Article

Early Detection of the Marathon Wall to Improve Pacing Strategies in Recreational Marathoners

Mohamad-Medhi El Dandachi [1], Veronique Billat [1], Florent Palacin [2] and Vincent Vigneron [1,*]

[1] EA 4526-Laboratoire IBISC Paris-Saclay, 91000 Evry-Courcouronnes, France;
    m.eldandachi@etu.emse.fr (M.-M.E.D.); veroniquelouisebillat@gmail.com (V.B.)
[2] Movement, Balance, Performance, and Health Laboratory (EA 4445), Université de Pau et des Pays de
    L'Adour, 65000 Tarbes, France; palacinflorent@gmail.com
[*] Correspondence: vincent.vigneron@univ-evry.fr

Abstract: The individual marathon optimal pacing sparring the runner to hit the "wall" after 2 h of running remain unclear. In the current study we examined to what extent Deep neural Network contributes to identify the individual optimal pacing training a Variational Auto Encoder (VAE) with a small dataset of nine runners. This last one has been constructed from an original one that contains the values of multiple physiological variables for 10 different runners during a marathon. We plot the Lyapunov exponent/Time graph on these variables for each runner showing that the marathon wall could be anticipated. The pacing strategy that this innovative technique sheds light on is to predict and delay the moment when the runner empties his reserves and 'hits the wall' while considering the individual physical capabilities of each athlete. Our data suggest that given that a further increase of marathon runner using a cardio-GPS could benefit of their pacing run for optimizing their performance if AI would be used for learning how to self-pace his marathon race for avoiding hitting the wall.

Keywords: marathon; deep neural; Lyapunov; hitting the wall; VAE

1. Introduction

Marathon running represents one of the most demanding tests of human endurance and strength, requiring participants to demonstrate resilience and mental fortitude over a distance of 42.195 km. Elite marathon runners showcase extraordinary capabilities, with the fastest male athletes averaging speeds of approximately 20 km/h and the fastest female athletes achieving around 19.3 km/h. These remarkable feats are not accomplished through a simple, unwavering pace. Instead, they result from carefully calibrated speed fluctuations designed to balance energy conservation and fatigue management. A key component of these performance strategies is the concept of critical speed. The maximum sustainable aerobic pace an athlete can maintain without rapidly accumulating fatigue. Running at or below this pace allows the body to sustain energy supply and meet the muscular demand for oxygen. Exceeding this velocity, however, forces the body's metabolism to shift towards anaerobic pathways, leading to accelerated fatigue. Through rigorous training, elite runners develop the capacity to approach or even surpass their critical speed for extended periods. By alternating between slower and faster speeds, they optimize their energy use, delay fatigue, and maximize performance [1,2].

This strategy enables them to sustain energy and delay fatigue, thereby optimizing their potential across the race distance without reaching their maximal oxygen uptake (O_2max), which cannot be sustained for a long time [3]. The latter represents the maximum

rate at which an individual can consume oxygen during intense exercise. Higher levels of O_2max are typically associated with enhanced endurance and aerobic capacity, enabling runners to maintain faster speeds over extended distances [4]. The combination of an athlete's O_2max and their tolerance to oxygen deficit (their ability to continue exercising when the demand for oxygen surpasses the supply) helps to inform their pacing strategy and predict their endurance potential. In contrast to the elite runners, recreational runners, an increasingly large and diverse demographic often adopt a rigid pacing approach, aiming to maintain a constant speed throughout the marathon. This approach leads them to achieve their O_2max [5]. This strategy, however, frequently results in a phenomenon known as the marathon wall, which is characterised by a sudden onset of extreme fatigue that typically occurs around the 26th kilometer [6]. The phenomenon is caused by a depletion of glycogen stores and an increased reliance on less efficient energy sources, which results in a significant reduction in speed. A sharp decline in speed is often observed in recreational runners who encounter this phenomenon, resulting in a final median speed that is considerably lower than their initial pace [7]. This pacing challenge emphasises the necessity for a more adaptive strategy that could assist runners in achieving their optimal performance without experiencing significant energy deficits in the latter stages of the race.

As the race progresses, particularly in the final 15 km, physiological indicators such as heart rate, oxygen consumption (O_2), and respiratory rate exhibit increasing variability or entropy [8,9]. In this context, entropy reflects the body's fluctuating physiological state as it strives to meet the escalating demands of prolonged exertion. Two primary types of entropy are relevant to marathon running:

1. Clausius Entropy (Thermodynamic Entropy): Clausius entropy, which is rooted in the principles of thermodynamics, reflects the disorder or heat accumulation within the body during prolonged exertion. As runners approach the final kilometers, their bodies generate a substantial amount of internal heat, particularly in the presence of elevated temperatures, with body temperatures frequently reaching over 40 °C. This rise in thermodynamic entropy can place additional strain on the body's cooling and energy systems, thereby exacerbating fatigue [10,11].

2. Shannon entropy (informational entropy) is defined as follows: Derived from information theory, Shannon entropy refers to the amount of predictability in the runner's physiological data. During the marathon, there is a reduction in Shannon entropy in the final 10 km, which indicates that the physiological responses become less varied and more predictable as the body reaches a state of fatigue. This reduction in informational entropy indicates that the body's capacity to adapt in a dynamic manner is impaired, thereby reducing the efficacy of self-regulatory mechanisms in the latter stages of the race [9].

One promising avenue for enhancing pacing strategies in real time is the use of a variable automatic encoder. This AI-based technology is capable of analysing intricate physiological signals and translating them into actionable feedback for the runner. By dynamically encoding data from variables such as heart rate, O_2, speed, and perceived exertion, a variable automatic encoder can provide a continuously updated representation of the runner's state. Integration of this encoder with a cardio-GPS device could facilitate the provision of personalized pacing adjustments to runners based on their real-time physical state. This would assist them in optimising speed fluctuations without exceeding their limits, and in developing an understanding of the relationship between their perceived exertion and their actual physiological response in terms of % of O_2max. Indeed, AI could assist with pacing in a manner that is commensurate with the Borg Scale for Perceived Exertion [12]. In structured training program, runners are frequently introduced to a variety of levels of exertion, which are typically evaluated using the Borg scale of perceived

exertion. The scale ranges from 6 to 20, with each level corresponding to a subjective feeling of effort and difficulty. The most commonly utilized levels for marathon training are as follows [13]:

The pace is designated as "Easy" and is equivalent to a speed of 5.5–6.4 km/h. The individual is comfortable and able to engage in conversation.

The 14th level of exertion, designated as "moderate pace", is characterized by a moderate level of effort that allows for sustained conversation. The level of difficulty is moderate, yet the pace is sustainable.

The seventeenth level is designated as "Hard Pace." This level of exertion is considerable, yet it can be sustained for relatively brief distances.

The Maximal Effort (20): This level of exertion is close to the point of exhaustion and is therefore not sustainable over extended periods of time.

Perceived exertion levels allow runners to adjust their pace according to the specific demands of a race. This study aims to determine whether an AI system can learn these exertion levels through calibration tests to establish a personalized energetic signature. Based on O_2max, oxygen deficit tolerance, and perceived exertion, this signature could guide AI-assisted pace adjustments during the race. By analyzing sophisticated physiological data, this pilot study explores whether an AI-powered system can offer more effective pacing strategies for marathon runners than traditional cardio-GPS devices. The goal is to design an adaptable pacing assistant that integrates and encodes each runner's unique physiological responses, enabling optimal pace adjustments while avoiding the rigidity of fixed speed maintenance. Such an approach could allow recreational runners to input personalized pacing profiles into their devices, with AI guiding them through optimal speed variations across a marathon distance. This innovation aims to support runners in achieving their personal best under safe, optimized pacing conditions, fostering greater enjoyment and sustainability, especially for diverse and aging populations. To demonstrate this potential, a pilot experiment investigates the capacity of a deep neural network to extract an energetic signature. Physiological modulations in a cohort of runners during a marathon were analyzed using extensive datasets, including heart rate and speed (Garmin 630), oxygen uptake (O_2), respiratory frequency, and metabolic data (Cosmed K5). The use of deep learning in multi-sensor sports data analysis remains uncommon, partly due to the high cost and variability of data acquisition. To address this, innovative strategies such as fractal methods and data augmentation techniques, including sliding windows, were applied to enhance temporal progression within datasets.

The study focuses on analyzing marathon runners' physiological parameters through deep neural networks to derive insights about performance and propose race strategy adjustments. Artificial intelligence can assist sports physiologists in prioritizing performance tests, uncovering details that are otherwise inaccessible to the human eye. A Variational Autoencoder (VAE), a generative statistical model, was used to create individual signatures sensitive to physiological variations and fatigue [14]. Additionally, Hölder exponents and multifractal spectrum analysis provided a deeper understanding of cardiac autoregulation during intense exercise [15]. While Lyapunov exponents have been used to characterize equilibrium plateau [16], their integration into a multivariable energetic context remains unexplored and could help identify exhaustion points and unsustainable pacing with greater precision.

The study seeks to elucidate the unique physiological signatures of marathon runners and explore the interpretability of Garmin and K5 data. Ultimately, it aims to detect fatigue-induced disruptions in race dynamics, advancing our understanding of marathon performance and supporting improved training and race strategies.

2. Materials and Methods

2.1. Subjects

Even if we started with 10 runners, one of them was excluded on account of incomplete data. This was due to a malfunction in the analyzers (a battery issue) that occurred at the half-marathon. Then, nine recreational but well-trained male marathon runners (mean age: 40.1 ± 10.6 years; weight: 72.7 ± 6.5 kg; height: 178.3 ± 7.5 cm) with performance representative of the average performance of non-elite runners but well-trained male marathon runners whose performance are in the first quartile of performance in popular marathon as the Paris marathon [17] (Table 1). To avoid introducing additional variables that could impact statistical analysis, we deliberately included only one gender in our investigation.

