

Perturbation recovery time as a valid measure of balance control in human running

Carlo von Diecken

Institute of Biomechanics and Orthopaedics, German Sport University Cologne, Cologne, Germany

<https://orcid.org/0009-0006-3286-3806>

Luca Braun

luca.braun@hs-offenburg.de

Institute of Biomechanics and Orthopaedics, German Sport University Cologne, Cologne, Germany

<https://orcid.org/0000-0001-5692-3785>

Bastian Anedda

Institute of Biomechanics and Orthopaedics, German Sport University Cologne, Cologne, Germany

<https://orcid.org/0000-0001-6102-4629>

Steffen Willwacher

Institute for Advanced Biomechanics and Motion Studies, Offenburg University of Applied Sciences, Offenburg, Germany <https://orcid.org/0000-0002-1303-3165>

Research Article

Keywords: Human locomotion, perturbation recovery, whole-body angular momentum, balance metrics

Posted Date: March 24th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9201516/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Perturbation recovery time as a valid measure of balance control in human running

Carlo von Diecken¹, Luca Braun^{1*}, Bastian Anedda¹, Steffen Willwacher²

Affiliations

- 1 Institute of Biomechanics and Orthopaedics, German Sport University Cologne, Cologne, Germany
- 2 Institute for Advanced Biomechanics and Motion Studies, Offenburg University of Applied Sciences, Offenburg, Germany

*Corresponding Author: Luca Braun

Department of Mechanical and Process Engineering, Offenburg University

Max-Planck Straße 1, 77652 Offenburg

Mail address: luca.braun@hs-offenburg.de

Tel.: +49 781 205-1135

Abstract

Balance control during running may be critical for performance and injury prevention, but the lack of valid and interpretable metrics limits research. This study introduces perturbation recovery time (PRT) as a novel measure of running-specific balance control, defined as the time required for whole-body angular momentum (WBAM) to return to statistically defined steady-state dynamics following a perturbation. Thirty recreational runners completed treadmill trials under normal running and two artificial balance impairment conditions (restricted arm swing and weighted vest loading). Participants were exposed to pseudorandomized anterior–posterior and mediolateral perturbations of varying magnitudes. PRT was computed using a functional data analysis approach based on 95% prediction bands derived from baseline WBAM dynamics. Peak deviation (PD), peak-to-peak range (PPR) and maximum Lyapunov exponent (MLE) were calculated for comparison. Detection accuracy was defined as the proportion of runners showing larger values in impaired vs. normal conditions and in higher vs. lower perturbation intensities. PRT demonstrated the highest overall detection accuracy ($80.8 \pm 10.5\%$) compared to PPR ($75.0 \pm 6.0\%$), PD ($73.0 \pm 11.8\%$) and MLE ($34.0 \pm 28.1\%$). PRT detected balance impairment induced by restricted arm swing in 82.8% and by weighted vest in 66.7% of runners, respectively. It also showed high sensitivity to perturbation intensity (86.8%). In contrast, deviation-based WBAM metrics failed to detect impairment in the weighted vest condition and MLE showed poor discriminatory ability in all comparisons. Notably, MLE was negatively associated with PRT, suggesting that high local dynamic stability predicts reduced ability to recover from large perturbations. These findings support PRT as a valid and interpretable measure of balance control in running, potentially capturing aspects of recovery dynamics not captured by existing metrics. PRT may provide a useful tool for assessing balance control in athletic and clinical populations, although further work is needed to establish its relationship with performance and injury risk.

Introduction

When running in real-world environments, balance control is challenged by internal and external perturbations such as sensorimotor or muscle force noise (Harris & Wolpert, 1998; Osborne et al., 2005), ground level changes (AminiAghdam et al., 2020; Ernst et al., 2019; Grimmer et al., 2008) and surface feature changes (e.g., more compliant/slippery than expected, Dufek et al., 2009). During off-road running events such as trail running races, runners are exposed to even larger perturbations as they navigate uphill and downhill slopes that feature uneven ground and discrete obstacles (e.g., rocks, roots etc.; Scheer et al., 2020). Similarly, players of contact-oriented team sports (e.g., soccer, rugby, American football etc.) must tolerate unexpected collisions (Naughton et al., 2018) or quickly alter direction in response to opponents while running at high speeds (Loturco et al., 2022; Young et al., 2015). When encountering any of these perturbations, a rapid and effective neuromuscular corrective response is required to re-enter stable equilibrium and avoid falling (Seethapathi & Srinivasan, 2019).

Successful recovery following a perturbation depends on the relationship between perturbation magnitude and an individual's intrinsic balance capacity (Wu et al., 2025). While perturbation characteristics (e.g., type, amplitude and timing within the gait cycle) determine the mechanical demand of a successful recovery, the intrinsic balance ability determines a person's ability to save themselves from a fall given a certain perturbation. Thus, a runner with high intrinsic balance capacity exposed to large external perturbations (e.g., fast downhill running on uneven terrain) may exhibit a similar likelihood of falling as an individual with lower balance capacity under less demanding conditions. Consequently, a quantitative measure of intrinsic balance ability during running could help identify individuals whose balance capacity is insufficient relative to the perturbation demands they encounter, enabling targeted interventions to reduce fall risk. There further have been research efforts to evaluate the effects of technologies such as footwear (Ekizos et al., 2017; Frank et al., 2019a; Hollander et al., 2021; Kettner et al., 2024; Von Diecken, Riedl, & Ueberschär, 2025) and prostheses (Arellano et al., 2015; Look et al., 2013; Sepp et al., 2019) on stability and balance control but research is currently limited by the lack of validated and interpretable balance metrics (Bruijn et al., 2013).

Currently, the most widely used stability metric in running is the maximum short-time Lyapunov exponent (MLE; Bruijn et al., 2013; Winter et al., 2023) which measures the system's robustness against small continuous perturbations or—practically—its local dynamic stability (LDS). Despite the measure's widespread use in running biomechanics as a stability metric, its practical relevance is currently unclear. There is some early research indicating that the measure may be related to running economy (Von Diecken, Riedl, Willwacher, et al., 2025) or may be useful as a screening tool for overuse injuries (Schütte et al., 2018), but these studies are far from conclusive. While in walking biomechanics, research on passive walker models has shown that MLE is related to global system stability (i.e., likelihood of falling given a certain perturbation; Bruijn et al., 2012), such a connection has not yet been established in running. Determining the relationship between LDS and global stability metrics would improve the interpretability of the MLE and help determine its relevance for fall and injury prevention.

During legged locomotion, balance maintenance requires tight regulation of whole-body angular momentum (WBAM) in sagittal and frontal movement planes (Herr & Popovic, 2008; Hinrichs, 1987; Popovic et al., 2004; Thielemans et al., 2014; Van Dieën et al., 2025a). This has motivated researchers to utilize WBAM control as well as discrete variables derived from WBAM, such as variation compared to baseline dynamics, peak WBAM over a stride, peak-to-peak range and integrated WBAM to characterize balance control and recovery in walking

(Adam et al., 2025; Golyski et al., 2025; Martelli et al., 2013; Shokouhi et al., 2024; Yamagata et al., 2024; Yamaguchi & Masani, 2022) and running (Sepp et al., 2019). While these WBAM-derived quantities show promise as measures of balance control—being sensitive to balance training (Adam et al., 2025) and changing with age (Yamaguchi & Masani, 2022)—they are largely unitless, depend on assessed movement plane and are generally very small in magnitude. These characteristics limit interpretability as well as comparability across studies. Recently, Wu et al. (2025) addressed these limitations by developing a single quantitative measure of intrinsic balance ability in walking, which they termed “perturbation recovery time” (PRT). The researchers define PRT as the time required after a perturbation for the system trajectory to return to steady-state conditions and demonstrate that this measure is well-suited to detect artificial balance impairments in healthy young adults. However, they employed a computationally expensive optimization approach to determine proper parameters and thresholds for the definition of steady-state dynamics and recovery, making the method difficult to replicate and adopt for other contexts such as running.

In this study, we adapt the concept of PRT to quantify running-specific balance control in healthy adult runners using standardized treadmill-based perturbations. Rather than employing an optimization framework, we apply a novel statistical approach based on functional data analysis to determine steady-state WBAM dynamics based on 95% prediction bands. We define PRT as the time until the final statistically significant deviation in WBAM across sagittal and frontal movement planes following a perturbation. We then evaluate the sensitivity of this measure to detect perturbations of different amplitudes and directions, as well as its ability to detect reductions in balance ability induced by two artificial balance impairments. Further, we compare PRT against other balance metrics derived from WBAM kinematics as well as the MLE with respect to detection of artificial balance control impairment. Finally, we determine the relationship between PRT and MLE to investigate whether LDS can predict global balance ability.

Methods

Participants and ethical declaration

Thirty recreational runners (11 females and 19 males, age = 38.9 ± 13.3 years, body mass = 71.9 ± 12.3 kg, $\text{VO}_{2\text{max}} = 52.7 \pm 7.9$ ml·kg⁻¹·min⁻¹) participated in this study. The study protocol was approved by the Institutional Review Board of Offenburg University of Applied Sciences (approval code: 06/23). All participants provided written informed consent before participation and the study was conducted in accordance with the Declaration of Helsinki.

Experimental protocol

During all trials, runners were equipped with a safety harness tethered to the ceiling to prevent a fall in the event that a runner would lose their balance. Participants first completed a standardized warm-up consisting of 5 min of steady treadmill running at 2.78 m/s (6:00 min/km). This was followed by three baseline trials of 5 min steady treadmill running at 2.78 m/s, consisting of normal running, running with restricted arm swing and running with 10% of body mass added to the trunk using a weighted vest (for further details see section “conditions”). Next, participants completed a 2 min familiarization protocol consisting of 14 anterior-posterior (AP) perturbations of increasing intensity with two perturbations per amplitude, followed by a single mediolateral (ML) perturbation (for further details see section “perturbations”). During this familiarization trial, perturbations were administered at a predictable 10 stride-interval and steadily increased in amplitude from 1.5 m/s during the first perturbation to 2.7 m/s during the

final perturbation. Pilot testing had indicated that the most severe perturbations (i.e., amplitude = 2.7 m/s) approached the limit of comfort and recoverability without falling for some runners. For this reason, we implemented the familiarization trial both to familiarize participants with the perturbations but also to ensure they could safely handle the full range of perturbation amplitudes in the subsequent trials. A total of six participants either did not feel confident or were judged by the operator to be unable to safely cope with the highest perturbation intensity. In these instances, all perturbations of the highest intensity were reduced to the second highest intensity during the following perturbation trials (i.e., from 2.7 m/s to 2.3 m/s), resulting in the same number of perturbations but lower peak perturbation intensity. After familiarization, participants completed three 8 min trials of perturbed treadmill running consisting of normal running, running with restricted arm swing and running with 10% of body mass added at the trunk. Condition order during the perturbation trials was independently randomized from baseline condition order. Perturbation trials began with 2 min of steady running during which participants were aware that no perturbations would occur. This unperturbed period of running served as a short familiarization to ensure kinematic stability prior to the following perturbations (Paquette et al., 2024). During the following 6 minutes, 12 AP perturbations consisting of 4 perturbations at 1.5 m/s (AP1), 2.3 m/s (AP2) and 2.7 m/s (AP3) amplitude, respectively, as well as 4 ML perturbations were applied in pseudo-randomized order, resulting in a total of 16 perturbation per trial. Timing, amplitude and direction of perturbations were pre-determined and specified in three different perturbation protocols which were independently randomized across perturbation trials and spread perturbation intensity and direction in a pseudo-randomized manner across 500 strides. Perturbations were triggered based on counted gait events detected in real-time. Each perturbation protocol included one ~60 s period during which no perturbations were applied, and one instance where two AP perturbations occurred within <10 s of one another to minimize the predictability of perturbation timing. The time point of these features differed across perturbation protocols and thus could not be predicted by participants. Therefore, perturbations were generally expected by participants but unpredictable in their timing. Lastly, participants completed an incremental treadmill running protocol to assess $\text{VO}_{2\text{max}}$ with step duration of 1 min and speed increments of 1 km/h until voluntary exhaustion.

