Validity and reliability of an inertial measurement unit to assess jump performance

Short Title – Comparison of IMU and force platform

Premraj Saini[1,2], Anugrah Suresh[2], Sandy Nair[3], Georgian Badicu[4], Rodrigo Ramirez-Campillo[5,6], Rohit K. Thapa[1]

[1] Symbiosis School of Sports Sciences, Symbiosis International (Deemed University), Pune, 412115, India.

[2] Department of Strength and Conditioning, Sports Authority of India-Lakshmibai National College of Physical Education, Trivandrum, 695581, India.

[3] Department of Physical Education, VTM NSS College, University of Kerala, Trivandrum, 695581, India.

[4] Department of Physical Education and Special Motricity, Transilvania University of Brasov, Brasov, 500068, Romania.

[5] Exercise and Rehabilitation Sciences Institute, Faculty of Rehabilitation Sciences, Universidad Andres Bello, Santiago, 7591538, Chile.

[6] Department of Physical Activity Sciences, Universidad de Los Lagos, Osorno, Chile.

Corresponding author:

Dr. Rohit K. Thapa, PhD

Symbiosis School of Sports Sciences

Symbiosis International (Deemed University)

Pune, 412115, India

Email: rohit.thapa@ssss.edu.in

Mobile: +91 9678324029

Abstract

The aim of this study was to investigate the validity and test-retest reliability of an IMU (*Output V2*) compared to a force plate (Kistler, Winterthur, Switzerland) for measuring countermovement jump (CMJ) and drop jump (DJ) performance metrics. In a randomized crossover study design, male athletes (n=17; age: 22.3±2.3 years) performed three CMJ and DJ maximal effort trials, with the force plate and the IMU concurrently recording jump performance data. The intraclass correlation coefficient (ICC [consistency]), Cronbach's alpha, coefficient of variation (CV), and standard error of measurement (SEM) were calculated to determine test-retest reliability. Validity (i.e., between-instruments comparison) was assessed using unpaired t-test, Pearson correlation, ICC (absolute agreement), concordance correlation coefficient (CCC), and Bland-Altman plots. Test-retest reliability of IMU was good to excellent for both tools (ICC lower limit ≥ 0.857, Cronbach's α ≥ 0.977, CV ≤ 2.3%, relative SEM ≤ 2.72%). Compared to the force plate, the IMU measured similar (between-instruments $p > 0.05$) CMJ height, DJ height, reactive strength index, and DJ contact time, with almost perfect Pearson correlations ($r \geq 0.983$), substantial CCC ($\rho_c \geq$ 0.952), and good to excellent ICC values (≥ 0.860). Bland–Altman plots showed negligible bias and strong agreement. The IMU *Output V2* was found to be valid and reliable when compared to a force plate for measuring CMJ and DJ performance metrics. The *Output V2* IMU's cost-effectiveness, portability, and ease of use via mobile devices make it a practical alternative for field assessments.

Keywords: plyometric exercise, athletic performance, sports technology, wearable sensors, stretch-shortening cycle, micro-electromechanical systems, IMU.

1. INTRODUCTION

Jump performance can impact athletic output in various sports, such as basketball, volleyball, and football.[1, 2] Traditionally, force plates are considered the gold standard technology for quantifying jump-related performance measures (e.g., force, impulse).[3] However, force plates can be expensive and not portable (e.g., plates embedded in the ground). They may face other logistical issues (e.g., highly-trained personnel needed to use the equipment).[4] Nonetheless, less expensive, portable, user-friendly measurement systems have been developed and validated (e.g., jump mats, laser-operated devices, motion capture systems, mobile applications), although these systems are limited in their ability to obtain real-time data during competition or training.[5]

However, more accessible technologies are available to obtain jump-related performance measures, such as inertial measurement units (IMU).[6] An IMU is a compact device that combines an accelerometer, gyroscope, and magnetometer to capture motion dynamics in real time, making it suitable for applications in biomechanics and sports performance analysis.[7, 8] The ability of IMUs to operate outside laboratory settings allows for data collection essentially anywhere, which is crucial for understanding athletic performance in real-world conditions.[9]

Custom-designed IMUs (i.e., not commercially available) have been investigated. For example, Keskinoğlu et al.[5] developed an IMU with the jump meter on an athlete's shorts to assess jump metrics using raw acceleration signals and demonstrated high accuracy compared to a jump mat. Similarly, Schleitzer et al.[10] developed a dual sensor IMU (mounted on the participant´s sternum and ankle) to detect jump take-off and landing events for estimating jump height using flight time. They demonstrated high reliability compared to force platforms.

Regarding commercially available IMUs, Comyns et al.[4] reported on the reliability and validity of the *Output V1* Sensor (Output Sports, Dublin, Ireland) for measuring countermovement CMJ-related performance metrics. However, Montoro-Bombú et al. [11] reported a poor reliability for the *Output V1* Sensor in measuring drop jump (DJ) related performance metrics (e.g., ground contact time, reactive strength index [RSI], jump height). An updated version of the *Output V1* Sensor was commercially released, the *Output V2* Sensor (Output Sports, Dublin, Ireland). Perrotta et al.[12] reported high test-retest reliability and a high level of agreement between the *Output V2* Sensor IMU and force plates (concurrent validity) to measure CMJ height in female athletes. However, considering the poor reliability of the *Output V1* Sensor in measuring DJ-related performance metrics,[11] the usability of the *Output V2* Sensor IMU to assess jump-related performance metrics derived from jump tests other than the CMJ is yet to be investigated. The lack of studies validating commercially available IMUs may be considered a critical issue requiring research attention, because it is important for athletes to be assessed using different jump tests (e.g., CMJ; DJ), as these can involve different biological and/or biomechanical factors related to athletic performance (e.g., muscle power; muscle-tendon stretch-shortening cycle [SSC]).[13,14]

