# The modelling of urban running races 

Ricardo Enguiça1* © and Nuno D. Lopes ${ }^{1}$

"Correspondence: ricardo.roque@isel.pt
${ }^{1}$ Centro de Matemática Computacional e Estocástica FC-ULisboa and Instituto Politécnico de Lisboa, Instituto Superior de Engenharia de Lisboa, Rua Conselheiro Emídio Navarro, 1 -1950-062 Lisboa, Portugal


#### Abstract

In this paper, we model mass running urban races, taking into consideration several conditioning factors. The main goal is to find ideal configurations of the start of the race, splitting it into several waves, reducing the density of athletes and the overall time lost, when comparing the normal race results with a race without density constraints. This model takes into account distinct realistic runners' profiles, changes in slope and width on the race course and its influence on the running pace. Moreover, density levels, dynamics of the start of the race and time between the departure of waves are also considered.


Keywords: Running race; Density; Wave splitting; Numerical model

## 1 Introduction

Urban races with a massive number of participants are common on most major cities worldwide. The logistics for preparing these events have to take into consideration several factors that ensure not only the fast return to the normal functioning of the traffic, but also the satisfaction of runners, concerning total time spent, race conditions, etc. There are lots of different race settings: 10 km races, half-marathons, full marathons, or loop races with a different number of laps for different groups, and the number of participants can easily reach up to 40,000 for the longer ones. Obviously, the costs associated to security, traffic management and staff can reach significant values, so the general configuration is something that concerns organizing teams. On the other hand, it is usual for these events to be sponsored by companies (sometimes the event even takes the name of the sponsor); so, in order to keep financial support, organizers need to maintain high standards of service quality. This problem was presented as a challenge in an European Study Group for Industry some years ago by a company that organizes running events, and they were specifically concerned with 10 km races, in particular, on the split of runners into waves. The first steps towards the design of the model that we present were taken during that event.

The fact that on smaller courses the delays associated to density issues have a larger proportional impact on the final times of runners led us to focus our simulations on 10 km races instead of longer races. Loop races are also a challenge in terms of evaluating the best starting points for the different departures and the corresponding timings, and is something that can be addressed in a future work.
© The Author(s) 2023. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

This paper attempts to reproduce the usual dynamic of an urban race taking into account several distinct aspects that influence the behaviour of runners. The developed model also aims at evaluating and comparing different starting configurations that could be adopted by the organization, and to point in the direction of some strategies that will allow to reduce jams and congestions during the race, without impacting the competitive environment, while not increasing significantly the race total time. The company that presented the challenge was more concerned with the departure configuration, as the data that they provided showed that the main congestions were during that phase of the race.
On the one hand, to split the race into several waves, has implicitly associated the idea that to minimize the number of delayed runners, the slower ones should start in the last wave. But on the other, this has an implication on the total time of the race, since the slower runners are going to be the last ones to start running. This reasoning implies that a large number of waves could be counterproductive, since it would take longer to clear the involved infrastructures. This cost-benefit trade-off has to be quantified and in our model we consider a metric that penalizes departure delays and also the increase in the total time of the race.

To make the model more realistic, we also take into consideration the influence of elevation gain/loss and the presence of slow and fast runners in all waves. The parameters taken on the metric will be justified mainly with empirical arguments but, whenever possible, we will try to provide some data validation.

Other parameters, like the weather, or the runners' tiredness can be taken into account. Rainy conditions influence a runner's performance, because a more careful approach to other runners decreases speed, or clothes getting soaked increases the weight, slowing runners down. Wind also influences speed (differences can be substantial, but are transversal to all runners and can be considered as a factor similar to the elevation). On the other hand, hot and humid conditions will also decrease general performances. Concerning tiredness, the main drive for this work are the congestions at the beginning of the race, but considering a factor of tiredness for each runner might have a substantial impact on the final times of runners. The data set analysed only considers final times of athletes, so when we implement a profile, this factor is already included, but diluted over the whole race.

We emphasize that the set of parameters used are the ones that better adjust to the analysed videos and runners data, but the model is easily adapted for other sets, and the simulations that follow can also be made without any extra effort.
The overall mathematical model is an initial value problem (one dimensional in space), given by a system of ordinary differential equations (one equation for each runner) relating their position and speed. The speeds change over time in response to the distinct runner's profiles and density levels as well as slope and width of the race course. Moreover, we also stress that the time-step based numerical simulations are obtained using time-integration schemes for the sets of ordinary differential equations that compose the proposed mathematical model (see, e.g., Lambert [6]). The code is available on the public GitHub repository mentioned in the Declarations section and it is possible to change directly the relevant parameters and to obtain the corresponding simulations.
Concerning the literature on this subject, it is possible to find several papers related, but the emphasis is different from the one taken in our approach. In Treiber et al. [15], a traffic model is analysed for cross-country ski marathons, which represents an approach
to a non-vehicular traffic flow model, with the existence of lanes and taking into account different fitness levels, gradients, and interactions between the athletes in all traffic situations. The same author takes another approach in Treiber [14], where running marathons are evaluated, using a macroscopic model for non-congested flow, which is coupled to a kinematic-wave model for congested flow (under which all runners behave similarly, unlike in our model). The performance (i.e., free-flow speed distributions) of the athletes is taken into account in the different starting groups, and a relatively small number of race configurations is considered. That paper is mainly concerned in avoiding jams on a bottleneck in the course and the departure procedure is the same for all athletes, considering only the capacity of the starting section. In our approach, the starting procedure is based on the speed of each runner and the speed of the slowest runners ahead, on the width of the course, and we also define a minimum necessary free space ahead of a participant in order to start moving.
In Farina et al. [2], jams are addressed on a 20 km race, but on the final sector, as this specific race ends in a stadium. The given strategy relies on separating the finish line in several lanes, where athletes are directed to different parts of the stadium. On urban races, it is nowadays standard to place the finish line on relatively open spaces that avoid this possibility of overcrowding, and participants are often driven through a lane where they pick participation medals, water, etc.