Measurements

All trials were continuously recorded using a marker-based motion capture system (100 Hz; 16 Vero cameras, Vicon, Oxford, UK) tracking 76 retroreflective markers placed according to a full-body marker set (Willwacher et al., 2016) as well as retroreflective marker patches attached to the treadmill belt to accurately measure treadmill belt velocity (Willwacher et al., 2021). Vertical ground reaction forces were continuously recorded using a force-instrumented treadmill (1000 Hz; GRAIL, Motek Medical, Amsterdam, The Netherlands). Inertial measurement units (IMUs; Ultium Motion, Noraxon Inc., Scottsdale, AZ, USA) were attached unilaterally at the foot, distal shank and sacrum of each participant, recording 3D angular velocities at 200 Hz. Additionally, participants were equipped with a face mask connected to a metabolic cart for recording of metabolic gas-exchange data as well as with 8 wireless EMG electrodes placed unilaterally on lower-limb musculature. Note that the unperturbed baseline trials, the $\text{VO}_{2\text{max}}$ test, the metabolic gas-exchange data and the EMG data were recorded as part of a larger research project and will not be analyzed in the present article.

Conditions

Runners completed the perturbation protocols in three different conditions which were randomized in order. Two of the conditions (restricted arm swing, RA; weighted vest, WV)

were designed to impair the runners' ability to regain balance following a perturbation, while the third condition (normal running, NR) served as a control condition of unimpaired balance. The concept of artificially impaired balance conditions to determine the effectiveness of a given balance metric is based on the work of Wu et al. (2025), who first applied this approach to determine the effectiveness of several steady-state and perturbation-recovery balance metrics in walking. We chose to restrict arm swing by having participants voluntarily hold their arms behind their backs as the first impairment condition. Arm swing in running serves to balance the angular momentum about the vertical axis by cancelling the angular momentum of the legs (Hinrichs, 1987) and provides important contributions to lateral stability (Arellano & Kram, 2011; Koo, 2025). Further, restricting arm swing has already been shown to increase perturbation recovery time in walking (Bruijn et al., 2010) and we expected a similar effect on running. The second impairment condition consisted of a weighted vest attached to the upper torso of the participants, adding 10% of the participants' body mass to the trunk, evenly split between the chest and back. Placing additional mass on the upper torso of the participants served to increase the body's moment of inertia, as the load was positioned above the COM. Pilot testing showed that the belt-based perturbations we applied during this study cause the body to start rotating forwards or sideways depending on perturbation direction and we reasoned that the same perturbation velocity would result in greater angular momentum, making it more challenging to counteract this rotation and thus increasing recovery difficulty.

Perturbations

To actively challenge participants' balance control, we applied pseudorandomized treadmill belt velocity perturbations of different directions and magnitudes that were unpredictable in their timing. These perturbations included AP perturbations that consisted of sudden belt accelerations as well as ML perturbations that consisted of lateral translations of the treadmill body. The AP perturbations were further scaled to three different velocity amplitudes (1.92 ± 0 m/s, 2.33 ± 0 m/s, 2.73 ± 0 m/s peak belt velocity above baseline running speed; Hipper et al., 2025; Fig. 1A,D), which were achieved by constant acceleration of 15 m/s^2 for varying durations (Fig. 1B,D). The ML perturbations were achieved through sudden lateral translation of the treadmill body by 0.1 m in ~ 800 ms, resulting in an average peak mediolateral velocity of 0.24 m/s. Due to technical limitations of only 0.1 m of lateral movement possible, we did not scale ML perturbations in their intensity. All perturbations were set to trigger at 30% of stance, which was achieved through real-time stride time calculation based on vertical ground reaction force data using a force threshold of 100 N. Due to inherent system delay, we predicted this timepoint and triggered the perturbation one stride in advance by continuously averaging stride time over 20 strides (Fig. 1D). Because of this prediction and the conservative force threshold—which was necessary due to noise in the vertical ground reaction force data—the onset timing was more variable than amplitude or duration (Fig. 1C). Still, we achieved an average onset timing of $30.7 \pm 3.4\%$ of stance and all perturbations triggered within the initial 50% of stance. All perturbation types were evenly split between left and right legs within each perturbation trial. To ensure that differences in perturbation parameters between conditions did not influence results, we tested differences in amplitude, duration and onset timing between conditions. While condition did not affect amplitude or duration, onset occurred significantly earlier in the WV condition than in NR (1.35%, $p < 0.001$) and RA (1.92%, $p < 0.001$). Notably, differences of 1.35% and 1.92% only correspond to an average timing difference of 9.8 and 13.9 ms, respectively (based on an average stride time of 722.0 ms). The discovery of earlier perturbation onset in the WV condition led us to test the systematic effect of body mass and, indeed, onset occurred significantly earlier with increasing body mass ($\beta = -0.14 \text{ \%stance per kg}$, $p = 0.003$). In absolute terms, this means that onset occurred 10.11 ms earlier for every 10 kg of body mass increase.

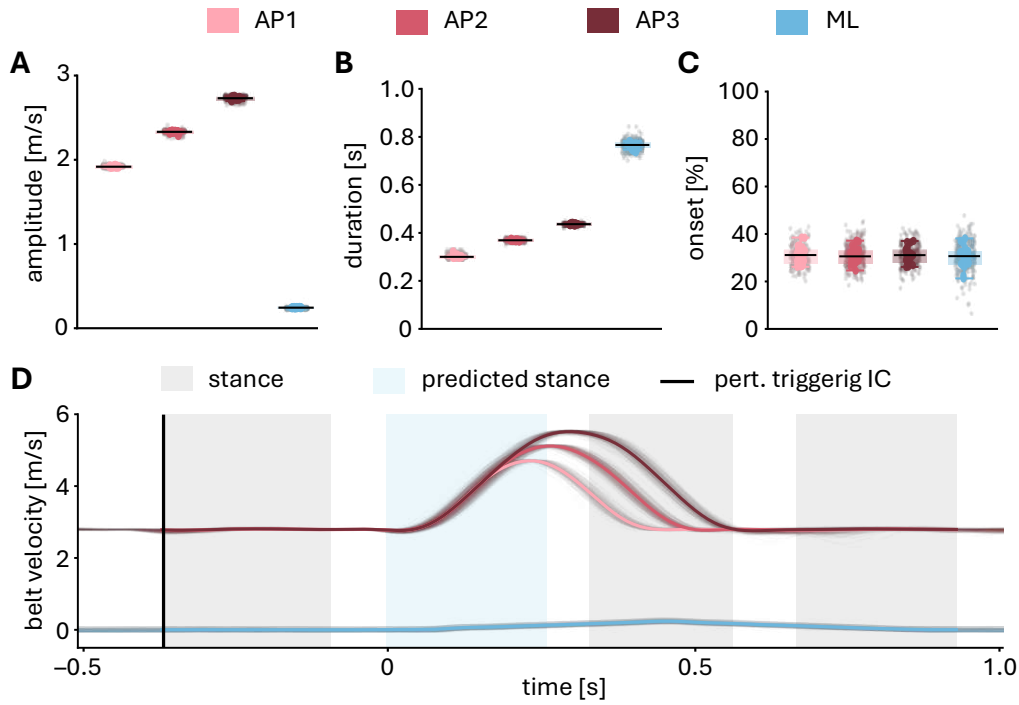


Fig. 1. Perturbation characteristics and timing. (A) Peak belt velocity, (B) Perturbation duration and (C) perturbation onset for AP perturbations across the three intensity levels (AP1–AP3) as well as ML perturbations (ML). (D) Belt velocity over time of AP and ML perturbations, aligned to IC preceding perturbation. Shaded regions indicate stance phase, with perturbations targeted at 30% of stance. Colors indicate perturbation types (AP intensities and ML). In (A–C), solid circles represent participant means and transparent circles represent individual perturbations. In (D), solid lines represent the overall average and transparent lines represent individual perturbations.

Data analysis

Kinematics

For each condition and participant, marker trajectories and ground reaction force data were exported for processing in OpenSim (Delp et al., 2007; Seth et al., 2018) (v4.5) and low-pass filtered using a recursive, fourth-order digital Butterworth filter with a matched cut-off frequency of 20 Hz (Mai & Willwacher, 2019). For the present study, a three-dimensional full-body musculoskeletal model derived from the original Gait2392 model (Braun et al., 2026; Delp et al., 1990; Hamner et al., 2010) was employed, comprising 27 degrees of freedom (DOF). These included six DOFs at the ground–pelvis joint, three DOFs at each hip joint, one DOF at each knee joint, one DOF at each ankle joint, three DOFs at each shoulder joint, one DOF at each elbow joint and three DOFs at the lumbar joint, with each DOF represented by a generalized coordinate. Owing to the large relative motion between the shoe and the foot during running, the metatarsophalangeal joint DOF was locked and therefore not tracked in the subsequent analyses. For all 30 participants, the generic model was scaled to individual anthropometry using the OpenSim Scale Tool by minimizing marker error between the distances of the experimental markers and the corresponding virtual markers on the generic model. For all trials of the WV condition, mass of the trunk segment was increased by 10%. Joint kinematics for each dynamic trial were subsequently computed using inverse kinematics, which minimized the weighted sum of squared errors between experimental and virtual marker positions. Based on segmental kinematics, whole-body angular momentum (WBAM) $\vec{H}$ about

the sagittal and frontal axes through the center of mass (COM) were computed as the sum of individual segment angular momenta according to

$$\vec{H} = \frac{\sum_{i=1}^n [(\vec{r}_i^{\text{COM}} - \vec{r}_{\text{body}}^{\text{COM}}) \times m_i (\vec{v}_i^{\text{COM}} - \vec{v}_{\text{body}}^{\text{COM}}) + I_i \vec{\omega}_i]}{MVH} \quad (1)$$

where n is the total number of body segments in the musculoskeletal model, $\vec{r}_i^{\text{COM}}$, $\vec{v}_i^{\text{COM}}$ and $\vec{\omega}_i$ are the position, velocity and angular velocity vectors of the COM of segment i , respectively, $\vec{r}_{\text{body}}^{\text{COM}}$ and $\vec{v}_{\text{body}}^{\text{COM}}$ are the position and velocity vectors of the whole-body COM, respectively and m_i and I_i denote the mass and inertia tensor of segment i , respectively. To minimize variance across participants, absolute WBAM in both axes was normalized by participants' mass (M), running speed (V ; 2.78 m/s) and height (H), resulting in one normalized dimensionless time-series vector per axis.

