Samuel Schleitzer · Svenja Wirtz · Ross Julian · Eric Eils

Institute of Sport and Exercise Sciences, University of Münster, Münster, Germany

Development and evaluation of an inertial measurement unit (IMU) system for jump detection and jump height estimation in beach volleyball

Introduction

Wearables are commonly implemented into many aspects of human performance and health. Likewise, the number of literature reports on these body-worn devices has surged and these devices collect a wide variety of motion-related data to provide wearers with automated evaluated biofeedback (Peake, Kerr, & Sullivan, 2018). Notably, the use of wearables is also changing rapidly in the clinical area where it has rapidly evolved from safe settings in laboratories to applied use in unsupervised practice (Picerno et al., 2021). Despite this rapid progress (including in sports), it is often unclear whether these devices meet the criteria of test validity, reliability, and objectivity. For example, one review article reported that only 5% of devices have been formally validated and tested in real-world settings. The authors recommend that future research should aim to validate wearable devices in ecological settings, to provide and improve measurement accuracy (Peake et al., 2018; Schmidt, Alt, Nolte, & Jaitner, 2020).

In sports, the use of wearables is already common in the individual fitness and running sectors, and these devices are also finding their way into complex sports games (Camomilla, Bergamini, Fantozzi, & Vannozzi, 2018). As they become more prevalent in complex game situations, these devices must fulfill greater requirements from a measurement and evaluation perspective, as players' move-ments are often more complex and can be influenced by teammates and opponents.

In recent years, the popularity of sand sports has increased immensely, which is highlighted by beach volleyball's consistent presence as an Olympic sport and the consideration of beach handball as another Olympic sand sport (Handball-World.news, 2021). Regarding wearables, sand surfaces place additional, increased demands on motion detection than do solid surfaces, since sand surfaces are uneven and changeable, i.e. it changes it shape with every contact. Transferring the technology used on a solid ground to the sand is challenging, and systems that are promising on solid ground may not necessarily be suitable for sand without specific adaptation.

However, sand sports, especially beach volleyball, could greatly benefit from a validated system for automated jump detection and jump height measurement, and both are important parameters of the load structure of beach volleyball (Künkler, 2009). In men's international match-play, defenders complete about 68 jumps and blockers 106 jumps per hour, respectively (Czimek, 2017). In addition to these times of high load during tournaments, player may also experience high loads when jumping during weekly training. The high load in beach volleyball can also be accompanied by typical overuse injuries (e.g. patellar tendinitis) (Bahr & Reeser, 2003), and in this context the amount of jump training may be a risk factor (Sprague, Smith, Knox, Pohlig, & Grävare Silbernagel, 2018). Therefore, coaches are advised to consider how much jumping their athletes are doing during training and play to protect them from overuse (Bahr & Bahr, 2014).

The analysis of jumps and jump loads in beach volleyball is usually time consuming and complex measurements in the laboratory (e.g. three-dimensional motion analysis, force plates) are often required. Similarly, manually counting jumps from a video-recorded match is also time consuming, impractical and lacks specific information like force production and jumping heights. Thus, a system that counts and categorizes jumps in an automated manner may help protect athletes from overuse injuries, rehabilitate them after injuries, and provide validated knowledge regarding overuse injuries (Jarning, Mok, Hansen, & Bahr, 2015). In addition, such a system would open new possibilities for optimizing performance and for diagnostics by allowing coaches to control athletes' load in training and during play. A highly desirable system would be one that can be used for screening to prevent injury and in rehabilitation, for performance diagnostics and training control in the field, and to also periodize training and maximize balance between load and recovery.

Commercial systems that count jumps automatically and measure jump height became available over the last few years and have been tested on rigid surfaces
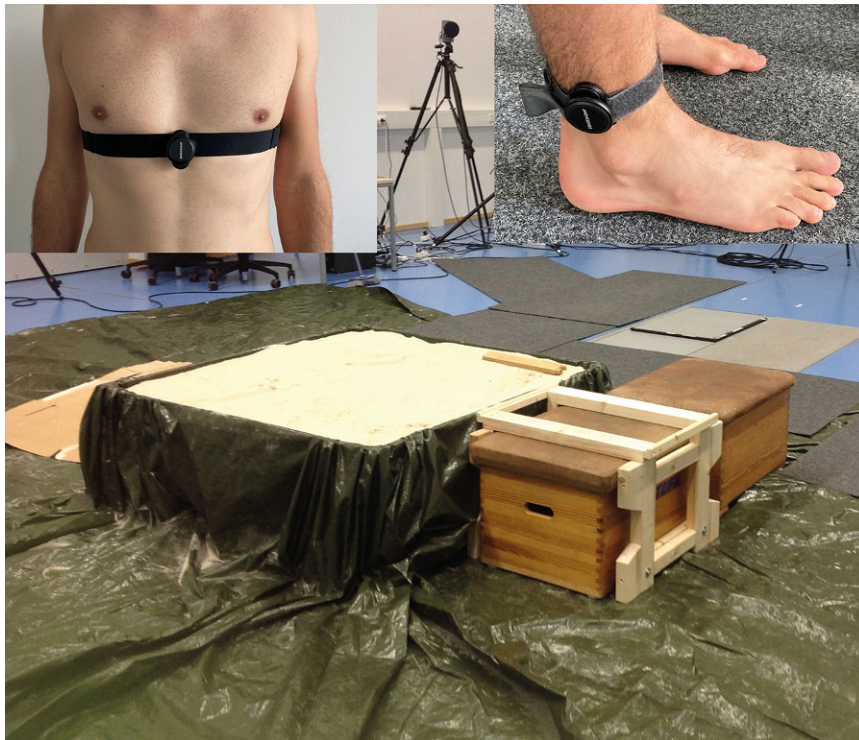
Fig. 1 ▲ Laboratory setup when measuring countermovement jumps in the sandbox, which was placed on three force plates. The figure also shows the placement of the sensors on the chest (*upper left*) and ankle (*upper right*) to detect jumps and to measure jumping height

