Drift Detection for Identifying Training Patterns Prior to Performance Improvement in Runners

Nathália Tito[1], Lucas G. Tavares[1], Glauco F. Amorim[1]

[1]Centro Federal de Educação Tecnológica Celso Suckow da Fonseca - CEFET/RJ

{nathalia.tito, lucas.tavares}@aluno.cefet-rj.br

glauco.amorim@cefet-rj.br

Abstract. *The growing popularity of running in Brazil has led to an increase in amateur runner participation, creating a demand for personalized recommendations to support these athletes. This study aimed to identify runners who showed performance improvements by detecting concept drift in their training performance time series. Training metrics were computed and compared between periods preceding performance improvements and declines. It was observed that, in cases of improvement, these metrics were higher compared to the periods preceding performance declines. These differences were statistically significant (p-value < 0.05), indicating which training patterns are associated with performance progression and representing a first step toward deeper insights.*

1. Introduction

Brazil follows the global trend of the increasing popularity of running as a sport. Between 2007 and 2015, there was a growth of over 90% in the number of road races held in Rio de Janeiro and São Paulo [Thuany et al., 2021]. As the number of races grows, there has been a shift in participation patterns, marked by the rise of recreational runners. This transformation in runner profiles also influences overall performance in competitions [Reusser et al., 2021]. Moreover, with the proliferation of smart devices such as smartwatches, it has become possible to understand and analyze the performance of runners with diverse characteristics [Berndsen et al., 2019]. The resulting personal data can provide valuable insights for longitudinal health studies, enabling the development of tools to uncover and understand internal patterns [Smyth, 2019].

Understanding running performance is a complex task, especially considering the wide variety of athlete profiles and the different pacing strategies adopted [Casado et al., 2021]. Nakayama et al. [2010] demonstrated, through statistical time series analysis, that stride interval variability in trained runners is significantly lower compared to non-runners. Complementing this perspective, a study by Besomi et al. [2018] identified distinct profiles of urban runners using cluster analysis, considering variables such as sociodemographic data, health conditions, motivations, training characteristics, and running-related behaviors.

By identifying how a runner's performance develops, it becomes possible to make more precise decisions for strategic adjustments. According to Komitova et al. [2022], time series analysis has strong potential in understanding performance-related phenomena, offering valuable methodologies to support evaluations and contributing to the development of new strategies in sports science. Furthermore, with the growing availability of temporal data, there is a promising opportunity for applying machine learning to sports analysis. Adaptive and personalized recommendations based on each athlete's profile during a race, for example, could lead the runner to achieve better results and support them in reaching their goals [Berndsen et al., 2019]. The study by Smyth [2019] highlights the importance of physical activity and the development of technologies that can motivate people toward healthier choices, by recognizing different routines and habits and using persuasive recommendations to transform harmful behaviors into healthier habits through intelligent management of training load and recovery time, for instance.

In this context, the present study aimed to explore two main objectives: (a) to analyze the historical performance time series of different runners, identifying those who showed improvement or decline over time through concept drift detection; and (b) to investigate which training characteristics differ between periods of performance progression and regression. The results revealed significant differences between the groups analyzed. As an initial step toward a broader understanding, these preliminary findings are expected to guide runners aiming to enhance their performance and contribute to more effective training strategies in the future.

2. Theoretical Framework

In streaming environments, data streams consist of continuous sequences of information that may exhibit shifts in their underlying distributions over time — a phenomenon known as concept drift [Agrahari and Singh, 2024]. Such drift challenges traditional modeling approaches, which often assume stationarity and rely on batch training, leading to degraded predictive performance when the data evolve [Webb et al., 2016]. In the context of time series, where data are sequentially indexed over time, concept drift can be understood as a type of event — akin to anomalies, motifs, or more specifically, change points — which represent significant changes in the statistical properties of the series [Ogasawara et al., 2025]. Event detection in time series encompasses retrospective analysis, real-time monitoring, and forecasting, all of which are impacted by the presence of concept drift and its variants.

