TRANSPORT PROBLEMS 2021 Volume 16 Issue 1
PROBLEMY TRANSPORTU DOI: 10.21307/tp-2021-010

Keywords: cyclists; road safety; fatigue; classification; pattern recognition

Gaetano BOSURGI, Orazio PELLEGRINO*
University of Messina, Department of Engineering
Contrada di Dio – Villaggio Sant'Agata, Messina, Italy
Joanna WACHNICKA, Romanika OKRASZEWSKA
Gdansk University of Technology, Department of Highway and Transportation Engineering
Narutowicza 11/13, 80-233 Gdańsk, Poland
**Corresponding author*. E-mail: opellegrino@unime.it

ROAD SAFETY FOR CYCLISTS BASED ON THE CALORIES NEEDED

Summary. Cyclists are a vulnerable group of road users, especially when no separate infrastructure for cyclists is provided. Then, road factors such as distance and altitude differences can indirectly affect cyclists' safety. Therefore, the authors proposed a procedure based on the geometric characteristics of the road that can determine riding difficulties for cyclists. The proposed procedure can be used both by the public authorities who manage cyclists' safety and as a method of classifying the road network for cycling.

The proposed procedure, based on the use of pattern recognition techniques, analyses data from a sample of nine riders who travelled on rural roads within the Municipality of Messina (Italy) to classify the roads according to their cycling difficulty. For each rider, duration, distance, road slope, altitude difference and spent calories have been measured and analysed.

The collected data were used for the development of a model capable of predicting the cyclist's physical effort as a function of the road alignment itself. Knowing the effort required to cycle along a route can contribute to a more complete assessment of road classification, commonly defined according to motor vehicles. Moreover, it may constitute a measure determining the safety of cycling by encouraging cyclists to travel along roads compatible with their physical abilities.

1. INTRODUCTION

Cyclists' safety has been discussed in recent years, especially as the use of this mode of transport has been increasing worldwide. While this success has been enthusiastically welcomed by those focusing on positive environmental effects, others indicate a huge increase in accidents, even beyond the official reports [1-3].

The increase in the number of bicycles has rarely determined adjustments of the road infrastructure, which is traditionally designed to satisfy the needs of motorized vehicles only [4, 5]. Even if the urban areas, mainly in Northern Europe, include a dense network of bicycle paths assuring higher safety for cyclists, the rural roads contain all the traffic components within the same cross-section, without any passive or active protection for vulnerable users [6].

In recent years, scientific research has investigated a series of factors that may influence and reduce the risk of accidents. For instance, separation of bike flows from motor vehicles has represented a strategical step towards safer infrastructures, even if critical issues remain in connection areas, where the various flows dangerously converge in extremely complex scenarios [7-10]. From a functional perspective, one of the best solutions is to introduce roundabouts, solving most of the safety problems of the various traffic components. Advanced Driver Assistance Systems (ADAS), i.e., electronic systems that may support the driver in danger or emergencies, may represent another helpful strategy towards reducing the probability of accidents. Preliminary applications, even for bicycles, are already available and the results appear to be positive [11].

When the road infrastructure is used by all the traffic components, the scenario becomes the most dangerous one and it may be further complicated by other factors, such as road pavement quality, road width, parked cars, the practice of cyclists staying in groups, access frequency and considerable speed differences between bicycles and motor vehicles [12].

Unlike the materials that make up road pavements, the human factors influencing the driving actions of cyclists and car drivers are characterized by high uncertainty and dispersion, due to the heterogeneous nature of the users [13-15]. Gildea and Simms [16] assessed the causes of collisions in Ireland. Although the involvement with motor vehicles is predominant (56%), the percentage of accidents with autonomous or other cyclists (29% + 8%, respectively) is high. The analysis was very thorough and focused not only on traffic and road characteristics but also some human factors of cyclists. Among them, it was found that men are more often involved in accidents than women, irrespective of factors such as familiarity with the infrastructure, wearing protective clothing, consuming alcohol or using mobile phones when cycling.

A psycho-physical state may be a further aggravating issue for cyclists' safety. High alcohol or drug levels have disrupting effects on attention capacity and reaction time, besides consequences on the physical capacity required for driving [17].

In terms of the physical state, some research focused on the role of fatigue in cyclists' safety and analysed the variables that mainly influence it, such as the path length or the road elevation trend [18-21]. In the literature, only specific experiments are reported that cannot be generalized to other contexts, probably owing to difficulties related to including the human factors in an analytical procedure [22].

Physical fatigue cannot be quantified by direct methods but can only be inferred from voluntary admission of the cyclist involved, which excludes all fatal accidents from the report.

Despite this, there is no doubt that the energy spent leads to a series of consequences on riding behaviour that can be extremely dangerous, such as deficit in attention [23, 24], slowness in reaction times, greater irregularity in trajectories, lower speed and, consequently, greater speed difference between a bicycle and any motor vehicle moving in the same direction.

In this respect, recent research has highlighted some factors that favour cycling and others that discourage it. Among the former, health improvements, the cost and time savings are noted, while the risk of accidents and adverse weather conditions stand out among the latter. Even if the study does not investigate the role of these variables, there is a strong link between the cyclist's physical abilities and the safety that he or she perceives [25].

Also, Bulsink et al. [26] and Kiewiet et al. [27] have carried out experiments on elderly cyclists, ascertaining that one reason for the disrupted stability involved a more complex and slower recovery manoeuvre than that of the younger cyclists, demonstrating that the contribution of physical resistance is necessary to ensure the safety of the journey. Elderly cyclists present not only physical problems but also psychological stress in particular situations of environmental complexity. In this respect, Vlakveld et al. [28] reported an increase in workload measurements compared to younger users and a simultaneous decrease in speed, a sign that the latter cannot always be taken as a reference for safety. In terms of the factors that influence cyclists' stress, Gadsby et al. [29] published a large study involving trials in the Netherlands and the United States. Among the causes reported by users, those relating to interaction with motor vehicles, pavement conditions and road geometry are the most commonly reported.