Maximum Lyapunov exponent

The MLE was calculated from a high-dimensional state space derived from three IMUs placed unilaterally on the foot, shank and the pelvis. From each IMU, angular velocity signals along the three orthogonal axes were extracted, resulting in a 9-dimensional signal vector at each time point. For each perturbation trial, all 500 strides of the perturbation protocol were extracted and time-normalized to 50,000 samples (500 strides $\times$ 100 samples per stride), exceeding the recommended minimum of 150 strides for stable MLE estimation (Brujin et al., 2009). A 9-dimensional state-space vector $s(t)$ was then reconstructed from the concatenated gyroscope signals using time-delay embedding as follows:

$$s(t) = [x(t), x(t + \tau), x(t + 2\tau) \dots, x(t + (d_E - 1)\tau)]$$

where $x(t)$ represents the 9-dimensional vector of time-normalized gyroscope signals, τ denotes the time delay, and d_E represents the embedding dimension. The time delay τ was determined from the first minimum of the average mutual information function, while the embedding dimension d_E was identified using the global false nearest neighbors approach (Winter et al., 2023). Mean integer values of these parameters across participants and conditions were used for the final embedding ($\tau = 9$, $d_E = 4$).

Divergence of neighboring trajectories was computed using the Rosenstein algorithm (Rosenstein et al., 1993). For each state vector, the nearest spatial neighbor that was not temporally adjacent was identified using a Theiler window of one stride to prevent autocorrelation. The logarithmic separation of neighboring trajectories was then tracked over time, and the short-term divergence rate was calculated as the slope of the divergence curve over the first 25% of a stride (Fohrmann et al., 2022). Larger MLE values indicate faster divergence of trajectories indicating lower LDS. In contrast to approaches that compute separate Lyapunov exponents for individual body segments, this framework estimates dynamic stability from a single high-dimensional representation of locomotor dynamics, thereby preserving the coupled behavior of the locomotor system (von Diecken et al., 2025).

Perturbation recovery time

Perturbation recovery time (PRT) was quantified based on deviations from baseline WBAM kinematics in frontal and sagittal axes. For this, initial ground contact events (IC) of left and right legs were determined based on vertical ground reaction forces (threshold = 50 N) and the vertical difference between left and right toe markers during the final six minutes of all

perturbation trials. WBAM time series were segmented into strides using consecutive unilateral ICs and time-normalized to 201 data points per stride, resulting in two time-normalized WBAM arrays per trial and axis (one from left IC to left IC and one from right IC to right IC). Perturbation onsets were detected based on optically tracked treadmill belt velocity as the time point at which belt velocity exceeded the mean ± 2 standard deviations of steady-state belt velocity (Hipper et al., 2025) and classified as left or right perturbation based on the stance leg at perturbation onset. Time-normalized WBAM time series were then separated into an unperturbed baseline array and a perturbation array. The baseline array included all strides from 2 min onwards that were at least 10 strides removed from a perturbation event (i.e., we excluded 10 strides prior to and 10 strides following each perturbation) and were used to define a participant-specific reference of unperturbed WBAM dynamics. For the perturbation array, time-normalized WBAM segments were extracted spanning the stride during which perturbation onset occurred as well as the six following strides. Extracted segments were side-specific, i.e., if the perturbation occurred during left stance phase, time-normalized strides from left IC to left IC were extracted and vice versa. This ensured consistency in the normalized timepoint of perturbation onset (i.e., % stride) in the perturbation array, regardless of perturbation side, and thus enabled comparability of perturbation recovery across sides. Perturbations that occurred within < 10 s of one another were removed from the perturbation array to ensure full recovery had been achieved prior to all included perturbations.

Based on the time-normalized baseline WBAM curves of each axis, we constructed 95% prediction bands around each WBAM axis for each participant. To achieve this, we fit a concurrent function-on-function regression model to each WBAM axis in the form of

$$Y_i(t) = \mu(t) + \varepsilon_i(t), \quad t \in [0,100] \quad (2)$$

with $Y_i(t)$ as the i -th observed functional curve, $\mu(t)$ as the population mean functional curve (Fig. 2A) and $\varepsilon_i(t)$ as the random error process capturing deviation of the i -th curve from the population mean (Fig. 2B). We then constructed 95% fast and fair simultaneous prediction bands (PB95; Fig. 2E, grey shaded area; Creutzinger et al., 2026) around the mean curve $\mu(t)$ in the form of

$$PB_{95}(t) = [\mu(t) \pm u(t)\sqrt{\hat{c}_{PB}(t,t)}] \quad (3)$$

with $u(t)$ as the time-adaptive critical value function (Fig. 1D, red curve) and $\sqrt{\hat{c}_{PB}(t,t)}$ as the pointwise prediction standard deviation (Fig. 1D, blue curve).

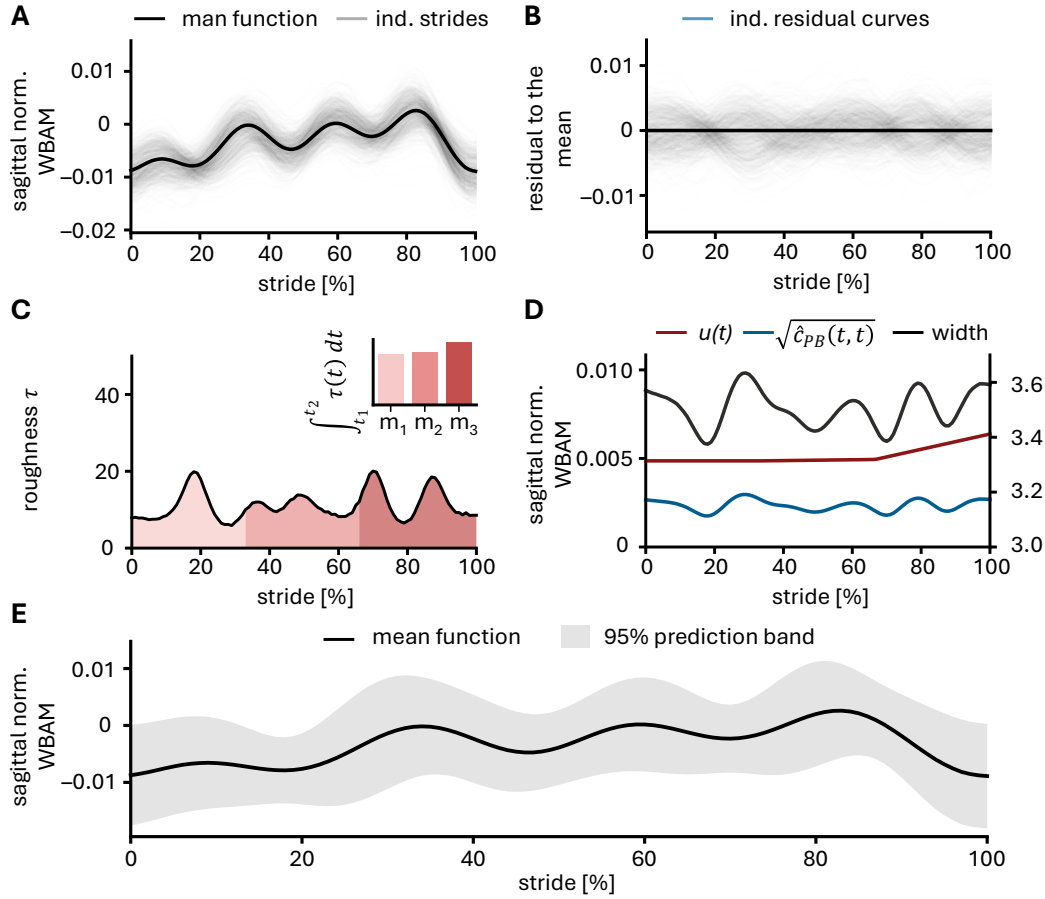


Fig. 2. Construction of participant-specific prediction bands. (A) Mean WBAM function $\mu(t)$ estimated from time-normalized baseline strides across participants, with individual strides shown in grey. (B) Residual functions $\varepsilon_i(t)$, representing deviations of individual baseline strides from the mean function. (C) Estimated roughness function $\tau(t)$, quantifying local oscillatory behavior of the residual process and informing the expected frequency of threshold crossings. (D) Components of the prediction band construction, including the time-adaptive critical value function $u(t)$ (red), the pointwise prediction standard deviation $\sqrt{\hat{c}_{PB}(t, t)}$ (blue) and their product defining the half-width of the prediction band (black). (E) Final 95% simultaneous prediction band (grey shaded area) around the mean function $\mu(t)$ (black), representing the range within which WBAM trajectories from an additional unperturbed stride are expected to lie with 95% probability.

The critical value function $u(t)$ is piecewise linear across $m = 3$ intervals such that the overall significance level α is allocated proportionally to the length of each interval, ensuring a fair distribution of type I error across the entire stride. The magnitude of $u(t)$ adapts locally according to the integral of the estimated roughness of the residual process $\tau(t)$ (Fig. 1C), which quantifies the local oscillatory behavior of the residual process and determines how frequently the curve is expected to cross a given threshold. Conceptually, $\sqrt{\hat{c}_{PB}(t, t)}$ determines relative band width across the stride cycle based on pointwise residual standard deviation and mean estimation error, while $u(t)$ scales this relative band width to achieve the desired α -level during each subinterval of the stride cycle. The half band width (Fig. 2D, black curve) is then the product of $\sqrt{\hat{c}_{PB}(t, t)}$ and $u(t)$. Since we constructed two prediction bands (one per WBAM axis), we applied a Bonferroni correction to the α -level of each band to preserve 95% overall confidence, resulting in an adjusted α -level of 0.025 per band. The resulting prediction band (Fig. 2E, shaded area) defines the range within which WBAM trajectories from an additional

stride of the same runner are expected to lie with 95% probability. Any portion of a trajectory falling outside this band at any time point is therefore considered a statistically significant deviation from steady-state behavior. The code to create fair and fast simultaneous prediction bands can be found at <https://github.com/creutzml/FunctionalPrediction>.

Following each perturbation, sagittal and frontal WBAM time-series data were continuously compared against their respective prediction bands. PRT was defined as the time from perturbation onset until all WBAM trajectories returned to and remained fully within their respective bands for at least one full stride (Fig. 3).