Table 1. Subjects Age, Personal Best Marathon Time and the Performance at the Sénart Marathon. * Best Personal Time reached at the Sénart Marathon.

N° Runners	Age (Years)	Faster Marathon Time (Years)	Sénart Marathon (2019)
1	47	03 h 12′48″ (2016)	03 h 31′34″
2	44	03 h 34′57″ (2019)	03 h 34′57″ *
3	22	03 h 22′40″ (2019)	03 h 22′40″ *
4	34	02 h 50′00″ (2019)	02 h 50′00″ *
5	47	02 h 59′22″ (2016)	03 h 32′07″
6	58	03 h 27′32″ (2013)	04 h 30′34″
7	29	02 h 57′03″ (2015)	03 h 14′24″
8	36	03 h 27′58″ (2017)	03 h 51′44″
9	43	02 h 44′00″ (2015)	03 h 13′53″

All participants volunteered and maintained their regular training routines without alterations. The selected runners had prior experience completing a minimum of two marathons. They had been engaged in consistent training, involving three to four sessions per week, covering a range of 50 to 80 km per week, for over 5 years. Every week, the participants incorporated a High-Intensity Interval Training session, involving 6 repetitions of 1000 m at intensities between 90% to 100% of their maximal heart rate, along with a tempo training session of 15 to 25 km at speeds ranging from 100% to 90% of their average marathon pace.

Ethical considerations were met, as the study's objectives and procedures received approval from an institutional review board (CPP Sud-Est V, Grenoble, France; reference: 2018-A01496-49). All participants were well-informed about the study and provided written consent to participate. Table 2 in the manuscript presents information regarding participants' ages, their personal best marathon completion times, and the year these performances were achieved. Notably, some of the runners achieved their personal best during the Sénart Marathon.

Table 2. Hyperparameters chosen for VAE-K5 (K5 Cosmed) and VAE-GAR (Garmin watch). The model was trained on a GPU-equipped system NVIDIA RTX 3090 with 32 GB RAM, and each training run took approximately 2 h for convergence. The VAE architecture consists of two fully connected encoder and decoder layers, with a latent dimension of 24 or 12, optimized using the Adam optimizer with a learning rate of 10-4.

	VAE-K5	VAE-GAR
input size	60	60
number of samples	10,844	10,844
Latent dim	24	12
Hidden Layers	3 (encoder)	3 (encoder)
Batch Size	32	64
Learning rate	10-4	10-4
KL div. weight	10-4	10-3
Epochs	350	200

2.2. The Marathon and Experimental Measures

In the context of the marathon event, all participants participated in an official race known as the Sénart Marathon in France. The race commenced at 9 a.m. under specific environmental conditions. On 1 May 2019, in Sénart, the weather included temperatures between 11 and 15 °C (from 9 a.m. to 1 p.m.), no precipitation, and an average humidity of 60%. Blood lactate levels were assessed using a finger-based lactate measurement device (Lactate PRO2 LT-1730; ArKray, Kyoto, Japan) immediately after a 15-min warm-up at a leisurely pace and three minutes after crossing the finish line.

Throughout the study, we collected continuous data on respiratory gases (oxygen uptake [O_2], ventilation [E], and respiratory exchange ratio [RER]) using a portable telemetric system (K5; Cosmed, Rome, Italy) that allowed breath-by-breath analysis. Additionally, a combination of a global positioning system (GPS) watch (Garmin, Olathe, KS, USA) and the K5 system was utilized to monitor heart rate and speed responses, with 5-s averaged data, during each trial. The data collection sampling frequency were the following:

- Garmin data: Collected at a frequency of 1 Hz (once per second), providing continuous recording of the athlete's pacing, heart rate, and GPS coordinates.
- Cosmed K5 data: Acquired at a sampling rate of 0.2 Hz (once every 5 s), capturing breath-by-breath physiological parameters, including oxygen uptake (VO_2), carbon dioxide production (VCO_2), and ventilation (VE).
- Outlier Treatment: Outliers were identified using statistical criteria based on physiological plausibility, particularly for heart rate, VO_2, and speed data. Heart rate values outside physiologically plausible ranges (below 40 bpm or above 220 bpm) were identified as outliers. VO_2 and speed data points were similarly screened using statistical thresholds (values beyond three standard deviations from the participant's mean). Approximately 1.5% of the collected data points were identified as outliers and subsequently removed from the dataset.
- Missing Data: Missing data (2% of the dataset) ocasionally arise due to sensor failures, communication issues with the athlete, or manual data-collection errors. Missing data were addressed using spline interpolation when the gap was short (less than 30 s), connecting the last known data point to the next known data point using polynomials for a smoother approximation. Longer gaps or substantial missing segments were excluded from the analysis to ensure data integrity. Kalman filter from Subway and Stoffer [18] were used in this paper for ensuring each data point is assigned one definitive imputed value.
- Scaling of Data: for any variable (e.g., heart rate), the standardized value calculated as, where denote the mean and standard deviation of x over all observations in the dataset, All physiological variables–heart rate, VO_2, and speed–were standardized using Z-score normalization before being provided as inputs into the VAE. Such uniform scaling also assists in dimensionality reduction and network training by preventing any one variable's scale from dominating the learning process. To prevent pacing-related influences, runners were encouraged to self-pace their runs while the cardio-GPS display was concealed.

Hydration and refreshment points were available every 5 km during the marathon, with additional stations offering sponges and sustenance every 5 km from 7.5 km onward. Runners could remove their masks at these points to consume food and beverages. Consistently, all participants drank one glass of water and consumed fruit at each hydration station along the route and at the Start/Finish area.

For metabolic assessment during exercise, the participants utilized COSMED reusable face masks constructed from silicone to prevent allergenic reactions. These masks were ergonomically designed to fit snugly and comfortably, maintaining a proper seal without

compromising data accuracy. In high-intensity exercises, the runners used masks with inspiratory valves to reduce resistance during inhalation, enhancing comfort.

The rate of Perception Exertion (RPE) was tracked using the Borg 6–20 scale [12], and participants reported their level of fatigue at least every kilometer or more frequently as needed. This scale was employed to correlate physiological stress indicators with marathon fatigue, and participants were familiarized with it during the two weeks preceding the race.

2.3. Mathematical Procedure
2.3.1. Variational Auto Encoder

Autoencoders are a class of neural networks widely employed in unsupervised learning tasks, particularly in the domain of dimensionality reduction, feature learning, and data generation. The fundamental architecture of an autoencoder consists of an encoder and a decoder, which work in tandem to learn a compressed representation of the input data. The encoder maps the high-dimensional input data into a lower-dimensional latent space, capturing essential features and patterns. Subsequently, the decoder reconstructs the original input data from this compressed representation (Figure 1).
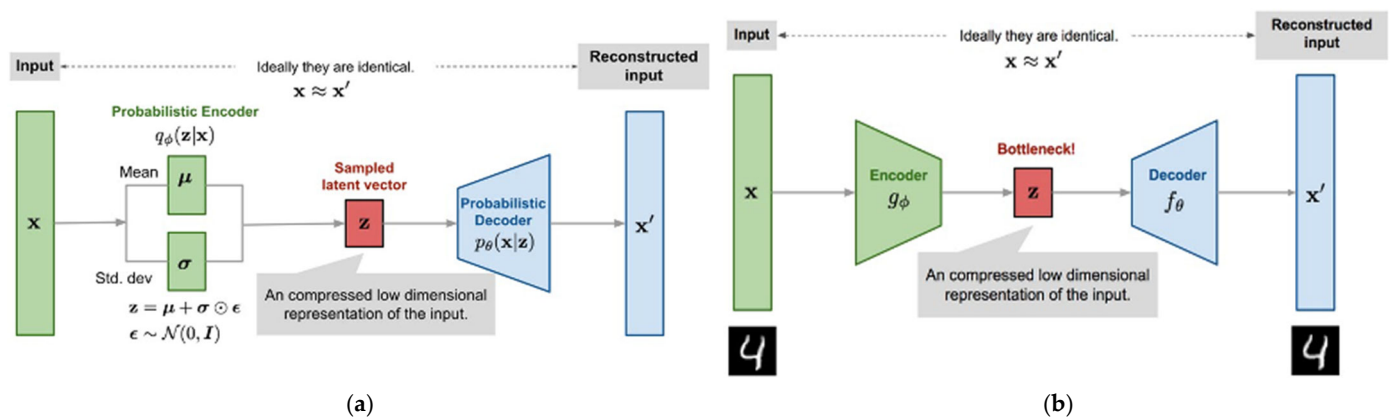

(**a**)   (**b**)

Figure 1. Principle of Autoencoders (**a**) Illustration of Variational Autoencoder (VAE) model architecture. (**b**) Illustration of Autoencoder (AE) model architecture; (**b**) Illustration of Variational Autoencoder (VAE) model architecture.

