one covering the morning rush hour in the direction RER → Saclay, the other covering the evening rush hour in the opposite direction, Saclay → RER. These datasets provide the values used in the tests of Section 3.1, excluding a different parameter every time, which is arbitrarily varied to assess its effect on the system.

The physical properties of the line, its size \( M \) and section lengths \( \{d_a\}_{a \in A} \), are known. The normal distribution of preferred speeds is estimated at \( \mu^b_{pref} = 50km/h \) and \( \sigma^b_{pref} = 2km/h \), which is a reasonable assumption since the line in part uses bus-only lanes and has priority at crossroads. User influx profiles originate from the data and, as expected, are widely different in the morning and evening. Buses are articulated three-door with a capacity of \( C = 120 \). The planned headway is \( H_0 = 5mn \), and departure times are not randomly perturbed in these simulations, i.e. \( \sigma_H = 0 \). The uniform distribution interval of recommended speeds is given by \( v_{rec}^0 = 40km/h \) and \( v_{com}^0 = 60km/h \). The boarding and exiting rates are \( \rho_{in} = 3s/passenger \) and \( \rho_{out} = 1s/passenger \), and we use the maximum-function variant of \( \tau \) with a fixed stopping overhead of \( t_0 = 5s \). This reflects the fact that boarding takes on average three times longer than exiting, as it is allowed only through one door, where passengers need to pay, whereas there are two other doors for immediate exit. The time step \( \Delta t \) is set to 1s and every run lasts 14,400 iterations, i.e. 4 hours, which is enough to cover each of the two rush periods between 7-10am and 4-7pm. Each test consists of 100 runs for statistical purposes.
3.1 Assessing disturbance factors

In a first step, we want to explore parameter space to find out how different factors may either encourage or control disturbances in the planned schedule. In a next part, we will propose preventive measures to alleviate these fluctuations. The quantities of interest are:

- the uniform distribution of section lengths $d$, parameterized by $d_-$ and $d_+$,
- the normal distribution of preferred speeds $v_{\text{pref}}$, parameterized by $\mu_{\text{pref}}$ and $\sigma_{\text{pref}}$,
- the power-law distribution of incoming users $P_i$, parameterized by $P_-, P_+$ and $\alpha$, and normalized by $P$.

Six tests were performed: three to study amplitude effects and three for heterogeneity effects. Each test consists in varying only one of the parameters and running the model 100 times for each value to obtain a more reliable average for the output and reduce stochastic bias. In 3.1.1, we focus on the effect of amplitude, thus we set lengths, speeds, and influx rates to the same values for all agents. In 3.1.2, we focus on the effect of heterogeneity, thus we set the mean of each distribution and try out different standard deviations. In both cases, the mean and the standard deviation of all outputs are calculated.

3.1.1 Assessing amplitude effects

In the first batch of three tests, the effect of input amplitude is assessed by giving homogeneous values to all the parameters:

- Test 1: section lengths are all equal to $d = d_- = d_+$, which can be 500, 1000, 1500, or 2000m
- Test 2: preferred speeds are all equal to $\mu_{v_{\text{pref}}}$, which can be 40, 45, 50, 55, or 60km/h, while $\sigma_{v_{\text{pref}}} = 0$
- Test 3: passenger influx profiles are given by the data (users were polled about their planned itineraries, recorded in $D^{ij}$), then renormalized by dividing by their sum and multiplying by total rate $P$, which can be 2, 4, 6, 8, 10, 12, 14, or 16 p/mn.

Among the output values described in Section 2.4, the results obtained for the headway standard deviation at the last bus stop, $\sigma_{Mh}$, are plotted in Fig. 2. Note that stops are renumbered according to direction, therefore the morning’s last stop is the evening’s first stop, and vice-versa. Headway standard deviation provides the best benchmark quantity, as it is not directly impacted by the varied inputs—in contrast to commercial speed, for example, which depends directly on preferred speed.