Cycling design standards, issued by some European countries or cities, lend very limited importance to the problems generated by vertical alignment of the road. The question is deliberately simplified because it is believed that the majority of cyclists are not interested in travelling on long and sloping roads and, in any case, the terrain of the countries of central Europe such as Holland, Belgium and Denmark, especially in urban sites, is basically flat [30-36].

However, even if it does not represent the majority of cases, it is always necessary to check the cyclists' safety on roads with high slopes as uphill terrain increases the energy spent considerably and reduces comfort and safety. On the contrary, downhill terrain leads to a considerable increase in speed and, therefore, a greater dispersion of the trajectories with negative consequences on safety.

In this paper, the authors attempt to overcome the limitation of the existing literature and propose a general procedure for allowing the analyst to forecast the physical performance required by the cyclist along a specific road. This procedure relies on the evaluation of a very simple indicator, the calories burned, determined considering several relevant indicators representing the main features of the road. This indicator might be used with others, such as construction and functional aspects of the road, to determine cyclists' safety during the ride. The authors want to highlight that in this research, there is a link between physical effort and road alignment that has never been considered in an analytical relationship in the literature summarized before. Furthermore, to generalize the results and make them useful for future applications, this link has been treated with statistical methodologies that will be illustrated in the Methods section.

2. METHODS

2.1. Premise

The authors, in line with what was reported in some studies, believe that cyclists' safety depends on physical and functional aspects of the road but, mainly, on psycho-physiological characteristics of the users. Unlike car drivers, for whom the workload is mainly related to mental effort, cyclists are strongly influenced by physiological fatigue. This variable is highly uncertain, as it is a function of the health status, age, physiology and the impact of the road construction characteristics on his or her performance.

Therefore, there is a need to find an indicator to represent the cyclist's stress, with a satisfactory level of accuracy. Whether the aim is to build a model with general valence and one that is easy to configure, it is important to identify the easily measurable input variable using low-cost devices.

From these statements, two fundamental considerations emerge:

- the input variables should be related to the infrastructure, not the cyclists, so as the model, once calibrated, can process different roads without novel cycling measures and
- the numerous uncertainties should be controlled as much as possible, but it is not possible to eliminate them completely.

First, normalization of the calories burned is required, to balance unavoidable physiological differences among the analysed cyclists. Second, it is convenient to classify this variable into a sufficient number of classes (three in this case), representing the difficulties, more or less evident, of the cyclist. However, all the other variables should be normalized too (in the [0,1] range), to avoid overbalancing in the analytical model towards those variables with the highest absolute values.

2.2. Some details of the experiments

To prepare an appropriate dataset for the forecasting model, the activity of a sample of nine cyclists moving along a rural road has been considered. During cycling experiments, several variables have been measured, such as elapsed time, total distance, grade, elevation and the calories burned (used as an output variable) as a function of the heart rate.

The definition of a specific analytical model allowed the analysts to estimate the effort for a medium-capacity cyclist along a road, knowing only the main geometrical characteristics of the road. The experiments were conducted on two different rural roads in the city of Messina (Italy), in a suburban area, characterized by low volumes of traffic. The main features of the road are as follows: 7 m wide, no shoulders, around 9 km long, average grade equal to 5% and pavement in good condition.

The users' sample is made up of amateur cyclists, with a similar level of fitness to each other, aged between 40 and 50 years, with a good health status (certified by medical documentation), who did not cycle together. The cyclists, unaware of the purpose of this scientific research, used road bikes equipped with computers able to measure some physiological, dynamic and physical variables (heart rate, burned calories, cadence, time, distance, speed, grade, elevation, etc.), without this instrumentation causing an impediment to normal riding. After the experiments, all the measured variables were entered into a notebook and processed using Pattern Recognition techniques in the Matlab® Statistical Toolbox environment. This methodology allowed the analysts to classify the various observations into three proper classes, related to different ranges of burned calories and, thus, the cycling fatigue. The classification of events alone would have represented only a modest scientific result. The advantage of such an analytical model is to provide a similar forecast on other roads not used to build the model too, considering only input data, without other contributions by actual cyclists.

Naturally, the same path may determine a slightly different consumption of calories for each cyclist, depending on his or her weight, training level, heart rate, etc. To take into account this aspect, the various data were normalized according to the following:

$$Cal_n = (Obs_n - Cal_{min})/(Cal_{MAX} - Cal_{min}) \quad (1)$$

where Cal_n are the normalized calories burned for the n-th observation, Obs_n is the absolute calories burned for the n-th observation, Cal_{min} is the calories burned at the beginning of the road and Cal_{MAX} is the calories burned at the end of the road. Cal_{min}, generally, is equal to zero but if the section of the road to be analysed is included within a longer section, it can be different from zero.

In this way, each cyclist has Cal values in the [0,1] range comparable with other users of the sample. The three classes into which the measures can be classified are as follows:

Class 1: $0 \leq Cal < 0.3$.

Class 2: $0.3 \leq Cal < 0.6$.

Class 3: $Cal \geq 0.6$.

The other variables have been normalized using a similar procedure as well.

Obviously, there are other factors that influence the consumption of calories, such as the weight of bicycles and cyclists and the presence of wind, as they determine some important resistance. In this case, the sample is very homogeneous, as the racing bikes used in this research weigh between 8 kg and 8.50 kg, and cyclists weighed between 70 and 75 kg. Furthermore, during the experimentation, there was no wind of at least appreciable intensity.

For different values of these quantities, the cyclist's calorie expenditure in absolute terms will vary proportionately. However, the final result is expressed in caloric expenditure classes, rather than in absolute values and, therefore, any accuracy errors are insignificant because their values are considerably lower than the class interval in the approximation of the result.

2.3. Brief note about the Linear Discriminant Analysis (LDA)

Our goal, therefore, is to verify that a particular observation (as a set of features) is representative of the output, which, in our case, is the calories spent.

Besides, when the input variables assume different values, the model must understand whether and in which way the output will change. Pattern recognition and, in particular, discriminating models, only work well if the final result is a class rather than a purely numerical value [37].