Indeed, CMJ is associated with sprinting,[15] maximal isometric strength,[16] change of direction performance (overall),[17] slow SSC (e.g., jumps involving foot-ground contact times >250 ms), profiling performance,[18] neuromuscular fatigue,[19] and assisting during the return to sport after injuries.[20] Whereas the DJ is associated with maximal dynamic strength, change of direction performance (deceleration-acceleration transitioning), endurance performance,[21] fast SSC (e.g., jumps involving foot-ground contact times <250 ms),[22, 23] and with the efficiency of the transitioning from eccentric to concentric muscle-tendon actions, usually taking into consideration the ratio between the foot-ground contact time and the flight time of the jump (i.e., reactive strength index [RSI]).[24]

The study that assessed the *Output V2* Sensor IMU included female participants.[12] The reliability and validity of IMU systems might vary between males and females.[25] Therefore, studies on males are required to reduce the uncertainty on the *Output V2* Sensor IMU to assess jump-related performance metrics in athletes (males and females). Further, although the CMJ-related performance results obtained with the *Output V2* Sensor IMU were compared against the data obtained with the *gold* standard (i.e., force plate), the statistical approach comparing results between measurement devices (i.e., mean differences, effect sizes, Pearson's *r*, Bland–Altman plots) may benefit from a complementary statistical perspective,[12] such as Lin's concordance correlation coefficient[26] and intraclass correlation coefficients (two-way random-effects, single measures, absolute agreement model), particularly based on the ICC lowest 95% CI.[27]

Therefore, this study aimed to investigate the validity and test-retest reliability of an IMU (*Output V2*) compared to a force plate (Kistler, Winterthur, Switzerland) for measuring both CMJ and DJ performance metrics in collegiate athletes.

2. MATERIAL & METHODS

2.1 Participants

The sample size estimation was performed using G*Power software 3.1.9.7. (University of Dusseldorf, Dusseldorf, Germany). For the correlation bivariate model, a two-tailed test with ρ H_1 = 0.982 (correlation obtained in a previous study[4] for CMJ height) and ρ H_0 = 0.90 (assumed null hypothesis correlation), with a significance level of $\alpha = 0.05$, the required sample size was determined to be $n = 13$ participants. Male college athletes ($n = 17$; age: 22.3 ± 2.3 years, height: 169.7 ± 8.0 cm, body mass: 64.8 ± 7.1 kg) from football (n = 6), basketball (n = 3), volleyball (n =3), athletics (n = 3), badminton (n = 1), and hockey (n = 1) volunteered to participate in this study. The inclusion criteria were as follows: (i) athletes who participated in inter-college competitions; (ii) a minimum of one year of formal training experience in their respective sports; (iii) active participation in their respective sport for at least five hours per week; (iv) free from lower limb injuries in the past six months that could restrict the participants from performing the jumps. Prior to recruitment, the participants were informed about the procedures of the study, associated benefits, and risks. Thereafter, written informed consent was obtained from each participant. The study was approved by the Sports Authority of India – Lakshmibai National College of Physical Education and was conducted according to the updated version of the Declaration of Helsinki.

2.2 Instrumentation

2.2.1 Force Platform

The criterion reference used in the study to assess jump performance was the Kistler 9286BA (600 mm × 400 mm × 35 mm; mass: 17.5 kg) force platform (Kistler, Winterthur, Switzerland). The force platform transmitted the raw data to a data acquisition system (DAQ, Type 5691A1, Kistler) using a connection cable (Type 1759A, Kistler). Thereafter, the data acquisition system converted the analog signals into digital data at a sampling rate of 1,000 Hz, which was connected to a laptop through a USB 2.0 cable where the digital data was automatically analyzed using Kistler BioWare software (version 5.3.0.7). The force platform has an excellent center of pressure accuracy, with maximum errors of around 2 mm in the x/AP and y/ML directions. The plate offers a wide measuring range (0-10 kN) and a high natural frequency (>400 Hz).[28]

2.2.2 Inertial Measurement Unit Device

The *Output V2* sensor (Output Sports, Dublin, Ireland) (50 mm × 33 mm × 15 mm; mass: 21 g) was worn on the left shoe as per the directions of the manufacturer.[11] The IMU device strap was readjusted manually after each maximal jump trial to ensure accurate measurements. The IMU device integrated an accelerometer (± 2g-16, 16 bit), a magnetometer (± 1300 Mt [x, y axis], ± 2500 µT [z axis]), and a gyroscope (± 125-2000 dps, 16 bit).[29] The IMU device recorded the data with a 1,000 Hz sampling frequency and provided real-time data to the Output Capture Application (version 2.7.2) that was installed on an Android device and connected via Bluetooth (version 5.2) with the IMU. The data of all participants were recorded in the application and later exported to the computer in a .csv file.

2.3 Procedure

2.3.1 Familiarization Session

Prior to the data collection, all participants underwent two familiarization sessions of the testing protocols for both the CMJ and DJ tests. During the first familiarization session, the techniques of CMJ and DJ were demonstrated and explained to all participants by a certified strength and conditioning coach. The participants were also made familiar with the warm-up protocol to be used during the study. The feedback on the correct technique (e.g., minimizing ground contact time during DJ) for both jumps was also provided. During the second familiarization session, both jumps were practiced, and minor corrections were made. No additional familiarization sessions were conducted as the participants were able to perform the jumps with the correct form and technique. The height and body mass of the participants were also collected during these sessions using a Seca 284 (Seca, Hamburg, Germany) scale.

