Technische Universität München
Fakultät für Sport- und Gesundheitswissenschaften

Lehrstuhl für Trainingswissenschaft und Sportinformatik

Validation of Electronic Performance Tracking Systems

Methodology, Design & Applications

Daniel M. Linke

Vollständiger Abdruck der von der Fakultät für Sport- und Gesundheitswissenschaften der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Philospohie (Dr. phil.)

genehmigten Dissertation.

Vorsitzender:
Prof. Dr. David Franklin

Prüfende der Dissertation:
1. Prof. Dr. Martin Lames
2. Prof. Dr. Arnold Baca
3. Prof. Jonathan Wheat, Ph.D.

Die Dissertation wurde am 23.01.19 bei der Technischen Universität München eingereicht und durch die Fakultät für Sport- und Gesundheitswissenschaften am 16.10.19 angenommen.

Sn lx e`lhkx `mc eqhdmcr vgn khrsdm vgdm H&l `mfqx+
gtf ld vgdm H&l r`c `mc k`tfg vhsg nq `s ld vgdmdudq onrrhakd-
Sg`mj you for being there!

Abstract

This dissertation is a compilation of four publications that seek to expand the scientific knowledge of performance analysis based on spatiotemporal position data from two different perspectives: (i) by developing and applying methods to assess the measurement accuracy of tracking technologies in team sports, and (ii) by applying innovative analysis methods that could potentially improve the interpretation of the time-dependent declines in team-sport specific match running activity.

Study One assessed the measurement accuracy of the most commonly used tracking technologies in professional team sports (i.e., video-based tracking technology (VBT), radar-based local positioning system (LPS), and global positioning system (GPS)). The position, speed, acceleration and distance measures of each technology were compared against simultaneously recorded measures of a ”gold standard” reference system (VICON). It was shown that most substantial differences between EPTS occurred at the spatial accuracy, whereas speed and acceleration errors of GPS were comparable to those of LPS. One crucial insight in this regard is the noticeably large error margin in the accuracy of key performance indicators KPIs that is independent of the respective system or technology, which we are still facing in EPTS in general. Especially in KPI categories with a high impact on practical decisions, such as high-speed performance in-

dicators, we found significant deviations from the gold standard, indicating that future research activities should aim at diminishing these inherent errors.

Study Two examined the impact of two anatomical reference points that are commonly used to represent the human body in space on kinematic tracking variables. EPTS traditionally rely on one of two body positions as the ultimate representative for the entire body in space: the center between the scapulae (COS) and the pelvic center (COP). Results showed that differences between COP and COS heavily depend on the underlying movement characteristic. Low-speed running showed the lowest deviations whereas accelerated movements and movements with sharp changes in direction lead to a significant increase in the observed differences. Results further showed that COS sprinting distance was on average -44.65% ($p < .001$) lower in comparison to COP. Similarly, the maximum speed obtained from COS was -2.94% ($p = .001$) lower in comparison to COP. On the contrary, maximum acceleration values of COS were on average 16.15% ($p = .02$) higher compared to COP. The study illustrates that the anatomical reference point used to represent the entire body in space needs to be carefully considered in the interpretation of tracking variables delivered by different EPTS.

Study Three applied an innovative approach to quantify the contribution of game interruptions to the fatigue-related declines in match running performance throughout a football match. Results showed a significant decline in effective playing time over the course of a match, from 66.3% of the total playing time in the first 15 minutes to 55.9% in the final 15 minutes of a match. Under consideration of the total playing time, match running performances decreased by 24.2%

on average; considering the effective playing time, they decreased on average by only 10.2%. It can, therefore, be concluded that more than half (57.9%) of the commonly reported decline in match running performance cannot be assigned to physical fatigue, but rather to an increase in game interruptions as the game progresses. This study demonstrated for the first time that the decline in players' match running performance during football matches is substantially amplified by a proven increase in game interruptions, indicating that there may be a tendency among practitioners to overestimate fatigue-induced performance declines.

Study Four described and analyzed the substitution tactics of an international field hockey team during competitive match play and identified the impact of bench periods on the physical output of players when re-entering active game-play. A local position measurement system recorded the physical performance of the players. An average of 58.0±4.6 substitutions was registered during a tournament match, where the average player performed 4.7±0.8 individual substitutions with an average pause duration of 5.4±1.2 min and an average on-field playing duration of 7.18±2.14 min. Within the first minute of substitution, re-entering players covered a significantly larger total distance (159.7±33.0 $m \cdot min^{-1}$) compared to the team average (139.4±33.3 $m \cdot min^{-1}$). A significant decrease in physical performance was observed within the fifth minute after substitution. These findings suggest a 'first-minute-rush-effect', wherein substitutes covered a significantly larger total distance compared with the team average. Further, re-entering field hockey players experienced distinguishable signs of fatigue within the fifth minute after substitution.

Contents

List of Figures

Acronyms

AOA	Angle of Arrival.
COM	Center of Mass.
COP	Center of Pelvis.
COS	Center of Shoulders.
EPTS	Electronic Performance Tracking Systems.
GDOP	Geometric Dilution of Precision.
GNSS	Global Navigation Satellite System.
GPS	Global Positioning System.
IMU	Inertial Measurement Unit.
KPI	Key Performance Indicator.
LPS	Local Positioning System.
RFID	Radio-frequency Identification.
RMSE	Root Mean Square Error.
SSG	Small-sided Game.
TDOA	Time Distance of Arrival.
TOF	Time of Flight.
VBT	Video-based Tracking.
WLAN	Wireless Local Area Network.

Publication list

The following publications are presented in support of this thesis:

I **Linke, D.**, Link, D., & Lames, M. (2018). Validation of electronic performance and tracking systems EPTS under field conditions. *PloS one, 13*(7), e0199519.

https://doi.org/10.1371/journal.pone.0199519

II **Linke, D.**, & Lames, M. (2018). Impact of sensor/reference position on player tracking variables: Center of scapulae vs center of pelvis. *Journal of Biomechanics , 83*(23), 319-323.

https://doi.org/10.1016/j.jbiomech.2018.11.046

III **Linke, D.**, Link, D., Weber, H., & Lames, M. (2018) Decline in match running performance in football is affected by an increase in game interruptions. *Journal of Sports Science & Medicine, 17*(4), 662-667.

https://www.jssm.org/hf.php?id=jssm-17-662.xml

IV **Linke, D.**, & Lames, M. (2016). Substitutions in elite male field hockey–a case study. *International Journal of Performance Analysis in Sport, 16*(3), 924-934.

https://doi.org/10.1080/24748668.2016.11868939

1 Introduction

1.1 Background

With a continuous increase in technological advances in athlete tracking technologies, player tracking has become one of the most critical components of load monitoring. Primarily used in professional team sports, electronic performance and tracking systems (EPTS) primarily track player (and ball) positions but can also be used in combination with inertial measurement units (IMUs) and heart-rate monitors to provide inertial load and other physiological information [19]. In particular, video-based tracking systems (VBT), radio frequency-based local positioning systems (LPS) and global positioning systems (GPS & GNSS) have become indispensable core technologies for assessing the physical demands of both training and competition. In fact, a recent survey on the current practices of high-level football clubs to monitor training load showed that of 41 clubs surveyed, 40 collected heart rate and GPS data for every player during every field training session [1].

Despite continuously improving the status of technological development at large, the accuracy evaluation of positional data that is obtained by EPTS permanently plays a crucial role in the delineation and characterization of physical performance and tactical behavior. To accurately interpret positional data, practitioners and scientists should first understand the limitations of the respective tracking device and its data. Therefore, valid and reliable information about the differences in quality between EPTS is highly warranted; yet it is hard to obtain. Excellent accuracy properties, as conventionally alleged by the manufacturers, often serve as a sales argument and should, therefore, be critically questioned by an independent appraiser.

Fundamentally, EPTS detect the kinematics of an athlete (i.e., momentary position, velocity, and acceleration). In order to asses the physical measurement accuracy of any EPTS, the kinematic variables of the EPTS are commonly compared against simultaneously recorded variables of a criterion reference. This rather straightforward validation approach is known as *criterion-based validity*, where the kinematic variables are validated against some gold standard measurement that has been accepted as a measure of the concept of interest [51].

While an athlete’s kinematic variables are ’directly observable’ and, in theory, easy to validate (e.g., spatial deviation from the ’true’ position in a coordinate system), sports practice almost never interprets the raw biomechanical data, but instead tends to derive ’performance indicators’, with some being referred to as ’key performance indicators’ (KPI). Examples include the total distance covered, number of accelerations or high-intensity efforts. More importantly, most performance analyses that are used routinely by sports scientists are based on the premise that the respective performance characteristic is not directly measurable as a variable, but can only be thought of as hypothetical constructs (e.g., fatigue-resistance or tactical intelligence) [4].

In the framework of this dissertation, this concept refers to as *construct validity* – the validity of some construct used to represent a property that is not directly observable. One of the most discussed theoretical frameworks derived from time–motion analysis is the construct of fatigue. Among mental and tactical factors encompassing fatigue, particular attention was paid to the element of physical fatigue [53], leading to a large body of published time-motion analyses of the match running performance in top-level football [15, 24, 25, 55]. This has provided comprehensive evidence of time-dependent declines in distances covered by players over the course of play, indicating an inability to consistently perform at the same level even at the highest standards of contemporary competition [17].

Despite the increasing availability of time-motion research, doubts have been cast as to the real importance of physical performance in relation to success and the extent to which time–motion metrics can be used to determine whether players experience fatigue during matchplay [17]. One particular concern with many of the studies that have reported the existence of transient or end-game fatigue in match-play is the failure to account for confounding factors such as temporal changes in a team's tactics, ball possession, the scoreline, and game interruptions. A further commonly neglected but critical feature is the change in the time the ball is in play throughout competition [40], which underlines the need for new methods that could potentially improve interpretation of the time-dependent declines in physical activity.

1.2 Motivation

Summarizing the above, the distinction between EPTS research involving constructs (e.g., fatigue) and variables (e.g., covered distance) inevitably has further implications for the validation approach undertaken in this thesis. In particular, it is apparent that a comprehensive EPTS validation approach comprises both the validation of the technological measurement accuracy (accuracy and precision of the raw biomechanical data) and the validation of hypothetical performance analysis constructs derived therefrom.

Due to these reasons, this thesis seeks to expand the sport-scientific knowledge of performance analysis based on spatiotemporal tracking data from two perspectives: (i) by developing and applying a method to assess the measurement accuracy of tracking technologies in team sports, and (ii) by applying innovative analysis methods that could potentially improve the interpretation of the time-dependent declines in team-sport specific match running activity.

2 Integration Into the Context of Sports Science

The thematic area of this dissertation broadly allocates to the context of *classical performance analysis*, which includes all research that involves the analysis of actual sports performance in training or competition [35]. The main reason for doing performance analysis is to use theoretical considerations, or practical experiences, to identify a set of abilities and skills essential to sports performance [41]. Within this framework, there exists an overlap between performance analysis and other disciplines such as technical, physical and psychological aspects, that are often being investigated within performance analysis investigations [51]. In a subsequent step, operational definitions for these traits are developed and empirical measurements are taken to assess the level of these traits, which may happen on the level of performance prerequisites (e.g., biomechanical or physiological measurements) or the level of competition behavior (e.g., fatigue-related declines in match running performance) [41].

Classical performance analysis can be further sub-categorized into a theoretical aim and a practical aim [44]. *Theoretical performance analysis* seeks to generate models whose empirical foundations provide useful information for sports practice by identifying the performance behaviors that are important for a given sports [41]. The empirical data on which these models are predicated may then be used to provide statistical norms for performance measures (in the case of this dissertation, e.g., norms for accuracy differences in player tracking technologies or differences in variables delivered by different anatomical reference points (*Study One and Two*)) or for necessary levels of performance prerequisites (e.g., norms for fatigue-related performance declines for soccer players (*Study Three*)) [41].

On the contrary, *practical performance analysis* aims at a coupling of sports performance with individual training processes [41]. Thus, sports performance is analyzed with the goal to identify useful information for the training process. Against this background, *Study Four* could be assigned to practical performance analysis, e.g., by identifying fatigue-related performance declines of a single team, and by allowing recommendations for the resulting training routines.

A similar approach to defining sports performance research was proposed by Atkinson & Nevill [5], who distinguish between '*basic*' and '*applied*' research questions. Basic, or theory-driven research, is designed to corroborate or discount theories of the underlying mechanisms of a particular phenomenon - a process usually involved in modeling physiological or psychological mechanisms based on classical hypothesis testing procedures [5].

Applied research, on the other hand, wishes to investigate factors affecting variables in a ’real-world’ setting. Atkinson & Nevill [5] argue that some sports performance research has been overly concerned with physiological predictors of performance (theory-driven research), at the expense of not providing a valid and reliable description of the exact nature of the task in question (applied research). Similarly, Drust and Green [26] state that although there is excellent progress in performance analysis research and dissemination in sports practice, the data available is frequently descriptive in nature and of little impact in the ’real world’.

They further highlight an insufficient distinction between basic and applied research and advocate a research model specific to applied problems in sports science. In line with the above, a review by Mackenzie and Cushion [46] has raised methodological concerns with the current positivist and key performance indicator driven research that has focused on attempting to predict successful future performance, despite the inherent problems associated with investigating a multifaceted and often uncontrollable phenomenon. They further note that there hasn't been enough research on performance analysis in practice and mention that a naturalistic and qualitative framework would be more appropriate.

The above-mentioned distinction between theoretical (theory-driven) and practical (applied) performance analysis as two different research areas is also known as the so-called *practical impact debate*. More specifically, the practical impact debate suggests that researchers should consider providing stronger rationales for conducting their research, given its integral role in the coaching process, and its potential to further our understanding of performance and impact professional practice [46].

While there is no doubt that the practical impact debate could benefit from a clear distinction between theoretical (theory-driven) and practical (applied) performance analysis, it would create a wrong impression if these two areas were perceived as being independent. Accordingly, this thesis attempts to fill the gap between theoretical and practical performance analysis by covering both a theory-driven approach to fatigue-related research in professional soccer *(Study Three)* and an applied approach to fatigue-related substitution tactics of a professional field hockey team *(Study Four)*.

3 Position Detection Systems in Sports

This chapter provides an overview of the different technologies in use for position detection in game sports. The focus here is mainly on technological details of raw data acquisition as well as signal processing concepts that are relevant for practical decisions in EPTS performance analysis.

In general, position detection technologies can be split according to the target setting (outdoors vs. indoors vs. mixed), the underlying technology (VBT vs. LPS vs. GNSS), the degree of automation (automatic vs. semi-automatic), the spatial resolution, and the center of origin (room centered vs. user-centered). In addition to these categories, the available technologies may vary regarding sampling rate, spatial resolution, accuracy, the degree of the dependency of the accuracy on current conditions, e.g., position of the athlete, and instantaneous dynamics of the athlete [43]. The following section provides an overview of the basic functioning of the most commonly used tracking systems in sports, categorized based on the type of technology they use to determine location.

The following section provides an overview of the basic functioning of the most commonly used tracking systems in sports, categorized based on the type of technology they use to determine location.

3.1 Wearables

Wearable position detection systems require each player to wear a receiver or transmitter, contrary to optical (video)-based systems, which do not require players to have any equipment attached to them.

3.1.1 Global Navigation Satellite Systems (GNSS)

The most prevalent global navigation satellite system (GNSS) is known as the Global Positioning System (GPS), developed initially in 1973 for the United States military and operated using 31 satellites orbiting the Earth [70]. The recent development of portable GNSS units has permitted broader application of this technology in a variety of settings, including sports, thus providing an additional means for describing and understanding the spatial context of physical activity [21].

Each of the dedicated GNSS satellites contains an atomic clock that transmits radio signals containing information about the exact time to the GNSS receiver [42]. The time taken for the signal to travel from the satellite to the receiver is calculated by comparing the time of signal transmission to the time of arrival. The distance of the satellite from the GNSS receiver can then be calculated by multiplying the signal travel time by the speed of light [42]. However, GNSS signals that are sent to receivers do not directly contain the position of the receiver. Instead, the systems track individual receivers using the principle of triangulation. In order for triangulation to work, each satellite must send a signal with a precise time-stamp. The receiver then calculates the time it took for the signal to reach it (also known as Time-of-Flight measurement TOF) to gather information about its position. If the receiver has received signals from more than three satellites, a precise location can be calculated, thus the term triangulation. Finally, a signal from a fourth satellite is needed to calculate the accurate timestamp, since GNSS receivers are not equipped with atomic clocks that can stay in sync with the satellites. Therefore, to acquire the accurate position of the GNSS receiver, a minimum of four satellites need to be in communication with the receiver. The position can then be trigonometrically determined [42].

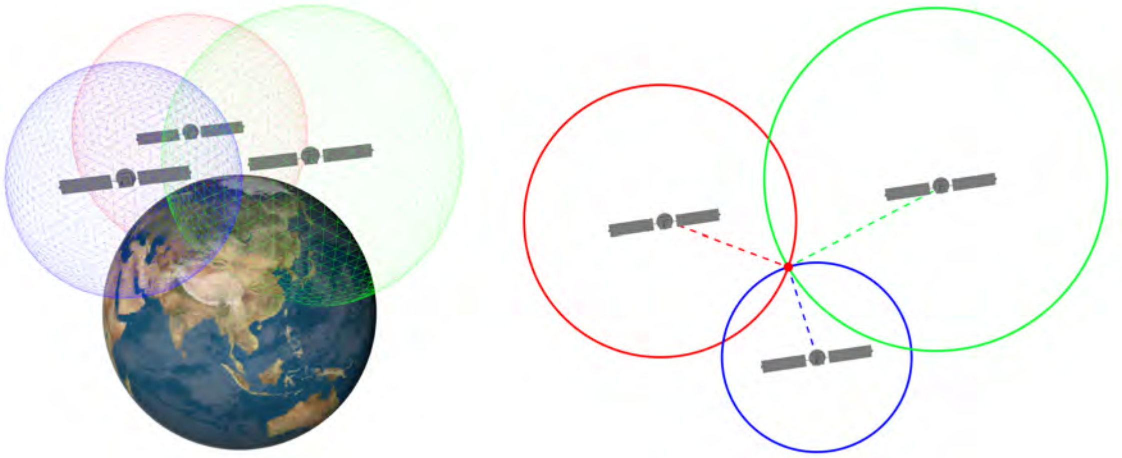


Figure 3.1: Visualisation of the trilateration Principle. Using three spheres/circles, trilateration can pinpoint a precise location. Each satellite is at the center of a sphere/circle and where they all intersect is the position of the GNSS receiver.

GPS technology was released for civilian use in 1983, however, the US Department of Defense applied an intentional 'error' measurement to the civilian satellite transmission (known as Selective Availability), to limit hostile forces using the system [64]. As a consequence, differential GPS was developed in order to reduce the errors associated with Selective Availability. This requires a stationary receiver at a known and calculated position allowing the GPS signal given by the satellite to be compared to a known point, thus establishing the corrective error of the signal [70].

In addition to the United State's GPS (31 satellites), Russia's Global Navigation Satellite System GLONASS (24 satellites), the European Union's Galileo Positioning System GALILEO (30 satellites), as well as China's BeiDou Navigation Satellite System can be used by the public for geospatial positioning. One important quality feature of sport-specific GNSS devices is the ability to receive signals from different GNSS in order to increase the number of available satellites for position detection.

It should further be mentioned that GNSS devices derive distance and velocity via two different methods - positional differentiation or Doppler shift. The GNSS devices calculate position (latitude and longitude) using information of the distance of each satellite to the device and then triangulating the devices location [47]. Subsequently, distance is calculated via positional differentiation (change in location over time), from which velocity can be derived (distance over time). Today's standard GNSS devices are typically accurate to within a 4.9 m radius under open sky [72]. Therefore, sport-specific GNSS systems typically use differential GPS that requires an additional receiver with a fixed and known position typically close to the pitch. With the help of the signals of this stationary receiver, one is able to clean the signal and get a better spatial resolution down to some centimeters. Velocity can also be calculated by measuring the change in frequency of the satellite emitted periodic signal (Doppler shift) [47], which provides an instant measure of velocity from which distance can be derived (velocity multiplied by time). Velocity calculated via Doppler-shift has shown a higher level of precision and less error compared with velocity calculated via positional differentiation during linear running at a range of velocities for 1-Hz GPS devices [70].

The determining factor for the potential measurement quality of GNSS systems is composed of the number of connected/available satellites and the strength of the current satellite configuration, known as dilution of precision (DOP), or geometric dilution of precision (GDOP).

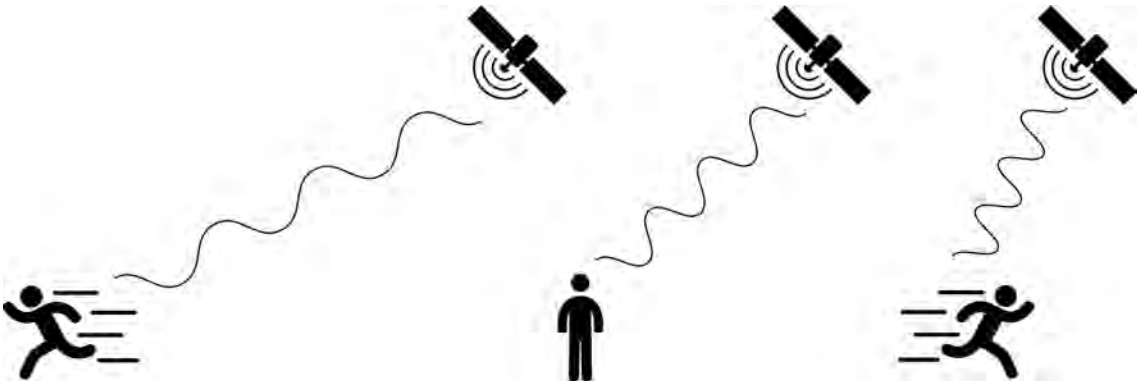


Figure 3.2: Doppler Effect: as the athlete is moving away from the satellite, the frequency is stretched relative to the observer. When moving towards the satellite, the wavelength shrinks and increases the frequency.

GNSS performance is hindered by anything that prevents the receiver from getting the satellite signals properly. Electromagnetic waves cannot penetrate through metal or concrete structures. Therefore, GNSS position detection is not suited for indoor events. Although there are some technical remedies underway, this remains a substantial limitation for tracking sports events. This is unaffected by the option of collecting data from GNSS sensors equipped with IMUs, that provide indoor data, also. When using GNSS outdoors, there are problems caused by so-called 'multipath effects' (reflected signals from buildings, metal constructions). Optimum conditions consist of a flat plane with an unobstructed view to the sky and no metal or concrete structures around. One may assume, for example, that GNSS works better on a training pitch than in a football stadium.

A significant advantage of GNSS applications in sports is its ready-to-use property. As soon as receivers are attached and switched on, the systems provide data. Also, GNSS is comparatively cheap, although costs may increase with increasing accuracy demands.

3.1.2 Local Positioning Systems (LPS)

Several different radio frequency-based technologies have been developed for local/indoor position detection systems LPS. These technologies include infrared, ultra-sound, radio-frequency identification, wireless local area network, bluetooth, and ultra wideband [32]. The most commonly used LPS sports-tracking technologies are ultra-wide-band systems (e.g., Ubisense's Real-Time Location System RTLS, Ubisense, Cambridge, UK; Inmotio's local position measurement system LPM, Abatec AG, Regau, Austria; ChyronHego's ZXY Arena, London, UK; Kinexon, Munich, Germany; RedFIR, Fraunhofer IIS, Nuremberg, Germany). Similar to GNSS, LPS systems require players to wear technical equipment. Contrary to passive GNSS receivers, local static base stations at known locations (also referred to as anchor nodes) receive electromagnetic waves that are emitted by active transmitters (also referred to as nodes).

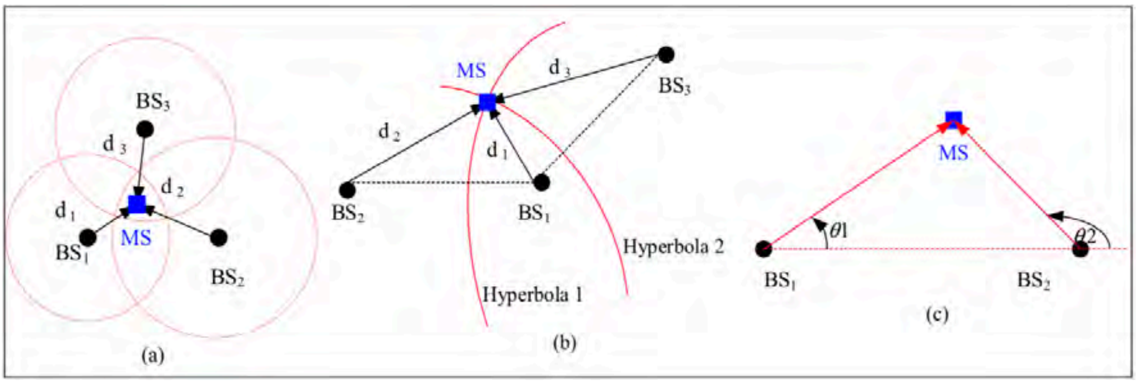


Figure 3.3: Position determination techniques: (a) time-of-flight (TOF); (b) time-difference of arrival (TDOA); (c) angle of arrival (AOA). Base Stations BS; Mobile Station MS (node/transmitter). [77]

The position of the player can be determined by calculating either the time-of-flight (TOF), time-difference of arrival (TDOA) and/or angle of arrival (AOA) of ultra-wide-band radio signals traveling from the transmitter to the base stations [67]. The TOA scheme estimates the player's location by measuring the arrival time of the radio signals coming from different anchor nodes, whiles the TDOA method measures the time difference between the arriving radio signals [77]. Using nonlinear regression, this information can be converted to the form of a hyperbola (see Fig. 3.3). Once enough hyperbolas have been calculated, the position of the target can be calculated by finding the intersection [31]. The AOA technique is conducted within the anchor nodes by observing the arriving angles of the signals coming from the player's node.

LPS systems typically have an overall sampling rate of up to 1 kHz (shared between the number of transmitters in use) and a static accuracy of approximately 0.1 m [43] whereas the dynamic accuracy ranges between 0.18-0.28 m [45, 61].

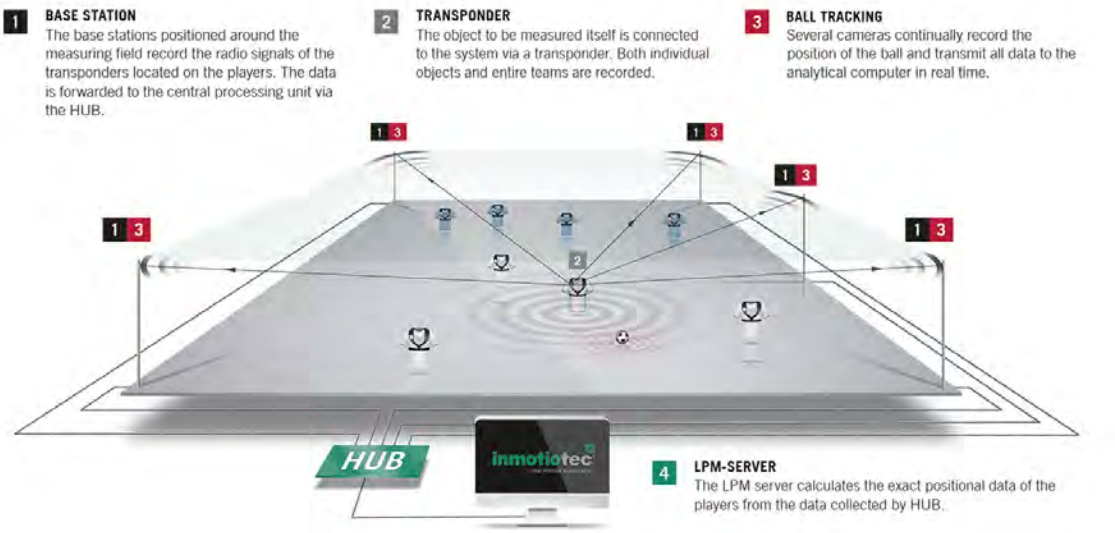


Figure 3.4: Exemplary arrangement of transponders, base stations, server, and the signal flow of Inmotio's LPM system.
Image retrieved from: https://www.inmotiotec.com/en/technology/inmotio-system

Systems that rely on distance measurements (or pseudo-distance measurements as is the case with Ubisense RTLS and Abatec LPM) between base stations and mobile devices also differ regarding the complexity of deployment, i.e., on-site installation of the base stations, calibration of the setup, and software configurations [43]. Depending on the system, this process can be rather complicated and time-consuming or somewhat straightforward and automated.

Despite superior accuracy, as indicated by recent studies [45, 66], the cost and effort of LPS installation is a significant drawback to wide-scale deployment in professional sports. Although the use of transponders removes the issue of occlusions, the specific surroundings of the pitch play a critical role in the data quality of radar-based systems. External interfering radio frequencies and high buildings near the pitch can distract the electromagnetic waves and decrease data quality. A player moving close to an advertising board or a pitch boundary is affected in the same way. Since the technology used is more demanding costs for LPS are typically higher by a factor of 2-3 compared to GNSS.

3.2 Video-Based Tracking (VBT)

As stated above, spatiotemporal motion capture can be achieved with wearable tracking devices attached to the player's body which is not always possible in sports domains, where the player movement can be affected, or technological equipment is not allowed [48]. As a solution, vision-based systems, which do not require players to have any equipment attached to them, allow the simultaneous tracking of all players, officials, and the ball through multiple cameras placed at fixed locations around the pitch. The cameras are located in a way to allow the entire playing surface to be recorded ensuring every player is visible for the total match duration [18]. SportVU (STATS LLC., Chicago, USA), Amisco (Nice, France), TRACAB (Stockholm, Sweden) and Venatrack (Slough, UK) are probably the best-known examples of vision-based tracking systems.

Detecting the position of players at a given moment in time is the first step in player tracking, and is also required in sports graphics systems for visualization of key moments of a game. Video-based techniques used to detect the position of players at a given moment in time range from those relying on a human operator to click on the feet of players in a calibrated camera image to automated techniques that use sophisticated algorithmic pattern detection techniques to identify regions that are likely to correspond to players [68].

Typically, multiple high definition cameras track players, which requires identification at the start of the tracking process and may need re-identification if the automatic tracking fails, usually through player occlusion [8]. These video-based systems are, therefore, classed as semi-automatic. The vision from each camera is simultaneously relayed to dedicated servers and converted to high-quality video files using proprietary software [71]. Sampling at 10-25 Hertz (Hz), the player trajectory is determined as x and y coordinates, usually measured in meters from the center spot (origin).

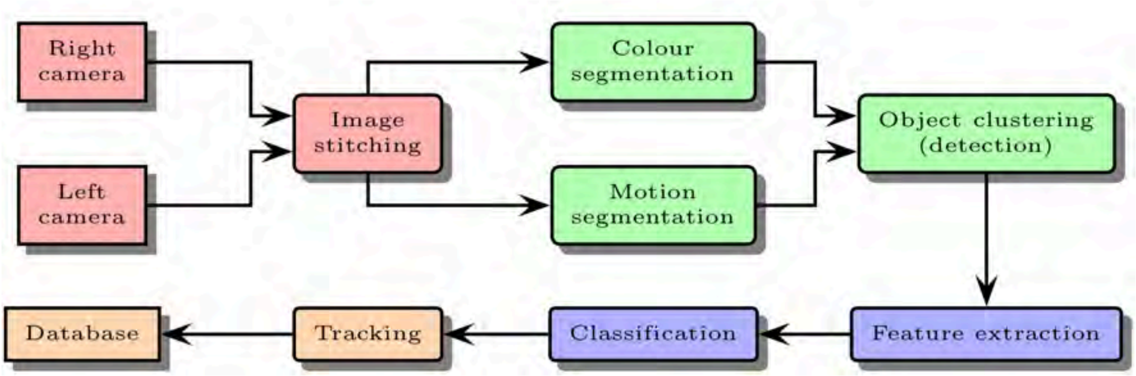


Figure 3.5: System chart of an exemplary image processing pipeline, using the example of a two-camera system. Colors indicate independent module topics [63].

While there are numerous optical-tracking manufacturers with just as many different hardware and software specifications, most of the optical-tracking systems rely on more or less the same production pipeline (see Fig.3.5). Typically, the basis of all later processing steps is a panoramic image composed of up to fifteen cameras with different viewing angles (see Fig.3.6). [37]. To translate image positions into field coordinates, the cameras and their pose are calibrated in advance. Compensation of radial lens distortion is applied within the stitching process and allows for a linear mapping (homography) from image to world coordinates [63, 78]. The homographies are then estimated from at least four point correspondences per camera (e.g., the corner points and the end points of the center line within the distortion corrected image). In a next step, a Direct Linear Transformation Algorithm is used to compute the transformations for the image in formations, which allow projecting image points from the panoramic coordinate systems (undistorted) to field coordinates [63]. Further pre-processing steps may include playfield detection, player detection, occlusion resolution, and appearance modeling [48].

Figure 3.6: Three different examples of semiautomatic optical tracking systems. *Left:* STATS's SportVU and Pixellot camera system. *Right:* ChyronHego's Tracab Gen 4 system in a weatherproof casing.

Semi-automated video tracking is a popular technique for assessing player movements. However, there are still limitations. An operator is still required, and subsequent data analysis is a time-consuming process often requiring a 12-48 hour turnaround. This delay limits the practical application of this technology in providing coaches and practitioners with information for post-match recovery and training sessions. Also, the nature and location of the array of cameras required to obtain the data mean these systems are not portable, limiting the analysis to games played at suitably equipped stadiums or training pitches [1].

In general, optical player detection and tracking are quite challenging due to many difficulties, such as similar appearance of players, complex interactions and severe occlusions, unconstrained outdoor environment, changing background, varying number of players with unpredictable movements, calibration inaccuracy, lack of pixel resolution especially of distant players or small objects (ball), clutter and motion blur [48]. Accordingly, VBT tends to rely on semi-automated approaches, with operators helping to initialize and correct the tracking data. Consequently, fully-automated tracking and labeling of players remains an open challenge.