Though our approach for the departure of the race is based on a microscopic approach, there are several studies concerning propagation of starting waves (see e.g. Tomoeda et al. [13]). Also related to running events, but on training strategies management, we also cite a very interesting recent paper by Roels [11], and from the physiological point of view, see e.g., Fiorini [3] and Pritchard [10]. In a recent survey (see Peckover et al. [9]) an analysis of congestion during events from the perspective of runners is made, and it is possible to see that $97 \%$ of runners have experienced congestion during the race start, which also motivates the more thorough analysis of the departure procedure of our model. In Haberman [4] there is a detailed analysis of the "Lighthill-Whitham-Richards" traffic model with cars that cannot overtake each other, done from the point of view of partial differential equations, considering a velocity field, the flow and density of cars (concerning numerical solution using finite differences methods, see Dazango [1]). This model is not suitable for our purpose since the velocity field approach does not allow different runners to have different speeds if they are in the same position of the course at a given instant, but from the point of view of the influence of density in the speed of runners, we take a similar approach. In this book there is also a section concerning the behaviour of cars after a traffic light turns green, which also has a correspondence in our model, when we define the delays and speeds at the race start. In May [7] we can also find the basic features of standard vehicle traffic models. A more recent approach using clustering analysis in high density situations is presented in Kerner and Konhäuser [5]. This paper is organized in the following way: in Sect. 2 we explain the details of the mathematical and numerical model; in Sect. 2.7 we establish a metric that will allow comparing the quality of different starting configurations; in Sect. 3, we analyse and compare different scenarios and, finally, in Sect. 4 we provide some conclusions based on our simulations and point in some directions in terms of future work.

## 2 Mathematical model

In this section we state the main details of our mathematical model for the speed of each runner. We consider a one-dimensional model in space, using a density function where the road width is considered, avoiding the two-dimensional case.
The mathematical model for all the runners is given by the following initial value problem (IVP):

$$
\left\{\begin{array}{l}
\mathbf{x}(0)=\mathbf{x}_{0},  \tag{1}\\
\dot{\mathbf{x}}(t)=\mathbf{v}(t), \quad \text { for } t>0 .
\end{array}\right.
$$

In order to numerically integrate system (1) of $N$ coupled ordinary differential equations we use a 2nd-order Adams-Bashforth-Moulton method (see, e.g., Lambert [6]). However, several other methods may be chosen in our computational code. For computational performance reasons, we implemented the model using the $C++$ language using the odeint library of the Boost $C++$ libraries (see, e.g., Schling [12]) where a large group of methods for IVP systems are available, including the several Adams-Bashforth-Moulton and Runge-Kutta methods.
Next we explain in more detail the several factors that influence our model.

### 2.1 Race start

We begin by displaying runners behind the starting line in rows of one person per meter, with the rows separated by half a meter. For example, if we take 4000 runners on a 10 m wide track, the starting wave would be spread along 200 m in 400 rows. Setting an initial velocity of $2.5 \mathrm{~ms}^{-1}$, and considering that a person only starts moving when there is 1 m of free space available in front, then each row takes about 0.4 s to start moving, which would imply that the last row would only move after 160 s .

This value of 0.4 s can also be associated to the natural human reaction time to a stimulation (see Nakamura [8]). For example, a typical professional sprinter has a reaction time to the starting signal between 0.1 s and 0.2 s . Considering that, after start running, all the runners run at speed $2.5 \mathrm{~ms}^{-1}$ until they cross the starting line, it would take 80 seconds more for the last runner to cross it, totalizing 240 s for the whole wave to be on course.
Considering a generic wave of $N$ runners on a $w$ meters wide track and an initial speed $v_{0}$, then the wave would spread on $N / 2 w$ meters (on $N / w$ rows), taking $N /\left(w v_{0}\right)$ seconds for the last row to start moving and $1.5 N /\left(w v_{0}\right)$ seconds to cross the starting line.

### 2.2 Density delay

In the normal development of the race, a runner is only going to be delayed by others if the number of runners in a certain space in front is large. We define vital space as being the region of the track with length $d$, starting at the position $x$ of the runner. Here we assume $d$ as the length in front of the runner that defines where the presence of other runners affects his speed. In Fig. 1 one can see the sketch of the vital space of the runner at the position $x$ (represented by the filled black circle). The runners in his vital space are represented by the red circles.
By the analysed videos, it seems reasonable to consider a rectangle of 1.33 m by 2 m as the minimum space that a runner needs free to keep his pace unchanged, and two rows of slower runners in front of someone with this minimum spacing would make it hard for


Figure 1 Sketch of the vital space of the runner
an overtake to be done smoothly. Considering the case of a constant width $w$ section, the vital space is a rectangle of $d w \mathrm{~m}^{2}$, where the presence of more than $1.5 w$ runners will be considered as an excessive density situation. Therefore, in our speed formula, we will only affect the speed of a runner due to density issues if the vital space has more that 1.5 w other runners.

If an untroubled runner has speed $v$, under density issues his speed will be affected. In this situation there are at least 1.5 w runners in the vital space, and we will consider that the level of influence of density is directly connected to the speed of the $w_{s}$ slowest runners in the vital area. In our model we considered $w_{s}=5$.

We denote the set of runners on the vital space of a given runner $i$ at time $t$ by $G_{i}(t)$, and the average speed of the 5 slowest elements in $G_{i}(t)$ by $v_{i}^{G}(t)$. The affected speed of the runner should be between $v$ and $v_{i}^{G}(t)$, and though it is expected that $v_{i}^{G}(t)<v$, the opposite can happen (for example if the runner had just been overtaken by several faster athletes). So, in order to avoid this situation, we consider the speed of the delaying runners as $v_{i}^{l}=$ $\min \left\{v, v_{i}^{G}(t)\right\}$, and the new speed of the runner is given by $v_{d}=(1-\rho) v+\rho v_{i}^{l}$, where $0 \leq$ $\rho \leq 1$. If the density is very large then the density weight $\rho$ will be closer to 1 , making the delayed speed similar to $v_{i}^{l}$. In our model, we will consider values of $\rho$ increasing gradually from 0.4 to 0.8 for a number of runners in $G_{i}(t)$ between $1.5 w$ and $2.5 w$, and $\rho=0.8$ if that number is above 2.5 w .
For future improvement of the model, a more thorough analysis of the impact of large densities could be made (possibly changing the maximum value of 0.8 , or having a nonlinear growth for mid-range values).

### 2.3 Probability distribution of runners' speed

Using the final times of the 10,000 runners in a 10 km race that took place in 2015 in Lisbon (São Silvestre of Lisbon race), we generate each runner's normal speed for a race simulation from the corresponding cumulative relative frequencies graphic (on the left of Fig. 2), which is obtained from the final times histogram, using a different bin per minute (on the right of the same Figure).
In order to avoid having many runners with the exactly same running profile, we use the polygonal line that approximates the corresponding continuous variable distribution function, instead of using the discrete version. On this continuous framework, to generate the speed of a runner from this distribution function, we just have to generate a random number $q$ on the interval [ 0,1 ], find the corresponding inverse through the continuous distribution function, and setting it as the normal speed of the runner. For example, if $q=0.4$, the corresponding finishing time for 10 km is about 55 minutes, and the speed $v=$ $\frac{10,000}{55 \times 60} \approx 3.03 \mathrm{~ms}^{-1}$. With this strategy, we get the speeds of runners in realistic proportions.