Headways are most disturbed by higher passenger demand In Figs. 2a, 2b (lengths and speeds), we observe that regularity is always better in the morning than the evening. However, comparing to Fig. 2c (passenger influx), the effect of increasing lengths and speeds on headways appears much weaker than the effect of increasing demand in the evening. Headway standard deviation increases quickly with demand along the evening profile of passenger influx. This stands in contrast to the morning profile, where this effect is much less pronounced. This stark discrepancy can be explained by the fact that the timetable is assumed to be respected at the departure point in the model, hence a higher demand in the morning should not affect it very much. But whereas passengers take a little time to exit buses, the time needed to board was estimated to be three times as high. Therefore, it is only natural that evening buses, which gradually fill up, experience more delays than morning buses, which are generally full at once. In conclusion, depending on the time of the day, increasing demand can have the strongest effect on regularity, essentially correlated with the time that buses spend at stops.
Service regularity loss in high-frequency feeder bus lines

(a) Test 4: section length interval.  
(b) Test 5: preferred speed deviation.  
(c) Test 6: passenger influx exponent.

Figure 3: Assessing heterogeneity effects: (a) Heterogeneity in section lengths does not disturb planning. (b) Heterogeneity in preferred speeds can disturb planning, but not in a realistic range (which should remain below ±4 km/h). (c) Heterogeneity in demand along stops is a clear cause of planning disturbance. See Section 3.1.2 for details.

3.1.2 Assessing heterogeneity effects

In the second batch of three tests, the effect of heterogeneity on the quality of service is studied by varying the standard deviation of the parameters’ distribution laws under constant mean:

- **Test 4**: distance dispersion \([d_-, d_+]\) is set to three ranges, \([1000, 1000]\), \([750, 1250]\), and \([500, 1500]\), keeping the mean at \((d_- + d_+)/2 = 1000\) m
- **Test 5**: preferred speed dispersion \(\sigma_{\text{pref}}^{\text{v}}\) is set to 0, 1, 2, 3, 4, and 5 km/h, while \(\mu_{\text{pref}}^{\text{v}} = 50\) km/h
- **Test 6**: passenger influx is increasingly diversified by setting \(\alpha\) to 0, 0.5, 1, 1.5, and \([P^-, P^+]\) to \([1, 1]\), \([0.1, 2.5]\), \([0.01, 8]\), and \([0.001, 20]\), respectively, then in each case renormalizing the \(P^n\)'s by \(P = 8p/\text{mn}\) and multiplying them by the \(D^n\)'s from the data.

As in the previous assessment, only the final headway standard deviation is examined (Fig. 3). Our observations follow.

**Heterogeneity in section lengths does not disturb planning**  Fig. 3a shows that there is practically no difference in the headway distribution when section lengths are diversified. Altogether, combined with Fig. 2a, we can conclude that section lengths do not have any significant effect on headways.

**Heterogeneity in preferred speeds can disturb planning, but not in a realistic range**  It can be seen in Fig. 3b that a higher dispersion among the preferred speeds of bus drivers is a clear factor of variability in headways, as one could expect. Compared to the default levels of Fig. 3a, however, this effect becomes noticeable only for \(\sigma_{\text{pref}}^{\text{v}} \geq 4\) km/h, which is unlikely to be the case in practice.

**Heterogeneity in demand along stops is a clear cause of planning disturbance**  In Fig. 3c, an increase of the exponent \(\alpha\) in the power-law distribution of passenger influx rates also creates more variable headways. Therefore, we conclude that more heterogeneity in passenger influx rates leads to less regularity.

3.2 Assessing preventive measures

In this section, we use our model to test self-driven measures, i.e. without the need for online central intervention, that may improve the quality of service on line 91.06C. We consider two main possibilities: *increasing the frequency of bus departure on the line, and implementing all-door boarding in buses*. The former is easily achieved in our model by reducing the value of parameter \(H_0\). The latter can be simulated through a drop in boarding time \(\rho\) in down to 1.5 s/passenger, i.e. half the level of the current one-door boarding scheme. Theoretically, it should be one third, since the articulated buses have three doors. Therefore, 1.5 s/p is a reasonable upper bound of the actual time per passenger. We look again at the effect on the final headway distribution via its standard deviation \(\sigma_M^h\). We also consider here its maximum \(h_{\text{max}}^M\), i.e. the maximum time interval without passing buses.