(Benson et al., 2020; Borges et al., 2017; Charlton, Kenneally-Dabrowski, Sheppard, & Spratford, 2017; Jarning et al., 2015; MacDonald, Bahr, Baltich, Whittaker, & Meeuwisse, 2017; Skazalski, Whiteley, Hansen, & Bahr, 2018). Recently, a proprietary system developed for indoor volleyball was validated for beach volleyball on a sand surface (Schmidt, Meyer, & Jaitner, 2021). The authors reported that the commercially available system provides good validity for jump count and jump height; however, the device had limitations as it detected a high number of false-positive jumps on the sand compared to a rigid surface (Charlton et al., 2017). In addition, jump heights were found to be inaccurate, with the direction of performance being unsystematically biased. One major problem in using proprietary systems is that the user normally has no access to the underlying algorithms and signal processing procedures. In addition, the algorithms were mainly programmed for a rigid surface interaction and the change to a sand surface may force the system in terms of event

identification to validly count jumps and measures jump height. IMU systems that are freely available and programmable are also present on the market. With the help of these systems, demands of beach volleyball can specifically be addressed and extended for future requirements.

Therefore, the first aim of the present study was to develop and validate an IMU system that detects jumps and measures jump height in the sand under standardized conditions. For this purpose, we used a customized sandbox positioned on force plates in an experimental laboratory setting. The experimental group consisted of physical education students. A second aim of the present study was to validate the jump detection of the IMU system under ecologically valid conditions in the field. For this purpose, we compared jumps detected by the IMU system with jumps detected in a synchronized video. The experimental group consisted of experienced beach volleyball players.

Methods

All subjects completed informed written consent to participate in the study before measurements began, and all procedures performed were in accordance with the Declaration of Helsinki and the local ethics committee (ID 2021-53-EE).

In previously comprehensive investigations and pilot studies, we ensured the measurement capability of the system with synchronous acceleration signals from the sensors on individual test subjects. In the present investigation, which has two parts, we first used the IMU to measure standardized countermovement jumps (CMJs) under laboratory conditions on sand surfaces and compared the measurements with those made from a gold standard system (force plates). We used standardized CMJs for two reasons: (i) to simplify CMJs and to reduce variability in movement executions due to arm swing and subsequently in acceleration signals of the sensors, (ii) CMJs with hand on the hips while jumping is the standardized form when performing performance diagnostics in jumping. Based on these results, we tested the suitability of the system in the field using complex beach volleyball-specific jump actions. Jump height detection was not performed in the field because of a missing second measuring system for validation under more ecological conditions.

Measurements under standardized laboratory conditions

Twenty physical education students (5 women and 15 men; 182 ± 10 cm; 75 ± 10 kg) participated in the first measurement. Before the tests, a warm-up was performed in a self-instructed manner. Each subject completed five CMJs on two consecutive days. Each subject had two trials to become familiar with the environment and with jumping on an unstable surface. Participants started in a hip wide stance with fully extended knees. Hands were placed on hips throughout the jump. The jump was initiated with a quick downward movement up to a 90°knee angle, followed by an immediate extension of the legs

producing the upward jump movement. Legs were required to be in a straight position while in the air, as well as ankle plantarflexion during landing. Incorrect executions were corrected immediately, and the jump had to be repeated.

The measurements were taken in a self-constructed sandbox (size 1.25 m × 1.25 m × 0.3 m) supported on three different force plates (Type: 9287C, Kistler, Winterthur, Switzerland). Thus, the ground reaction forces could be recorded, and the jump height could be calculated in sand under laboratory conditions. Simultaneously with the force measurements, acceleration data of an IMU system were recorded. The IMU system consisted of two IMUs from Movesense (Suunto Oy, Finland), which were attached to participants' sternum and lateral malleolus (◘ Fig. 1). Data acquisition of this system was performed using a specific software that also synchronized acceleration with video data (Data Collector, Kassa solution GmbH, Düsseldorf, Germany). The measurement frequency of the force plates was 1500 Hz and the measurement frequency from each sensor was 52 Hz (highest possible and stable frequency when using several sensors at the time of the study).

The sand fulfilled the specifications of the German beach volleyball federation for indoor sand (grain size: 0.1–1.0 mm; grain shape: from round edges to rounded; grain distribution: even; $CaCO_3 \leq 2$–3%; $SiO_2 \geq 95$–98%) (Borrmann et al., 2009).

For data processing, the vertical component of the ground reaction force and the acceleration data from the sensors were processed and analysed using MATLAB (R2019b, MathWorks Inc., Natick, MA, USA). For the force data, the flight phase of the CMJ was detected using the raw unfiltered data (force value less than or greater than 20 N). A threshold of 20 N was used to account for oscillations in the vertical force data due to swinging of the sandbox. For the IMU data, synchronous video and acceleration data were first used to subdivide acceleration data into portions with relevant information. Overall, the algorithm searched for sequences of extrema in the acceleration data from the

S. Schleitzer · S. Wirtz · R. Julian · E. Eils

Development and evaluation of an inertial measurement unit (IMU) system for jump detection and jump height estimation in beach volleyball

