Between-unit agreement of GPS-derived mechanical work and power: time to trust the numbers?

Martin Buchheit, [1 2 3 4 5] **Andres Lopez Sagarra** [6]

[1]Type 3.2 Performance, Montvalezan, France
[2]Optimo Sport Performance Center, Estepona, Spain
[3]INSEP, Paris, France
[4]HIIT Science, Revelstoke, Canada
[5]Athletica, Revelstoke, Canada
[6]SPORT Research Group (CTS-1024), CIBIS Research Center, University of Almería, Almería, Spain

GPS reliability | **Between-unit agreement** | **Mechanical work** | **Mechanical power** | **Standardized error** | **Coefficient of variation** | **WIMU Pro** | **ADI** | **Systematic error** | **Football**

Headline

GPS-derived mechanical work and mechanical power represent a new generation of player tracking metrics that combine position and velocity data to quantify the mechanical demands of locomotion beyond traditional distance and speed measures (Buchheit & Simpson, 2017; Buchheit et al., 2026). These metrics, derived through advanced signal processing of raw GPS data (e.g., Athletic Data Innovations, ADI, Hudl), have gained increasing interest for training load monitoring, return-to-performance assessment, and match analysis. However, while the test-retest reliability of some of these metrics has been examined in field conditions (Buchheit et al., 2018), their between-unit agreement, the measurement error introduced when different GPS units are used to track the same physical activity, has not been independently assessed.

Between-unit agreement is a distinct and practically important aspect of measurement quality. In team sport settings, it is often impractical to assign the same GPS unit to the same player across all sessions. When units are swapped, any difference between devices becomes a source of noise in the longitudinal tracking of individual players. Early work on between-unit variability using 50 GPS units on a custom-built sled revealed substantial between-unit differences, with CV values ranging from 1% for peak speed to 56% for deceleration counts, and some units measuring 2 to 6 times more acceleration events than others (Buchheit et al., 2014). These findings raised serious concerns about the usefulness of acceleration-derived GPS metrics at the time.

Since then, GPS hardware has evolved considerably. Current-generation devices such as the WIMU PRO EVO (Hudl) operate with multi-constellation satellite reception (GPS, GLONASS, Galileo), improved chipsets, and advanced filtering algorithms (Gomez-Carmona et al., 2019; Rodriguez-Fernandez et al., 2025). Whether these technological advances have translated into meaningfully improved between-unit agreement, particularly for the newer ADI-derived mechanical metrics, remains to be established.

A further limitation of the existing literature is the near-universal reliance on the coefficient of variation (CV%) as the primary indicator of between-unit reliability. While straightforward to compute, CV% lacks an external reference point: a 10% CV on PlayerLoad may sound concerning, but without knowing how this compares to the typical between-player differences observed in competitive matches, practitioners cannot determine whether the measurement noise could genuinely alter their interpretation of an individual player's data. This practical gap motivated the development of a standardized error approach in the present study, where between-unit measurement error is expressed relative to the between-player variability observed in elite football matches.

Aim

To quantify the between-unit agreement of 10 WIMU PRO EVO GPS units for both traditional GPS 2.0 metrics and ADI-derived GPS 3.0 metrics (mechanical work, mechanical power, and movement vector classification), to introduce a standardized error metric that expresses measurement noise relative to match-level between-player variability, and to decompose the total between-unit error into systematic and random components to inform practical unit management decisions.

Methods

Experimental setup

Ten WIMU PRO EVO units (Hudl, Lincoln, US; 10 Hz GNSS, 100 Hz accelerometer) were mounted on a rigid trolley using adhesive tape, ensuring all units shared an identical trajectory (Figure 1). The trolley was manually pulled through 12 tasks on outdoor pitches in Doha, Qatar (April 2026). Tasks included constant-speed running (416 m athletics track lap), high-speed running efforts at near-maximal velocity (two variants, including one with a curve), random changes of direction at moderate speed (two variants with different operators), curve and linear maximal acceleration and deceleration runs, and five pitch-dimension shuttles of varying length (16 m, central circle circumference, pitch length, pitch width). A full session combining all tasks (16 min, 1660 m) was also recorded and analyzed as a single continuous block, analogous to what practitioners would observe over a training session or a segment of match play. All units maintained excellent satellite reception throughout (mean 28.6 ± 2.1 satellites; mean Position Dilution of Precision (PDOP) 0.957 ± 0.081).

