Validation of an Inertial Sensor System for Swing Analysis in Golf †

Paul Lückemann [1], Daniel M. Haid [1], Philip Brömel [1], Stefan Schwanitz [2,*] and Christian Maiwald [3]

[1] School of Mechanical Engineering, Chemnitz University of Technology, 09107 Chemnitz, Germany; paul.lueckemann@physik.hu-berlin.de (P.L.); daniel-matthias.haid@s2013.tu-chemnitz.de (D.M.H.); philip.broemel@s2015.tu-chemnitz.de (P.B.)

[2] Department of Sports Equipment & Technology, Chemnitz University of Technology, 09107 Chemnitz, Germany

[3] Department of Food and Nutrition, and Sport Science, University of Gothenburg, 405 30 Gothenburg, Sweden; christian.maiwald@googlemail.com

* Correspondence: stefan.schwanitz@mb.tu-chemnitz.de; Tel.: +49-371-531-36657

† Presented at the 12th conference of the International Sports Engineering Association, Brisbane, Queensland, Australia, 26–29 March 2018.

Abstract: Wearable inertial sensor systems are an upcoming tool for self-evaluation in sports, and can be used for swing analysis in golf. The aim of this work was to determine the validity and repeatability of an inertial sensor system attached to a player's glove using a radar system as a reference. 20 subjects performed five full swings with each of three different clubs (wood, 7-iron, wedge). Clubhead speed was measured simultaneously by both sensor systems. Limits of Agreement were used to determine the accuracy and precision of the inertial sensor system. Results show that the inertial sensor system is quite accurate but with a lack of precision. Random error was quantified to approximately 12%. The measurement error was dependent on the club type and was weakly negatively correlated to the magnitude of clubhead speed.

1. Introduction

One main aspect of playing golf is to hit the ball as far and as accurately as possible, at least for many of the drives taken during a round of golf. To achieve maximum carry distance, clubhead speed is one of several decisive factors in the golf swing that influence ball trajectory. The slightest variation of swing or addressing the ball can dramatically change the ball's trajectory. Available technologies have made the analysis of golf clubhead swing path and ball flight trajectory more affordable and more popular, not only in golf coaching, but also in TV coverage of pro golf events.

Previous biomechanical research on the golf swing has mostly been carried out in the laboratory using dynamometry, video analysis or other motion tracking systems like electro-goniometers [1,2]. However, it is questionable whether the measurement results of these studies can be transferred to the field. Wearable inertial sensor systems offer the opportunity to measure swing parameters directly on the golf course and to give individual players instant feedback on their actual golf swing. Inertial sensors have been used in various ways to measure swing parameters [2–7], and biomechanical models have been developed based on this data [6]. One main problem of using inertial sensors for swing analysis is that algorithms have to be developed which estimate the key variables (e.g., clubhead speed, clubhead path) based on accelerometer and gyroscope data not

captured on the clubhead itself, but attached to e.g., the club shaft, the club handle, or the golfer's hand. The displacement of the sensor relative to the clubhead, induced by shaft flexion and torsion, and the interaction of glove, hand, and club may massively influence the validity of these algorithms. Thus, the validity of inertial sensor systems in swing analysis has been the subject of previous studies. The study by Lyu and Smith [3] found that bat swing speed in baseball was underestimated by 8% on average using wireless inertial sensors, likely due to sensor saturation at 16 g. But also, a difference in exact definition of clubhead position between the inertial sensor and the reference system might lead to a bias in measuring clubhead speed.

Radar based swing analysis systems have been established as the gold standard for measuring kinematic properties of the golf swing [8]. The technologies like the one incorporated in the TrackMan series of products, offer the significant advantage of direct velocity measuring and being field ready and cost-effective tools for capturing swing properties under realistic conditions [8].