The separation of the detected observations (or objects) can be achieved with linear separation surfaces or, more precisely, with straight lines in a two-dimensional space, planes in a 3D space or, moreover, hyperplanes in an nD space (where n is the number of input variables).

In general, it is possible to state that all the surveyed observations, reported in a data set X of [m × n] dimension (where m is the number of features and n is the number of objects), will be assigned to the C classes that, in this case, represent different functional classes.

Therefore, N_i objects belonging to the class ω_i are obtained. Each object x is described by the corresponding values taken by the m variables: $x = (x_1, x_2, \ldots, x_m)$.

Then, the purpose is to project the objects belonging in X on a C-1 dimension hyperplane called Y, where it is more convenient to perceive the separation between the classes.

For example, in the simplest case where C=2, the m-dimension data set will have a number of N_1 samples belonging to ω_1 and N_2 belonging to ω_2. Thus, the goal is to obtain any scaling for projecting the x observations on an opportune straight line:

$$y = w^T x \quad (2)$$

where

$$x = \begin{bmatrix} x_1 \\ \cdots \\ x_m \end{bmatrix} \quad (3)$$

and

$$w = \begin{bmatrix} w_1 \\ \cdots \\ w_m \end{bmatrix} \quad (4)$$

The following procedure will be applied to obtain the axis that guarantees the best separability between the classes.
First, it is necessary to identify some indices measuring the class separation, as the mean vectors in x and y (Eq. (5) and Eq. (6)), as follows:

$$\mu_i = \frac{1}{N_i}\sum_{x\in\omega_i} x \quad (5)$$

$$\widetilde{M}_i = \frac{1}{N_i}\sum_{y\in\omega_i} y = \frac{1}{N_i}\sum_{x\in\omega_i} w^T x = w^T \frac{1}{N_i}\sum_{x\in\omega_i} x = w^T \mu_i \quad (6)$$

The distance between projected averages (similar to the distance between centroids), calculated through Eq. (6), represents an acceptable criterion for the final decision:

$$J(w) = |\tilde{\mu}_1 - \tilde{\mu}_2| = |w^T\mu_1 - w^T\mu_2| = |w^T(\mu_1 - \mu_2)| \quad (7)$$

However, in this way, there will be no information about dispersion within the classes. For this reason, Fisher [38] introduced into the above-mentioned objective function J(w) normalization with respect to a measured representative of this dispersion, called scatter. $\tilde{S}_i^2$ represents, in this way, the variability within the class ω_i after projecting it along the new y axis, as in Eq. (8):

$$\tilde{S}_i^2 = \sum_{y\in\omega_i}(y - \tilde{\mu}_i)^2 \quad (8)$$

The sum $\tilde{S}_1^2 + \tilde{S}_2^2$corresponds to the variability in the two classes after the projection on the new y axis and the analyst's aim is to find the $w^T x$ linear function that maximizes the function J(w), reported in Eq. (9):

$$J(w) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{S}_1^2 + \tilde{S}_2^2} \quad (9)$$

In conclusion, in the ideal representation on the new axis, the observations belonging to the same class are very close to each other and, at the same time, the averages between the different classes are as far as possible. It is necessary to express J(w) as an explicit function of w to find the function maximum. Then, it is possible to define a scatter in the multivariate space x, like Eq. (10) and Eq. (11):

$$S_i = \sum_{x\in\omega_i}(x - \mu_i)(x - \mu_i)^T \quad (10)$$

$$S_W = S_1 + S_2 \quad (11)$$

where S_i is the covariance matrix of the ω_i class and S_w is the within-class scatter matrix.

The scatter of the projection on y, expressed as a function of the scatter matrix in the x space, is expressed through Eq. (12) and Eq. (13):

$$\tilde{S}_i^2 = \sum_{y\in\omega_i}(y - \tilde{\mu}_i)^2 = \sum_{x\in\omega_i}(w^T x - w^T\mu_i)^2 =$$
$$\sum_{x\in\omega_i} w^T(x - \mu_i)(x - \mu_i)^T w = w^T S_i w \quad (12)$$

$$\tilde{S}_1^2 + \tilde{S}_2^2 = w^T S_1 w + w^T S_2 w = w^T(S_1 + S_2)w = w^T S_w w = \tilde{S}_w \quad (13)$$

where $\tilde{S}_w$ is the scatter matrix for the class projected on the y axis.

Similarly, it is possible to derive the differences between averages projected on the y axis in terms of averages in the original x space, as reported in Eq. (14):

$$(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = (w^T\mu_1 - w^T\mu_2)^2 = w^T(\mu_1 - \mu_2)(\mu_1 - \mu_2)^T w = w^T S_B w = \tilde{S}_B \quad (14)$$

The S_B matrix represents the scatter between the class of the original observations, while $\tilde{S}_B$ is that reported on the y axis.

Therefore, Eq. (8) becomes Eq. (15):

$$J(w) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2} = \frac{w^T S_B w}{w^T S_W w} \tag{15}$$

where J(w) is a measure of the difference between the considered classes means, normalized by the value of the within-class scatter matrix.

The derivative equal to zero, as is known, yields the maximum of the function. The final expression (bypassing the other steps) is represented by Eq. (16):

$$S_W^{-1} S_B w - J(w) w = 0 \tag{16}$$

By solving the eigenvalue problem, Eq. (17) is derived:

$$S_W^{-1} S_B w = \lambda w \tag{17}$$

where λ= J(w) is a scalar.

The so-called linear discriminant yields the optimal solution, reported in Eq. (18):

$$w^* = \underset{w}{arg\,max}\, J(w) = \underset{w}{arg\,max} \left(\frac{w^T S_B w}{w^T S_W w}\right) = S_W^{-1}(\mu_1 - \mu_2) \tag{18}$$

If the classes are more than 2 (for example C), there will be C-1 projection vectors w_i (instead of only y), but the procedure will be the same.

