



# Development and Preliminary Evaluation of an IoT Enabled Ultrasonic System for Volleyball Jump Measurement

Rubiyatno<sup>1</sup>, Putra Sastaman<sup>1</sup>

<sup>1</sup>Sports Coaching Education, Faculty of Teacher Training and Education, Universitas Tanjungpura, Pontianak, Indonesia

Corresponding Author: Rubiyatno, [Rubiyatno@fkip.untan.ac.id](mailto:Rubiyatno@fkip.untan.ac.id)

## Abstract

**Objectives:** To develop an Internet of Things (IoT) enabled ultrasonic system for volleyball attack and block jump measurement and to examine its preliminary measurement performance and expert-rated feasibility.

**Materials and Methods:** A research and development study was conducted at Universitas Tanjungpura and Yuso Kalbar Volleyball Academy, Indonesia. The study documentation identified 60 purposively selected novice volleyball athletes aged 15–20 years; available results described a five-athlete pilot and a 20-athlete field trial. Five experts evaluated the prototype. Sensor and manual measurements were summarized using descriptive statistics, relative errors, and reported Pearson correlations.

**Results:** In the pilot, mean reference and sensor readings were 300.4 and 296.8 cm, respectively, with a mean absolute difference of 3.6 cm and a reported mean relative error of 1.20%. In the field trial, mean IoT and manual values were 65.2 and 64.2 cm for attack jumps and 56.4 and 55.4 cm for block jumps. Reported relative errors were 1.56% and 1.81%, respectively. Pearson correlations were 0.999 and 0.997 (both  $p < .001$ ). The reported overall expert score was 4.25/5. The relationship between pilot readings and displacement-based field outcomes was not documented.

**Conclusions:** The prototype showed promising preliminary feasibility and strong associations with manual measurements. Agreement, repeatability, and sensitivity to training-related change require further evaluation before the system can be considered interchangeable with reference methods or effective for improving athletic performance.

**Keywords:** volleyball; vertical jump; ultrasonic sensor; Internet of Things; measurement validity

## Introduction

Attack and block actions make vertical jumping an important component of volleyball assessment. However, the value of a jump measurement depends on the task performed and the quantity being measured. Volleyball-specific attack and block jump tests provide information that should be interpreted alongside, rather than automatically substituted for, standardized countermovement and squat jump tests. Sattler et al. (2012) demonstrated high reliability for volleyball-specific jumping tests and identified

differences associated with playing position. These findings support the assessment of attack and block actions within their own technical contexts.

Digital measurement tools offer practical alternatives for collecting jump information outside a laboratory. Their usefulness nevertheless depends on validation against an appropriate reference. In elite youth volleyball players, Borges et al. (2017) evaluated a wearable jump monitor against a jump-and-reach reference and observed systematic overestimation. Skazalski et al. (2018) subsequently found that a wearable device counted volleyball jumps accurately but that its height estimates were less suitable when precise maximal-jump assessment was required. Thus, successful jump detection does not necessarily establish accurate height measurement.

Recent validation studies also show that the performance of a digital tool depends on the outcome, calculation method, and testing conditions. Vieira et al. (2023) supported smartphone-based estimates of jump height while cautioning against extending those conclusions to velocity and power. Medeiros et al. (2024) evaluated a smartphone application specifically in elite women beach volleyball players on sand. These studies illustrate the importance of evaluating a device for its intended population, surface, and measured variable rather than assuming that validity transfers between settings.

Ultrasonic sensing provides another approach to tracking vertical position. Baumgart et al. (2017) compared an ultrasonic positioning system with force-plate and video methods and reported high trial-to-trial reliability alongside systematic differences in measured height. Therefore, ultrasonic jump measurement itself is an established approach. A useful development question is whether a locally implemented ultrasonic device, combined with internet-based recording and a website interface, can support the assessment of volleyball-specific jumps in a particular coaching environment.