Data processing

Raw GPS files were processed through two independent pipelines. The native WIMU software (Hudl Signal) provided GPS 2.0 metrics: total distance, high-speed running distance (>19.8 km/h), peak speed, maximum acceleration and deceleration, PlayerLoad (the only metric derived from the tri-axial accelerometer, quantifying the cumulative tri-planar acceleration load expressed in arbitrary units; AU), and high-intensity acceleration and deceleration event counts. The same raw files were independently reprocessed through ADI (Athletic Data Innovations, Hudl) to derive GPS 3.0 metrics: total mechanical work, mechanical work segmented by movement phase (linear and directional components), and mechanical power (mean and peak, normalized to body mass). All mechanical work and power calculations are derived from GPS position and velocity data; the inertial measurement unit is not involved in these computations. Directional sub-components (cuts left/right;

arcs left/right) were aggregated into combined cuts and combined arcs metrics for reporting, as individual left/right values are unlikely to be monitored separately in practice. A detailed description of the ADI mechanical metrics and their validation against instrumented treadmill data has been published previously (Buchheit et al., 2015; Buchheit et al., 2018).


Fig. 1. **Trolley setup with 10 WIMU PRO EVO units mounted**

Statistical analysis

Between-unit agreement

For each task and metric, the between-unit standard deviation (SD), coefficient of variation (CV% = SD/mean × 100), and worst-case range (maximum minus minimum across 10 units) were computed. While all 12 tasks were analyzed individually to explore task-specific patterns, the primary results reported in the tables are from the full session, as the aggregation across multiple movement types provides a more ecologically valid estimate of between-unit error. For maximum acceleration and deceleration, values were excluded from task-specific analyses when the group mean was below 3 m/s^2, as these represent tasks with no intent to produce maximal efforts and the apparent variability reflects noise in near-zero values rather than genuine measurement error. This issue does not affect the full session results, where the aggregation naturally includes tasks with genuine maximal efforts.

Standardized measurement error

The between-unit SD was standardized against the between-player SD observed in elite football matches: Standardized Error (SE) = between-unit SD / between-player match SD. Hopkins magnitude thresholds were applied: <0.10 trivial, 0.10-0.30 small, 0.30-0.60 moderate, >0.60 large (Hopkins et al., 2009). Reference between-player match SDs were: total distance 990 m and HSR distance 171 m (Collins et al., 2025; MLS, N = 1243 matches, HSR defined as 19.8-25.2 km/h), and peak speed 1.4 km/h (Silva et al., 2024; Portuguese League, N = 20 players, 34 matches). For PlayerLoad (~120 AU), maximum acceleration and deceleration (~0.6 m/s2), and high-intensity acceleration and deceleration event counts (~15 and ~14, respectively), no single published source reports between-player match SDs directly; these values were estimated from the typical ranges and variability reported across multiple studies and should be interpreted with appropriate caution.

Standardization of GPS 3.0 metrics

For the ADI-derived metrics where no published between-player match SDs exist, a regression-based approach was used. Mechanical work was regressed on its traditional GPS constituent metrics (distance, HSR, maximum acceleration/deceleration, high-intensity event counts) using competitive match data from LOSC Lille (N = 56 player-match observations, 8 matches). The regression coefficients ($R^2 = 0.83$, distance as the dominant predictor) were combined with the published between-player SDs of the constituent metrics, accounting for inter-metric correlations, to project a match-level between-player SD for mechanical work (147 J). For mechanical power, the observed between-player SDs from the LOSC match data were used directly (15.15 W/kg for peak, 0.46 W/kg for mean). Directional sub-components were scaled proportionally from the total mechanical work projection based on their relative contribution in match data.