2.3.2 Warm-up

Before the testing session, the participants performed a 10-minute standard RAMP (i.e., raise, activate, mobilize, and potentiate) warm-up protocol,[30] which consisted of four minutes of light to moderate intensity jogging on a synthetic athletic track, followed by one minute of static stretching, three minutes of dynamic movements through the full range of motion (walking on toes, heels, knee to chest, hip internal & external rotation, hip hinge, lunge & rotate, deep lunge, body weight squats), and two-minute low-intensity jumps (standing pogo, CMJ, DJ).

2.3.3 Testing Protocol

The participants were asked to refrain from any extraneous physical activity 24 hours before the testing session. Data collection occurred on a single day for all participants under a controlled laboratory environment (temperature: 25 °C; time: 9 AM – 2 PM). Each participant performed three maximal effort trials for the CMJ and the DJ in a randomized crossover manner (using the research randomization tool available at www.randomizer.org), with a rest interval of one minute between trials.

2.3.3.1 Countermovement Jump

Participants were instructed to stand with their feet shoulder-width apart on the force platform. The participants were instructed to keep their hands on their waist and were not allowed to move them to minimize the influence of the arm swing. Upon receiving a verbal cue, the participants performed a downward movement (eccentric contraction) followed by an upward movement (concentric contraction), leading to a vertical jump. No flexion was allowed during the flight. If flexion was observed, the trial was rejected, and a new trial was performed. Vertical jump height was concurrently recorded using both the IMU and the force platform.

2.3.3.2 Drop Jump

Participants performed a DJ from a standard height (30 cm) and were instructed to keep their hands on their waist and to jump as high as possible upon landing. The participants were also verbally instructed to minimize the ground contact time on the first landing prior to the vertical jump. Participants were asked to drop using only the left leg to make the protocol consistent. The participants were not allowed to flex their knees during the flight phase of the jump. Drop jump height, RSI, and contact time (CT) were concurrently recorded using both the IMU and the force platform. For the IMU and the force plate, the RSI was calculated as jump height (m) ÷ contact time (s).[24]

2.4 Statistical Analysis

The descriptive statistics of variables are reported as means and standard deviations. The normality of data was verified using the Shapiro-Wilk test. The three maximal effort trials for the CMJ and the DJ were used for reliability and validity analyses. Test-retest reliability was assessed using the intraclass correlation coefficient (ICC) using a two-way random-effects, single measures, consistency model.[31] This assessment corresponds to the ICC (C,1) model according to McGraw and Wong [32] and uses the formula

$$\text{ICC} = \frac{MS_R - MS_E}{MS_R + (k-1)\, MS_E} \quad (1)$$

where MS_R– mean square for rows, MS_E – mean square for error, k – number of trials)

for the ICC (3,1) model as described by Shrout and Fleiss [33]. The interpretation was based on the lower bound of the 95% confidence interval as poor (<0.5), moderate (0.5–0.75), good (>0.75–0.9), or excellent (>0.9).[27] The ICC's consistency model was used to assess the trial-to-trial reliability of jumps, evaluating the degree to which repeated trials ranked participants consistently while allowing for potential systematic differences in absolute scores between trials. Internal consistency of repeated trials was further evaluated using Cronbach's α

$$\text{ICC} = (\frac{k}{k-1}(1 - \frac{\sum_{i=1}^{k} \sigma_{Y_i}^2}{\sigma_X^2}) \quad (2)$$

where k – number of items, $\sigma_{Y_i}^2$ – variance of item i, σ_X^2 – variance of the total scores across all items), indicating the extent to which the trials measured the same underlying performance construct, with values ≥0.70 considered acceptable. Measurement error was quantified via the coefficient of variation (CV), calculated as $(\frac{SD}{Mean})\times 100$, where values <10% were deemed acceptable.[34] ICC assumes homoscedasticity (i.e., constant variance across the measurement range) and can be inflated by between-subject variability. Therefore, it was supplemented with the standard error of measurement (SEM), which is independent of between-subject variability.[35] It was derived from repeated measures ANOVA as $\sqrt{MS_{error}}$ (mean square error) for absolute SEM and $(\frac{SEM}{Mean})\times 100$ for relative SEM for relative SEM.[31]

Absolute agreement between measurement systems was examined using a two-way random-effects, single measures, absolute agreement ICC model.[31] It corresponds to ICC (A,1) notation by McGraw and Wong [32] and ICC (2,1)

$$\text{ICC} = \frac{MS_R - MS_E}{MS_R + (k-1)\, MS_E + \frac{k\,(MS_C - MS_E)}{n}}, \quad (3)$$

where MS_R– mean square for rows,MS_C – mean square for columns, MS_E – mean square for error, k – number of trials, n – number of participants) according to Shrout and Fleiss [33] with the same interpretation thresholds as reliability. The ICC's absolute agreement model was used to assess the agreement between the two measurement devices, evaluating both the consistency of rankings and the closeness of the actual measurement values. Additionally, the concordance correlation coefficient (CCC) was computed to assess reproducibility, following Lin's method,[26] with the formula:

$$\text{CCC} = \left(\frac{2\rho\sigma_x\,\sigma_y}{\sigma_x^2 + \sigma_y^2 + (\mu_x - \mu_y)^2}\right), \quad (4)$$

where ρ is the Pearson correlation coefficient, σ represents standard deviations, and μ denotes means, where ρ is the Pearson correlation coefficient, σ represents standard deviations, and μ denotes means. The CCC was interpreted based on the lower bound of the 95% confidence interval as poor (<0.90), moderate (0.90–0.95), substantial (>0.95–0.99), or almost perfect (>0.99).[36] Systematic bias between instruments was evaluated using unpaired t-tests supplemented by Hedge's g effect size (ES), categorized as trivial (<0.2), small (0.2–0.6), moderate (>0.6–1.2), or large (>1.2–2.0).[37] Pearson's correlation coefficient (r, $\frac{\sum_{i=1}^{n}(x_i - \bar{x})\,(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$) was also calculated to assess the strength and direction of the linear relationship between the measurements obtained from the two measurement devices, with magnitudes classified as trivial (≤0.1), low (>0.1–0.3), moderate (>0.3–0.5), high (>0.5–0.7), very high (>0.7–0.9), or almost perfect (>0.9–1.0).[37] Bland-Altman plots were generated to illustrate agreement between the force plate and the IMU.[38] All statistical analyses were performed using SPSS 23.0.0 (IBM, New York, USA). The

CV, SEM, and ES were computed in Microsoft Excel (Microsoft, California, USA). The significance level was established at $p \leq 0.05$.