The present development project addressed this practical question at Universitas Tanjungpura and Yuso Kalbar Volleyball Academy in West Kalimantan, Indonesia. The intended contribution was a portable measurement-and-recording workflow linking ultrasonic readings with athlete information and session records. Its novelty lies in the prototype and its contextual application; the available evidence does not support a claim that it is the first ultrasonic or internet-connected jump measurement system. The objectives were to develop the system, evaluate expert-rated feasibility, and examine preliminary relationships and differences between sensor and manual measurements during attack and block jump assessment.

## **Materials and Methods**

### **Study Design**

The study used a research and development approach organized around the ten-stage Borg and Gall development framework described in the source manuscript (Gall et al., 1996). Development was described as extending across two years, with prototype construction and feasibility assessment in the first year and an effectiveness phase in the second. Observations, interviews, and documentation supported identification of practical requirements and development of the system. The available numerical evidence comprises pilot measurements, field-trial summaries, expert ratings, and comparisons with manual measurements.

The source documentation also described a quasi-experimental pretest–posttest control-group component. However, it did not provide group sizes, baseline and follow-up

results, intervention exposure, or group-by-time estimates. Consequently, the present article reports the available development and measurement-comparison findings and does not infer an intervention effect from between-method differences.

### **Study Participants**

The documented recruitment pool comprised 60 novice volleyball athletes associated with the Faculty of Teacher Training and Education, Universitas Tanjungpura, Pontianak, and Yuso Kalbar Volleyball Academy. Purposive selection criteria were regular participation in training, age between 15 and 20 years, and basic competence in attack and block techniques. Pilot results were presented for five athletes and field-trial results for 20 athletes. Whether these subsets overlapped or how they related to the remaining recruited athletes was not specified; the reported recruitment total should therefore not be interpreted as the denominator for every analysis.

Five experts participated in prototype appraisal: two volleyball coaches, two information technology specialists, and one specialist in testing and measurement. Participant sex distribution, anthropometric characteristics, recruitment dates, sample-size justification, and expert selection criteria were not supplied in the available documentation. Ethics approval, participant consent, and parental permission procedures for participants younger than 18 years require confirmation from the authors before submission.

### **Study Organization**

The prototype combined ultrasonic sensing with an IoT connection and a website-based interface. Its intended functions were to acquire measurements during attack and block jumps, associate records with individual athletes, and make session data accessible to coaches and athletes. Portability and internet-based storage were described as product features. Transmission latency, missing-data frequency, processing speed, and long-term durability were not quantified in the reported trials.

Development activities included identifying measurement needs through observation and interviews, constructing the prototype, comparing its readings with a manual reference, obtaining expert feedback, and conducting pilot and field testing. Manual measurement served as the comparison method; the make, model, and measurement protocol of the reference instrument were not specified. The sensor model, mounting configuration, sampling frequency, peak-detection algorithm, calibration procedure, and conversion from distance readings to reported jump outcomes were also not sufficiently detailed for independent replication.

The pilot table contained five paired reference and sensor readings in the range of approximately 294–305 cm. These are retained as recorded heights because the documentation did not establish whether they represent absolute reach, another positional measure, or vertical displacement. The field table reported attack and block outcomes in a range of approximately 52–70 cm. These two sets of values are presented separately; their underlying measurement definitions must be verified before they can be interpreted as equivalent outcomes.

Experts rated functional suitability, accuracy and reliability, ease of use, and technological innovation on a five-point scale. The reported category for each aspect was “good.” These ratings represent expert appraisal of the prototype, including perceived measurement quality; they are distinct from empirical estimates of repeatability or

instrument agreement. The category thresholds and item-level rating data were not provided.

### Statistical Analysis

Analyses were described as being conducted in IBM SPSS Statistics, although the source abstract identified version 26 and the methods identified version 31. No software version is assigned here pending verification. Descriptive outcomes included means, standard deviations, minima, and maxima. Pilot absolute differences were obtained as the absolute value of the sensor reading minus the reference reading. The pilot relative error was consistent with  $100 \times |\text{sensor} - \text{reference}| / \text{reference}$ , averaged across the five observations. Field relative errors are reproduced as reported because individual field measurements were unavailable.