Figure 2 Cumulative relative frequencies (left graph) and final times histogram (right graph). Time in minutes is represented in the horizontal axis in both graphs


Figure 3 Linear regression of speed deviation by meter of elevation per km for a runner. Each color (from blue to red) of the scatter data stands for a a race made by the athlete

As stated before, we will assume this normal speed to be constant during the race, but if we were to consider a tiredness factor, this value would be a function of time instead. In the following subsection, we adapt this "normal" speed of the runner to the elevation gain/loss of the track.

### 2.4 Influence of elevation gain/loss

Analysing a considerable amount of real runners profiles, concerning their time per km and the accumulated elevation gain and loss per km, we were able to detect the most common deviations from the averaged race pace in the presence of a track with slope. Making a (through origin) linear regression between the difference in elevation (in meters) and the deviation of the pace in percentage, the deviations of the line were small, so we associate one of these regression lines to each runner. Obviously, this linear regression takes into consideration that there exists a realistic maximum absolute value of the slope, and we only consider races with slopes within this range of common slopes. A logistic regression might be a better approximation, but for the sake of simplicity, we only considered the linear case. In Fig. 3 one can see an example of such regression line for a given runner.

For this specific runner, each meter in elevation per km represents a decrease of $0.236 \%$ in the pace. For our runners database, we will use decreases in percentage from $0.12 \%$ to
$0.25 \%$ per meter of elevation in each km , in a range of -100 m to 100 m of elevation per km.

We could have taken a more rigorous model concerning the elevation gain/loss of the track by analyzing separately the influence of positive and negative slopes (there might be some slight differences in the total time in a km with constant slope equal to 0 and a km with 50 m of elevation gain followed by 50 m of elevation loss, or vice-versa).

### 2.5 Runners separation into several waves

By separating runners into several waves, departing at different times, we will be able to decrease the occurrences of large density situations. If this split is done taking into consideration the expected final time of the runners, this decrease would be much more evident. This strategy has been widely implemented, in particular it was applied on the race that we used as a model.
It is very common that runners are allocated to waves that do not correspond to their capacities: on one hand because they do not have official records in the previous year (starting in the last wave), and on the other hand because they managed to get a bib allowing to start in a wave with lower official time. We take this under consideration in our model, because it is this dynamic that creates more entropy within the waves, and also when runners of different waves meet more frequently (the slower of the front wave and the faster in the back wave). In this framework, we assume the existence of a percentage of runners belonging to a certain wave but that should be allocated to another one. Let us denote this percentage of runners that should be in wave $j$ but start in wave $i$ by $p_{i j}$. In order to minimize the values of $p_{i j}$, instead of asking for a proof of recent results on the type of the race, the organization could implement a strategy where the runners were driven to make an honest prediction of their final times (for example, a prediction failing by more that a certain amount of time would imply that in a future race that runner would have to start on the last wave). This strategy would not only avoid slower runners trying to start on a wave of faster runners, but would also allow faster runners that didn't have official times in the previous year to start in the corresponding wave.
In the configurations analysed ahead, we will be able to see the difference in time lost per runner if the runners are split randomly or taking into consideration the expected final times.
A suggestion that could have a positive impact on the degree of satisfaction of the participants would be to have a last wave for "charity runners", that is those that care more about "atmosphere and course features". The lost time of runners on this wave would not have any impact on the metric (influencing only the total time of the race).

### 2.6 Detailed mathematical model

In Table 1 we summarize the notations used.
The mathematical model given by the initial value problem (1) is such that $\mathbf{v}(t)$ is the speed vector with coordinates $v_{i}(t),(i=1, \ldots, N)$ defined by:

$$
v_{i}(t):= \begin{cases}0, & \text { if } t<\bar{d}_{i}^{w},  \tag{2}\\ v_{0, i}, & \text { if } t \geq \bar{d}_{i}^{w} \text { and } x_{i}(t)<0, \\ {\left[1-\rho_{i}(t, \mathbf{x})\right] v_{i}^{p}\left(x_{i}(t)\right)+\rho_{i}(t, \mathbf{x}) v_{i}^{l}(t),} & \text { otherwise }\end{cases}
$$

Table 1 Table of Notations

| Notation | Description | Units |
| :---: | :---: | :---: |
| N | number of runners in the race |  |
| i | runner's identification number ( $i \in\{1, \ldots, N\}$ ) |  |
| L | course's length | m |
| $t$ | time in seconds (race starts at $t=0$ ) | s |
| X | vector of the positions of all runners at time $t\left(\mathbf{x}:=\left(x_{1}(t), \ldots, x_{N}(t)\right)\right.$ | $m^{N}$ |
| $x$ | position on the course | m |
| $x_{0, i}$ | starting position of runner $i\left(x_{0, i} \leq 0\right)$ | m |
| $s(x)$ | slope of the course at position $x$ |  |
| $w(x)$ | width of the course at position $x$ | m |
| $m_{i}$ | deviation of the speed of runner $i$ per unit of slope | $\mathrm{ms}^{-1}$ |
| d | length (along the course) of the vital space | m |
| $\delta_{t}$ | time step | s |
| $G_{i}(t, \mathbf{x})$ | set of runners in the vital space of runner $i$ at time $t$ |  |
| $D_{i}(t, \mathbf{x})$ | number of runners per unit of area in the vital space |  |
| $v_{i}^{G}(t, \mathbf{x})$ | average speed of the 5 slowest runners of $G_{i}(t, \mathbf{x})$ | $\mathrm{ms}^{-1}$ |
| $v_{0, i}$ | speed of runner $i$ at departure | $\mathrm{ms}^{-1}$ |
| $\bar{v}_{i}$ | average speed of runner $i$ (fixed) | $\mathrm{ms}^{-1}$ |
| $\bar{V}_{0}^{w_{i}}$ | maximum speed at departure for runners in the wave of runner $i$ | $\mathrm{ms}^{-1}$ |
| $v_{i}^{p}(x)$ | speed profile of runner $i$ evaluated at position $x$ | $\mathrm{ms}^{-1}$ |
| $\underline{V}_{i}^{\prime}(t, \mathbf{x})$ | runner $i$ speed component due to density at time $t$ | $\mathrm{ms}^{-1}$ |
| $\bar{d}_{i}^{w}$ | departure time delay for runner $i$ (depends on wave and starting position) | s |
| $\rho_{i}(t, \mathbf{x})$ | density weight for runner $i$ at time $t$ |  |