3. RESULTS

Table 1 presents the descriptive statistics, comparative analysis, ICC, and Pearson correlation values for the jump performance variables assessed using the force platform and the IMU *Output V2*. For the jump performance metrics CMJ height, DJ height, DJ RSI, and DJ CT, no difference was found between the values obtained with the force platform compared to the IMU *Output V2* ($p \geq 0.441$, ES = 0.08 to 0.15 [trivial]), with good to excellent ICC (0.860 to 0.957), nearly perfect correlation ($r \geq 0.983$; $p < 0.001$) and determination coefficient r^2 (Figure 1).

*****Insert Table 1 about here*****

*****Insert Figure 1 about here*****

The force platform demonstrated excellent reliability across all jump metrics (Table 2), with ICC ranging from 0.912 to 0.954, Cronbach's alpha values ≥0.986, and acceptable CV values (<1.8%). Similarly, the IMU *Output V2* exhibited good to excellent reliability for all jump metrics, with ICC values ranging from 0.857 to 0.923, Cronbach's alpha values ≥0.977, and acceptable CV values (<2.3%). The low absolute and relative SEM demonstrated by the IMU *Output V2* indicated high test-retest reliability across all measures (Table 2).

*****Insert Table 2 about here*****

The Bland–Altman plots (Figure 2) indicated minimal bias and narrow limits of agreement between the force platform and the IMU *Output V2* for all four jump performance measures. The plots demonstrated significant agreement between the devices, with average differences falling within the 95% limits of agreement (i.e., average difference ± 1.96 standard deviations of the difference). The distribution of scores in all plots was centered around the bias line, with no evidence of proportional bias. Furthermore, Lin's CCC showed substantial agreement between measures of force platform and the IMU for all metrics: CMJ height (ρ_c = 0.978, 95% CI: 0.963–0.987), DJ height (ρ_c = 0.990, 95% CI: 0.982–0.994), DJ RSI (ρ_c = 0.971, 95% CI: 0.952–0.983), and DJ CT (ρ_c = 0.985, 95% CI: 0.974–0.991).

*****Insert Figure 2 about here*****

4. DISCUSSION

The IMU *Output V2* is valid and reliable when compared to a force plate (i.e., gold standard) for measuring CMJ and DJ performance metrics. Indeed, the IMU *Output V2* allows measures of jump performance with an acceptable (i.e., *low*) coefficient of variation. Additionally, the CMJ height, DJ height, DJ RSI, and DJ CT measures were similar using either the force plate or the IMU *Output V2* (all comparisons p>0.05, trivial ES), with no proportional bias observed in the Bland-Altman plots. These findings highlight the validity and reliability of the IMU *Output V2* to measure jump performance, comparable to a force plate, although with the advantage of the IMU cost-effectiveness, portability, and ease of use via mobile devices, making the IMU *Output V2* a practical alternative for jump performance field assessments. The innovation of the present study lies in its methodological design and scope. Unlike previous studies that focused solely on CMJ[4,

[12] or reported inconsistent findings in DJ assessment using older IMU models,[11] the current study concurrently assessed both CMJ and DJ metrics using the updated *Output V2* IMU against a gold-standard force platform. This study is the first study, to the author's knowledge, to validate and test the reliability of *Output V2* IMU across both jump types in collegiate athletes.

The test-retest reliability results of CMJ height of the present study (ICC 0.97, 95%CI 0.92-0.99; ≤1.1% CV) align with thresholds suggested to achieve robust reliability (e.g., ICC ≥0.8; CV ≤10.0%).[39] Further, current results are consistent with those from earlier investigations by Comyns et al.[4] and Perrotta et al.[12] Comyns et al.[4] reported an ICC of 0.98 (95%CI 0.96-0.99), and ≤3.8% CV, whereas Perrotta et al. [12] reported ICC of 0.85 (95% CI 0.79-0.89) and 5.4% CV. Additionally, the test-retest reliability of the Kistler force platform is in line with a previous study by Barefoot. [40] Furthermore, when the CMJ height measured with the IMU *Output V2* was compared against the force platform, an $r = 0.98$ and $r^2 = 0.97$ were obtained, which is in line with previous studies by Comyns et al.[4] that reported $r = 0.98$ and $r^2 = 0.96$, and by Perrotta et al.[12] that reported $r = 0.87$ and $r^2 = 0.76$.