Finally, the installation and use of these systems is expensive and often incur a service charge for the analysis of match data. Therefore, the systems are predominantly used by clubs in the top professional soccer leagues and are not accessible to moderate or amateur level teams.

4 Methodology

Conducting EPTS validation studies requires several methodological decisions, which can be made more or less appropriately. These decisions depend on the intention or purpose of the validation (e.g., system optimization, system comparison, demonstrating benchmarks), and thus are by no means trivial.

In order to arrive at a comprehensive accuracy testing of EPTS, several different methodological problems need to be solved. Taken together, the considerations to these problems constitute the methodological standards for validation studies of EPTS. The most relevant aspects are mentioned in this chapter to facilitate a critical reception of EPTS validation studies.

4.1 Gold Standards for Position Tracking in Sports

The most commonly used criterion methods include predefined circuits with known spatial arrangements (to evaluate distance measurement accuracy) [16, 30, 38, 56], timing gates (to evaluate sprint times and average speeds) [16, 30, 38, 57, 71], and radar/laser-based speed measurements for the evaluation of instantaneous running speed [54, 66, 73]. However, all the methods that have been mentioned have specific drawbacks, which should not be underestimated. First, it should be critically questioned whether a human test subject is capable of completing a predefined circuit without deviation from the previously calculated course (e.g., errors introduced by shortening/lengthening of movement trajectories during curved runs or simply due to the body sway). For a visualization of the deviation between pre-calculated and actual distance covered see Fig 4.1 and supplementary material *Video 1*.

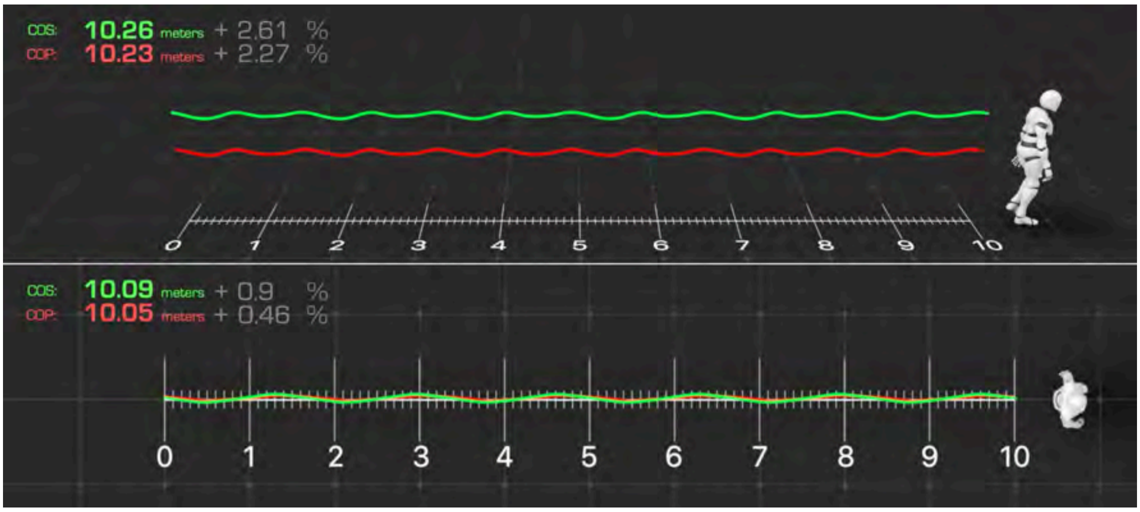


Figure 4.1: Visualization of the percentage deviation between pre-calculated (10 m) and actual distance covered by the center of pelvis COP and center of shoulders COS. Analysis conducted in 3D (top) and 2D (bottom). Image retrieved from *Video 1*.

Distance references that are based on predefined running circuits are inevitably susceptible to errors that are introduced by the unsteady nature of human motion. Second, although they are widely accepted as a high precision criterion measure for time measurements, timing gates are of limited use only as a speed reference [7], the reason for this being that the calculated speed that is derived from the known distance and time between the timing gates can only be considered as a mean speed value, rather than peak or instantaneous speed measures. Third, while radar/laser guns are capable of measuring the instantaneous speed of an object with high accuracy, they are suitable only when it comes to validate linear movements. The question here is not only whether a human subject can ever move in a perfectly straight line (see *Video 1*), but more importantly, whether such linear movements are representative of the arbitrary multi-directional motion sequences in team sports. Therefore, the actual positional data obtained by EPTS should ideally be compared with the instantaneous positional and speed data of a two or three-dimensional reference system with known error estimates [52].

4.2 Terminology

4.2.1 Position Data

In the context of team sports, position data can be divided into two subcategories: (i) object trajectories that represent the movement of players or the ball (spatiotenporal data), and (ii) event data that contains the location and time of match events, such as passes, shots on target or fouls.

Spatiotemporal and event data are complementary in that they describe different aspects of play (physical vs. tactical), and can provide a richer explanation of the game when used in combination [33]. This thesis, however, mainly covers the aspect of the validity of spatiotemporal data, the defining characteristic of which is that it is a sequence of samples containing the time-stamp and location of a player in the horizontal plane (XY-coordinates) [33].

4.2.2 Accuracy and Precision

The critical metric for evaluating an EPTS is the accuracy, defined as how much the reported location deviates from the actual location. The accuracy is denoted by an accuracy value and precision value (e.g., 20 cm accuracy over 95% of the time). The accuracy of an EPTS is often used to determine whether the chosen system is appropriate for a specific application [50]. Precision, on the other hand, indicates how often we expect to get at least the given accuracy (see Fig. 6.2). For example, GNSS-based EPTS apply to a broad spectrum of applications, but their spatial accuracy has shown to be highly variable. In other words, the distance between the true location and the reported location may fluctuate, sometimes considerably [50].

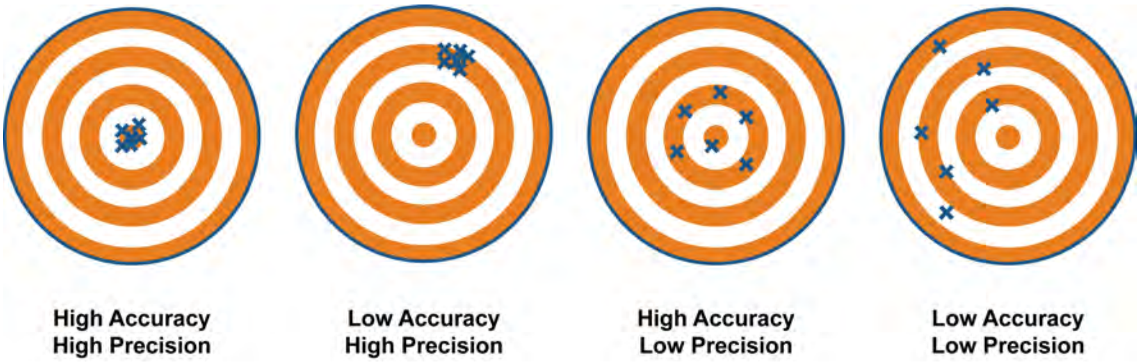


Figure 4.2: Dart-board metaphor of the concept of accuracy and precision.

4.2.3 Relative and Absolute Position Accuracy

In the framework of EPTS validations, it is essential to distinguish between two different types of accuracy definitions: relative accuracy and absolute accuracy.

Relative Accuracy is defined as the measure of how position estimations are positioned relative to each other. For example, the distance measurement between two positions delivered by an EPTS and the actual distance measurement between the two positions in the 'real' world must be within certain relative accuracy. In other words, relative accuracy is independent of coordinate translation.

In contrast, absolute accuracy defines how close a measured value is to a known absolute reference frame such as Universal Transverse Mercator (UTM) coordinate system. To demonstrate the concept of absolute and relative accuracy, Figure 4.3 illustrates three hypothetical accuracy scenarios.

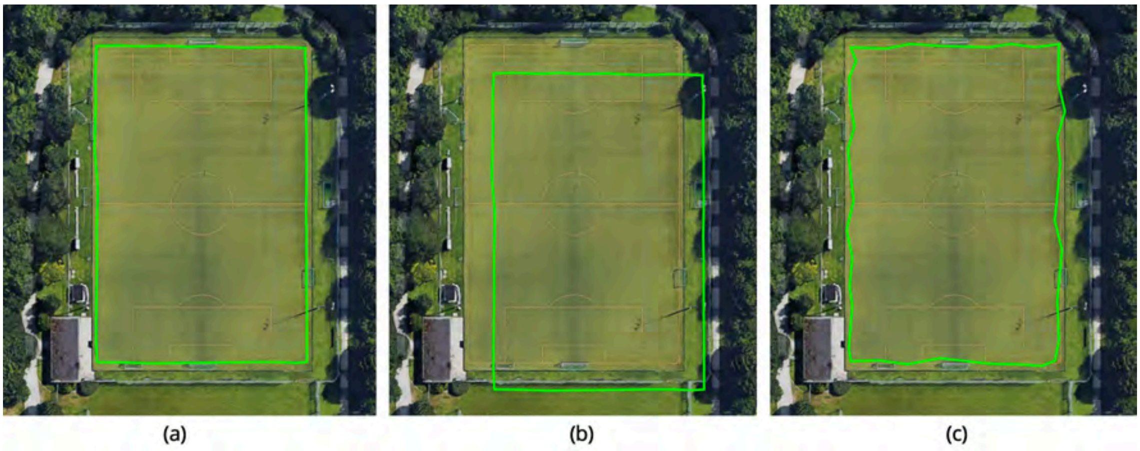


Figure 4.3: Demonstration of absolute and relative accuracy. In this hypothetical example, a GNSS unit is worn by a subject that walks along the sidelines of a soccer pitch. **a:** good absolute and relative accuracy. **b:** poor absolute accuracy / good relative accuracy. **c:** good absolute accuracy / poor relative accuracy.

In the hypothetical example shown in Figure 4.3 (b), the GNSS system could have an absolute accuracy of only 15 m. The covered distance, however, could be accurate to within centimeters. This raises the question of whether absolute and relative accuracy should be equally weighted in the context of EPTS validations. Ideally, a tested EPTS should demonstrate both absolute and relative accuracy requirements, especially if the coordinates are used to analyze and visualize tactical behavior. However, it could be argued that the accuracy of derived KPIs is independent of the absolute accuracy.

4.2.4 Validity and Reliability

The central aspect of EPTS accuracy assessments is the concept of *validity* - the extent to which an instrument (such as an EPTS) accurately measures the elements that it is supposed to measure [6, 69]. Validity can be further categorized into (i) logical validity, (ii) criterion validity, and (iii) construct validity. Criterion validity allows for an objective measure of validity, and there are two types of criterion validity: concurrent and predictive criterion validity [69]. Concurrent validity concerns the agreement between the observed value and the true or criterion value of a measure [34]. In the case of EPTS validations, this means that the delivered tracking variables correlate with the tracking variables delivered by a criterion reference technology with superior accuracy. In the case of this dissertation, EPTS are supposed to measure the positions and movements of players (and ball) during a match and to provide reliable KPIs. In other words, validity concerning position measurements represents the degree of agreement between a test result and an accepted reference value (criterion standard or gold standard), which is also referred to as accuracy [29].

Some authors argue that sports performance is not a ’variable’ in the classical sense, but rather a ’construct’ [4, 22]. Construct validity refers to the degree in which a protocol measures a hypothetical construct, for example, fatigue assessed with declines in running performance.

Reliability is the measurement criterion that denotes the precision of a measurement. Concerning EPTS measurements, reliability is typically assessed using measures of intra- and inter-observer agreement, which comprises not only the consistency (or stability) of the measure but also the consistency (or stability) of the entity that is being measured [41]. Atkinson and Nevill [6] have reported two types of reliability or reproducibility: absolute reliability (the degree to which repeated measurements vary for individuals) and relative reliability (the degree to which individuals maintain their position in a sample with repeated measurements). When tests are used to discriminate among individuals (cross-sectional assessment), parameters of relative reliability should be used (e.g., intraclass correlation coefficient, ICC) [36]. On the other hand, parameters of absolute reliability (e.g., standard error of measurement, SEM) are required for evaluative tests to monitor changes over time (longitudinal assessment) [23]. The *test-retest reliability* of an EPTS can be described by the ability of the EPTS to consistently provide the same measure [59]. As a consequence, EPTS measures can be reliable without being valid. However, for it to be considered valid, it must be reliable [9]. A sports scientist or practitioner may use EPTS variables to monitor the changes in the physical performance of an individual or compare activity profiles between teams. Therefore it is vital that the EPTS is valid and reliable so true changes can be identified [27].

4.2.5 Criterion Reference

A customary approach to asses the validity of EPTS is to determine the strength of the relationship between the output of a particular EPTS and an independent criterion that is accepted as a standard against which the EPTS could be judged, which is also referred to as a criterion-based validity concept [20]. The notion of a diagnostic reference standard or 'gold standard' pertains to the best available method for establishing the presence or absence of a condition of interest [13]. There are, however, no commonly accepted accuracy guidelines for a criterion standard to be valid as a suitable reference method, nor are there recommendations for acceptable measurement errors in the domain of EPTS [65]. Accordingly, the definition of a gold standard method remains to a large extent subject to interpretation. The International Organization for Standardization (ISO) defines a criterion standard or gold standard as a value or measuring method that serves as an agreed-upon reference for comparison, which is derived as a theoretical or established value, based on scientific principles [29]. Versi [74] further states that a gold standard is not necessarily the perfect test but merely the best available test under reasonable conditions. In the case of EPTS validations, one could argue that an appropriate gold standard should be able to detect fundamental differences in the EPTS' measurement accuracy. To make this possible, a 'true' EPTS gold standard should have a proven accuracy with measurement errors of at least one order of magnitude smaller than the EPTS to be validated. To elaborate, if we assume that, for example, there is an error of a few decimeters in a specific EPTS, the gold standard method of choice should have a proven accuracy of a few centimeters. This allows to draw valid conclusions regarding the absolute measurement accuracy of individual EPTS and also on accuracy differences between EPTS.

In addition to the aspect of measurement accuracy, the gold standard procedure should also be taken into account, as different movements may require different kinds of gold standards. For instance, timing gates could provide a reliable gold standard reference for average velocities on predefined distances, whereas the state-of-the-art systems for soccer-specific movements are represented by kinematic 3D motion capture systems [45]. These systems are commonly used in the field of biomechanics and are widely accepted as the gold standard for human movement analysis [28, 52].

Although modern 3D motion-capture systems generally satisfy the accuracy and precision requirements quite easily, problems could arise from the fact that these systems are designed as laboratory devices that require specific light conditions, operate on a limited measurement volume, and require elaborate calibration when used as a mobile device.

Unfortunately, a gold standard with sub-centimeter accuracy that covers the entire pitch of most team sports does not yet exist or, more explicitly, would require an unreasonable amount of resources. For the time being, EPTS accuracy studies employing a gold standard are limited to smaller measurement volumes than complete pitches. Nevertheless, the aim should be to apply measurement volumes that provide, to some extent at least, the option to validate realistic match-like movements. During several pilot studies, it became apparent that the available VICON equipment allowed a well-covered measurement area of 30x30m. The question arises whether this volume enables players to reach their full performance potential. In the case of professional football, 95% of sprints are covered with a distance of fewer than 30m [3]. It can, therefore, be expected that the applied 30x30m volume allows simulating the full spectrum of football-specific movement intensities.

4.3 Types of Validation Studies

A look at previous research shows that different types of EPTS validation studies can be distinguished: (i) comparative studies between different EPTS without an agreed upon reference (the interest lies in the sole difference between the measurements rather than the absolute accuracy), and (ii) accuracy studies with an agreed upon reference (all analyzed EPTS can be compared to the 'truth' by using a superior reference method).

It should also be emphasized that scientifically approved validation studies should ideally be conducted by an independent third party, rather than by the manufacturer of the EPTS itself. In this context, due consideration should also be given to the party to whom the validation results could be of interest (customer, manufacturer, or scientific community).

4.4 Measurement Site

The measurement site should provide optimum conditions for the tested systems as well as for the gold standard. A second criterion is the ecological validity - EPTS should be scrutinized under conditions they are usually used in for practical application. These claims create problems when EPTS using different technologies are to be compared. The fact that each system has its specific optimum site (indoor, outdoor, stadium) leads to compromises required when all systems are to be tested simultaneously and - consequential - in one place. A mandatory condition for GNSS measurements is a direct line of sight with as many satellites as possible to ensure minimum horizontal dilution of precision (HDoP) [76]. Satellites that cannot be acquired due to landscape feature obstructions weakens and increases the HDoP. Therefore, stadiums with roofed stands near the pitch, which are the typical features of modern football stadiums, are unsuitable. Indoor venues might not be covered by the GNSS at all.

LPS systems demand evenly distributed antennas/anchors mounted around the pitch [67]. This requires sufficient space all around the measurement area, which could be a problem for indoor studies. Electromagnetic radiation may be perturbed by the absorbing/reflecting properties of the environment. Thus, experience has demonstrated that metal constructions such as fences, tribunes, or benches are likely to compromise the transmission quality.

Image-based tracking systems require to be mounted at a sufficient height and distance from the pitch to provide an optimal viewing angle, [48] thereby reducing overlays of objects as good as possible. This condition is ideally met in new stadiums that are exclusively designed for football matches.

The considerations made so far reveal that finding a venue for a comparison of systems is not easy, based on the three main technologies of EPTS, because some requirements are rather exclusive (e.g., free line of sight to the satellites vs. high above and close to the pitch). As a side note, it may be recommended for customers of EPTS to make the provider demonstrate the accuracy of the system at the venues where the customer intends to apply the system, because the quality of measurements is dependent on the venue.

Last but not least, the criterion standards could require special conditions that have to be met in validation studies. For instance, infrared-based systems that represent the state-of-the-art system as the gold standard might not be used in direct sunlight. Furthermore, additional interfering components such as wet and/or infrared reflecting surfaces should be omitted. Outdoor infrared recordings must be conducted after sunset with the help of artificial light (floodlights). This could, in turn, create a conflict with the demand for sufficient light conditions of video-based systems.

Figure 4.4: Exemplary VICON test setup at the Duesseldorf Esprit Arena.

4.5 Location on the Pitch

A further decisive factor is the spatial arrangement of exercises on the playing field, as previous studies demonstrated that, depending on the underlying core technology, EPTS show location-dependent accuracy fluctuations. With regard to image-based multi-camera tracking systems, the coverage of the field has to be optimized in advance for each venue [48]. Differences in the coverage are unavoidable, which lead to different accuracy results in different areas of the pitch. Moreover, one might expect problems with objects moving in the line of sight of a camera compared to the trajectories that are perpendicular to the line of sight. This problem could occur in cases where the cameras are not mounted all around the stadium but at one spot, as is common practice in VBT. Similarly, the measurement accuracy of the radar-based LPM system is reported to decrease at the margins of the observation field [52] and could be influenced by the constructions at the testing site, which possibly leads to different accuracy in different pitch areas. GNSS technologies are known to have issues with inconsistent satellite signals such as total signal loss or poor accuracy due to signal reflection off, or shading by, buildings [62]. Depending upon its location on the playing field and distance from interfering constructions, the signal quality can vary at different parts of the pitch.

As a consequence, tests should not (only) be conducted in one confined space, but rather spread all over the playing field. Although this is quite costly, not only is it indispensable for a comprehensive check of the validity but would also provide valuable information for practical use, e.g., the places where the cameras should be mounted or where the exercises should be conducted with the highest demand on precision.

4.6 Test Exercises

A key factor of a professional standard for EPTS is the selection of exercises under comparison. Those should ideally include exercises under controlled and defined conditions in order to produce the entire range of movement patterns. In addition to these strictly controlled exercises, the test design should include 'free' and complex game scenarios to mimic the conditions under which tracking systems are generally used - sports competition and training. Accordingly, three different types of exercises should be considered to scrutinize the performance of a particular EPTS with the aim of pinning down the strengths and weaknesses:

Static testing. The most elementary exercise is static testing. For some time, objects that are completely static could be measured. This test provides valuable information on the baseline error of each system and for data processing later on.

Movements under controlled conditions. Elementary movements should be analyzed, detecting the respective capabilities of the system, which provides information about the dynamic accuracy and the adaptation to speed and change in direction.

Free and complex game scenarios. Finally, exercises with the highest ecological validity are small-sided games (SSGs). Typically, in sports such as football, movements are not predictable and thus create the demand for this scenario. Unfortunately, to the best of our knowledge, a gold standard for full pitch testing does not exist to date. Therefore, the best possible alternative are SSGs, which are played on reduced pitch areas, often using modified rules and involving a smaller number of players than traditional football.

4.7 Hierarchy of Tracking Variables

The most basic information that is provided by EPTS is the xy-position of the tracked object. This information is given at a particular measurement frequency specific to the system. On the other hand, there are other variables derived from this information at an intermediate level like speed and acceleration, which are typically derived from xy-positions - except for GNSS also using Doppler shift. As these data require additional data processing and are subject to different measurement problems, it is recommended to account for this intermediate level in a validation, also. Finally, performance indicators are derived from tracking data like distance covered or number and duration of stays in certain intensity intervals. Consequently, three different levels of tracking variables should be included in a validation study, because at each level different error characteristics could occur and different accuracy demands are to be met:

Fundamental Tracking Data
Raw trajectories (spatio-temporal XYZ-positions)

Kinematic variables for each point in time
Momentary speed and acceleration

Key Performance Indicators (KPI)
e.g. distance covered, sprints, accelerations, etc

It should also be noted that manufacturer proprietary software often use data-processing algorithms that are subject to intellectual property protection, and their specific algorithms are not disclosed to the end user [47]. However, without detailed knowledge of the software's internal processes, it is not possible to draw conclusions as to the causes of the occurring deviations from the gold standard.

Therefore, to achieve a transparent validation procedure, and to facilitate appropriate interpretation and replication by others, this thesis pursued the approach to calculate the KPIs based on the provided raw data independently independent from the manufacturer's data processing.

4.8 Reference Position on the Human Body

The fundamental focus of interest is the movement of the human body in space. At first sight, this definition seems unambiguous. A prime question, however, concerns to what part(s) of the human body best reflects the entirety of a human body and its motion through space. In biomechanics, the single representative for the entire human body is commonly referred to as its center of mass (COM) and is located somewhere between 55% (women) and 57% (men) of standing height (approximately at the waist level) [75]. Under the assumption that each EPTS endeavors to detect the position of the human body as a whole, the COM (or rather the COM's XY-position that is projected on the ground plane) could be considered a valid criterion measure. In the case of wearable tracking devices, however, the systems detect the position of the sensors that are fixed to the players (usually attached between the shoulder blades or on top of the shoulders). In video-based systems, objects are tracked by image segmentation using different techniques of image recognition [10]. Typically a rectangle is identified enclosing segmented parts of the player, and a weighted estimate of the body parts locates the body's center. Collectively, EPTS traditionally rely on one of two body positions as the ultimate representative for the entire body in space: the upper torso between the scapulae (GNSS and LPS) or the estimated center of mass (VBT and some LPS systems).

Tracking the EPTS sensors with the gold standard focuses on internal validity of the measurement, while comparing the best estimate of the COM with the results of the EPTS compares it to the practically most relevant information, thus putting the accent on ecological validity. To estimate COM, six adhesive marker mounts can be glued on the participant's skin (right shoulder (RSHO), left shoulder (LSHO), left anterior superior iliac spine (LASI), right anterior superior iliac spine (RASI), sacrum (SACR), and front/umbilicus (FRON)) (see Fig 4.5). The reflective markers are then fixed to the mounts through a tight-fitting compression shirt. COM is then estimated utilizing the reconstructed pelvis method [60], defined as the spatial center of RASI, LASI, SACR, and FRON. COS is defined as the spatial center of RSHO and LSHO.

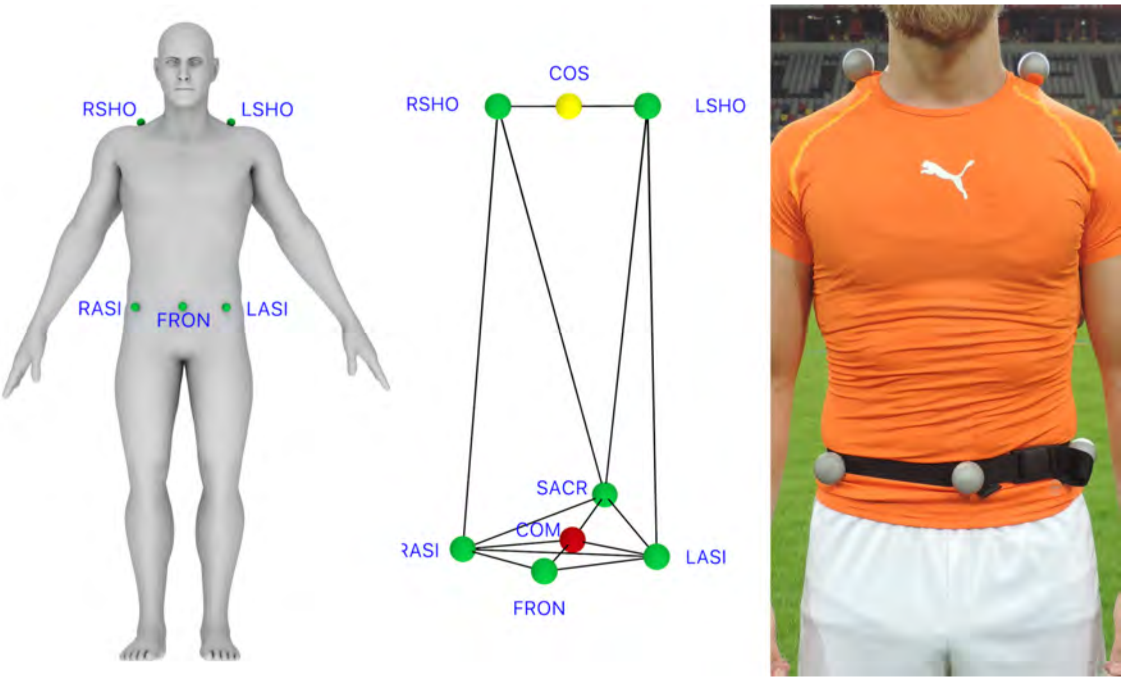


Figure 4.5: VICON marker positions on the human body (green), center of shoulders COS (yellow), center of mass COM (red).

During this dissertation, it became obvious that the classical definition of COM cannot be measured very precise by the use of hip markers (solely under the condition that the subject is standing upright. However, depending on the body position/lean/tilt, the COM can even be outside the body). Therefore, a more suitable term/representative for the body's center seems to be the center of pelvis COP.

4.9 Vicon Proof of Accuracy

A question that is frequently not addressed in EPTS validation studies is controlling and documenting the accuracy of the applied gold standard [49]. In most cases, it is not sufficient to report the manufacturer's specification. Instead, it is recommended that researchers conduct their own measurements. This holds especially when the gold standard is not applied in typical situations like a gait lab for VICON but in a natural environment like a stadium with specific light conditions and measurement volume stretched to its highest possible extent. A comprehensive validation approach, therefore, should provide unbiased evidence of the measurement accuracy of the applied reference technology because no test (even if the best available) is perfect and has some bias (error) attached to it.

4.9.1 Validation Method

To demonstrate the (relative) measurement accuracy of the applied VICON setup, a rigid calibration object with precisely known dimensions is moved in increasing concentric circles to cover the entire measurement area. As the markers on the calibration object remain at accurately known distances to each other at any given time, the distances between the markers that are delivered by the VICON software, which are calculated in retrospect, can be used to demonstrate the measurement accuracy. Measurement errors in 3D space were estimated by means of the root mean square error (RMSE).

4.9.2 Motion Capture Technology

During the course of this thesis, an infrared camera-based motion capture system (VICON, Oxford, UK) was utilized to determine the criterion movement position, speed, and acceleration data. The setup consisted of 33 cameras in total (Six *Vantage 5* cameras (16.0 mm), nine *Bonita* cameras (8.5 mm), 12 *MX T10-S* (8.5 - 12.5 mm) and six *MX T10* (8.5 - 12.5 mm), Software: Nexus, Version 2.3). To achieve sufficient coverage of the entire measurement area, each camera had to be equipped with special lenses (focal length between 8.5 mm & 16.0 mm) and was stationed at its previously calculated position. The cameras were evenly distributed around the observation field and calibrated using a custom-made rigid calibration object with three markers at precisely known spatial distances from one another (200.0 & 400.0 mm). Retro-reflective markers with a diameter of 38.0 mm were used to assure stable recognition of the markers within the entire measurement area (30.0 x 30.0 m, 900.0 m^2).

4.9.3 Vicon Accuracy Results

Due to the large-scale measurement area under consideration, the quality of the measurement accuracy varied (with decreasing accuracy toward the peripheral regions of the measurement area) (see Fig. 4.6).

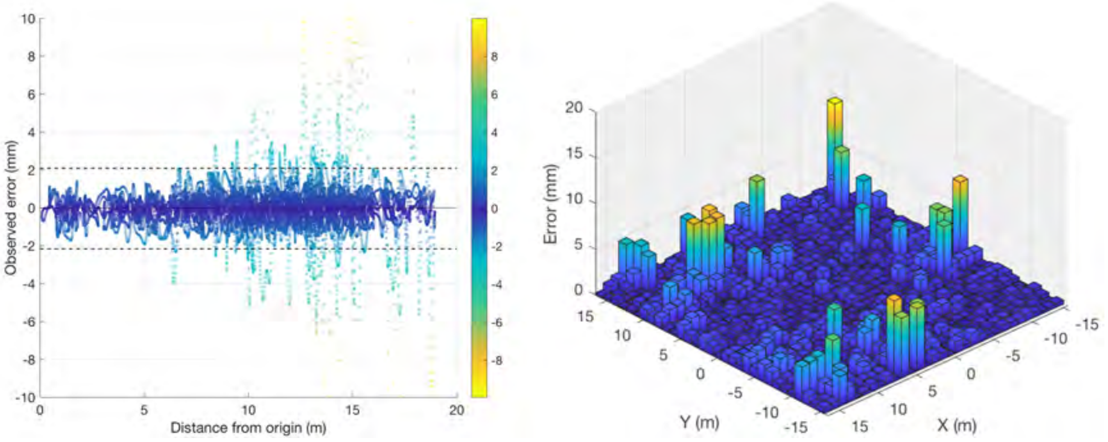


Figure 4.6: *Left:* Visualization of VICON measurement errors in relation to the distance from the origin. *Right:* Spatial distribution of the observed errors over the 30x30m measurement area.

Overall, the calibrated VICON setup achieved an accuracy of 1.16 mm root mean square error RMSE (95% CI [-2.09 mm, +2.17 mm]) at a measurement frequency of 100 Hz, thus meeting the aforementioned requirements for a gold standard reference with errors being more than one magnitude smaller than the errors of the EPTS being tested (ca. >0.2 m). In Fig. 4.6, the observed measurement errors are plotted against the distance from the origin.

4.10 Data Processing

Preliminary results showed that, despite the proven ability to provide marker positions with sub-centimeter accuracy, raw motion-capture data needs further customization to serve as an appropriate criterion method. The exemplary visualization of an athlete's speed signal (see Fig. 4.7) serves to demonstrate the need for further adaptations of the gold standard.

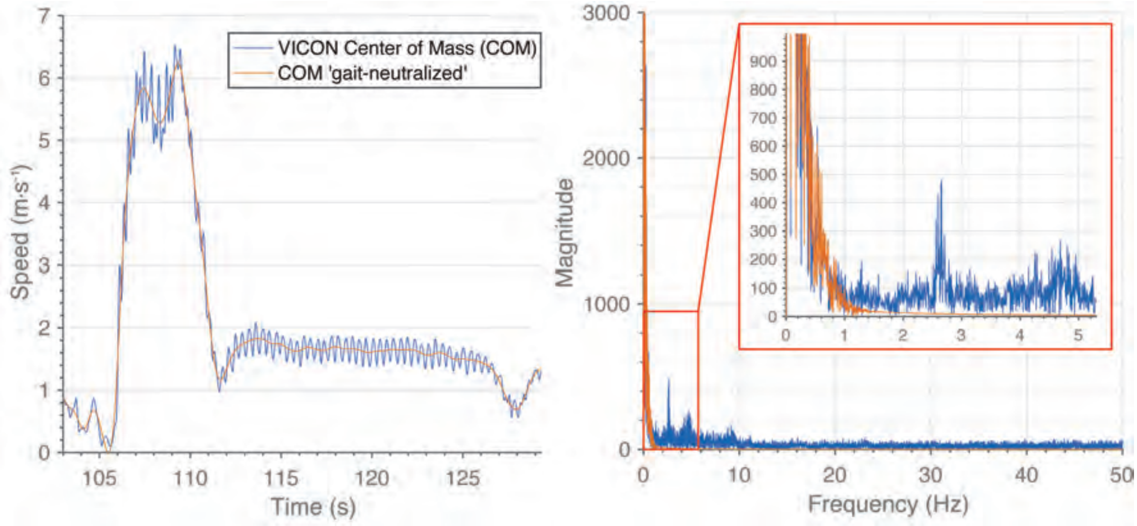


Figure 4.7: *Left:* VICON speed signal of a sprint, followed by low speed cruising. The blue line represents the raw VICON speed signal. The red line represents the 2 Hz low-pass filtered speed signal (criterion speed signal, gait-neutralized). *Right:* Speed signal comparison of an accelerating-cruising-walking sequence before (blue) and after (red) 2 Hz low-pass filtering (criterion measure, gait-neutralized).

When humans walk and run, between the heel-strike and mid-stance, the forward speed of the COM decreases and between the mid-stance and toe-off, it increases within each instance of ground contact of each leg [75]. This results in a ”true” horizontal speed curve that looks like a sine wave oscillating around the mean horizontal speed (see Fig. 4.7). Since most EPTS do not have the capability to assess intra-cyclic speed or acceleration fluctuations, a comparison with a gold standard that does have this capability would be ”unfair” in the sense that first, there is an increased deviation because intra-cycle speed is not achievable in the case of these systems, and second, EPTS are specifically designed to assess the gross movements of players.