with

$$
\begin{align*}
& v_{0, i}=\min \left\{\bar{v}_{i}, \bar{v}_{0}^{w_{i}}\right\}, \quad \text { (initial speed) },  \tag{3}\\
& v_{i}^{p}(x)=\bar{v}_{i}+m_{i} s(x), \quad(\text { speed profile })  \tag{4}\\
& v_{i}^{l}(t, \mathbf{x})=\min \left\{v_{i}\left(t-\delta_{t}\right), v_{i}^{G}(t, \mathbf{x})\right\} \quad \text { (density speed) } \tag{5}
\end{align*}
$$

and

$$
\rho_{i}(t, \mathbf{x})= \begin{cases}0, & D_{i}(t, \mathbf{x})<0.375  \tag{6}\\ \frac{1}{0.625}\left(D_{i}(t, \mathbf{x})-0.125\right), & 0.375 \leq D_{i}(t, \mathbf{x}) \leq 0.625 \\ 0.8, & D_{i}(t, \mathbf{x})>0.625\end{cases}
$$

This last formula defines a density weight $\rho_{i}$ directly in terms of the density $D_{i}$, which is defined by:

$$
\begin{equation*}
D_{i}(t, \mathbf{x})=\frac{\#\left\{j \in\{1, \ldots, N\}: x_{i}(t) \leq x_{j}(t) \leq x_{i}(t)+d\right\}}{\int_{x_{i}(t)}^{x_{i}(t)+d} w(x) \mathrm{d} x} . \tag{7}
\end{equation*}
$$

Equation (3) sets the speed of runner $i$ before crossing the starting line as the minimum between his normal speed and the maximum speed of his wave; equation (4) computes the speed of a runner at a given point $x$ of the course; and finally, equation (5) states that the speed of the runner $i$ is affected by the speeds of other runners in his vital space. Moreover, this component of the speed also depends on the speed of runner in the previous time step $\left(t-\delta_{t}\right)$.

### 2.7 Metric for the quality of the split

In order to evaluate the quality of the split of the race into several waves, we use for comparison the race with the same runners in the same initial configuration (connecting waves one after another) without any density delays (that is, removing from the speed model the impact of density). These delays can be split into the lost time in the departure, and the lost time after crossing the starting line.
For a race split into several waves, the official final time of each runner is the time taken between crossing the starting line and crossing the finish line. So when we compare runners' times with or without density taken into account, it is on the difference of these official times that we focus. Though the time that a runner takes until they reach the starting line is not taken into consideration for the official record, this delay is always a source of stress, so we also used it in our metric.
Just taking into account the total time lost due to density delays (adding the delays of all runners) might seem like a reasonable metric, but when a provided service has complaints, the overall impact is not proportional to the degree of dissatisfaction, taking also into consideration the amount of complaints. To include these two features, we use in our metric the following weights $\left(w_{d}\right)$ : for the time lost below $t_{l}=30 \mathrm{~s}$, this value $t_{l}$ is multiplied by $w_{d}=2$ (we denote the total time lost in these circumstances by $T_{1}$ ); for the values $30 \leq t_{l}<60$, then $w_{d}=1.5$ (total time denoted by $T_{2}$ ); for $60 \leq t_{l}<120, w_{d}=1.25$ (total time denoted by $T_{3}$ ); and for $t_{l} \geq 120$, then $w_{d}=1$ (total time denoted by $T_{4}$ ). So 20 runners with a delay of 25 s each contribute with a total of $20 \times 25 \times 2=1000$ units of delay in our metric, while 10 runners with 20 s of delayed time each contribute with a total of $10 \times(30 \times 2+20 \times 1.5)=900$ units and 4 runners with 125 s of delayed time contribute with $4 \times(30 \times 2+30 \times 1.5+60 \times 1.25+5 \times 1)=740$ units (note that the total delay in all three examples is the same, but the metric value is smaller for the cases where less runners are delayed).
Allocating a runner in a wave other than the first one creates a natural tension associated to the fact that they know that there are already runners in action. To put this aspect in our metric, for each runner that doesn't start in the first wave, we apply an extra delay of 5 s multiplied by the number of waves that start before his one. This penalizes a larger number of waves.
The general degree of upsetness of a runner starting on a later wave could be tackled from a different perspective in a future development of this model: other options such as having a more individualized penalty could be more accurate and can possibly provide a more realistic metric than the used one.
For the time between the official start of a wave and the moment that each runner crosses the starting line (as mentioned before, this time is not added to the official time of the runner), we assigned a smaller impact in the metric, assigning it the weight $w=0.2$ (we denote the total time lost in this situation by $T_{0}$. For example, if a runner takes 60 seconds to cross the starting line, then only 12 units are added on the metric.

The total span time of the race is also included as a factor. We define the span time of the race $T$ as the total time of the race without time gap between waves (in this case, we settle a gap of 1 second in the data provided in the tables ahead). Considering several waves, let $p$ represent the proportion of extra time taken due to the splitting runners into waves. Then the metric value is also increased, but by a factor of $1+p / 2$ (since smaller waves will also
imply a shorter delay in the runners of the last wave, we reduce the impact of this feature by a half).
Finally, for the final value of the metric, we rescale the total obtained sum for each runner, that is, we divide by the total number of runners.
With the notations introduced above, the formula for the metric is the following:

$$
\begin{equation*}
\frac{1}{N}\left(0.2 T_{0}+2 T_{1}+1.5 T_{2}+1.25 T_{3}+T_{4}+\sum_{i=2}^{k} 5(i-1)\left(\# w_{i}\right)\right)\left(1+\frac{p}{2}\right) \tag{8}
\end{equation*}
$$

where $\# w_{i}$ is the number or runner on wave $i$ and $k$ is the total number of waves.