There exists a wide range of concept drift detectors, each with distinct approaches and sensitivities. Among the most explored in the literature are: DDM (Drift Detection Method), which monitors the error rate and its standard deviation to identify abrupt or gradual changes; HDDM (Hoeffding Drift Detection Method), which enhances this approach by using Hoeffding's inequality to compare error means across batches; ADWIN (Adaptive Windowing), which dynamically adjusts the size of a sliding window by splitting it into sub-windows to detect changes; ECDD (EWMA for Concept Drift Detection), which employs exponentially weighted moving averages to provide smoother responses and sensitivity to continuous variations; and finally, CUSUM (Cumulative Sum Control Chart) and Page-Hinkley, which use a sequential approach that accumulates residuals or differences from the historical mean, making them sensitive to persistent or gradual drifts [Ogasawara et al., 2025]. The selection of the most suitable concept drift detector is directly related to the characteristics of the data stream and the analytical objective.

3. Related Work

The use of time series based methods has gained increasing relevance in the sports domain, especially given the growing availability of collected data. The study by Komitova et al. [2022] highlights the importance of time series analysis in the sports context, emphasizing the application of advanced temporal data mining techniques such as anomaly detection, motif identification, and trend analysis for extracting meaningful patterns. These approaches not only enhance the understanding and automatic recognition of activities but also contribute to detailed analysis and prediction of athletic performance over time.

With the growing popularity of wearable devices, athletes have begun to generate data continuously during physical activities, enabling performance analysis through time series. Stival et al. [2023] proposed a method to detect changes in runners' time series data generated by such devices, including heart rate and speed, aiming to monitor physical exertion or potential health issues. In a related context, Yung et al. [2024] evaluated event detection in historical time series during an athlete's rehabilitation process, identifying changes in performance metrics that signaled progress and supported physicians in making decisions about the athlete's return to full training. In Chang et al. [2023], wearable devices were used to collect time series data of speed, acceleration, and joint angle during running, applying deep learning techniques to identify stages of fatigue.

Regarding performance, numerous studies have investigated various factors associated with improvement or decline. Physical attributes, training frequency, and running experience have all been significantly associated with performance [Suwankan et al., 2024]. A previous study found a high correlation between marathon finish times and specific training indicators, such as volume and intensity [Tanda, 2011]. In Knechtle et al. [2011], in addition to anthropometric variables, training intensity variables were also associated with half-marathon times and demonstrated significant relevance.

Despite advances in time series analysis techniques in sports, the use of concept drift based approaches to investigate improvements or declines in performance remains underexplored in the literature, especially in running. Furthermore, studies that integrate runners training history with their individual characteristics and translate these into actionable recommendations are exceedingly rare. A previous preliminary study indicated that more experienced runners tend to exhibit fewer change points in their individual training series [Tito et al., 2024]; however, this work did not consider speed levels and did not further investigate long-term performance progression. The results of this study are also expected to serve as a supporting approach for classifying runners who have shown improvement in their performance, enabling the identification of factors associated with such progress. Analyzing these factors can help improve training routines and provide personalized recommendations for athletes aiming to enhance their performance.

4. Method

This section provides an overview of the main steps involved in the development of the analysis. It describes the process of data acquisition and preprocessing, as well as the methodologies employed for concept drift detection. Additionally, it details the approaches used to interpret the processed data, with the aim of extracting the most relevant analyses and results.

4.1. Data Collection

A survey was conducted with runners of varying levels of experience, consisting of questions related to anthropometric data, occupational background, sports history, and training habits. In addition to the questionnaire, participants also shared their historical activity data extracted from the Strava app[1], which records a wide range of time series data during runs. A total of 22 runners participated in the study, 64% of whom were male. No age or prior running experience criteria were established for participation. Of the 22 respondents, 19 provided their training histories.

4.1.1. Ethical Approval

This study was approved by the Ethics Committee of the Federal Institute of Education, Science and Technology of Rio de Janeiro (IFRJ) under the CAAE (Certificate of Presentation for Ethical Consideration) number 85859024.5.0000.5268. The research was conducted in accordance with local legislation and institutional requirements. Prior to participation, all volunteers provided informed consent for the use of their data for research purposes.