For this reason, comparisons with ”gait-neutralized” speed and acceleration of the gold standard is advisable. To achieve this goal, we studied typical speed signals of football-specific movements through spectral analysis using Fast Fourier Transform (FFT) (Fig 4.7).

The occupied bandwidth, as a measurement of the frequency bandwidth that contains 99.0% of the total power of the speed signal, is located at approximately 2 Hz. A further noticeable peak in the spectrum, at around 2.5 Hz, most probably corresponds to the intra-cyclic variations of movement speed. Therefore, the gait-neutralized reference speed was calculated using a 4th order 2 Hz Butterworth low pass filter on the raw VICON speed (change in position divided by change in time). The gait-neutralized acceleration was calculated using finite differentiation of the gait-neutralized speed (difference in speed divided by change in time).

Further analysis of the VICON data showed that the projection of the COM travels considerably even if there is no perceivable movement of a player. This is due natural body sway or postural changes, which do not result in discernible changes of position (for a visualization, see *Video 2*). Therefore, the gold standard's positional data was additionally processed using a 'waypoint method' to account for these microscopic movements that are partially detectable only in the case of highly sensitive devices, but not for EPTS, and should be excluded from the assessment of the athlete's gross motion (see Fig. 4.8).

The waypoint method assumes that only after a distance traveled between any two tracking points exceeds a certain threshold, typically one step length, these tracking points can be considered for distance calculation. With these remaining points (support points) a new trajectory was calculated using cubic spline interpolation. A threshold of 60.0 cm used as investigations showed that this is a reasonable estimate for the COM displacement during a walk cycle, thus aiming to exclude COM displacements that are smaller than a single step in a way that they are excluded from the measurements of the gross motions of players. It should be stressed here, that the waypoint method should only be used to obtain

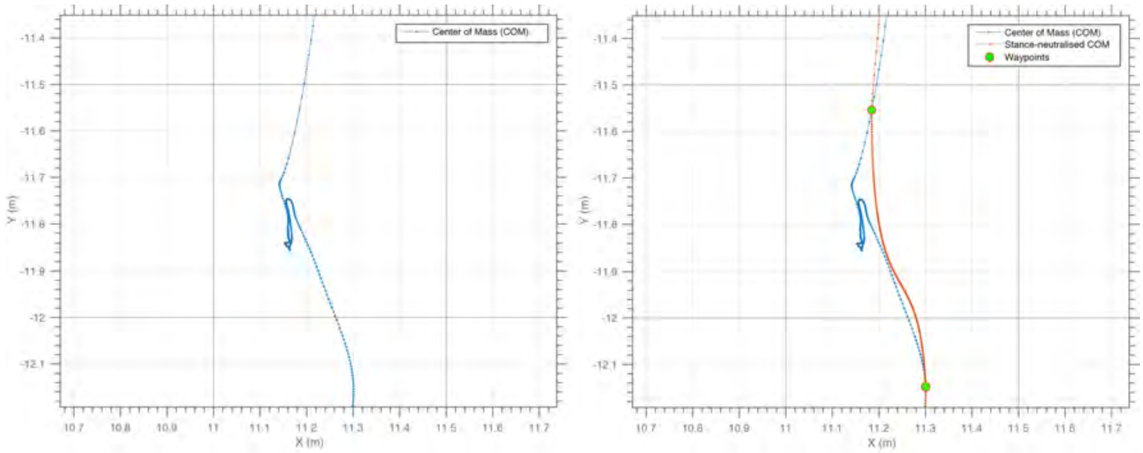


Figure 4.8: Stance-neutralization with the waypoint method. *Left:* The blue trajectory represents the actual horizontal movement of the center of mass (COM). In the example shown, the athlete moves from south to north and briefly stops in the center. Even if the athlete is standing still (both feet on the ground for a certain time), the COM trajectory would lead to artificial accumulating distances that are measured. *Right:* The waypoint method (red trajectory) suppresses the micro movements that are not relevant for the gross motion.

the aggregated distance references, whereas the spatial accuracy (XY-position in space) of each system should be validated against raw VICON positions (4th order 10 Hz Butterworth low pass filter applied to the raw positions).

4.11 Statistical Analysis

When comparing the time series of the measurements, the question of the most appropriate statistics arises in the context of EPTS' validation studies.

In the past, correlation measures (Pearson's r or the intraclass correlation (ICC)) have been a popular choice as the measure of agreement in validation studies. Whereas the ICC avoids the well-known problem of the linear relationship being mistaken for agreement, it does not avoid other problems that are associated with correlation coefficients in this context. Correlation demonstrates a dependency on the range of the measurement and the relation between the actual scale of measurement and the size of error [12]. Moreover, it is sensitive to the outliers and heterogeneous samples. It is therefore apparent that the commonly used ICC is prone to exactly the same constraints as Pearson's r, in that it includes the variance term for individuals and is therefore affected by the sample heterogeneity to such a degree that a high correlation could still imply unacceptable measurement error pertaining to certain analytical goals [6]. In this context, previously reported agreements with regard to r >.98 should be interpreted with caution.

It is more advisable, therefore, to use an absolute measure for the differences by means of the Root Mean Square Error (RMSE), which is commonly used in biomechanical analysis to measure the difference between the values that are observed by a criterion measure and by tested systems. In the formula stated below, X_{gs} represent the observed values of the gold standard and X_{syst} represent the observed values of the respective system at the time i.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{gs} - X_{syst})^2}{n}}$$

To further obtain a relative error measure, a simple percentage deviation of the observed value obtained from the criterion cannot be recommended as over-estimations and under-estimations might cancel each other out and give the false impression of an acceptable agreement. Instead, it is suggested to calculate the percentage root mean square error (% RMSE), which is calculated from the absolute RMSE, as its percentage of the average total distance measured by the criterion. This provides a value that is comparable to a coefficient of variation (CV) and may be taken as relative error. However, it needs to be emphasized that the % RMSE is dependent on the size of the measured value (e.g. a measurement error of 1m over a distance of 2 m results in a % RMSE of 50%, whereas a measurement error of 1m over a distance of 100m results in a % RMSE of 1%). Thus,

% RMSE values should only be compared within the same categories.

In addition to the numerical methods stated above, Altman and Bland [2] introduced the method of 'limits of agreement', an indicator of absolute reliability like standard error of measurement (SEM) and CV. In this graphical method, the differences (or the ratios) between the two techniques are plotted against the averages of the two techniques. Alternatively, the differences can be plotted against one of the two methods, if this method is considered a reference or gold standard method [39]. Horizontal lines are drawn at the mean difference and at the limits of agreement, which are defined as the mean difference plus and minus 1.96 times the standard deviation of the differences [2]. A modified version of the traditional Bland-Altman Plot is shown in Figure 4.9.

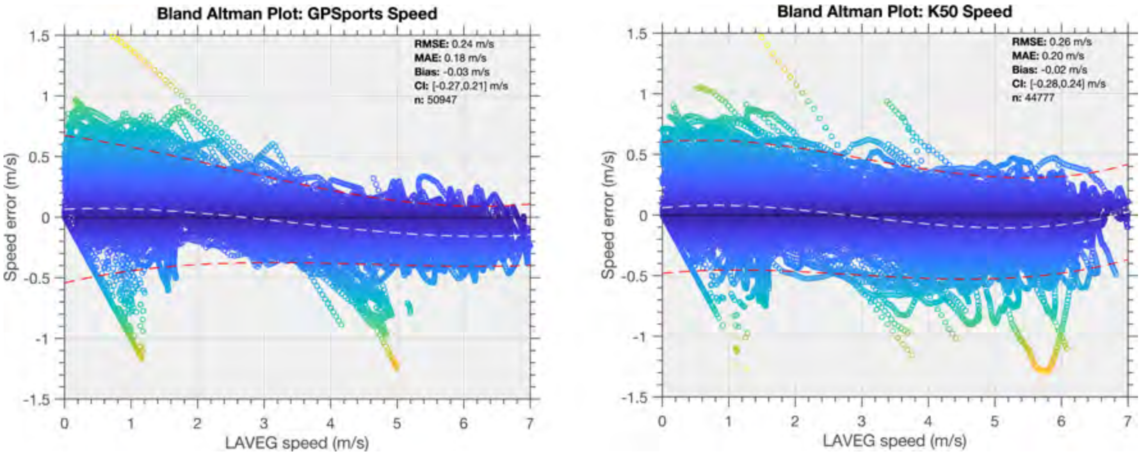


Figure 4.9: Exemplary Bland-Altman analysis of speed measurements of two different GPS systems (GPSports (left) and K50 (right)). The deviations between the reference method (Laveg laser) and the GPS system are plotted against the Laveg speed. The bias line (white) and random error lines forming the 95% limits of agreement (red) are also presented on the plot.

Another visual method is to display the probability distribution of errors in a histogram. A histogram provides a visual interpretation of numerical data by showing the number of data points that fall within a specified range of values (called 'bins'). An example is shown in Figure 4.10.

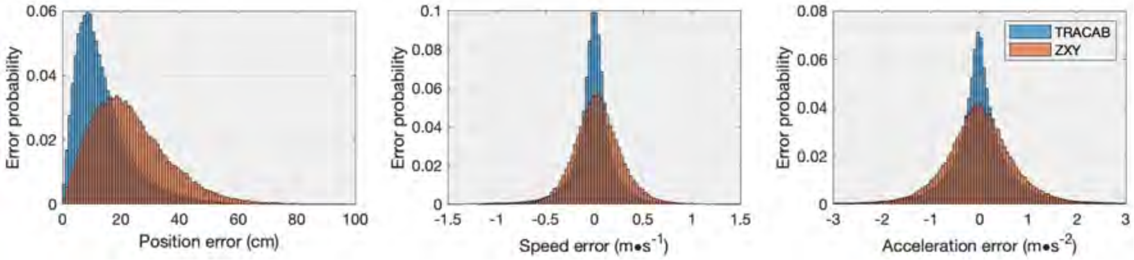


Figure 4.10: Exemplary Error Histogram of two different EPTS. The histograms show the error distributions of position (left), speed (middle), and acceleration (right).

Collectively, this chapter discussed a number of methodological considerations and highlighted the levels of complexity and intricacy involved in EPTS accuracy testing. It is also evident that previous research did not always consider all of the above-mentioned aspects. Hence the need for a more comprehensive EPTS validation approach.

5 Publications

The following chapter presents the publications underlying this thesis. Thematically, the studies are divided into two-subject areas. The first two publications deal with the subject of validity and comparability of tracking data provided by different tracking technologies. To achieve this objective, *Study One* applied a criterion-based validation method to analyze the measurement accuracy of the most commonly used EPTS technologies in sport. To gain a better understanding of the occurring deviations of tracking results delivered by different EPTS, *Study Two* investigated the influence of the respective anatomical reference position on resultant tracking variables.

Whereas the first two studies are assigned to the field of measurement science and technology, study three and four intended to put spatiotemporal tracking data into sport-scientific practice. In particular, *Study Three* initiated an innovative approach to quantify the contribution of game interruptions to the fatigue-related declines in match running performance over the course of a football match. Finally, *Study Four* described and analyzed the substitution tactics of an international field hockey team during competitive match play and identified the impact of bench periods on the physical output of players when re-entering active gameplay.

PLOS | ONE

RESEARCH ARTICLE

Validation of electronic performance and tracking systems EPTS under field conditions

Daniel Linke*, Daniel Link, Martin Lames

Chair of Performance Analysis and Sports Informatics, Technical University Munich, Munich, Germany

* daniel.linke@tum.de

Citation: Linke D, Link D, Lames M (2018) Validation of electronic performance and tracking systems EPTS under field conditions. PLoS ONE 13(7): e0199519. https://doi.org/10.1371/journal.pone.0199519

Editor: Luca Paolo Ardigò, Universita degli Studi di Verona, ITALY

Received: September 26, 2017

Accepted: June 9, 2018

Published: July 23, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The publication of the study was supported by the German Research Foundation (DFG, www.dfg.de) and the Technical University of Munich (TUM, www.tum.de) in the framework of the Open Access Publishing Program. This research was also supported by a research fund that was received from STATS LLC (www.stats.com). The Technical University of Munich, represented by the Chair of Training Science and

Abstract

The purpose of this study was to assess the measurement accuracy of the most commonly used tracking technologies in professional team sports (i.e., semi-automatic multiple-camera video technology (VID), radar-based local positioning system (LPS), and global positioning system (GPS)). The position, speed, acceleration and distance measures of each technology were compared against simultaneously recorded measures of a reference system (VICON motion capture system) and quantified by means of the root mean square error RMSE. Fourteen male soccer players (age: 17.4±0.4 years, height: 178.6±4.2 cm, body mass: 70.2±6.2 kg) playing for the U19 Bundesliga team FC Augsburg participated in the study. The test battery comprised a sport-specific course, shuttle runs, and small sided games on an outdoor soccer field. The validity of fundamental spatiotemporal tracking data differed significantly between all tested technologies. In particular, LPS showed higher validity for measuring an athlete's position (23±7 cm) than both VID (56±16 cm) and GPS (96±49 cm). Considering errors of instantaneous speed measures, GPS (0.28±0.07 $m \cdot s^{-1}$) and LPS (0.25±0.06 $m \cdot s^{-1}$) achieved significantly lower error values than VID (0.41±0.08 $m \cdot s^{-1}$). Equivalent accuracy differences were found for instant acceleration values (GPS: 0.67±0.21 $m \cdot s^{-2}$, LPS: 0.68±0.14 $m \cdot s^{-2}$, VID: 0.91±0.19 $m \cdot s^{-2}$). During small-sided games, lowest deviations from reference measures have been found in the total distance category, with errors ranging from 2.2% (GPS) to 2.7% (VID) and 4.0% (LPS). All technologies had in common that the magnitude of the error increased as the speed of the tracking object increased. Especially in performance indicators that might have a high impact on practical decisions, such as distance covered with high speed, we found >40% deviations from the reference system for each of the technologies. Overall, our results revealed significant between-system differences in the validity of tracking data, implying that any comparison of results using different tracking technologies should be done with caution.

Introduction

Electronic performance and tracking systems (EPTS) primarily track player (and ball) positions and have become one of the most important components to monitor a player's overall external (locomotor) load [1]. In particular, semi-automatic multiple-camera video systems

Sports Informatics, functioned as an independent third party for conducting this validation study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The terms of this arrangement have been reviewed and approved by the Technical University of Munich in accordance with its policy on objectivity in research.

Competing interests: This research was supported by a research fund that was received from STATS LLC (www.stats.com). The Technical University of Munich, represented by the Chair of Training Science and Sports Informatics, functioned as an independent third party for conducting this validation study. The funders did not exhibit influence on study design, data collection and analysis, decision to publish, or preparation of the manuscript. These conditions do not alter our adherence to PLOS ONE policies on sharing data and materials. We declare that with STATS LLC, no relationship like employment, consultancy, patents, products in development, marketed products, etc. exists. Further, the authors do not own sticks or shares of STATS LLC and the authors are not paid for employment of consultancy.

(VID), radio-based local positioning systems (LPS) and global positioning systems (GPS) have become indispensable core technologies for assessing the physical and tactical behaviour of both training and competition [2, 3]. As a matter of fact, it is not uncommon for some players to be tracked by two or three different EPTS during a regular week, considering that GPS systems and/or LPS systems are often used during training sessions, while most teams obtain positional data from official matches from semi-automatic camera systems [4]. Consequently, validity, interchangeability and agreement between different EPTS are of key importance to allow for a substantiated assessment of a player's overall locomotor load and to integrate the data of different systems in a meaningful way.

A review of the literature on the subject of EPTS' validity reveals that previous studies differ with regard to the number of tested core technologies (single technology [5±15] vs. multiple technology studies [3, 16, 17]), the choice of exercises (predefined movement patterns [3, 5, 7, 11, 13, 14, 17, 18] vs. complex and free movements scenarios [6, 8, 19]), and, most importantly, the utilized criterion method. The most commonly used criterion methods include predefined movement circuits with known spatial arrangements (to evaluate distance measurement accuracy) [3, 7, 11, 12, 18], timing gates (to evaluate average speed) [3, 5, 7, 9, 12, 18], and radar/laser-based speed measurements for the evaluation of instantaneous running speed [10, 13, 17].

However, all these methods have specific drawbacks. First, distance references that are based on predefined running circuits are inevitably susceptible to errors introduced by the participants (e.g. errors introduced by postural sway or the difficulty for participants to follow the marked course as precisely as possible [7]). Second, timing gates are only of limited suitability as a speed reference [20], the reason for this being that this approach only determines average speed based on limited sampling points [21]. Third, while radar/laser guns are capable of measuring the instantaneous speed of an object with high accuracy, they are suitable only when it comes to validating linear running movements without changes in direction [17].

Therefore, the actual positional data obtained by EPTS should ideally be compared with the instantaneous positional and speed data of a two or three-dimensional reference system with known error estimates [8]. However, to our knowledge, merely four validation studies used a kinematic analysis approach to evaluate the validity of EPTS. Specifically, Duffield et al. [6] and Vichery et al. [19] used a VICON motion analysis system to validate GPS systems in field-based team sports, while Ogris et al. [8] and Stevens et al. [22] investigated the accuracy of a radar-based LPS-system during soccer-specific movements. Limitations of the aforementioned studies include a lack of instantaneous accuracy measures for both speed and position (rather than merely average differences of the mean aggregated data) [6, 19, 22], missing information on the specific data processing steps [6, 19], a lack of realistic game scenarios [6, 19, 22], an insufficient size of the test area [6, 19, 22], as well as a lack of direct comparison between different technologies [6, 8, 19, 22].

A review of the literature further reveals that previous EPTS validation studies can be divided into three categories, according to the examined parameters. The first category contains studies that analyzed position accuracy (spatial coordinates) [8, 17]. Others examined the accuracy of instantaneous speed and acceleration data [10, 23]. Errors in this category could result from either a poor quality of position data or inadequate processing algorithms [21]. Eventually, an accumulation of errors in the first two categories can lead to errors in the third category: key performance indicators (KPI) that are aggregated from the continuous data (e.g. distance covered, mean or peak speed, peak accelerations, etc.) [6, 19, 22]. Consequently, aiming at a comprehensive accuracy assessment of EPTS requires comparisons in three different categories, because in each category different problems could occur and different accuracy demands are to be met.

Furthermore, up-to-date information on the spatial accuracy of sport-specific GPS technologies is still missing in the current literature. Considering the fact that various studies made use of GPS-based spatial coordinates to answer relevant scientific questions [24, 25], as well as the fact that several commercial GPS-systems determine distance via positional differentiation and speed via Doppler shift [21], information on the spatiotemporal accuracy of sport-specific GPS-systems is still scarce.

Therefore, the purpose of the current study was to assess the accuracy of the most commonly used tracking technologies in professional team sports under field conditions (i.e., semi-automatic multiple-camera video technology, radar-based LPS technology, and GPS technology). Measures of each technology were compared to that of a reference system (*VICON*). This was done for test runs along predefined tracks, shuttle runs, and small sided games. The results could contribute to an improved understanding of performance parameters provided by EPTS.

Methods

Participants

Fourteen male soccer players (age: 17.4±0.4 years, height: 178.6±4.2 cm, body mass: 70.2±6.2 kg) playing for the German Bundesliga team FC Augsburg participated in the study. Prior to participation, all players received comprehensive verbal and written explanations of the study, which was conducted within a period of two consecutive days. On each single day, 10 players participated. On the second day, four players from the first day had to be substituted. Therefore, fourteen different individual players participated in total. Voluntarily signed informed consent to wear GPS/LPS sensors and VICON markers and to participate in the collection of spatiotemporal tracking data was provided by both the players and their parents. Institutional board approval for the study was obtained from the Ethics Commission of the Technical University of Munich. To ensure confidentiality, all performance data were anonymized. This study conformed to the recommendations of the Declaration of Helsinki.

Validated systems

The following EPTS were included in the validation study:

Video Technology (VID). STATS SportVU (three-camera HD system, cameras: 3 x BASLER acA2500-14gc, 2560*1500 pixels, 16 frames per second). Software: STATS SportVU version 2.12.0, build # 12351. The camera elevation angle ranged from 22Ê(close sideline) to 11Ê (rear sideline).

Global Positioning System (GPS). GPSports (GPSports Sports Performance Indicator (SPI) Pro X, Canberra, Australia). This version of the SPI Pro provides raw position, instant speed and distance data at 15 Hz (5 Hz interpolated to 15 Hz). Software: Team AMS firmware: R1 2015.10. All GPS devices were activated 15 min prior to the data collection to allow the acquisition of satellite signals. Unfortunately, horizontal dilution of precision (HDoP) information cannot be retrieved with the provided Team AMS software. After making a request to the manufacturer in this regard, we were informed that the internal code automatically rejects data with HDoP values >4, which is well below the maximum value of 50 [26].

Local Positioning System (LPS). Inmotio (LPM system, 1 kHz, Inmotio Object Tracking BV, Amsterdam, Netherlands). Software: Inmotio Client, firmware: v3.7.1.153. 11 base stations were set up and calibrated under the supervision of an expert of the Inmotio company. During data collection, 22 transponders were activated to simulate a real match situation in terms of the number of transponders that were active at the same time, which resulted in an individual sampling rate of 45.45 Hz (1 kHz/22 transponders). LPM data was filtered with the integrated weighted Gaussian average filter set at 85%, as recommended by the manufacturer.

To ensure optimal device positioning on the body and minimization of crosstalk between GPS and LPS, athletes wore only one device of each system simultaneously. Using the harness provided by the manufacturers, GPS devices were positioned on the upper thoracic spine between the scapulae. LPS devices were worn in a vest containing a transponder located on the back that was connected to two antennas, one on top of each shoulder. The position of the athlete is then calculated as the spatial center of both antennas (manufacturer information).

Reference system

VICON system specifications. An infrared camera-based motion capture system (VICON, Oxford, UK) was utilized to determine criterion position, speed, and acceleration data. The setup comprised 33 cameras in total (six *Vantage 5* cameras (16.0 mm), nine *Bonita* cameras (8.5 mm), 12 *MX T10-S* (8.5Ð12.5 mm) and six *MX T10* (8.5Ð12.5 mm), Software: Nexus, Version 2.3). Retro-reflective markers with a diameter of 38.0 mm were used to assure stable recognition of the markers within the entire measurement area (30.0 x 30.0 m, 900.0 m^2).

VICON measurement accuracy. To demonstrate the spatial accuracy of the applied VICON setup, a rigid calibration object was moved through the VICON area, spiraling from the center to the edges of the measurement volume. As the markers on the calibration object remained at accurately known distances to each other at any given time, the distances between the markers that are delivered by the VICON software, which were calculated in retrospect, can be used to describe the crucial aspect of measurement accuracy (see S1 Dataset). Overall, the average error of the calibrated VICON setup was 0.0 mm (SD = 1.0 mm, 95% CI [-1.9 mm, +2.0 mm]), resulting in an RMSE of 1.0 mm at a frequency of 100 Hz.

Comparison criteria: Center of mass (COM). Under the assumption that each EPTS endeavors to detect the position of the human body as a whole, the center of mass (COM) (or rather the XY-position of the body's center that is projected on the ground plane) was considered a valid criterion measure. However, in the case of wearable tracking devices, the systems actually detect the position of the sensors that are fixed to the players (usually attached between the shoulder blades or on top of the shoulders). In video-based systems, objects are tracked by image segmentation using different techniques of image recognition [27]. Typically a rectangle is identified enclosing segmented parts of the player, and a weighed estimate of the body parts locates the body's center.

Eventually, the choice of the most suitable reference position on the human body should not be prescribed by the technological prerequisites of the respective EPTS, but rather by biomechanical considerations. We, therefore, advocate the idea that the ultimate reference position for each EPTS should be the COM, irrespective of where the respective transponder/receiver is attached to the human body. To estimate COM, five adhesive marker mounts were glued on each participant's skin (right shoulder (RSHO), left shoulder (LSHO), left anterior superior iliac spine (LASI), right anterior superior iliac spine (RASI), and sacrum (SACR)) (see Fig 1). The reflective markers were then fixed to the mounts through a tight-fitting compression shirt. COM is then estimated by means of the reconstructed pelvis method [28], defined as the spatial center of the RASI, LASI, and SACR.

Venue and satellite reception

Measurements took place at Rosenaustadion (Augsburg, Germany). This particular stadium is characterized by low stands (12.0 m maximum height at 50.0 m distance from the sideline). In

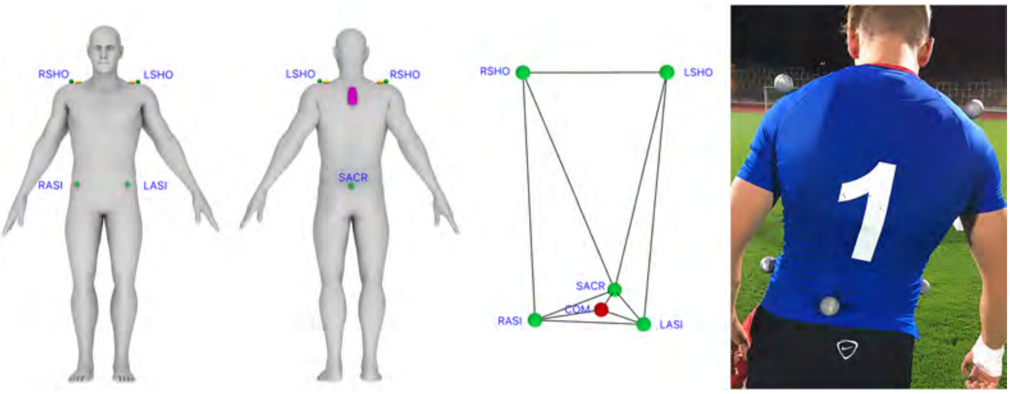


Fig 1. Device positions. Device positions on the human body. VICON markers (green), GPS receiver between the scapulae (purple), LPS antennas on top of both shoulders (orange), center of mass COM (red).

https://doi.org/10.1371/journal.pone.0199519.g001

addition, the pitch (105.0 m x 67.0 m) is surrounded by a tartan track. To meet the standard requirements for the camera system (sufficient height to obtain the required viewing angle), an additional platform had to be built on top of the stands (see Fig 2). During the entire measurement period, the number of connected GPS satellites was 10.1 ± 0.8, which is in the range of previous validation studies (e.g. 8 ± 1 [19], 9.5 ± 2 [3] and 12.3 ± 0.3 [13]). Thus, for all technologies involved, the minimum requirements were met. Data was recorded after sunset using floodlights. The weather was dry and windless with temperatures around 8ÊCelsius.

Exercises

Sport-specific course (SSC). A predefined circuit with prescribed movement intensities (Fig 3) was used to analyze elementary movements under controlled conditions, e.g. curved runs and runs with sharp turns. Within each trial, six distinct elementary movement patterns were performed: (1) 15 m sprint into 5 m deceleration, (2) 20 m sprint into 10 m backward running into 10 m forward running, (3) 505 agility test, (4) two rapid 90Êturns, (5&6) curved runs toward and away from the camera (see S2 Video). The beginning and end of each individual section was marked with two flat pylons, which in turn were equipped with reflective VICON markers. This enabled us in hindsight to detect the starting and endpoint of each section by means of the players' XY-position (a player was located within/outside a certain section if his COM crossed the line between the two start/end points).

20 m shuttle run test (SHU). Players repeatedly ran 20 m shuttles with 180Êchanges of direction at 11 $km{\cdot}h^{-1}$ for a period of two minutes. Subjects ran in groups consisting of ten players each. The shuttle run test was performed to obtain controlled test conditions including change of directions.

Small-sided game (SSG). Finally, exercises with the highest ecological validity are matches that take place on a full-sized pitch. Unfortunately, to the best of our knowledge, a gold standard for full pitch testing does not exist to date. Therefore, the best possible alternative are SSGs with fewer players competing on a smaller sized field. In our case, 5vs5 small-sided games were played, without goals, as collective possession play (see S2 Video). The format of the game-play comprised repeated 2-min bouts interspersed with 1-min of passive rest. Each drill was performed in a continuous regime, under the supervision, coaching, and

Fig 2. Venue. (Top) VICON test location on the pitch; (bottom right) scaled 3D model of the Rosenaustadion Augsburg. VICON area (blue), VID camera position (orange) at 21.6 m height and 82.0 m distance from the center spot. Pitch size: 105.0 x 67.0 m; (bottom left) additional camera platform.

https://doi.org/10.1371/journal.pone.0199519.g002

motivation of the coaches to maintain a high work-rate. The ball was always available owing to prompt replacement any time it was hit out of the measurement area.

Data analysis

Parameters for analysis. As indicated in the introduction, the validation of EPTS should be implemented through analysis of (i) position data, (ii) instant speed and acceleration data, and (iii) KPIs. It should also be noted here, that modern GPS systems derive speed and acceleration data based on the Doppler shift effect, instead of differentiation of position data [21]. We procured fundamental and derived data from the export option of each tracking system (XY-data, instant speed, and acceleration). Instead of using the KPIs as provided by the manufacturers' proprietary software, we deliberately decided to re-calculated these metrics, allowing us to use exactly the same algorithms for all tested systems. Manufacturer proprietary software often use data-processing algorithms that are subject to intellectual property protection, and their specific algorithms are not disclosed to the end user [21]. Therefore, to achieve a transparent validation procedure, and to facilitate appropriate interpretation and replication by others, it was decided to independently calculate the KPIs based on the provided raw data. Running intensities were divided into the following speed thresholds: standing (<1 km·h^{-1}), low speed (≥ 1 to <6 km·h^{-1}), moderate speed (≥ 6 to <15 km·h^{-1}), elevated speed (≥ 15 to

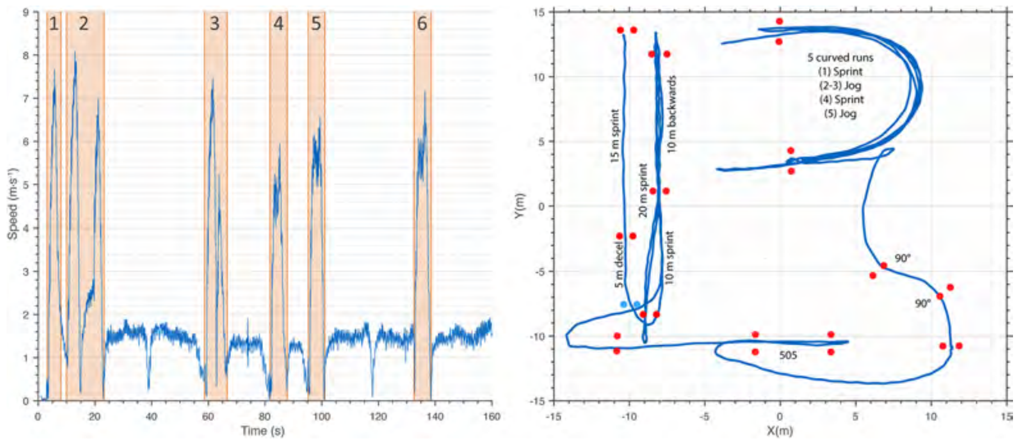


Fig 3. Sport-specific course (SSC) illustrated with exemplary VICON data. Left: chronological sequence of movement patterns (1 = 15 m sprint into 5 m deceleration; 2 = 20 m sprint into 10 m backward running into 10 m forward running; 3 = 505 agility test; 4 = two rapid 90Ê turns; 5&6 = curved runs toward and away from the camera). **Right**: spatial representation of movement patterns. Starting position in the top-left corner.

https://doi.org/10.1371/journal.pone.0199519.g003

<20 km·h^{-1}), high speed (≥20 to <25 km·h^{-1}), and very high speed (≥ 25 km·h^{-1}). Peak speed was defined as the highest measured speed value. High acceleration and deceleration thresholds were set at ≥3 m·s^{-2}, and ≤3 m·s^{-2}, respectively.

Data processing. To produce an evenly sampled time series among the systems prior to accuracy analysis, each data set was up-sampled to 100 Hz. The timing offset between the data sets was estimated by means of a cross-correlation procedure. Each coordinate system was then aligned with the VICON coordinate system via a generalized Procrustes analysis (GPA, euclidean similarity transformation, i.e. translation and rotation). After spatial and temporal synchronization of all systems involved, the VICON time code served as the ultimate reference for detecting the EPTS's start and end points of the respective exercise/section.

Data processing of raw VICON data consisted of filtering using a 4th order 10 Hz Butterworth low pass filter. Gaps in the data of 1 to <10 ms were filled using spline interpolation. Gaps that were ≥10 ms were excluded from analysis. XY-positions for spatial accuracy analysis were directly derived from the 100 Hz VICON data. The third dimension (Z-coordinate) was neglected in the calculations.

Raw VICON data needs further adjustments in order to serve as an appropriate reference. When humans walk and run, between the heel-strike and mid-stance, the forward speed of the COM decreases and between the mid-stance and toe-off, it increases within each instance of ground contact of each leg [29]. This results in a ªtrueº horizontal speed curve that looks like a sine wave oscillating around the mean horizontal speed (see Fig 4 left).