As with any scientific assessment of new measurement devices, data reliability and validity are core criteria for usability. Reliability of a new device is commonly assessed using repeated measurements of the same quantity (true value) under identical conditions. This assumes the true value to be constant across replications. In the case of golf swings, there are no exact repetitions of any golf swing, so the true value may be subject to change across repetitions. Thus, reliability of measurements in the context of repeated human movements is always confounded with naturally occurring motor variability. Since there is no way to determine the true value of every single repetition of human movement, it makes sense to compare measurements of the device in question to those of a gold standard device to assess measurement error. Data from the gold standard device will serve as the closest estimate of quantity to be measured. Measurement error of the device in question can then be divided into systematic and random error components. They are commonly referred to as accuracy and precision, which both have an influence on agreement [9,10].

The goal of the present study was to compare the measurement results of an inertial sensor-based mobile device for analysing golf swing (Zepp Golf 2, Zepp Labs Inc., San Jose, CA, USA) to a gold standard device (TrackMan 4, TrackMan Golf, Vedbæk, Denmark), and to analyse the accuracy and precision of the mobile device's data. Clubhead speed was chosen as the key variable, since it is provided by both systems. Possible confounding variables, like type of club used and the golfer's handicap, were analysed for their potential influence on the measurement results of the mobile device. For the factor club type, we hypothesized that different club types have an effect on the systematic measurement error due to their different lengths and degrees of lie.

2. Materials and Methods

2.1. Participants

To ensure adequate skill levels of subjects, only golfers with a EGA-handicap [11] of 54 or better were included in the study. Subjects' golf handicaps ranged from 0 (professional) to 54 (beginner) (Quartiles Q25/Q50/Q75: 11.8/20.0/45.0). Subjects were excluded if they suffered from any physical injury or any other sickness, which would affect their golf swing. Altogether 20 subjects (aged 37 ± 13 years, 178.4 ± 7.5 cm tall, 3 female, 17 male) were included in the study.

2.2. Materials

The inertial sensor system to be evaluated was the Zepp Golf 2. It was attached to each golfer's glove at the backhand of the leading arm, 3–5 cm distal from the wrist. With a size of 25.4 × 25.4 × 12.3 mm and a weight of 6.25 g, it was presumed to have no influence on the golf swing. The device consisted of two accelerometers, two gyroscopes, and a lithium ion battery. More detailed specifications are not provided by the founding company. The obtained data was sent via Bluetooth to the Zepp Golf Swing Analyzer App (Version 3.4.1) on a tablet computer positioned nearby. The radar system TrackMan 4 was used as the reference system for the measurements. Measurements were made for three different types of golf clubs. To represent different categories of clubhead speed, a driving club (either driver or 3-wood), a 7-iron, and a wedge were used for analysis. Subjects were

allowed to choose their favorite woods and wedges out of their own club sets, but were required to use the same 7-iron, which was instrumented with additional sensors for another study conducted simultaneously. Subjects used their own gloves for all swings, and the Zepp Golf 2 sensor was attached to the glove with its universal clamping mechanism.

2.3. Procedure

Prior to the measurements, each subject was briefed about the procedure and the aim of the study, and also gave written informed consent to volunteer in the experiment. Subjects performed an individual warm up until they felt ready. The Zepp Golf 2 sensor was then mounted to the subjects' glove, and after selecting the club type in the Zepp Golf Swing Analyzer App, it performed a three-second calibration while subjects took their ball addressing position. Subjects were then instructed to hit all golf balls with their usual swing technique. Each subject performed five valid shots with each of the three club types. Subjects were instructed to report any perception of a mishit for each shot. If subjects felt the ball was hit improperly, or if ball flight indicated a bad hit, the research team declared the shot as invalid and asked the subject to repeat the shot until five valid shots per club had been recorded.