4.2. Data Processing and Cleaning

The data were processed using the Python programming language. For reading the activity data, the libraries Gpxo, Fitdecode, and Tcxreader were employed, each specific to handling data in GPX, FIT, and TCX formats, respectively. To ensure the robustness of the results, a data cleaning step was applied, in which activities with speed events exceeding 30 km/h and distances shorter than 1.5 km were excluded, as they were unlikely to represent running workouts. Additionally, training sessions that did not contain speed information were also discarded.

4.3. Drift Detection

After the data processing and cleaning steps, the average speed of each individual workout was calculated for each runner. These averages were organized in temporal order, forming historical performance time series. For concept drift detection in these series, the Page-Hinkley method[2], available in the River library, was employed. This method was selected due to its widespread use in the literature and its ability to detect both increases (drift up) and decreases (drift down) in the series behavior, as illustrated in Figure 1. In this study, the hyperparameter threshold was set to 30, based on preliminary experiments that balanced sensitivity and robustness — that is, the method flags a drift only after deviations accumulate beyond this value, reducing false positives from small variations. This threshold represents the maximum deviation tolerated before signaling a change, where lower values make the method more sensitive and higher values, more stable [Ogasawara et al., 2025].

The segment of each runner's training history used to compute and compare training indicators — both for improvement and decline cases — was defined as follows: in the case of improvement, the first improvement drift point over time was identified. If the athlete had previously exhibited a decline drift point, the segment ranged from that decline point to the improvement point, in order to understand which factors or patterns preceded the performance enhancement. If no prior decline point was present, the segment began

[1] Available at: https://www.strava.com/. Last accessed on: April 25, 2025.

[2] Available at: https://riverml.xyz/dev/api/drift/PageHinkley/. Last accessed on: May 4, 2025.

at the start of the time series. For decline cases, the same logic was applied, using the interval from the most recent prior improvement point (if available) to the decline point. Indicators such as average total distance, standard deviation of speed, and the number of training sessions per week and per month were calculated for both segments.

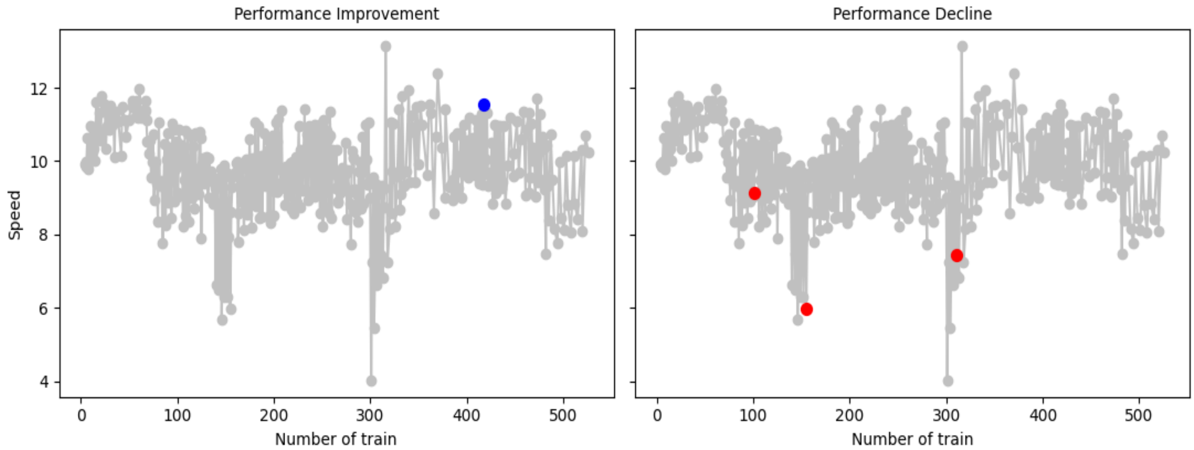

Figure 1. Drift detection in the training history series

5. Results

Among the 19 participants who shared their training history, concept drift was detected in 8 of them. In 3 participants, only drift associated with performance improvement was identified, while in 4 others, only drift related to performance decline was observed. For one participant, both types of drift were detected, indicating the presence of distinct periods of improvement and decline over time (used as an example in Figure 1).