**Increasing the departure frequency reduces the maximum headway but not the dispersion**  Fig. 4a shows that smaller \(H_0\) values provoke a sharp decrease in maximum headway, but hardly modify headway variability. It means that this measure would not be effective: despite adding more buses to the line, the target frequency could still be missed. Moreover, supply was increased unilaterally in this virtual test, whereas in real life, economic considerations would condition it by an increase in user demand—which could eventually worsen the situation.
All-door boarding can improve service during the evening peak  In Fig. 4b we see that the effect on headway dispersion is significant in the evening, but inexistent in the morning. Implementing all-door boarding also slightly increases the average speed of buses in the evening. Therefore, this measure seems an effective way of combating irregularity in the case where there are many boarding passengers. These results suggest that opening all doors can alleviate the inherent loss of performance due to boarding, and restore some balance.

4 Discussion

We have proposed an agent-based model and simulation of bus transportation, focusing on the regularity of transit and avoidance of “clumping” as an essential criterion of quality of service. We studied the influence of various factors on this quantity, including distances between stops, preferred speeds of drivers, and planned frequency of bus departures. Based on real data, our statistical results indicate that user demand, represented by the influx of passengers at the bus stops, was the parameter with the most significant effect on the variability of time intervals, or “headways”, between vehicles at the final stop on the line. By contrast, section lengths or driving styles showed little to no impact on regularity.

Similar to ongoing research in this domain [6], we applied this model to a simulation of real-world behavior and the evaluation of self-regulation measures, in the sense that no central supervision or control were needed in real time. We identified all-door boarding to be the most helpful in minimizing the loss of regularity. The great advantage of this measure is its low-cost implementation, as it only involves installing fare-collector equipment at the bus doors and/or bus stops. One drawback, however, could be an increase in “free riders”, as the absence of direct control might give users an incentive to bend the rules. A cost-benefit analysis would need to be conducted to assess the weight of this potential issue. On the other hand, increasing the frequency of departure at the start of the line does not appear to be effective. It is also not economically feasible as it would need to be justified by an overall increase in demand volume, therefore probably cancelling any advantage it might bring.

The available data from line 91.06C allowed us to set most of the parameters of our model to realistic values. The hypotheses we made about speed levels are in accordance with the features of this particular line. Bus 91.06C partially runs on bus-only lanes, specially constructed for that purpose, and wherever it shares the road with other vehicles, traffic is never heavy in the directions and times of day that we considered here, therefore it could be neglected. It also seemed a reasonable assumption to model driving styles by a normal distribution, as it is a characteristic law of human behavioral diversity. In that case, we observed that the average preferred speed had no effect on regularity, while the width of its dispersion was only moderately influential. We also assumed that passengers flowed in at independent times, so they could be modeled via a Poisson process. Although it was verified by the data on this line, it may not always be a faithful representation of other types of discrete events, in particular the sudden and simultaneous appearance of dozens of users from the same location (such as an office building at closing time). Finally, other possible external factors were not taken into account in this model and should be the focus of future work: for example, more realistic interference from concurrent traffic on the roads and at intersections, through a bias applied on the average section speeds, or by using a secondary agent-based model to simulate the cars.
Acknowledgments

This research work was conducted under the leadership of the Technological Research Institute SystemX, supported by public funds within the scope of the French Program “Investissements d’Avenir”. It is part of the MIC project on Multimodal Transportation Systems in the context of the Systems of Systems program. Simulations were carried out with software by The CoSMo Company. We thank the Communauté d’agglomération du Plateau de Saclay (CAPS) and Albatrans (bus operator) for the provided data.

References