## 3 Numerical tests

Settled the metric, we now pass to the comparison of different scenarios, concerning the initial disposition of the runners in waves. The objective of our simulations relies on evaluating the impact of the course design (namely changes in slope and width), the number of waves, the unevenness of the wave splitting, the degree of mixture of runners' performance in each wave and the time gap between the start of each wave. Several scenarios will be simulated and the useful information for the analysis of each of them will be stated in a corresponding table. In all these scenarios we assume $N=10,000, L=10,000$ and $d=4$.
We note that all the simulations in this work were made on a computer running Linux Ubuntu 18.04 equipped with an AMD Ryzen 9 3950X processor, and each simulation of a full race took approximately 80 s . Moreover, all the pre and post processing was made using Python scripting language.
In Tables 2-9 of the following subsections, the first column gives us the time gap (in seconds) between the last runner of a wave crossing the starting line and the first runner of the next wave to start running. When this gap is equal to 1 , this basically means that the waves are fictitious and that all the runners start at the same time. The second column characterizes the first wave and the following have the same information for the other waves. If the number of waves is equal to 2 , the runners are split into two groups according to their abilities, separated by an associated quantile value. Each column associated to a wave is split in two columns labeled "mixture" and "wdt". In the "mixture" column there is a vector associated to the mixture level. The first value in this vector is the number of runners of the first group that are assigned to the first wave, the second value is the number of runners of the second group on the first wave (for the cases with more waves the reasoning will be the same). In the "wdt" column we show the wave departure time, i.e., the time passed since the beginning of the race for the runners of this wave to start running. Since we evaluate different gap times for the same level of mixture, we just mention the $w d t$ for the lines corresponding to gaps different than 1 . On the top of each column related to a wave it is also stated the maximum speed (wms) of the runners of this wave before the starting line.

On the third-to-last column it is stated the average time lost per runner when comparing with the same race without the density delay factor, on the second-to-last one we have the total time of the race and on the last, the metric value.

At the end of the section we make a brief description of the procedure to find an ideal configuration.

### 3.1 Course design 1: 10 km wide flat course

We start by analyzing the case of a 10 km course, with constant width of 10 m and without elevation gain/loss.
In the Tables 2 and 3 we have the information concerning a split into two waves: in Table 2, two waves of 5000 runners; and in Table 3, a first wave with 2500 runners and a second with 7500 .

Obviously, the best value for the metric is attained when the waves don't have any mixture and no gap time between then, but this is an unrealistic scenario since it is not possible

Table 2 Scores as function of time gaps and wave mixtures - 10 m wide flat course. wms - wave maximum speed before the starting line; wdt - wave departure time; tlpr - time lost per runner

| Gap [s] | Wave $1(\mathrm{wms}=3.34)$ |  | Wave $2(\mathrm{wms}=2.92)$ |  | tlpr | total time | metric score |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | mixture | wdt | mixture | wdt |  |  |  |
| 1 | $(5000,0)$ | 0 | $(0,5000)$ | 288 | 82.5 | 6594 | 154.3 |
| 300 |  | 0 |  | 587 | 82.5 | 6893 | 157.8 |
| 1 | $(4900,100)$ | 0 | $(100,4900)$ | 319 | 85.1 | 6625 | 157.7 |
| 60 |  | 0 |  | 378 | 84.0 | 6684 | 157.1 |
| 180 |  | 0 |  | 498 | 84.1 | 6804 | 157.8 |
| 300 |  | 0 |  | 618 | 83.7 | 6924 | 159.5 |
| 1 | $(4000,1000)$ | 0 | $(1000,4000)$ | 347 | 93.6 | 6653 | 168.4 |
| 60 |  | 0 |  | 406 | 89.0 | 6712 | 163.8 |
| 180 |  | 0 |  | 526 | 84.4 | 6832 | 159.8 |
| 300 |  | 0 |  | 646 | 82.6 | 6952 | 159.2 |
| 1 | $(3000,2000)$ | 0 | $(2000,3000)$ | 347 | 128.8 | 6653 | 207.5 |
| 60 |  | 0 |  | 406 | 114.6 | 6712 | 193.0 |
| 180 |  | 0 |  | 526 | 95.1 | 6832 | 172.6 |
| 300 |  | 0 |  | 646 | 87.6 | 6952 | 165.0 |
| 1 | $(2500,2500)$ | 0 | $(2500,2500)$ | 347 | 157.3 | 6653 | 237.1 |
| 60 |  | 0 |  | 406 | 142.8 | 6712 | 223.0 |
| 180 |  | 0 |  | 526 | 112.9 | 6832 | 193.1 |
| 300 |  | 0 |  | 646 | 94.3 | 6952 | 177.4 |

Table 3 Scores as function of time gaps and wave mixtures - 10 m wide flat course. wms - wave maximum speed before the starting line; wdt - wave departure time; tlpr - time lost per runner

| Gap [s] | Wave 1 (wms = 3.34) |  | Wave 2 (wms = 2.92) |  | tlpr | total time | metric score |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | mixture | wdt | mixture | wdt |  |  |  |
| 1 | $(2500,0)$ | 0 | $(0,7500)$ | 139 | 103.8 | 6611 | 185.5 |
| 300 |  | 0 |  | 438 | 103.8 | 6910 | 189.7 |
| 1 | $(2475,25)$ | 0 | $(25,7475)$ | 156 | 104.2 | 6628 | 186.0 |
| 60 |  | 0 |  | 215 | 104.2 | 6687 | 186.8 |
| 180 |  | 0 |  | 335 | 104.5 | 6807 | 188.8 |
| 300 |  | 0 |  | 455 | 104.0 | 6927 | 190.0 |
| 1 | $(2000,500)$ | 0 | $(500,7000)$ | 169 | 113.1 | 6641 | 195.9 |
| 60 |  | 0 |  | 228 | 110.4 | 6700 | 183.6 |
| 180 |  | 0 |  | 348 | 108.3 | 6820 | 193.0 |
| 300 |  | 0 |  | 468 | 106.3 | 6940 | 192.5 |
| 1 | $(1500,1000)$ | 0 | $(1000,6500)$ | 169 | 123.5 | 6641 | 208.2 |
| 60 |  | 0 |  | 228 | 117.4 | 6700 | 201.5 |
| 180 |  | 0 |  | 348 | 110.5 | 6820 | 195.5 |
| 300 |  | 0 |  | 468 | 106.9 | 6940 | 193.4 |
| 1 | $(1250,1250)$ | 0 | $(1250,6250)$ | 169 | 135.7 | 6641 | 221.9 |
| 60 |  | 0 |  | 228 | 124.1 | 6700 | 209.2 |
| 180 |  | 0 |  | 348 | 113.4 | 6820 | 198.6 |
| 300 |  | 0 |  | 468 | 109.4 | 6940 | 196.0 |