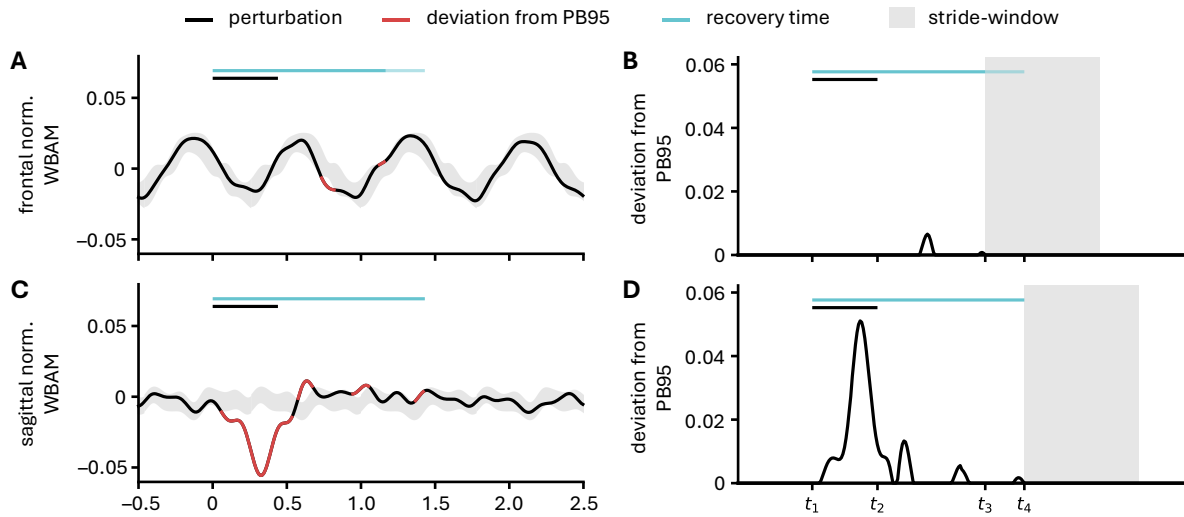


Fig. 3. Determination of perturbation recovery time from WBAM trajectories. Example sagittal (A) and frontal (B) WBAM trajectories (black) with prediction bands (grey shaded areas). Perturbation onset (t_1) marks the beginning of the perturbation, followed by perturbation end (t_2). During and after the perturbation, WBAM deviates outside the prediction band (red segments), indicating statistically significant departures from baseline dynamics. (C–D) Corresponding deviation signals used for recovery detection, illustrating excursions outside the prediction band and subsequent re-entries. The time point t_3 denotes when the first WBAM trajectory re-enters and is fully contained within the prediction band, while t_4 denotes when the second trajectory also re-enters and is fully contained. Recovery is only achieved once both trajectories remain within the prediction band for a full stride cycle (highlighted grey region). Perturbation recovery time (PRT) is therefore defined as the elapsed time from perturbation onset (t_1) to recovery (t_4).

To test the validity of the proposed PRT as a balance control metric, we compared PRT across two domains. Firstly, we compared PRT across perturbation intensities. As PRT is designed to measure the duration it takes to regain balance following a perturbation, stronger perturbations (i.e., greater velocity amplitude) should naturally correspond to longer PRTs. We calculated perturbation intensity accuracy as the proportion of runners who had shorter PRT in AP1 vs. AP2, AP2 vs. AP3 and AP1 vs. AP3. Secondly, we compared PRT across running conditions (i.e., across normal, restricted arm swing and weighted vest conditions) and calculated the participant-level detection accuracy as the proportion of runners who had shorter PRT in normal running vs. each impairment condition. We further calculated analogous detection accuracies for discrete WBAM-derived values of peak deviation following perturbation (PD) and peak-to-peak range over six strides following perturbation (PPR) and calculated condition detection accuracy of MLE.

Statistical analysis

All statistical analyses were conducted in RStudio (v. 2026.01.0, PBC, Boston, MA) with an alpha-level of 0.05. For a detailed analysis of perturbation direction, magnitude and condition effects of PRT, PD and PPR, we employed linear mixed-effects models accounting for the hierarchical structure of our data with subject included as a random intercept in all models. The fixed-effects structure was defined a priori as perturbation type and condition, and we evaluated whether adding the perturbation type by condition interaction improved model fit. Interaction terms were retained only when supported by likelihood-ratio model comparison. Otherwise, the corresponding main-effects model was reported. Inference for fixed effects was based on Type III ANOVA from the final selected model. For significant categorical effects, we computed estimated marginal means (EMMs) and Tukey-adjusted pairwise contrasts to identify differences between levels. Average values of perturbation type and condition are reported as EMMs [95% CI] averaged across all other levels and in case of significant pairwise contrasts, percentage change is reported for better interpretability.

A separate trial-level analysis was conducted for MLE. First, MLE differences across conditions were tested with a linear mixed-effects model including condition as a fixed effect and participant as a random intercept. Second, to test whether MLE predicted recovery behavior, mean PRT was aggregated at the trial level and modeled as a function of MLE with participant as random intercept and condition adjustment. For this analysis, the same reporting framework was used, including fixed effects, Type III tests, EMMs of linear trends and conditional and marginal R^2 values.

Results

Detection accuracy of all measures

PRT achieved the best overall detection accuracy with $80.8 \pm 10.5\%$, followed by PPR with $75.0 \pm 6\%$, PD with $73.0 \pm 11.8\%$ and MLE with $34.0 \pm 28.1\%$ (Fig. 4).

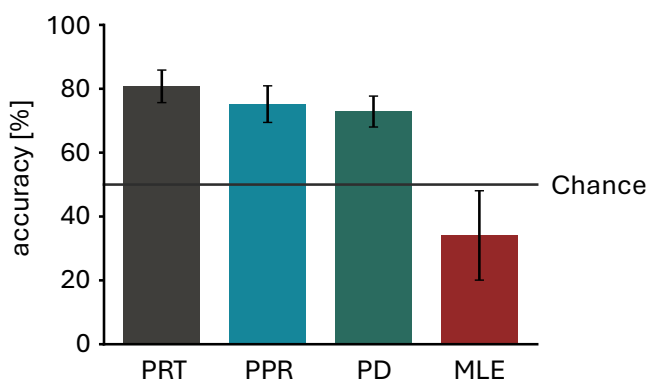


Figure 4. Detection accuracy of perturbation outcome measures. Comparison of detection accuracy across all evaluated measures. Bars represent mean detection accuracy (%) across subjects, with error bars indicating standard deviation. The horizontal line indicates chance level (50%).

While global accuracies were comparable between PRT, PPR and PD, condition accuracies were substantially different. PRT achieved good accuracy with $82.8 \pm 20.2\%$ in the comparison

between RA and NR and moderate accuracy with $66.7 \pm 25.0\%$ in the comparison between WV and NR (Fig. 5). Both PPR and PD achieved good accuracy in the RA vs. NR comparisons with $89.4 \pm 19.6\%$ and $79.1 \pm 22.0\%$, respectively, but achieved very poor accuracy in the WV vs. NR comparisons with $<20\%$, indicating that most runners had both lower PD and lower PPR during the WV condition. Perturbation intensity accuracies differed less between the three measures, with PRT achieving an overall accuracy of 86.8% , whereas PD and PPR achieved excellent accuracies of $100 \pm 3.8\%$ and $98.7 \pm 5.5\%$, respectively.

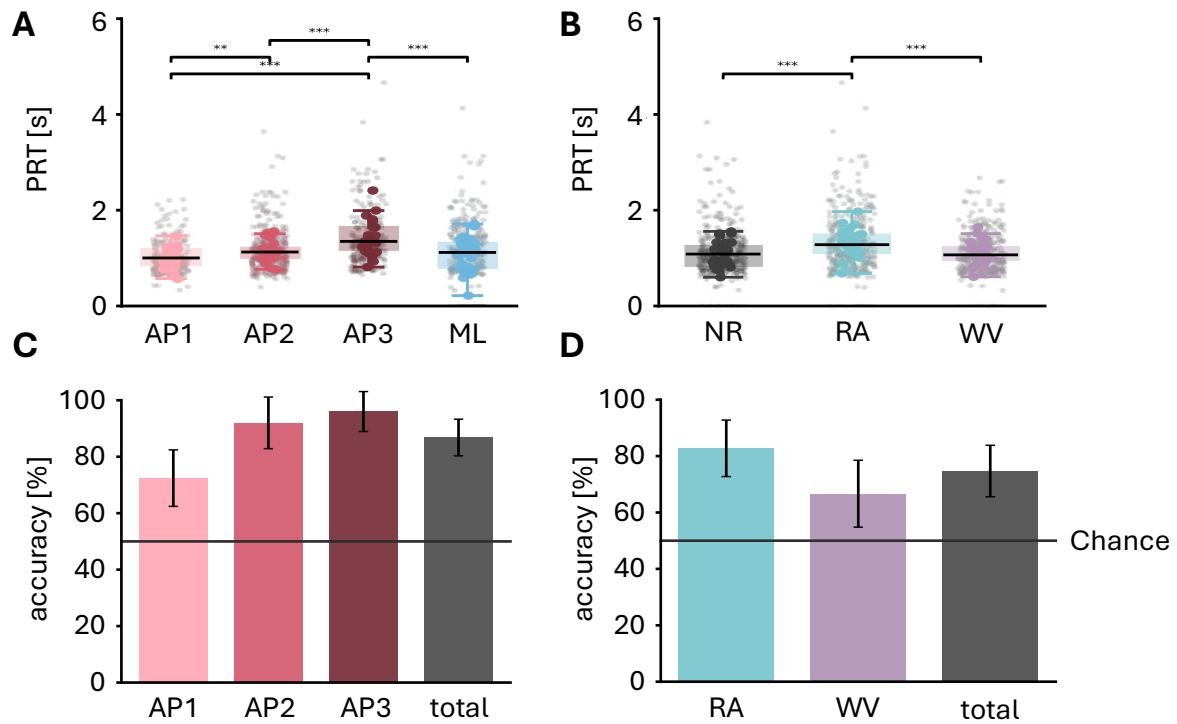


Figure 5. Perturbation recovery time and corresponding detection accuracies across conditions and perturbation intensities. (A) Perturbation recovery times (PRT) across perturbation types (AP1–AP3), shown as participant means (solid circles) and individual data points (transparent circles). (B) PRT across running conditions. Horizontal lines indicate the median across conditions or perturbation types, respectively, and asterisks denote significant differences. (C–D) Detection accuracies derived from PRT. (C) Accuracy across perturbation intensities, demonstrating high discriminability between intensity levels. (D) Accuracy across conditions, showing good discrimination between RA and Normal and moderate discrimination between WV and Normal. Bars represent mean accuracy (%) with standard deviation, and the horizontal line indicates chance level (50%). * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

PRT is calculated as the time from perturbation onset until the final deviation in frontal or sagittal WBAM from their respective prediction bands. As such, PRT following any perturbation is ultimately determined either by frontal or by sagittal WBAM deviation. Fig. 6 shows the percentage of perturbation during each perturbation and condition pair that were determined by frontal (darker color) and sagittal (lighter color) deviation. During NR and WV conditions, AP perturbations were more likely to be determined by sagittal WBAM deviation, whereas ML perturbations were more likely to be determined by frontal WBAM deviation. Only in the RA condition were both AP and ML perturbations more likely to be determined by frontal WBAM deviation.