In the processing phase, this general methodology should be calibrated to minimize the percentage of misclassified elements, by changing the number and quality of the input variables, number of output classes and the discriminant technique. This iterative procedure is needed because the choice of the optimum configuration is not always obvious, as it may also depend on the size and typology of the acquired data.

3. RESULTS

The experiments consisted of cycling on different days on the two roads shown in Figure 1a and 1b, respectively, 9 and 10 km long and with an average grade of 5%. The first road (named "Gesso") was used to train the model and the second road (named "Salice") was used to test it. In both cases, before cycling on the reference path, the cyclists cycle for 20 km at moderate speed for warm-up.

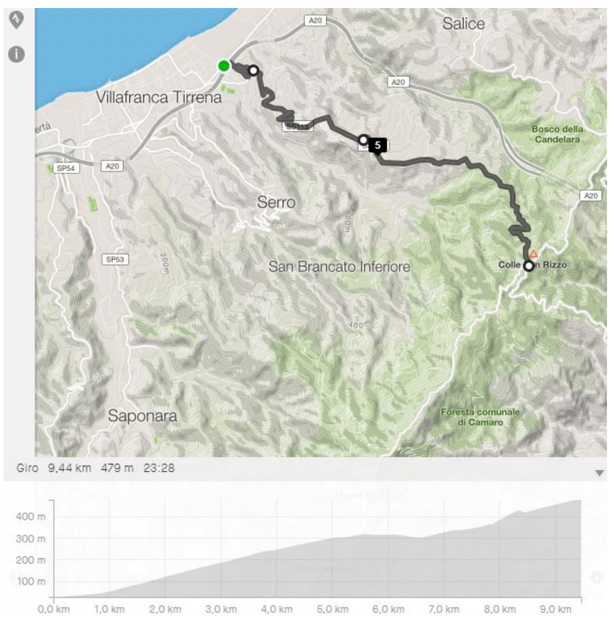

Fig. 1a. The two roads selected for the experiments: a) "Gesso" road, from "Villafranca Tirrena" to "Colle San Rizzo"

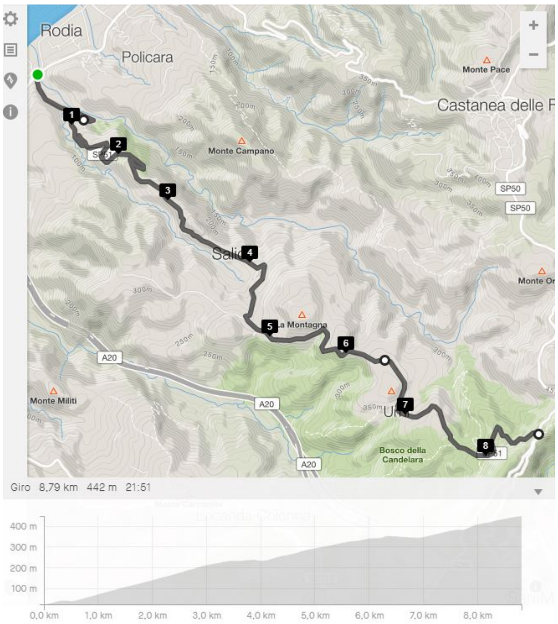

Fig. 1b. The two roads selected for the experiments: b) “Salice” road, from “Rodia” to “Portella”

Table 1 lists part of the variables recorded by the bike computer during cycling on the first road (Fig. 1a). According to the sampling frequency of the internal GPS, a measure every 5 seconds has been recorded. About 3300 observations have been recorded, for an average cycling time equal to 16,500 seconds that, divided for the involved cyclists (9), provides an average cycling time of 1,833 seconds per user (around 30 minutes).

The LDA model (defined as specified in the Methods section) reached a global accuracy level of 90.5% in identifying the class of calories burned for each observation. The 1st class includes observations with very low values of burned calories (<0.3), the 2nd class refers to the range [0.3-0.6], while the 3rd class includes observations for which the caloric consumption has been relevant (more than 0.6). All these values are dimensionless as ratios of quantities with the same unit of measure (calories).

Table 1

Short extract from the database powered by the experimentation for the 1st road

Time(min)	Distance (m)	Grade (%)	Altitude (m)	Calories (Cal)
…	…	…	…	…
351,93	29491,80	7,61	505,23	180,93
352,00	29502,50	7,56	506,05	181,38
352,07	29512,50	7,15	506,82	181,82
352,13	29522,90	7,50	507,63	182,27
352,20	29533,90	7,01	508,40	182,72
352,27	29546,50	6,72	509,14	183,17
352,33	29559,60	5,98	509,80	183,61
352,40	29571,00	5,58	510,42	184,03
352,47	29581,40	5,75	511,05	184,45
352,53	29592,90	4,92	511,57	184,85
352,60	29604,20	4,34	512,00	185,24
352,67	29614,40	4,83	512,45	185,64
…	…	…	…	…

Even if the problem is multidimensional, for greater clarity it was decided to represent the graphs in 2D by comparing only two variables at a time. Figures 2a, 3a and 4a show some representations of the results for the 1st road. After this training phase, the test phase was performed, on the 2nd road, considering only 2 cyclists, for evaluating the forecasting capacity of the model. In Table 2, an extract of the novel dataset is provided: in this case, the burned calories are not reported, as the related class must be estimated by the model. Naturally, the values were properly measured during cycling and used as references for testing the model.

Table 2

Short extract from the database powered by the experimentation regarding the 2nd road

Time(min)	Distance (m)	Grade (%)	Altitude (m)
...	...	...	...
820,0	54139	6	372
820,6	54150	6	373
820,0	54160	7	373
820,0	54170	6	374
820,8	54180	6	375
820,0	54191	6	375
820,0	54202	7	376
821,0	54213	7	377
821,0	54220	6	377
821,0	54226	7	378
821,2	54234	7	379
821,0	54244	7	380
821,0	54255	7	380,8
...	...	...	...

Figures 2b, 3b and 4b are related to the classification of the novel observations for the 2nd road, based on the forecasting capacity of the model. As stated, direct data measured by the bike computers have been used to test the efficiency of the LDA model: the test guaranteed an accuracy value equal to 89,9% (number of properly classified observations in the total observations).