Table 4 Scores as function of time gaps and wave mixtures - flat 10 m wide course. wms - wave maximum speed before the starting line; wdt - wave departure time; tlpr - time lost per runner

| Gap [s] | Wave $1(\mathrm{wms}=3.34)$ |  | Wave 2 ( $\mathrm{wms}=2.92$ ) |  | Wave 3 (wms = 2.50) |  | tlpr | total time | metric score |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | mixture | wdt | mixture | wdt | mixture | wdt |  |  |  |
| 1 | (3333, 0, 0) | 0 | (0,3333, 0) | 186 | (0,0,3334) | 384 | 58.4 | 6593 | 117.3 |
| 300 |  | 0 |  | 485 |  | 982 | 58.0 | 7191 | 118.7 |
| 1 | $(3000,166,167)$ | 0 | $(166,3000,167)$ | 213 | $(167,167,3000)$ | 426 | 66.9 | 6635 | 129.4 |
| 60 |  | 0 |  | 272 |  | 544 | 65.7 | 6753 | 128.7 |
| 120 |  | 0 |  | 332 |  | 664 | 63.9 | 6873 | 127.7 |
| 180 |  | 0 |  | 392 |  | 784 | 63.1 | 6993 | 127.8 |
| 300 |  | 0 |  | 512 |  | 1024 | 61.3 | 7233 | 127.7 |
| 1 | $(2333,500,500)$ | 0 | $(500,2333,500)$ | 215 | $(500,500,2334)$ | 442 | 84.2 | 6651 | 150.3 |
| 60 |  | 0 |  | 274 |  | 560 | 75.3 | 6769 | 140.5 |
| 120 |  | 0 |  | 334 |  | 680 | 69.8 | 6889 | 135.0 |
| 180 |  | 0 |  | 394 |  | 800 | 65.9 | 7009 | 131.5 |
| 300 |  | 0 |  | 514 |  | 1040 | 60.8 | 7249 | 127.3 |

to control the mixture level of the waves. For example, with only 100 mixed runners, the best performance is obtained for a gap of 60 s (which corresponds to the departure of the second wave 378 s after the first), but for a mixture of 2500 runners (this would roughly match a random selection for the runners in each wave) the lowest result for the metric is obtained for a gap of 300 s , that corresponds to almost 11 minutes between the start of both waves. With this random selection mixture, we can see how much time is lost in average if this factor is not taken into account.

With this, we can say that it is crucial to separate runners by pace with the smallest possible degree of mixture, as the average time lost can almost double if runners aren't separated at all, as we can see at the bottom of both tables.
With the same level of mixture, the longer the gap, the less concentration issues appear, but it becomes clear that there is a threshold where the impact becomes almost negligible: for mixtures above $35 \%$, separating waves more than 180 seconds is less effective in terms of average time lost.

Comparing Tables 2 and 3 shows that uneven waves are obviously more delayed in average (roughly around $25 \%$ on this particular case), but we can also see that the impact of the degree of mixture is about the same than for the even sized waves case.

In Tables 4 and 5 we present the information for the same flat course for the case of three waves instead of two, also on both scenarios where the waves are split evenly and unevenly.

Through a quick analysis of Tables 4 and 5, we can infer that the presence of a third wave allows the average time lost per runner to decrease about $30 \%$. The fact that the waves are smaller makes the impact of mixtures slightly more evident on the 3 waves case. The difference on the total time of the race is also clear, increasing it slightly.

Concerning the influence of the width of the course for the flat race without any wave splitting, Fig. 4 allows us to evaluate the changes on the total time spent and also on the metric.

### 3.2 Course design 2: 10 km with variable width and slope

Next we analyse the influence of having slopes and narrow parts on the course. We used a course whose properties are sketched on Fig. 5.

Table 5 Scores as function of time gaps and wave mixtures - flat 10 m wide course. wms - wave maximum speed before the starting line; wdt - wave departure time; tlpr - time lost per runner

| Gap [s] | Wave 1 (wms = 3.34) |  | Wave 2 (wms = 2.92) |  | Wave 3 (wms = 2.50) |  | tlpr | total time | metric score |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | mixture | wdt | mixture | wdt | mixture | wdt |  |  |  |
| 1 | (2500, 0, 0) | 0 | (0, 2500, 0) | 139 | (0,0,5000) | 284 | 67.9 | 6600 | 132.2 |
| 300 |  | 0 |  | 438 |  | 882 | 67.9 | 7198 | 138.2 |
| 1 | $(2200,100,200)$ | 0 | $(100,2200,200)$ | 170 | $(200,200,4600)$ | 331 | 75.5 | 6647 | 142.8 |
| 60 |  | 0 |  | 229 |  | 449 | 73.5 | 6765 | 141.5 |
| 120 |  | 0 |  | 289 |  | 569 | 72.1 | 6885 | 141.0 |
| 180 |  | 0 |  | 349 |  | 689 | 71.0 | 7005 | 140.8 |
| 300 |  | 0 |  | 469 |  | 929 | 69.9 | 7245 | 141.8 |
| 1 | $(1900,200,400)$ | 0 | $(200,1900,400)$ | 170 | $(400,400,4200)$ | 336 | 80.4 | 6652 | 148.8 |
| 60 |  | 0 |  | 229 |  | 454 | 76.2 | 6770 | 144.7 |
| 120 |  | 0 |  | 289 |  | 574 | 72.4 | 6890 | 141.2 |
| 180 |  | 0 |  | 349 |  | 694 | 70.8 | 7010 | 140.3 |
| 300 |  | 0 |  | 469 |  | 934 | 67.9 | 7250 | 138.9 |

Figure 4 Total time lost and metric value as a function of the width of the course. Total race times in seconds are also shown for each course width. Flat constant width courses



Figure 5 Characterization of the course used in the simulations of Tables 6 and 7. Elevation profile of the course at the left graph and the width profile at the right graph

As we can see, this course has slopes taking values from $-4 \%$ to $4 \%$ and the width varies from 2 to 11 meters. This width is very narrow on a zone of maximum slope (around 4000 m after the starting line), and this potentiates concentration issues, in order to test our model under rough conditions.
Table 6 gives us the data for this course with the runners split into 3 even waves.
Comparing the results with the corresponding ones from the flat race, this particular changes in slope and width imply an increase on the overall time lost of about $50 \%$, and on Table 7, we can see that this value is slightly less on case of uneven waves. So it is clear that the design of the course has a determinant impact on the activation of the concentration delays that the metric uses, implying significant changes on the runners' performances.