2.4. Data Analysis

The measurement error for the Zepp Golf 2 system was analysed using Bland and Altman's limits of agreement (LoA) [10]. Since every shot was subject to natural motor variability, the limits of agreement procedures for non-constant true values were employed [10]. We assumed the golf shots of each subject to be mutually independent, since the variable in question was agreement between devices (measurement error) and subjects were unaware of error magnitudes throughout their testing session. Thus, the measurement error (d) of the Zepp Golf 2 sensor was calculated by subtracting the TrackMan 4 club head speed from the Zepp Golf 2 club head speed for every shot ($d = v_Z - v_T$). Pooled d were plotted against the corresponding averaged clubhead speeds ($v = \frac{1}{2}(v_Z + v_T)$) in a Bland-Altman diagram to check for heteroscedasticity [9]. We calculated Pearson's correlation coefficient r to investigate the relationship between the magnitude of clubhead speed (v) and its measurement error (d). Since only weak correlations were found, constant limits of Agreement (LoA) were calculated using the average across d as mean difference (MD) and the standard deviation (SD) to be inserted into the LoA formula (LoA = MD ± 1.96 × SD). SD was assessed using Bland and Altman's variance formula for pairwise replicated data,

$$\text{SD}^2 = MS_w + \frac{(n-1)\sum m_i}{(\sum m_i)^2 - \sum m_i^2} \times (MS_b - MS_w), \tag{1}$$

with MS_w being the mean within-subject variation and MS_b being the mean between-subject variation. Both were determined using a one-way ANOVA as described in [9]. Sample size $n = 20$ and number of replications $m_i = 5$ yields equation (1) to

$$\text{SD} = \sqrt{\frac{1}{5}(4\,MS_w + MS_b)}. \tag{2}$$

LoA represent the precision of Zepp Golf 2, and 95% of the deviations between Zepp Golf 2 and TrackMan 4 can be expected to lie within the LoA. Accuracy was estimated by MD, which indicates the average (systematic) bias of the Zepp Golf 2 clubhead speed measures compared to TrackMan 4 readings.

Measurement errors (d) were then averaged across the five swings per club type to assess the true value of measurement error (d_{avg}) for each subject and club type. Box plots of d_{avg} were used to visually assess normal distribution of measurement errors and to illustrate differences in distributions between the three club types. Due to slight differences in distributions, Friedman's test was used to check the hypothesis of different errors being associated with different club types. When applicable, post-hoc Wilcoxon's signed rank tests were carried out to identify statistically significant differences in error magnitude between club types. All significance levels were set to $\alpha = 0.05$.

3. Results

The measurement error of Zepp Golf 2, represented by the differences of clubhead speed (*d*), was plotted in Figure 1a–c for each club type. Swings of one subject were plotted as data points with corresponding numbers 1 to 20. In general, subject swings appeared as grouped data points, for example subjects 12 and 19 in Figure 1c. However, the within-subject variation for measurement error and clubhead speed differed substantially between subjects (e.g., subject 1 vs. 12). Figure 1d aggregates the individual data into box plots of distributions of d_{avg} per club type across all subjects. It illustrates that the variations and central tendencies of measurement errors differed slightly between club types. Table 1 contains the numerical summary of measurement errors: averaged clubhead speeds for each club type, mean within-subject (MS_w) and between-subject variations (MS_b), LoA (given as absolute and relative LoA), and correlation coefficient between *d* and clubhead speed.
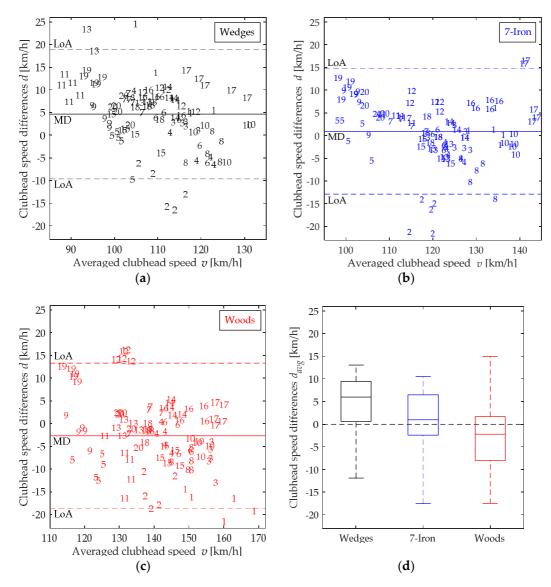

Figure 1. (**a**–**c**) Bland-Altman-diagrams showing measurement error *d* vs. clubhead speed *v*, averaged via Trackman 4 and Zepp Golf 2. The 5 swings per subject are denoted as subject numbers (1 to 20). MD and LoA are represented by solid and dashed lines, respectively. Corresponding values, including correlation coefficients, are noted in Table 1. (**d**) Box plots show the distribution of subject-averaged measurement errors d_{avg}, grouped by the different club types. The dashed line suggests a measurement error of zero.