Pearson coefficients are reported as associations between methods, with their source-reported significance levels. The source also reported t tests with  $p < .05$ , but did not provide test statistics, degrees of freedom, exact p values, or sufficient information to establish whether measurements were paired. Those results are identified explicitly as unverified source-reported comparisons. They are not treated as evidence of training effectiveness or method agreement. Following the distinction between correlation and agreement described by Bland and Altman (1986), the absence of limits of agreement, agreement-based intraclass correlations, and repeatability estimates restricts the interpretation of measurement validity.

## Results

### Pilot Measurements

**Table 1.** Pilot comparison of reference and sensor readings

| Athlete | Reference (cm) | Sensor (cm) | Absolute difference (cm) | Relative error (%) |
|---------|----------------|-------------|--------------------------|--------------------|
| A       | 298            | 295         | 3                        | 1.01               |
| B       | 305            | 301         | 4                        | 1.31               |
| C       | 302            | 298         | 4                        | 1.32               |
| D       | 297            | 294         | 3                        | 1.01               |
| E       | 300            | 296         | 4                        | 1.33               |
| Mean    | 300.4          | 296.8       | 3.6                      | 1.20               |

Note.  $n = 5$ . Values reproduce the source table. The recorded height construct was not defined sufficiently to classify these readings as jump displacement.

Sensor readings were lower than reference readings for every pilot observation. Absolute differences ranged from 3 to 4 cm. The reported mean relative error of 1.20% corresponds to approximately 98.8% when subtracted from 100%; this arithmetic complement should not be interpreted as a comprehensive measure of instrument accuracy

### Field Measurements

**Table 2.** Descriptive field-trial outcomes by jump type and method

| Parameter | Attack IoT | Attack manual | Block IoT | Block manual |
|-----------|------------|---------------|-----------|--------------|
| Mean (cm) | 65.2       | 64.2          | 56.4      | 55.4         |

|                         |      |      |      |      |
|-------------------------|------|------|------|------|
| SD (cm)                 | 2.96 | 2.95 | 2.71 | 2.69 |
| Minimum (cm)            | 60   | 59   | 52   | 52   |
| Maximum (cm)            | 70   | 69   | 62   | 61   |
| Reported mean error (%) | 1.56 | —    | 1.81 | —    |

Note. The source describes a field trial involving 20 athletes. Individual paired data and the calculation procedure for field relative error were unavailable. SD = standard deviation.

Mean IoT readings exceeded manual readings by 1.0 cm for both jump types. The reported dispersions were similar across methods, while relative error was slightly greater for block than attack measurements. Similar standard deviations describe between-athlete dispersion and do not, by themselves, establish test–retest reliability.

### Expert Appraisal

**Table 3.** Expert appraisal of the prototype

| Assessment aspect         | Mean score (1–5) | Reported category |
|---------------------------|------------------|-------------------|
| Functional suitability    | 4.30             | Good              |
| Accuracy and reliability  | 4.20             | Good              |
| Ease of use               | 4.25             | Good              |
| Technological innovation  | 4.30             | Good              |
| Overall score as reported | 4.25             | Good              |

Note. Five experts participated. The reported overall score is retained. The simple unweighted average of the four displayed aspect means is 4.2625; item-level data or weighting information are needed to verify the reported aggregate.

All four assessed aspects were classified as good. Functional suitability and technological innovation received the highest displayed means. The accuracy and reliability rating is an expert judgment rather than a calculated reliability coefficient.

### Reported Method Comparisons

**Table 4.** Source-reported differences between manual and IoT measurements

| Jump type | Manual mean (cm) | IoT mean (cm) | Reported t-test result |
|-----------|------------------|---------------|------------------------|
| Attack    | 64.2             | 65.2          | $p < .05$              |
| Block     | 55.4             | 56.4          | $p < .05$              |

Note. Significance thresholds are reproduced from the source and could not be independently verified. Test statistics, degrees of freedom, exact p values, pairing information, and raw data were not supplied.