Since most EPTS do not have the capability to assess intra-cyclic speed or acceleration fluctuations, a comparison with a gold standard that does have this capability would be ªunfairº in the sense that first, there is an increased deviation because intra-cycle speed is not achievable in the case of these systems, and second, EPTS are only meant for assessing the gross movements of players. For this reason, comparisons with ªgait-neutralizedº sped and acceleration of the gold standard is advisable. To achieve this goal, we studied typical speed signals of football-specific movements through spectral analysis using Fast Fourier Transform (FFT) (Fig 4 right). The occupied bandwidth, as a measurement of the frequency bandwidth that contains

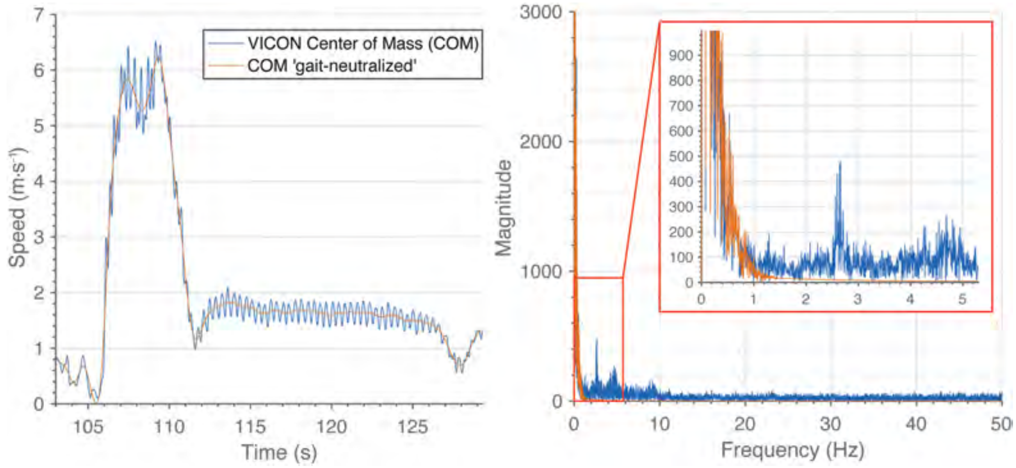


Fig 4. Gait-neutralization. Left: Speed signal comparison of an accelerating-cruising-walking sequence before (blue) and after (red) 2 Hz low-pass filtering (criterion measure, gait-neutralized). **Right**: Spectral analysis of a center of mass (COM) speed signal recorded with VICON. Blue line: unfiltered speed signal; red line: 2 Hz low-pass filtered speed signal.

https://doi.org/10.1371/journal.pone.0199519.g004

99.0% of the total power of the speed signal, is located at approximately 2 Hz. A further noticeable peak in the spectrum, at approximately 2.5 Hz, most probably corresponds to the intra-cyclic variations of movement speed. Therefore, the gait-neutralized reference speed was calculated using a 4th order 2 Hz Butterworth low pass filter on the raw VICON speed (change in position divided by change in time). The gait-neutralized acceleration was calculated using finite differentiation of the gait-neutralized speed (change in speed divided by change in time). Further analysis of the VICON data showed that the projection of the COM travels considerably even if there is no perceivable movement of a player. This is due natural body sway (see S1 Video for a graphical illustration), or postural changes, which do not result in discernible changes of position. Therefore, the gold standard's positional data was additionally processed using a ªwaypoint methodº to account for these microscopic movements that are partially detectable only in the case of highly sensitive devices, but not for EPTS, and should be excluded from the assessment of the athlete's gross motion. The waypoint method assumes that only after a distance traveled between any two tracking points exceeds a certain threshold, typically one step length, these tracking points can be considered for distance calculation. With these remaining points (support points) a new trajectory was calculated using cubic spline interpolation (see Fig 5). We used a threshold of 60.0 cm as our investigations showed that this is a good estimate for the COM displacement during a walk cycle, thus aiming to exclude COM displacements that are smaller than a single step in a way that they are excluded from the measurements of the gross motions of players. It should be stressed here, that the waypoint method was only used to obtain the aggregated distance references, whereas the spatial accuracy (XY-position in space) of each system was validated against the raw VICON positions (4th order 10 Hz Butterworth low pass filter applied to the raw positions).

Statistical analysis

Accuracy of fundamental XY-position data was estimated by means of the root mean square error (RMSE). Since we also analyzed the error pertaining to speed and acceleration

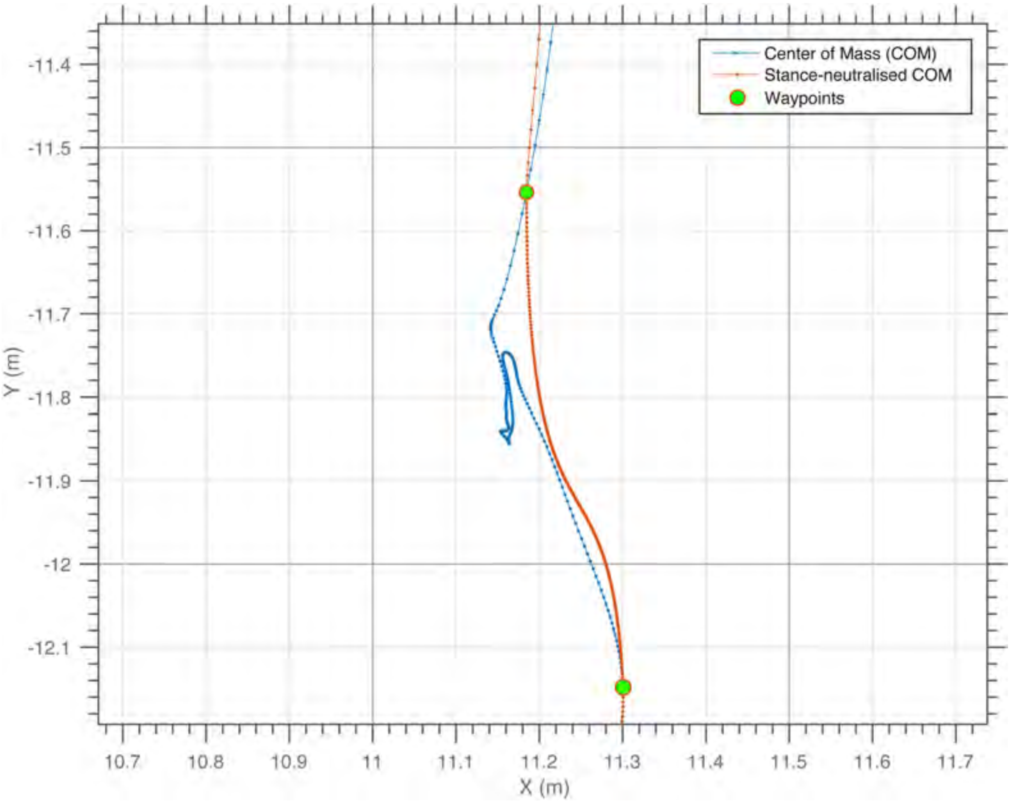


Fig 5. Stance-neutralization with waypoint method. The blue trajectory represents the actual horizontal movement of the center of mass (COM). In the example shown, the athlete moves from south to north and briefly stops in the center. Even if the athlete is standing still (both feet on the ground for a certain time), the COM trajectory would lead to artificial accumulating distances that are measured. The waypoint method (red trajectory) suppresses the micro movements that are not relevant for the gross motion.

https://doi.org/10.1371/journal.pone.0199519.g005

measurements, we distinguish between three types of RMSE: *dRMSE* (m): distance root mean square error (horizontal 2D accuracy); *vRMSE* ($m \cdot s^{-1}$): instant speed root mean square error, and *aRMSE* ($m \cdot s^{-2}$): instant acceleration root mean square error.

To analyze the accuracy of fundamental (XY-position) and derived (instant speed and acceleration) measures, single-sample t-tests were conducted to determine if the mean of the resulting RMSEs of an individual EPTS was statistically significantly different from zero. Two-tailed paired t-tests were used to compare the aggregated (numerical) metrics derived by the respective EPTS with that derived from the reference system. Inter-system differences in accuracy levels were tested using repeated-measures one-way analysis of variance (ANOVA). Bonferroni's post hoc analyses were used when significant differences were found. A Shapiro-Wilk test was applied for testing the normality of the residuals and a Levene's test was used to test the homoscedasticity. In cases where data failed the normality test, non-parametric test procedures were used to analyze the data (Wilcoxon Signed-Ranks test and KruskalĐWallis test by ranks). Effect sizes (ES) were quantified to indicate the meaningfulness of the differences in the mean values. Cohen's *d* effect sizes for the t-tests was classified as trivial (0-0.19), small

Table 1. Sample Size.

	GPS	LPS	VID(total)
Sport-specific course (SSC)	6	12	26
Shuttle run (SHU)	20	10	37
Small-sided game (SSG)	38	50	134
Sum	**64**	**72**	**197**

Sample size: valid trials (single observations) included for analysis. GPS = Global Positioning System; LPS = Local Positioning System, VID = video system (synonymous with the total number of trials).

https://doi.org/10.1371/journal.pone.0199519.t001

(0.20±0.49), medium (0.50±0.79) and large (>0.80) [30]. Eta squared (η^2) ES for the analysis of variance were classified as small (0.02±0.12), medium (0.13±0.25) and large (>0.26) [30]. Since pre-screening of results revealed skewed error distributions and frequent outliers, descriptive statistics have been presented as the median (Med) and standard deviation (SD). Statistical significance for all calculations was set at $p < 0.05$.

Sample size

Table 1 summarizes the number of single observations included for analysis. The number of observations for each system varied due to organizational reasons, which were mainly caused owing to incomplete data sets, time restrictions and the fact that only one wearable technology could be analyzed at the same time (whereas the VID system recorded all trialsÐirrespective of which wearable system was measured at the same time). The total number of exercises comprised 26 SSCs, 4 SHUs, and 14 SSGs. The total number of trials included for analysis results from the sum of participating players per exercise (see Table 1). For GPS, LPS, and VID, four, three, and 13 data files contained data gaps. Accordingly, the relative loss of data sets due to measurement errors was 6.3%, 4.2%, and 4.6%, respectively.

Results

Fundamental data (position accuracy)

Table 2 and Fig 6 report the measurement error of EPTS in the respective category. Overall, smallest errors of fundamental spatial accuracy (dRMSE) were achieved by the radar-based LPS system (SSC: 27±5 cm; SHU: 22±13 cm; SSG: 23±5 cm; pooled: 23±7 cm), followed by the image-based VID system (SSC: 57±9 cm; SHU: 59±28 cm; SSG: 56±12 cm; pooled: 56±16 cm), and the GPS system (SSC: 88±22 cm; SHU: 133±54 cm; SSG: 81±51 cm; pooled: 96±49 cm). GPS showed noticeable exercise-dependent fluctuations in spatial accuracy. In particular, GPS demonstrated lower spatial accuracy during the shuttle runs (133±54 cm). For VID, we found significant differences in the X and Y dRMSE accuracy (X: 28±13 cm, Y: 50±15 cm) [F(1, 392) = 247.40, p<.001]. Post hoc analysis of the ANOVA revealed no homogeneous subsets, implying that the spatial error (dRMSE) differs significantly between all tested systems.

Derived data (instant speed and acceleration)

Lowest errors in vRMSE were achieved by the radar-based LPS system (SSC: 0.29±0.05 $m \cdot s^{-1}$; SHU: 0.20±0.04 $m \cdot s^{-1}$; SSG: 0.25±0.06 $m \cdot s^{-1}$; pooled: 0.25±0.06 $m \cdot s^{-1}$), followed by GPS (SSC: 0.26±0.01 $m \cdot s^{-1}$; SHU: 0.30±0.09 $m \cdot s^{-1}$; SSG: 0.28±0.06 $m \cdot s^{-1}$; pooled: 0.28±0.07 $m \cdot s^{-1}$) and VID (SSC: 0.36±0.07 $m \cdot s^{-1}$; SHU: 0.47±0.08 $m \cdot s^{-1}$; SSG: 0.41±0.08 $m \cdot s^{-1}$; pooled: 0.41±0.08 $m \cdot s^{-1}$) (Table 2). It is also apparent that each system's speed accuracy depends on the respective exercise. Whereas LPS presented the lowest speed error during the shuttle run trials, GPS and

Table 2. RMSE results.

		LPS		GPS		VID		ANOVA		
		Median	***±SD***	**Median**	***±SD***	**Median**	***±SD***	**p**	**ES**	**Sign. diff. groups**
dRMSE (m)	Sport-specific course (SSC)	0.27	*0.05*	0.88	*0.22*	0.57	*0.09*	***	large	ALL
	Shuttle run (SHU)	0.22	*0.13*	1.33	*0.54*	0.59	*0.28*	***	large	ALL
	Small-sided game (SSG)	0.23	*0.05*	0.81	*0.41*	0.56	*0.12*	***	large	ALL
vRMSE ($m \cdot s^{-1}$)	Sport-specific course (SSC)	0.35	*0.06*	0.32	*0.01*	0.41	*0.07*	**	large	GPS&VID / LPS&VID
	Shuttle run (SHU)	0.31	*0.04*	0.39	*0.08*	0.52	*0.08*	***	large	GPS&VID / LPS&VID
	Small-sided game (SSG)	0.36	*0.06*	0.39	*0.06*	0.47	*0.08*	***	large	GPS&VID / LPS&VID
	Standing (pooled)	0.34	*0.17*	0.18	*0.20*	0.23	*0.26*	***	medium	GPS&LPS / VID&LPS
	Low speed (pooled)	0.26	*0.07*	0.26	*0.13*	0.33	*0.12*	***	medium	GPS&VID / LPS&VID
	Moderate speed (pooled)	0.25	*0.07*	0.27	*0.07*	0.43	*0.10*	***	large	GPS&VID / LPS&VID
	Elevated speed (pooled)	0.34	*0.12*	0.37	*0.19*	0.49	*0.24*	***	small	GPS&VID / LPS&VID
	High speed (pooled)	0.39	*0.13*	0.37	*0.25*	0.50	*0.30*			
	Very high speed (pooled)	0.37	*0.12*	0.39	*0.13*	0.61	*0.43*	*	medium	GPS&VID / LPS&VID
aRMSE ($m \cdot s^{-2}$)	Sport-specific course (SSC)	0.69	*0.16*	1.18	*0.14*	0.78	*0.16*	***	large	GPS&LPS / GPS&VID
	Shuttle run (SHU)	0.58	*0.10*	0.56	*0.17*	0.80	*0.15*	***	large	GPS&VID / LPS&VID
	Small-sided game (SSG)	0.69	*0.13*	0.69	*0.14*	0.97	*0.19*	***	large	GPS&VID / LPS&VID

Presented as median ± standard deviation (SD). Inter-system differences in accuracy levels were tested using repeated-measures one-way analysis of variance (ANOVA). p values are presented as * ($p \leq 0.05$), ** ($p \leq 0.01$) and *** ($p \leq 0.001$). η^2 effect sizes (ES) for the analysis of variance were classified as *small* (0.02±0.12), *medium* (0.13±0.25) and *large* (>0.26). Homogeneous subsets are listed if Bonferroni's post hoc analysis did not result in significant difference between individual groups. GPS = Global Positioning System; LPS = Local Positioning System, VID = video system; dRMSE = distance root mean square error; vRMSE = velocity root mean square error; aRMSE = acceleration root mean square error.

https://doi.org/10.1371/journal.pone.0199519.t002

VID showed the lowest speed error during the sport-specific course trials. The point that all systems have in common is that the speed error increases as the speed increases.

Overall, it is apparent that the acceleration error (aRMSE) follows a similar pattern as the vRMSE categories. Due to the necessary step of derivation, an increase in the overall error is recognizable for all systems. For GPS, the error of the instantaneous acceleration (aRMSE) is twice as high during the sport-specific course (1.18±0.14 $m \cdot s^{-2}$) when compared to the shuttle run (0.56±0.17 $m \cdot s^{-2}$).

Post hoc analysis of the ANOVA revealed notably frequent homogeneous subsets of GPS and LPS, implying that the vRMSE and aRMSE of GPS and LPS did not differ significantly (with the only exception of aRMSE for SSCs and vRMSE for standing (see Table 2)).

Sport-specific course

Results of the specific categorization into fundamental movement patterns during the SSC trials are presented in Table 3. The point that all the systems have in common is that dRMSE, vRMSE, and aRMSE were lowest during low speed location changes. Compared to GPS and LPS, VID showed significantly lower speed accuracy values during linear sprint exercises (15 m sprint into 5 m acceleration and backward into forward sprints), which were both aligned at a 90Êangle (perpendicular to the camera view). However, in the opposite direction, (505 agility test, movements parallel to the camera view), VID showed smaller errors (0.32±0.23 $m \cdot s^{-1}$) than both LPS (0.51±0.07 $m \cdot s^{-1}$) and GPS (0.53±0.11 $m \cdot s^{-1}$).

Results of key performance indicators (KPI). The percentage difference in KPIs between the respective EPTS and the criterion measure are presented in Table 4.

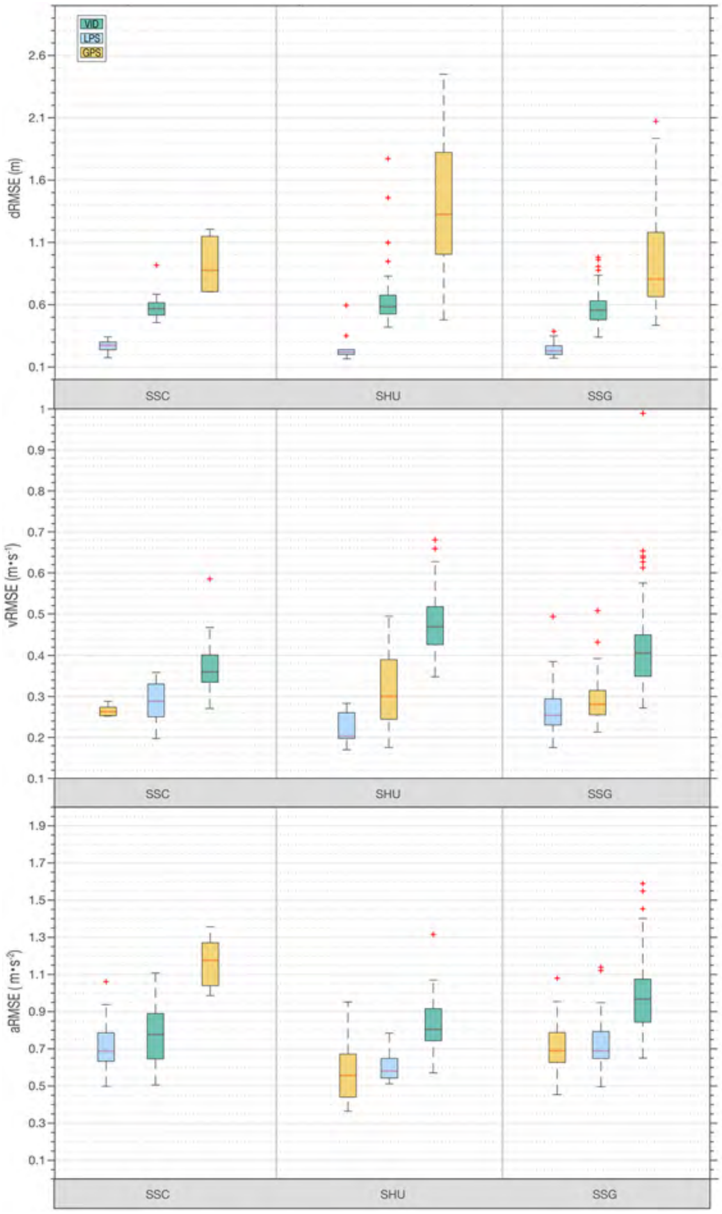


Fig 6. RMSEs by exercise. dRMSE (top), vRMSE (middle) and aRMSE (bottom). Box plots are used based on the five-number summary: minimum, first quartile, median, third quartile, and maximum. The central rectangle spans the first quartile to the third quartile. A red line inside the rectangle represents the median and the whiskers above and below the box show the locations of the maximum and minimum value. Red crosses indicate outliers. GPS = Global Positioning System; LPS = Local Positioning System, VID = video system; dRMSE = distance root mean square error; vRMSE = velocity root mean square error; aRMSE = acceleration root mean square error.

https://doi.org/10.1371/journal.pone.0199519.g006

Table 3. Results of the sport-specific course SSC.

		LPS		GPS		VID		ANOVA		
		Median	***±SD***	**Median**	***±SD***	**Median**	***±SD***	**p**	**ES**	**Sign. diff. groups**
dRMSE (m)	Low speed location change	0.21	*0.04*	0.79	*0.18*	0.45	*0.09*	***	large	LPS&GPS / LPS&VID
	15m sprint / 5m deceleration	0.34	*0.16*	0.99	*0.29*	0.63	*0.71*	**	large	LPS&GPS / LPS&VID
	Backward / forward sprint	0.29	*0.05*	1.14	*0.42*	0.70	*0.22*	***	large	LPS&GPS / LPS&VID
	505 agility test	0.42	*0.12*	0.99	*0.15*	0.92	*0.20*	***	large	LPS&GPS / LPS&VID
	90¡ turns	0.45	*0.10*	0.92	*0.16*	0.75	*0.07*	***	large	LPS&GPS / LPS&VID
	Curved run I (towards)	0.41	*0.17*	1.21	*0.32*	0.70	*0.15*	***	large	LPS&GPS / LPS&VID
	Curved run II (away)	0.48	*0.12*	1.27	*0.50*	0.60	*0.18*	***	large	LPS&GPS / GPS&VID
vRMSE (m·s^{-1})	Low speed location change	0.20	*0.05*	0.16	*0.02*	0.22	*0.05*	**	medium	LPS&GPS / GPS&VID
	15m sprint / 5m deceleration	0.32	*0.23*	0.27	*0.04*	0.63	*0.39*	***	large	LPS&VID / GPS&VID
	Backward / forward sprint	0.38	*0.07*	0.25	*0.04*	0.63	*0.20*	***	large	LPS&VID / GPS&VID
	505 agility test	0.51	*0.07*	0.53	*0.11*	0.32	*0.23*			NONE
	90¡ turns	0.44	*0.10*	0.35	*0.09*	0.79	*0.15*	***	large	LPS&VID / GPS&VID
	Curved run I (towards)	0.47	*0.14*	0.52	*0.10*	0.52	*0.09*			NONE
	Curved run II (away)	0.46	*0.14*	0.54	*0.19*	0.46	*0.15*			NONE
aRMSE (m·s^{-2})	Low speed location change	0.49	*0.15*	0.93	*0.20*	0.44	*0.14*	***	large	LPS&GPS / GPS&VID
	15m sprint / 5m deceleration	0.93	*0.86*	1.22	*0.35*	1.33	*0.70*			NONE
	Backward / forward sprint	0.87	*0.14*	1.20	*0.18*	1.23	*0.42*	**	large	LPS&GPS / LPS&VID
	505 agility test	1.24	*0.22*	2.07	*0.37*	0.85	*0.50*	***	large	LPS&GPS / GPS&VID
	90¡ turns	1.34	*0.37*	1.45	*0.33*	1.65	*0.49*			NONE
	Curved run I (towards)	0.94	*0.16*	1.55	*0.24*	1.02	*0.28*	**	large	LPS&GPS / GPS&VID
	Curved run II (away)	1.04	*0.43*	1.61	*0.35*	0.78	*0.26*	**	large	LPS&GPS / GPS&VID

Presented as median ± standard deviation (SD). p values are presented as * ($p \leq 0.05$), ** ($p \leq 0.01$) and *** ($p \leq 0.001$). η^2 effect sizes (ES) for the analysis of variance were classified as *small* (0.02±0.12), *medium* (0.13±0.25) and *large* (>0.26). Homogeneous subsets are listed if Bonferroni's post hoc analysis did not result in a significant difference between individual groups. GPS = Global Positioning System; LPS = Local Positioning System, VID = video system; dRMSE = distance root mean square error; vRMSE = velocity root mean square error; aRMSE = acceleration root mean square error.

https://doi.org/10.1371/journal.pone.0199519.t003

Discussion

Results showed that largest accuracy differences between EPTS were present in the first data category (fundamental XY-position in space). In particular, LPS had higher accuracy than both VID and GPS for measuring an athlete's position in space. However, our results also revealed that in the second category (instant speed and acceleration) errors of GPS are comparable to those of LPS, most likely related to the fact that GPS uses two fundamentally different measurement principles to determine an athlete's position and speed. In the third data category (KPIs), differences between technologies were not as pronounced as in the first and second data category, yet all technologies had in common that the magnitude of the error increased as the speed of the tracking object increased.

Position accuracy

The radar-based LPS system demonstrated the highest spatial accuracy with a dRMSE ranging from 22 cm (SHU) to 27 cm (SSC) (see Table 2). These findings are in accordance with previous research by Ogris et al. [8] (23 cm) and Siegle et al. [17] (24 cm). The sport-specific course has, however, also revealed that the spatial accuracy of the LPS system is dependent on instantaneous dynamics. In particular, fast changes of direction can lead to a significant increase of

Table 4. Results of key performance indicators (KPI).

		LPS					GPS					VID				
Test	**Metric**	**MeanGS**	**RMSE**	**RMSE%**	**p**	**ES**	**MeanGS**	**RMSE**	**RMSE%**	**p**	**ES**	**MeanGS**	**RMSE**	**RMSE%**	**p**	**ES**
SSC	Standing (m)	*0.81*	0.68	83.94	**	*medium*	*0.86*	3.68	429.85	*	*large*	*2.52*	5.89	234.23	***	*large*
	Low speed (m)	*102.93*	5.40	5.24	-	-	*82.30*	6.30	7.66	*	*trivial*	*96.96*	11.24	11.60	***	*small*
	Moderate speed (m)	*87.80*	8.04	9.15	-	-	*92.55*	7.94	8.58	-	-	*92.22*	15.52	16.83	**	*small*
	Elevated speed (m)	*51.80*	7.22	13.94	**	*large*	*43.80*	6.39	14.58	-	-	*45.86*	19.16	41.78	*	*small*
	High speed (m)	*52.76*	11.60	21.98	***	*medium*	*51.61*	9.35	18.11	-	-	*51.36*	16.34	31.82	-	-
	Very high speed (m)	*18.55*	5.31	28.65	-	-	*18.51*	9.46	51.12	*	*large*	*13.18*	12.91	97.94	-	-
	High acceleration (m)	*17.56*	6.60	37.58	***	*large*	*14.91*	9.71	65.14	*	*large*	*24.35*	8.50	34.90	-	-
	High deceleration (m)	*24.83*	3.93	15.82	-	-	*23.55*	10.94	46.46	-	-	*13.63*	4.35	31.94	-	-
	Total distance (m)	*314.64*	7.31	2.32	***	*small*	*289.63*	3.54	1.22	-	-	*301.59*	3.59	1.19	*	-
	Top speed ($m \cdot s^{-1}$)	*7.61*	0.34	4.51	-	-	*7.69*	0.31	4.03	*	*large*	*7.50*	0.51	6.81	-	-
SHU	Standing (m)	*0.15*	0.18	116.23	*	*large*	*0.09*	0.73	856.39	***	*large*	*0.38*	2.28	597.83	***	*large*
	Low speed (m)	*25.86*	1.30	5.01	-	-	*21.69*	12.41	57.22	***	*large*	*31.52*	13.69	43.42	***	*large*
	Moderate speed (m)	*440.09*	4.98	1.13	-	-	*427.69*	26.93	6.30	-	-	*389.35*	51.50	13.23	***	*small*
	Elevated speed (m)	*1.63*	3.38	207.14	-	-	*51.35*	39.81	77.53	*	*small*	*28.43*	73.71	259.25	***	*small*
	High speed (m)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Very high speed (m)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	High acceleration (m)	*5.12*	3.68	71.85	**	*large*	*9.26*	3.24	34.99	-	-	*14.06*	10.87	77.30	**	*small*
	High deceleration (m)	*4.67*	4.84	103.66	**	*large*	*12.09*	7.33	60.60	**	*small*	*3.61*	2.39	66.08	*	*trivial*
	Total distance (m)	*466.88*	3.45	0.74	-	-	*485.94*	21.44	4.41	-	-	*446.17*	19.77	4.43	**	*trivial*
	Top speed ($m \cdot s^{-1}$)	*4.07*	0.46	11.32	-	-	*4.44*	0.22	5.01	-	-	*4.34*	0.71	16.37	***	*medium*
SSG	Standing (m)	*0.29*	0.27	95.00	***	*large*	*0.27*	3.27	1194.41	***	*large*	*0.40*	1.83	455.30	***	*large*
	Low speed (m)	*34.11*	2.73	7.99	-	-	*42.46*	7.80	18.37	***	*small*	*36.92*	6.21	16.83	**	*trivial*
	Moderate speed (m)	*106.87*	6.59	6.16	***	*trivial*	*156.89*	5.87	3.74		-	*113.20*	10.34	9.13	-	-
	Elevated speed (m)	*13.74*	2.98	21.67	-	-	*22.20*	8.58	38.65	***	*small*	*16.33*	8.05	49.30	***	*trivial*
	High speed (m)	*5.91*	2.59	43.77	-	-	*4.12*	4.01	97.44	***	*medium*	*5.72*	5.58	97.62	-	-
	Very high speed (m)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	High acceleration (m)	*3.68*	3.05	82.87	***	*medium*	*5.64*	2.83	50.25	***	*small*	*6.41*	5.83	90.87	***	*medium*
	High deceleration (m)	*3.26*	2.34	71.92	***	*medium*	*9.12*	8.50	93.25	***	*medium*	*2.96*	3.00	101.11	***	*medium*
	Total distance (m)	*153.38*	6.05	3.95	***	*trivial*	*224.10*	4.90	2.18	**	*trivial*	*165.00*	4.60	2.79	***	*trivial*
	Top speed ($m \cdot s^{-1}$)	*4.81*	0.34	7.09	*	*trivial*	*5.47*	0.33	6.08	***	*small*	*4.86*	0.42	8.64	-	-

Deviation from the criterion standard presented as root mean square error (RMSE) and the percentage RMSE (RMSE in relation to the mean total distance in the respective category measured by the gold standard *MeanGS*). The table shows accuracy results for the covered distances in various intensity zones as well as the measured peak speed. *p* values are presented as * ($p \leq 0.05$), ** ($p \leq 0.01$) and *** ($p \leq 0.001$). Cohen's *d* effect sizes (ES) for the t-tests was classified as trivial (0-0.19), small (0.20±0.49), medium (0.50±0.79) and large (>0.80). GPS = Global Positioning System; LPS = Local Positioning System, VID = video system, SSC = sport-specific course; SHU = shuttle run; SSG = small-sided game.

https://doi.org/10.1371/journal.pone.0199519.t004

the spatial error (e.g. 0.45 cm during 90Êturns, see Table 3). Rapid speed and direction changes seem to be a challenge for the underlying Kalman filter, which is generally based on linear dynamical systems, thus suppressing rapid movement changes.

The only previous study that analyzed the spatial accuracy of the VID system reported an error of 73 cm dRMSE [17] (vs. 56-59 cm in the present study, see Table 2). Such a difference could be caused by either technological advancements in camera gear, different viewing angles, or the used criterion reference (LAVEG vs. VICON). For VID, we found significantly higher spatial errors in the 505 agility test of the sport-specific course (Table 3). Since players tend to lower their upper body to counteract accelerations occurring at the turning point, we assume that the visual tracking algorithm detects the center of the athletes' body at a lower height, thus leading to a spatial position shift in the vertical Y-axis.

To the best of our knowledge, information about the spatial accuracy of sport-specific GPS systems has not been reported prior to this study. This could be due to the fact that GPS systems are predominantly used to evaluate physical performance metrics (rather than spatial/tactical behavior). Nevertheless, it is incomprehensible why only limited information is available on the spatial accuracy of GPS systems, especially against the background that various studies made use of GPS coordinates to analyse spatial motion behaviour (e.g. position-specific centroids, team centroids and distance between centroids [24, 25]), as well as the fact that several commercial GPS systems determine distance metrics via differentiation of position data [21]. This study shows that the average spatial measurement error of GPS was 96 cm, almost twice as high as that of its nearest competitor VID (56 cm).

Instantaneous speed and acceleration accuracy

Considering that GPS exhibited the highest spatial errors, one would think that this error pattern should exert its influence on the vRMSE/aRMSE categories. Contrastingly, it is found that GPS speed errors were not significantly different from those of LPS (see Table 2). Whereas the vision-based and radar-based technology utilize differentiation of position data over time for speed determination, most commercially available GPS systems circumvent the problem of error propagation from fundamental to derived data by using two completely different measurement principles. Modern GPS systems can determine speed by measuring the rate of change in the satellites' electromagnetic signal frequency, also known as the Doppler effect [31, 32]. Doppler measurements are immune to cycle slips (temporary signal anomalies or low signal-to-noise ratio caused by obstructions such as buildings, trees, etc.) [33]. Thus, research works dealing with GPS speed measurement reveal that using GPS Doppler measurements can provide greater speed accuracy than indirect measurement, which is based on error-prone position data [34]. As a consequence, despite comparably inferior spatial accuracy values, GPS systems are capable of measuring instantaneous speed, and consequently acceleration, with comparatively higher accuracy.

As depicted in the results, errors of the VID system were lower in movements in the X-axis when compared to movements in the Y-axis. Thus, lowest vRMSE errors for section 3 (505) were achieved by VID. This specific test was carried out in parallel alignment to the camera view (X-axis). Apart from that, it is apparent that instant speed an acceleration errors of the LPS and GPS technology are fairly consistent whereas the errors of the video technology have proven to be considerably higher. These results demonstrate the importance of the most accurate possible detection of position in space. Any inaccuracy on the fundamental data (XY-positions) will otherwise lead to increased error propagation in the derived data category (instantaneous speed and acceleration).

KPI accuracy

Overall, lowest deviations can be observed in the total distance category. RMSE% ranged from 1.2% (VID during SSC) to 4.9% (GPS during SSGs). These differences are in line with previous literature on GPS (1.9%) [35] and LPS (1.6Ð2.0%) [7, 22]. Given a total distance of approximately 11.4 ± 1.0 km in professional soccer matches [36], an error of 4.9% would correspond to a discrepancy of 560 m, which in turn is more than half a standard deviation (1.0 km). It is therefore questionable to what extent EPTS with an apparently small error of e.g. 4.9% for total distance can sufficiently describe the performance hierarchy between players. In agreement with previous studies [6, 15], we found evidence that GPS units are capable of accurately measuring distances with low and moderate speed (see Table 4), whereas they still have problems with regard to tracking movements involving high-speed direction changes (e.g. 90-180Ê turns, see Table 3). GPS had the lowest sampling rate in this study (GPSports 15 Hz units are actually 5 Hz with interpolated data). Our results again confirm that a 5 Hz sampling rate only partially captures high-intensity movements involving frequent changes of direction. Similar findings have been identified by previous GPS validation studies [10, 15]. The generally high deviations in the lowest speed category of all systems (standing, <1 km·h^{-1}) can be attributed to the fact that standing phases practically never occurred during the exercises, and thus the values for the total standing distance were considerably low. Minor differences could therefore lead to high deviations. The same applies to high-intensity categories such as high speed distance or high acceleration distance. It should also be noted that the RMSE% increases significantly as the movement intensity increases. This characteristic error pattern is particularly obvious in the case of high-speed categories. Considering that the percentage deviation increases considerably in these relevant performance categories (e.g. high-speed distance during small-sided games: RMSE% ranged from 43.8% (LPS) to 97.6% (VID)), the present results confirm that to this day EPTS may not be accurate enough to measure high-speed and acceleration distances with a reasonable degree of accuracy [13]. The RMSE in the peak speed category ranged between 0.22 m·s^{-1} (GPS during shuttle runs) and 0.71 m·s^{-1} (VID during shuttle runs). These values reveal the technology-dependent accuracy variations of the VID system. As the movement direction of the shuttle run was conducted in the vertical (perpendicular) camera axis (Y-axis), VID tended to overestimate the peak speed during shuttle runs.