Abstract

Wearables are commonly used in practice for measuring and monitoring performance in high-level sports. That being said, they are often designed and intended for use during sports conducted on rigid surfaces. As such, sports that are conducted on sand, e.g. beach volleyball, lack equipment that can be specifically applied in the field. Therefore, the aim of this study was to develop and validate an inertial measurement unit (IMU)-based system for automatic jump detection and jump height measurement in sand. The system consists of two IMUs, which were attached to different parts of the athletes' bodies. For validation under laboratory conditions, 20 subjects each performed five jumps on two consecutive days in a sandbox placed on force plates. Afterwards, five beach volleyball athletes performed complex combinations of beach volleyball-specific movements and jumps wearing the IMUs whilst being video recorded simultaneously. This was conducted in an ecologically valid setting to determine the validity of the IMU to correctly detect jumping actions. The results of the laboratory tests show excellent day-to-day reliability (intraclass correlation coefficient [ICC] = 0.937, two-way mixed effects, single measurement, consistency) and excellent concurrent validity (ICC = 0.946, two-way mixed effects, single rater, absolute agreement) compared to the gold standard (force plates). The accuracy in jump detection of the IMU was 100 and 97.5% in the laboratory and ecologically valid settings, respectively. Although there are still some aspects to consider when using such devices, the current findings provide recommendations regarding best practice when using such a device on a variable and unstable surface. Collectively, such a device could be applied in the field to provide coaches and practitioners with direct feedback to monitor training or match play.

Keywords

Jump detection · Sand surface · Validation · Wearable · Diagnostics

different sensors in different directions (◘ Fig. 2). First, it searched for take-offs and landings separately using each sensor, and then time-conforming jumps and landings of both sensors were linked to possible jumps. Information of both sensors were then used to classify take-off and landings to finally calculate jump height over the flight time using the following formula:

$$jump\ height\ h = 9.81 \frac{m}{s^2} \cdot \frac{t^2}{8}$$

where t is flight time in seconds.

For statistical analysis of the laboratory-based measurements, day-to-day reliability (test–retest on two consecutive days) was performed using the intraclass correlation coefficient (ICC; two-way mixed effects, single measurement, consistency) for both the force plate and the IMU system jump heights. In addition, the Pearson correlation coefficient was determined. The interrater reliability between the two measurement systems

(force plates as the gold standard) was determined using the intraclass correlation coefficient ICC (two-way mixed effects, single rater, absolute agreement). Additional parameters (bias, limit of agreements) and Bland–Altman plots were used to represent the deviations of the measurement systems (Bland & Altman, 1986).

The reliability determinations were carried out with different types of data (individual measurements and average measurements). For the individual measurements, one time we considered all measured trials without exception, and another time we considered all measurements where the coefficient of variation (CV) of the jump heights of one subject did not exceed the 10% mark. The purpose of selecting the data in this way was to exclude some unrealistic outliers that arose because of the measurement changeability of the sand and the uneven sand surface after take-off and before landing and algorithm's inadequacies.
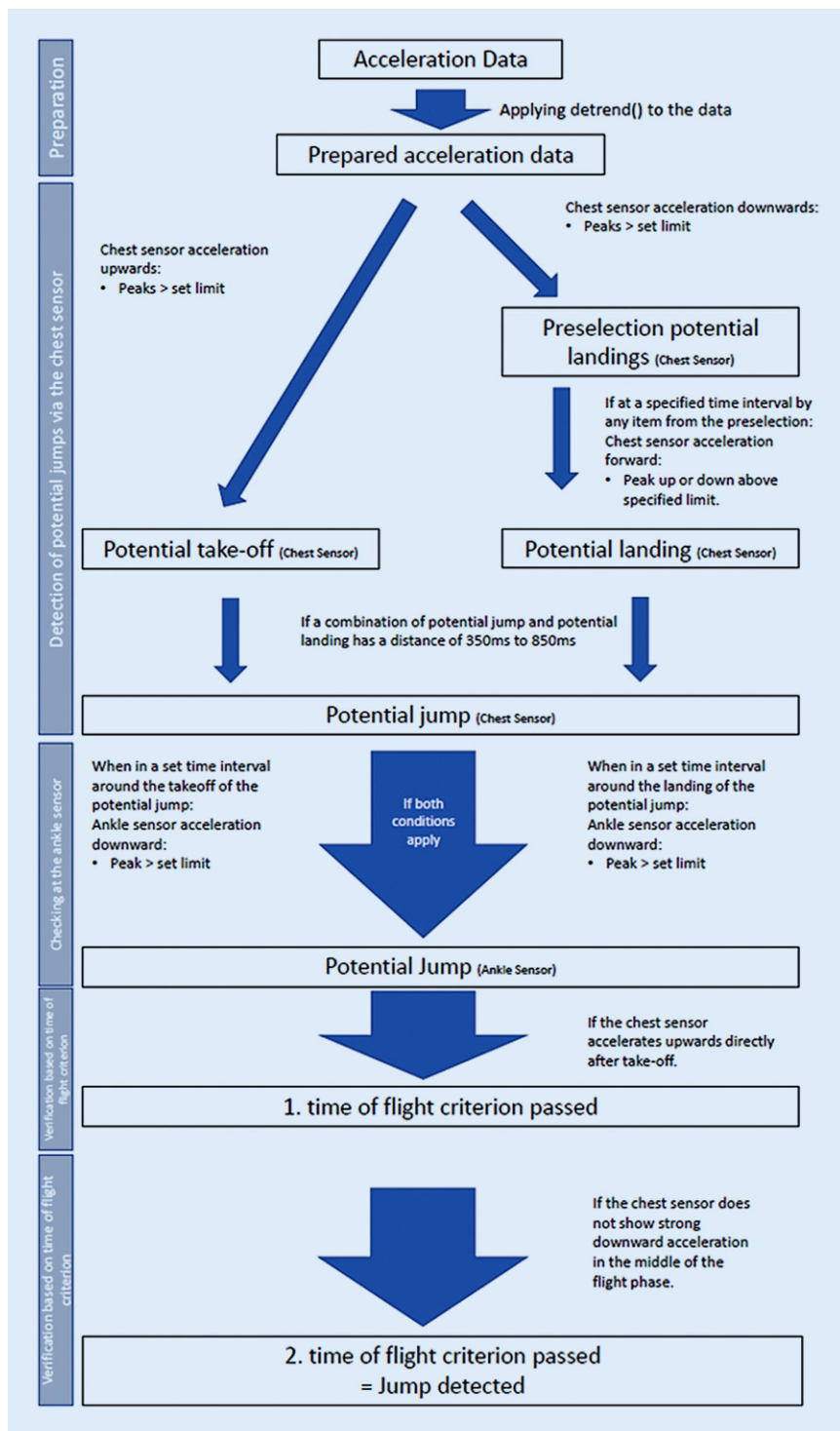
Fig. 2 ▲ Operating principles of the Matlab algorithm for jump detection

For the average measurements, the reliability analysis was determined one time with the mean of all 5 trials and one time with the mean of the 3 middle trials (without the maximum and minimum trials).