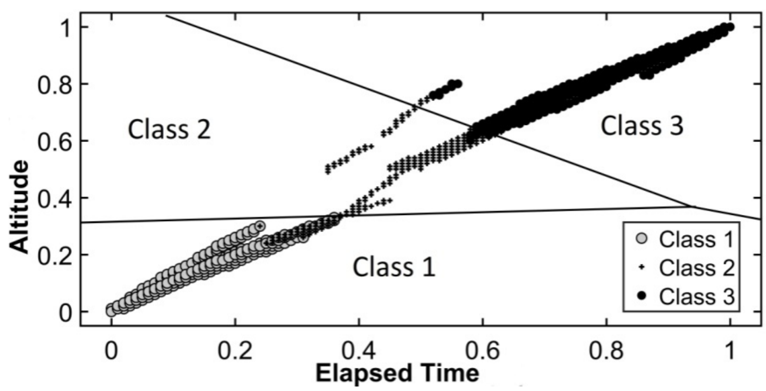

Fig. 2a. Distribution of the observations in a 2-D space (elapsed time vs. altitude): trials (1st road)

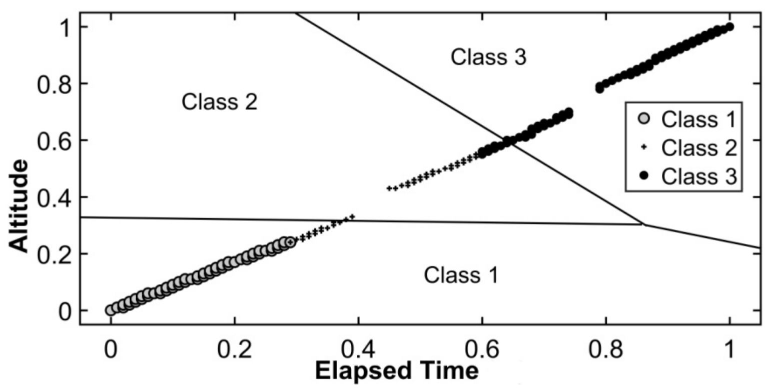

Fig. 2b. Distribution of the observations in a 2-D space (elapsed time vs. altitude): new observations (2nd road)

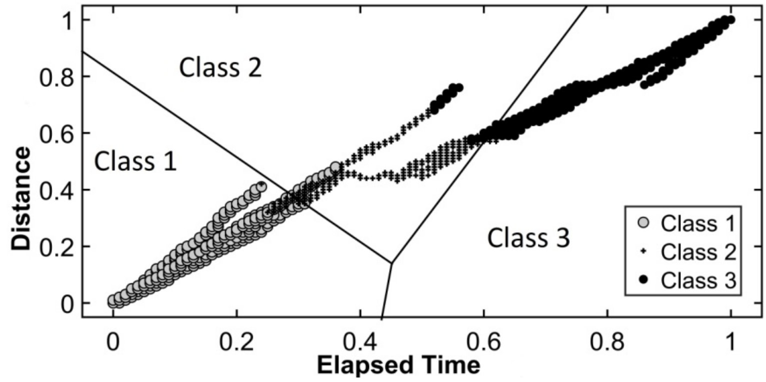

Fig. 3a. Distribution of the observations in a 2-D space (elapsed time vs. distance): trials (1st road)

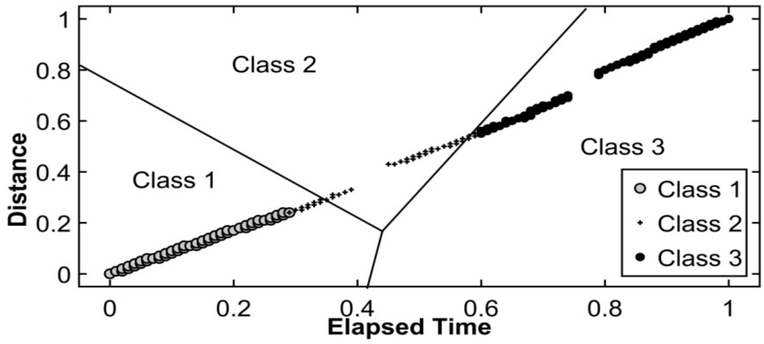

Fig. 3b. Distribution of the observations in a 2-D space (elapsed time vs. distance): new observations (2nd road)

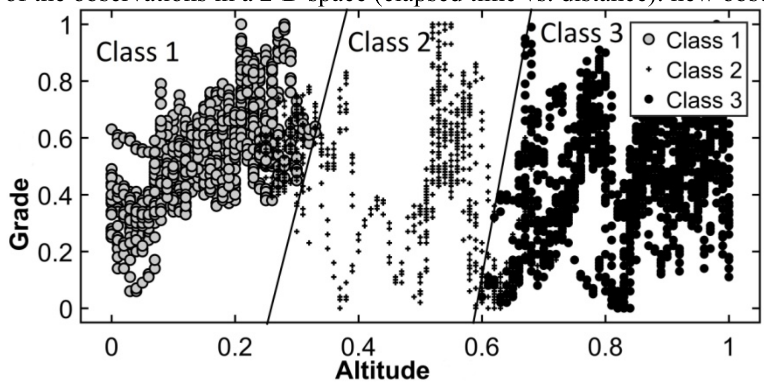

Fig. 4a. Distribution of the observations in a 2-D space (altitude vs. grade): trials (1st road)