The source reported statistically significant differences between methods. These comparisons concern measurement values and contain no reported pretest-to-posttest change scores. They therefore do not establish that use of the prototype improved attack or block jumping performance.

**Table 5.** Source-reported Pearson correlations between methods

| Jump type | r     | p        | Statistical meaning              |
|-----------|-------|----------|----------------------------------|
| Attack    | 0.999 | $< .001$ | Very strong positive association |
| Block     | 0.997 | $< .001$ | Very strong positive association |

Note. Coefficients and p values are retained from the source. Raw paired observations and analysis-specific sample sizes were not available for verification. Correlation does not establish interchangeability.

Both reported coefficients indicated very strong positive relationships between IoT and manual measurements. Agreement-based statistics and confidence intervals were not provided, so the magnitude of individual measurement disagreement cannot be established from these coefficients alone.

## **Discussion**

### **Interpretation of the Preliminary Findings**

The prototype combined ultrasonic measurement with an internet-connected recording workflow and received favorable expert ratings. The pilot and field summaries suggest that sensor readings tracked the comparison measurements, but they also reveal an important calibration question: the pilot showed lower sensor readings, whereas the field means were higher for the IoT method. Because the pilot height construct was not defined and the trial protocols were incompletely described, this reversal cannot confidently be attributed to improved calibration, different jump techniques, or a change in the measured variable. It identifies a question for verification rather than evidence of consistent bias across settings.

The closest technological comparison is the ultrasonic system examined by Baumgart et al. (2017). Their results demonstrated that a method can be repeatable while producing height estimates that differ systematically from other approaches. The present findings are compatible with that general measurement concern, but do not establish equivalent device performance. The earlier study described its sensor positioning, sampling, calibration, and height definition in detail. Reporting those elements for the present prototype is necessary to determine whether its recorded quantity is absolute reach, body-position displacement, or another operational definition of jump height.

### **Comparison with Other Digital Measurement Studies**

Research in volleyball shows why a small mean difference should be assessed together with individual error. Borges et al. (2017) found that a wearable jump monitor systematically overestimated attack and block measurements relative to a jump-and-reach reference. Skazalski et al. (2018) distinguished excellent jump-count accuracy from the precision needed to evaluate maximal jumping ability. In the present study, the 1 cm mean differences describe average offsets only. They do not reveal whether some athletes experienced substantially larger discrepancies or whether the system could identify small changes over successive training sessions.

Jimenez-Olmedo et al. (2023) provide a recent example of strong association coexisting with limited concordance. In highly trained female volleyball players, their comparison of an inertial measurement unit with motion capture produced a strong relationship but substantial systematic differences. This supports a restrained interpretation of the present Pearson coefficients: athletes with larger manual values also tended to have larger sensor values, but the two methods cannot yet be treated as interchangeable. A significant difference between methods is likewise compatible with systematic measurement bias rather than a beneficial intervention effect.

Smartphone studies offer useful methodological comparisons without validating the present ultrasonic hardware. Bogataj et al. (2020) evaluated concurrent validity and repeatability for several jump tasks in recreationally active adults. Vieira et al. (2023) demonstrated that validity conclusions were outcome-specific, with stronger support for

jump height than for several derived mechanical variables. Accordingly, the current device should remain described as a jump-measurement prototype. The available results do not establish that it measures lower-limb force, power, or other mechanical variables.

Medeiros et al. (2024) assessed jump measurements on sand in elite female beach volleyball players and included both observer agreement and comparisons with a force platform. Their design highlights the importance of surface and population specificity. The present trial involved novice athletes from an Indonesian university and academy setting, so its findings should not automatically be generalized to beach volleyball, elite teams, or other age groups. Murawa et al. (2024) further showed that changing the calculation approach within smartphone-based assessment can alter agreement with a force platform. Thus, clear definitions and algorithms are essential even when a digital interface appears straightforward.