Field measurements

Five subjects (3 women and 2 men; 177 ± 4 cm; 68 ± 4 kg) with comprehensive beach volleyball experience (tournament experience at the highest regional level) participated in the ecologically valid measurements on an outdoor beach field. Prior to testing, subjects warmed up independently, and subsequently the IMU system was attached to participants' sternum and lateral malleolus and connected to an iPad (6th generation, Apple Inc., CA, USA). Subjects then performed 32 different trials consisting of attack jumps, block jumps, jump serves, as well as dives or beach volleyball-typical combinations. For example, one complex exercise consisted of serving, sprinting to the net, and then performing a blocking action. The last four trials of each subject were performed individually, i.e. subjects performed self-invented beach volleyball specific rallies, leading to a different number of jump action for each player. The complex movement sequences were stored on the iPad together with the synchronous video recordings previously described. The videos were used to categorize the movements and clearly assign them to the important sections of the acceleration signals measured by the IMU system. The measuring frequency of the IMU systems was also 52 Hz and the sand also fulfilled the specifications of the German beach volleyball federation for indoor sand like mentioned in part one of the study.

For data processing, the aforementioned algorithm was used, further developed and optimized using pilot data from the field to identify and count jumps in the field.

For statistical analysis of the ecologically valid measurements, the jumps recognized by the IMU were compared to the video recorded jumps. According to Charlton et al. (2017) the jumps we categorized as true positive (TP; a jump detected on video and by the IMU system), false positive (FP; a jump detected by the IMU system, but not on video) and false negative (FN; a jump detected on video, but not on the IMU system). These values were used to represent TP, FP and FN proportionately as percentages according to Skazalski et al. (2018).

Table 1 Day-to-day reliability of jumping height on sand surface

	Jump height		Bland–Altman parameters			Correlation	
	Day 1	Day 2	Bias	LoA–	LoA+	ICC (3,1) Consistency [CI]	Pearson [CI], *p*-value
Force plate measurements							
All trials (*n* = 100)	28.7 ± 5.5	30.2 ± 5.7	−1.44	−7.17	4.29	0.866 [0.817–0.902]	0.866 [0.807–0.908], *p* < 0.001
All trials without CV > 10% (*n* = 100)	28.7 ± 5.5	30.2 ± 5.7	−1.44	−7.17	4.29	0.866 [0.817–0.902]	0.866 [0.807–0.908], *p* < 0.001
Mean of all 5 trials (*n* = 20)	28.7 ± 5.4	30.2 ± 5.7	−1.44	−5.26	2.37	0.934 [0.872–0.971]	0.940 [0.851–0.976], *p* < 0.001
Mean of 3 middle trials (*n* = 20)	28.7 ± 5.4	30.2 ± 5.7	−1.44	−5.26	2.37	0.939 [0.872–0.971]	0.940 [0.860–0.976], *p* < 0.001
IMU system measurements							
All trials (*n* = 100)	30.7 ± 5.7	30.9 ± 6.6	−0.14	−8.44	8.15	0.767 [0.690–0.827]	0.774 [0.681–0.842], *p* < 0.001
All trials without CV > 10% (*n* = 88)	30.8 ± 5.7	30.8 ± 6.0	0.16	−6.62	6.95	0.827 [0.766–0.873]	0.831 [0.753–0.887], *p* < 0.001
Mean of all 5 trials (*n* = 20)	30.7 ± 5.3	30.9 ± 6.3	−0.14	−4.73	4.44	0.922 [0.839–0.963]	0.932 [0.832–0.973], *p* < 0.001
Mean of 3 middle trials (*n* = 20)	30.7 ± 5.4	30.6 ± 6.2	0.04	−4.09	4.18	0.937 [0.868–0.970]	0.942 [0.857–0.977], *p* < 0.001

Please note that all 100 jumps on each day were analysed in the row "all trials." Some trials analysed with the inertial measurement unit (IMU) system showed unrealistic deviations from the remaining 4 trials (coefficient of variation > 10%) that did not derive from variations within the jumping performance but from unrealistic outliers. These trials were excluded in the row "all trials without CV > 10%." The mean of all 5 trials or the mean of the 3 middle trials (without minimum and maximum values) were used for test–retest reliability analysis of subjects' mean values

LoA– lower limit of agreement resp. lower limit of the 95% confidence interval, *LoA+* upper limit of agreement resp. upper limit of the 95% confidence interval, *ICC* Intraclass correlation coefficient, *CV* Coefficient of variation, *CI* Confidence interval

Table 2 Reliability between measuring systems—force plates (the gold standard) and the IMU system—for jumping height on a sand surface (day 2)