As shown in Figure 2, the selected indicators tend to exhibit higher values in segments associated with performance improvement. The results of the Mann–Whitney tests comparing these segments are presented in Table 1, with all observed differences being statistically significant (p-value < 0.05).

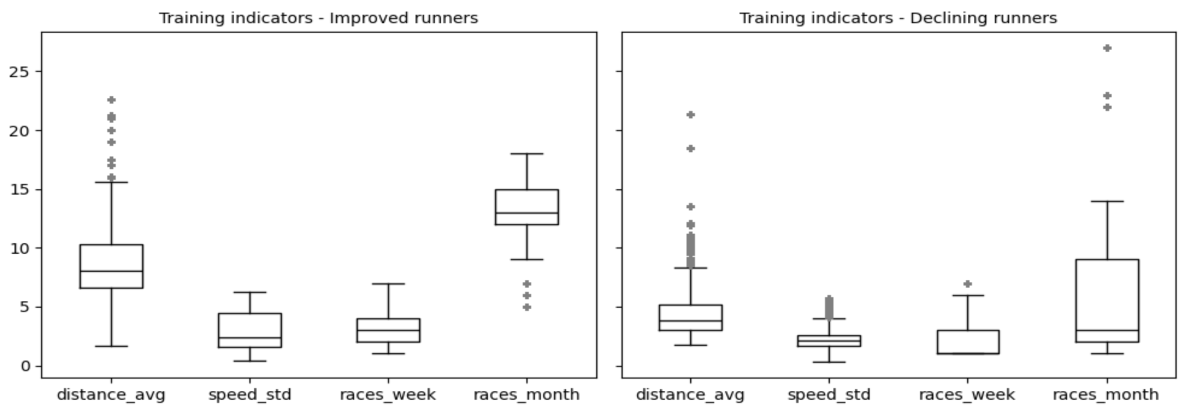

Figure 2. Distribution of training indicators

Table 1. Test results for the training indicators

Training indicators (avg)	Improved runners	Declining runners	*p*-value
Distance (km)	8.46	5.29	0.00***
Variation in training speed (km/h)	2.76	2.18	0.01*
Training sessions per week (freq.)	2.87	1.95	0.00***
Training sessions per month (freq.)	11.59	4.90	0.00*

***p < 0.0001; **p < 0.001; *p < 0,01.

These results are consistent with previous studies, which indicate that training intensity is a relevant factor in achieving faster race times [Knechtle et al., 2011; Tanda, 2011; Suwankan et al., 2024]. Additionally, as a complementary analysis based on the information provided in the survey, it was observed that among the runners who showed performance improvement, most reported engaging more frequently in other types of exercise. However, understanding how these individual factors and characteristics relate to performance improvement requires further investigation.

6. Limitations

This study has some limitations. The number of participants was relatively small, which may have affected the robustness of the results and limited their generalizability. Additionally, this study employed only a single method for concept drift detection, and it is expected that comparing different approaches could contribute to more accurate and robust detection. Furthermore, the feature data collected during the study provide valuable information that could be more extensively explored through prior segmentation aimed at understanding performance based on different runner profiles, using machine learning algorithms. A previous study analyzed different groups of runners based on biological, training, socioeconomic, and psychological factors — such as fear of failure — and identified two distinct groups: “amateur runners” and “recreational runners” [Thuany et al., 2020]. These findings highlight how such variables can be valuable for detecting and understanding different runner profiles.

7. Final Considerations

This study aimed to analyze runners’ performance over time by using historical training time series to identify patterns of improvement or decline through concept drift detection. It also investigated which training features differ the most during periods of performance progression and regression. The findings showed statistically significant results and the proposed approach is expected to offer practical guidance not only to individual runners based on their training history but also to other athletes with similar characteristics.