Table 1. Corresponding parameters to Figure 1 including mean variation within-subjects (MS_w) and between-subjects (MS_b). LoA given as absolute values and normalized to averaged clubhead speed. Correlation coefficients r given for linear model of d vs. clubhead speed.

Club Type	Averaged Clubhead Speed	MS_w [km²/h²]	MS_b [km²/h²]	LoA Absolute and Relative to Averaged Clubhead Speed		Pearson's r
Wedge	110.6 km/h	15.6	202.2	4.6 ± 14.3 km/h	4.2 ± 12.9%	−0.41
Iron 7	120.2 km/h	7.8	216.3	0.9 ± 13.8 km/h	0.7 ± 11.4%	−0.23
Wood	138.0 km/h	7.1	303.8	−2.7 ± 16.0 km/h	−2.0 ± 11.6%	−0.37

Precision of Zepp Golf 2 was assessed using the upper and lower LoA boundaries. The three club types showed equal widths of the absolute LoA interval. Normalizing the LoA interval to the averaged clubhead speeds resulted in highest precision for woods, followed by 7-iron, and wedges.

Accuracy of Zepp Golf 2 was assessed using the mean differences (MD). MDs differed across club types (wedges: +4.6 km/h, 7-iron: +0.9 km/h, woods: −2.7 km/h). Friedman's test indicated statistically significant differences between club types with $p = 3 \times 10^{-5}$. Post-hoc Wilcoxon's signed rank tests were applied, which resulted in $p \leq 0.023$ for each combination of the three club types. Therefore, the alternative hypothesis, that measurement error is influenced by different club types, was accepted.

A weak correlation between d and clubhead speed was found, as Pearson's r for any club type was −0.37 on average. This finding corresponds to decreasing MDs with increasing club length and higher clubhead speeds. Additionally, a weak positive correlation was found between d_{avg} and the subjects' handicaps.

4. Discussion

The aim of this study was to evaluate the validity of the Zepp Golf 2 inertial sensor system using the TrackMan 4 radar system as the reference, and using the LoA method to quantify the agreement between the two systems' clubhead speed outputs. LoA define the interval in which 95% of the deviations between Zepp Golf 2 and TrackMan 4 can be expected across individuals. Hence, for a swing from an individual, representing the studied population, the clubhead speed measurement error of Zepp Golf 2 was expected to be 4.6 ± 14.3 km/h for wedges, 0.9 ± 13.8 km/h for 7-iron, and −2.7 ± 16.0 km/h for woods. The systematic bias (accuracy, MD) was small in comparison to the random error (precision, LoA width). A directed bias of −8%, as found by Lyu and Smith for a bat swing [3], was not reproduced in our study. Most of the other studies on golf swing kinematics focused on variables of the swing path (position and orientation) [1,4–6] and cannot be compared directly to our clubhead speed measurements.

We found a lack of precision for the Zepp Golf 2 sensor system, since with any club type potential measurement errors of up to 27 km/h can occur, which are in the same order of magnitude as the interindividual differences between clubhead speeds in our sample. The results suggest that Zepp Golf 2 can only deliver accurate clubhead speeds on average across multiple shots, but not for individual shots. This limits the usability of the system for golf swing analyses focusing on clubhead speed. Furthermore, although no correlation was observed between measurement error and clubhead speed within individual club types, it appears that MD is associated with club length and is different between club types. Figure 1 indicates that the longer the club, the more the average clubhead speed is underestimated by Zepp Golf 2. While there is a positive bias in average clubhead speed for wedges, this bias almost disappears for 7-iron, and reverses for woods. Length, design, and weight distribution change with club type, and individual differences in swing characteristics across club types may not be adequately reflected by the algorithms used in Zepp Golf 2.