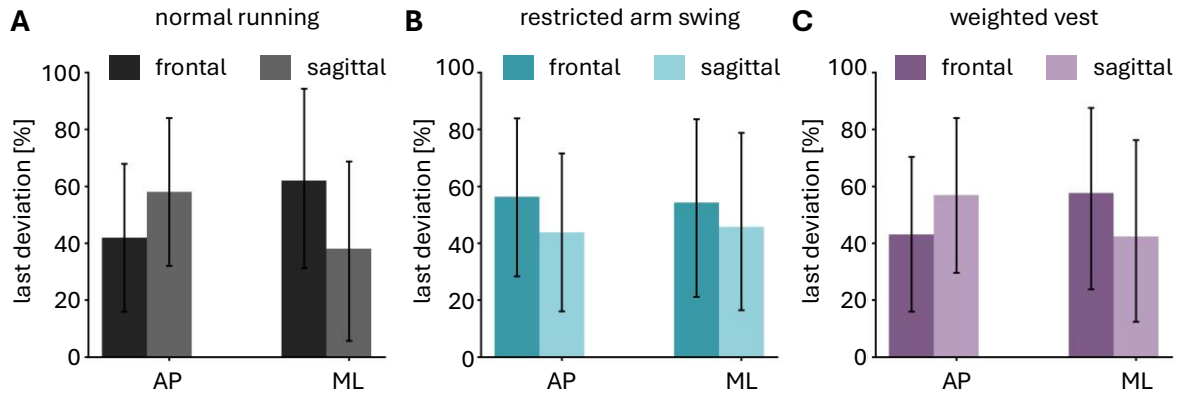


Fig. 6. Contribution of frontal and sagittal WBAM deviations to perturbation recovery time (PRT). Percentage of perturbations for which PRT was determined by the final deviation in either frontal (darker bars) or sagittal (lighter bars) WBAM across perturbation types and running conditions. Panels show results for (A) Normal running, (B) reduced arm swing and (C) weighted vest. Bars represent mean percentages across subjects with standard deviation.

The linear-effects model showed that both perturbation type ($F_{3,1153.1} = 34.19, p < 0.001$) and condition ($F_{2,1151.9} = 21.54, p < 0.001$) had significant effects on PRT (Fig. 5A, B). EMMs showed an increase in recovery times in RA (1.28 [1.19, 1.38] s) of 18.5% compared to NR (1.08 [0.99, 1.17] s, $p < 0.001$) and 16.6% compared to WV (1.10 [1.01, 1.19] s, $p < 0.001$), while NR and WV did not differ ($p = 0.752$). Across perturbation types, recovery time increased with AP intensity. AP3 (1.41 [1.32, 1.51] s) resulted in 23.9% greater PRT than AP2 (1.14 [1.05, 1.23] s, $p < 0.001$), 41.6% greater PRT than AP1 (1.00 [0.90, 1.10] s, $p < 0.001$) and 33.3% greater PRT than ML1 (1.06 [0.97, 1.15] s, $p < 0.001$). AP2 resulted in 14.2% longer recovery times than AP1 ($p = 0.006$) while AP2 and AP1 did not differ from ML1 ($p > 0.143$). Marginal R^2 was 0.10, while conditional R^2 was 0.23.

Peak deviation

For PD in the frontal plane, model fit was improved by including the perturbation type–condition interaction ($\chi^2(6) = 13.84, p = 0.031$). PD was significantly affected by perturbation type ($F_{3,1142.6} = 4.35, p = 0.005$), condition ($F_{2,1142.5} = 121.90, p < 0.001$) and their interaction ($F_{6,1141.3} = 2.32, p = 0.031$; Fig. 7A, B). EMM contrasts showed that frontal plane PD were highest in RA (0.0228 [0.0214, 0.0242]), second highest in NR (0.0210 [0.0195, 0.0224]) and lowest in WV (0.0196 [0.0182, 0.0211]). Pairwise contrasts showed that RA significantly increased frontal PD compared to NR (+8.6%, $p < 0.001$) and WV (+16.3%, $p < 0.001$), while WV significantly reduced frontal PD compared to NR (−6.7%, $p < 0.001$). All condition contrasts were significant across all perturbation types (all $p < 0.023$). Marginal R^2 was 0.084, while conditional R^2 was 0.651. For sagittal PD, the interaction was not supported ($\chi^2(6) = 8.23, p = 0.222$). Perturbation type ($F_{3,1143.7} = 2129.68, p < 0.001$) and condition ($F_{2,1143.7} = 13.65, p < 0.001$) significantly affected sagittal PD (Fig. 7C, D). Compared to WV (0.0338 [0.0317, 0.0360]), sagittal PD were 5.9% higher in NR (0.0358 [0.0337, 0.0379]), $p < 0.001$) and 5.6% higher in RA (0.0357 [0.0336, 0.0379], $p < 0.001$), while NR and RA did not differ ($p = 0.976$). Marginal R^2 was 0.719, while conditional R^2 was 0.870.

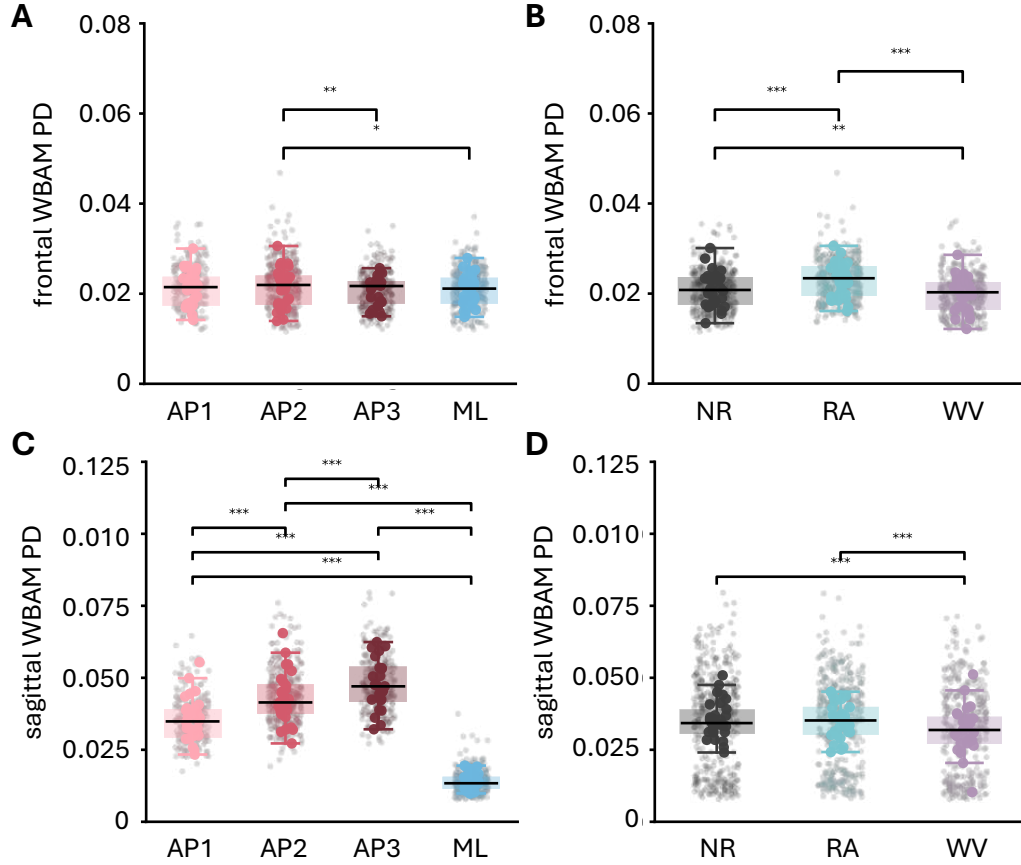


Figure 7. Peak WBAM deviation (PD) across perturbation types and running conditions. (A–B) Frontal plane PD across perturbation types (A) and running conditions (B), shown as boxplots (median, interquartile range, and whiskers) with individual data points overlaid. (C–D) Sagittal plane PD across perturbation types (C) and running conditions (D). Asterisks indicate significant pairwise comparisons. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

Peak-to-peak range

For PPR, both axes supported interaction models (frontal plane: $\chi^2(6) = 17.62, p = 0.007$; sagittal plane: $\chi^2(6) = 14.97, p = 0.020$). Frontal PPR was significantly affected by condition ($F_{2,1142.0} = 214.86, p < 0.001$) and perturbation type–condition interaction ($F_{6,1141.2} = 2.96, p = 0.007$), while perturbation type ($F_{3,1142.0} = 2.30, p = 0.076$) had no effect (Fig. 8A, B). Marginal condition EMMs were greatest in RA (0.0429 [0.0403, 0.0454]), followed by NR (0.0391 [0.0366, 0.0417]) and WV (0.0363 [0.0337, 0.0388]). All pairwise contrasts between conditions were significant and effects were consistent across all perturbation types. Compared to NR, RA led to a 9.7% increase in frontal PPR ($p < 0.001$), while WV led to a 7.8% decrease ($p < 0.001$). Marginal R^2 was 0.106, while conditional R^2 was 0.737. In the sagittal plane, PPR was affected by perturbation type ($F_{3,1144.0} = 2211.12, p < 0.001$), condition ($F_{2,1143.8} = 16.69, p < 0.001$) and perturbation type–condition interaction ($F_{6,1141.54} = 2.51, p < 0.020$; Fig. 8C, D). Marginal condition EMMs were greatest in RA (0.0429 [0.0403, 0.0454]), followed by NR (0.0391 [0.0366 to 0.0417]) and WV (0.0363 [0.0337, 0.0388]). Pairwise contrasts indicated condition differences depended on perturbation type. AP3 showed no condition differences, whereas AP1, AP2 and ML1 showed larger ranges in RA than WV (10.3%, 9.2% and 12.2%, respectively) and AP2 showed a larger range in NR than in WV (7.3%). Across all conditions, sagittal PPR increased with AP perturbation intensity and was

greater in AP than in ML direction. Across EMMs, values were 0.0459 [0.0433, 0.0485] for AP1, 0.0548 [0.0522, 0.0573] for AP2, 0.0645 [0.0619, 0.0671] for AP3 and 0.0194 [0.0168, 0.0219] for ML. Across marginal EMMs, AP2 was 19.4% higher than AP1, AP3 was 40.5% higher than AP1 and AP3 was 17.7% higher than AP2. Relative to ML, AP1 was 136.6% higher, AP2 was 182.5% higher and AP3 was 232.5% higher. Marginal R^2 was 0.735, while conditional R^2 was 0.874.