A previous study by Montoro-Bombú et al.[11] reported poor test-retest ICC for DJ RSI (95% CI lower bound = 0.151) using a previous version of IMU used in the present study,[11] whereas the results of the present study obtained good test-retest ICC (95% CI lower bound = 0.86). Furthermore, Montoro-Bombú et al.[11] reported a CV ≤32.2% for the IMU, which is unacceptable for test-retest reliability.[34] Of note, in the Montoro-Bombú et al. study,[11] the CV for the force platform was ≤38.8%, suggesting that the participants' performances varied significantly across trials, probably due to uncontrolled methodological issues (e.g., sub-optimal warm-up; data mixed from participants of different sexes and athletic levels). Additionally, the Bombú et al. study[11] reported poor test-retest ICC for CT (95% CI lower bound = 0.291) and moderate ICC for DJ height (95% CI lower bound = 0.773) using the previous version of IMU (i.e., *Output V1*). These results contrast with the present study's findings, which showed an excellent ICC for CT (95% CI lower bound = 0.957) and DJ height (95% CI lower bound = 0.957). Moreover, CV ≤21.1% and ≤25.6% were reported for DJ CT and DJ height using the IMU *Output V1*,[11] whereas the present study reported CV values of ≤1.6% for both DJ CT and DJ height. The potential reason for the improved reliability in the DJ outcomes may be the update in the hardware version in the present study from the study by Montoro-Bombú et al.[11] Montoro-Bombú et al.[11] used the older version of the IMU used in the present study (i.e., *Output V1*), whereas the present study incorporated the IMU *Output V2*, which includes several updates (e.g., reduced mass and volume; magnetometer with a refined range, potentially contributing to improved signal stability and power efficiency; Bluetooth 5.2). Although specific algorithmic updates were not disclosed by the manufacturer, these updates appear to improve signal quality, enhance event detection and metric computation, and integrate data from tri-axial accelerometers, gyroscopes, and magnetometers by applying machine learning models trained on larger, more diverse datasets.

The field of sports science is continuously evolving, and so is the technology used in sports science equipment. However, despite this rapid growth, it remains unclear whether these devices meet the criteria for reliability and validity. A review article by Peake et al.[41] reported that only 5% of the devices have been formally validated and tested in real-world settings. Therefore, it is essential for researchers to validate these sports science tools in various environments to enhance the accuracy of measurements.[10] Of note, selecting the appropriate statistical treatment to evaluate the reliability and validity of the equipment is critical. Reporting only mean differences and correlations is insufficient; a thorough analysis is needed to reveal consistency,[34] agreement,[38] reproducibility,[26] and measurement error.[31] Inadequate statistical treatment can lead to overlooking bias, overestimating reliability,[27, 39] or misrepresenting a device's practical usefulness. ICC is often used as a measure of reliability;[31] however, its value is highly dependent on between-subjects variability; a large ICC can mask poor trial-to-trial consistency if the sample is very heterogeneous, while a low ICC can occur despite small measurement error if the sample is very homogeneous.[31, 35] In contrast, the SEM provides an absolute index of reliability, quantifying the expected trial-to-trial noise in the data in the original units of measurement. Because the SEM is not affected by between-subjects variability, its inclusion alongside the ICC provides a more complete and stable estimate of measurement error, offering practitioners insight into the precision of an individual score.[31, 35] Proper methods ensure that differences or agreements reflect true performance rather than chance, which is vital for informed decisions in athletic monitoring and performance assessment.

The practical implications of this research are greatly significant for coaches and practitioners who are looking for effective and accurate methods to evaluate athletic performance. The IMU *Output V2* portability, affordability, validity, and reliability make it a viable alternative to conventional, laboratory-based in-ground force platforms. These attributes allow practitioners to perform frequent assessments (pre-, mid-, and post-training) in field conditions without the need for costly or cumbersome equipment.[42] This portability means that coaches can seamlessly incorporate jump performance testing into regular training routines, providing athletes with real-time feedback. This ability can assist in customizing training programs to meet individual athletes' requirements, tracking progress, and adjusting training loads to enhance performance. Additionally, by consistently monitoring neuromuscular fatigue and the risk of injury, practitioners can mitigate the chances of overtraining or injury, thereby improving both athlete longevity and performance.[43] The capability to conduct quick and accurate evaluations using the IMU *Output V2* also may empower coaches and trainers to make data-driven choices in competitive situations, where prompt insights can result in performance improvements and strategic benefits. The user-friendliness and accuracy of the IMU present a valuable resource for both routine athlete monitoring and high-performance decision-making. Further, a post-hoc power analysis in G*Power software 3.1.9.7. (University of Dusseldorf, Dusseldorf, Germany), with a two-tailed test, $\rho\ H_1 = 0.983$ (correlation obtained in the present study's results for CMJ height), $\rho\ H_0 = 0.92$ (null hypothesis correlation assumed), $\alpha = 0.05$, and total $n = 17$, indicated an 85.3% probability to correctly detecting a true correlation between measurement instruments. However, the number of participants per sport in the present study varied between 1 and 6. Therefore, current findings should not be extrapolated to athletes from specific sports, particularly for athletes with different key features (e.g., age, sex, competitive level). Future research is advised to confirm current findings, with larger samples of participants.

5. CONCLUSION

The IMU *Output V2* exhibited excellent test-retest reliability with the data being comparable to the force platform to measure CMJ and DJ metrics, with no significant differences between instruments. Additionally, the high correlation values indicated that the IMU *Output V2* is a reliable and valid instrument to measure CMJ and DJ performance. These results suggest that the IMU *Output V2* can be an alternative to traditional laboratory-based equipment, which is costeffective, portable and easily operable using a mobile device via Bluetooth for field-based assessment of CMJ and DJ performance. However, future research should explore the application of this approach across different athletic populations and various performance metrics.

DECLARATIONS

Acknowledgement: The authors would like to thank the participants who volunteered for the study.

Author contribution: P.S., A.S., and R.K.T. conceived and designed the study. P.S., A.S., and S.N. were involved in the data collection process. P.S. and R.K.T. conducted the formal analysis of the data. P.S., A.S., S.N., G.B., R.R.C., and R.K.T. wrote/revised the final draft of the manuscript. All authors approved the final draft of the manuscript for publication.

Funding: This research received no external funding.

Data availability: All data generated or analyzed during this study are included in the published article as Table(s) and Figure(s). Any other data requirement can be directed to the corresponding author upon reasonable request.

Ethical approval and consent to participate: The study was approved by the Internal Review Board of Sport Authority of India – Lakshmibai National College of Physical Education with approval number SAI-LNCPE/IRB/DSC/2023-24/02. Prior to recruitment, the participants were informed about the procedures of the study, associated benefits, and risks. Thereafter, written informed consent was obtained from each participant.