Limitations

It is regrettable that at the time being, there is no gold standard for a full-size pitch of team-based sports. Since the natural field of application of EPTS are official matches, this leads to the fact that they might not be validated in the scenarios that are most relevant to them.

We could provide optimum environmental conditions for LPS and near-optimum conditions for GPS but could only meet the minimum requirements in case of the the VID system. It can be assumed that results of the VID system improve under optimum conditions such as in stadiums with steep stands in close proximity to the pitch.

It is worth noting that our results are based on untreated raw data, as provided by the manufacturer's proprietary software. Therefore, it is to be expected that the validity of the tested EPTS could be further improved by additional data filtering procedures.

Finally, this study did not examine the inter-unit agreement, i.e. systematic or random differences between different sensors in GPS and LPS systems. This is important though, in case valid comparisons between different players and sessions are of interest in the sensor-based systems [15].

Conclusion

Collectively, results of this study revealed that largest differences between EPTS occurred at the spatial accuracy, whereas speed and acceleration errors of GPS were comparable to those of LPS. Yet one important insight in this regard is the noticeably large error margin in the third data category (accuracy of KPIs) that is independent of the respective system or technology, which we are still facing in EPTS in general. Especially in KPI categories that might have a high impact on practical decisions, such as high speed performance indicators, we found significant deviations from the gold standard. Thus, the primary aim of future activities should be aimed at diminishing these inherent errors. Until then, it is recommended that practitioners do not make direct comparisons between KPIs collected by different EPTS. Since there are typically different systems at work in competition and training, we encourage any development toward a standardization of internal algorithms. In case there is no hint available at different operational definitions for filtering techniques or KPIs in different systems, this means the sports practice is led astray. For the time being, a consequence in this regard is to conduct comparisons between EPTS on the level of XY-data, instantaneous speed, and acceleration data, in addition to merely comparing calculated KPIs.

Supporting information

S1 Video. Body sway visualization. Exemplary 3D animation of the center of mass (COM) displacement. Despite being static (the whole body is not traveling any distance in the conventional sense), the animation demonstrates that COM is constantly in motion, thus leading to unintended accumulation of travel distance. This example demonstrates the need for a distance calculation method that compensates this effect.
(MP4)

S2 Video. Visualization of the test procedures. Exemplary animation of the scouting video and the motion capture data. Red dots and lines indicate the center of mass (COM) that was projected to the ground plane. Colored lines represent the position as reported by the respective tracking technology.
(MP4)

S1 Dataset. Calibration run.
(CSV)

S2 Dataset. System database.
(XLSX)

Acknowledgments

The authors would like to thank the coaches and players from FC Augsburg for their supportive collaboration. We also thank Sebastian Wenninger, Thomas Blobel, Philipp Hofmann, and Philipp Wanner (from the Chair for Training Science and Sports Informatics) for their participation during the preparation and execution of the test. This research was supported by a research fund that was received from STATS LLC. The Technical University of Munich, represented by the Chair of Training Science and Sports Informatics, functioned as an independent third party for conducting this validation study. The terms of this arrangement have been reviewed and approved by the Technical University of Munich in accordance with its policy on objectivity in research.

Author Contributions

Conceptualization: Daniel Linke.

Data curation: Daniel Linke.

Formal analysis: Daniel Linke.

Funding acquisition: Daniel Linke, Martin Lames.

Investigation: Daniel Linke.

Methodology: Daniel Linke.

Project administration: Daniel Linke.

Resources: Daniel Linke, Martin Lames.

Software: Daniel Linke.

Supervision: Martin Lames.

Validation: Daniel Linke.

Visualization: Daniel Linke.

Writing ± original draft: Daniel Linke, Martin Lames.

Writing ± review & editing: Daniel Link, Martin Lames.

References

1. Akenhead R, Nassis G. Training load and player monitoring in high-level football: current practice and perceptions. International journal of sports physiology and performance. 2016; 11(5):587±593. https://doi.org/10.1123/ijspp.2015-0331 PMID: 26456711
2. Carling C, Bloomfield J, Nelsen L, Reilly T The role of motion analysis in elite soccer. Sports Medicine. 2008; 38(10):839±862. https://doi.org/10.2165/00007256-200838100-00004 PMID: 18803436
3. Buchheit M, Allen A, Poon TK, Modonutti M, Gregson W, Di Salvo V. Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and GPS technologies. Journal of Sports Sciences. 2014; 32(20):1844±1857. https://doi.org/10.1080/02640414.2014.942687 PMID: 25093242
4. Buchheit M, Simpson BM. Player-Tracking Technology: Half-Full or Half-Empty Glass?. International journal of sports physiology and performance. 2017; 12(Suppl 2):2±35. https://doi.org/10.1123/ijspp.2016-0499
5. Valter DS, Adam C, Barry M, Marco C. Validation of Prozone®: A new video-based performance analysis system. International Journal of Performance Analysis in Sport. 2006; 6(1):108±119. https://doi.org/10.1080/24748668.2006.11868359
6. Duffield R, Reid M, Baker J, Spratford W. Accuracy and reliability of GPS devices for measurement of movement patterns in confined spaces for court-based sports. Journal of Science and Medicine in Sport. 2010; 13(5):523±525. https://doi.org/10.1016/j.jsams.2009.07.003 PMID: 19853507
7. Frencken WG, Lemmink KA, Delleman NJ. Soccer-specific accuracy and validity of the local position measurement (LPM) system. Journal of Science and Medicine in Sport. 2010; 13(6):641±645. https://doi.org/10.1016/j.jsams.2010.04.003 PMID: 20594910
8. Ogris G, Leser R, Horsak B, Kornfeind P, Heller M, Baca A. Accuracy of the LPM tracking system considering dynamic position changes. Journal of Sports Sciences. 2012; 30(14):1503±1511. https://doi.org/10.1080/02640414.2012.712712 PMID: 22906154
9. Redwood-Brown A, Cranton W, Sunderland C. Validation of a real-time video analysis system for soccer. International Journal of Sports Medicine. 2012; 33(08):635±640. https://doi.org/10.1055/s-0032-1306326 PMID: 22510800
10. Varley MC, Fairweather IH, Aughey RJ. Validity and reliability of GPS for measuring instantaneous velocity during acceleration, deceleration, and constant motion. Journal of Sports Sciences. 2012; 30(2):121±127. https://doi.org/10.1080/02640414.2011.627941 PMID: 22122431
11. Rawstorn JC, Maddison R, Ali A, Foskett A, Gant N. Rapid directional change degrades GPS distance measurement validity during intermittent intensity running. PloS one. 2014; 9(4):e93693. https://doi.org/10.1371/journal.pone.0093693 PMID: 24733158
12. KöklüY, Arslan Y, Alemdaroğlu U, Duffield R. Accuracy and reliability of SPI ProX global positioning system devices for measuring movement demands of team sports. The Journal of sports medicine and physical fitness. 2015; 55(5):471±477. PMID: 25303067
13. Rampinini E, Alberti G, Fiorenza M, Riggio M, Sassi R, Borges T, et al. Accuracy of GPS devices for measuring high-intensity running in field-based team sports. International journal of sports medicine. 2015; 36(01):49±53. https://doi.org/10.1055/s-0034-1385866 PMID: 25254901
14. Beato M, Bartolini D, Ghia G, Zamparo P. Accuracy of a 10 Hz GPS Unit in Measuring Shuttle Velocity Performed at Different Speeds and Distances (5±20 M). Journal of human kinetics. 2016; 54(1):15±22. https://doi.org/10.1515/hukin-2016-0031 PMID: 28031753
15. Scott MT, Scott TJ, Kelly VG. The validity and reliability of global positioning systems in team sport: a brief review. The Journal of Strength & Conditioning Research. 2016; 30(5):1470±1490. https://doi.org/10.1519/JSC.0000000000001221

16. Randers MB, Mujika I, Hewitt A, Santisteban J, Bischoff R, Solano R, et al. Application of four different football match analysis systems: A comparative study. Journal of sports sciences. 2010; 28(2):171±182. https://doi.org/10.1080/02640410903428525 PMID: 20391091
17. Siegle M, Stevens T, Lames M. Design of an accuracy study for position detection in football. Journal of Sports Sciences. 2013; 31(2):166±172. https://doi.org/10.1080/02640414.2012.723131 PMID: 22994162
18. Hoppe MW, Baumgart C, Polglaze T, Freiwald J. Validity and reliability of GPS and LPS for measuring distances covered and sprint mechanical properties in team sports. PloS one. 2018; 13(2):e0192708. https://doi.org/10.1371/journal.pone.0192708 PMID: 29420620
19. Vickery WM, Dascombe BJ, Baker JD, Higham DG, Spratford WA, Duffield R. Accuracy and reliability of GPS devices for measurement of sports-specific movement patterns related to cricket, tennis, and field-based team sports. The Journal of Strength & Conditioning Research. 2014; 28(6):1697±1705. https://doi.org/10.1519/JSC.0000000000000285
20. Aughey RJ. Applications of GPS technologies to field sports. International Journal of Sports Physiology and Performance. 2011; 6(3):295±310. https://doi.org/10.1123/ijspp.6.3.295 PMID: 21911856
21. Malone JJ, Lovell R, Varley MC, Coutts AJ. Unpacking the black box: applications and Considerations for using GPS devices in Sport. International Journal of Sports Physiology and Performance. 2017; 12(2):18±26.
22. Stevens TG, de Ruiter CJ, van Niel C, van de Rhee R, Beek PJ, Savelsbergh GJ. Measuring acceleration and deceleration in soccer-specific movements using a local position measurement (LPM) system. International Journal of Sports Physiology and Performance. 2014; 9(3):446±456. https://doi.org/10.1123/ijspp.2013-0340
23. Akenhead R, French D, Thompson KG, Hayes PR The acceleration dependent validity and reliability of 10 Hz GPS. Journal of Science and Medicine in Sport. 2014; 17(5):562±566. https://doi.org/10.1016/j.jsams.2013.08.005 PMID: 24041579
24. Aguiar M, Gonçalves B, Botelho G, Lemmink K, Sampaio J. Footballers' movement behaviour during 2-, 3-, 4-and 5-a-side small-sided games. Journal of Sports Sciences. 2015; 33(12):1259±1266. https://doi.org/10.1080/02640414.2015.1022571 PMID: 25782702
25. Gonçalves BV, Figueira BE, Maças V, Sampaio J. Effect of player position on movement behaviour, physical and physiological performances during an 11-a-side football game. Journal of Sports Sciences. 2014; 32(2):191±199. https://doi.org/10.1080/02640414.2013.816761 PMID: 24016056
26. Townshend AD, Worringham CJ, Stewart IB. Assessment of speed and position during human locomotion using nondifferential GPS. Medicine & Science in Sports & Exercise. 2008; 40(1):124±132. https://doi.org/10.1249/mss.0b013e3181590bc2
27. Beetz M, Kirchlechner B, Lames M. Computerized real-time analysis of football games. IEEE pervasive computing. 2005; 4(3):33±39. https://doi.org/10.1109/MPRV.2005.53
28. Saini M, Kerrigan DC, Thirunarayan MA, Duff-Raffaele M. The vertical displacement of the center of mass during walking: a comparison of four measurement methods. Journal of Biomechanical Engineering. 1998; 120(1):133±139. https://doi.org/10.1115/1.2834293 PMID: 9675692
29. Virmavirta M, Isolehto J. Determining the location of the body's center of mass for different groups of physically active people. Journal of Biomechanics. 2014; 47(8):1909±1913. https://doi.org/10.1016/j.jbiomech.2014.04.001 PMID: 24742487
30. Cohen J. A power primer. Psychological bulletin. 1992; 112(1):155. https://doi.org/10.1037/0033-2909.112.1.155 PMID: 19565683
31. Schutz Y, Herren R. Assessment of speed of human locomotion using a differential satellite global positioning system. Medicine and Science in Sports and Exercise. 2000; 32(3):642±646. https://doi.org/10.1097/00005768-200003000-00014 PMID: 10731007
32. Zhang J, Zhang K, Grenfell R, Deakin R. Short note: On the relativistic Doppler effect for precise velocity determination using GPS. Journal of Geodesy. 2006; 80(2):104±110. https://doi.org/10.1007/s00190-006-0038-8
33. Bahrami M, Ziebart M. Instantaneous Doppler-aided RTK positioning with single frequency receivers. In: Position Location and Navigation Symposium (PLANS), 2010 IEEE/ION. IEEE; 2010. pp. 70±78.
34. Chalko TJ. High accuracy speed measurement using GPS (Global Positioning System). NU Journal of Discovery. 2007; 4:1±9.
35. Johnston RJ, Watsford ML, Kelly SJ, Pine MJ, Spurrs RW. Validity and interunit reliability of 10 Hz and 15 Hz GPS units for assessing athlete movement demands. The Journal of Strength & Conditioning Research. 2014; 28(6):1649±1655. https://doi.org/10.1519/JSC.0000000000000323
36. Di Salvo V, Baron R, Montero FC, Bachl N, Pigozzi F. Performance characteristics according to playing position in elite soccer. International Journal of Sports Medicine. 2007; 28(3):222±227. https://doi.org/10.1055/s-2006-924294 PMID: 17024626

Journal of Biomechanics 83 (2019) 319–323

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

Short communication

Impact of sensor/reference position on player tracking variables: Center of scapulae vs center of pelvis

Daniel Linke *, Martin Lames

Department of Sport and Health Sciences, Chair of Performance Analysis and Sports Informatics, Technical University of Munich, Germany

ARTICLE INFO

Article history:
Accepted 26 November 2018

Keywords:
Athlete tracking systems
Center of pelvis
Body reference
Tracking variables
Wearable sensors

ABSTRACT

Electronic performance and tracking systems (EPTS) traditionally rely on one of two body positions as the ultimate representative for the entire body in space: the upper torso between the scapulae (GPS- and radar-based systems) or the body's estimated center (optical and some radar-based systems). The aim of this study was to quantify the impact of the respective reference point upon the resulting kinematic tracking variables. We present a marker-based method comparing center of pelvis (COP) derived tracking variables with center of scapulae (COS) derived tracking variables in a 30×30 m (900 m^2) VICON measurement area. Fourteen male soccer players completed a running circuit with prescribed team-sport specific movements. Results showed that differences between COP and COS heavily depend on the underlying movement characteristic. Low-speed running showed the lowest deviations whereas accelerated movements and movements with sharp changes in direction lead to a significant increase in the observed differences. Results further showed that COS sprinting distance was on average −44.65% ($p < 0.001$) lower in comparison to COP. Similarly, maximum speed obtained from COS was −2.94% ($p = 0.001$) lower in comparison to COP. On the contrary, maximum acceleration values of COS were on average 16.15% ($p = 0.02$) higher compared to COP. Our work illustrates that the anatomical reference point used to represent the entire body in space needs to be carefully considered in the interpretation of tracking variables delivered by different EPTS.

1. Introduction

Monitoring players' locomotor demands during both training and competition is a common practice in professional team sports (Buchheit et al., 2014; Carling, 2013; Stevens et al., 2014). As a matter of fact, it is not uncommon for some players to be tracked by two or three different electronic performance and tracking systems (EPTS) during a regular week (Buchheit and Simpson, 2017). Consequently, interchangeability and agreement between different EPTS are of key importance to allow for a substantiated assessment of a player's external load and to integrate the data of different EPTS in a meaningful way. As a result, scientific literature provides various validation studies that aim at quantifying the difference or agreement between different EPTS (Linke et al., 2018; Buchheit et al., 2014; Johnston et al., 2014; Randers et al., 2010; Siegle et al., 2013). The key finding of these studies is that a between-systems agreement in terms of locomotor activity (e.g. distance covered, number of sprints and accelerations) proved problematic. Causes for these inconsistencies between different EPTS can primarily be traced to fundamental differences in the underlying technology (global positioning systems (GPS), radar-based local positioning systems (LPS), or semiautomatic camera systems) and data processing procedures, which are often chosen at the discretion of the manufacturer (Malone et al., 2017).

Apart from this, less attention was given to the aspect of the chosen anatomical reference point on the human body. Whereas wearable sensor systems (GPS and LPS units) are either attached to the upper thoracic spine between the scapulae (center of scapulae COS) or, in some cases, to the lumbosacral region of the spine, semiautomatic camera systems typically identify a pixel-based bounding box to estimate the body's center, hereinafter referred to as center of pelvis (COP) (Manafifard et al., 2017). Obviously, using different anatomical reference positions provides additional potential for inconsistencies to occur when comparing metrics provided by different EPTS. Therefore, the aim of this study was to quantify the differences between COP and COS derived tracking variables. The results could contribute to an improved understanding of performance parameters provided by EPTS.

---

* Corresponding author at: Department of Sport and Health Sciences, Chair of Training Science and Sports Informatics, Technical University of Munich, Uptown München - Campus D, Georg-Brauchle Ring 60/62, D-80992 Munich, Germany.
E-mail address: daniel.linke@tum.de (D. Linke).

https://doi.org/10.1016/j.jbiomech.2018.11.046
0021-9290/

2. Methods

2.1. Participants

Fourteen male elite youth soccer players (age: 17.4 ± 0.4 years, height: 178.6 ± 4.2 cm, body mass: 70.2 ± 6.2 kg) playing for the German Bundesliga team FC Augsburg participated in the study. Voluntarily signed informed consent to participate in the study was provided by both the players and their parents. Institutional board approval for the study was obtained from the Ethics Commission of the Technical University of Munich.

2.2. Protocol

Five adhesive marker mounts were glued on each participant's skin (acromio-clavicular joint of the right (RSHO) and left (LSHO) shoulder, left anterior superior iliac spine (LASI), right anterior superior iliac spine (RASI), and first sacrum vertebra (SACR)) (Fig. 1). The reflective markers were then fixed to the mounts through a tight-fitting compression shirt. COP was estimated by means of the reconstructed pelvis method (Saini et al., 1998), defined as the spatial center of RASI, LASI, and SACR. COS was defined as the spatial center of RSHO and LSHO.

A circuit with seven different exercises was used to analyze elementary movements of the kind typically encountered in field-based team sports: (1) 15 m sprint followed by 5 m deceleration, (2) 20 m sprint, (3) 10 m backward running followed by 10 m forward sprinting, (4) 505 agility test (maximum 180° turning ability), (5) 90° turns, (6) curved runs, and (7) low-speed recovery runs between exercises (<6 km·h^{-1}).

2.3. Data acquisition and processing

A 33-camera motion capture system (VICON, Oxford, UK) was utilized to determine marker positions at a frequency of 100 Hz. Retro-reflective 38.0 mm markers were used to assure stable recognition of the markers within the measurement area (30.0 × 30.0 m, 900.0 m^2). Raw displacement data was low pass filtered (4th order Butterworth) with a cut-off frequency of 10 Hz. Speed (scalar magnitude of velocity, as per rate of change in position) and acceleration (as per rate of change in speed) were derived from the filtered displacement data. Vertical (Z) displacement was neglected since EPTS typically report an athlete's motion in the horizontal plane (XY-coordinates). Kinematic deviations between COP and COS (momentary position, speed, and acceleration) were calculated both from an entire circuit (pooled) as well as for each of the seven exercises individually. Key performance indicators (KPI) were additionally calculated from the displacement data of an entire circuit (i.e. maximum speed, maximum acceleration, and distance traveled while running with different ranges of speed (low-speed (<6 km·h^{-1}), moderate-speed (6–15 km·h^{-1}), elevated-speed (15–20 km·h^{-1}), high-speed (20–25 km·h^{-1}), and sprinting (>25 km·h^{-1}))). Data were analyzed with MATLAB (Release 2017b, The MathWorks, Inc., Natick, MA) based on marker trajectories processed with Vicon Nexus (Version 2.3).

2.4. Statistical analysis

Deviations between COP and COS are presented as root mean square error (RMSE): dRMSE (m): distance RMSE (horizontal distance between COP and COS); vRMSE (m·s^{-1}): instant speed RMSE, and aRMSE (m·s^{-2}): instant acceleration RMSE. Each participant completed the circuit twice. Final analysis includes the mean of the resulting values. Non-parametric test procedures were used because the data was not normally distributed. Wilcoxon signed-ranks tests were used to test if the RMSEs between COP and COS differ from zero. Non-parametric Friedman tests were used to test whether the RMSEs differ between exercises. Dunn's test was used to for post-hoc comparisons of Friedman's results. Cohen's d and Kendall's W coefficient of concordance were calculated to measure the size of effects (Cohen, 1992). Descriptive statistics are presented as mean and standard deviation (SD). Statistical significance was set at $p < 0.05$.

3. Results

3.1. Position, speed and acceleration deviation

Wilcoxon Signed-Ranks tests showed that dRMSE ($Z = 3.29$, $p < 0.001$, $d = 0.88$), vRMSE ($Z = 3.33$, $p < 0.001$, $d = 0.89$) and

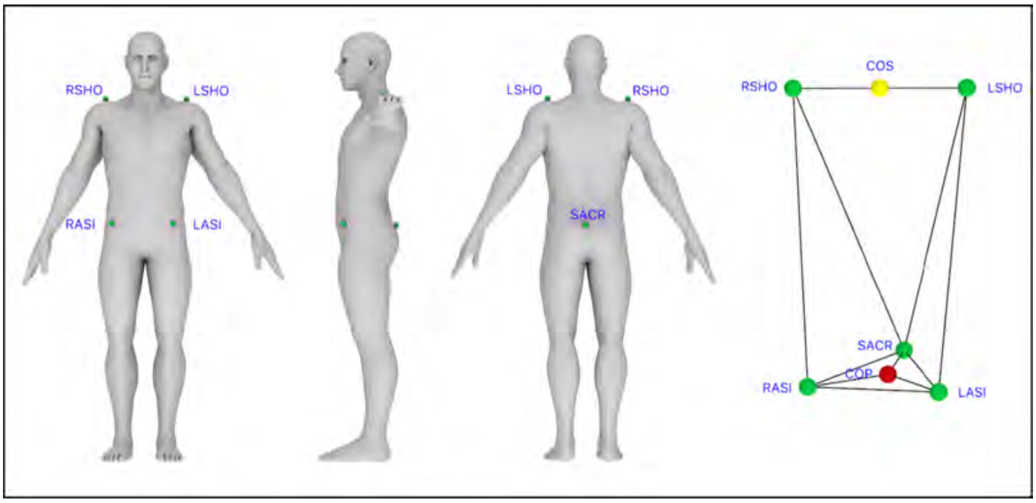


Fig. 1. Marker positions on the human body. Center of pelvis (COP) in red, center of scapulae (COS) in yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Mean deviation (root mean square error RMSE) between COP and COS based on all data points of each subject (n = 14). One-sample Wilcoxon Signed Rank Tests were used to test if the deviations between COP and COS were significantly different from zero. Standardized (Z) test statistic and resulting effect sizes (Cohen's d) are also displayed. CI = 95% confidence intervals. dRMSE = distance root mean square error (spatial horizontal distance between COP and COS); vRMSE = instant speed root mean square error; aRMSE = instant acceleration root mean square error.

	Mean ± SD		5–95% CI	Z	p	d
dRMSE (cm)	11.32	±1.61	[11.00, 12.32]	3.29	< 0.001	0.88
vRMSE (m·s^{-1})	0.24	±0.02	[0.24, 0.27]	3.33	< 0.001	0.89
aRMSE (m·s^{-2})	4.68	±0.50	[4.47, 5.03]	3.29	< 0.001	0.88

aRMSE ($Z = 3.29$, $p < 0.001$, $d = 0.88$) were significantly different from zero (Table 1). Non-parametric Friedman tests of RMSE differences among exercises rendered Chi-square values of 68.48 (dRMSE, $p < 0.001$, $W = 0.82$), 67.69 (vRMSE, $p < 0.001$, $W = 0.81$), and 61.96 (aRMSE, $p < 0.001$, $W = 0.74$). dRMSEs were smallest during recovery runs (7.31 ± 1.75 cm) and highest during 20 m sprints (22.65 ± 2.56 cm) and 505 agility tests (dRMSE = 24.75 ± 1.73 cm). Deviations in instantaneous speed and acceleration were also smallest during recovery runs (vRMSE = 0.16 ± 0.02 m·s^{-1}; aRMSE = 2.79 ± 0.32 m·s^{-2}) but highest during 90° turns (vRMSE = 0.52 ± 0.05 m·s^{-1}; aRMSE = 9.91 ± 1.25 m·s^{-2}) (Fig. 2). vRMSE was significantly higher during exercises with multidirectional movements (90° turns, curved runs, and 505 agility tests) compared to straight line movements (Fig. 2).

3.2. Key performance indicators KPI

Percentage deviations between KPIs derived by either COP or COS ranged from −0.32% (total distance) to −44.65% (sprinting distance), indicating that the magnitude of deviations increases with increasing speed of motion, which is also supported by large effect sizes ($d > 0.80$) in the elevated, high-speed, and very high-speed categories (Table 2).

4. Discussion

This study examined the impact of two anatomical reference points that are commonly used to represent the human body in space on kinematic tracking variables. Results show that the applied reference point has a significant influence on the athlete's momentary position, speed, and acceleration. Consequently, resultant KPIs also differ between COP and COS, sometimes considerably. Highest percentage deviations were found in the sprinting distance and maximum acceleration category. Since these KPIs are amongst the most important variables to be tracked in sports practice (Buchheit and Simpson, 2017), it is advised that coaches and practitioners consider the component of the technology-related anatomical reference point in their interpretation of EPTS derived tracking variables.

4.1. Spatial deviation dRMSE

Highest deviations in the horizontal plane were observed during movements with high accelerations/decelerations and sharp changes in direction (e.g. 505 agility tests, see Fig. 2), which is most probably an effect of the upper body's inclination angle. During the initial stage of sprinting, athletes tilt their upper body forward to direct ground reaction forces in the horizontal direction. As they approach their maximum speed, the torso is lifted gradually into an upright position (Mero et al., 1992). Accordingly, COS is positioned in front or behind COP, depending on the upper body's tilt. Similarly, sharp turns are associated with a lateral inclination of the upper torso (to counteract the occurring centrifugal forces) and COS is shifted laterally in relation to COP. Collectively, these circumstances indicate that the spatial deviation between COP and COS heavily depends on the underlying movement characteristic, resulting in an average dRMSE of 11.6 ± 1.6 cm (Fig. 2).

4.2. Momentary speed and acceleration deviation vRMSE & aRMSE

Momentary speed and acceleration deviation between COP and COS increased with speed and change in direction, which could be explained by the fact that the coupling strength between thorax and pelvis is inversely related to the movement frequency (Lamoth et al., 2002; Beek et al., 1995). Accordingly, the in-phase coordination present at low walking speeds gradually evolves towards antiphase coordination as the speed increases (Lamoth et al., 2002), resulting in an average deviation of 0.26 ± 0.03 m•s^{-1} in speed, and 4.78 ± 0.65 m•s^{-2} in acceleration, respectively (Fig. 2).

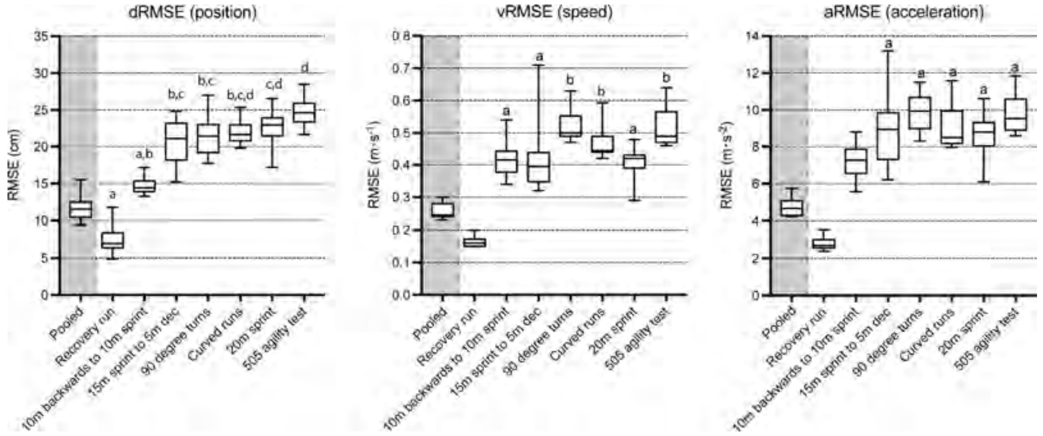


Fig. 2. Spatial deviation (dRMSE), instant speed (vRMSE) and instant acceleration (aRMSE) deviation depending on the exercise. Dunn's post hoc analyses were used to identify homogeneous exercise subsets (a, b, c, d: groups with the same letter designation are not significantly different. All other groups differ significantly from each other).

Table 2
Results of non-parametric Wilcoxon signed-ranks tests of KPI differences between COP and COS (n = 14). The percentage difference (% Diff) is presented from the COS perspective, i.e. a negative %Diff means that COS underestimates the respective KPI in comparison to COP. d = Cohen's d effect sizes. Z = standardized test statistic.

Metric	COP		COS					
	Mean	±SD	Mean	±SD	Z	p	d	% Diff
Total distance (m)	327.00	±15.25	326.42	±14.96	2.98	0.003	0.80	−0.32%
Low-speed distance (m)	89.46	±22.30	91.30	±22.45	2.35	0.018	0.63	2.01%
Moderate-speed distance (m)	114.02	±22.64	114.86	±21.82	1.29	0.119	0.30	0.73%
Elevated-speed distance (m)	48.46	±2.91	53.85	±4.12	3.29	0.001	0.88	10.00%
High-speed distance (m)	55.36	±4.76	52.45	±5.32	2.98	0.003	0.80	−5.56%
Very high-speed distance (m)	19.27	±1.80	13.32	±3.01	3.29	0.001	0.88	−44.65%
Maximum speed ($m{\cdot}s^{-1}$)	8.01	±0.15	7.78	±0.20	3.30	0.001	0.88	−2.94%
Maximum acceleration ($m{\cdot}s^{-2}$)	24.93	±4.07	29.73	±2.71	3.11	0.019	0.83	16.15%
Maximum deceleration ($m{\cdot}s^{-2}$)	−23.30	±2.27	−22.36	±2.55	1.01	0.265	0.29	−4.23%

4.3. Key performance indicators KPI

Maximum speed derived by COP was on average 3% higher than maximum speed derived by COS, possibly caused by the torso's tilting behavior during the transition from accelerated to decelerated running. Due to the relatively short sprint distance in this study (20 m), subjects reached maximum speed during the phase transition from acceleration to deceleration. At the same time, the torso's forward lean present in the acceleration phase transitions into a posterior lean during deceleration (Hewit et al., 2011), shifting COS posterior in relation to COP, which results in an earlier deceleration of COS compared to COP. Our results, therefore, suggest that COS underestimates maximum speed values during acceleration phases.

It was further shown that maximum accelerations derived by COS were on average 16% higher than those derived by COP, most probably due to the fact that, during acceleration, the flexion and extension of the trunk results in an increased range of motion of the upper torso compared to the pelvis (Schache et al., 1999). Contrastingly, maximum decelerations derived by COS were on average 4% lower than those derived by COP. Since the upper body plays an important role in damping gait-related oscillations during locomotion (Kavanagh et al., 2004), higher ground contact-induced decelerations in the pelvis region are well comprehensible.

The difference in the total distance covered category was statistically significant but rather trivial with respect to the percentage deviation (<1%). Similar results were found in low-speed (2%) and moderate-speed categories (<1%). However, deviations tended to increase as speed increased. COS derived distances in the elevated speed category were on average 10% higher than those derived by COP. On the contrary, COS derived distances in the highest speed categories (high speed & sprinting) were on average significantly smaller (−6% & −45%). A possible explanation for the significant underestimation of high-speed distances derived by COS could be grounded in COS's underestimation of maximum speed. As COP generally spent more time in high-speed categories, it is not surprising that the aggregated distances covered with high-speed differ considerably. Since high-speed distances account for only a small proportion of the total distance, small differences in this category can ultimately lead to high percentage deviations between COP and COS.

It should be mentioned that the raw data characteristics and data processing steps used in this study may not necessarily be comparable to data processing steps commonly applied in the field to EPTS. Whereas we decided to use a measurement device with the best possible spatiotemporal accuracy to investigate the original research question, there are no generally accepted definitions of required signal properties, optimal sampling rate or algorithmic specifications in the field of EPTS. As a result, tracking data obtained from different EPTS may vary, sometimes considerably, and thus transferability of results inevitably depends on the particular data processing procedures of the respective EPTS.

5. Conclusions

This study quantified the impact of two anatomical reference points (COP and COS) on resultant kinematic parameters. While only trivial deviations (<1%) were observed in total distance measures, high deviations were found in high-speed running distance (45%), maximum speed (3%), maximum acceleration (16%) and deceleration (4%) categories. Our results illustrate that the anatomical reference point used to represent the entire body in space needs to be carefully considered in the interpretation of tracking variables delivered by different EPTS.

Conflicts of interest

The authors declare no financial or personal relationships with other persons or organizations that might inappropriately influence our work presented herein.