	Jump height		Bland-Altman parameters			Correlation	
Reliability between systems	Force plates	IMU system	Bias	LoA–	LoA+	ICC (2,1) Abs. Agreement [CI]	Pearson [CI], *p*-value
All trials (*n* = 100)	30.2 ± 5.7	30.9 ± 6.6	−0.71	−7.06	5.64	0.859 [0.807–0.897]	0.872 [0.815–0.912], *p* < 0.001
All trials without CV > 10% (*n* = 93)	30.2 ± 5.7	30.8 ± 6.0	−0.44	−5.13	4.26	0.916 [0.885–0.939]	0.918 [0.879–0.945], *p* < 0.001
Mean of all 5 trials (*n* = 20)	30.2 ± 5.7	30.9 ± 6.3	−0.71	−4.66	3.24	0.939 [0.873–0.972]	0.948 [0.870–0.980], *p* < 0.001
Mean of 3 middle trials (*n* = 20)	30.2 ± 5.7	30.6 ± 6.2	−0.47	−4.28	3.33	0.946 [0.888–0.975]	0.950 [0.875–0.980], *p* < 0.001

Please note that all 100 jumps on each day were analysed in the row "all trials." Some trials analysed with the inertial measurement unit (IMU) system showed unrealistic deviations from the remaining 4 trials (coefficient of variation > 10%) that did not derive from variations within the jumping performance but from unrealistic outliers. These trials were excluded in the row "all trials without CV > 10%." The mean of all 5 trials or the mean of 3 middle trials (without minimum and maximum value) were used for test–retest reliability analysis of subjects' mean values

LoA– lower limit of agreement resp. lower limit of the 95% confidence interval, *LoA+* upper limit of agreement resp. upper limit of the 95% confidence interval, *ICC* Intraclass correlation coefficient, *CV* Coefficient of variation, *CI* Confidence interval

Results

Day-to-day reliability on sand surfaces in the lab

The results of the reliability measurements in the laboratory for both individual and mean jump height values of physical education students are shown in ◘ Table 1.

Results for the force plate measurements show good to excellent correlation coefficients for analyses of both the individual values and the average. Reliability improved when moving from single to mean values. The range in the Bland–Altman parameters decreased when averages were used instead of individual tests.

An influence on the offset cannot be seen overall.

Results for the IMU system show good day-to-day reliability for individual values and excellent day-to-day reliability for averages. Reliability improved when moving from individual to mean values. The IMU results had larger variability; however, this was reduced when values greater than 10% of the coefficient of variation were removed. Using the mean values, we also observed an increase in the day-to-day reliability when only the three middle trials were used instead of all five individual trials per subject (such that maximum and minimum trials were excluded). The range for the Bland–Altman parameters decreased when the mean values were used instead of the individual trials.

Concurrent validity

◘ **Figure 3** shows the Bland–Altman plot of all jumps measured with both measuring systems on day 2 (day 2 was chosen, due to higher familiarization of the subject with the sand). The limits of agreement defined a tolerance interval from −7.06 to 5.64 cm (range 12.7 cm). The bias between the measurement systems is about 0.58 cm (the IMU system overestimates the results of the force plate). Overall, unrealistic outliers occur with the IMU system (values outside the tolerance range).
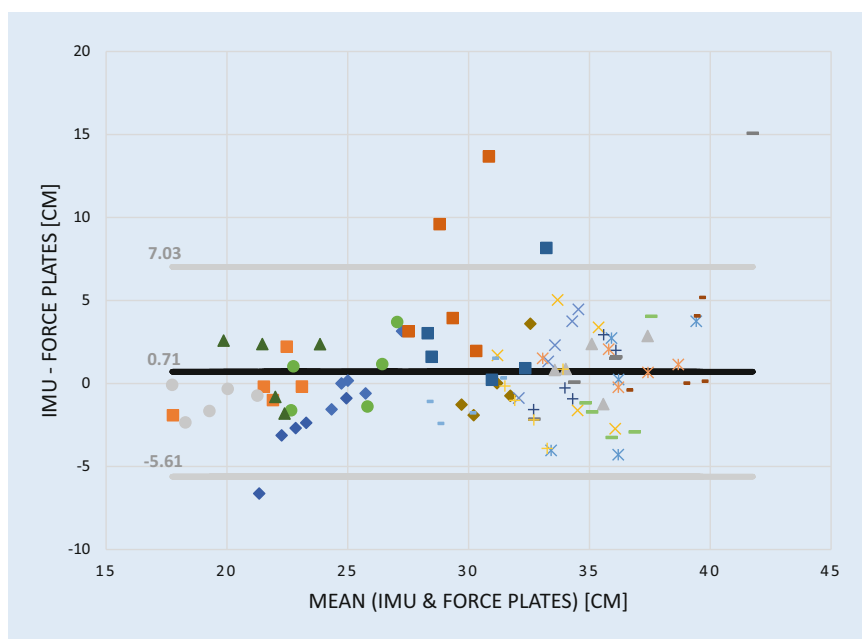
Fig. 3 ▲ Bland–Altman plot showing the association and limits of agreement between jumping heights measured with force plates and inertial measurement unit (IMU) system. The *horizontal black line* indicates the bias, and the *shaded lines* delimit the 95% confidence interval (limits of agreement)

ICCs for concurrent validity show good to excellent correlation coefficients for individual values and excellent concurrent validity for the averages. Concurrent validity improved when moving from individual to mean values. When using the mean values, we also observed an increase in concurrent validity when only the three middle trials were used instead of all five individual trials per subject (such that maximum and minimum trials were excluded). The range of the Bland–Altman parameters was reduced when mean values were used instead of single values. Overall, this did not recognizably influence the bias (◘ **Table 2**).

Jump detection in the laboratory and in the field

Jump detection of CMJ in the laboratory and different jump types in the field are presented in ◘ **Table 3**. The results show that all CMJs in the laboratory were correctly identified as jumps. In the field, a total of 95.9% (306) of the 319 jumps were correctly identified as jumps. A large proportion of the nonrecognized jumps (false negatives) occurred during serves.

◘ **Table 4** shows the jump detection results in the field divided by subject. The rate of detected jumps in the field for each subject ranged from 91.8 to 100%. In addition to the detected jumps, the system classified an additional 14 jumps as jumps that were not classified as jumps in the video (false positives). A majority (71%) of all false positives were performed by subject 3. The rest were distributed among all other test subjects.