The encoder and decoder components of a standard autoencoder are typically implemented using feedforward neural networks. During training, the autoencoder aims to minimize the difference between the original input and the reconstructed output. This process encourages the network to learn a meaningful encoding of the input data in the latent space. Autoencoders have found applications in various fields, including image denoising, anomaly detection, and feature extraction. While conventional AEs are effective at learning compact representations of data, they lack a probabilistic interpretation, making them limited in tasks that require uncertainty estimation or generative capabilities. Variation Autoencoders (VAEs) address this limitation by introducing probabilistic modeling into the autoencoder framework. VAEs reinterpret the latent space as a probability distribution, allowing for the generation of new data samples by sampling from this distribution. The key principle of VAEs is to impose a constraint on the latent space's distribution to encourage it to follow a specific prior distribution, often a Gaussian distribution. This constraint is typically enforced using the Kullback-Leibler (KL) divergence, which encourages the learned distribution to be like the chosen prior (Table 2). During training, VAEs aim to minimize two primary components of the loss function: the reconstruction loss, which ensures faithful data reconstruction, and the KL divergence, which aligns the learned latent distribution with the desired prior distribution. Balancing these two components enables

VAEs to generate new data samples that exhibit both meaningful variations and adherence to the underlying data distribution [19].

Here we detail the experimental setup, including the architecture and hyperparameters of the VAEs trained in the frame of the project. The selection of these parameters was guided by both prior research in the field and empirical evaluation on our specific dataset. We use the MSE model to calculate the reconstruction loss, while we systematically tuned the hyperparameters to achieve optimal performance and convergence during training.

The Variational Autoencoder (VAE) was trained from scratch, without any pre-trained weights or transfer learning, ensuring that all latent representations emerged solely from the given race data. The following hyperparameters were chosen based on a combination of heuristics and experiments and are listed in Table 2. To validate the chosen parameters, we conducted a sensitivity analysis by systematically varying key hyperparameters while monitoring performance metrics. The results indicated that the selected parameters yielded optimal trade-offs between model complexity, convergence, and reconstruction quality. To ensure reproducibility, a public code repository (GitHub) will be provided upon publication, containing (1) data preprocessing scripts to replicate input transformations (2) model architecture and training configurations for consistent setup (3) hyperparameter settings and evaluation metrics.

The Variational Autoencoder (VAE) was chosen for this study due to its ability to perform unsupervised learning and latent space modeling, making it particularly well-suited for capturing the intrinsic structure of marathon race data. Unlike supervised models that require labeled data, the VAE learns meaningful representations without explicit labels, which is crucial given the limited number of runners. The model allows for dimensionality reduction, encoding race dynamics into a latent space where variations in runner performance, including the "hitting the wall" phenomenon, can be analyzed.

The choice of a VAE is motivated by its ability to learn meaningful latent structures in an unsupervised manner, which is crucial given the limited number of runners and the absence of explicit labels. The VAE facilitates dimensionality reduction while preserving the variability in race dynamics, allowing for an interpretable latent space. Supervised models such as CNNs are primarily designed for classification tasks and require labeled data, which is not available in this context. While recurrent neural networks (RNNs) and long-short term memory models (LSTMs) are effective for time-series prediction, this study does not aim to predict the race timeline but rather to explore underlying latent variables. Moreover, recurrent models often require large datasets and significant computational resources, whereas VAEs efficiently extract meaningful representations with limited data.

Alternative approaches such as standard autoencoders could have been considered; however, VAEs offer a probabilistic framework that enhances generalization and prevents overfitting. A potential comparison with other models, such as GANs or deterministic autoencoders, could further validate the VAE's capacity to learn structured latent spaces. The primary objective is not classification or forecasting, but rather uncovering hidden dynamics in race behavior, making the VAE the most suitable choice for this study.

2.3.2. Model Validation

Given the unsupervised nature of the task, traditional supervised validation techniques such as accuracy or F1-score are not directly applicable. Instead, the model's performance was assessed using reconstruction error, latent space coherence, and consistency of learned representations across different runners. To ensure robustness, a holdout validation approach was used, where the dataset was split into 80% for training and 20% for testing. The reconstruction loss (Mean Squared Error and Kullback-Leibler divergence) was monitored to prevent overfitting. Additionally, the VAE was tested on an independent

subset of race segments, verifying that latent variables remained stable and interpretable across different runners.

Further evaluation involved visualizing the latent trajectories, revealing meaningful groupings related to physiological events like "hitting the wall". To confirm the VAE's reliability, multiple training runs with different initialization seeds were performed, ensuring that the latent space structure was consistent across experiments. While standard supervised benchmarks are not applicable here, these validation strategies provide strong evidence that the VAE captures relevant race dynamics in a reproducible and generalizable manner.

The small-sample size constraint is mitigated by the methodology used to extract overlapping temporal windows from race signals, significantly increasing the number of training samples. By segmenting race data into smaller windows, the VAE learns local patterns and variations within individual performances, rather than being constrained to whole-race trajectories. This approach enhances the effective sample size, improving the model's ability to generalize. While a larger and more diverse dataset would undoubtedly strengthen the conclusions, the primary objective of this study is not to make broad population-level predictions, but rather to explore latent representations of race dynamics. The findings serve as a proof of concept, demonstrating that the VAE can encode meaningful variations in running performance.

2.3.3. Construction of the Learning Dataset

- Principle of sliding window segmentation with overlapping for dataframe:

To use the datasets at hand while increasing the amount of data for training and considering the temporality of the race, we chose to implement a sliding window on our pre- existing datasets. The sliding window technique is commonly employed in signal processing and data analysis, particularly in scenarios where data is acquired progressively or in a continuous stream. This technique enables data croping in a "sliding" manner.

In scenarios involving progressive data acquisition, such as streaming sensor data or online monitoring systems, it is often impractical to process the entire dataset at once due to memory and computational constraints. The sliding window segmentation technique addresses this challenge by considering a subset of the most recent data points, which are continuously updated as new data arrives. This approach maintains a fixed-size window over the data stream, allowing for the analysis of temporal trends, patterns, or features within the window's span.

- Implementation in the frame of our study:
  (1) Initialization: Select an appropriate window size, denoted as Lw which determines the number of data points included in the window. At first, we consider the array formed with first Lw points of a given run dataset and add it to the learning dataset. (We work in a multi-variable environment; the sliding window goes over all the variables measured by a specific acquisition system over time: the output of such a process will thus be matrixes of size 'Lw $\times$ 12' for K5, 'Lw $\times$ 4' for GARMIN) [20].
  (2) Window Update: The ΔT oldest data points are removed from the beginning of the window as we inject the ΔT ones. (ΔT: Time Shift) (Figure 2).
  (3) Outcome: We stop window sliding further when we are at less than ΔT points away from the end of the dataset. The last array is stored and added to the learning dataset.
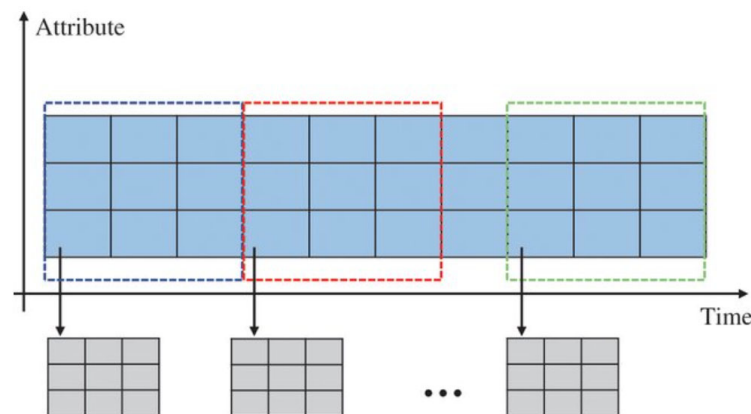
Figure 2. Illustration of the fixed length sliding window slicing process for a progressive single-data acquisition: the abscissa represents time, the *y*-axis the observed modality.

Choosing an appropriate window size Lw is critical, as a smaller window may capture rapid changes but overlook long-term trends, while a larger window may smooth out important fluctuations. By experience, we can show that $L_w = 60$, $\Delta T = 1$ (or a window of 10 min and a shift of 10 s) give satisfactory results in the context of our study. The resultant learning dataset thus contains over 8000 2D samples for K5, and over 10,000 2D samples for GARMIN.

2.3.4. Marathon Individual Signature

- Introducing t-SNE: To visually comprehend the intricate dynamics and patterns inherent in marathon races, we employ t-Distributed Stochastic Neighbor Embedding (t-SNE) in conjunction with our Variational Autoencoder (VAE). t-SNE is a nonlinear dimensionality reduction technique that excels at preserving local structures in high-dimensional data when projected onto a lower-dimensional space. The primary goal of t-SNE is to map high-dimensional data points into a 2D or 3D space while maintaining the relationships and distances between these points as closely as possible. This makes t-SNE particularly effective at revealing clusters, patterns, and disparities in complex datasets [21].

- Visualizing 2D signatures of Marathon races: In our study, the VAE compresses the high-dimensional feature space of marathon race data into a lower-dimensional latent space. However, to gain intuitive insights and visually represent the runners' behaviors, we extend the analysis by integrating t-SNE. By applying t-SNE to the latent representations learned by the VAE, we further reduce the dimensionality and capture intricate structures that might not be apparent in the original feature space. The 2D t-SNE signatures provide a human-interpretable representation of the complex dynamics observed during marathon races. These signatures allow for exploring individual runner trajectories within the reduced space. This level of interpretability is particularly valuable for our study's objectives, where understanding the variations and patterns in runners' behaviors contributes to comprehensive insights into race dynamics.

- Marathon signature viability: The choice of the hyperparameters in Table 2 was ultimately based upon the cleanliness of the signatures obtained with the t-SNE. To evaluate their viability, we have decided to rely upon their continuity. There are several methods available to quantify the "continuity" of a discrete dynamic system. Each method offers a distinct perspective on how to measure the smoothness or regularity of transitions within the system. The study of the variation rate—measuring the rate of change between consecutive values in a time series [22]—or even entropy measures [23] and Fractal Analysis [24] are means available to rate our satisfaction of

the signatures obtained. While this criterion is important and must be looked upon, we suppose in the frame of this project that the signatures we visualized on are decent enough to work with.