### **Practical Contribution and Application**

The practical contribution of the prototype is the integration of measurement with athlete-linked digital records. Such a workflow could help coaches retrieve previous results and organize repeated assessments without manually transferring every reading to a separate record. However, the present study did not quantify time savings, user error, data completeness, internet dependence, or cost. These remain potential operational benefits to investigate rather than demonstrated advantages over existing systems.

For provisional use, a coach would need a documented and stable assessment protocol covering the jump task, reference position, sensor alignment, calibration, and scoring rule. Attack and block results should be stored separately because the actions have different technical demands, consistent with the task-specific approach discussed by Sattler et al. (2012). Comparing repeated results from the same protocol may be more defensible than combining measurements obtained from different devices or calculation methods. Any interpretation of small changes should await estimates of measurement error and minimum detectable change.

### **Limitations and Future Research**

Several limitations constrain the conclusions. The relationship between the 60 recruited athletes and the five- and 20-athlete trial samples was not documented. Sex distribution and anthropometric characteristics were absent, and the pilot and field outcomes appeared to use different height scales without an explicit conversion rule. The manual comparison method was insufficiently specified and should not be assumed to represent a criterion standard. Raw field data were unavailable, preventing verification of correlations, error calculations, and significance tests. Expert ratings supported perceived feasibility, but could not substitute for measured reliability or agreement.

The source described a controlled pretest–posttest phase without providing the corresponding longitudinal results. Therefore, no causal conclusion about improved jumping performance can be drawn. Hardware specifications, repeatability, environmental sensitivity, signal loss, network latency, and durability also remain unquantified. The expert aggregate requires reconciliation with the displayed aspect scores. These limitations concern the completeness of the evidence as well as the scope of the initial evaluation.

Future validation should obtain simultaneous paired measurements with a clearly defined reference, provide the individual observations, and prespecify acceptable error for the

intended coaching decision. Agreement should be examined through mean bias and limits of agreement, with repeated trials used to evaluate within-session and between-session repeatability (Bland & Altman, 1986). Testing should distinguish absolute reach from jump displacement and include both attack and block tasks. A separate controlled longitudinal study would be needed to determine whether access to the digital feedback changes training decisions or improves performance.

## Conclusions

An IoT-enabled ultrasonic prototype was developed for volleyball jump measurement and digital recording. Favorable expert appraisal and strong reported relationships with manual measurements support continued development and a more complete validation study. The main contribution is a locally implemented workflow connecting jump assessment with athlete records in a coaching setting. Current findings do not establish interchangeability with reference instruments, reliable detection of small performance changes, or improved jumping ability. Further work should standardize the measurement definition and protocol, quantify agreement and repeatability, and evaluate the practical value of the digital feedback during routine training.

## Acknowledgments

The authors thank the athletes, coaches, and expert reviewers who contributed to prototype appraisal and testing at Universitas Tanjungpura and Yuso Kalbar Volleyball Academy.

## Conflict of Interest

The authors declare no conflicts of interest related to this study.

## Funding

This research was funded by an internal research grant from the Institute for Research and Community Service (LPPM), Universitas Tanjungpura, Indonesia

## References

- Baumgart, C., Honisch, F., Freiwald, J., & Hoppe, M. W. (2017). Differences and trial-to-trial reliability of vertical jump heights assessed by ultrasonic system, force-plate, and high-speed video analyses. *Asian Journal of Sports Medicine*, 8(4), Article e12921. <https://doi.org/10.5812/asjms.12921>
- Bland, J. M., & 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](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Bogataj, Š., Pajek, M., Andrašić, S., & Trajković, N. (2020). Concurrent validity and reliability of My Jump 2 app for measuring vertical jump height in recreationally active adults. *Applied Sciences*, 10(11), Article 3805. <https://doi.org/10.3390/app10113805>
- 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>
- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction* (6th ed.). Longman.