To achieve this broader goal, this work will be extended in future research phases, including a deeper exploration of the information provided by athletes, aiming to understand how individual factors also influence performance trajectories. This investigation could enable prior segmentation based on runner profiles, facilitating the early detection of improvement patterns and contributing to a better understanding of the relationships between these characteristics and performance developments over time.

References

Agrahari, S. and Singh, A. K. (2024). Comparison based analysis of window approach for concept drift detection and adaptation. *Applied Intelligence*, 55(1).

Berndsen, J., Smyth, B., and Lawlor, A. (2019). Pace my race: recommendations for marathon running. In *Proceedings of the 13th ACM Conference on Recommender Systems*, page 246–250, NY, USA. Association for Computing Machinery.

Besomi, M., Leppe, J., Di Silvestre, M. C., and Setchell, J. (2018). Serun study: Development of running profiles using a mixed methods analysis. *PLOS ONE*, 13(7):1–18.

Casado, A., Hanley, B., Jiménez-Reyes, P., and Renfree, A. (2021). Pacing profiles and tactical behaviors of elite runners. *Journal of Sport and Health Science*, 10(5):537 – 549.

Chang, P., Wang, C., Chen, Y., Wang, G., and Lu, A. (2023). Identification of runner fatigue stages based on inertial sensors and deep learning. *Frontiers in Bioengineering and Biotechnology*, 11.

Knechtle, B., Knechtle, P., Barandun, U., Rosemann, T., and Lepers, R. (2011). Predictor variables for half marathon race time in recreational female runners. *Clinics*, 66(2):287–291.

Komitova, R., Raabe, D., Rein, R., and Memmert, D. (2022). Time Series Data Mining for Sport Data: a Review. *IJCSS*, 21(2):17 – 31.

Nakayama, Y., Kudo, K., and Ohtsuki, T. (2010). Variability and fluctuation in running gait cycle of trained runners and non-runners. *Gait posture*, 31:331–5.

Ogasawara, E. S., Salles, R., Porto, F., and Pacitti, E. (2025). *Event Detection in Time Series*. Synthesis Lectures on Data Management (SLDM). Springer.

Reusser, M., Sousa, C., Villiger, E., Alvero Cruz, J. R., Hill, L., Rosemann, T., Nikolaidis, P., and Knechtle, B. (2021). Increased participation and decreased performance in recreational master athletes in “berlin marathon” 1974–2019. *Frontiers in Physiology*, 12:631237.

Smyth, B. (2019). Recommender systems: A healthy obsession. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):9790–9794.

Stival, M., Bernardi, M., and Dellaportas, P. (2023). Doubly-online changepoint detection for monitoring health status during sports activities. *The Annals of App. Statistics*, 17.

Suwankan, S., Suwankan, S., Theanthong, A., and Kemarat, S. (2024). Prediction of 10 km running time by physical and training characteristics in recreational runners.

Tanda, G. (2011). Prediction of marathon performance time on the basis of training indices. *Journal of Human Sport and Exercise*, 6(3):511 – 520.

Thuany, M., Gomes, T. N., and Almeida, M. (2020). Is there any difference between "amateur" and "recreational" runners? a latent class analysis. *Motriz. Revista de Educação Física*.

Thuany, M., Gomes, T. N., Estevam, L., and Almeida, M. (2021). *Crescimento no número de corridas de rua e perfil dos participantes no Brasil*.

Tito, N., Paixão, B., Tavares, L., Ogasawara, E., and Amorim, G. (2024). Diferenciando perfis de corredores por meio de pontos de mudança nos treinos. In *Anais do XXXIX Simpósio Brasileiro de Bancos de Dados*, pages 834–840, Porto Alegre, Brasil. SBC.

Webb, G., Hyde, R., Cao, H., Nguyen, H.-L., and Petitjean, F. (2016). Characterizing concept drift. *Data Mining and Knowledge Discovery*, 30.

Yung, K. K., Teune, B., Ardern, C. L., Serpiello, F. R., and Robertson, S. (2024). A change-point method to detect meaningful change in return-to-sport progression in athletes. *International Journal of Sports Physiology and Performance*, 19(9):943 – 948.