2.4. Use of Lyapunov Exponents for Anticipating the "Marathon Wall"

2.4.1. Lyapunov Exponents

Lyapunov exponents are mathematical quantities used to characterize the behavior of dynamic systems—whether discrete or continuous, particularly in the context of chaos theory. They provide insights into the sensitivity of trajectories within a system to small perturbations, aiding in identifying chaotic or unpredictable behavior.

For our analysis, we focus on the first Lyapunov exponent, denoted as λ_1. Its value is particularly insightful for quantifying chaos within dynamic systems. It characterizes the rate at which initially close trajectories in the phase space diverge exponentially, a hallmark of chaotic behavior. Higher values of λ_1 indicate more robust chaos [25], suggesting that the system is highly sensitive to initial conditions and prone to unpredictability.

In the context of our study analyzing marathon runners' behavior during a race, the calculation of the first Lyapunov exponent serves several vital purposes. Firstly, it quantifies the extent of variability and unpredictability exhibited in runners' behavior throughout the race. Additionally, an elevated first Lyapunov exponent could signify the presence of intricate chaotic patterns within runners' pacing and positioning, reflecting nuanced interactions among diverse influencing factors. Moreover, this analytical approach offers insights into performance dynamics, illuminating how minor deviations in initial conditions can yield diverse outcomes among individual runners. Lastly, shifts in the behavior of the Lyapunov exponent could correspond to critical junctures within the race, such as the commencement, culmination, or demanding segments, thereby elucidating runners' responses to specific race dynamics. In essence, calculating the first Lyapunov exponent within our marathon study plays an essential role in characterizing variability, detecting chaotic tendencies, comprehending performance dynamics, and unveiling pivotal transition points in the race.

2.4.2. Computing λ_1 with the Wolf ODE Algorithm

Consider a dynamic system described by the vector differential equation:

$$\frac{dx}{dy} = f(x) \tag{1}$$

where represents the state vector and represents the vector field governing the system's dynamics. The Wolf ODE algorithm estimates λ_1 using the concept of tangent vectors and their evolution and consists of several key steps detailed in Table 3 [26]:

Table 3. Computation of the First Lyapunov Exponent Using the Wolf ODE Algorithm.

Mathematical Formulation	Description	Step
$T_0 = I_n$	Initialize the tangent matrix to identity, representing orthonormal tangent vectors.	1. Initialization
$T_{t+1} = J_t \times T_t$	Compute the Jacobian matrix J_t of the system at each time step.	2. Jacobian Update
$T_{t+1}^{orth} = GS(T_{t+1})$	Apply Gram-Schmidt orthogonalization to T_{t+1} to maintain orthonormality.	3. Orthogonalization
$\delta x_{proj,t} = \delta x(t) . T_t^{orth}$	Project the separation vector onto the tangent space to quantify divergence.	4. Projection
$\lambda_1 = \frac{1}{N\Delta_t} \sum_{t=1}^{N} ln(\delta x_{proj,t})$	Compute the logarithm of the norm of projected vectors and accumulate.	5. Accumulation

In the same manner as in Section 2.3.2, we constructed a sliding window over each race of chosen size Lw to plot the over time. We then compute with the Wolf algorithm method detailed above over the Lw window points to obtain the plot for the entire race. After optimization, we found that Lw = 30 is optimal, large enough for to make sense and small enough to get precise race sensitivity information.

2.5. Mann-Whitney U Test for Comparing the Appearance of Fatigue-Induced Cracks in the K5 and Garmin Datasets

The Mann-Whitney U test has proven to be a valuable tool in our comparative analysis of the appearance of fatigue-induced cracks in the K5 and Garmin datasets.

3. Results

3.1. Individual Marathon Signatures

In this section, we present the distinctive runner signatures we derived from our analysis. These signatures encapsulate the unique behavioral profiles of marathon participants during the race.

3.1.1. Convergence Analysis of the Variational Autoencoder (VAE)

To assess the convergence behavior and the effectiveness of our VAE in capturing the underlying structure of marathon race data, we conducted a rigorous convergence analysis. This analysis involved monitoring key performance metrics as the training process unfolded.

Loss Function Convergence: One crucial indicator of VAE convergence is the behavior of the loss function over epochs. As training progressed, we tracked the reconstruction loss (rloss) and the KL divergence loss (klloss). rloss decreased as expected following an inverted exponential tendency, going from 0.9980 at the beginning of the training to as low as 0.0682 after 350 epochs. The reduction in these losses signifies the VAE's ability to reconstruct input data faithfully while learning informative latent representations.

Training stability: We closely observed the training stability, ensuring that the VAE did not exhibit signs of overfitting or under-fitting. Validation datasets were used after the training to ensure these conditions were respected while changing the number of neurons in hidden layers to optimize our machine. For the final hyperparameters regrouped in Table 2, the validation error is 0.1316 for VAE-K5 (relative error of 8%), which is acceptable is the frame of our study.

3.1.2. Signatures Obtained with the K5/GARMIN Datasets

Before visualizing the signatures, runners were ranked based on their race finishing time. This ranking allows us to order the presentation of runner signatures in a structured manner, pro- viding a clear progression from the highest-performing to the lowest-performing participants. This approach enables us to observe potential correlations between performance and behavioral patterns, offering a more insightful perspective on the dynamic behaviors exhibited during the marathon race. Each participant is associated with a letter, going from A (highest performance) to H (lowest performance). Runner no1 (see Table 1) K5 acquisition system did not consistently measure the entirety of the race, and for the sake of homogeneity of the results its signature won't be considered in this paper.

Figure 3 regroups the signatures obtained with the t-SNE method on our VAE-K5 latent space after training. We have implemented a color-coded scheme to represent the progression of the marathon race visually. This gradual shift in colors allows us to track the temporal aspect of each runner's performance intuitively. Blue signifies the early stages of the race, while yellow represents the concluding moments, enabling us to observe how runners adapt and evolve their behaviors throughout the race duration. It is important to note that the orientation of the runner signatures presented here is not a critical factor in our analysis. Instead, our

focus lies in the intricate details of these signatures, their shape, loops, and ruptures. While the X and Y axes of the t-SNE representation provide a visual layout, the true essence of these signatures emerges from the patterns and behaviors they encapsulate. Our observations suggest that the continuity and smoothness of the signature, rather than its orientation, may indicate performance. Interestingly, a more fragmented or "broken" signature appears to correlate with a higher level of performance, implying that certain deviations from uniformity may signify adaptive strategies employed by runners. Indeed, the 4 best runners seem to have a similar "triangular" signature with an increased number of loops and rupture, with a more visible "break" at ±60% of their race. On the other hand, the 4 lowest performances seem to have more of a "linear" shape, with fewer twists and turns than their counterparts. These turns might signify adaptive pacing strategies, efficient navigation of the course's challenges, or a heightened ability to respond to varying race dynamics. This observation underscores the value of examining the nuanced behaviors encoded within these signatures as potential indicators of marathon performance.

Visualization of K5 signatures:

(a)

(b)

(c)

(d)

(e)

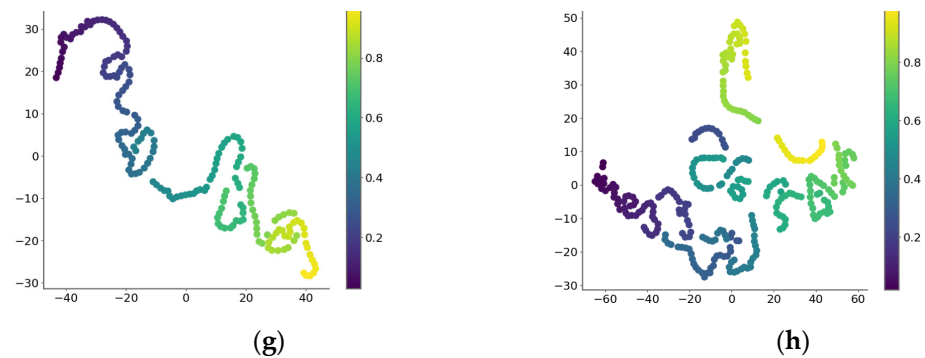(f)

Figure 3. *Cont.*

(**g**)