Table 6 Scores as function of time gaps and wave mixtures - non flat and non constant width course. wms - wave maximum speed before the starting line; wdt - wave departure time; t/pr - time lost per runner

| Gap [s] | Wave 1 (wms = 3.34) |  | Wave 2 (wms = 2.92) |  | Wave 3 (wms = 2.50) |  | tlpr | total time | metric score |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | mixture | wdt | mixture | wdt | mixture | wdt |  |  |  |
| 1 | $(3333,0,0)$ | 0 | $(0,3333,0)$ | 186 | $(0,0,3334)$ | 384 | 89.9 | 6655 | 158.2 |
| 300 |  | 0 |  | 485 |  | 982 | 89.4 | 7253 | 165.3 |
| 1 | $(3000,166,167)$ | 0 | $(166,3000,167)$ | 213 | $(167,167,3000)$ | 426 | 110.4 | 6697 | 183.6 |
| 60 |  | 0 |  | 272 |  | 544 | 108.5 | 6815 | 183.0 |
| 120 |  | 0 |  | 332 |  | 664 | 106.9 | 6935 | 182.6 |
| 180 |  | 0 |  | 392 |  | 784 | 105.2 | 7055 | 182.1 |
| 300 |  | 0 |  | 512 |  | 1024 | 102.3 | 7295 | 181.7 |
| 1 | $(2333,500,500)$ | 0 | $(500,2333,500)$ | 215 | $(500,500,2334)$ | 442 | 146.5 | 6713 | 222.4 |
| 60 |  | 0 |  | 274 |  | 560 | 140.4 | 6831 | 217.8 |
| 120 |  | 0 |  | 334 |  | 680 | 134.8 | 6951 | 213.6 |
| 180 |  | 0 |  | 514 |  | 1040 | 129.2 | 7071 | 209.4 |
| 300 |  | 0 |  | 634 |  | 1280 | 119.3 | 7311 | 201.8 |



Figure 6 A snapshot of the last simulation of the Table 6

Table 7 Scores as function of time gap and mixtures - non flat and non constant width course. wms - wave maximum speed before the starting line; wdt - wave departure time; tlpr - time lost per runner

| Gap [s] | Wave 1 (wms = 3.34) |  | Wave 2 ( $\mathrm{wms}=2.92$ ) |  | Wave 3 (wms = 2.50) |  | tlpr | total time | metric score |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | mixture | wdt | mixture | wdt | mixture | wdt |  |  |  |
| 1 | (2500, 0, 0) | 0 | (0,2500, 0 ) | 139 | (0,0,5000) | 284 | 102.7 | 6688 | 175.7 |
| 300 |  | 0 |  | 438 |  | 882 | 102.8 | 7286 | 183.7 |
| 1 | (2200, 100, 200) | 0 | $(100,2200,200)$ | 170 | $(200,200,4600)$ | 331 | 122.2 | 6735 | 199.7 |
| 60 |  | 0 |  | 229 |  | 449 | 120.9 | 6853 | 199.8 |
| 120 |  | 0 |  | 289 |  | 569 | 118.7 | 6973 | 199.2 |
| 180 |  | 0 |  | 349 |  | 689 | 117.0 | 7093 | 198.9 |
| 300 |  | 0 |  | 469 |  | 929 | 114.1 | 7333 | 198.6 |
| 1 | $(1900,200,400)$ | 0 | $(200,1900,400)$ | 170 | $(400,400,4200)$ | 336 | 136.5 | 6740 | 215.4 |
| 60 |  | 0 |  | 229 |  | 454 | 133.9 | 6858 | 214.5 |
| 120 |  | 0 |  | 289 |  | 574 | 130.5 | 6978 | 212.5 |
| 180 |  | 0 |  | 349 |  | 694 | 127.7 | 7098 | 211.2 |
| 300 |  | 0 |  | 469 |  | 934 | 120.6 | 7338 | 206.4 |

In Fig. 6 we can see a snapshot of the simulation in the last line of in Table 6. This snapshot is taken at time 43 min 37 s , and the runners of each wave are sketched with points of different colors. This shows how the three different waves mix during the race.

In Table 7 we present the results for the uneven waves on this course, where again it becomes clear that this increases the time lost and the metric values.

Table 8 Scores as function of time gaps and wave mixtures - São Silvestre course. wms - wave maximum speed before the starting line; wdt - wave departure time; tlpr - time lost per runner

| Gap [s] | Wave 1 (wms = 3.34) |  | Wave 2 (wms = 2.92) |  | Wave 3 (wms = 2.50) |  | tlpr | total time | metric score |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | mixture | wdt | mixture | wdt | mixture | wdt |  |  |  |
| 1 | (3333, 0, 0) | 0 | (0,3333, 0) | 117 | (0,0,3334) | 241 | 64.8 | 6460 | 120.3 |
| 300 |  | 0 |  | 416 |  | 839 | 65.0 | 7058 | 126.2 |
| 1 | $(3000,166,167)$ | 0 | $(166,3000,167)$ | 134 | $(167,167,3000)$ | 268 | 77.2 | 6487 | 136.7 |
| 120 |  | 0 |  | 253 |  | 506 | 72.7 | 6725 | 133.4 |
| 300 |  | 0 |  | 433 |  | 866 | 68.9 | 7085 | 131.9 |
| 1 | $(2333,500,500)$ | 0 | $(500,2333,500)$ | 135 | $(500,500,2334)$ | 277 | 109.7 | 6494 | 174.0 |
| 120 |  | 0 |  | 254 |  | 515 | 94.2 | 6732 | 158.4 |
| 300 |  | 0 |  | 434 |  | 875 | 78.1 | 7092 | 143.0 |



Figure 7 A snapshot of the last simulation of the Table 8

### 3.3 Course design 3: São Silvestre of Lisbon race

An approximate course of the race where we based our model, the São Silvestre of Lisbon race, is analysed in this subsection.

In Table 8 we have the data for 3 waves, split evenly. The uneven cases in the following will no longer be considered, since it already became clear that it increases unnecessarily the race time and delays.

In Fig. 7, as in the previous course, we provide a snapshot of the simulation of this race at time 33 min 45 s . On this picture we can also see the width of the course and the slope. This race starts with a 16 meter wide course, with a steep descent, and after 1000 m this width already decreased to 6 m . There is a long part of the course with zero slope, and on the final 3500 m , the course has a first part with a steep ascent and a final part with a steep descent until the finish line.

### 3.4 Course design 4: inverted São Silvestre of Lisbon race

In order to analyse the possible effect of starting and ending the race ascending instead of descending, we have also computed the simulated results for the São Silvestre course, but inverting the direction of the runners.
The obtained results are fairly similar to the ones from the original course. The fact that the race starts ascending makes the start of the race slower, but since all the runners are slower on those circumstances, the delays are about the same in average. The total time of the race is not really affected by the inversion, which makes us believe that it is mostly the difference between the elevation gain and loss that has an evident impact on the time lost.

Table 9 Scores as function of time gaps and wave mixtures - inverted São Silvestre course. wms wave maximum speed before the starting line; wdt - wave departure time; tlpr - time lost per runner



Figure 8 Inverted course S. Silvestre: A snapshot of the last simulation of Table 9

Finally, in Table 9 we evaluated the case where we split the runners into 4 even waves of 2500 runners (on the inverted São Silvestre course).
The time lost in average by each runner decreases about $20 \%$, but the total time of the race increases on the same scale. The metric values are smaller, even though the total time race is significantly larger, so if this is an aspect that should be avoided, we can increase the penalty for the total time, or simply ask for the total time race to be below a given threshold, and therefore not allowing the split into to many waves. In Fig. 8 one can see a snapshot of the last simulation of the previous table at time 34 min 12 s .