Systematic vs. random error decomposition

To determine whether between-unit differences are consistent across tasks (systematic) or fluctuate unpredictably (random), an ICC was computed for each metric using z-scores across the 12 tasks, treating units as subjects and tasks as conditions. The ICC quantifies the proportion of between-unit variance that is stable across tasks. Higher ICC values indicate a greater proportion of the between-unit variance is stable across tasks (i.e., systematic), while lower values indicate predominantly random random fluctuations.

Results

GPS 3.0 mechanical metrics (ADI)

Table 1 presents the between-unit agreement for ADI-derived GPS 3.0 metrics during the full session. Total mechanical work showed a CV of 0.39% and a trivial standardized error of 0.006. Mean mechanical power (CV = 0.38%, SE = 0.025) and peak mechanical power (CV = 3.52%, SE = 0.058) were similarly trivial. The aggregated linear component of mechanical work showed a trivial standardized error (SE = 0.016), while the directional components (cuts and arcs) showed small standardized errors (SE = 0.11-0.12). All six GPS 3.0 metrics showed purely random error structure (ICC $<$ 0.16), confirming that units can be freely interchanged without introducing systematic bias into any of these measures.

Table 1. Between-unit agreement for ADI GPS 3.0 metrics (Full Session, N = 10 units).

Metric	Mean	Unit SD	CV%	Ref SD	Std Err	Mag.	Range	Rng Std	ICC	Err. type	Ref
MechWork (J)	223	0.86	0.39	147	0.006	Trivial	2.7	0.019	0.01	Random	4
MW Linear (J)	135	1.47	1.08	89.3	0.016	Trivial	5.1	0.057	0.16	Random	6
MW Cuts (J)	25.8	1.80	6.99	17.0	0.106	Small	4.8	0.28	0.10	Random	6
MW Arcs (J)	16.6	1.27	7.66	11.0	0.116	Small	3.9	0.36	<0.01	Random	6
MechP mean (W/kg)	2.98	0.01	0.38	0.46	0.025	Trivial	0.04	0.08	0.01	Random	5
MechP max (W/kg)	25.0	0.88	3.52	15.2	0.058	Trivial	2.5	0.16	0.07	Random	5

Ref: 4 = Projected via regression from LOSC Lille match data (N = 56, R^2 = 0.83); 5 = Observed between-player SD from LOSC match data; 6 = Proportional scaling from total MechWork projection. Range = max minus min across 10 units. Range Std = range/reference between-player SD. Magnitude thresholds: <0.10 trivial, 0.10-0.30 small, 0.30-0.60 moderate.

Traditional GPS 2.0 metrics (WIMU native)

Table 2 presents the between-unit agreement for traditional GPS metrics during the full session. Distance (CV = 0.36%, SE = 0.006), HSR (CV = 1.92%, SE = 0.018), peak speed (CV = 0.21%, SE = 0.04), and PlayerLoad (CV = 10.1%, SE = 0.03) all showed trivial standardized errors. Maximum acceleration and deceleration showed moderate standardized errors (SE = 0.48), though this was observed only during tasks involving genuine maximal efforts. High-intensity event counts showed small standardized errors (SE = 0.13).

PlayerLoad provides a clear illustration of why standardized error is preferable to CV% for practical interpretation. Its CV of 10.1% would traditionally be classified as poor between-unit agreement, yet when expressed relative to the 120 AU between-player SD observed in elite matches, the measurement noise is trivial (SE = 0.03). Units can therefore be swapped without meaningful impact on PlayerLoad-based training load monitoring. Conversely, maximum acceleration shows a more modest CV of 4.7%, but the between-player match SD is only $\sim$0.6 m/s^2, making the standardized error moderate (SE = 0.48). These two metrics invert their priority ranking entirely depending on how the error is expressed.

Error was predominantly random for most metrics (ICC $<$ 0.15). PlayerLoad (ICC = 0.61) and high-intensity deceleration counts (ICC = 0.73) showed a degree of systematic consistency. PlayerLoad is the only metric in this study derived from the tri-axial accelerometer rather than GPS position data; the systematic pattern likely reflects unit-specific accelerometer sensitivity characteristics. For these specific metrics, maintaining fixed unit-player assignment may further optimize longitudinal data quality.