Discussion

The aim of the present study was to develop and validate an IMU system that detects jumps and measures jump height in the sand under standardized conditions and detects jumps in the field under real conditions. The results show good to excellent day-to-day reliability and concurrent validity for jumping height, an excellent 100 and 96% true positive detection rate and a low 0 and 4% false positive detection rate in the laboratory and the field, respectively. Therefore, the system is highly suitable for performance diagnostics under standardized conditions and in detecting jumps in the field on a sand surface.

Jump detection

Jump detection in indoor sports have long been in use and jump detection rate ranges from 95 to 99.8% (Charlton et al., 2017; Gageler, Wearing, & James, 2015; Skazalski et al., 2018), mainly using the commercially available VERT™ system (VERT™, Fort Lauderdale, FL, USA) on a rigid indoor surface. False positive detection rate in the field varied from 2.3% (Skazalski et al., 2018) to 12.1% (Charlton et al., 2017) in indoor volleyball. A recent investigation reported a jump detection (true positive) and a false positive rate of 97.5 and 10.7% with the VERT™ on a sand surface, respectively (Schmidt et al., 2021). Given these ranges, the jump detection of 95.9 and 4.4% of our IMU system appears to be excellent in detecting true and false positive jumps on a sand surface and is within the range or better than commercially available systems. The low false positive detection rate of the IMU system used in the current study in comparison to the VERT™ may be due to the number of sensors used within both systems. The VERT™ system consists of one IMU that is attached to an elastic waistband directly below the navel, whereas our system consists of two sensors attached to the sternum and the ankle. The combination of these two signals allows to differentiate between acceleration of the leg and the upper body and thus to separate sprints or changes of direction from jumps.

Despite the high success in jump detection rate, jump type and individual results (◘ **Tables 3 and 4**) also play an important role in that context. One drawback is that during serving, a lesser percentage of jumps were detected (82.6%) and 17.4% of the serves were not detected (in contrast, the technique of attack jump is always executed from the same sequence of steps and the recognition rate is around 96.9%). However, it is important to note that 71% of misclassified jump actions originated from one subject. This could potentially be due to individual techniques of performing jumping actions. For example, the difficulty in detecting the jump serve arose from the many ways participants can execute

Table 3 Identification of jumps during laboratory and in-field measurements

	Jump type	Video	True positives (n, %)	False negatives (n, %)
Laboratory	CMJ	200	200 (100.0%)	0 (0.0%)
Field	Block	137	133 (97.1%)	4 (2.9%)
	Attack	159	159 (96.9%)	5 (3.1%)
	Serve	23	19 (82.6%)	4 (17.4%)
Total		519	506 (97.5%)	13 (2.5%)

Please note that the number of jump serves is lower than block or attack jumps due to the specific structure of the complex exercises, where one trial contained only one serve but several block or attack jumps

CMJ countermovement jumps

Table 4 Identification of jumps during in-field measurements separated by subjects

	Video	True positives (n, %)	False negatives (n, %)	False positives (n, %)
Subjects				
1	55	55 (100.0%)	0 (0.0%)	1 (1.8%)
2	60	58 (96.7%)	2 (3.3%)	1 (1.7%)
3	73	67 (91.8%)	6 (8.2%)	10 (13.0%)
4	66	62 (93.9%)	4 (6.1%)	0 (0.0%)
5	65	64 (98.5%)	1 (1.5%)	2 (3.0%)
Total	319	306 (95.9%)	13 (4.1%)	14 (4.4%)

the serve (different ways of approaching the serve, throwing the ball, and then hitting it with float/rotation). As such, for one individual who performed jump serves, the IMU system had the lowest recognition rate and the highest number of false positives due to this individual's technique; specifically, "false positives" were triggered by this individual's particularly strong sidestep movement. To improve the robustness of the algorithm, jump types and individual technical ranges should be considered. Within the sample, this individual also had the lowest overall performance level. This suggests that the algorithm may more easily detect complex movements when performed at higher expertise levels but in general, algorithms must be adapted to individual performance characteristics at all levels of play. Schmidt et al. (2021) also reported different level of recall for jump types and increased limits of agreement for individual participants showing that this is a general problem to capture complex movements with IMU systems.

Day-to-day reliability

The overall day-to-day reliability of sand measurements in a sandbox on force plates and the IMU system can be rated as good or excellent (Koo & Li, 2016), depending on which data (individual values or mean values) is used (■ Table 3). In test–retest analyses in the literature, both approaches can be found (Stanton, Wintour, & Kean, 2017), and test–retest reliability improves when mean values are used instead of individual values. Whether single or mean scores should be used to test reliability also depends, among other factors, on the research question. In the present study, reliability measurements were therefore performed with different types of data (individual and mean values) to describe the changes and to provide the user with a basis for decision-making. Overall, when interpreting the reliability parameters, it must also be considered that the jumping analyses did not take place on a solid surface but on a sand surface, which increases the variability in the jumping performance and, thus, influences the quality of the reliability.

Although the results of the individual measurements on the force plate and with the IMU system (■ Table 3) are considered good regarding reliability (Koo & Li, 2016), the confidence intervals ranges drop to 0.82 for the force plates and under 0.75 for the IMU system indicating only moderate reliability. This is some-

how surprising, since the force plates were considered as gold standard. During the measurements, we ensured that factors influencing variance were minimized (same surface, same investigators, same instructions, same time of the measurements, subjects did not perform any training between the two tests). Therefore, this slightly reduced reliability on sand compared to solid ground may be attributed to the changeability of the sand surface. Assuming this influence, one has to find strategies to cope with this increased variation to ensure meaningful statements about jumping performance of sand surfaces.