### 3.5 Procedure to find ideal departure configuration

We finish by presenting an algorithm that synthesises the procedure to be taken by race organizers in order to find an ideal configuration.

1. As initial inputs, the course configuration and a maximum for the total time of the race ( $T_{\max }$ ) should be defined.
2. By running the model for an increasing number of waves separated by 300 seconds without mixture and comparing $T_{\max }$ with the simulation total time, find a
maximum number of waves $M$ (it is not realistic to consider a number of waves greater than 5).
3. Define the expected levels of mixture between waves for a number of waves between 2 and $M$.
4. For the considered mixture levels, run the simulation for several gap values, and for a number of waves between 2 and $M$ (stopping increasing the gap time when the total race time exceeds $T_{\max }$ or if the metric value is evidently increasing).
5. Find the configuration that gives the lowest metric value.

## 4 Conclusions and future work

With the data provided in the previous section, it became clear that the level of mixture of runners in different waves is a crucial factor for the metric values. This is not a factor that can be directly controlled and it should be minimized by the organization by taking measures in that direction. An ideia that can increase the general degree of satisfaction is to set a last wave with the assumed non-competitive runners and also to establish conversion times from races of different lengths made by a runner in the previous year to indicate in which wave should they start. This would help to decrease the variance on the competitive waves.

Concerning the number of waves, with the metric that we developed, by taking 4 waves the results were better, making this split the best alternative. But for different parameters, a smaller number of waves could become a better solution, for example, if the increase on the total time of the race is penalized by a larger factor, the ideal number of waves can be 3 or even 2 . Also if the value added to the metric associated to a runner being assigned to a later wave is larger, this could imply that the number of waves should be smaller. The unevenness of the split should always be avoided, since the best results were obtained for waves with the same number of runners.

The ideal gap between waves depends significantly on the level of mixture of runners and also on the number of waves, but for the analysed cases, a gap smaller than 180 s usually provided worse results.

It also became clear that the course design severely affects the metric performance, and the delays are much smaller for the more regular courses (without many changes in slope and course width). Following the results of the "inverted" São Silvestre race, we were also able to infer that the metric results are only affected by the total amount of elevation gain/loss, and not by the order in which elevation changes.

Overall, the implemented model replicates the most important dynamics on this type of races, but in the future, improvements on the set of runners' profiles can be made, namely concerning tiredness, the nonlinear influence of slopes or the race momentum. but these changes would be mainly related to parameter tuning, rather that a significant change in the model design.
As stated before, loop races also have specific issues that are of interest to study, as the evaluation of the best starting points and timings, and is something that can be addressed in a future work. Longer races, like marathons also move thousands of participants around the world, and the fact that the long distance somehow changes the focus from the organizing decisions to the runners decisions implies a different type of model.

## Acknowledgements

The authors would like to express their gratitude to the other members of Challenge 3 in the 140th European Study Group with Industry (Barreiro, Portugal, June 4-8, 2018) for the productive discussions, which were the embryo for the work presented in this article. This work is financed by national funds through FCT projects UIDB/04621/2020 and UIDP/04621/2020 of CEMAT at FC-Universidade de Lisboa.

## Funding

This research received no external funding.

## Availability of data and materials

The code for simulations and supporting data are available at the repository https://github.com/ndlopes-github/BoostRace.

## Declarations

## Competing interests

The authors declare no competing interests.

## Author contributions

The authors contributed to the manuscript equally. All authors have read and approved the manuscript.
Received: 30 December 2022 Accepted: 11 July 2023 Published online: 02 August 2023

## References

1. Daganzo C. The cell transmission model: network traffic. Transp Res, Part B. 1995;29(2):79-93.
2. Farina R, Kochenberger G, Obremski T. The computer runs the bolder boulder: a simulation of a major running race. Interfaces. 1989;19(2):48-55. https://doi.org/10.1287/inte.19.2.48
3. Fiorini C. Optimization of running strategies according to the physiological parameters for a two-runners model. Bull Math Biol. 2017;79(1):143-62. https://doi.org/10.1007/s11538-016-0230-9. PMID: 27826878. Epub (2016).
4. Haberman R. Mathematical models. Englewood Cliffs: Prentice Hall; 1977.
5. Kerner BS, Konhäuser P. Structure and parameters of clusters in traffic flow. Phys Rev. 1994;50:54-83.
6. Lambert JD. Numerical methods for ordinary differental systems: the initial value problem. New York: Wiley; 1991.
7. May A. In: Traffic flow fundamentals. Englewood Cliffs, N.J. 1990.
8. Nakamura H. An experimental study of reaction time of the start in running a race. Res Q Am Phys Educ Assoc. 1934;5(sup1):33-45. https://doi.org/10.1080/23267402.1934.10761656.
9. Peckover S, Raineri A, Scanlan A. An analysis of congestion during running events from the perspective of runners: prevalence, impact on safety and satisfaction, and preferred controls. Event Manag. 2022;26(5):967-78.
10. Pritchard WG. Mathematical models of running. SIAM Rev. 1993;35(3):359-79. http://www.jstor.org/stable/2132424.
11. Roels G. High-performance practice processes. Manag Sci. 2020;66(4):1509-26. https://doi.org/10.1287/mnsc.2019.3286.
12. Schling B. The Boost C++ Libraries. XML Press; 2011.
13. Tomoeda A, Akiyasu D, Imamura T, Nishinari K. Propagation speed of a starting wave in a queue of pedestrians. Phys Rev E. 2012;86(3):036113.
14. Treiber M. Crowd flow modeling of athletes in mass sports events: a macroscopic approach. In: Chraibi M, Boltes M, Schadschneider A, Seyfried A, editors. Traffic and granular flow '13. Cham: Springer; 2015. https://doi.org/10.1007/978-3-319-10629-8_3.
15. Treiber M, Germ R, Kesting A. From drivers to athletes: modeling and simulating cross-country skiing marathons. In: Chraibi M, Boltes M, Schadschneider A, Seyfried A, editors. Traffic and granular flow'13. vol. 9. Cham: Springer; 2015. https://doi.org/10.1007/978-3-319-10629-8_2.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Submit your manuscript to a SpringerOpen ${ }^{\bullet}$ journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at $>$ springeropen.com