Satellite signal quality

All 10 units maintained excellent satellite conditions throughout the session (mean 28.6 ± 2.1 satellites, range 25.3-31.0; mean PDOP 0.957 ± 0.081, range 0.881-1.028). The uniformly high signal quality across all units confirms that the measurement conditions were optimal and equivalent for all devices. Between-unit differences in satellite count and PDOP showed no association with measurement deviation, suggesting that the small differences observed in this study reflect unit-specific hardware and signal processing characteristics rather than satellite geometry. It should be noted that Doha is not covered by a Satellite-Based Augmentation System (SBAS), which may result in slightly lower positional accuracy compared with regions where SBAS is available (e.g., Europe, North America). Between-unit agreement could therefore be even tighter in SBAS-supported locations.

Table 2. Between-unit agreement for traditional GPS 2.0 metrics (Full Session, N = 10 units).

Metric	Mean	Unit SD	CV%	Ref SD	Std Err	Mag.	Range	Rng Std	ICC	Err. type	Ref
Distance (m)	1661	6.0	0.36	990	0.006	Trivial	16.0	0.016	0.14	Random	1
HSR (m)	160	3.1	1.92	172	0.018	Trivial	10.3	0.060	0.23	Mixed	1
Peak speed (km/h)	25.1	0.05	0.21	1.4	0.04	Trivial	0.19	0.14	0.05	Random	2
Peak accel (m/s^2)	6.23	0.29	4.70	0.6	0.48	Moderate	0.90	1.49	0.03	Random	3
Peak decel (m/s^2)	7.61	0.29	3.75	0.6	0.48	Moderate	0.78	1.29	0.02	Random	3
Player Load (AU)	30.7	3.1	10.1	120	0.03	Trivial	9.7	0.08	0.61	Systematic	3
HI accel (n)	35	2.0	5.52	15	0.13	Small	6	0.40	0.11	Random	3
HI decel (n)	32	1.9	5.89	14	0.13	Small	5	0.36	0.73	Systematic	3

Ref: 1 = Collins et al. 2025 (MLS, N = 1243 matches); 2 = Silva et al. 2024 (Portuguese League); 3 = Estimated from multiple published sources. See Table 1 footnote for other abbreviations.

Discussion

The primary finding of this study is that GPS-derived mechanical work and mechanical power show trivial between-unit measurement error when standardized against match-level between-player variability. Total mechanical work (SE = 0.006), mean mechanical power (SE = 0.025), and peak mechanical power (SE = 0.058) can all be considered highly reliable between units. Furthermore, the error is purely random (ICC < 0.08), meaning that the small between-unit differences fluctuate unpredictably across tasks rather than reflecting consistent device-specific characteristics. Practitioners can therefore use these metrics with confidence, even when units are rotated between players across sessions.

GPS 3.0 mechanical metrics: first between-unit evidence

While the test-retest reliability of some of the ADI-derived mechanical metrics has been previously examined in field conditions (Buchheit et al., 2018), where CVs of 11% for vertical stiffness and 7% for the velocity-load/force-load ratio were reported, the present study provides the first independent assessment of their between-unit agreement. The distinction is important: test-retest reliability confounds device noise with biological variability (day-to-day fluctuations in an athlete's neuromuscular status, stride pattern, and effort level), whereas the trolley protocol isolates the pure between-device component by ensuring all units measure an identical mechanical input. The CV of 0.39% for total mechanical work highlights how much of the previously reported variability was biological rather than device-related, which is reassuring for practitioners considering the adoption of these metrics.