Acknowledgements

This work was supported, in part, by a research fund received from STATS LLC. The authors would like to thank the coaches and players from FC Augsburg for their supportive collaboration. We also thank Sebastian Wenninger, Thomas Blobel, Philipp Hofmann, and Philipp Wanner (from the Chair for Training Science and Sports Informatics) for their participation during the preparation and execution of the test.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbiomech.2018.11.046.

References

Beek, P.J., Peper, C.E., Stegeman, D.F., 1995. Dynamical models of movement coordination.
Buchheit, M., Allen, A., Poon, T.K., Modonutti, M., Gregson, W., Di Salvo, V., 2014. Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and GPS technologies. J. Sports Sci. 32 (20), 1844–1857.
Buchheit, M., Simpson, B.M., 2017. Player-tracking technology: half-full or half-empty glass? Int. J. Sports Physiol. Performance 12 (Suppl 2). S2-35-S32-41.
Carling, C., 2013. Interpreting physical performance in professional soccer match-play: should we be more pragmatic in our approach? Sports Med. 43 (8), 655–663.
Cohen, J., 1992. A power primer. Psychol. Bull. 112 (1), 155.
Hewit, J., Cronin, J., Button, C., Hume, P., 2011. Understanding deceleration in sport. Strength Conditioning J. 33 (1), 47–52.

Johnston, R.J., Watsford, M.L., Kelly, S.J., Pine, M.J., Spurrs, R.W., 2014. Validity and interunit reliability of 10 Hz and 15 Hz GPS units for assessing athlete movement demands. J. Strength Condition. Res. 28 (6), 1649–1655.

Kavanagh, J.J., Barrett, R., Morrison, S., 2004. Upper body accelerations during walking in healthy young and elderly men. Gait Posture 20 (3), 291–298.

Lamoth, C., Beek, P.J., Meijer, O.G., 2002. Pelvis–thorax coordination in the transverse plane during gait. Gait and Posture 16 (2), 101–114.

Linke, D., Link, D., Lames, M., 2018. Validation of electronic performance and tracking systems EPTS under field conditions. PloS one 13 (7), e0199519.

Malone, J.J., Lovell, R., Varley, M.C., Coutts, A.J., 2017. Unpacking the black box: applications and considerations for using GPS devices in sport. Int. J. Sports Physiol. Perform. 12 (Suppl 2). S2-18-S12-26.

Manafifard, M., Ebadi, H., Moghaddam, H.A., 2017. A survey on player tracking in soccer videos. Comput. Vision Image Understand. 159, 19–46.

Mero, A., Komi, P., Gregor, R., 1992. Biomechanics of sprint running. A review. Sports Med. 13 (6), 376–392.

Randers, M.B., Mujika, I., Hewitt, A., Santisteban, J., Bischoff, R., Solano, R., Zubillaga, A., Peltola, E., Krustrup, P., Mohr, M., 2010. Application of four different football match analysis systems: a comparative study. J. Sports Sci. 28 (2), 171–182.

Saini, M., Kerrigan, D.C., Thirunarayan, M.A., Duff-Raffaele, M., 1998. The vertical displacement of the center of mass during walking: a comparison of four measurement methods. J. Biomech. Eng. 120 (1), 133–139.

Schache, A.G., Bennell, K.L., Blanch, P.D., Wrigley, T.V., 1999. The coordinated movement of the lumbo–pelvic–hip complex during running: a literature review. Gait Posture 10 (1), 30–47.

Siegle, M., Stevens, T., Lames, M., 2013. Design of an accuracy study for position detection in football. J. Sports Sci. 31 (2), 166–172.

Stevens, T.G., de Ruiter, C.J., van Niel, C., van de Rhee, R., Beek, P.J., Savelsbergh, G.J., 2014. Measuring acceleration and deceleration in soccer-specific movements using a local position measurement (LPM) system. Int. J. Sports Physiol. Perform. 9 (3), 446–456.

©Journal of Sports Science and Medicine (2018) **17,** 662-667
http://www.jssm.org

Research article

Decline in Match Running Performance in Football is affected by an Increase in Game Interruptions

Daniel Linke [1]✉, Daniel Link [1], Hendrik Weber [2] and Martin Lames [1]
[1] Technical University Munich, Chair of Performance Analysis and Sports Informatics, Munich, Germany; [2] DFL Deutsche Fußball Liga GmbH, Frankfurt, Germany

Abstract
This study quantified the contribution of game interruptions to the fatigue-related declines in match running performance over the course of a football match. Using a semi-automatic multiple camera system, the running activity of 792 individual German Bundesliga performances was divided into pre-defined 15-minute intervals and subsequently analysed under two prerequisites: with (effective playing time) and without (total playing time) consideration of game interruptions. Results showed a significant decline in effective playing time over the course of a match, from 66.3% of the total playing time in the first 15 minutes to 55.9% in the final 15 minutes of a match. Under consideration of the total playing time, match running performances decreased by 24.2% on average; considering the effective playing time, they decreased on average by only 10.2%. It can, therefore, be concluded that more than half (57.9%) of the commonly reported decline in match running performance cannot be assigned to physical fatigue, but rather to an increase in game interruptions as the game progresses. In conclusion, this study demonstrated for the first time that the decline in players' match running performance during football matches is substantially amplified by a proven increase in game interruptions, indicating that there may be a tendency among practitioners to overestimate fatigue-induced performance declines.

Key words: Football, fatigue, match running performance, effective playing time.

Introduction

Over the past two decades, fatigue development in football has become one of the primary research areas in the broader field of football physiology (Bangsbo et al., 1994; Mohr et al., 2003; 2010; Reilly, 1997; Bradley et al., 2013; Waldron and Highton, 2014). Among mental and tactical factors encompassing fatigue, particular attention was paid to the element of physical fatigue (Paul et al., 2015), leading to a wealth of time–motion analyses pertaining to match running performance in top-level football (Bradley et al., 2009; Di Mascio and Bradley, 2013; Di Salvo et al., 2009; Rampinini et al., 2007; Weston et al., 2011). To evaluate physical fatigue in terms of a decline in players' match running performance, the approach taken by most studies has been to compare temporal running patterns in the early and later stages of a match. Using such segmentation methods, studies have demonstrated a decline in running performance between the first and second halves of football matches, particularly in total distance covered (TD), time spent in high intensity running (HIR) and number of sprints (Di Salvo et al., 2009; Mohr et al., 2003; Rampinini et al., 2009). Others have examined match-running performance across 15-min periods, reporting significant reductions of HIR and acceleration efforts over the course of a match (Akenhead et al., 2013; Bradley et al., 2010). A more detailed analysis on minute-by-minute observations revealed that by eight minutes into the second half the median distance per minute had already substantially decreased in comparison to the corresponding median distance in the first half (Barros et al., 2007).

To gain practice-oriented recommendations from such observations, sports scientists need to consider the work rate-specific characteristic of the sport. Football is an intermittent sport that involves frequent but brief periods of high-intensity movement, interspersed with lower intensity running (Bangsbo, 1994) and matches are usually composed of a series of play periods randomly interspersed with game stoppages, such as when the referee has called an infringement, or the ball is off the playing field (Wallace and Norton, 2014). Accordingly, a study of game interruptions in elite football showed that matches are halted on average for 38% of the total match time (Siegle and Lames, 2012). More importantly, evaluations of the time the ball is in play over predefined match periods have indicated that the duration of game interruptions increases towards the end of a match (Carling and Dupont, 2011), indicating that an increase in game interruptions towards the end of a match could have an impact on match running performance. However, only one study has analyzed football-specific match running performance while considering the effective playing time (T_{eff}), and this was based only on an entire match without consideration of fatigue development (Castellano et al., 2011). Thus, to the best of our knowledge, no study has evaluated the contribution of game interruptions to the decline in match running performance in professional football.

Therefore, the aims of this study were twofold: (1) to test whether a verifiable decrease in effective playing times occurs over the course of a match, and (2) to quantify the contribution of game interruptions to the decline in match running performance as the game progresses.

Methods

Participants and sample size

The sample comprised positional data from 51 matches of the German Bundesliga across the 2012/2013 (n = 21) and 2013/2014 (n = 30) season. Analysis included data for outfield players (goalkeepers excluded) who completed the

Received: 09 April 2018 / Accepted: 08 October 2018 / Published (online): 01 December 2018

full duration of the match, excluding substitutes, as their performance differs significantly from the performance of the players they replace (Carling et al., 2010). In addition, only matches with a narrow end result (one-goal difference or tie games) were considered for analysis. Based on these criteria, the match running performances of 792 players (single observations) were examined. It was a condition of players' employment that data such as those used in this study could be obtained for routine assessment of their performance during the competitive Bundesliga season. Hence, the usual ethics committee approval was not required (Winter and Maughan, 2009). All subject identifiers were removed to ensure confidentiality. This study conformed to the recommendations of the Declaration of Helsinki.

Data collection

Analysis included the official match running performance data of the Deutsche Fußball Liga GmbH, which were assessed by a computerized multiple-camera tracking system (TRACAB®, Stockholm, Sweden) operating at 25 Hz. This tracking system semi-automatically assesses the match running performance data of all players, the position of the ball, and corresponding match events (such as game stoppages), allowing the quantification of effective playing time (T_{eff}) as the total playing time (T_{tot}) minus all game stoppages such as for fouls, goals, free kicks, substitutions and injuries, i.e. the total time during the match that the ball is in play (Castellano et al., 2011).

Using a football-specific criterion measure approach pioneered by Reilly and Thomas (1976) (for a review, see Carling et al., 2005; Reilly, 2003), an independent study on the validity and reliability of the TRACAB® tracking system reported average measurement errors of 2% for measures of distance covered.

Match analysis

To investigate temporal patterns in match running performance, data were divided into six pre-defined 15-minute match periods (three per half). Periods of extra time at the end of the first and second halves were excluded from the analysis. The categorical independent variables were (i) the match period, and (ii) the method of time registration (T_{eff} vs. T_{tot}). The continuous dependent variable was the match running performance data. In accordance with previous research (Bradley et al., 2009; Carling et al., 2008; Carling and Dupont, 2011; Mohr et al., 2003), match running data were divided into the following categories: total distance, walking (0.7–7.2 km·h^{-1}), jogging (7.2–14.4 km·h^{-1}), running (14.4–19.8 km·h^{-1}), high-speed running (19.8–25.1 km·h^{-1}), and sprinting (>25.2 km·h^{-1}). Match running performance parameters calculated for the T_{eff} condition included only the performances that took place when the ball was in play, whereas the parameters calculated for the T_{tot} condition included the performance over the entire period of the match. To allow sound comparison of physical performance in both T_{eff} and T_{tot}, all distance categories were converted to a relative analysis per unit of time (m·min^{-1}).

The chosen procedure to analyse the contribution of game interruptions to the decline in match running performance consisted of the following concept: for each match period, each player's match running performance was calculated under two conditions (T_{eff} and T_{tot}). Under the assumption that T_{eff} decreases towards the end of a match, it is to be expected that the observed difference in match running performance between the two conditions (T_{eff} and T_{tot}) increases equally. Therefore, the hypothesis was tested whether the difference in match running performance between the two conditions (T_{eff} and T_{tot}) increases towards the end of a match.

Statistical analysis

A one-way analysis of variance (ANOVA) was conducted to test the hypothesis whether (i) a verifiable decrease T_{eff} occurs over the course of a match, and (ii) whether the observed difference in match running performance in the two conditions (T_{eff} and T_{tot}) increases over the course of a match. Eta-squared η^2 effect sizes were classified as small (0.01–0.05), medium (0.06–0.14) and large (>0.14), in accordance with the recommendation of Cohen (1988). Bonferroni's post hoc analyses were used when significant differences were found to compare performance parameters in the first (1'-15') and final (76'–90') match periods. Cohen's d effect sizes for Bonferroni's t-tests were classified as trivial (0-0.19), small (0.20–0.49), medium (0.50–0.79) and large (>0.80). Because preliminary analyses indicated that the distributions of match running performances were skewed, log transformation was used, which yielded normally distributed data. Statistical significance was set at $p < .05$. Data are presented as the mean ± standard deviation (SD). Statistical analyses were performed using the software package SPSS version 23.0, (IBM Corp., Armonk, NY, USA).

Results

Effective playing time

There was a statistically significant difference in T_{eff} between the six match periods as determined by one-way ANOVA [$F(5,480) = 9.844$, $p < 0.001$, $\eta^2 = 0.134$]. Pairwise comparisons of the means using Bonferroni's post-hoc procedure indicated that T_{eff} was significantly lower ($p < 0.001$, $d = 1.12$) in the last match period (33.5 ± 4.9 s·min^{-1}, 55.9% of T_{tot}) compared to the first match period (39.8 ± 6.3 s·min^{-1}, 66.3% of T_{tot}). There was no significant difference between the second, third, and fourth match period ($p = 0.242$). These results affirm the hypothesis that length and frequency of game interruptions do, in fact, increase towards the end of a match, therefore leading to a significant reduction of T_{eff} (see Figure 1).

Match running performance

A one-way ANOVA was conducted to compare the effect of the progressive match period on the difference in match running performance between the two conditions (T_{eff} and T_{tot}). There was a significant effect of the match period on difference in total distance covered per player [$F(5, 9204) = 75.7$, $p < 0.001$, $\eta^2 = 0.08$]. Bonferroni's post hoc comparisons indicated that the mean difference in the first (50.6 ± 18.0 m·sec^{-1}) match period was significantly lower than

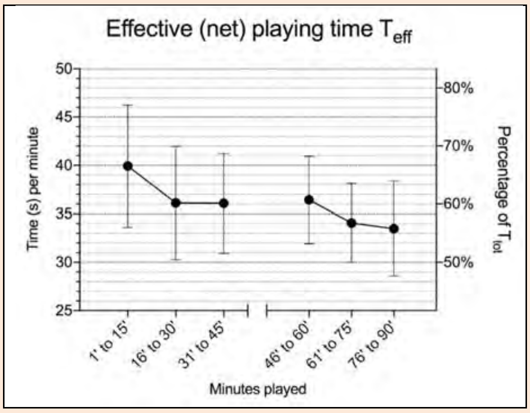


Figure 1. **Development of effective playing time T_{eff} over the course of a match. Data were obtained for six 15-min periods of match-play (presented as means and standard deviation).**

the difference in the last (61.9 ± 14.2 $m \cdot sec^{-1}$) match period [t(767) = 15.8, $p < 0.001$, d = 0.69] (see Table 1 and Figure 2). Under consideration of T_{tot}, total distance decreased from 99.1 ± 16.0 $m \cdot min^{-1}$ during the first match period to 78.0±13.8 $m \cdot min^{-1}$ during the final match period ($p < 0.001$, d = 1.41), corresponding to a decline of 21.2%. In comparison, under consideration of T_{eff}, total distance decreased from 149.7 ± 14.9 $m \cdot min^{-1}$ during the first match period to 139.8 ± 16.8 $m \cdot min^{-1}$ during the last match period ($p < 0.001$, d = 0.62), corresponding to a decline of merely 6.6 %. Comparable results were found for the jogging, running, high- speed running, and sprinting category (see Table 1 and 2). It should be noted, however, that the amplifying effect decreases with increasing running speeds, expressed in lower effect sizes in higher speed categories.

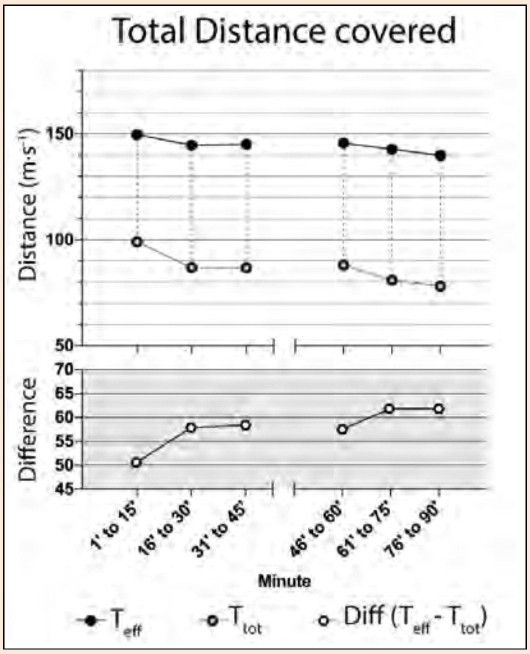


Figure 2. **Development of match running performance over the course of a match (using the example of total distance covered).**

Table 1. **Results from one-way ANOVA and Bonferroni's post hoc procedure showing the effect of the match period on the difference in the respective performance category between T_{eff} and T_{tot}.**

	ANOVA of Differences			Bonferroni's post hoc procedure				
	F	< p	η^2	1' to 15'	76' to 90'	Change	<p	d
Total Distance	75.7	.001	.08	50.6 ± 18.0	61.9 ± 14.2	22.2%	.001	.69
Walking	225.6	.001	.20	10.7 ± 3.5	15.7 ± 3.6	48.1%	.001	1.45
Jogging	32.5	.001	.03	24.1 ± 8.5	27.9 ± 7.6	15.8%	.001	.47
Running	16.5	.001	.02	10.7 ± 5.7	12.1 ± 5.2	13.5%	.001	.27
High Speed Running	21.2	.001	.02	3.9 ± 2.7	4.6 ± 2.8	18.0%	.001	.26
Sprinting	7.9	.001	.01	1.4 ± 1.5	1.7 ± 1.7	22.8%	.006	.20

Table 2. **Descriptive statistics: difference in match running performance ($m \cdot min^{-1}$) between the first (minute 1' to 15') and last (minute 76' to 90') match period.**

	Effective playing time T_{eff}					Total playing time T_{tot}				
	1' to 15'	76' to 90'	Change	< p	n	1' to 15'	76' to 90'	Change	< p	n
Total Distance	149.7 ± 14.9	139.8 ± 16.8	-6.6%	.001	.62	99.1 ± 16.0	78.1 ± 13.8	-21.2%	.001	1.41
Walking	32.3 ± 6.3	36.0 ± 6.6	11.4%	.001	.57	21.6 ± 6.0	20.2 ± 5.1	-6.7%	.001	.25
Jogging	71.3 ± 13.0	62.9 ± 14.1	-11.9%	.001	.62	47.4 ± 11.2	35.2 ± 9.5	-25.7%	.001	1.17
Running	30.5 ± 10.4	26.8 ± 9.6	-12.0%	.001	.37	20.0 ± 6.6	14.8 ± 5.3	-25.6%	.001	.87
High speed running	11.1 ± 5.2	10.3 ± 5.3	-7.1%	.001	.15	7.2 ± 3.2	5.7 ± 2.8	-20.8%	.001	.50
Sprinting	3.3 ± 3.2	2.8 ± 3.1	-13.6%	.003	.16	2.1 ± 1.9	1.5 ± 1.6	-27.6%	.001	.34

Discussion

Based on the assumption that game interruptions increase towards the end of a football match, this study examined the resulting effects on the frequently reported decline in match running performance in professional football. In contrast to previous studies focusing on the decline of match running performance, this study compared the magnitude of the observed declines - considering both effective and total playing times.

As a preliminary result, this study confirms that an increase in game interruption time does, in fact, lead to a 10.4% decline in effective ball-in-play time towards the end of a match. Although changes of T_{eff} over the course of a match have rarely been documented, these findings are in good accordance with previous research by Carling and

Dupont (2011), who reported an overall decline of eight percentage points of T_{eff} in French Ligue 1 matches (from 66% T_{eff} in the first period (0–5 min) to 58% in the last period (85–90 min)).

Direct comparison of the decline in match running performance with previous research is only possible to a limited extent, as only few studies have provided detailed results of individual performance indicators in each respective match period. In addition, different timeframes (1', 5', 15' or 45' periods) as well as different definitions of speed categories have been chosen for investigating match related fatigue patterns.

Under the T_{tot} condition, total distance decreased by 21.2% over the course of a match, whereas a similar study reported a 14.3% decline of total distance in UK Premier League players between the first and last 5-min period (Weston et al., 2011). In another study, Carling and Dupont (2011) found that high-intensity running distance (comprising high speed running and sprinting distance) covered by French Ligue 1 midfielders decreased by 24.2% between the first and last 5-min period. Similarly, Bradley et al. (2010) found that Premier League players covered 17.8% less high-intensity running distance in the last 15-min period compared to the first. Players in the present study covered 20.8% less high-speed running distance and 27.6% less sprinting distance in the same period. It could, therefore, be assumed that the results of the present study are generally consistent with those of previous research on match-related fatigue, as far as T_{tot}-based analyses are concerned. However, a unique element of this investigation was the ability to differentiate between performance declines based on both T_{eff} and T_{tot}.

As shown in Table 2, walking distance increased by 11.4% under the T_{eff} condition. Contrastingly, walking distance decreased by 6.7% under the T_{tot} condition. In consequence, higher gains of walking distance under the T_{eff} condition could indicate fatigue-related gains in walking distance during active gameplay. The decrease of walking distance under the T_{tot} condition, by contrast, probably results from an increase of game interruptions, leading to more standing phases, and thus also less distance covered in total.

The key finding of this study was that all other investigated performance indicators decreased by significantly less when game interruptions were considered. This was evident in both percentage declines of match running performance and a significant increase of the difference in match running performance between the two conditions (T_{eff} and T_{tot}). With respect to the total playing time, jogging, running, high intensity running, and sprinting distance decreased by 24.2% on average, whereas they decreased by only 10.2% with respect to the effective playing time. In other terms, the reported decline in match running performance under T_{tot} is more than twice as high as under the T_{eff} condition, indicating that approximately 57.9% of the decline in match running performance observed for T_{tot} is caused by an increase of game interruptions and thus cannot be related to physical fatigue.

It is noteworthy, however, that there were notable effect size differences for the different performance indicators. Specifically, smaller effect sizes were found in higher speed categories (see Table 1). Collectively, these results demonstrate that the influence of game interruptions on the decline in match running performance decreases with increasing movement intensity. Only small or even no effect sizes were found in the categories with a predominantly high-intensity nature (high intensity running and sprinting). In contrast, medium effect sizes were present in the medium-intensity categories such as total distance, walking, jogging and running distance. This may be explained by the predominantly intermittent nature of activity patterns of football, with players switching between brief bouts of high-intensity running and longer periods of low-intensity exercise (Rampinini et al., 2007). Thus, a majority of the total distance is covered at low or medium intensities, resulting in strong effect sizes for the associated movement categories.

However, it can be presumed that the occurrence of high-intensity movement is highly dependent on a variety of randomly occurring factors, such as sudden opportunities that require short and intense efforts to gain an advantage over the opponents, which can occur at any time during the match. It can be concluded that, regardless of physical fatigue towards the end of a match, players are able to perform at high intensity at any time, whenever necessary.

From a statistical perspective, the magnitude of match-related fatigue can be quantified more effectively for low and medium intensity categories. The often-described decrease in high-intensity movement categories (Bradley et al., 2009; 2010; Mohr et al., 2003) appears to be attenuated considerably when considering the effective playing time.

The limitations of this study include a lack of control for position-specific subdivisions of players, seasonal variation, match importance, and international differences. The observed patterns may, therefore, be a reflection of this specific league. In particular, match status (winning, drawing or losing) is a factor that has attracted increasing attention in the scientific literature, with some studies suggesting that the current score of a match has a considerable influence on match-related performance outcomes (Taylor et al., 2008). Bradley and Noakes (2013) demonstrated that players were able to maintain their high-intensity running performance in the second half of matches in which they were losing heavily, but this was not evident in matches where they were well ahead in score; this, in turn, could indicate players had a pacing strategy in an attempt to avoid unnecessary fatigue during clear results (Edwards and Noakes, 2009).

It should further be mentioned that concerns have been expressed about the attempted to quantify accumulated fatigue by comparing match-running activity during the first game period with that of the final game period (Carling, 2013). Reasons for such concerns are based on the frantic nature of the very first phases of gameplay in which teams show engagement to register their presence with the opposition (Bangsbo et al., 1991). Such purely psychological factors could, therefore, be mistakenly interpreted as physical fatigue. Overall, physical performance in football is influenced by a great number of factors, all of which can hardly be considered collectively (Carling, 2013). It seems likely that no single study would be able to comprehensively measure and control for all extraneous influences (Paul et al., 2015). Thus, caution is needed before attributing our findings to the nature of football. Finally, future work could investigate the influence of the effective playing time on the proven decline in high-intensity running immediately after the most intense 5-min period (Bradley et al., 2009).

Nevertheless, the intention of this study was to provide a basic overview of the influence of game interruptions on the physical performance of football players, with a focus on the decline in running performance. For sports scientists and coaches, knowledge of fatigue patterns considering the effective playing time provides a more accurate representation of competitive physical demands, and this, in turn, can be applied in training to develop practice drills that are more closely tailored to actual match requirements (Castellano et al., 2011).

Conclusion

This study demonstrated for the first time that the decline in players' match running performance during football matches is substantially amplified by a proven increase in game interruptions as the game progresses, indicating that there may be a tendency among practitioners to overestimate fatigue-induced performance declines. Previous studies with the objective to quantify a reduction in match running performance should, therefore, be interpreted with caution, as game interruptions are often omitted from these studies.

Acknowledgements

This work was supported by the German Research Foundation (DFG) and the Technical University of Munich within the funding program Open Access Publishing. The reported experiments comply with the current laws of the country; in which they were performed. The authors have no conflicts of interests to declare.

References

Akenhead, R., Hayes, P.R., Thompson, K.G. and French, D. (2013) Diminutions of acceleration and deceleration output during professional football match play. *Journal of Science and Medicine in Sport* **16(6),** 556-561.

Bangsbo, J. (1994) The physiology of soccer-with special reference to intense intermittent exercise. *Acta Physiologica Scandinavica. Supplementum* **619,** 1-155.

Bangsbo, J., Nørregaard, L. and Thorsøe, F. (1991) Activity profile of competition soccer. *Canadian Journal of Sports Sciences* **16(2),** 110–116.

Barros, R.M., Misuta, M.S., Menezes, R.P., Figueroa, P.J., Moura, F.A., Cunha, S.A., Anido, R. and Leite, N.J. (2007). Analysis of the distances covered by first division Brazilian soccer players obtained with an automatic tracking method. *Journal of Sports Science and Medicine* **6(2),** 233-242.

Bradley, P.S., Di Mascio, M., Peart, D., Olsen, P. and Sheldon, B. (2010) High-intensity activity profiles of elite soccer players at different performance levels. *Journal of Strength and Conditioning Research* **24(9),** 2343-2351.

Bradley, P.S. and Noakes, T.D. (2013) Match running performance fluctuations in elite soccer: indicative of fatigue, pacing or situational influences? *Journal of Sports Sciences* **31(15),** 1627-1638.

Bradley, P.S., Sheldon, W., Wooster, B., Olsen, P., Boanas, P. and Krustrup, P. (2009) High-intensity running in English FA Premier League soccer matches. *Journal of Sports Sciences* **27(2),** 159-168.

Carling, C. (2013) Interpreting physical performance in professional soccer match-play: should we be more pragmatic in our approach? *Sports Medicine* **43(8),** 655-663.

Carling, C., Bloomfield, J., Nelsen, L. and Reilly, T. (2008). The role of motion analysis in elite soccer: contemporary performance measurement techniques and work rate data. *Sports Medicine* **38(10),** 839-862.

Carling, C. and Dupont, G. (2011) Are declines in physical performance associated with a reduction in skill-related performance during professional soccer match-play? *Journal of Sports Sciences* **29(1),** 63-71.

Carling, C., Espie, V., Le Gall, F., Bloomfield, J. and Jullien, H. (2010) Work-rate of substitutes in elite soccer: a preliminary study. *Journal of Science and Medicine in Sport* **13(2),** 253-255.

Carling, C., Williams, A.M. and Reilly, T. (2005) *The Handbook of Soccer Match Analysis*. Taylor & Francis: London.

Castellano, J., Blanco-Villasenor, A. and Alvarez, D. (2011) Contextual variables and time-motion analysis in soccer. *International Journal of Sports Medicine* **32(6),** 415-421.

Cohen, J. (1988) *Statistical power analysis for the behavioural sciences.* 2nd ed.. Hillsdale, N.J.: L. Erlbaum Associates.

Di Mascio, M. and Bradley, P.S. (2013) Evaluation of the most intense high-intensity running period in English FA premier league soccer matches. *Journal of Strength and Conditioning Research* **27(4),** 909-915.

Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P. and Drust, B. (2009) Analysis of high-intensity activity in Premier League soccer. *International Journal of Sports Medicine* **30(3),** 205-212.

Edwards, A.M. and Noakes, T.D. (2009) Dehydration: cause of fatigue or sign of pacing in elite soccer? *Sports Medicine* **39(1),** 1-13.

Liu, H.Y., Hopkins, W., Gomez, M.A. and Molinuevo, J.S. (2013) Inter-operator reliability of live football match statistics from OPTA Sportsdata. *International Journal of Performance Analysis in Sport* **13(3),** 803-821.

Mohr, M., Krustrup, P. and Bangsbo, J. (2003) Match performance of high-standard soccer players with special reference to development of fatigue. *Journal of Sports Sciences* **21(7),** 519-528.

Mohr, M., Mujika, I., Santisteban, J., Randers, M.B., Bischoff, R., Solano, R., Hewitt, A., Zubillaga, A., Peltola, E. and Krustrup, P. (2010) Examination of fatigue development in elite soccer in a hot environment: a multi-experimental approach. *Scandinavian Journal of Medicine and Science in Sports* **20(Suppl 3),** 125-132.

Paul, D.J., Bradley, P.S. and Nassis, G.P. (2015) Factors affecting match running performance of elite soccer players: shedding some light on the complexity. *International Journal of Sports Physiology and Performance* **10(4),** 516-519.

Rampinini, E., Coutts, A.J., Castagna, C., Sassi, R. and Impellizzeri, F.M. (2007) Variation in top-level soccer match performance. *International Journal of Sports Medicine* **28(12),** 1018-1024.

Rampinini, E., Impellizzeri, F.M., Castagna, C., Coutts, A.J. and Wisloff, U. (2009) Technical performance during soccer matches of the Italian Serie A league: effect of fatigue and competitive level. *Journal of Science and Medicine in Sport* **12(1),** 227-233.

Reilly, T. (1997). Energetics of high-intensity exercise (soccer) with particular reference to fatigue. *Journal of sports sciences* **15(3)**, 257-263.

Reilly, T. (2003) Motion analysis and physiological demands. In: *Science and Soccer.* EDS: Reilly, T. & Williams, A.M. Routledge: London. 9-72).

Reilly, T. and Thomas, V. (1976). A motion analysis of work-rate in different positional roles in professional football match-play. *Journal of Human Movement Studies* **2,** 87-97.

Siegle, M. and Lames, M. (2012) Game interruptions in elite soccer. *Journal of Sports Sciences* **30(7),** 619-624.

Taylor, J.B., Mellalieu, S.D., James, N. and Shearer, D.A. (2008) The influence of match location, quality of opposition, and match status on technical performance in professional association football. *Journal of Sports Sciences* **26(9),** 885-895.

Waldron, M. and Highton, J. (2014) Fatigue and pacing in high-intensity intermittent team sport: an update. *Sports Medicine* **44(12),** 1645-1658.

Wallace, J.L. and Norton, K.I. (2014) Evolution of World Cup soccer final games 1966-2010: game structure, speed and play patterns. *Journal of Science and Medicine in Sport* **17(2),** 223-228.

Weston, M., Drust, B. and Gregson, W. (2011) Intensities of exercise during match-play in FA Premier League referees and players. *Journal of Sports Sciences* **29(5),** 527-532.

Winter, E.M. and Maughan, R.J. (2009) Requirements for ethics approvals. *Journal of Sports Sciences* **27(10),** 985.

Key points

- The effective ball-in-play time decreases from 66.3% of the total playing time in the first 15 minutes to 55.9% in the last 15 minutes of a match.
- Under consideration of the absolute playing time, match running performances decreased by 24.2% on average, whereas they decreased by only 10.2% under consideration of the effective playing time.
- Accordingly, 57.9% of the commonly reported decline in match running performance cannot be assigned to physical fatigue, but rather to an increase in game interruptions as the game progresses.

AUTHOR BIOGRAPHY

Daniel LINKE
Employment
Sports Scientist
Technical University Munich
Degree
PhD student
Research interests
Team sports performance analysis, validity and reliability of athlete tracking technologies
E-mail: daniel.linke@tum.de

Daniel LINK
Employment
Senior Lecturer
Technical University Munich
Degree
PhD
Research interests
Team sport performance analysis, sport analytics, sports data quality
E-mail: daniel.link@tum.de

Hendrik WEBER
Employment
CEO Sportec Solutions GmbH
(Deutsche Fußball Liga DFL)
Degree
PhD
Research interests
Team sports performance analysis, sports validation methods, performance indicators in football
E-mail: hendrik.weber@dfl.de

Martin LAMES
Employment
Professor
Technical University Munich
Degree
PhD
Research interests
Sports and exercise science, game sports research
E-mail: martin.lames@tum.de

✉ **Daniel Linke**
Technical University Munich, Chair of Performance Analysis and Sports Informatics, Georg-Brauchle-Ring 60/62, Munich, Germany

Substitutions in elite male field hockey – a case study

Daniel Linke & Martin Lames

Technical University Munich, Department for Training and Computer Science in Sports, Georg-Brauchle-Ring 60/62, Munich, Germany.

Abstract

The purpose of this study was to (i) describe and analyse the substitution tactics of an international field hockey team during competitive match play and (ii) to identify the impact of bench periods on the physical output of players when re-entering active game play.

Data were collected from thirteen professional male field hockey players during three international tournament matches. The physical performance of the players was recorded by a high-frequency local position measurement system. The number and length of substitutions were obtained by observation.

An average of 58.0 ± 4.6 substitutions were registered during the course of a tournament match, where the average player performed 4.7 ± 0.8 individual substitutions with an average pause duration of 5.4 ± 1.2 min and an average on-field playing duration of 7.18 ± 2.14 min. Within the first minute of substitution, re-entering players covered a significantly larger total distance (159.7 ± 33.0 m · min^{-1}) compared with the team average (139.4 ± 33.3 m · min^{-1}). A significant decrease in physical performance was observed within the fifth minute after substitution.