(**h**)
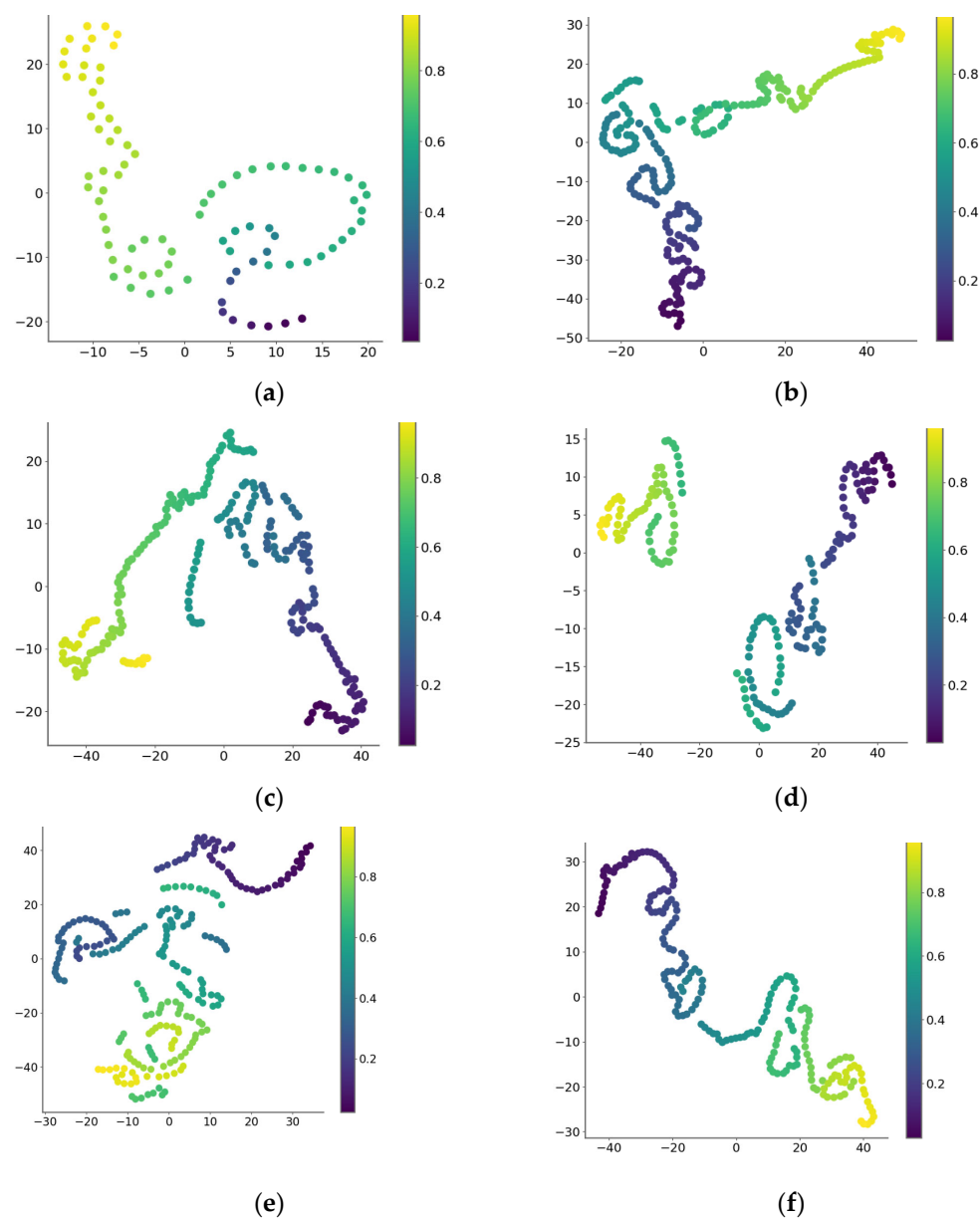
Figure 3. Runners signatures for the K5 dataset classed from Highest (**a**) to lowest (**h**) performance. Runner A, who completed the marathon in 2 h and 50 min, was the fastest among all participants, finishing approximately 20 to 30 min ahead of the others (see Table 1). However, the significantly lower number of data points in Runner A's graph is not solely attributable to this faster pace. The main cause is the premature termination of data collection by the K5 analyzer, which stopped recording at around 70% of the race (approximately 30 km) due to battery depletion. (**a**) runner A, (**b**) runner B, (**c**) runner C, (**d**) runner D, (**e**) runner E, (**f**) runner F, (**g**) runner G, (**h**) runner H.

Figure 4 this time regroups the signatures obtained on the GARMIN dataset (VAE-GAR), containing the acquisition of 4 variables (Speed, HR, Pace, Stride Length).

Visualization of GARMIN signatures:

(**a**)

(**b**)

(**c**)

(**d**)

Figure 4. *Cont.*

(**e**)

(**f**)

(**g**)

(**h**)
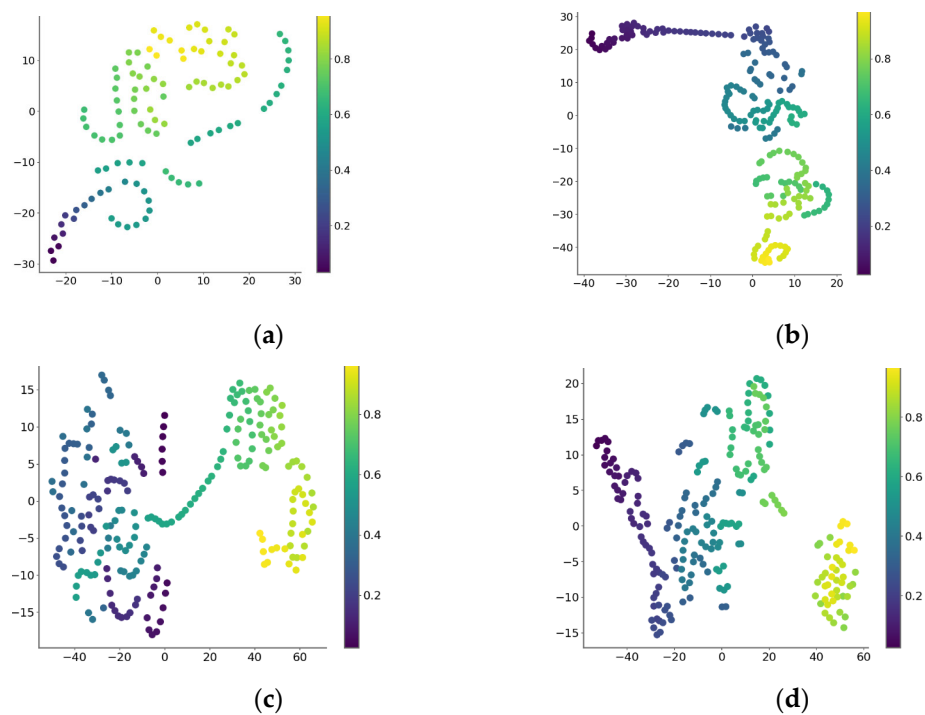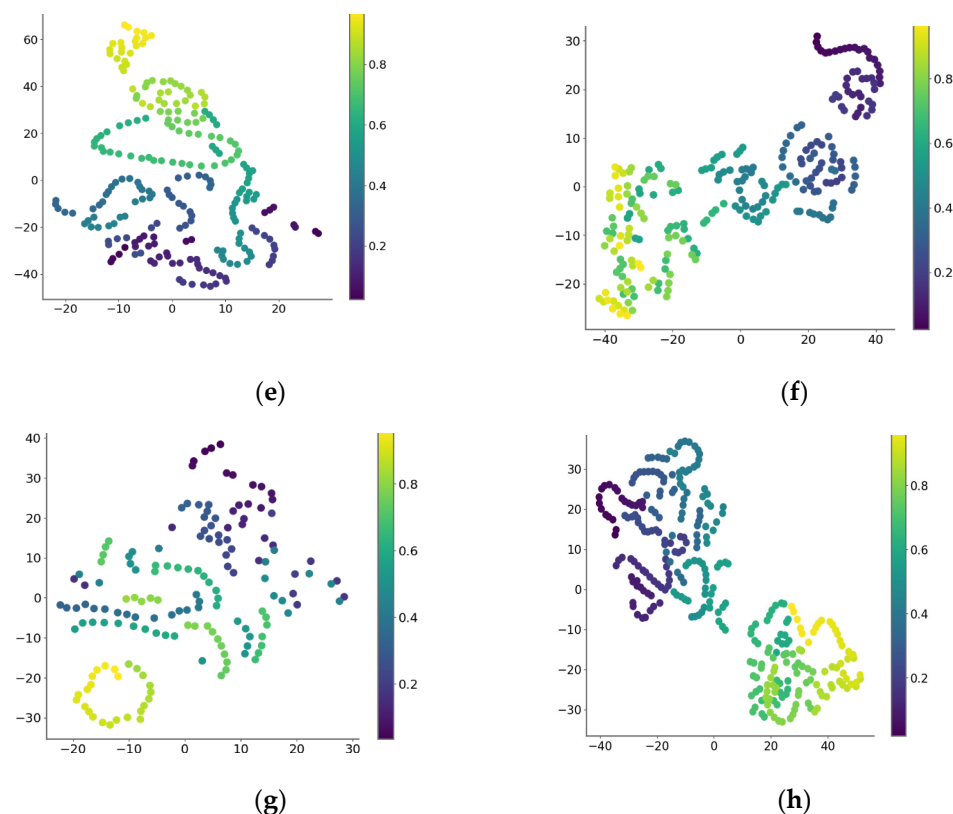
Figure 4. Runners signatures for the GARMIN dataset classed from Highest (**a**) to lowest (**h**) performance. Runner A's dataset contains substantially fewer data points than those of the other runners, whose recordings continued through to the finish line. (**a**) runner A, (**b**) runner B, (**c**) runner C, (**d**) runner D, (**e**) runner E, (**f**) runner F, (**g**) runner G, (**h**) runner H.

3.1.3. Comparison K5 and GARMIN Signatures

Notably, as we can observe in Figure 5, the GARMIN signatures present a distinct challenge in terms of visual interpretation compared to those derived from the K5 dataset. The visual complexity and sparser nature of the GARMIN signatures introduce a heightened level of intricacy, with only a limited number of discernible similarities apparent at first glance.