The directional sub-components of mechanical work (aggregated cuts and aggregated arcs) showed small standardized errors (SE 0.11-0.12) and random error structure. While their CVs (7.0-7.7%) are higher than the total mechanical work, this partly reflects the smaller absolute values of these sub-components and the sensitivity of movement phase classification to the precise GPS trajectory during curved and lateral movements. The physical separation between units on the trolley (∼10-20 cm) means that during turns and curves, units on opposite sides follow slightly different arcs with different centripetal accelerations, which may affect the classification of movement into cuts versus arcs. This is a design limitation of the trolley protocol rather than a device limitation, and the between-unit differences observed for these directional metrics likely overestimate what would be seen in body-worn conditions where all measurements come from a single device location.

Comparison with previous-generation GPS units

The present results demonstrate substantial progress in GPS between-unit agreement compared with previous-generation technology. Using a similar trolley-based protocol with 50 GPS units (15 Hz, GPSports), Buchheit et al. (2014) reported between-unit CV values ranging from 1% for peak speed to 56% for deceleration counts, with some units measuring 2 to 6 times more acceleration/deceleration events than others. Those findings led the authors to question the usefulness of acceleration-derived GPS indices.

In the present study, the CV for distance has decreased from ∼3-5% to 0.36% (a 4- to 10-fold improvement), and peak speed reliability has tightened from 1% to 0.21%. For high-intensity event counts, the improvement is equally notable: CV values of 5-6% compared with 15-56% previously. This progress likely reflects advances in GPS chipset technology (i.e., WIMU PRO EVO chipset), multi-constellation support (GPS, GLONASS, Galileo), and improved filtering algorithms (Rodriguez-Fernandez et al., 2025; Gomez-Carmona et al., 2019) rather than the sampling rate, as the current units operate at 10 Hz compared with 15 Hz in the earlier study. It is also worth noting that the present data were collected in Doha, which is not covered by a Satellite-Based Augmentation System (SBAS). The between-unit agreement observed here may therefore represent a conservative estimate compared with what could be expected in SBAS-supported regions such as Europe or North America.

Beyond CV%: the case for standardized error

A central methodological contribution of this study is the standardization of between-unit measurement error against the between-player variability observed in elite football matches. This reframes the reliability question from "how noisy is this metric?" to the more practically relevant "could this noise change my interpretation of a player's data?"

The PlayerLoad example illustrates this most clearly. A CV of 10% would traditionally be classified as poor between-unit agreement. Yet when expressed relative to the $\sim$120 AU between-player SD in elite matches, the measurement noise is trivial (SE = 0.03). A unit swap would shift a player's PlayerLoad by less than 3% of the typical range observed between players, far below the threshold for any meaningful misinterpretation.

This approach builds on the conceptual framework proposed by Hopkins (2000) for expressing reliability statistics relative to the smallest worthwhile change, as applied in previous work (Buchheit et al., 2018; Haugen and Buchheit, 2016), and extends it to the between-unit context by using published between-player match SDs as the reference denominator. For newer metrics such as mechanical work where direct published references do not yet exist, we have demonstrated a regression-based projection approach that derives the reference SD from the metric's constituent components.

Sources of between-unit differences

Despite all 10 units sharing an identical trajectory on the trolley, measurable between-unit differences persist. Several mechanisms likely contribute. First, even on a rigid trolley, the physical separation between units ($\sim$10-20 cm) means that during turns and curves, each unit follows a slightly different arc. For the central circle task (radius $\sim$9 m), a 10 cm lateral offset produces approximately a 1% path length difference, consistent with the observed CV of 0.58% for distance during that task. This geometric effect also explains the higher CVs for directional metrics, where even small differences in centripetal trajectory can influence the classification of movement into linear, cut, or arc phases.

Additionally, the position of each unit within the trolley may influence accelerometer-derived metrics such as PlayerLoad. Units placed at different locations on the rigid tray experience slightly different vibration patterns, resonance frequencies, and shock transmission depending on their proximity to the wheels, the pulling attachment, or the edges of the structure. These subtle differences in the mechanical environment could explain the systematic between-unit pattern observed for PlayerLoad (ICC = 0.61), as each unit's position on the trolley remained fixed across all 12 tasks.