It is still debated whether to use the best or mean values for performance diagnostics or when monitoring changes over time (e.g. Al Haddad, Simpson, & Buchheit, 2015). However, on a sand surface, we believe that mean values instead of best values are more suitable to describe the performance level of athletes because the changeability and inconsistency of the sand surface increases variation in the data. This is shown by a large range of limits of agreement within Bland–Altman plots in the present study but also in the literature (Schmidt et al., 2021). To illustrate this influence, imagine a situation where one subject performs nearly two identical jumps with different landing locations (e.g. 10 cm apart). This is not of interest on a rigid surface, but on a sand surface subjects compress the sand under their feet for several centimetres during take-off and may land on a different level leading to a decreased flight time and thus a decreased jumping height. One strategy to reduce this effect is the exclusion of unrealistic values or using mean values like described above. The results clearly show that this strategy increases day-to-day reliability and concurrent validity to an excellent level. Another reason to use mean instead of best values in beach volleyball is that it is more important to jump high consistently instead particularly high once.

Concurrent validity

When focusing on concurrent validity of the IMU system in comparison to the

VERT™ system (Schmidt et al., 2021), both systems show a high concurrent validity in jump height when compared to a gold standard. Schmidt et al. (2021) compared the jump height results of the VERT™ system with the results of a marker-based motion analysis and reported a bias of 2.6 cm for the block jump and 7.7 cm for the spike jump between both systems. The limits of agreements were for both jump types around 15 cm. In the actual study, flight times derived from the force plates were compared to flight times determined using the acceleration signals of the IMUs. The results show a much smaller bias of −0.7 cm and a reduced but similar range of limit of agreements of 12.5 cm. Larger biases (3.6–4.3 cm) have also been reported indoors by Charlton et al. (2017), indicating that this is probably not a surface-based effect. It may be an effect of the evaluation method of the jumping height since this differs between the two mentioned studies and the actual investigation. We compared calculated flight time data within both systems, whereas Schmidt et al. (2021) and Charlton et al. (2017) compared flight-time-based acceleration data with motion analysis-based estimated jump height. The latter overestimates the jump height that is detected via flight time and increases the difference between jump heights thus increasing the bias between systems. In consensus with Schmidt et al. (2021), high range of limits of agreement of all trials is based on the varying results within some single subjects, indicating that strategies to reduce this variation (e.g. exclusion of unrealistic values, using mean values) may help to increase validity of the IMU system. It is also obvious that individual large variation in jump height is not related to performance level of the subjects as presented in ◘ Fig. 3.

Limitations

The IMU system can measure at different frequencies (52, 104, 208 and 416 Hz), but in pilot studies we found that measurement frequencies above 52 Hz led to increased transmission errors. For newer versions of IMUs and with the changeover to Bluetooth 5.0, it may be possible to increase sample frequency to 104 Hz or even 208 Hz without issues. In principle, this would allow more accurate detection of the jump and the landing, so that the differences in the jump height compared to the force plate can be reduced even further.

A more accurate detection of flight times is also possible when accelerometers are placed closer to the ground, as this can improve the quality of the detected signals with respect to the jump height. Yet, we highlight two factors to consider: First, extreme uncoordinated, sometimes single leg jumps, or landings lead to inaccurate jump height measurements because acceleration is only measured on one leg; second, different angles in the ankle joint at the time of landing affect the quality of the signals in sand and, thus, the jump height calculation. Applying an additional sensor to the second foot and changing the location of the sensors from the ankle joint to the dorsal surface of the foot would diminish both problems and improve the quality of the acceleration signals when landing. However, this would also increase the impact on the athletes.

The decreased concurrent validity on the first day compared to the second day could suggest that the subjects were not yet properly acclimated to the sand condition. As such, a longer habituation period should be provided in future investigations. Instead of using only two familiarization jumps as we did in this study, an additional two to three jumps are recommended.

Practical applications

Monitor training and control athlete load on a sand surface. Complex match-specific actions, such as sprints, fast changes of direction pose an extreme challenge when attempting to disentangle the signal from the noise of many different acceleration signals, which were measured within the current study. Thus, the high success rates our sensor system achieved under these conditions indicate that this jump detection system can already be used in practice, for example, to monitor training and control athlete load. It is also a time efficient alternative to video analyses to count jumps.

Suitable for standardized performance diagnostics on a sand surface. The system is highly suitable for performance diagnostics using CMJs under standardized conditions on a sand surface.

Freely available and adjustable system. The used IMU system is freely available and programmable. The big advantage of open systems compared to proprietary systems is that one can adapt the underlying algorithms and signal processing procedures to the demands of the sport, different performance level or individual techniques. It is also possible to use up to 5 different sensors within the system.

Conclusions

An inertial measurement unit (IMU)-based system for detection of jumps and jump height in beach volleyball on a sand surface was successfully developed. The device demonstrates excellent detection rates of jump types (block, attack, serve) and excellent concurrent validity for jump height detection (standardized countermovement jumps). Results can be used to for screening to prevent injury and in rehabilitation, for performance diagnostics and training control in the field, and to also periodize training and maximize balance between load and recovery. In future studies, it would be advantageous to evaluate the system with a greater number of athletes during training and within competition to determine the ecological validity of the IMU system. Here, including artificial intelligence seems advantageous to enable continuous automated improvements.

Corresponding address

Prof. Dr. Eric Eils
Institute of Sport and Exercise Sciences, University of Münster
Horstmarer Landweg 62b, 48149 Münster, Germany
eils@uni-muenster.de

Declarations

Conflict of interest. S. Schleitzer, S. Wirtz, R. Julian and E. Eils have no relevant financial or nonfinancial interests to disclose. The authors have no conflicts of interest to declare that are relevant to the content of this article. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or nonfinancial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

For this article, all persons gave their informed consent prior to their inclusion in the study. All human studies have been approved by the appropriate ethics committee and have therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

References

Al Haddad, H., Simpson, B. M., & Buchheit, M. (2015). Monitoring changes in jump and sprint performance: best or average values? *International Journal of Sports Physiology and Performance*, *10*(7), 931–934. https://doi.org/10.1123/ijspp.2014-0540.