(**a**) runner B—GARMIN and K5 signatures

Figure 5. *Cont.*
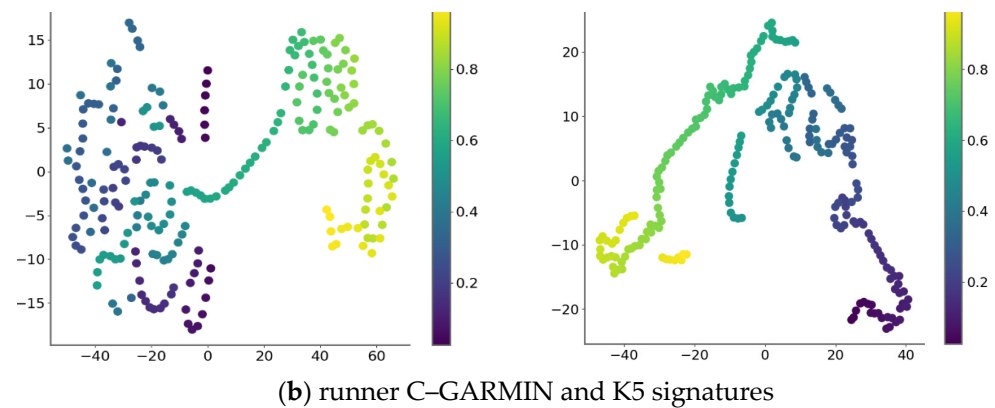
(**b**) runner C–GARMIN and K5 signatures
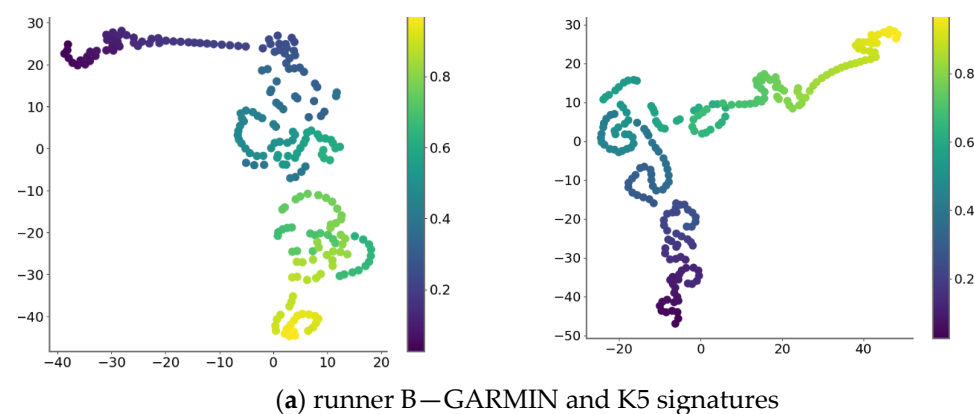
Figure 5. Comparative analysis of runners B and C signatures: We can retrieve the "triangular" shape and overall pattern for Runner B. On the other hand, the signatures are visually very different for Runner C.

3.1.4. From a Visual to a Mathematical Characterization of Performance

Visually, we tend to say that K5 and GARMIN signatures are different and do not permit a parallel interpretation between them. Excluding the particular 'triangular' shape that we can find in both Runner B signatures, K5 seems to contain much more information on the dynamics of race, characterized by the superior number of loops and turns that it presents. Furthermore, the more continuous and clean aspect of K5 signatures leads us to believe that they better represent the reality of dynamics observed during the marathon. However, this hypothesis still needs to be confirmed under precise mathematical conditions.

While the runner signatures provide valuable visual insights into the race dynamics of individual runners, it's crucial to emphasize that a more thorough quantification of their "control" over the race requires robust mathematical modeling. Thus, we turn our attention to the upcoming section on Lyapunov Exponents. Here, we will employ mathematical models to assess the chaotic nature of marathon race dynamics rigorously. By quantifying the sensitivity of trajectories and exploring how small perturbations in initial conditions can lead to divergent outcomes, we aim to elucidate the intricate and adaptive behaviors exhibited by marathon runners during races. Additionally, it could potentially serve as an opening for individual performance enhancement. Understanding the underlying chaos and variability in race dynamics may empower runners and coaches to develop more effective training strategies, pacing techniques, and race management approaches. By harnessing this knowledge, runners can strive for improved and more controlled performances to pursue personal and competitive success.

3.2. Unveiling the Chaos: Exploring Marathon Race Dynamics Through Lyapunov Exponents Individual Marathon Signatures

3.2.1. Graphical Representation of Chaos Dynamics

Similarly, to the signatures, these Lyapunov Exponents are directly computed from the latent space, providing a quantifiable measure of chaos within the runners' performances. We thus have plotted λ_1 throughout the race for all runners for both K5 and GARMIN datasets thanks to the Wolf ODE algorithm, presented in Section 2.3.2. Figure 6 regroups the graphs obtained for three runners, ranked by increasing order of performance. The proposed arrangement suggests a progressive flattening of the K5 curves, accompanied by a reduction in oscillations in the GARMIN curves. Additionally, it is observed that the top runners tend to have their λ_1 curve positioned above the *x*-axis. However, further observations are needed to confirm this trend. These Lyapunov plots have been made for all runners of the cohort except Runner G, who admittedly didn't run this marathon as he
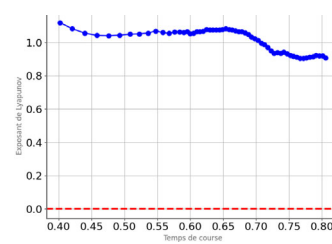
would usually have, a.k.a running the entirety of the race at a slow pace. The choice to not represent his λ_1 tendency has been made not to falsify further results.
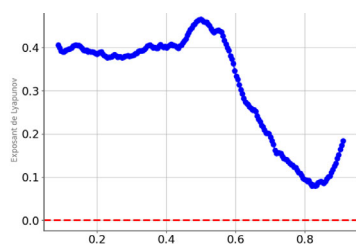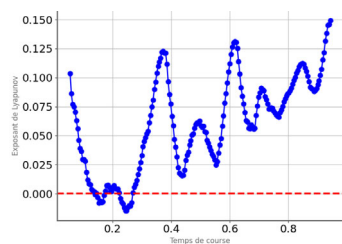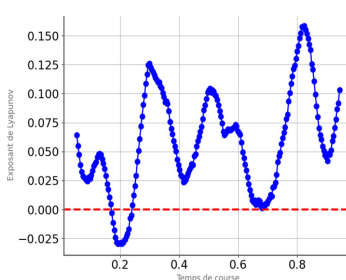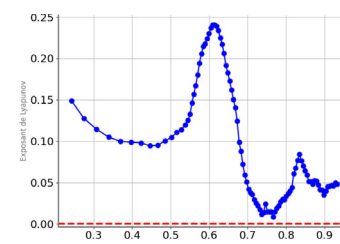

Runner C-K5
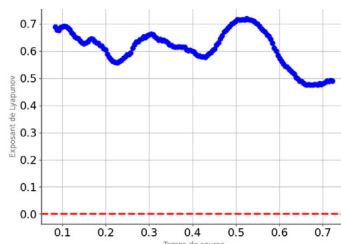
Runner D-K5

Runner H-K5

Runner C-GARMIN

Runner D-GARMIN

Runner H-GARMIN

Figure 6. Dynamics of λ_1 over time for three runners going from the highest (**left**) to lowest (**right**) performance. Note that Runner D optimized his performance compared to his last race.

Within the context of marathon racing, the Lyapunov exponent graph unveils not only the control exhibited by runners but also provides insights into moments of potential fatigue and the phenomenon often referred to as "hitting the wall". A crucial aspect of our analysis involves recognizing the distinctive patterns in these graphs, particularly when encountering a discernible crack or deviation from a steady Lyapunov exponent trend. These cracks in the Lyapunov exponent graph often coincide with points in the race where runners experience heightened fatigue and reduced control. It is during these moments that the chaotic nature of the race dynamics may intensify, leading to fluctuations in pacing and performance. Identifying and understanding these critical junctures, where runners may temporarily lose their grip on race control and endure the rest of the race, provides valuable insights into the challenges faced during marathon races. To accurately characterize the λ_1 plots for each runner, we decided to visually evaluate the percentage in the race from which λ_1 sharply and heavily drops, never to increase back up again, characterizing the "wall" and the irreversible fatigue of the runner. Table 4 regroups the values where these cracks appear in each Lyapunov graph for K5 and GARMIN.

Table 4. First appearance of Lyapunov 'crack' in the percentage of the race classed from the highest (A) to lowest performance (H).

N° Runner	K5	GARMIN
A	65%	61%
B	54%	20%
C	50%	38%
D	52%	12%
E	60%	21%
F	49%	32%
H	49%	13%

With a calculated U-value of 6 (compared to a critical value of 9) and a corresponding z-score of 2.29 (resulting in a *p*-value of 0.0214), our findings suggest a statistically significant difference in the occurrence of these cracks between the two datasets. This underscores the test's utility in quantifying disparities and highlights that fatigue patterns, as indicated by Lyapunov exponent graphs, diverge significantly between the K5 and GARMIN datasets. Such insights are pivotal for understanding the distinctive characteristics of race dynamics captured by different tracking technologies. If we can demonstrate that K5 effectively encapsulates and represents the nuanced race dynamics, it underscores the superior utility of K5 as a data source for understanding and analyzing marathon performances. This distinction emphasizes the importance of selecting data sources that robustly capture the intricacies of race dynamics and highlights the limitations that may be associated with alternative tracking technologies like GARMIN. Such insights enhance our ability to make informed choices when studying and interpreting marathon race data.

3.2.2. Alignment of Fatigue Indicators and Performance Degradation

Table 5 provides a side-by-side comparison of the appearance of Lyapunov cracks, the occurrence of RPE (Rating of Perceived Exertion) reaching 15 during the race, and the identification of significant speed drops at specific points in the marathon event. The table offers insights into potential correlations and patterns among these key race dynamics indicators, shedding light on their interplay during the race.

Table 5. Appearance of the first λ_1 'crack', RPE 15 and significant speed drop in distance traveled for each runner, classed from the highest (A) to lowest performance (H). Runner A is highlighted because the K5 acquisition stops at about 70% of the race, which is before the noticeable speed drop observed during the marathon.