Our findings suggest a 'first-minute-rush-effect', wherein substitutes covered a significantly larger total distance compared with the team average. Further, re-entering field hockey players experienced distinguishable signs of fatigue approximately four minutes after substitution.

Key words: hockey, substitution, time-motion analysis, physical exertion, LPM

1. Introduction

A characteristic feature of the rules in field hockey is the possibility of unlimited substitutions, meaning that there is no limit to the number of players who can be substituted at the same time or to the number of times a single player can be substituted throughout the course of a match. With sixteen players eligible to play per match and eleven players on the pitch, the coach is allowed to frequently substitute up to five players simultaneously, which allows for spontaneous tactical changes as well as the counteraction of the inevitable physical fatigue that occurs over the course of the game. The counteraction of fatigue may primarily be achieved through the optimal relation between the active playing time spent on the pitch and the recovery time off the pitch after being substituted, also known as the work-rest ratio.

Despite the possibility of multiple substitutions, previous studies have reported a significant in-game decrease in the work load of professional hockey players (Lythe and Kilding, 2011). Physical fatigue in particular results in reduced performance in high-speed running activities (Jennings *et al.*, 2012a). Therefore, sophisticated management of both the number and length of substitutions appears to be of key importance when attempting to optimise the level of performance across a team. In soccer, for example, the physical outputs of substitute players are significantly higher than those of players who have played the entire match (Mohr *et al.*, 2003). In professional field hockey, however, little scientific knowledge is available on the practical use of substitutions and its influence on the performance of the athletes involved. To our knowledge, only one study has aimed to address the issue of substitution strategies by determining the effects of different substitution frequencies on the physical and technical outputs of strikers (Lythe and Kilding, 2013). The results showed that a higher frequency of substitutions appears to offset decreases in physical outputs owing to fatigue while increasing the technical impact on the game. While the aforementioned investigation provided an initial insight with respect to the position of the strikers, no study to date has investigated the actual substitution conditions of a professional field hockey team during competitive match play.

Therefore, to provide a better understanding on the subject of substitutions in field hockey, this study aimed to (i) describe the substitution tactics of an international field hockey team during competitive match play and (ii) identify the effect of a substitution on the physical output of players when re-entering active game play after bench periods.

2. Methods

2.1.Experimental Approach to the Problem

The objective measures selected for this study were classified into two categories: (i) The average time players spent on the pitch during active game play or off the pitch (during bench periods). This included both the entire match duration and the respective substitution periods. The data were measured along with the total number of substitutions per match for each positional group. (ii) The match distance covered relative to the individual time on the field ($m \cdot min^{-1}$). Only data representing the actual game play were included in the analysis. The data set excluded the positional data pertaining to warm-up, warm-down, breaks and bench periods. To further investigate the question of whether the performance enhancing effect of substitutions existed, and also its duration, the absolute data were corrected to a minute-by-minute value (distance divided by time played) to enable an equal comparison of re-entering players and the average team performance across the entire game.

A substitution (either on or off) was operationally defined to occur when a player crossed the side touchline in exchange for a teammate. Moreover, substitutions were only recognised if the re-entering player had previously actively participated in the game, meaning that the first appearances of players were excluded from the analysis. A total of 174 substitutions were included for statistical analysis based on this definition. For each substitution event, we registered: (i) the tactical position of the player; (ii) the enumerated substitution number per individual player; (iii) the active playing time on the field after the substitution; (iv) the pause duration (i.e. the length of time the player spent on the bench) before the respective substitution; and (v) the distance covered during the first, second, third, fourth and fifth minute after re-entering active play.

2.2. Subjects

The investigation involved thirteen male field hockey players belonging to the German National field hockey team. Their physical parameters (mean ± SD) were as follows: age, 25.4 ± 2.5 years; height, 183.4 ± 5.4 cm; body mass, 83.1 ± 9.7 kg. All participants received a clear explanation of the study. The Faculty for Medicine ethics committee of the Technical University of Munich approved all procedures. Athletes and coaches gave consent for any data collected during the tournament to be used for audit purposes and for that data to be submitted for publication.

2.3. Procedures

Time-motion analyses were conducted during the ERGO Masters 2014 (4 Nations Cup Duesseldorf), which occurred ten days after the end of the German Field Hockey Season within a preparatory training course for the Hockey World Cup 2014 to be played two weeks later in Den Haag. All tournament matches were performed on water-based synthetic pitches under International Hockey Federation rules. The tournament consisted of three matches per team that were played over a period of four days, with a rest day separating matches two and three. Three distinct positional groups (4 strikers, 5 inside forwards and 4 halves) separated the players for positional comparison. The low number of substitutions for goalkeepers and fullbacks excluded them from the analysis. All participants played in their customary position during the investigation.

2.4. Apparatus

A radar-based local position measurement (LPM) system (Inmotio Object Tracking BV, Amsterdam, Netherlands) collected the positional data of the players. Active transponders worn by the athletes in a chest strap sent signals to fixed passive base stations placed around the pitch, as described by Stelzer *et al.* (2004). Data were collected at 1000 Hz and then divided by the number of transponders in use, resulting in an individual player frequency of 62.5 Hz. The LPM data provided information on speed, distance and position. Several authors have reported very good accuracies when using the radar-based LPM technology to measure the movement of football players, including Frencken *et al.* (2010), who investigated the accuracy and validity of the LPM system in football-specific conditions and reported an average positional error of

1–3 cm for transponders worn by the athletes. Siegle *et al.* (2013) reported an average positional error of 24 cm for the LPM system and concluded that the radar-based system is more valid in detecting x, y positions than a common image based system that produces a positional error of 73 cm. As both football and field hockey can be described as competitive, high-intensity, intermittent field-based team games (Gabbett, 2010; Lythe & Kilding, 2011) with equivalent movement patterns on the pitch, it can be presumed that the above mentioned studies of accuracy remain also valid for the field hockey environment.

2.5. Statistical Analysis

Descriptive data are presented as mean ± standard deviations. Data from all three games were averaged. The differences in positions (strikers, inside forwards and halves), covered distances and pause durations were assessed using the analysis of variance (ANOVA) method. Duncan's new multiple range test (MRT) was used to identify differences post-hoc. The statistical significance was set at $p < 0.05$. A statistical power analysis was conducted post hoc. Given the effects size of the present study ($d = 0.29$) and an n of 174, we observed a power level of 0.89.

3. Results

An average of 58.0 ± 4.6 substitutions per match was measured in the three tournament games, during which the average player performed 4.7 ± 0.8 individual substitutions with a pause duration of 5.4 ± 1.2 min and an on-field playing duration of 7.18 ± 2.14 min (Table 1). Players at the half position averaged significantly ($p<0.05$) higher pause durations (6.2 ± 1.3 min) compared to strikers (5.2 ± 1.1 min) and inside forwards (5.2 ± 1.2 min). Among the selected positional groups, the mean total match time for each individual player was 42.5 ± 5.5 min (61 ± 7.9% out of 70 min of the total match duration).

Table 1 Mean time on pitch and bench durations per individual player, presented for positional groups.

	Strikers	Inside Forwards	Halves	Mean
Time on pitch per match (min)	44.67 ± 6.04	43.07 ± 5.42	39.78 ± 4.45	42.52 ± 5.45
Bench duration per match (min)	25.33 ± 3.06	26.93 ± 4.46	30.22 ± 3.54	27.48 ± 4.21
Work:rest ratio per match	1.76	1.60	1.32	1.55
Substitutions per match	4.50 ± .80	5.00 ± .70	4.58 ± .90	4.70 ± .81
Time on pitch per substitution (min)	7.25 ± 1.59	7.24 ± 2.19	7.02 ± 2.23	7.18 ± 2.14
Bench duration before substitution (min)	5.19 ± 1.08	5.21 ± 1.23	6.16 ± 1.26 ^{a}	5.37 ± 1.23

a - Significant difference exists between positional groups ($p < 0.05$).

Within the first minute of substitution, re-entering players covered a significantly ($p<0.05$) larger total distance (159.7 ± 33.0 $\mathrm{m \cdot min^{-1}}$), compared to the second (136.7 ± 34.9 $\mathrm{m \cdot min^{-1}}$), third (140.1 ± 35.3 $\mathrm{m \cdot min^{-1}}$), fourth (141.1 ± 35.2 $\mathrm{m \cdot min^{-1}}$) and fifth (130.5 ± 37.8 $\mathrm{m \cdot min^{-1}}$) minutes (Figure 1). We also detected a significant ($p < 0.05$) decrease in the total distance covered between the fourth and fifth minute after substitution. The mean total distance covered by each individual player during an on-field period was 139.4 ± 33.3 m · min^{-1}.

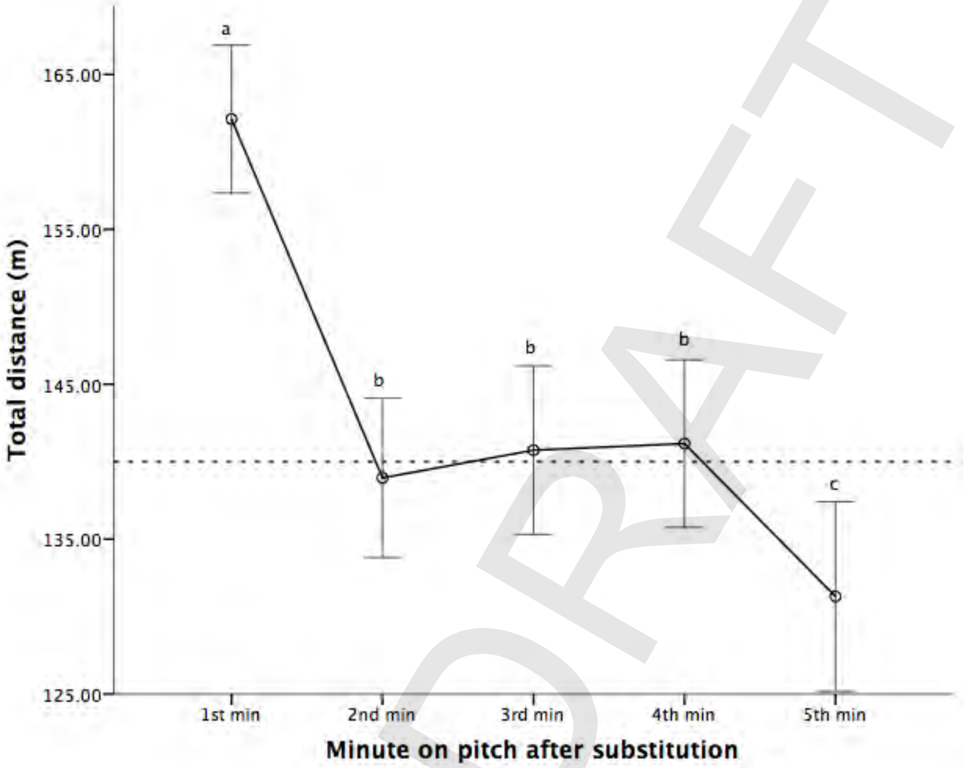


Figure 1. Total distance (m · min^{-1}) covered during the 1st, 2nd, 3rd, 4th and 5th minute after substitution. The dashed line shows the team average of an on-field period across the entire match. Duncan's MRT was used to identify homogeneous subsets (^{a,b,c} groups with the same letter designation are not significantly different ($p < 0.05$)).

The individual pause durations before substitution increased with each consecutive substitution number, beginning with an average pause duration of 5.0 ± 1.6 min before the first substitution and rising to 6.3 ± 1.7 min before the sixth substitution (Figure 2). A decreasing trend was observed between the third (5.9 ± 1.5 min) and fourth substitutions (5.8 ± 1.1 min). In this respect, the pause durations increased by 21.6% over the course of a match.

The mean total distance covered during substitution periods decreased significantly ($p<0.05$) between the first (157.0 ± 30.6 m · min^{-1}) and second (143.0 ± 32.8 m · min^{-1}) substitutions. Though not significant, a decreasing trend of total distance was also observed in the subsequent periods.

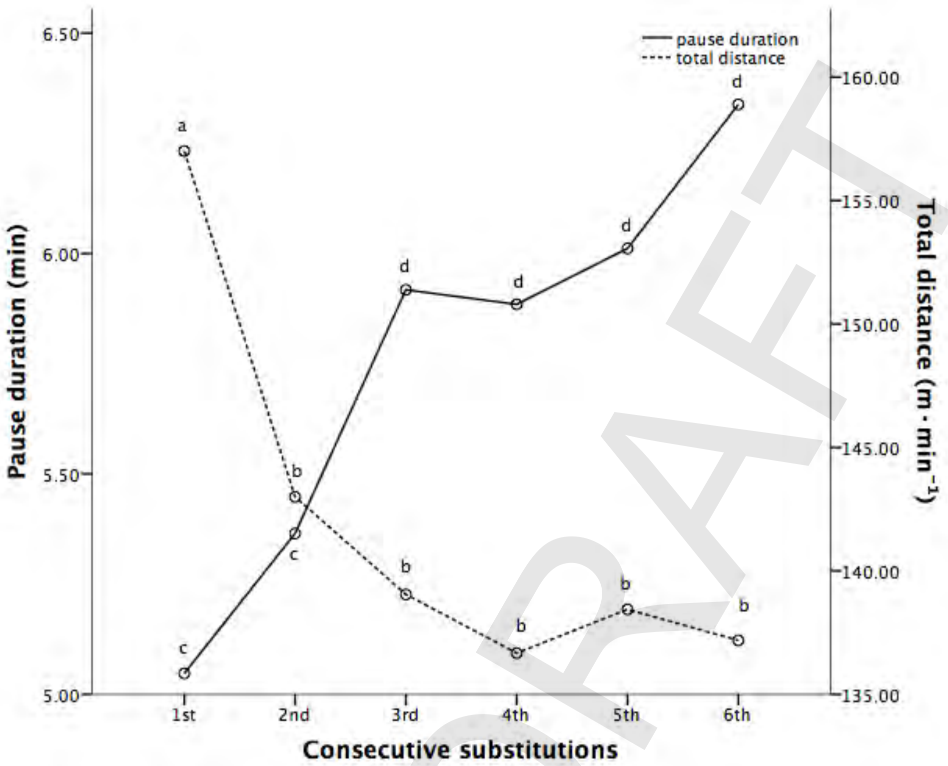


Figure 2. Mean total distance (m · min^{-1}) covered in the respective substitution period per match in comparison with the average pause duration before substitution–depending on the consecutive substitution count. Duncan's MRT was used to identify homogeneous subsets (^{a,b,c,d} groups with the same letter designation are not significantly different ($p < 0.05$)).

4. Discussion

Despite the possible use of unlimited substitutions, an approximate 6% decrease in the physical output of players has been reported when comparing the first and second half of a field hockey game (Lythe and Kilding, 2011; Jennings *et al.* 2012a). Furthermore, Jennings *et al.* (2012b) reported a noticeable within game decrease in the physical output of elite field hockey athletes, where fatigue is indicated by a reduction in high speed running performance. Hence, there is a need to identify sophisticated substitution strategies in order to optimise the team's exercise intensity throughout the course of a match.

Data presented in this study showed that halves averaged significantly higher pause durations prior to time on pitch periods than strikers and inside forwards (Table 1). Though not significant, other results suggested that halves spent less total time on the pitch, more total time on the substitute's bench and averaged the lowest playing duration per individual substitution period, whereas the differences between strikers and inside forwards were only marginal.

These findings were in contrast with recent studies examining the exercise intensities during elite men's field hockey competition, in which halves or midfielders averaged the longest time on pitch compared to the other positional groups (Lythe and Kilding, 2011; Jennings *et al.* 2012a) (with the exception of fullbacks, who were excluded from this analysis due to a low number of substitutions).

The mean total time on pitch for each individual player (42.5 min) was lower than previously reported by Lythe and Kilding (2011) (51.9 min) , White and MacFarlane (2013) (48 min) and Buglione *et al.* (2013) (58.2 min), primarily because the present study excluded fullbacks, who play the highest amount of minutes per match compared to the other positional groups, from the analysis (Lythe and Kilding, 2011; Jennings *et al.* 2012a). Regarding the position of strikers, Jennings *et al.* (2012a) reported an average match playing time of 44.34 min, which is comparable to the findings in the present study (44.67 ± 6.0 min), whilst Lythe and Kilding (2011) reported an average playing time of 50.4 ± 15.4 min. A meaningful comparison between the studies is limited, however, since the aforementioned studies relating hockey specific substitution parameters applied varying classifications for positional groups (strikers, halves, inside forwards and fullbacks (Lythe and Kilding, 2011), or defenders, midfielders and strikers (Jennings *et al.* 2012a), respectively. In addition, each study analysed only a single team, which suggests that the observed differences could also be due to specific tactics for individual teams.

To the authors' knowledge, only one study has investigated the number of substitutions in elite hockey matches (18 ± 0.4, 31 ± 0.8 and 17 ± 1.5 for strikers, midfielders, and defenders, respectively (Jennings *et al.* 2012a)). However, the study only investigated the total number of rotations in each positional group, and not the number of rotations per individual athlete, as was the case in the present study. When presented as the total number per positional group, the present study observed 18 ± 1.0, 22 ± 3.6 and 18 ± 1.0 rotations per match for strikers, inside forwards and halves, respectively. We found no significant position related differences in the average number of rotations per individual athlete (Table 1).

The players' physical performance, as described by the total distance covered per minute of play, declined by ~14% between the first and second minute after substitution (Figure 1). Within the second, third and fourth minute after substitution, re-entering players were able to maintain a level of performance comparable to the team average across the entire game. These results indicated the presence of a 'first-minute-rush-effect', in which re-entering players exceeded the average team performance by ~13%. Similarly, Mohr *et al.* (2003) reported that the amount of high-intensity running in the last 15 min of a soccer game is 25% higher for substitutes in comparison with players who have played the entire match. In field hockey, however, special emphasis should be placed on the present findings that re-entering players covered a significantly larger total distance compared to the team average, yet only within the first minute after substitution. The fact that elite field hockey substitutes approach average team values after a relatively short time could be due to the coaches' use of a high number of interchanges, thereby enabling the overall team exercise intensity to be maintained at a consistently high level. In addition, the rules in field hockey state that time shall not be stopped for substitutions, which means that re-entering players have to occupy their customary position during active game play. In order to maintain the tactical formation of the team, re-entering players therefore will try to find their assigned position on the pitch as fast as possible, which could explain why the running performances of substitutes were exceptionally high within the first minute after the substitution. In this regard, there should be a distinction between tactical substitutions and substitutions due to fatigue, since not every rolling interchange is performed with the intention to enable a recovery break, but rather to adapt to spontaneous changes in the course of play. Nonetheless, a significant decrease in the players' performance was apparent between the fourth and fifth minute after entering the field as a substitute, which indicated that elite field hockey players experienced distinguishable signs of fatigue approximately four minutes after entering the field as a substitute. At present, we do not know if this trend continues in the subsequent on-field period following the focused five-minute window. However, a comparison of the mean value of an entire on-field period (~139 $\text{m} \cdot \text{min}^{-1}$) and the mean value during the first five minutes (~142 $\text{m} \cdot \text{min}^{-1}$) appeared to imply that the performance did not decrease further in the period between five minutes after re-entering and leaving the pitch in exchange for a teammate.

In soccer, a decrease in running performance across the course of a match is thought to identify the physiological impairment of a player, indicative of fatigue (Mohr *et al.* 2003). Similar findings have been made in rugby league (Waldron *et al.* 2013), rugby union (Lacome *et al.* 2014) and American Football matches (Wisbey *et al.* 2010). It should however be questionable to transfer these findings to a single on-field period of a hockey player, in particular taking into account the shorter on-field periods in field hockey. Instead, whilst some form of fatigue is likely to occur during a single on-field period, it might be more appropriate to relate this phenomenon to a complex pacing strategy on behalf of the player (Waldron and Highton, 2014). As proposed by Edwards and Noakes (2009), brief periods of peak intensity are constantly followed by a significant brief period of reduced intensity in an attempt to distribute the energy resources that optimize match-running performance.

The present study revealed a linear relationship between the average pause duration prior to re-entering and the consecutive substitution count (Figure 2), indicating that coaches are likely well aware of the occurrence of physical fatigue throughout the course of a match and thus attempt to minimise this trend by extending recovery times as the game progresses. However, despite the aforementioned substitution strategy, our results showed a decreasing trend of total distance covered within each subsequent substitution period. As reported in earlier studies (Lythe and Kilding, 2011,2013; Jennings *et al.* 2012a), our findings confirmed a decrease in the physical output of field hockey athletes during games.

This case study included a relatively small number of players who all played for the same team, and thus the information obtained in this study could have been a reflection of this particular team only. In addition, a larger sample size could potentially allow for a more detailed analysis. For example, variations in the players' physical performance that is dependent on their individual work-to-rest-ratio could be a matter of great interest for coaches as they seek to obtain a substitution strategy that is matched to both the physical demands of the tactical role and the physical capacity of the individual athlete. Furthermore, future studies could consider the number of available players per tactical position on the pitch. A higher number of available substitutes allows for a higher substitution frequency, and vice versa. Therefore, variations in the tactical starting line-up of both the team and opponent should be explored further, as the tactical composition of substitutes and team strategies has an impact on the substitution parameters in terms of frequency and physical output.

The results of this study provide new information for coaches and scientists of elite field hockey. Specifically, this study was the first to analyse the number and length of substitutions during field hockey competition. Our data demonstrated the presence of a 'first-minute-rush-effect' of substitutes, in which re-entering players covered a significantly larger total distance in comparison to the team average. This confirmed our hypothesis that a recovering effect during bench periods results in a temporarily enhanced performance when players re-enter active game play. However, a significant below-average performance of substitutes was apparent approximately four minutes after entering the field of play. Despite the lengthening of bench durations towards the end of the game, the performance of substitutes decreased constantly within each consecutive re-entry to the field as a substitute period. Further, the results showed that halves were the most unique of the positional groups as they spent less time on the pitch and more time on the bench than strikers and inside forwards. Collectively, it appeared that field hockey players experienced temporary fatigue during substitution periods and throughout the course of a game. Despite the possible use of unlimited substitutions, we recognised significant temporary differences between the performance of substitutes and the average team performance. Further research is required to identify peak performances of individual players that are dependent on their physical abilities and varying work-to-rest-ratios.

5. Conclusions

Despite the limitation that the results presented here are transferable to other teams only to a limited extent, this case study provides preliminary exploration and understanding of substitutions in elite men's field hockey. It is thus intended to generate novel theories and deliver an initial approach for analysing field hockey performance in the context of the underlying substitution strategies. In an attempt to maximise the overall team performance, coaches and sport scientists may be able to use these indications to optimise existing substitution strategies. In this context, it is worth mentioning that novel theory developed from case study research is likely to have important strengths like novelty, testability, and empirical validity, which arise from the intimate linkage with empirical evidence (Eisenhardt, 1989). Specifically, this case study demonstrated that field hockey players experience distinguishable performance fluctuations both during a single on-field period and over the course of a game. As a conclusion, we suggest that physical performance data from field hockey should be interpreted in the context of the underlying substitution strategies, as differing substitution tactics considerably influence the physical output of the athletes.

6. References

1. Buglione A., Ruscello B., Milia R., Migliaccio G.M., Granatelli G., D'Ottavio S. (2013). Physical and Physiological demands of elite and sub-elite Field Hockey players. **International Journal of Performance Analysis in Sport**, 13(3), 872–884.
2. Edwards A.M., Noakes T.D. (2009). Dehydration: cause of fatigue or sign of pacing in elite soccer? **International Journal of Sports Medicine**, 39(1),1-13.
3. Eisenhardt, K.M. (1989). Building Theories from Case-Study Research. **Academy of Management Review**, 14(4), 532-550.
4. Frencken W.G., Lemmink K.A., Delleman N.J. (2010). Soccer-specific accuracy and validity of the local position measurement (LPM) system. **Journal of Science and Medicine in Sport**, 13(6), 641–645.
5. Gabbett T.J. (2010). GPS analysis of elite women's field hockey training and competition. **Journal of Strength and Conditioning Research**, 24(5),1321-1324.

6. Jennings D.H., Cormack S.J., Coutts A.J., Aughey R.J. (2012a). GPS analysis of an international field hockey tournament. **International Journal of Sports Physiology and Performance**, 7(3), 224–231.
7. Jennings D.H., Cormack S.J., Coutts A.J., Aughey R.J. (2012b). International field hockey players perform more high-speed running than national-level counterparts. **Journal of Strength and Conditioning Research**, 26(4), 947–952.
8. Lacome M., Piscione J., Hager J.P., Bourdin M. (2014). A new approach to quantifying physical demand in rugby union. **Journal of Sports Sciences**, 32(3), 290-300.
9. Lythe J., Kilding A.E. (2011). Physical demands and physiological responses during elite field hockey. **International Journal of Sports Medicine**, 32(7), 523–528.
10. Lythe J., Kilding A.E. (2013). The effect of substitution frequency on the physical and technical outputs of strikers during field hockey match play. **International Journal of Performance Analysis in Sport**, 13(3), 848–859.
11. Mohr M., Krustrup P., Bangsbo J.(2003). Match performance of high-standard soccer players with special reference to development of fatigue. **Journal of Sports Sciences**, 21(7), 519–528.
12. Siegle M., Stevens T., Lames M. (2013). Design of an accuracy study for position detection in football. **Journal of Sports Sciences**, 31(2), 166–172.
13. Stelzer A., Pourvoyeur K., Fischer A. (2004). Concept and application of LPM - a novel 3-D local position measurement system. **IEEE Transactions on Microwave Theory and Techniques**, 52(12), 2664–2669.
14. Waldron M., Highton J. (2014). Fatigue and pacing in high-intensity intermittent team sport: an update. **International Journal of Sports Medicine**, 44(12), 1645-1658.
15. Waldron M., Highton J., Daniels M., Twist C. (2013). Preliminary evidence of transient fatigue and pacing during interchanges in rugby league. **International Journal of Sports Physiology and Performance**, 8(2), 157–164.
16. White A.D., MacFarlane N. (2013). Time-on-pitch or full-game GPS analysis procedures for elite Field Hockey? **International Journal of Sports Physiology and Performance**, 8(5), 549–555.
17. Wisbey B., Montgomery P.G., Pyne D.B., Rattray B. (2010). Quantifying movement demands of AFL football using GPS tracking. **Journal of Science and Medicine in Sport**, 13(5), 531-536.

6 Discussion

This chapter can be seen as an addition to the discussions in the manuscripts of the conducted studies. Caused by the emphasis on the methodology development of EPTS validations, the insights derived from the findings predominately focus on the resulting implications of the applied validation method.

6.1 Summary

Four studies have been conducted to develop and apply a method to assess the measurement accuracy of tracking technologies in team sports, and to further adopt innovative analysis methods to answer questions from sports practice concerning issues with positional data.

Study One assessed the measurement accuracy of the most commonly used tracking technologies in professional team sports by comparing the position, speed, acceleration and distance measures of each technology against simultaneously recorded measures of a criterion reference. One of the key insights of this study was the noticeably large error margin in the accuracy of key performance indicators KPIs that is independent of the respective system or technology, which we are still facing with EPTS in general.

Study Two showed that the anatomical reference point used to represent the body in space needs to be carefully considered in the interpretation of tracking variables delivered by different EPTS.

Study Three demonstrated for the first time that the decline in players' match running performance during football matches is substantially amplified by a proven increase in game interruptions, indicating that there may be a tendency among practitioners to overestimate fatigue-induced performance declines.

Study Four described and analysed the substitution tactics of an international field hockey team and found a 'first-minute-rush-effect', wherein substitutes covered a significantly larger total distance compared to the team average.

6.2 EPTS Validation Methodology

The central issue of this thesis was the development and application of a comprehensive validation method for EPTS. The following section, therefore, summarizes the main methodological aspects that were obtained during the course of this dissertation.

In general, there are two basic designs for EPTS validation studies: Studies that compare the results between different EPTS and others that have a 'ground truth' involved, i.e. a gold standard or criterion reference. The former only provides the scope for recording the differences between various systems. Buchheit et al. [16] conducted a study comparing the three relevant technologies that have been elaborated in the previous sections. The results of this study show that there are considerable discrepancies between different systems, and performance indicators are not comparable between systems at a practically acceptable level. Consequently, the authors included transformation equations but had to do so for different pitch sizes and KPIs. Additionally, the sometimes very poor correlations between even the for pitch and KPI adjusted results lead to the conclusion that, at present, the research focus should lie on improving the systems themselves instead of looking for the most appropriate way to make their results more comparable. Basic improvements could be stimulated by providing feedback involving a gold standard and by conducting the comparisons not only at the level of performance indicators but also on the basic XY-data, instantaneous speed, and acceleration.

Another challenge that needs to be solved is finding an appropriate testing venue. The task of supplying sufficient conditions for all tested systems as well as for the gold standard can only be a compromise, because there are partially conflicting demands of different EPTS. For instance, GPS systems need open spaces to perform best, whereas video-based systems need an ideal elevation angle that is, in principle, only achievable with constructions around the pitch. Sometimes, the only solution is to erect additional scaffolding on top of the stadium stands [45].

The test exercises that were used during the validation contained a representative spectrum of relevant exercises. Controlled exercises were administered sequentially in a sports-specific run, which is an economic way of testing. Shuttle run and 5vs5 are common training exercises; the latter involves game-like complex movements.

The employed gold standards (marker-based infrared 3D position detection) allowed to study the basic exercises such as linear runs with accelerations and decelerations as well as specific exercises with a controlled trajectory and game-like complex movements. The precision and coverage of the infrared gold standard were evaluated using a calibration rod with markers at fixed and known distances. The results with a 95 % Confidence Interval (CI) for the RMSE [-1.9 mm, +2.0 mm] achieved throughout the measurement volume demonstrate that the well-established precision of infrared position detection for (a smaller number of) fixed cameras in laboratory settings was successfully extrapolated to a field setting. To the best of the authors' knowledge, a measurement volume of 30x30 meters was not achieved before in the field of EPTS' validation studies. It allows not only to conduct complex basic exercises such as a sport-specific course and training exercises such as shuttle runs but a 5vs5 exercise also with typical game-specific, but also free and complex movements. Nevertheless, the maximum speed that can be achieved in full-size matches (approximately 7.9 - 8.6 $m \cdot s^{-1}$ m/s, [14, 24]) is somewhat higher than the speeds attained in this thesis (6.6 $m \cdot s^{-1}$ in 5vs5; 7.9 $m \cdot s^{-1}$ in the SSC), potentially leading to higher errors values in full-sized matches.

Further, while judging the applied gold standard, one has to be aware that it takes a lot of time (about 4 hours) and personal expertise (a specialist from the manufacturer was signed up) for mounting and calibrating the system. Moreover, the weather conditions must be favorable and the measurements can be obtained only after sunset. Thus, several conditions had to be met in order to conduct the validation study.

The data processing in the study involved several standard tasks pertaining to comparing the time series of measurements. Re-sampling of the different systems - each working with a different sampling rate - was achieved by spline interpolation. Cross-correlation provides time synchronization as it is a method that is favorable for the tested systems, since among the list of any possible delays, the one that maximizes the agreement is chosen. Spatial alignment requires translation and rotation of the XY-data into a common coordinate system. Moreover, some problems that are specific to EPTS validation studies had to be solved. Gait neutralization becomes necessary when a gold standard is too precise in the sense that it captures intra-cycle speed of walking and running that closely resembles a sine curve. Since the tested devices are not able and not meant to do so, gait neutralization was introduced by filtering the gold standard speed data with a cut-off frequency of 2 Hz. Since there are no similar experiences from previous studies, this methodological step needs further verification and refinement.

Various reasons lead to the fact that a player, who is certainly standing still, accumulates distance even in these moments. Since the reasons include measurement errors as well as natural body sway or a change of posture, it is appropriate for the purpose of this thesis to procure information only about the gross displacement of a player, which is achieved by stance neutralization. The waypoint method was adapted to the problem under scrutiny, which introduced a threshold of about one COM-step length (60 cm) for a 'real' displacement. Only then can it be ensured that the criterion distance is representative of the athlete's gross motion.

6.3 EPTS Measurement Accuracy *(Study 1)*

Collectively, results of *Study One* revealed that largest differences between EPTS occurred at the spatial accuracy, whereas speed and acceleration errors of GPS were comparable to those of LPS. Yet one important insight in this regard is the noticeably large error margin in the third data category (accuracy of KPIs) that is largely independent of the respective system or technology, which we are still facing in EPTS in general. Especially in KPI categories that might have a high impact on practical decisions, such as high-intensity performance indicators, we found significant deviations from the gold standard.

In line with previous research, *Study One* also confirmed that the radar-based technology showed higher validity for measuring an athlete's position (0.23 m RMSE) than both computer vision (0.56 m RMSE) and GPS (0.96 m RMSE), thereby seemingly confirming the general perception of LPS being the technology of choice when it comes to position accuracy.

In a recent unpublished study, however, we applied the same validation setup to assess the measurement accuracy of ChyronHego's tracking technologies TRACAB Gen 4 and ZXY Arena. Compared with *Study One*, ZXY accuracy results (RMSE: 0.27 m) were in good accordance with the accuracy of Inmotio's LPM system (RMSE: 0.25 m). Conversely, the TRACAB system detected a player's position with 0.17 m accuracy RMSE, and thus for the first time providing evidence that computer-vision is seemingly capable of measuring a player's position with higher accuracy than LPS technologies. A screen-shot of the TRACAB validation is presented in Figure 6.1 (single image of supplementary material *Video 2*).

It should, however, be noted that TRACAB's advantage in measurement accuracy is less pronounced, sometimes even reversed, in the instantaneous speed and acceleration categories (see Fig. 6.2). TRACAB's instantaneous speed and acceleration errors were particularly large during movements that involved sharp turns and short lunges. Since both speed and acceleration measures are directly derived from the position data, these results suggest that data processing adjustments could potentially increase TRACAB's speed and acceleration accuracy.