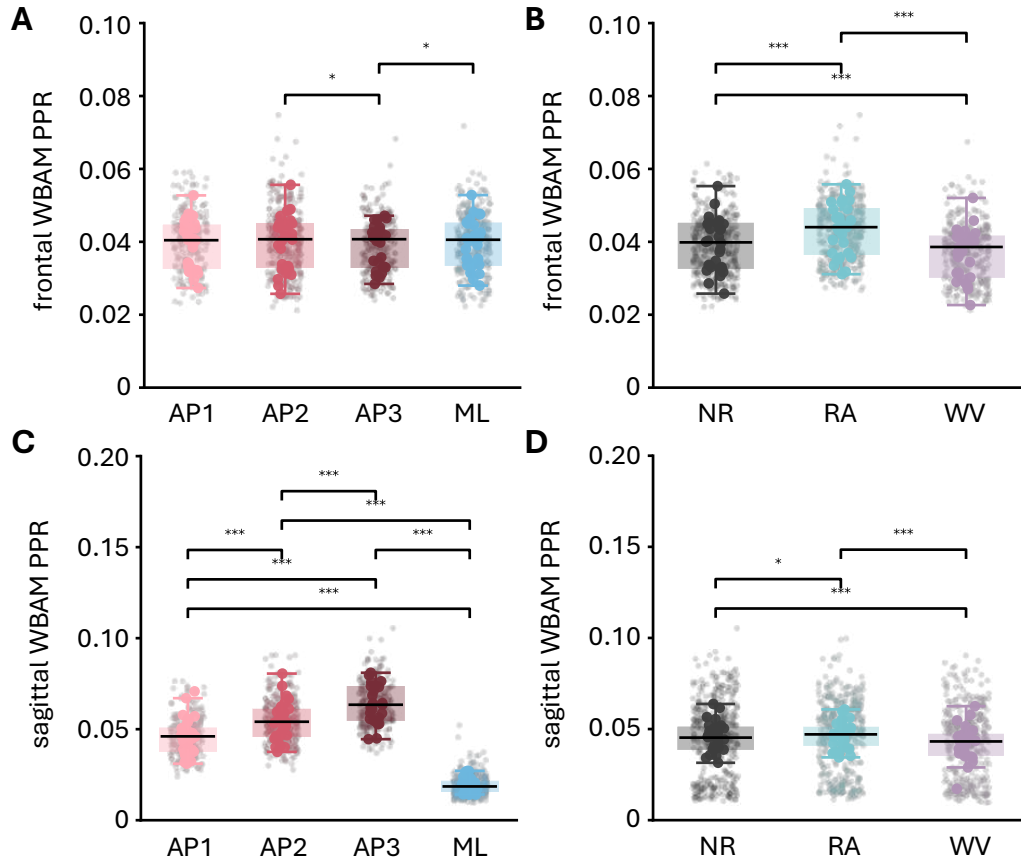


Figure 8. Peak-to-peak WBAM range (PPR) across perturbation types and running conditions. (A–B) Frontal plane PPR across perturbation types (A) and running conditions (B), shown as boxplots (median, interquartile range, and whiskers) with individual data points overlaid. (C–D) Sagittal plane PPR across perturbation types (C) and running conditions (D). Asterisks indicate significant pairwise comparisons. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

Maximum Lyapunov Exponent

MLE showed a significant condition effect ($F_{2,55.8} = 4.11, p = 0.022$; Fig. 9A). EMMs were 2.32 [2.23, 2.42] in NR, 2.17 [2.07, 2.26] in RA and 2.30 [2.20, 2.41] in WV. Tukey-adjusted pairwise contrasts indicated lower MLE in RA than NR ($p = 0.030$), with no significant difference between NR and WV ($p = 0.946$) or RA and WV ($p = 0.086$). Marginal R^2 was 0.098, while conditional R^2 was 0.279. In the recovery-time prediction analysis, both condition ($F_{2,56.5} = 8.07, p < 0.001$) and MLE ($F_{1,74.7} = 6.80, p = 0.011$) were significant predictors. The MLE slope was negative ($\beta = -0.239$), indicating that higher MLE was associated with shorter recovery time (Fig. 9B). Marginal R^2 was 0.169, while conditional R^2 was 0.632.

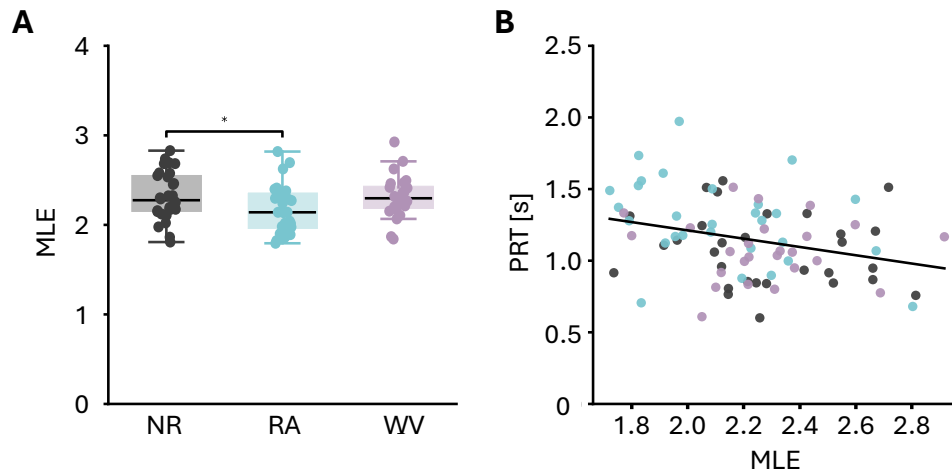


Figure 9. Maximum Lyapunov exponent (MLE) across conditions and its relationship with perturbation recovery time. (A) MLE across running conditions (normal running; NR, restricted arm swing; RA and weighted vest; WV), shown as boxplots (median, interquartile range, and whiskers) with individual data points overlaid. (B) Relationship between MLE and perturbation recovery time, shown as individual data points with fitted linear regression line. Asterisk indicates significant pairwise comparisons. * = $p < 0.05$.

Discussion

In this study, we applied treadmill-based perturbations during running to assess the accuracy of perturbation recovery time (PRT), peak WBAM deviation (PD) and peak-to-peak WBAM range (PPR) following perturbation and maximum Lyapunov exponent (MLE) to detect two artificial balance impairments and to separate perturbations of different magnitudes. Our novel measure of PRT had a participant-level condition accuracy of 74.8%, correctly identifying the restricted arm swing condition in 82.8% and the weighted vest condition in 66.7% of participants, respectively. PRT further correctly identified AP perturbation intensity in 86.6% of participants, combining to a global accuracy of 80.8% across all comparisons. These findings suggest that PRT can detect balance impairments during running and thus may represent a valid measure of balance ability. As such, it may have future utility as a screening tool for individuals whose athletic performance or injury risk is influenced by their capacity to recover from external perturbations.

In comparison to PRT, PPR and PD had lower global accuracies of 75% and 73%, respectively, when combining frontal and sagittal axes. The deviation-based metrics both had excellent accuracy in identifying AP perturbation intensity and good accuracy in detecting the restricted arm swing condition, but had poor accuracy in detecting the weighted vest condition. Values of <20% accuracy indicated that most runners had lower PPR and PD in the weighted vest (WV) condition compared with the normal running (NR) condition. As such, these findings—viewed in isolation—would lead to the conclusion that additional trunk mass improves balance ability. While it appears that the increased inertia of the upper body due to the additional mass dampened the initial change in angular momentum following the perturbation, it likely also made counteracting the forwards or sideways rotation more challenging, which is perhaps what delayed the return to steady-state kinematics and caused longer PRT in two-thirds of participants. PRT as a measure is designed to capture differences in the temporal pattern of the recovery, which may capture differences that are not visible in peak values or ranges. Still, PRT only had moderate accuracy in detecting the WV condition and average PRT did not differ

between WV and NR, indicating that some runners may be able to compensate for the increased demands of perturbation recovery due to the additional mass. Based on the proportional increase of frontal plane deviations determining PRT in RA compared to NR and WV conditions (Fig. 6), it appears that restricting arm swing reduced runners' ability to control WBAM in the frontal plane and return it quickly to baseline kinematics. This interpretation is further supported by 9.1% greater PD and 9.7% greater PPR in frontal-plane WBAM following AP perturbations compared to normal running (Figs. 7 and 8), whereas sagittal WBAM deviations and ranges did not differ between RA and NR conditions. These findings support previous studies indicating that arm swing contributes to lateral stability of running (Arellano & Kram, 2011; Koo, 2025) and plays an important role in perturbation recovery (Bruijn et al., 2010).

While the development of a valid measure of balance control is a crucial first step towards targeted training or technological intervention strategies for improvement of athletic performance or prevention of injuries through better balance ability, the link between balance ability and performance or injury risk first needs to be established. There is currently a lack of studies focusing on balance or perturbation recovery in running (Tokur et al., 2020), even though it seems intuitive that effective perturbation recovery would have important performance and injury prevention benefits across many running disciplines and team sports. In trail running, for example, it is speed on the downhill sections in particular that separates faster finishers from slower finishers, possibly owing to an increased ability to cope with the constant perturbations of technically challenging terrain (Genitrini et al., 2022). Trail runners further sustain a considerable number of traumatic injuries during races—primarily ankle sprains and fall-related skin injuries (Hamill et al., 2022; Viljoen et al., 2022)—which may, at least, be partially caused by an insufficient ability to cope with the increased balance demands of uneven terrain. In team sports, there is already research linking performance in static and dynamic balance tasks to occurrence of traumatic injuries (Oshima et al., 2018; P. Plisky et al., 2021; P. J. Plisky et al., 2006). However, these balance tests are far-removed from in-game actions and perhaps a stronger link between balance ability and injury rates could be established with more specific balance tests. For example, measuring PRT when applying external mechanical perturbations to the pelvis of players while they run on a treadmill (e.g., using the Bump'Em perturbation system developed by Tan et al. (2020)) could be a more applicable test of balance control. Future studies might address whether using PRT in conjunction with externally valid perturbation could be a useful predictor of injury occurrence or performance in a variety of athletic contexts.

For the first time, we found that MLE was significantly lower during running with restricted arm swing compared to normal running (Fig. 9A) and that lower MLE values (i.e., higher LDS) were related to longer recovery times following perturbation (Fig. 9B). MLE is used in running biomechanics as a measure of the runners' resistance to small continuous perturbations that occur during every step (Frank et al., 2019b; Hoenig et al., 2019; Promsri, 2022, for a review see Winter et al., 2023) and is typically interpreted as dynamic stability. However, dynamic stability could also be defined as a runners' ability to maintain or regain steady locomotion following a large external perturbation, a definition that is conceptually related to that of LDS and very similar to the way we calculated PRT. As such, one would expect that MLE and PRT would show a positive correlation with higher LDS being associated with quicker recoveries. However, we found the opposite, with the restricted arm swing condition showing the highest LDS out of all conditions and higher LDS predicting longer recovery times. These findings indicate that greater LDS is associated with lower resistance against large external perturbations and suggest that MLE and PRT measure fundamentally different aspects of running stability. One possible explanation for the divergence of MLE and PRT is the fact that MLE is primarily a measure of steady-state stability, whereas PRT depends both on steady-state stability and

active recovery following a perturbation—two characteristics that may not necessarily be linked. In fact, increasing steady-state stability may be a compensatory preventive strategy that individuals choose when external circumstances (e.g., when running with restricted arm swing or expecting external perturbations) or intrinsic factors (e.g., poor balance ability, running under fatigue) increase the perceived likelihood of a fall. If this were the case, then runners who have poor reactive balance abilities (indicated by long recovery times following large perturbations) or are artificially impaired may run with higher LDS to reduce the reactive compensation necessary to regain balance. This preventive strategy could also partially explain the counterintuitive finding that runners increase their LDS when running under severe fatigue during an exhaustive time-trial (Hoenig et al., 2019). This compensatory strategy may, however, come at the cost of an increased cost of transport (Thomas et al., 1999; Von Diecken, Riedl, Willwacher, et al., 2025) and may thus impede running performance. While this preventive compensation would explain the results found in the present study, additional work is necessary to elucidate whether this is truly the case.