Bahr, M. A., & Bahr, R. (2014). Jump frequency may contribute to risk of jumper's knee: a study of interindividual and sex differences in a total of 11,943 jumps video recorded during training and matches in young elite volleyball players. *British Journal of Sports Medicine*, *48*(17), 1322–1326. https://doi.org/10.1136/bjsports-2014-093593.

Bahr, R., & Reeser, J. C. (2003). Injuries among world-class professional beach volleyball players. The Fédération Internationale de Volleyball beach volleyball injury study. *The American Journal of Sports Medicine*, *31*(1), 119–125. https://doi.org/10.1177/03635465030310010401.

Benson, L. C., Tait, T. J., Befus, K., Choi, J., Hillson, C., Stilling, C., . . ., & Emery, C. A. (2020). Validation of a commercially available inertial measurement unit for recording jump load in youth basketball players. *Journal of Sports Sciences*, *38*(8), 928–936. https://doi.org/10.1080/02640414.2020.1737360.

Bland, M. J., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *327*(8476), 307–310. https://doi.org/10.1016/s0140-6736(86)90837-8.

Borges, T. O., Moreira, A., Bacchi, R., Finotti, R. L., Ramos, M., Lopes, C. R., & Aoki, M. S. (2017). Validation of the VERT wearable jump monitor device in elite youth volleyball players. *Biology of Sport*, *34*(3), 239–242. https://doi.org/10.5114/biolsport.2017.66000.

Borrmann, D., Breuer, G., Kummetat, U., Morbach, A., Münster, H., Mund, J., . . ., & Palmen, M. (2009). *BISp-Orientierungshilfe* (2nd edn.). Bundesinstitut für Sportwissenschaft.

Camomilla, V., Bergamini, E., Fantozzi, S., & Vannozzi, G. (2018). Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: a systematic review. *Sensors*. https://doi.org/10.3390/s18030873.

Charlton, P. C., Kenneally-Dabrowski, C., Sheppard, J., & Spratford, W. (2017). A simple method for quantifying jump loads in volleyball athletes. *Journal of Science and Medicine in Sport*, *20*(3), 241–245. https://doi.org/10.1016/j.jsams.2016.07.007.

Czimek, J. (2017). *Volleyball – Training & Coaching: Vom Jugend- zum Leistungsvolleyballer: kombinierte Rahmentrainingskonzeption Volleyball und Beach-Volleyball des Deutschen Volleyball-Verbandes*. Meyer & Meyer.

Gageler, H. W., Wearing, S., & James, A. D. (2015). Automatic jump detection method for athlete monitoring and performance in volleyball. *International Journal of Performance Analysis in Sport*, *15*(1), 284–296. https://doi.org/10.1080/24748668.2015.11868793.

Handball-World-com (2021). New hope for olympic dream: beach handball might make its appearance at the 2024 Olympic Games in Paris. https://www.handball-world.news/en/news-2-1-24-134099.html. Accessed 23 Oct 2021.

Jarning, J. M., Mok, K.-M., Hansen, B. H., & Bahr, R. (2015). Application of a tri-axial accelerometer to estimate jump frequency in volleyball. *Sports Biomechanics*, *14*(1), 95–105.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Künkler, A. (2009). *Erfassung leistungsbestimmender Merkmale des Sportspiels Beachvolleyball T4 – Analysen, Erkenntnisse und Strategieentwicklungen gegen die weltbesten Teams*. Dissertation. Duisburg-Essen: Universität Duisburg-Essen. https://d-nb.info/995237905/34

MacDonald, K., Bahr, R., Baltich, J., Whittaker, J. L., & Meeuwisse, W. H. (2017). Validation of an inertial measurement unit for the measurement of jump count and height. *Physical Therapy in Sport*, *25*, 15–19.

Peake, J. M., Kerr, G., & Sullivan, J. P. (2018). A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations. *Frontiers in Physiology*, *9*, 743. https://doi.org/10.3389/fphys.2018.00743.

Picerno, P., Iosa, M., D'Souza, C., Benedetti, M. G., Paolucci, S., & Morone, G. (2021). Wearable inertial sensors for human movement analysis: a five-year update. *Expert Review of Medical Devices*. https://doi.org/10.1080/17434440.2021.1988849.

Schmidt, M., Alt, T., Nolte, K., & Jaitner, T. (2020). Comment on "hurdle clearance detection and spatiotemporal analysis in 400 meters hurdles races using shoe-mounted magnetic and inertial sensor". *Sensors*, *20*(10), 2995. https://doi.org/10.3390/s20102995.

Schmidt, M., Meyer, E., & Jaitner, T. (2021). Quantifying jump-specific loads in beach volleyball by an inertial measurement device. *International Journal of Sports Science & Coaching*, *16*(2), 391–397.

Skazalski, C., Whiteley, R., Hansen, C., & Bahr, R. (2018). A valid and reliable method to measure jump-specific training and competition load in elite volleyball players. *Scandinavian Journal of Medicine & Science in Sports*. https://doi.org/10.1111/sms.13052.

Sprague, A. L., Smith, A. H., Knox, P., Pohlig, R. T., & Grävare Silbernagel, K. (2018). Modifiable risk factors for patellar tendinopathy in athletes: a systematic review and meta-analysis. *British Journal of Sports Medicine*, *52*(24), 1575–1585. https://doi.org/10.1136/bjsports-2017-099000.

Stanton, R., Wintour, S.-A., & Kean, C. O. (2017). Validity and intra-rater reliability of MyJump app on iPhone 6s in jump performance. *Journal of Science and Medicine in Sport*, *20*(5), 518–523. https://doi.org/10.1016/j.jsams.2016.09.016.