These recent results illustrate that, while accurate position data undoubtedly provides the basis for any convincing performance diagnostics, the actual value for any coach, expert or sports scientist lies in the interpretation of the processed data. In other words, the most accurate position data could prove inappropriate if the derived tracking variables do not reliably reflect a player's performance. Since

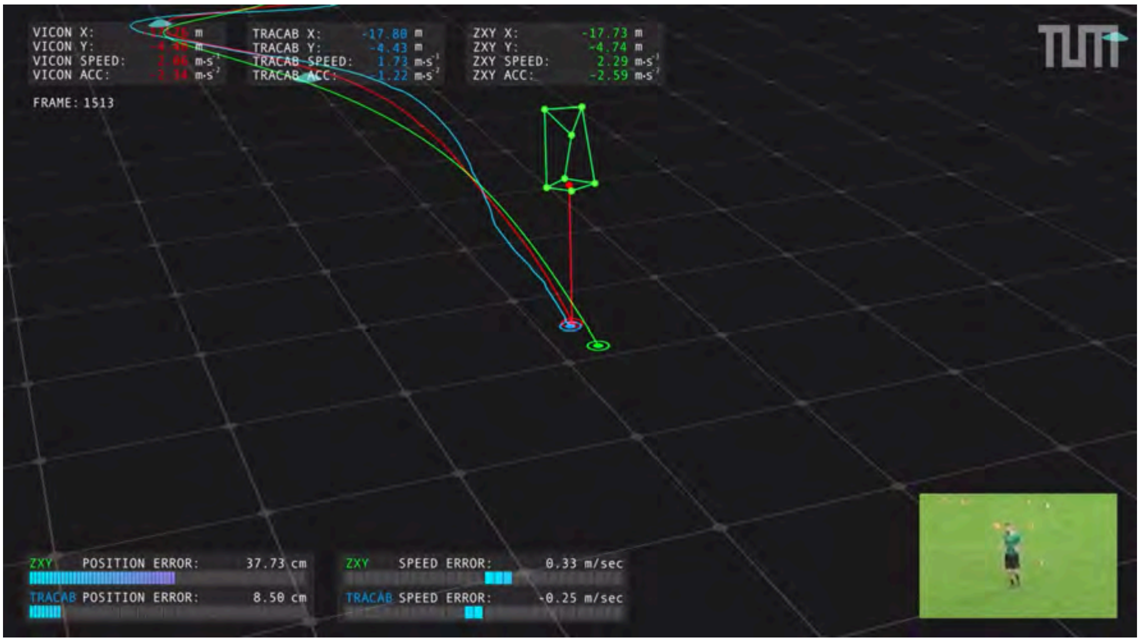


Figure 6.1: Exemplary visualization of the TRACAB validation. The red dot represents the center of pelvis projected onto the ground plane. The volored dots represent ChyronHego's tracking technologies TRACAB Gen 4 (blue), and ZXY Arena (green).

all relevant tracking variables (instant speed, acceleration, and KPIs) greatly depend on the underlying data characteristics and the subsequently applied processing procedures, the choice of the appropriate filtering methods and KPI algorithms should be carefully considered and documented.

6.4 Comparability of Tracking Variables Delivered by Different Anatomical Reference Positions *(Study 2)*

In order to meet the overreaching goal of comparability and interchangeability of tracking variables delivered by different EPTS, *Study Two* analyzed the impact of different anatomical reference points on the resulting kinematic variables

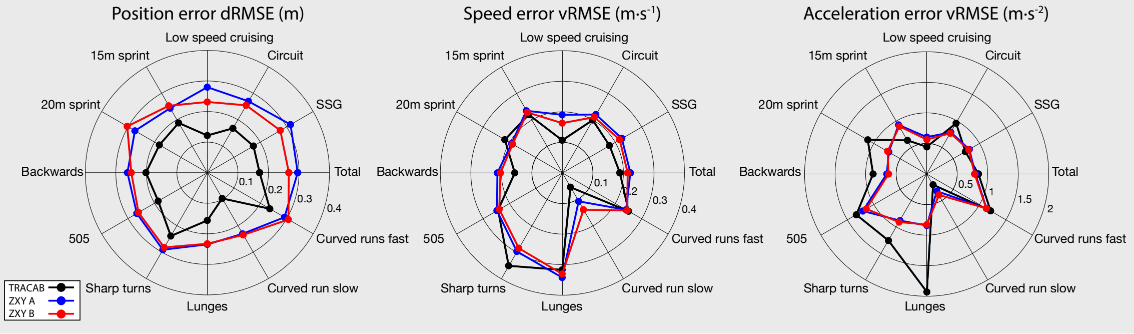


Figure 6.2: Accuracy results of Chyron Hego's tracking systems TRACAB Gen4 and ZXY Arena. The results are categorized into position, speed ans acceleration accuracy. Each category if further sub-categorized into different exercises.

and KPIs obtained therefrom (for a visual animation of the test procedure, see supplementary material Video 2). Results showed that COS sprinting distance was on average 44.65% lower in comparison to COP. Similarly, maximum speed obtained from COS was 2.94% lower in comparison to COP. On the contrary, maximum acceleration values of COS were on average 16.15% higher compared to COP. Contrastingly, top decelerations derived by COS were on average 5% lower than those derived by COM.

In addition to the published manuscript, visual inspection of the VICON data indicates that highest accelerations of COS occur during the early acceleration phase of a sprint from a standing start (see *Video 4* & Figure 6.3), caused by high-frequency movements in the upper torso. Highest decelerations, on the other hand, occurred during the deceleration phase of a sprint effort. Since the upper body plays an important role in damping gait-related oscillations during locomotion to ensure the head is stabilized, higher ground contact-induced decelerations in the pelvis region are well comprehensible.

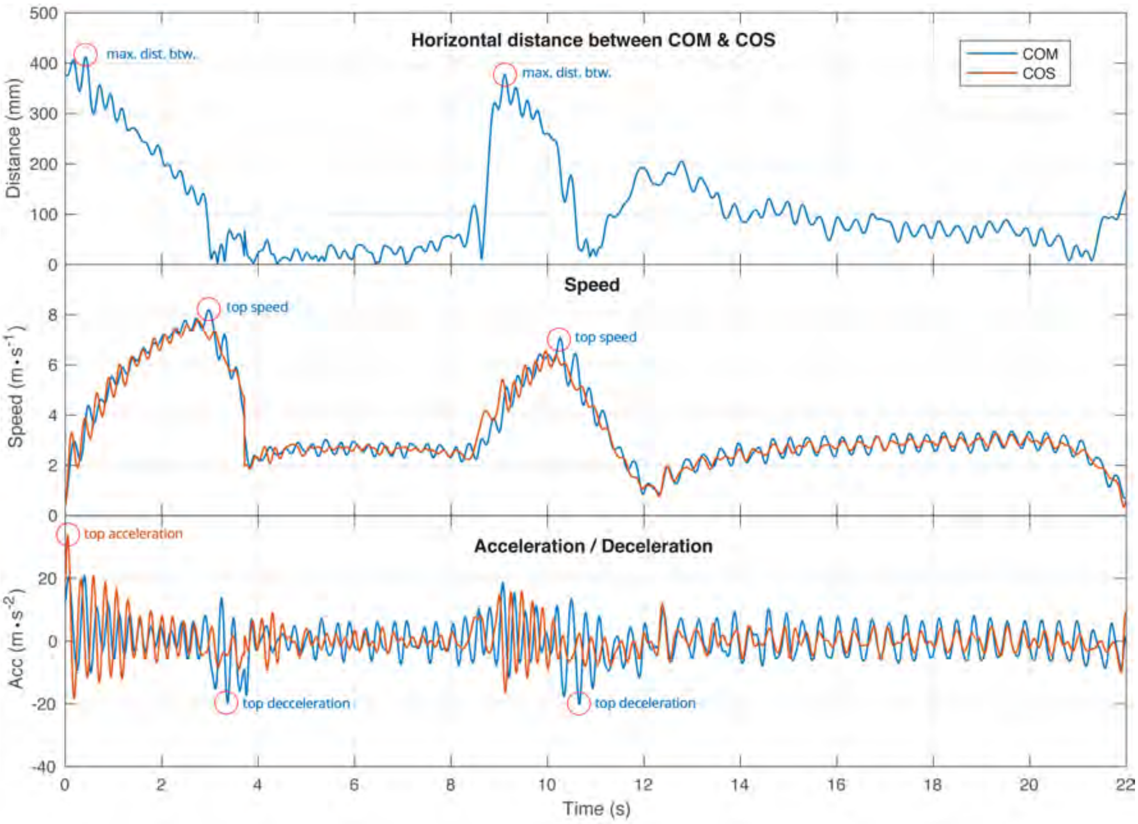


Figure 6.3: Exemplary visualization of the deviations between COM and COS. The figure shows a specific section of the circuit (sprint from a standing start, low-speed cruising, sprint from low-speed cruising, low-speed cruising).

6.5 Performance Declines in Team Sports (Study 3&4)

Whereas *Study One and Two* could be assigned to a criterion-oriented validation procedure (quality and comparability of kinematic tracking variables delivered by different EPTS),*Study Three and Four* focused on one of the most discussed theoretical frameworks derived from time–motion analysis data - the construct validity of physical fatigue.

One particular concern with recent research that reported the existence of transient or end-game fatigue was the failure to account for gameplay interruptions. Accordingly, *Study Three* intended to close this gap by applying an innovative approach to quantify the contribution of game interruptions to the fatigue-related declines in match running performance throughout a football match. Results showed that more than half (57.9%) of the commonly reported decline in match running performance could not be assigned to physical fatigue, but rather to an increase in game interruptions as the game progresses. These results demonstrate that time–motion analysis data should not be used in isolation to identify the occurrence of player fatigue. Instead, the context of the game should be considered to allow for a more pragmatic interpretation of match running activities.

The much discussed ’practical impact debate’ argues that sports performance research has been overly concerned with physiological predictors of performance (theory-driven research), at the expense of not providing a valid and reliable description of the exact nature of the task in question (applied research). In this context, Mackenzie and Cushion [46] argued that variables have been measured as a result of availability rather than to develop a deeper understanding of performance in applied settings. However, the value of research to a community is the extent to which its findings are (i) used as recommended practices in the preparation of practitioners, and (ii) incorporated by practitioners in everyday practice [46]. Therefore, in order to enable a more meaningful impact on professional practice, *Study Four* tried to connect the approaches of both theoretical and practical performance analysis. Within the scope of a funded research project, we supported the German national field hockey team with applied performance analysis based on position data. One practical question arising during this cooperation was the analysis of the team’s substitution strategy. The findings suggest a ‘first-minute-rush-effect’, wherein substitutes covered a significantly larger total distance compared with the team average. Further, re-entering players experienced apparent signs of fatigue approximately four minutes after substitution. Collectively, these results could potentially serve as a starting point for developing individual recovery intervals which could be an advantage in international competition.

6.6 Limitations

It is worth noting that the results of *Study One* are based on untreated raw data, as provided by the manufacturer's proprietary software. Therefore, it is to be expected that the validity of the tested EPTS could be further improved by additional data filtering procedures to the EPTS data. Further, it is regrettable that at the time being, there is no gold standard for a full-size pitch of team-based sports. Until this problem will be solved we have the regrettable fact that EPTS might not be validated in the scenarios that are most relevant to them.

It should also be mentioned that the raw data characteristics and data processing steps used in *Study Two* may not necessarily be comparable to data processing steps commonly applied in the field to EPTS. Since there are no generally accepted definitions of required signal properties, tracking data obtained from different EPTS may vary, sometimes considerably, and thus transferability of results inevitably depends on the particular data processing procedures of the respective EPTS.

The limitations of *Study Three* include a lack of control for position-specific subdivisions of players, seasonal variation, match importance, and international differences. The observed patterns may, therefore, be a reflection of this specific league (Deutsche Bundesliga).

Study Four included a relatively small number of players who all played for the same team, and thus the information obtained in this study could have been a reflection of this particular team only. In addition, a larger sample size could potentially allow for a more detailed analysis for variations in the players' physical performance that is dependent on their individual work-to-rest-ratio.

7 Outlook

Looking at current developments of athlete tracking technologies, it becomes evident that the EPTS landscape is subject to constant updates and innovations. Driven by a continuous goal to extend their market position, tracking providers are forced to enhance their products on a regular basis. Apart from purely technological improvements of the tracking hardware (e.g., increase in sampling frequency, number of available satellites, increased battery life, minimization of the product size, and improved memory storage), Artificial Intelligence, Machine Learning, and Deep Learning are current topics of considerable interest in EPTS software developments. These recent developments in information technology have led to a realistic perspective for an automated analysis of match behavior based on positional information of players and ball. Such methods are working on a very large data base, ’learn’ the underlying rules for attributing spatiotemporal configurations to certain events, and finally could be able with an acceptable error margin to automatically identify individual players and game events without the need of a human operator. However, as has been pointed out by other authors [58], many of the novel ’big data’ methods are performed by computer science researchers. Accordingly, a closer collaboration between computer and sports science may hold the key to apply such complex but also promising approaches in a more relevant manner.

Similarly, a relatively new approach in the field of sport-specific GNSS positioning is the concept of 'sensor fusion', which refers to the use of an inertial navigation system (INS) to correct or calibrate GNSS-derived position data. INS is a navigation approach that uses triaxial accelerometers, magnetometers, and gyroscopes to continuously calculate the position, orientation, and velocity of a moving object without the need for external references ppgroves2013principles. An integrated INS/GNSS system combines the advantages of both techniques by reducing INS errors and continuously providing reliable navigation data [11]. Fusing GNSS and IMU could potentially combine the long-term stability of the GNSS for slow movements with the short-term stability of the IMU for fast movements leading to a significant gain in accuracy as well as robustness.

A disadvantage resulting from the ongoing hardware and software developments is that EPTS validation results are only valid for a limited time. Notable examples of this include the generalization of measurement accuracy of the 'core' technologies (VBT vs. LPS vs. GNSS). Until recently, the view prevailed that LPS technologies have superior measurement accuracy than both VBT and GNSS. This dissertation showed, however, that this assumption is no longer valid. Recent enhancements in camera gear (e.g., sensor size) and the use of several camera angles lead to a significant increase in tracking accuracy.

As highlighted throughout this dissertation, the major challenge for future research will nevertheless remain in establishing reliable interchangeability and comparability of tracking variables delivered by different EPTS. Especially in KPI categories that have a high impact on practical decisions, such as high-speed performance indicators, every tested EPTS showed significant deviations from the gold standard - regardless of the quality of the underlying raw data. Since there is no clear consensus on the appropriate anatomical reference point, reporting standards, data filtering procedures, and KPI definitions, future directions of EPTS-related research should be aimed at diminishing these inherent inconsistencies.

References

(1) Akenhead, R., & Nassis, G. P. (2016). Training load and player monitoring in high-level football: current practice and perceptions. *International Journal of Sports Physiology and Performance, 11*(5), 587-593.

(2) Altman, D. G., & Bland, J. M. (1983). Measurement in medicine: the analysis of method comparison studies. *The Statistician*, 307-317.

(3) Andrzejewski, M., Chmura, J., Pluta, B., & Konarski, J. M. (2015). Sprinting activities and distance covered by top level Europa League Soccer Players. *International Journal of Sports Science & Coaching, 10*(1), 39-50.

(4) Atkinson, G. (2002). Sport performance: variable or construct?. *Journal of Sports Sciences, 20(*4), 291-292.

(5) Atkinson, G., & Nevill, A. M. (2001). Selected issues in the design and analysis of sport performance research. *Journal of Sports Sciences, 19*(10), 811-827.

(6) Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine, 26*(4), 217-238.

(7) Aughey, R. J. (2011). Applications of GPS technologies to field sports. *International Journal of Sports Physiology and Performance, 6*(3), 295-310.

(8) Bartlett, R., Button, C., Robins, M., Dutt-Mazumder, A., & Kennedy, G. (2012). Analysing team coordination patterns from player movement trajectories in soccer: methodological considerations. *International Journal of Performance Analysis in Sport, 12*(2), 398-424.

(9) Baumgartner, T. A. (1989). Norm-referenced measurement: reliability. *Measurement Concepts in Physical Education and Exercise Science, 20,* 45-7.

(10) Beetz, M., Kirchlechner, B., & Lames, M. (2005). Computerized real-time analysis of football games. *IEEE Pervasive Computing, 4*(3), 33-39.

(11) Bhatt, D., Aggarwal, P., Devabhaktuni, V., & Bhattacharya, P. (2014). A novel hybrid fusion algorithm to bridge the period of GPS outages using low-cost INS. *Expert Systems with Applications, 41*(5), 2166-2173.

(12) Bland, J. M., & Altman, D. G. (1990). A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. *Computers in Biology and Medicine, 20*(5), 337-340.

(13) Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., ... & de Vet, H. C. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clinical Chemistry and Laboratory Medicine, 41*(1), 68-73.

(14) Bradley, P. S., Di Mascio, M., Peart, D., Olsen, P., & Sheldon, B. (2010). High-intensity activity profiles of elite soccer players at different performance levels. *The Journal of Strength & Conditioning Research, 24*(9), 2343-2351.

(15) Bradley, P. S., Sheldon, W., Wooster, B., Olsen, P., Boanas, P., & Krustrup, P. (2009). High-intensity running in English FA Premier League soccer matches. *Journal of Sports Sciences, 27(*2), 159-168.

(16) Buchheit, M., Allen, A., Poon, T. K., Modonutti, M., Gregson, W., & Di Salvo, V. (2014). Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and GPS technologies. *Journal of Sports Sciences, 32*(20), 1844-1857.

(17) Carling, C. (2013). Interpreting physical performance in professional soccer match-play: should we be more pragmatic in our approach?. *Sports Medicine, 43*(8), 655-663.

(18) Castellano, J., Alvarez-Pastor, D., & Bradley, P. S. (2014). Evaluation of research using computerised tracking systems (Amisco® and Prozone®) to analyse physical performance in elite soccer: A systematic review. *Sports Medicine, 44*(5), 701-712.

(19) Ciuti, G., Ricotti, L., Menciassi, A., & Dario, P. (2015). MEMS sensor technologies for human centred applications in healthcare, physical activities, safety and environmental sensing: a review on research activities in Italy. *Sensors, 15*(3), 6441-6468.

(20) Colman, A. M. (2015). *A dictionary of psychology.* Oxford University Press, USA.

(21) Cummins, C., Orr, R., O'Connor, H., & West, C. (2013). Global positioning systems (GPS) and microtechnology sensors in team sports: a systematic review. *Sports Medicine, 43*(10), 1025-1042.

(22) Currell, K., & Jeukendrup, A. E. (2008). Validity, reliability and sensitivity of measures of sporting performance. *Sports Medicine, 38*(4), 297-316.

(23) de Vet, H. C., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology, 59(*10), 1033-1039.

(24) Di Mascio, M., & Bradley, P. S. (2013). Evaluation of the most intense high-intensity running period in English FA premier league soccer matches. *The Journal of Strength & Conditioning Research, 27*(4), 909-915.

(25) Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P., & Drust, B. (2009). Analysis of high intensity activity in Premier League soccer. *International Journal of Sports Medicine, 30*(03), 205-212.

(26) Drust, B., & Green, M. (2013). Science and football: evaluating the influence of science on performance. J*ournal of Sports Sciences, 31*(13), 1377-1382.

(27) Drust, B., Atkinson, G., & Reilly, T. (2007). Future perspectives in the evaluation of the physiological demands of soccer. *Sports Medicine*, 37(9), 783-805.

(28) Duffield, R., Reid, M., Baker, J., & Spratford, W. (2010). Accuracy and reliability of GPS devices for measurement of movement patterns in confined spaces for court-based sports. *Journal of Science and Medicine in Sport, 13*(5), 523-525.

(29) International Organization for Standardization. (1994). *ISO 5725-1: 1994: Accuracy (Trueness and Precision) of Measurement Methods and Results-Part 1: General Principles and Definitions.* Geneva, Switzerland: International Organization for Standardization.

(30) Frencken, W. G., Lemmink, K. A., & Delleman, N. J. (2010). Soccer-specific accuracy and validity of the local position measurement (LPM) system. *Journal of Science and Medicine in Sport, 13*(6), 641-645.

(31) Golden, S. A., & Bateman, S. S. (2007). Sensor measurements for Wi-Fi location with emphasis on time-of-arrival ranging. *IEEE Transactions on Mobile Computing, 6*(10).

(32) Gu, Y., Lo, A., & Niemegeers, I. (2009). A survey of indoor positioning systems for wireless personal networks. *IEEE Communications Surveys & Tutorials, 11*(1), 13-32.

(33) Gudmundsson, J., & Horton, M. (2017). Spatio-temporal analysis of team sports. *ACM Computing Surveys (CSUR) 50*(2), 22.

(34) Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine 30*(1), 1–15.

(35) Hughes, M., & Franks, I. (2007). *The essentials of performance analysis: an introduction.* Routledge.

(36) Impellizzeri, F. M., & Marcora, S. M. (2009). Test validation in sport physiology: lessons learned from clinimetrics. *International Journal of Sports Physiology and Performance, 4*(2), 269-277.

(37) Iwase, S., & Saito, H. (2004, August). Parallel tracking of all soccer players by integrating detected positions in multiple view images. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference* on (Vol. 4, pp. 751-754). IEEE.

(38) Köklü, Y., Arslan, Y., Alemdaroğlu, U., & Duffield, R. (2015). Accuracy and reliability of SPI ProX global positioning system devices for measuring movement demands of team sports. *The Journal of Sports Medicine and Physical Fitness, 55*(5), 471-477.

(39) Krouwer, J. S. (2008). Why Bland–Altman plots should use X, not (Y+ X)/2 when X is a reference method. *Statistics in Medicine, 27*(5), 778-780.

(40) Lago-Peñas, C., Rey, E., & Lago-Ballesteros, J. (2012). The influence of effective playing time on physical demands of elite soccer players. *Open Sports Science Journal 5*, 188–92.

(41) Lames, M., & McGarry, T. (2007). On the search for reliable performance indicators in game sports. *International Journal of Performance Analysis in Sport, 7*(1), 62-79.

(42) Larsson, P. (2003). Global positioning system and sport-specific testing. *Sports Medicine, 33*(15), 1093-1101.

(43) Leser, R., Baca, A., & Ogris, G. (2011). Local positioning systems in (game) sports. *Sensors, 11*(10), 9778-9797.

(44) Letzelter, H., & Letzelter, M. (1982). *Die Struktur sportlicher Leistungen als Gegenstand der Leistungsdiagnostik in der Trainingswissenschaft.*

(45) Linke, D., Link, D., & Lames, M. (2018). Validation of electronic performance and tracking systems EPTS under field conditions. *PloS one, 13*(7), e0199519.

(46) Mackenzie, R., & Cushion, C. (2013). Performance analysis in football: A critical review and implications for future research. J*ournal of Sports Sciences, 31*(6), 639-676.

(47) Malone, J. J., Lovell, R., Varley, M. C., & Coutts, A. J. (2017). Unpacking the black box: applications and considerations for using GPS devices in sport. *International Journal of Sports Physiology and Performance, 12*(Suppl 2), S2-18.

(48) Manafifard, M, Ebadi, H., and Moghaddam, H. A. (2017). A survey on player tracking in soccer videos. *Computer Vision and Image Understanding 159*, 19–46.

(49) Merriaux, P., Dupuis, Y., Boutteau, R., Vasseur, P., & Savatier, X. (2017). A study of Vicon system positioning performance. *Sensors, 17*(7), 1591.

(50) Muthukrishnan, K. (2009). Multimodal Localisation Analysis, Algorithms and Experimental Evaluation. *Ph.D. Dissertation, University of Twente, Enschede, The Netherlands*.

(51) O'Donoghue, P. (2009). *Research methods for sports performance analysis.* Routledge.

(52) Ogris, G., Leser, R., Horsak, B., Kornfeind, P., Heller, M., & Baca, A. (2012). Accuracy of the LPM tracking system considering dynamic position changes. *Journal of Sports Sciences 30*(14), 1503–1511.

(53) Paul, D. J., Bradley, P. S., & Nassis, G. P. (2015). Factors affecting match running performance of elite soccer players: Shedding some light on the complexity. *International Journal of Sports Physiology and Performance, 10*(4), 516-519.

(54) Rampinini, E., Alberti, G., Fiorenza, M., Riggio, M., Sassi, R., Borges, T. O., & Coutts, A. J. (2015). Accuracy of GPS devices for measuring high-intensity running in field-based team sports. *International Journal of Sports Medicine, 36*(01), 49-53.

(55) Rampinini, E., Coutts, A. J., Castagna, C., Sassi, R., & Impellizzeri, F. M. (2007). Variation in top level soccer match performance. *International Journal of Sports Medicine, 28*(12), 1018-1024.

(56) Rawstorn, J. C., Maddison, R., Ali, A., Foskett, A., & Gant, N. (2014). Rapid directional change degrades GPS distance measurement validity during intermittent intensity running. *PloS one, 9*(4), e93693.

(57) Redwood-Brown, A., Cranton, W., & Sunderland, C. (2012). Validation of a real-time video analysis system for soccer. *International Journal of Sports Medicine 33*(8), 635–640.

(58) Rein, R., & Memmert, D. (2016). Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. *SpringerPlus 5*(1), 1410.

(59) Robertson, S. J., Burnett, A. F., & Cochrane, J. (2014). Tests examining skill outcomes in sport: a systematic review of measurement properties and feasibility. *Sports Medicine, 44*(4), 501-518.

(60) Saini, M., Kerrigan, D. C., Thirunarayan, M. A., & Duff-Raffaele, M. (1998). The vertical displacement of the center of mass during walking: a comparison of four measurement methods. *Journal of Biomechanical Engineering, 120*(1), 133-139.

(61) Sathyan, T., Shuttleworth, R., Hedley, M., & Davids, K. (2012). Validity and reliability of a radio positioning system for tracking athletes in indoor and outdoor team sports. *Behavior Research Methods, 44*(4), 1108-1114.

(62) Schipperijn, J., Kerr, J., Duncan, S., Madsen, T., Klinker, C. D., & Troelsen, J. (2014). Dynamic accuracy of GPS receivers for use in health research: a novel method to assess GPS accuracy in real-world settings. *Frontiers in Public Health, 2*, 21.

(63) Schlipsing, M., Salmen, J., Tschentscher, M., & Igel, C. (2017). Adaptive pattern recognition in real-time video-based soccer analysis. *Journal of Real-Time Image Processing 13*(2), 345–361.

(64) Schutz, Y., & Chambaz, A. (1997). Could a satellite-based navigation system (GPS) be used to assess the physical activity of individuals on earth?. *European Journal of Clinical Nutrition, 51*(5), 338.

(65) Scott, M. T., Scott, T. J., & Kelly, V. G. (2016). The validity and reliability of global positioning systems in team sport: a brief review. *The Journal of Strength & Conditioning Research 30*(5), 1470–1490.

(66) Siegle, M., Stevens, T., & Lames, M. (2013). Design of an accuracy study for position detection in football. *Journal of Sports Sciences 31*(2), 166–172.

(67) Stelzer, A., Pourvoyeur, K., & Fischer, A. (2004). Concept and application of LPM-a novel 3-D local position measurement system. *IEEE Transactions on Microwave Theory and Techniques 52*(12), 2664–2669.

(68) Thomas, G., Gade, R., Moeslund, T. B., Carr, P., and Hilton, A. (2017). Computer vision for sports: current applications and research topics. *Computer Vision and Image Understanding 159*, 3–18.

(69) Thomas, J. R., Silverman, S., & Nelson, J. (2015). *Research methods in physical activity, 7E.* Human kinetics.

(70) Townshend, A. D., Worringham, C. J., & Stewart, I. B. (2008). Assessment of speed and position during human locomotion using nondifferential GPS. *Medicine & Science in Sports & Exercise 40*(1), 124–132.

(71) Valter, D. S., Adam, C., Barry, M., & Marco, C. (2006). Validation of Prozone®: A new video-based performance analysis system. *International Journal of Performance Analysis in Sport 6*(1), 108–119.

(72) van Diggelen, F., & Enge, P. (2015, September). The worlds first gps mooc and worldwide laboratory using smartphones. In *Proceedings of the 28th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2015)* (pp. 361-369).

(73) Varley, M. C., Fairweather, I. H., & Aughey, R. J. (2012). Validity and reliability of GPS for measuring instantaneous velocity during acceleration, deceleration, and constant motion. *Journal of Sports Sciences, 30*(2), 121-127.

(74) Versi, E. (1992). " Gold standard" is an appropriate term. *BMJ: British Medical Journal, 305*(6846), 187.

(75) Virmavirta, M., & Isolehto, J. (2014). Determining the location of the body′ s center of mass for different groups of physically active people. *Journal of Biomechanics, 47*(8), 1909-1913.

(76) Witte, T., & Wilson, A., (2004). Accuracy of non-differential GPS for the determination of speed over ground. *Journal of Biomechanics, 37*(12), 1891–1898.

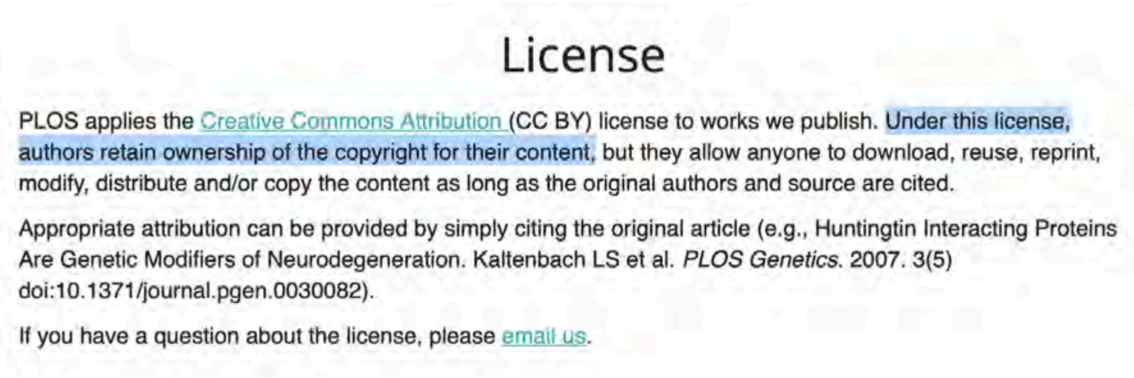
(77) Zaidi, M., Tourki, R., & Ouni, R. (2010, March). A new geometric approach to mobile position in wireless LAN reducing complex computations. In *Design and Technology of Integrated Systems in Nanoscale Era (DTIS)*, (pp. 1-7). IEEE.

(78) Zhang, N., Duan, L. Y., Li, L., Huang, Q., Du, J., Gao, W., & Guan, L. (2012). A generic approach for systematic analysis of sports videos. *ACM Transactions on Intelligent Systems and Technology (TIST), 3*(3), 46.

Supplementary Material *(USB Stick)*

- **Video 1:** Deviation between pre-calculated and actual distance covered by the center of pelvis COP and center of shoulders COS
- **Video 2:** Center of mass displacement during standing
- **Video 3:** VICON small-sided game validation example
- **Video 4:** VICON circuit validation example
- **Video 5:** VICON center of pelvis vs. center of mass visualization

PLOS Permission to Publish



License

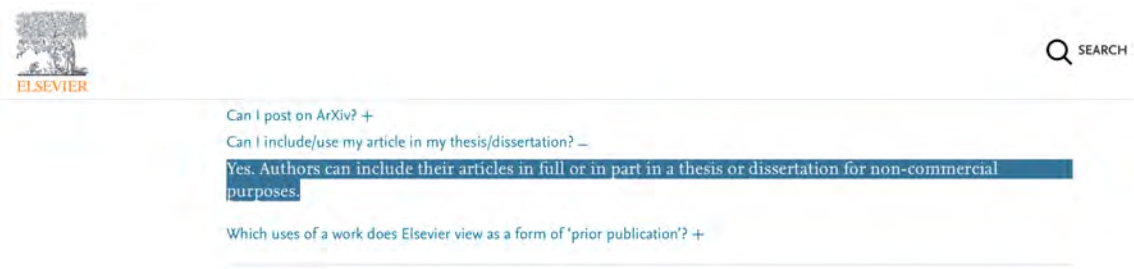
PLOS applies the Creative Commons Attribution (CC BY) license to works we publish. Under this license, authors retain ownership of the copyright for their content, but they allow anyone to download, reuse, reprint, modify, distribute and/or copy the content as long as the original authors and source are cited.

Appropriate attribution can be provided by simply citing the original article (e.g., Huntingtin Interacting Proteins Are Genetic Modifiers of Neurodegeneration. Kaltenbach LS et al. *PLOS Genetics*. 2007. 3(5) doi:10.1371/journal.pgen.0030082).

If you have a question about the license, please email us.

https://www.plos.org/license

Elsevier Permission to Publish



Can I post on ArXiv? +

Can I include/use my article in my thesis/dissertation? –

Yes. Authors can include their articles in full or in part in a thesis or dissertation for non-commercial purposes.

Which uses of a work does Elsevier view as a form of 'prior publication'? +

https://www.elsevier.com/about/policies/copyright/permissions

JSSM Permission to Publish

This work is licensed under a (cc) BY-NC-ND Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

http://creativecommons.org/licenses/by-nc-nd/4.0/

IJPAS Permission to Publish

O'Donoghue, Peter <PODonoghue@cardiffmet.ac.uk>
Linke, Daniel
Monday, October 29, 2018 at 11:35 PM
Show Details

Hi Daniel
The way round this is to put the last draft of the paper rather than the edited proof into the thesis
POD

From: Linke, Daniel <daniel.linke@tum.de>
Sent: 29 October 2018 13:33:57
To: O'Donoghue, Peter
Subject: Publication of Journal Article in Dissertation

Dear Peter

We published this article in 2016: **Linke, Daniel, and Martin Lames. "Substitutions in elite male field hockey–a case study."** ***International Journal of Performance Analysis in Sport*** **16.3 (2016): 924-934.**

The article is a part of my dissertation, which I am writing at the moment, and I would like to include the full article in the dissertation (in the Appendix).
Since the dissertation will be accessible to the public, this could be a violation of the Journal's copy rights.

Therefore, I would like to ask under what circumstances it would be possible to append the article in the appendix of my dissertation.

Thank you for your support.

Best regards

Daniel