Runner	First Major λ_1 Crack (km)	RPE 15 (km)	Major Speed Drop (km)
A	27	35	36
B	25.2	30	26.3
C	22.9	27	24
D	24	26	25
E	26.5	25	27
F	21.9	24	22.7
H	22.5	26	23

Table 6 provides a the R2 score of 0.975 indicates that about 97.5% of the variance in the major speed drop (km) is explained by the variance in the first major λ_1 crack (km) under this linear model. The *p*-value for slope is reported as 0.000, strongly suggesting a significant linear relationship between the two variables at the 95% confidence level and a slope spanning [0.76, 1.21]. Runner A was omitted because it behaved as an outlier relative to the other runners. Specifically, its values–27 km for the first major $\lambda_1 \backslash lambda\_1 \lambda_1$ crack and 36 km for the speed drop–lie substantially beyond the overall trend of the other data points. Including such an outlier can overly influence the regression line and artificially inflate or diminish the strength of the relationship among the rest of the observations.

Table 6. Parameter estimation of the linear model linking first major crack with major speed drop. the intercept is not significantly different from 0, meaning there's no strong evidence that the regression line crosses the y-axis at a non-zero value. the slope is very close to **1**, indicating that for each unit increase in the predictor, the outcome increases by roughly 0.98 on average.

	Coef	Std Err	t	$p > |t|$	[0.025	0.975]
intercept	1.1998	1.900	0.631	0.562	−4.076	6.476
slope	0.9846	0.08	12.377	0.000	0.764	1.206

Despite the acknowledged limitations in statistical power and generalizability, this preliminary dataset serves as a valuable proof-of-concept to illustrate the potential link between the timing of Lyapunov cracks and impending performance decrements. Moreover, small-scale pilot studies are a common, pragmatic step before launching larger-scale investigations. By examining a limited number of well-instrumented runners, we can test and refine data-collection protocols, ensure the feasibility of measuring relevant physiological indicators, and gather initial insights into underlying mechanisms.

As we can visualize in Figure 7, the analysis has unveiled a noteworthy linear correlation between the appearance of Lyapunov cracks and the occurrence of major speed drops during the marathon race. This finding suggests a clear relationship between the presence of chaotic behavior, as indicated by the Lyapunov cracks, and the moments when runners experience significant reductions in their running speed. Interestingly, our data further reveals that the Lyapunov crack tends to manifest approximately 1 km ahead of the observed speed drop and just before the point when the Rating of Perceived Exertion (RPE) reaches 15. This temporal alignment implies that the Lyapunov crack may serve as an early indicator of impending physical exertion and slowing pace, offering valuable insights into the dynamic interplay between runner behaviors, perceived effort, and performance fluctuations during the race.
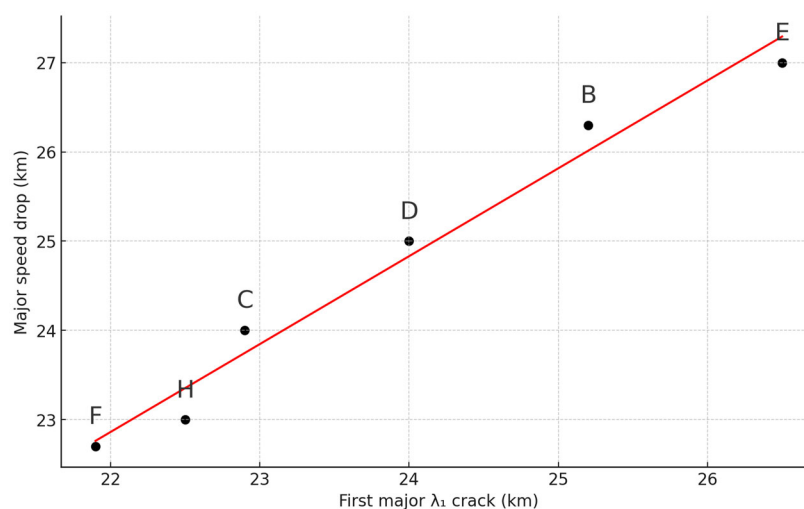

Figure 7. Scatter plot showing the first λ_1 crack compared to the speed drop observed for each runner (except Runner A whose K5 acquisition stops at about 70% of the race). Capital letters are put to refer to the runners. The coefficient of determination R2 = 0.975, indicating a strong linear relationship between these 2 variables.

4. Discussion

Marathon running, as explored in this study, presents a unique interplay between physiological limits, pacing strategies, and emerging AI technologies. Even if the IA community could highlights the importance of comparing our proposed VAE based deep neural network approach with another cutting-edge neural network research in a comparative context that could strengthen the relevance and impact of our findings as fault location in cloud data center interconnections, multi-fault location in 5G radio and optical wireless networks, or neural network mapping in optical Network-on-Chips, this approach primarily address optimization problems distinct from our focus on individualized athletic pacing strategies.

Indeed, the objective of our research is to specifically target physiological data-driven approaches, with a view to emphasizing personalized energy reserve management for the purpose of avoiding the drastic marathon speed that is known to appear in the so-

called "wall" close to the 30th km according to the individuals. Therefore, the primary objective was to address the critical challenge of balancing energy conservation with fatigue management, particularly for recreational runners who often struggle with rigid pacing strategies. By integrating advanced data analysis and AI tools, we aimed to uncover insights that could inform more adaptable pacing strategies, thereby enhancing performance and reducing the risk of energy deficits. Our study has demonstrated significant insights into the intricate dynamics of marathon performance, emphasizing the interplay between physiological regulation, pacing strategies, and emerging AI technologies. Through a combination of innovative analytical methods and advanced data acquisition techniques, we have provided a new perspective on how runners manage their pace and adapt to the physical demands of a marathon. The findings open avenues for further exploration and practical applications.

Research has demonstrated that muscular power output is regulated in an anticipatory manner to prevent uncontrolled disruptions in physiological homeostasis [27]. Pacing has a significant impact on energy production from both aerobic and anaerobic energy systems. The goal of the pacing strategy is to optimize these energy systems accordingly. Although the effects of various physiological regulators overlap, the conscious brain integrates their net input using the rating of perceived exertion (RPE) [12,28–30].

Changes in homeostatic status, reflected by momentary RPE, allow for alteration of pacing strategy (power output) in both an anticipatory and responsive manner based on pre-exercise expectations and peripheral feedback from different physiological sensors [31,32]. Recent studies have examined the continuous physiological response and RPE during marathons, revealing a similar decrease in the ratio between RPE and speed, heart rate, and VO_2 for all recreational runners [13].

Our current understanding of how runners adapt their marathon pace to account for various cardiorespiratory and biomechanical factors remains incomplete. One hypothesis proposes a strong connection between the rating of perceived exertion (RPE) and a physiological and mechanical message that is crucial for maintaining accurate adjustments in running speed. More specifically, it is essential to achieve a balance between stride amplitude and stride frequency, in the same way that a cyclist needs to dose his equipment. Furthermore, the signal must possess a sufficient level of uncertainty to effectively convey information.

In this comprehensive study, we have delved into the intricate dynamics of marathon races, employing novel analytical approaches to gain deeper insights into runner behaviors and performance. Our findings have shed light on several key facets of marathon race dynamics, prompting valuable discussions and considerations for future research and applications. One noteworthy revelation from our analysis is the contrast between the K5 and GARMIN datasets in terms of their ability to accurately describe race dynamics and predict fatigue. K5 emerges as a robust data source, effectively encapsulating the nuances of marathon race behaviors. However, this effectiveness comes at the cost of invasiveness, raising concerns regarding the democratization of the method. While K5 offers rich insights, the accessibility of such technology remains a challenge for widespread application. Addressing this limitation warrants exploration of alternative, less invasive data sources that still capture race dynamics effectively.

To ensure the coherence of race signatures with the continuity criteria, adopting mathematical models becomes imperative. Our study acknowledges the importance of modeling in assessing the fidelity of the observed signatures in adhering to the principles of continuity. Future research endeavors should consider incorporating mathematical frameworks that quantify the continuity of race behaviors, offering a more rigorous and objective assessment of the dynamic race strategies observed. In our pursuit of a deeper understanding of

marathon race dynamics, the utilization of a temporal Variational Autoencoder (t-VAE) proved instrumental. Unlike classic VAEs, t-VAEs are tailored to better consider the temporality of the race, allowing for more nuanced insights into the evolving behaviors of runners. This innovative approach offers promising avenues for future investigations into the dynamic interplay between various factors influencing marathon performance [33–36].

Our analysis unveils a compelling relationship between the appearance of Lyapunov cracks applied on the new variables synthetized with the VAE process. and critical moments in the marathon race, such as significant speed drops and the RPE reaching 15. This temporal alignment suggests the potential for a proactive race strategy to delay the Lyapunov crack. Runners and coaches can explore adaptive pacing techniques and interventions aimed at optimizing performance and mitigating the impact of fatigue-induced fluctuations. This strategic adaptation could lead to enhanced race outcomes and improved performance control thanks to real time race monitoring from the future physiological sensors relied on the phone [37].

5. Conclusions

In conclusion, our study demonstrates the intricate and multifaceted nature of marathon race dynamics, as revealed through advanced data analysis techniques. While K5 is an effective method for identifying race behaviours, its intrusive nature presents challenges in terms of its applicability to a wider audience. The necessity for mathematical modelling to evaluate coherence with continuity criteria, strategies to delay the Lyapunov crack, and the utilisation of t-VAEs for temporal considerations collectively contribute to a rapidly expanding field of research aimed at optimising marathon performance and advancing our comprehension of the dynamics inherent to long-distance running.