Also, each GPS chipset independently computes its position fix using its own clock, antenna characteristics, and signal processing pipeline. Even with the same satellite constellation, two chipsets may weight individual satellites differently, apply different atmospheric correction models, or resolve carrier-phase ambiguities differently. This creates small differences in position estimation that propagate into speed and distance calculations.

Finally, all units maintained excellent satellite reception (25-31 satellites, PDOP 0.88-1.03), and the between-unit differences in satellite metrics did not predict measurement deviation. This confirms that in the favorable signal conditions of the present test, the observed differences reflect unit-specific hardware and processing characteristics rather than satellite geometry.

Error type: systematic vs. random

The decomposition of total between-unit error into systematic and random components has direct practical implications. When error is systematic, the recommendation is to maintain fixed unit-player assignment so that any persistent characteristics cancel when tracking within-player changes over time. When error is random, units can be freely interchanged, as the noise averages out over repeated sessions.

In the present study, the vast majority of metrics showed predominantly random error (ICC < 0.20), including all ADI GPS 3.0 mechanical metrics, distance, peak speed, and maximum acceleration/deceleration. Only PlayerLoad (ICC = 0.61) and high-intensity deceleration counts (ICC = 0.73) showed substantial systematic components. PlayerLoad is notably the only metric derived from the tri-axial accelerometer rather than GPS position data, suggesting that accelerometer-specific calibration characteristics drive the systematic pattern. For these specific metrics, maintaining fixed unit-player assignment may further optimize data quality over time.

A practically important perspective emerges from the worst-case range analysis. For most metrics, the standardized range between the lowest- and highest-reading units remained trivial to small (< 0.36), confirming that unit swaps have negligible practical impact. For peak acceleration, the standardized range was larger (1.49), but critically, this error is purely random (ICC = 0.03): the same unit will not produce the same bias on a different occasion. This means that while any single session may contain more noise for peak acceleration than for other metrics, this noise tends to cancel out over repeated sessions and does not accumulate into a systematic drift. Fixing unit-player assignment would not reduce this type of error.

Reporting the full session

The primary results in this study are reported for the full session ($\sim$16 min, $\sim$1660 m) rather than for individual tasks. This choice is deliberate. A full session that combines walking, jogging, high-speed running, sprints, changes of direction, and recovery periods provides a more ecologically valid estimate of between-unit error than any single isolated task. In practice, GPS data are typically analyzed over entire training sessions or matches, where the variety of movement types naturally dilutes task-specific artifacts. For example, peak deceleration showed a CV of 60% during a constant-speed lap where no genuine deceleration occurred, but only 0.63% during a dedicated acceleration/deceleration task. The full session, which includes both contexts, produced a CV of 3.75%, better representing the between-unit error practitioners would encounter in operational conditions. All 12 tasks were analyzed individually to explore task-specific patterns and are available for further exploration in the supplementary interactive report.

Limitations

Several limitations should be acknowledged. First, the trolley protocol eliminates the body-worn vibration and soft-tissue artifact that contribute to real-world measurement variability, isolating only the device-to-device error component. Second, the trolley setup introduced several constraints that may have inflated the observed between-unit variability for some metrics: the physical separation between units ($\sim$10-20 cm) meant each unit followed a slightly different arc during curves. These design artifacts suggest that the between-unit variability reported here, particularly for directional sub-components and accelerometer-derived metrics, likely overestimates what would be observed in body-worn conditions. Third, the reference between-player SDs used for standardization were sourced from published literature and may not perfectly represent all competition contexts. Fourth, the regression-based projection of match-level SDs for mechanical work was dominated by distance, and the projected SD may be conservative. Fifth, Doha is not covered by a Satellite-Based Augmentation System (SBAS), and the favorable satellite conditions (PDOP < 1.03 for all units) mean that the present results may represent a conservative estimate of between-unit agreement compared with SBAS-supported regions or environments with poorer reception."

Conclusions

ADI-derived mechanical work and mechanical power show trivial and purely random between-unit measurement error. Practitioners can trust these metrics for training load monitoring, return-to-performance assessment, and longitudinal player tracking, even when units are rotated between players. Traditional GPS metrics are similarly reliable. The present results using a WIMU PRO EVO chipset represent a substantial improvement over previous-generation GPS units. Standardizing measurement error against match-level between-player variability, rather than relying on CV% alone, provides a more actionable framework for evaluating GPS reliability and informing unit management decisions.