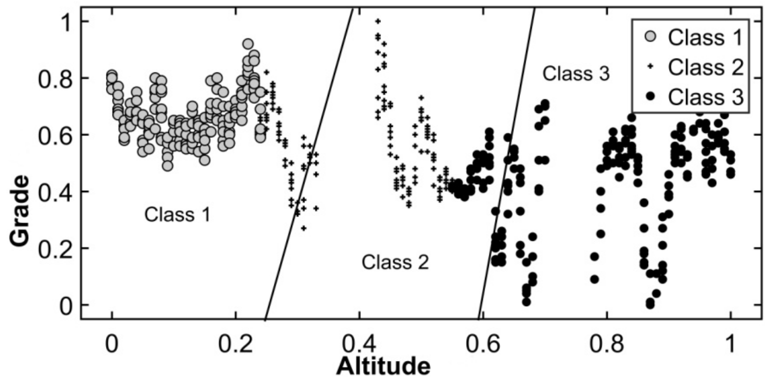

Fig. 4b. Distribution of the observations in a 2-D space (altitude vs. grade): new observations (2nd road)

4. DISCUSSION

All the trials were performed on a rural road. In the introduction, it has been underlined that urban scenarios face problems related to considerable encounters between cyclists and drivers of motorized vehicles, especially in the intersections. The rural roads in suburban areas, on the contrary, are generally characterized by proper length and grade and, thus, a different element for a cyclist's safety is considered, i.e., the state of fatigue.

The selected roads (Fig. 1a and 1b) have common features with roads in Eastern Sicily, as they connect coastal towns to mountainous ones. Then, it is possible to consider the validity of experimentation as generally reliable for this part of Sicily. Further, 9 km length and 5% average grade are typical of paths that can be covered by amateur cyclists and, thus, represent the ideal scenario for the aim of our research.

It should be noticed that data in Table 1 and Table 2 are not normalized to show the real value assumed by the variables. The sampling frequency (one measure every 5 seconds) represents an acceptable balance between the need to limit the database size and to obtain, in any case, a relevant number of observations representative of the observed phenomenon.

Although it was not reported in the Results section, a preliminary model included cadence and speed as input variables. Using 6 input variables, the model appeared slightly more accurate, but it was not considered because a higher number of variables is always more onerous (in data collection and computation). When this increase determines better performance, the related "costs" may be acceptable, but, in this case, the gain in accuracy was negligible (+1.3% in absolute value). Moreover, additional variables (cadence and speed) depend on the cyclist's activity, while the authors aim to apply the model to other roads, avoiding novel tests for cyclists, but relying only on infrastructure variables.

The final accuracy level, further, is good also because the linear PR model used (LDA) is reliable but does not represent the most refined solution from an analytical perspective. However, this choice is preferable not only for its simplicity and quickness but also for normal-size datasets and the advantages for use by non-expert analysts.

A specific test of the model was performed on a novel dataset, related to a different road (Fig. 1b). This test consisted of forecasting the calories burned for the novel observation on the 2nd road, considering the measured input variables. The proposed model classified the observations into three classes, and the results (correct classification for 89.9% of records) confirm the reliability of the analytical model, despite the different causes of uncertainty included in the problem.

The accuracy of the results, i.e., the capacity of the model to classify the various observations, may be verified in Figures 2a, 3a and 4a, in which only a few elements are misclassified. The meaning of these figures is relative, as the numerous observations in the correct class are overlapped, while those erroneously classified are, paradoxically, more separated and, thus, more evident. However, it is interesting to notice the relationship between the two variables provided in the charts. Figure 2a provides the classification in the 2D view Altitude - Elapsed Time. The relationship is substantially linear as the vertical geometry of the road is characterized by a constant grade along the entire path. This trend is very similar in Figure 2b too, for the 2nd road, which is, however, characterized by the same features and elevation trend. In this case, the observation distribution is less spread because only 2 cyclists have been included for testing, but also in the 1st case, the trend is defined enough.

The same considerations may be extended to Figure 3, in which observations are presented in the chart Distance - Elapsed Time. The trend is, naturally, linearly increasing, but in Figure 3a, there is a slight dispersion from the linear trend owing to the different performance of the nine users; this is less evident for the 2nd road, along which the cyclists showed similar performance. Figure 4 provides classification results in the chart Grade – Altitude and it is more interesting as this representation is not immediately intuitive. In truth, no relationship is expected between these variables, but the LDA classifier perfectly determined the boundaries of the areas related to the three classes. It is important to underline that also the input variables were normalized, unlike what is reported in Tables 1 and 2 for the reasons already mentioned, considering that grade, min and max values are very close (between 4.5% and 5.5%) and there is no remarkable standard deviation. However, when normalization led to the spread of the values in the total range scale, there was a gradation of this percentage point on a scale ranging from 0 to 1. This is the reason for the specific dispersion of the values in vertical axes, apparently without any relation with the x-axes. In truth, the y values have a very low oscillation around the average. The same considerations may be extended to the testing phase on the 2^{nd} road.

In conclusion, this methodology represents an analytical instrument to identify a certain level of physical effort required by cyclists for cycling along a specific road. This, in turn, is only one of the variables influencing the cyclists' safety, with those related to infrastructure and traffic. Since these variables will also be affected by some uncertainty, it is unnecessary and not convenient to pursue higher precision. Second, this procedure can be useful to classify the road for cycling use. The roads, in traditional standards, are functionally classified as a function, almost exclusively, of the needs and requirements of motorized vehicles. If their use is mixed, i.e., includes vulnerable users too, a classification methodology taking into account this means of transport is also reasonable.

These results represent real progress compared to the scientific findings of recent years, since the performance of cyclists had never been analytically linked to the geometric characteristics of the road.

However, this research does not aim to trivialize the problem of cyclists' safety. Many authors [2, 3, 7, 9], rightly, identify other situations that can cause great dangers, such as interaction with motor vehicles, the complexity of urban environments and traffic volumes.

Unlike the above, authors have evaluated a human factor variable that, as such, suffers from great uncertainty and difficulty in generalization to other contexts. These problems have been solved with appropriate statistical techniques and with the acceptance of a small level of approximation.

Road administrators, with the proposed methodology, can rely on fully objective and evolved decision support. As already reported, the variables related to the vertical alignment of roads are generally neglected by European regulations, since there are only a few prescriptions to be applied when the slopes exceed a certain threshold, without further comments on or considerations of the consequences for the cyclist in terms of safety and comfort.