In conclusion, this study addresses the following key questions:

1. It is therefore pertinent to enquire whether an AI system utilising a variable automatic encoder and learned energetic signatures could assist runners in more effectively managing their pace.
2. Does the use of AI-assisted pacing reduce the probability of experiencing the phenomenon known as the "marathon wall"?
3. It would be beneficial to ascertain whether AI pacing can enhance performance outcomes and mitigate health risks, particularly in the context of increasingly hot conditions.

In response to these three questions, it can be stated that the mathematical model is not yet sufficiently developed to allow for the accurate calculation of time. Furthermore, it would be prudent to ascertain the suitability of this approach before advocating it to runners who may be reluctant to forego the use of cardio-GPS monitoring during their races. It would be more beneficial to encourage runners to trust in their sensations, as measured by the Borg scale, rather than relying on speed in km/h or pace in miles per kilometer. Furthermore, it was observed that the cardio-GPS provides only a partial representation of the physiological response when compared to the full metabolic response.

Of course, the limited sample size of nine runners, which indeed restricts the generalizability of our findings. However, it is important to highlight that obtaining comprehensive cardiorespiratory data during an official marathon event is exceptionally challenging. This difficulty arises due to the high sensitivity of metabolic analyzers, constraints posed by marathon organization protocols, and official rules that significantly restrict equipment usage and data collection processes. Consequently, expanding the sample size under official competition conditions is often logistically complex and technically constrained. Nevertheless, future research efforts will aim to overcome these barriers and incorporate a larger, more diverse sample to enhance the applicability of the results.

References

1. Foster, C.; Schrager, M.; Snyder, A.C.; Thompson, N.N. Pacing strategy and athletic performance. *Sports Med.* **1994**, *17*, 77–85. [CrossRef]
2. Joyner, J.R.; Hunter, S.K.; Lucia, A.; Jones, A.M. Physiology and fast marathons. *J. Appl. Physiol.* **2020**, *128*, 1065–1068. [CrossRef]
3. Midgley, A.W.; McNaughton, L.R.; Wilkinson, M. Is there an optimal training intensity for enhancing the maximal oxygen uptake of distance runners? *Sports Med.* **2006**, *36*, 117–132. [CrossRef]
4. Joyner, M.J.; Coyle, E.F. Endurance exercise performance: The physiology of champions. *J. Physiol.* **2008**, *586*, 35–44. [CrossRef]
5. Molinari, C.A.; Edwards, J.; Billat, V. Maximal time spent at VO_2max from sprint to the marathon. *Int. J. Environ. Res. Public Health* **2020**, *17*, 9250. [CrossRef]
6. Skorski, S.; Abbiss, C.R. The manipulation of pace within endurance sport. *Front. Physiol.* **2017**, *8*, 102. [CrossRef]
7. Hanley, B. Pacing profiles and pack running at the IAAF World Half Marathon Championships. *J. Sports Sci.* **2016**, *34*, 1653–1659. [CrossRef]
8. Périard, J.D.; Eijsvogels, T.M.H.; Daanen, H.A.M. Exercise under heat stress: Thermoregulation, hydration, performance implications, and mitigation strategies. *Physiol. Rev.* **2021**, *101*, 1873–1979. [CrossRef]
9. Schmitt, S.; Haken, H.; Balagué, N.; Hristovski, R.; Davids, K.; Newell, K.M.; Seifert, L.; Button, C.; Renson, R.; McGarry, T.; et al. Complexity and nonlinear tools in sport and exercise physiology: A narrative review of applications and practical recommendations. *Eur. J. Appl. Physiol.* **2022**, *122*, 1067–1081.
10. Cheung, S.S.; Sleivert, G.G. Multiple triggers for hyperthermic fatigue and exhaustion. *Exerc. Sport Sci. Rev.* **2004**, *32*, 100–106. [CrossRef]
11. González-Alonso, J.; Teller, C.; Andersen, S.L.; Jensen, F.B.; Hyldig, T.; Nielsen, B. Influence of body temperature on the development of fatigue during prolonged exercise in the heat. *J. Appl. Physiol.* **1999**, *86*, 1032–1039. [CrossRef]
12. Borg, G. *Borg's Perceived Exertion and Pain Scales*; Human Kinetics: Champaign, IL, USA, 1998; 104p.
13. Christian, R.J.; Bishop, D.J.; Billaut, F.; Girard, O. The role of sense of effort on self-selected cycling power output. *Front. Physiol.* **2014**, *5*, 115. [CrossRef]
14. Altini, M.; Casale, P.; Penders, J.; Amft, O. Personalized cardiorespiratory fitness and energy expenditure estimation using hierarchical bayesian models. *J. Biomed. Inform.* **2015**, *56*, 195–204. [CrossRef]
15. Peng, C.K.; Buldyrev, S.V.; Havlin, S.; Simons, M.; Stanley, H.E.; Goldberger, A.L. Mosaic organization of DNA nucleotides. *Phys. Rev. E* **1994**, *49*, 1685–1689. [CrossRef]
16. Megro, F. Une Analyse Dimensionnelle de L'équilibre Précaire Chez L'homme. Ph.D. Thesis, Paris-Sud University, Paris, France, 2002.
17. Tous Les Chiffres du Marathon de Paris 2015. Lepape-Info. Available online: https://www.lepape-info.com/actualite/tous-les-chiffres-du-marathon-de-paris-2015/ (accessed on 15 February 2024).
18. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications with R Examples*, 4th ed.; Springer: Berlin/Heidelberg, Germany, 2017.
19. Le Cun, Y. *Quand la Machine Apprend: La Révolution des Neurones Artificiels et de L'apprentissage Profond*; Odile Jacob: Paris, France, 2023.
20. Diaz, M.; Charbonnel, P.E.; Chamoin, L. A new kalman filter approach for structural parameter tracking: Application to the monitoring of damaging structures tested on shaking-tables. *Mech. Syst. Signal Process.* **2022**, *182*, 109529. [CrossRef]
21. Wattenber, M.; Viégas, F.; Johnson, I. How to use t-SNE effectively. *Distill* **2016**, *1*. [CrossRef]
22. Bahamonde, A.; Montes, R.; Cornejo, P. Usefulness and limitations of convergent cross sorting and continuity scaling methods for their application in simulated and real-world time series. *R. Soc. Publ.* **2023**, *10*, 221590. [CrossRef]

23. Cao, Y.; Tung, W.; Gao, J.; Protopopescu, V.; Hively, L. Detecting dynamical changes in time series using the permutation entropy. *Phys. Rev. E* **2004**, *70*, 046217. [CrossRef]

24. Cervantes de la Torre, F.; Gonzales, J.; Real-Ramirez, C.; Hoyos-Reyes, L. Fractal dimension algorithms and their application to time series associated with natural phenomena. *J. Phys. Conf. Ser.* **2013**, *475*, 012002. [CrossRef]

25. Van Opstall, M. Quantifying chaos in dynamical systems with lyapunov exponents. *Furman Univ. Electron. J. Undergrad. Math.* **1998**, *4*, 1–8.

26. Filali, H.; Chouchane, K. *Sur un Algorithme Pour le Calcul des Exposants de Lyapunov d'un Système Dynamique*; Institute of Sciences and Technology, University Center Abdelhafid Boussouf: Mila, Algeria, 2020.

27. De Koning, J.J.; Foster, C.; Bakkum, A.; Kloppenburg, S.; Thiel, C. Regulation of pacing strategy during athletic competition. *PLoS ONE* **2011**, *6*, e15863. [CrossRef]

28. Ulmer, H.V. Concept of an extracellular regulation of muscular metabolic rate during heavy exercise in humans by psychophysiological feedback. *Experientia* **1996**, *52*, 416–420. [CrossRef]

29. Lambert, E.V.; St Clair Gibson, A.; Noakes, T.D. Complex system model of fatigue: Integrative homeostatic control of peripheral systems during exercise in humans. *Br. J. Sports Med.* **2005**, *39*, 52–62. [CrossRef]

30. St Clair Gibson, A.; Lambert, E.V.; Rauch, L.H.G.; Tucker, R.; Baden, D.A. The role of information processing between the brain and peripheral physiological systems in pacing and perception of effort. *Sports Med.* **2006**, *36*, 705–722. [CrossRef]

31. Tucker, R.; Noakes, T.D. The physiological regulation of pacing strategy during exercise: A critical review. *Br. J. Sports Med.* **2009**, *43*, e1. [CrossRef]

32. Tucker, R. The anticipatory regulation of performance: The physiological basis for pacing strategies and the development of a perception-based model for exercise performance. *Br. J. Sports Med.* **2009**, *43*, 392–400. [CrossRef]

33. Zhang, L.; Chen, H.; Wang, Q. Accurate Fault Location Using Deep Neural Evolution Network in Cloud Data Center Interconnection. *IEEE Trans. Cloud Comput.* **2022**, *10*, 335–348.

34. Liu, Q.; Zhao, J.; Feng, Y. Efficient Hybrid Multi-Fault Location Based on Hopfield Neural Network in 5G Coexisting Radio and Optical Wireless Networks. *IEEE Trans. Wirel. Commun.* **2023**, *22*, 1964–1978.

35. Li, Y.; Song, M.; Wu, X. Efficient O-Type Mapping and Routing of Large-Scale Neural Networks to Torus-Based ONoCs. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **2021**, *40*, 4233–4246.

36. Liu, F.; Guo, X. A Multi-Objective Evolutionary Algorithm for Neural Network Architecture Search in Fault Detection. *Neurocomputing* **2022**, *488*, 265–275.

37. Wang, Z.; Li, P.; Du, P. Adaptive Resource Allocation for Hybrid 5G-Optical Networks Using Reinforcement Learning. *Opt. Express* **2023**, *31*, 16412–16425.