There are several limitations to consider when interpreting results from this study. Firstly, we calculated PRT as the time to recover back to steady-state WBAM kinematics and as such, results depend critically on how steady state running is defined. In our data, mean band width differed significantly between conditions when creating prediction bands separately for each condition. As this would artificially bias results towards shorter recovery times in the condition with larger band width even if the WBAM kinematics following perturbation are identical, we decided to create the prediction bands based on all unperturbed strides across conditions. This approach results in a single steady state band per WBAM axis and participant. While this improves comparability between conditions and resulted in higher impairment detection accuracies in our dataset, it may reduce absolute accuracy and reliability if different conditions are included in the dataset. Secondly, we selected frontal and sagittal WBAM as our principal variables to describe the balance state of the runner and determine the point of full recovery. While these variables have a solid foundation within the balance control research of locomotion (Herr & Popovic, 2008; Hinrichs, 1987; Negishi & Ogihara, 2023; Popovic et al., 2004; Van Dieën et al., 2025b), additional state variables like COM position and velocity could provide valuable information and enhance accuracy of the PRT calculation. Future studies should focus on finding the ideal set of variables to accurately describe the balance state during running. Lastly, as this was the first validation study of PRT in running, we did not include a reliability analysis within this article. As a result, the inter-day reliability of PRT in running within participants is currently unknown and should be subject of future studies.

Conclusion

This study proposes perturbation recovery time (PRT) as a novel metric to quantify balance control during running and provides initial evidence supporting its validity. By defining recovery as the return to statistically derived steady-state WBAM dynamics, PRT offers a direct and interpretable measure of how effectively runners respond to external perturbations. Across all evaluated comparisons, PRT demonstrated the highest detection accuracy and was the only metric that consistently identified both perturbation intensity and experimentally induced balance impairments. In contrast, traditional WBAM-derived measures based on peak values or ranges captured only partial aspects of the perturbation response. The maximum Lyapunov exponent further showed limited practical relevance in this context and was inversely related to perturbation recovery performance, suggesting that steady-state stability and reactive balance control represent distinct constructs in running. These findings indicate that inclusion of perturbation recovery dynamics when assessing balance control adds valuable information not—or at least not fully—contained in steady-state measures. PRT captures the temporal evolution

of the recovery process and therefore may provide a more functionally meaningful representation of balance control. While promising, the broader applicability of PRT remains to be established. Future research should evaluate its reliability, sensitivity to different perturbation modalities and its relationship to real-world outcomes such as performance and injury risk. Establishing these links will be essential to determine whether PRT can serve as a practical tool for screening, monitoring and intervention in athletic and clinical populations.

Literature

- Adam, M. D., McElvain, D., George Hornby, T., Hyngstrom, A. S., & Schmit, B. D. (2025). Whole body angular momentum characterizes reactive balance adaptations and perturbation intensity. *Journal of Biomechanics*, 179, 112474. <https://doi.org/10.1016/j.jbiomech.2024.112474>
- AminiAghdam, S., Blickhan, R., & Karamanidis, K. (2020). The influence of sagittal trunk leans on uneven running mechanics. *Journal of Experimental Biology*, jeb.228288. <https://doi.org/10.1242/jeb.228288>
- Arellano, C. J., & Kram, R. (2011). The effects of step width and arm swing on energetic cost and lateral balance during running. *Journal of Biomechanics*, 44(7), 1291–1295. <https://doi.org/10.1016/j.jbiomech.2011.01.002>
- Arellano, C. J., McDermott, W. J., Kram, R., & Grabowski, A. M. (2015). Effect of Running Speed and Leg Prostheses on Mediolateral Foot Placement and Its Variability. *PLOS ONE*, 10(1), e0115637. <https://doi.org/10.1371/journal.pone.0115637>
- Braun, L., Umberger, B. R., Joshi, V., Denis, Y., Anedda, B., & Willwacher, S. (2026). Humans choose trunk lean angles consistent with minimal metabolic costs—An integrative experimental and musculoskeletal simulation approach. *Proceedings of the Royal Society B: Biological Sciences*, 293(2063), 20252717. <https://doi.org/10.1098/rspb.2025.2717>
- Bruijn, S. M., Bregman, D. J. J., Meijer, O. G., Beek, P. J., & Van Dieën, J. H. (2012). Maximum Lyapunov exponents as predictors of global gait stability: A modelling approach. *Medical Engineering & Physics*, 34(4), 428–436. <https://doi.org/10.1016/j.medengphy.2011.07.024>
- Bruijn, S. M., Meijer, O. G., Beek, P. J., & Van Dieën, J. H. (2010). The effects of arm swing on human gait stability. *Journal of Experimental Biology*, 213(23), 3945–3952. <https://doi.org/10.1242/jeb.045112>
- Bruijn, S. M., Meijer, O. G., Beek, P. J., & Van Dieën, J. H. (2013). Assessing the stability of human locomotion: A review of current measures. *Journal of The Royal Society Interface*, 10(83), 20120999. <https://doi.org/10.1098/rsif.2012.0999>

- Bruijn, S. M., Van Dieën, J. H., Meijer, O. G., & Beek, P. J. (2009). Statistical precision and sensitivity of measures of dynamic gait stability. *Journal of Neuroscience Methods*, 178(2), 327–333.
<https://doi.org/10.1016/j.jneumeth.2008.12.015>
- Creutzinger, M. L., Todo, X., Liebl, D., Mensinger, T., & Sharp, J. L. (n.d.). *Adaptive Simultaneous Prediction Bands in Concurrent Functional Linear Regression*.
- Delp, S. L., Anderson, F. C., Arnold, A. S., Loan, P., Habib, A., John, C. T., Guendelman, E., & Thelen, D. G. (2007). OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement. *IEEE Transactions on Biomedical Engineering*, 54(11), 1940–1950.
<https://doi.org/10.1109/TBME.2007.901024>
- Delp, S. L., Loan, J. P., Hoy, M. G., Zajac, F. E., Topp, E. L., & Rosen, J. M. (1990). An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. *IEEE Transactions on Biomedical Engineering*, 37(8), 757–767. <https://doi.org/10.1109/10.102791>
- Dufek, J. S., Mercer, J. A., & Griffin, J. R. (2009). The Effects of Speed and Surface Compliance on Shock Attenuation Characteristics for Male and Female Runners. *Journal of Applied Biomechanics*, 25(3), 219–228. <https://doi.org/10.1123/jab.25.3.219>
- Ekizos, A., Santuz, A., & Arampatzis, A. (2017). Transition from shod to barefoot alters dynamic stability during running. *Gait & Posture*, 56, 31–36.
<https://doi.org/10.1016/j.gaitpost.2017.04.035>
- Ernst, M., Götze, M., Blickhan, R., & Müller, R. (2019). Humans adjust the height of their center of mass within one step when running across camouflaged changes in ground level. *Journal of Biomechanics*, 84, 278–283. <https://doi.org/10.1016/j.jbiomech.2018.12.036>
- Fohrmann, D., Hamacher, D., Sanchez-Alvarado, A., Potthast, W., Mai, P., Willwacher, S., & Hollander, K. (2022). Reliability of Running Stability during Treadmill and Overground Running. *Sensors*, 23(1), 347. <https://doi.org/10.3390/s23010347>
- Frank, N. S., Prentice, S. D., & Callaghan, J. P. (2019a). Local dynamic stability of the lower extremity in novice and trained runners while running intraditional and minimal footwear. *Gait & Posture*, 68, 50–54. <https://doi.org/10.1016/j.gaitpost.2018.10.034>

- Frank, N. S., Prentice, S. D., & Callaghan, J. P. (2019b). Local dynamic stability of the lower extremity in novice and trained runners while running intraditional and minimal footwear. *Gait & Posture*, 68, 50–54. <https://doi.org/10.1016/j.gaitpost.2018.10.034>
- Genitrini, M., Fritz, J., Zimmermann, G., & Schwameder, H. (2022). Downhill Sections Are Crucial for Performance in Trail Running Ultramarathons—A Pacing Strategy Analysis. *Journal of Functional Morphology and Kinesiology*, 7(4), 103. <https://doi.org/10.3390/jfmk7040103>
- Golyski, P. R., Swaich, N. K., Panizzolo, F. A., & Sawicki, G. S. (2025). American Society of Biomechanics Young Scientist Pre-doctoral Award 2022: Effects of an elastic hip exoskeleton on stability quantified by mechanical energetics and whole-body angular momentum during walking with treadmill belt speed perturbations. *Journal of Biomechanics*, 188, 112784. <https://doi.org/10.1016/j.jbiomech.2025.112784>
- Grimmer, S., Ernst, M., Günther, M., & Blickhan, R. (2008). Running on uneven ground: Leg adjustment to vertical steps and self-stability. *Journal of Experimental Biology*, 211(18), 2989–3000. <https://doi.org/10.1242/jeb.014357>
- Hamill, J., Hercksen, J., Salzano, M., Udofa, A., & Trudeau, M. B. (2022). The prevalence of injuries in trail running: Influence of trails, terrains and footwear. *Footwear Science*, 14(2), 113–121. <https://doi.org/10.1080/19424280.2022.2048901>
- Hamner, S. R., Seth, A., & Delp, S. L. (2010). Muscle contributions to propulsion and support during running. *Journal of Biomechanics*, 43(14), 2709–2716. <https://doi.org/10.1016/j.jbiomech.2010.06.025>
- Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, 394(6695), 780–784. <https://doi.org/10.1038/29528>
- Herr, H., & Popovic, M. (2008). Angular momentum in human walking. *Journal of Experimental Biology*, 211(4), 467–481. <https://doi.org/10.1242/jeb.008573>
- Hinrichs, R. N. (1987). Upper Extremity Function in Running. II: Angular Momentum Considerations. *International Journal of Sport Biomechanics*, 3(3), 242–263. <https://doi.org/10.1123/ijsb.3.3.242>

Hipper, M., Braun, L., Denis, Y., Anedda, B., Helwig, J., Von Diecken, C., & Willwacher, S. (2025). A New Method to Accurately Describe Belt-Based Treadmill Perturbations. SSRN.

<https://doi.org/10.2139/ssrn.5076247>

Hoenig, T., Hamacher, D., Braumann, K., Zech, A., & Hollander, K. (2019). Analysis of running stability during 5000 m running*. *European Journal of Sport Science*, 19(4), 413–421.