This research at present, however, cannot be considered complete as the cyclists' safety must be further investigated, at least in the directions suggested as follows:

- The proposed procedure is useful on road sections where physical fatigue is greater, i.e., uphill roads. However, in terms of road safety, downhill sections are just as dangerous. In this case, it is not fatigue that is the prevailing variable but speed that should be studied together with the trajectories.
- The interaction with other traffic components is also interesting, especially with motor vehicles, given that a large proportion of accidents are due to this cause.

Therefore, the next steps of the research will aim to investigate the aforementioned conditions.

5. CONCLUSIONS

The safety of a cycling road depends on numerous variables, such as the pavement condition, horizontal and elevation geometry, transversal section, higher speeds of motorized vehicles and different sight-distances. These variables are easily identifiable and have the advantage of remaining almost constant for long periods.

However, despite the higher complexity, the performance of the cyclist, especially in terms of physical effort and fatigue, should be taken into consideration. This information influences the possibility to cycle on specific paths and, indirectly, users' safety. The issue, in general, is not practical to be solved to account for the huge uncertainty related to the human component that, unlike for car drivers, involves physical fatigue, in turn influenced by the user's health status, age, diseases, etc.

In this study, the physical effort and fatigue required along a certain road have been determined using a specific model, based on Pattern Recognition techniques, able to identify an appropriate class of effort according to some variables related to the features of roads: distance, elapsed time, grade and elevation. A sample of nine cyclists has been monitored, measuring also the burned calories (as a function of the heart rate) during cycling on the selected road. The aim was to assess if any dependence between the burned calories and the road characteristics exists. The excellent results, in terms of classification accuracy (around 90%), evidence the possibility to estimate the physical effort of the cyclist, by simply considering the main geometrical characteristics of the road.

In any case, this procedure may allow the infrastructure manager to identify a suitable class that contains information on the difficulty of travelling. If the level of achieved safety or expected class is not satisfactory, it is possible to implement a series of mitigation measures, such as roadside renovation, a parking ban, specific information and signage indicating the presence of vulnerable users, widening of the cross-section, separation of traffic currents, appropriate signs and so on.

Acknowledgment

The article has been presented at the 13th International BRD GAMBIT 2020 Conference.
This research was co-funded by the Science Excellence programme of the Ministry of Science and Higher Education.

References

1. Juhra, C. & Wieskotter, B. & Chu, K. & Trost, L. & Weiss, U. & Messerschmidt, M. & Malczyk, A. & Heckwolf, M. & Raschke, M. Bicycle accidents – Do we only see the tip of the iceberg? A prospective multi-centre study in a large German city combining medical and police data. *Injury, International Journal of Care Injured*. 2012. Vol. 43. P. 2026-2034.
2. Billot-Grasset, A. & Amoros, E. & Hours, M. How cyclist behavior affects bicycle accident configurations? *Transportation Research Part F: Traffic Psychology and Behaviour*. 2016. Vol. 41. P. 261-276.
3. Schepers, P. & Stipdonk, H. & Methorst, R. & Olivier, J. Bicycle fatalities: Trends in crashes with and without motor vehicles in The Netherlands. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2017. Vol. 46. P. 491-499.
4. D'Andrea, A. & Cappadona, C. & La Rosa, G. & Pellegrino O. A functional road classification with data mining techniques. *Transport*. 2014. Vol. 29. No. 4. P. 419-430.
5. Bosurgi, G. & Pellegrino, O. & Sollazzo, G. Road functional classification using pattern recognition techniques. *Baltic Journal of Road and Bridge Engineering*. 2019. Vol. 14. No. 3. P. 360-383.
6. Mehta, K. & Mehran, B. & Hellinga, B. A methodology to estimate the number of unsafe vehicle-cyclist passing events on urban arterials. *Accident Analysis and Prevention*. 2019. Vol. 124. P. 92-103.
7. Zangenehpour, S. & Strauss, J. & Miranda-Moreno, L.F. & Saunier, N. Are signalized intersections with cycle tracks safer? A case–control study based on automated surrogate safety analysis using video data. *Accident Analysis and Prevention*. 2016. Vol. 86. P. 161-172.
8. Laureshyn, A. & de Goede, M. & Saunier, N. & Fyhri, A. Cross-comparison of three surrogate safety methods to diagnose cyclist safety problems at intersections in Norway. *Accident Analysis and Prevention*. 2017. Vol. 105. P. 11-20.
9. Ng, A. & Debnath, A.K. & Heesch, K.C. Cyclist' safety perceptions of cycling infrastructure at un-signalised intersections: Cross-sectional survey of Queensland cyclists. *Journal of Transport & Health*. 2017. Vol. 6. P. 13-22.
10. Kovácsová, N. & de Winter, J.C.F. & Hagenzieker, M.P. What will the car driver do? A video-based questionnaire study on cyclists' anticipation during safety-critical situations. *Journal of Safety Research*. 2019. Vol. 69. P. 11-21.
11. Silla, A. & Leden, L. & Rämä, P. & Scholliers, J. & Van Noort, M. & Bell, D. Can cyclist safety be improved with intelligent transport systems? *Accident Analysis and Prevention*. 2017. Vol. 105. P. 134-145.
12. Beck, B. & Chong, D. & Olivier, J. & Perkins, M. & Tsay, A. & Rushford, A. & Li, L. & Cameron, P. & Fry, R. & Johnson, M. How much space do drivers provide when passing cyclists? Understanding the impact of motor vehicle and infrastructure characteristics on passing distance. *Accident Analysis and Prevention*. 2019. Vol. 128. P. 253-260.