Consent for publication: Not applicable

Competing interest: The authors declare no competing interests.

REFERENCES

1. Liu L, Xi L, Yongshun W, et al. More Jump more health: Vertical jumping learning of chinese children and health promotion. *Front Psychiatry* 2022; 13.

2. Karatrantou K, Gerodimos V, Voutselas V, et al. Can sport-specific training affect vertical jumping ability during puberty? *Biol Sport* 2019; 36: 217-224.

3. McMahon J, Suchomel T, Lake J, et al. Understanding the key phases of the countermovement jump force-time curve. *Strength Cond J* 2018; 40: 96-106.

4. Comyns T, Murphy J and O'Leary D. Reliability, usefulness, and validity of field-based vertical jump measuring devices. *J Strength Cond Res* 2023; 37: 1594-1599.

5. Keskinoğlu C, Özgünen KT and Aydın A. Designing and implementation of IMU-based wearable real-time jump meter for vertical jump height measurement. *Trans Inst Meas Control* 2023; 45: 2648-2657.

6. Kumar S, Singh J, Pradhan P, et al. Validity and reliability of an inertial measurement unit (BTS G-Walk) for measurement of countermovement jump height: A pilot-study. *J Anthropol Sport Phys Educ* 2023; 7: 19-22.

7. Adesida Y, Papi E and McGregor AH. Exploring the role of wearable technology in sport kinematics and kinetics: A systematic review. *Sensors* 2019; 19: 1597.

8. Liang W, Wang F, Fan A, et al. Extended application of inertial measurement units in biomechanics: from activity recognition to force estimation. *Sensors* 2023; 23: 4229.

9. Ekdahl M, Loewen A, Erdman A, et al. Inertial measurement unit sensor-to-segment calibration comparison for sport-specific motion analysis. *Sensors* 2023; 23: 7987.

10. Schleitzer S, Wirtz S, Julian R, et al. Development and evaluation of an inertial measurement unit (IMU) system for jump detection and jump height estimation in beach volleyball. *Ger J Exerc Sport Res* 2022; 52: 228–236.

11. Montoro-Bombú R, Field A, Santos AC, et al. Validity and reliability of the Output sport device for assessing drop jump performance. *Front Bioeng Biotechnol* 2022; 10.

12. Perrotta AS, Day BD, Wafai I, et al. Concurrent validity and reliability of photoelectric and accelerometer technology for calculating vertical jump height in female athletes. *Sports Engineering* 2023; 26: 48.

13. Xu J, Turner A, Comfort P, et al. A Systematic review of the different calculation methods for measuring jump height during the countermovement and drop jump tests. *Sports Med* 2023; 53: 1055-1072.

14. Klavora P. Vertical-jump tests: A critical review. *Strength Cond J* 2000; 22: 70-75.

15. Dietze-Hermosa M, Montalvo S, Gonzalez MP, et al. Association and Predictive ability of jump performance with sprint profile of collegiate track and field athletes. *Sports Biomech* 2024; 23: 2137-2156.

16. Muehlbauer T, Gollhofer A and Granacher U. Association of balance, strength, and power measures in young adults. *J Strength Cond Res* 2013; 27: 582-589.

17. Ranisavljev I, Matić M and Jankovic N. The relationship between maximal strength, vertical jump, acceleration and change of direction performance. *Facta Universitatis, Series: Physical Education and Sport* 2020; 17: 591-599.

18. Mitchell A, Holding C and Greig M. Factors influencing optimum countermovement jump performance and movement strategy in Championship professional football players: implications for player profiling. *Res Sports Med* 2022; 30: 30-40.

19. Claudino JG, Cronin J, Mezêncio B, et al. The countermovement jump to monitor neuromuscular status: A meta-analysis. *J Sci Med Sport* 2017; 20: 397-402.

20. Bishop C, Jordan M, Torres-Ronda L, et al. Selecting metrics that matter: comparing the use of the countermovement jump for performance profiling, neuromuscular fatigue monitoring, and injury rehabilitation testing. *Strength Cond J* 2023; 45: 545-553.

21. Jarvis P, Turner A, Read P, et al. Reactive strength index and its associations to measures of physical and sports performance: A systematic review with meta-analysis. *Sports Med* 2022; 52: 301-330.

22. Thapa RK, Sarmah B, Chaware U, et al. Fast and slow jump training methods induced similar improvements in measures of physical fitness in young females. *Women Sport Phys Act J* 2024; 32: wspaj.2023-0071.

23. Thapa RK, Chaware U, Sarmah B, et al. The effects of single and combined jump exercises utilizing fast and slow stretch-shortening cycle on physical fitness measures in healthy adult males: A randomized controlled trial. *Monten J Sports Sci Med* 2024; 13: 65-74.

24. Ramirez-Campillo R, Thapa RK, Afonso J, et al. Effects of plyometric jump training on the reactive strength index in healthy individuals across the lifespan: A systematic review with meta-analysis. *Sports Med* 2023; 53: 1029-1053.

25. Miqueleiz U, Aguado-Jimenez R, Lecumberri P, et al. Reliability of Xsens inertial measurement unit in measuring trunk accelerations: a sex-based differences study during incremental treadmill running. *Front Sports Act Living* 2024; 6: 1357353.

26. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989; 45: 255-268.

27. Koo TK and Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; 15: 155-163.

28. Multicomponent Force Plate Type 9286B https://kistler.cdn.celum.cloud/SAPCommerce_Download_original/000-713e.pdf (2008).

29. Output V2 Sensor Specifications https://www.outputsports.com/performance/product#howitworks.

30. Jeffreys I. Warm-up revisited: The ramp method of optimizing warm-ups. Professional Strength and Conditioning. (6) 12-18. *Professional Strength and Conditioning* 2007: 12-18.

31. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005; 19: 231-240.

32. McGraw KO and Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1996; 1: 30-46.

33. Shrout PE and Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86: 420-428.

34. Cormack SJ, Newton RU, McGuigan MR, et al. Reliability of measures obtained during single and repeated countermovement jumps. *Int J Sports Physiol Perform* 2008; 3: 131-144.

35. Claudino JG, Cardoso Filho CA, Boullosa D, et al. The role of veracity on the load monitoring of professional soccer players: A systematic review in the face of the big data era. *Appl Sci* 2021; 11: 6479.

36. McBride GB. *A proposal for strength-of-agreement criteria for Lin's Concordance Correlation Coefficient*. 2005.

37. Hopkins WG, Marshall SW, Batterham AM, et al. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc* 2009; 41: 3-13.

38. Bland JM and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307-310.

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

40. Barefoot M, Lamont H and Smith JC. Comparison of the reliability of four different movement thresholds when evaluating vertical jump performance. *Sports* 2022; 10: 193.

41. Peake JM, Kerr G and Sullivan JP. A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations. *Front Physiol* 2018; 9: 743.

42. Marković S, Dopsaj M, Tomazic S, et al. Can IMU provide an accurate vertical jump height estimate? *Appl Sci* 2021; 11: 12025.

43. Alba-Jiménez C, Moreno-Doutres D and Peña J. Trends assessing neuromuscular fatigue in team sports: A narrative review. *Sports* 2022; 10: 33.

Table 1. Comparison of force platform with IMU to assess jump performance.

Variable	Force platform	IMU *Output V2*	Unpaired t-test	MD ± SD	ES	*r*	ICC
	Mean ± SD		*p-value*		Hedge's *g*		(95%CI)
CMJ (cm)	35.42 ± 2.36	35.17 ± 2.35	0.606	0.24 ± 0.42	0.10	0.983	0.979 (0.945-0.990)
DJ (cm)	31.16 ± 2.67	30.93 ± 2.65	0.662	0.23 ± 0.31	0.08	0.993	0.990 (0.957-0.996)
DJ RSI (m/s)	1.161 ± 0.114	1.143 ± 0.114	0.441	0.017 ± 0.02	0.15	0.984	0.973 (0.860-0.990)
DJ CT (s)	0.269 ± 0.023	0.272 ± 0.024	0.646	0.002 ± 0.003	-0.12	0.989	0.985 (0.957-0.993)

CI – confidence interval, CMJ – countermovement jump, CT – contact time, DJ – drop jump, ES – effect size, ICC – intraclass correlation coefficient, IMU – inertial measurement unit, MD – mean difference, *r* – Pearson correlation, RSI – reactive strength index, SD – standard deviation.

Table 2. Reliability statistics.

Variable	Force platform					IMU *Output V2*				
	ICC (95% CI)	α	CV (%)	Absolute SEM	Relative SEM (%)	ICC (95% CI)	α	CV (%)	Absolute SEM	Relative SEM (%)
CMJ (cm)	0.979 (0.954-0.992	0.993	0.797	0.35	0.99	0.965 (0.923-0.986)	0.988	1.045	0.45	1.28
DJ (cm)	0.967 (0.928-0.987)	0.989	1.277	0.49	1.58	0.957 (0.906-0.983)	0.985	1.552	0.56	1.81
DJ RSI (m/s)	0.959 (0.912-0.984)	0.986	1.731	0.03	2.72	0.933 (0.857-973)	0.977	2.275	0.03	2.62
DJ CT (s)	0.967 (0.927-0.987)	0.989	1.458	0.004	1.60	0.950 (0.892-980)	0.983	1.585	0.005	1.97

α – Cronbach's alpha, CI – confidence interval, CMJ – countermovement jump, CT – contact time, CV – coefficient of variation, DJ – drop jump, ICC – intraclass correlation, IMU – inertial measurement unit, RSI – reactive strength index, SEM – standard error of measurement.