We found a weak dependency of measurement error on the subjects' golf handicaps. However, we did not assess a representative sample of golfers, the handicaps were self-reported and not verified, and our distribution of handicaps was skewed towards less-skilled golfers. Future studies may focus on possible effects of skill level on measurement error of glove-mounted systems, like the Zepp Golf 2 system.

High variation in measurement error for wedges and woods could partially be caused by variation in club specifications. Since subjects used different types of woods and wedges, but an identical 7-iron, the generic club types provided by the Zepp Golf 2 database may have induced systematic bias for wedges and woods. It is unlikely that individually chosen wedges and woods were the causal factor for imprecise measurements, since the random variation was comparable for all club types. Minimal effects on measurement error may have occurred by allowing subjects to wear their own gloves, so that the attachment of the Zepp Golf 2 sensor might have been slightly different for each subject. Although these factors may limit the accuracy and precision of Zepp Golf 2 measurements, they provide an estimate on practicability issues when using this device in the field, where the inertial sensor system can be used by any individual with any type of equipment.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hume, P.A.; Keogh, J.; Reid, D. The Role of Biomechanics in Maximising Distance and Accuracy of Golf Shots. *Sports Med.* **2005**; *35*, 429–449, doi:10.2165/00007256-200535050-00005.
2. Nam, C.N.K.; Kang, H.J.; Suh, Y.S. Golf swing motion tracking using inertial sensors and a stereo camera. *IEEE Trans. Instrum. Meas.* **2014**, *63*, 943–952, doi:10.1109/TIM.2013.2283548.
3. Lyu, B.; Smith, L.V. Evaluation of wireless bat swing speed sensors. *Sports Eng.* **2017**, *20*, doi:10.1007/s12283-017-0224-3.
4. King, K.; Yoon, S.W.; Perkins, N.C.; Najafi, K. Wireless MEMS inertial sensor system for golf swing dynamics. *Sens. Actuators A* **2008**, *141*, 619–630, doi:10.1016/j.sna.2007.08.028.
5. Lai, D.T.H.; Hetchl, M.; Wei, X.C.; Ball, K.; McLaughlin, P. On the difference in swing arm kinematics between low handicap golfers and non-golfers using wireless inertial sensors. *Procedia Eng.* **2011**, *13*, 219–225, doi:10.1016/j.proeng.2011.05.076.
6. Ghasemzadeh, H.; Loseu, V.; Jafari, R. Wearable coach for sport training: A quantitative model to evaluate wrist-rotation in golf. *J. Ambient Intell. Smart Environ.* **2009**, *1*, 1–12, doi:10.3233/AIS-2009-0021.
7. Seaman, A.; McPhee, J. Comparison of Optical and Inertial Tracking of Full Golf Swings. *Procedia Eng.* **2012**, *34*, 461–466, doi:10.1016/j.proeng.2012.04.079.
8. Leach, R.J.; Forrester, S.E.; Mears, A.C.; Roberts, J.R. How valid and accurate are measurements of golf impact parameters obtained using commercially available radar and stereoscopic optical launch monitors? *Measurement* **2017**, *112*, 125–136, doi:10.1016/j.measurement.2017.08.009.
9. Atkinson, G.; Nevill, A.M. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.* **1998**, *26*, 217–238, doi:10.2165/00007256-199826040-00002.
10. Bland, J.M.; Altman, D.G. Agreement between methods of measurement with multiple observations per individual. *J. Biopharm. Stat.* **2007**, *17*, 571–582, doi:10.1080/10543400701329422.
11. European Golf Association. Handicap System 2016–2019. Available online: http://www.ega-golf.ch/sites/default/files/epub_hcp_booklet_2016_3.11.2015.pdf (accessed on 5 December 2017).