- Jimenez-Olmedo, J. M., Pueo, B., Mossi, J. M., & Villalon-Gasch, L. (2023). Concurrent validity of the inertial measurement unit Vmaxpro in vertical jump estimation. *Applied Sciences*, 13(2), Article 959. <https://doi.org/10.3390/app13020959>
- Medeiros, A. I. A., Silva, G. M. da, Neto, F. O., Simim, M., Banja, T., Coswig, V. S., Afonso, J., Ramos, A., & Mesquita, I. (2024). Validity and reliability of My Jump 2® app to measure the vertical jump on elite women beach volleyball players. *PeerJ*, 12, Article e17387. <https://doi.org/10.7717/peerj.17387>
- Murawa, M., Krakowiak, W., & Kabaciński, J. (2024). Validity and reliability of a smartphone app for vertical jump height assessment using the marker displacement time method. *Applied Sciences*, 14(11), Article 4843. <https://doi.org/10.3390/app14114843>
- Sattler, T., Sekulic, D., Hadzic, V., Uljevic, O., & Dervisevic, E. (2012). Vertical jumping tests in volleyball: Reliability, validity, and playing-position specifics. *Journal of Strength and Conditioning Research*, 26(6), 1532–1538. <https://doi.org/10.1519/JSC.0b013e318234e838>
- 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*, 28(5), 1578–1585. <https://doi.org/10.1111/sms.13052>
- Vieira, A., Ribeiro, G. L., Macedo, V., de Araújo Rocha Junior, V., Baptista, R. de S., Gonçalves, C., Cunha, R., & Tufano, J. (2023). Evidence of validity and reliability of Jumbo 2 and MyJump 2 for estimating vertical jump variables. *PeerJ*, 11, Article e14558. <https://doi.org/10.7717/peerj.14558>

## Author Verification Notes

Editorial notes only. Remove this section after resolving the items below and before journal submission. The article has been rewritten to match the supplied template while retaining the source data and avoiding unsupported effectiveness claims.

**Participant accounting.** Confirm whether the five-athlete pilot and 20-athlete field samples overlap, how they relate to the stated total of 60, and the analysis-specific  $n$  for each correlation and test. Supply sex distribution, anthropometrics, dates, and recruitment details.

**Measured quantity and protocol.** Clarify what the 294–305 cm pilot readings represent and how the 52–70 cm field outcomes were calculated. Supply standing-reach values where applicable, sensor and reference instrument specifications, mounting, sampling frequency, calibration, algorithm, jump instructions, number of attempts, rest intervals, measurement order, and scoring rule.

**Statistical evidence.** Provide raw paired field data and original SPSS output. Confirm whether tests were paired or independent. The reported  $p < .05$  cannot be verified from the displayed summaries; an independent comparison with  $n = 20$  per method would not reproduce that threshold using the reported means and SDs. Do not describe these differences as improved performance.

**Longitudinal effectiveness.** Provide experimental and control group sizes, allocation method, intervention content and duration, pretest/posttest means and SDs, and appropriate between-group change analyses if retaining an effectiveness claim. Otherwise keep the preliminary measurement-evaluation framing used here.

**Expert aggregate.** Verify the reported overall score of 4.25. The unweighted average of 4.30, 4.20, 4.25, and 4.30 is 4.2625. Explain weighting or correct the aggregate using original ratings; provide the rating rubric and category thresholds.

**Ethics and declarations.** Supply ethics committee name, approval number and date, consent procedures, and permission for minors where applicable. Confirm funding, conflicts, product ownership, acknowledgments, author affiliation wording, and the corresponding author.

**Software and documentation.** Resolve SPSS version 26 versus 31. Verify the actual development stages and research dates. Supply original device photographs or a technical schematic if a product figure is required. No illustrative hardware image has been invented.

**References.** Ten journal references have DOI links matched to publisher/indexed records; the development-method book has no DOI supplied. Unrelated or poorly matched original references have been replaced. DOI metadata identifies an article but does not establish the quality of the present dataset.