Practical applications

- Mechanical work and mechanical power show trivial between-unit error with no systematic component. Units can be freely rotated between players without compromising these metrics.
- Distance, HSR, peak speed, and PlayerLoad are similarly robust to unit swaps when standardized against match-level variability.
- For PlayerLoad and HI deceleration counts, maintaining fixed unit-player assignment may further optimize longitudinal tracking.
- For peak acceleration and deceleration, confirming that unit assignment has not changed between comparison sessions is advisable, particularly for return-to-performance decisions.
- CV% should not be used as the sole indicator of measurement quality. Standardizing against between-player match variability provides a more practically meaningful evaluation.
- Current-generation GPS technology has improved substantially over previous-generation units, with 4- to 10-fold reductions in between-unit variability for distance-based metrics.
- The present data were collected in a region without SBAS coverage. Between-unit agreement may be even tighter in SBAS-supported regions (e.g., Europe, North America).

Acknowledgements

The authors thank Archie Ogden for his assistance during data collection.

References

1. Buchheit M, Al Haddad H, Simpson BM, Palazzi D, Bourdon PC, Di Salvo V, Mendez-Villanueva A. Monitoring accelerations with GPS in football: time to slow down? Int J Sports Physiol Perform. 2014;9(3):442-445.

2. Buchheit M, Gray A, Morin JB. Assessing stride variables and vertical stiffness with GPS-embedded accelerometers: preliminary insights for the monitoring of neuromuscular fatigue on the field. J Sports Sci Med. 2015;14(4):698-701.

3. Buchheit M, Lacome M, Cholley Y, Simpson BM. Neuromuscular responses to conditioned soccer sessions assessed via GPS-embedded accelerometers: insights into tactical periodization. Int J Sports Physiol Perform. 2018;13(5):577-583.

4. Buchheit M, Lopez Sagarra A, Boskovic A, Komino P, Norman D, Hader K. GPS 3.0: from distance into zones toward better proxies of internal neuromuscular load in elite football. Sports Performance & Science Reports. 2026 Feb;280:v1.

5. Buchheit M, Simpson BM. Player-Tracking Technology: Half-Full or Half-Empty Glass? Int J Sports Physiol Perform. 2017 Apr;12(Suppl 2):S235-S241. doi: 10.1123/ijspp.2016-0499. Epub 2016 Dec 14. PMID: 27967285.

6. Collins JJ, Fernandez Navarro J, McRobert AP, Silvers-Granelli H, Malone S, Collins KD. The physical demands of Major League Soccer match-play with specific reference to high-intensity activity by position, venue and opposition quality. PLoS ONE. 2025;20(10):e0334460. doi:10.1371/journal.pone.0334460

7. Gomez-Carmona CD, Bastida-Castillo A, Garcia-Rubio J, Ibanez SJ, Pino-Ortega J. Published 2019, volume 233, pages 238-248.

8. Haugen T, Buchheit M. Sprint running performance monitoring: methodological and practical considerations. Sports Med. 2016;46(5):641-656.

9. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med. 2000;30(1):1-15.

10. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. Med Sci Sports Exerc. 2009;41(1):3-13.

11. Rodriguez-Fernandez A, Suarez-Iglesias D, Vaquera A, Leicht AS, Rodriguez-Marroyo JA. Inter-system and inter-unit reliability of Polar Team Pro and WIMU PRO devices during external load measurements indoors. Proc Inst Mech Eng Part P J Sports Eng Tech. 2025;239(3). doi:10.1177/17543371221147857

12. Silva H, Nakamura FY, Loturco I, Ribeiro J, Marcelino R. Analyzing soccer match sprint distances: a comparison of GPS-based absolute and relative thresholds. Biol Sport. 2024;41(3):223-230. doi:10.5114/biolsport.2024.133663