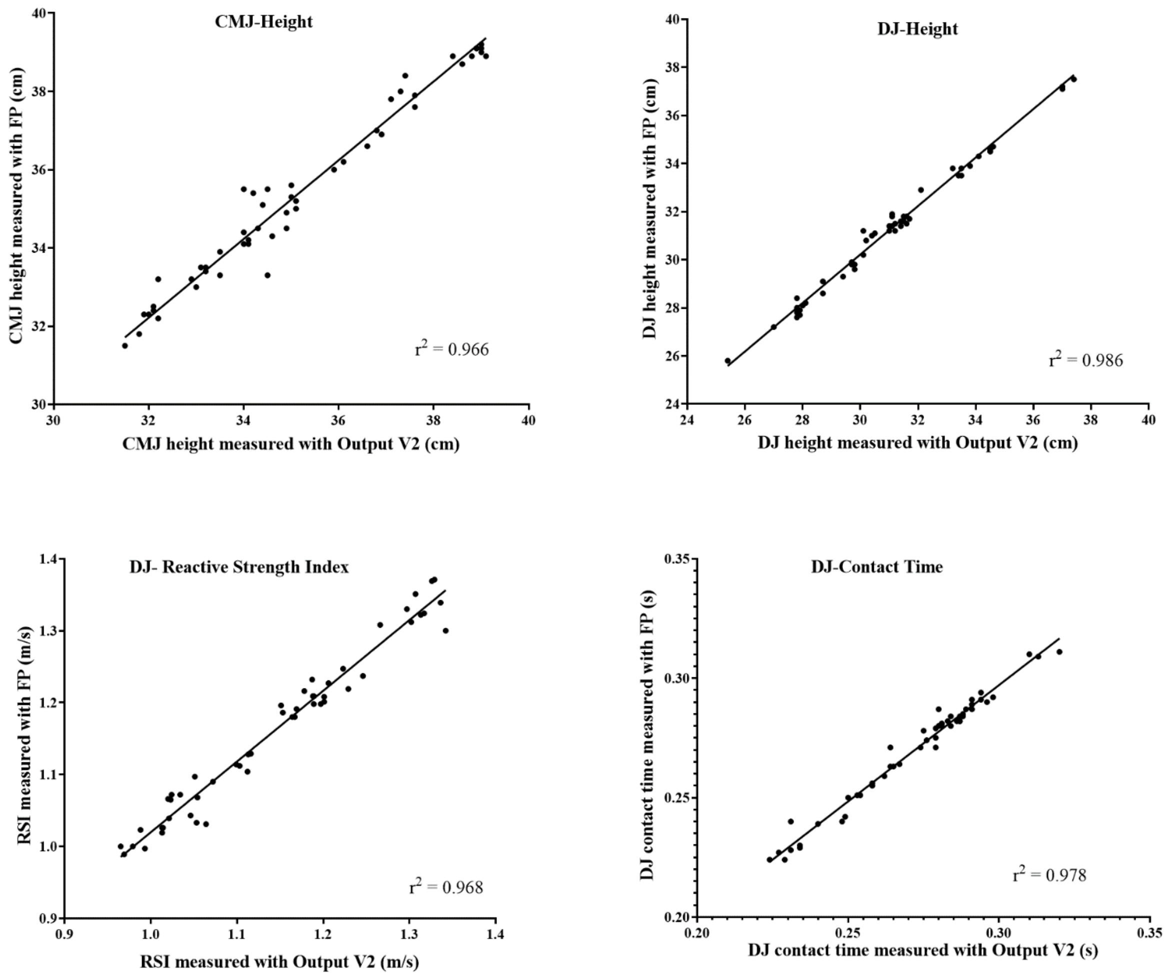

Figure 1. Scatter plot depicting the relationship between the inertial measurement unit *Output V2* and the force platform. CMJ – countermovement jump, DJ – drop jump, RSI – reactive strength index, CT – contact time.

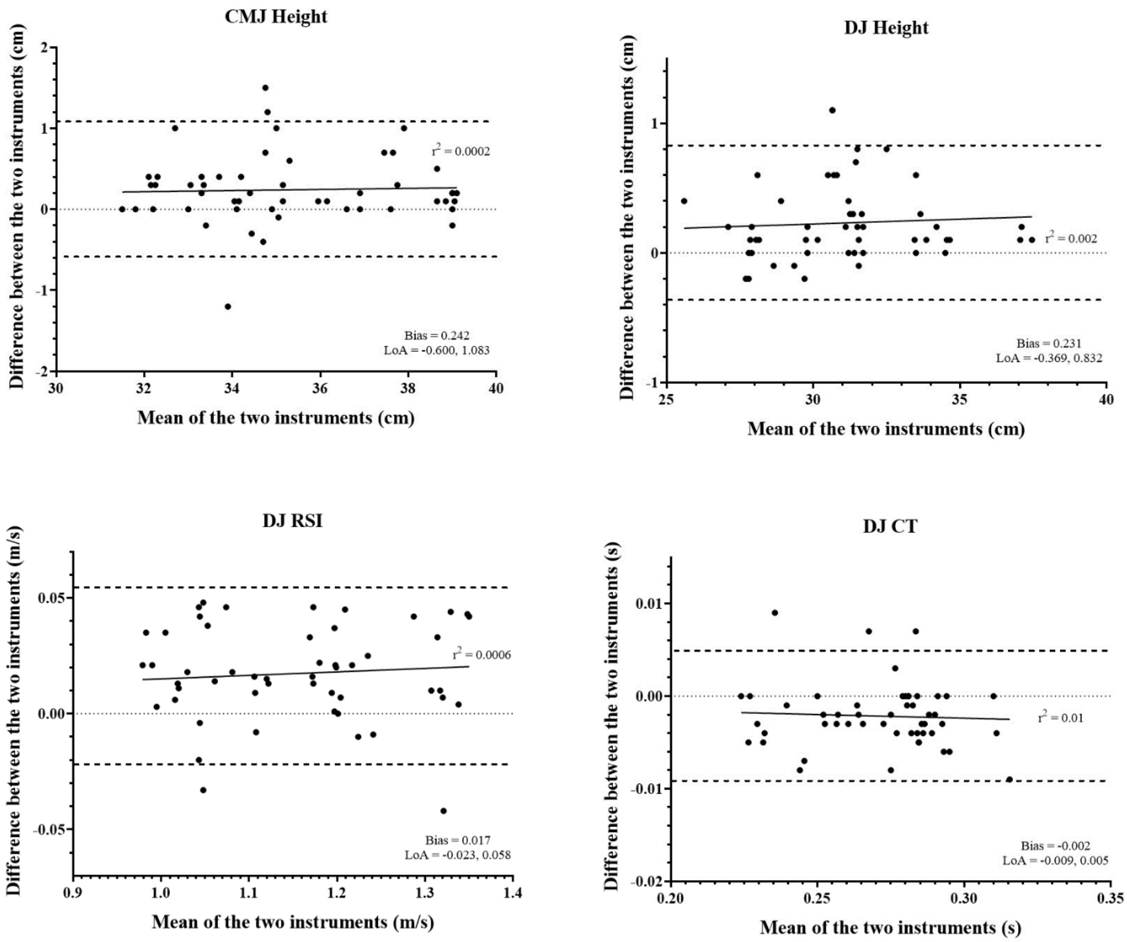

Figure 2. Bland–Altman plots representing countermovement jump (CMJ) height, drop jump (DJ) height, DJ reactive strength index (RSI), and DJ contact time (CT) data.