<https://doi.org/10.1080/17461391.2018.1519040>

Hollander, K., Hamacher, D., & Zech, A. (2021). Running barefoot leads to lower running stability compared to shod running—Results from a randomized controlled study. *Scientific Reports*, 11(1), 4376. <https://doi.org/10.1038/s41598-021-83056-9>

Kettner, C., Stetter, B., & Stein, T. (2024). *The effects of running shoe stack height on running style and stability during level running at different running speeds.*

<https://doi.org/10.1101/2024.11.19.624278>

Koo, Y.-J. (2025). *Active Arm Swing During Running Improves Rotational Stability of the Upper Body and Metabolic Energy Efficiency.*

Look, N., Arellano, C. J., Grabowski, A. M., McDermott, W. J., Kram, R., & Bradley, E. (2013). Dynamic stability of running: The effects of speed and leg amputations on the maximal Lyapunov exponent. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(4), 043131.

<https://doi.org/10.1063/1.4837095>

Loturco, I., Pereira, L. A., Reis, V. P., Abad, C. C. C., Freitas, T. T., Azevedo, P. H. S. M., & Nimphius, S. (2022). Change of Direction Performance in Elite Players From Different Team Sports. *Journal of Strength & Conditioning Research*, 36(3), 862–866.

<https://doi.org/10.1519/JSC.0000000000003502>

Mai, P., & Willwacher, S. (2019). Effects of low-pass filter combinations on lower extremity joint moments in distance running. *Journal of Biomechanics*, 95, 109311.

<https://doi.org/10.1016/j.jbiomech.2019.08.005>

- Martelli, D., Monaco, V., Luciani, L. B., & Micera, S. (2013). Angular Momentum During Unexpected Multidirectional Perturbations Delivered While Walking. *IEEE Transactions on Biomedical Engineering*, 60(7), 1785–1795. <https://doi.org/10.1109/TBME.2013.2241434>
- Naughton, M., Miller, J., & Slater, G. J. (2018). Impact-Induced Muscle Damage and Contact Sports: Etiology, Effects on Neuromuscular Function and Recovery, and the Modulating Effects of Adaptation and Recovery Strategies. *International Journal of Sports Physiology and Performance*, 13(8), 962–969. <https://doi.org/10.1123/ijsp.2017-0268>
- Negishi, T., & Ogihara, N. (2023). Regulation of whole-body angular momentum during human walking. *Scientific Reports*, 13(1), 8000. <https://doi.org/10.1038/s41598-023-34910-5>
- Osborne, L. C., Lisberger, S. G., & Bialek, W. (2005). A sensory source for motor variation. *Nature*, 437(7057), 412–416. <https://doi.org/10.1038/nature03961>
- Oshima, T., Nakase, J., Kitaoka, K., Shima, Y., Numata, H., Takata, Y., & Tsuchiya, H. (2018). Poor static balance is a risk factor for non-contact anterior cruciate ligament injury. *Archives of Orthopaedic and Trauma Surgery*, 138(12), 1713–1718. <https://doi.org/10.1007/s00402-018-2984-z>
- Paquette, M. R., Melaro, J. A., Smith, R., & Moore, I. S. (2024). Time to stability of treadmill running kinematics in novel footwear with different midsole thickness. *Journal of Biomechanics*, 164, 111984. <https://doi.org/10.1016/j.jbiomech.2024.111984>
- Plisky, P. J., Rauh, M. J., Kaminski, T. W., & Underwood, F. B. (2006). Star Excursion Balance Test as a Predictor of Lower Extremity Injury in High School Basketball Players. *RESEARCH REPORT*.
- Plisky, P., Schwartkopf-Phifer, K., Huebner, B., Garner, M. B., & Bullock, G. (2021). Systematic Review and Meta-Analysis of the Y-Balance Test Lower Quarter: Reliability, Discriminant Validity, and Predictive Validity. *International Journal of Sports Physical Therapy*, 16(5). <https://doi.org/10.26603/001c.27634>
- Popovic, M., Hofmann, A., & Herr, H. (2004). Angular momentum regulation during human walking: Biomechanics and control. *IEEE International Conference on Robotics and Automation*, 2004.

Proceedings. ICRA '04. 2004, 2405-2411 Vol.3.

<https://doi.org/10.1109/ROBOT.2004.1307421>

Promsri, A. (2022). Sex Difference in Running Stability Analyzed Based on a Whole-Body Movement:

A Pilot Study. *Sports*, 10(9), 138. <https://doi.org/10.3390/sports10090138>

Rosenstein, M. T., Collins, J. J., & De Luca, C. J. (1993). A practical method for calculating largest

Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, 65(1–2), 117–

134. [https://doi.org/10.1016/0167-2789\(93\)90009-P](https://doi.org/10.1016/0167-2789(93)90009-P)

Scheer, V., Basset, P., Giovanelli, N., Vernillo, G., Millet, G. P., & Costa, R. J. S. (2020). Defining Off-

road Running: A Position Statement from the Ultra Sports Science Foundation. *International*

Journal of Sports Medicine, 41(05), 275–284. <https://doi.org/10.1055/a-1096-0980>

Schütte, K. H., Seerden, S., Venter, R., & Vanwanseele, B. (2018). Influence of outdoor running fatigue

and medial tibial stress syndrome on accelerometer-based loading and stability. *Gait &*

Posture, 59, 222–228. <https://doi.org/10.1016/j.gaitpost.2017.10.021>

Seethapathi, N., & Srinivasan, M. (2019). Step-to-step variations in human running reveal how

humans run without falling. *eLife*, 8, e38371. <https://doi.org/10.7554/eLife.38371>

Sepp, L. A., Baum, B. S., Nelson-Wong, E., & Silverman, A. K. (2019). Dynamic balance during running

using running-specific prostheses. *Journal of Biomechanics*, 84, 36–45.

<https://doi.org/10.1016/j.jbiomech.2018.12.016>

Seth, A., Hicks, J. L., Uchida, T. K., Habib, A., Dembia, C. L., Dunne, J. J., Ong, C. F., DeMers, M. S.,

Rajagopal, A., Millard, M., Hamner, S. R., Arnold, E. M., Yong, J. R., Lakshmikanth, S. K.,

Sherman, M. A., Ku, J. P., & Delp, S. L. (2018). OpenSim: Simulating musculoskeletal dynamics

and neuromuscular control to study human and animal movement. *PLOS Computational*

Biology, 14(7), e1006223. <https://doi.org/10.1371/journal.pcbi.1006223>

Shokouhi, S., Sritharan, P., & Lee, P. V.-S. (2024). Recovering whole-body angular momentum and

margin of stability after treadmill-induced perturbations during sloped walking in healthy

young adults. *Scientific Reports*, 14(1), 4421. <https://doi.org/10.1038/s41598-024-54890-4>

- Tan, G. R., Raitor, M., & Collins, S. H. (2020). Bump'em: An Open-Source, Bump-Emulation System for Studying Human Balance and Gait. *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 9093–9099. <https://doi.org/10.1109/ICRA40945.2020.9197105>
- Thielemans, V., Meyns, P., & Bruijn, S. M. (2014). Is angular momentum in the horizontal plane during gait a controlled variable? *Human Movement Science*, 34, 205–216. <https://doi.org/10.1016/j.humov.2014.03.003>
- Thomas, D. Q., Fernhall, B., & Granat, H. (1999). *Changes in Running Economy During a 5-km Run in Trained Men and Women Runners*.
- Tokur, D., Grimmer, M., & Seyfarth, A. (2020). Review of balance recovery in response to external perturbations during daily activities. *Human Movement Science*, 69, 102546. <https://doi.org/10.1016/j.humov.2019.102546>
- Van Dieën, J. H., Bruijn, S., Lemaire, K. K., & Kistemaker, D. A. (2025a). *Simultaneous stabilizing feedback control of linear and angular momentum in human walking*. Biophysics. <https://doi.org/10.1101/2025.01.09.632072>
- Van Dieën, J. H., Bruijn, S., Lemaire, K. K., & Kistemaker, D. A. (2025b). *Simultaneous stabilizing feedback control of linear and angular momentum in human walking*. Biophysics. <https://doi.org/10.1101/2025.01.09.632072>
- Viljoen, C., Janse Van Rensburg, D. C. (Christa), Van Mechelen, W., Verhagen, E., Silva, B., Scheer, V., Besomi, M., Gajardo-Burgos, R., Matos, S., Schoeman, M., Jansen Van Rensburg, A., Van Dyk, N., Scheepers, S., & Botha, T. (2022). Trail running injury risk factors: A living systematic review. *British Journal of Sports Medicine*, 56(10), 577–587. <https://doi.org/10.1136/bjsports-2021-104858>
- Von Diecken, C., Riedl, M., & Ueberschär, O. (2025). Advanced footwear technology does not affect local dynamic stability during treadmill running. *Footwear Science*, 17(sup1), 142. <https://doi.org/10.1080/19424280.2025.2489675>

- Von Diecken, C., Riedl, M., Willwacher, S., & Ueberschär, O. (2025). Runners with lower dynamic stability exhibit better running economy. *Scientific Reports*, 15(1), 38117.
<https://doi.org/10.1038/s41598-025-26008-x>
- Willwacher, S., Kurz, M., Menne, C., Schrödter, E., & Brüggemann, G.-P. (2016). Biomechanical response to altered footwear longitudinal bending stiffness in the early acceleration phase of sprinting. *Footwear Science*, 8(2), 99–108. <https://doi.org/10.1080/19424280.2016.1144653>
- Willwacher, S., Oberländer, K. D., Mai, P., Mählich, D., Kurz, M., Koopmann, T., Fohrmann, D., Kantarev, A., & Kersting, U. G. (2021). A new method for measuring treadmill belt velocity fluctuations: Effects of treadmill type, body mass and locomotion speed. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-81951-9>
- Winter, L., Taylor, P., Bellenger, C., Grimshaw, P., & Crowther, R. G. (2023). The application of the Lyapunov Exponent to analyse human performance: A systematic review. *Journal of Sports Sciences*, 41(22), 1994–2013. <https://doi.org/10.1080/02640414.2024.2308441>
- Wu, J., Raitor, M., Tan, G. R., Staudenmayer, K. L., Delp, S. L., Liu, C. K., & Collins, S. H. (2025). Detecting artificially impaired balance in human locomotion: Metrics, perturbation effects and detection thresholds. *Journal of Experimental Biology*, 228(10).
<https://doi.org/10.1242/jeb.249339>
- Wu, J., Raitor, M., Truong, T. E., Liu, C. K., & Collins, S. H. (n.d.). *Perturbation Recovery Time Identifies Subtle Human Balance Impairments and Features*.
- Yamagata, S., Yamaguchi, T., Shinya, M., Milosevic, M., & Masani, K. (2024). Comparison of sensitivity among dynamic balance measures during walking with different tasks. *Royal Society Open Science*, 11(1), 230883. <https://doi.org/10.1098/rsos.230883>
- Yamaguchi, T., & Masani, K. (2022). Effects of age on dynamic balance measures and their correlation during walking across the adult lifespan. *Scientific Reports*, 12(1), 14301.
<https://doi.org/10.1038/s41598-022-18382-7>
- Young, W. B., Dawson, B., & Henry, G. J. (2015). Agility and Change-of-Direction Speed are Independent Skills: Implications for Training for Agility in Invasion Sports. *International*

Journal of Sports Science & Coaching, 10(1), 159–169. <https://doi.org/10.1260/1747-9541.10.1.159>