13. Westerhuis, F. & de Waard, D. Reading cyclist intentions: Can a lead cyclist's behaviour be predicted? *Accident Analysis and Prevention*. 2017. Vol. 105. P. 146-155.
14. Zeuwts, L.H.R.H. & Vansteenkiste, P. & Deconinck, F.J.A. & Cardon, G. & Lenoir, M. Hazard perception in young cyclists and adult cyclists. *Accident Analysis and Prevention*. 2017. Vol. 105. P. 64-71.
15. Haworth, N. & Heesch, K.C. & Schramm, A. Drivers who don't comply with a minimum passing distance rule when passing bicycle riders. *Journal of Safety Research*. 2018. Vol. 67. P. 183-188.
16. Gildea, K. & Simms, C. Characteristics of cyclist collisions in Ireland: Analysis of a self-reported survey. *Accident Analysis and Prevention*. 2021. Vol. 151. No. 105948.
17. de Waard, D. & Houwing, S. & Lewis-Evans, B. & Twisk, D. & Brookhuis, K.A. Bicycling under the influence of alcohol. *Transportation Research Part F*. 2016. Vol. 41. P. 302-308.
18. Wei, F. & Lovegrove, G. An empirical tool to evaluate the safety of cyclists: Community based, macro-level collision prediction models using negative binomial regression. *Accident Analysis and Prevention*. 2013. Vol. 61. P. 129-137.
19. Dozza, M. & Werneke, J. Introducing naturalistic cycling data: What factors influence bicyclists' safety in the real world? *Transportation Research Part F*. 2014. Vol. 24. P. 83-91.
20. Kamel, M.B. & Sayed, T. & Bigazzi, A. A composite zonal index for biking attractiveness and safety. *Accident Analysis and Prevention*. 2020. Vol. 137. No. 105439.
21. Bongiorno, N. & Bosurgi, G. & Pellegrino, O. & Sollazzo, G. A Methodology to Identify Critical Road Section by means of Cyclists' Fatigue. *Advances in Trasportation Studies: an international Journal*. 2020. Section B50. P. 95-110.
22. Bigazzi, A.Y. & Gehrke, S.R. Joint consideration of energy expenditure, air quality, and safety by cyclists. *Transportation Research Part F*. 2018. Vol. 58. P. 652-664.
23. Boele-Vos, M.J. & Van Duijvenvoorde, K. & Doumen, M.J.A. & Duivenvoorden, C.W.A.E. & Louwerse, W.J.R. & Davidse R.J. Crashes involving cyclists aged 50 and over in the Netherlands: Anin-depth study. *Accident Analysis and Prevention*. 2017. Vol. 105. P. 4-10.
24. Boele-Vos, M.J. & Commandeur, J.J.F. & Twisk D.A.M. Effect of physical effort on mental workload of cyclists in real traffic in relation to age and use of pedelecs. *Accident Analysis and Prevention*. 2017. Vol. 105. P. 84-94.
25. Useche, S.A. & Alonso, F. & Montoro, L. & Sanmartin, J. Healthy but risky: A descriptive study on cyclists' encouraging and discouraging factors for using bicycles, habits and safety outcomes. *Transportation Research Part F: Psychology and Behaviour*. 2019. Vol. 62. P. 587-598.
26. Bulsink, V.E. & Kiewiet, H. & van de Belt, D. & Bonnema, G.M. & Koopman B. Cycling strategies of young and older cyclists. *Human Movement Science*. 2016. Vol. 46. P. 184-195.
27. Kiewiet, H. & Bulsink, V.E. & Beugels, F. & Koopman H.F.J.M. The co-contraction index of the upper limb for young and old adult cyclists. *Accident Analysis and Prevention*. 2017. Vol. 105. P. 95-101.
28. Vlakveld, W.P. & Twisk, D. & Christoph, M. & Boele, M. & Sikkema, R. & Remy, R. & Schwa, A.L. Speed choice and mental workload of elderly cyclists on e-bikes in simple and complex traffic situations: A field experiment. *Accident Analysis and Prevention*. 2015. Vol. 74. P. 97-106.
29. Gadsby, A. & Hagenzieker, M. & Watkins, K. An international comparison of the self-reported causes of cyclist stress using quasi-naturalistic cycling. *Journal of Transport Geography*. 2021. Vol. 91. No. 102932.
30. Design Manual for Bicycle Traffic, CROW, Netherlands. Available at: https://www.crow.nl/publicaties/design-manual-for-bicycle-traffic-(1).
31. Vademecum Fietsvoorzieningen, Flanders (Belgium). Available at: http://www.mobielvlaanderen.be/vademecums/vademecumfiets01.php. [In Flemish: Vademecum Cycling Facilities]
32. Qualitätsstandards für Radschnellverbindungen, Baden-Württemberg. Available at: https://www.fahrradland-bw.de/radverkehr-in-bw/infrastruktur/radschnellverbindungen/. [In German: Quality standards for rapid cycle connections]
33. London Cycling Design Standards, Greater London. Available at: https://tfl.gov.uk/corporate/publications-and-reports/streets-toolkit#on-this-page-2.

34. *Håndbog supercykelstier anlæg og planlægning,* Denmark. Available at: http://vejdirektoratet.dk/DA/vejsektor/vejregler-og-tilladelser/vejregler/h%C3%B8ringer/Documents/H%C3%B8ringer%202016/16-01540-3%20H%C3%A5ndbog_Supercykelstier_h%C3%B8ring%203507636_1_1.pdf. [In Danish: Handbook of superbike trails construction and planning]
35. *Réseau cyclable à haut niveau de service. Objectifs et principes d'aménagement*, Cerema France. Available at: http://voiriepourtous.cerema.fr/fiches-produites-par-le-cerema-a1505.html. [In French: *High level of service cycling network. Planning objectives and principles*]
36. Manual for the design of cyclepaths in Catalonia, Spain. Available at: http://territori.gencat.cat/web/.content/home/01_departament/documentacio/territori_mobilitat/bicicleta/manual_per_al_disseny_de_vies_ciclistes_a_catalunya/pdf/vies_ciclistes_ang_tcm32-45418.pdf.
37. Recker, W.W. & McNally, M.G. Travel/activity analysis: Pattern recognition, classification and interpretation. *Transportation Research Part A: General*. Vol. 19/4. P. 279-296.
38. Fisher, R.A. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*. 1936. Vol 7. No. 2. P. 179-188.

Received 14.11.2019; accepted in revised form 16.02.2021