The Reliability of Technical and Tactical Tagging Analysis Conducted by a Semi-Automatic VTS In Soccer

by
Marco Beato¹, Mikael Jamil¹, Gavin Devereux¹

The Video Tracking multiple cameras system (VTS) is a technology that records two-dimensional position data (x and y) at high sampling rates (over 25 Hz). The VTS is of great interest because it can record external load variables as well as collect technical and tactical parameters. Performance analysis is mainly focused on physical demands, yet less attention has been afforded to technical and tactical factors. Digital.Stadium® VTS is a performance analysis device widely used at national and international levels (i.e. Italian Serie A, Euro 2016) and the reliability evaluation of its technical tagging analysis (e.g. shots, passes, assists, set pieces) could be paramount for its application at elite level competitions, as well as in research studies. Two professional soccer teams, with 30 male players (age 23 ± 5 years, body mass 78.3 ± 6.9 kg, body height 1.81 ± 0.06 m), were monitored in the 2016 season during a friendly match and data analysis was performed immediately after the game ended. This process was then replicated a week later (4 operators conducted the data analysis in each week). This study reports a near perfect relationship between Match and its Replication. R² coefficients (relationships between Match and Replication) were highly significant for each of the technical variables considered (p < 0.001). In particular, a high score of interclass correlation and a small coefficient of variation were reported. This study reports meaningless differences between Match and its Replication (intra-day reliability). We concluded that the semi-automatic process behind the Digital.Stadium® VTS was more than capable of recording technical tagging data accurately.

Key words: team sports, time-motion analysis, match analysis.

Introduction
Notational analysis is a convenient and practical procedure commonly used in soccer, in order to examine the effects of different game approaches (e.g. formations, game styles), attacking movements, technical actions (e.g. shots, crosses, passing sequences) as well as possession and the spatial locations of players (Barris and Button, 2008). Reliability is a key pillar of notational analysis which itself can exhibit large variations amongst observers (inter-observer reliability) (Duthie et al., 2003). Furthermore, reliability of notational analysis can be affected by several variables such as the observer’s level of experience, the total number of observers and the quality of the observer/s view (Duthie et al., 2003).

The Video-Tracking System (VTS) is a semi-automatic technology that significantly increases the accuracy and reliability of performance analysis and technical tagging procedures compared to traditional notational analysis and it is commonly used at the professional level (Barris and Button, 2008). This technology is able to analyse several players simultaneously (via a multiple camera system), providing a pool of data that can support coaching decisions and tactical strategies implemented in soccer (Carling et al., 2008).

The VTS is a technology that records two-dimensional position data (x and y) at high sampling rates (over 25 Hz). The VTS is used to
quantify external load variables during official competitions (Barris and Button, 2008; Buchheit et al., 2014; Carling et al., 2008). More specifically, the VTS is used to collect and analyse time-motion data such as total distance covered (TD), number of changes of direction, acceleration and deceleration activities, as well as time spent at high speed running (Di Salvo et al., 2006). Previous literature has reported that soccer players, on the basis of their roles and game level, usually cover a range of distances from 10 to 13 km (Di Salvo et al., 2009). From a coaching perspective, physical data should be utilised in conjunction with technical and tactical information, in order to have better comprehension of match demands (Beato et al., 2016b; Carling et al., 2008; Christopher et al., 2016). Performance analysis mainly focuses on physical demands and much less attention has been given to technical and tactical factors such as probability of scoring from different locations on the pitch, and other details relating to interactions between players (Hewitt et al., 2016). Moreover, it was previously reported that technical, rather than physical indicators differentiated between league rankings and/or competitive standards in elite soccer. This statement strongly supports the integration of technical and tactical tagging in performance analysis (Carling, 2013).

The VTS is of great interest because it can be used to record external load variables as well as collect additional technical parameters without the necessity of any additional device to be worn (Buchheit et al., 2014). Generally, the VTS has been revealed to have some advantages compared to other monitoring systems as the global position system (GPS) (Di Salvo et al., 2006). Players when wearing GPS devices need to carry extra weight or specialised equipment during official matches, which could hinder their performance and/or be uncomfortable (Cummins et al., 2013). Moreover, GPS validity and reliability could be affected by signal quality, which might be obstructed by the atmosphere and objects in the local environment (e.g., tall buildings and stadium) (Beato et al., 2016a; Cummins et al., 2013). FIFA has recently allowed the use of the GPS during official matches, however, the use of such technology is still limited due to the reasons reported above.

The VTS is one of the most popular choices for performance analysis in elite soccer during official matches (Di Salvo et al., 2006; Stevens et al., 2014). Previous research has however revealed a large amount of variability in the reliability of different VTSs (produced by different companies) (Rawstorn et al., 2014). The production companies responsible for creating a VTS should independently validate their own system (e.g. Prozone Sports Ltd®, OPTA sportsdata) (Liu et al., 2013). Many of the systems on the market nowadays, however, still require scientific validation (Carling et al., 2008). In the recent years, a significant amount of research has been conducted on VTS validity and reliability, and ambiguous data has been discovered during certain actions such as curvilinear and high-intensity movements, as well as in ecological context studies (e.g. during official matches) (Barris and Button, 2008; Carling et al., 2008). One possible explanation for these ambiguous findings can be attributed to human intervention. In the semi-automatic VTS, part of the analysis is done by operators that perform manual corrections. It is well known that these human operators face many difficulties when tracking several players in congestion, inevitably reducing the quality of the system (Barros et al., 2007; Barris and Button, 2008). The human operators level of skill is dependent on their level of training and personal experience, and this is particularly important when technical and tactical counts are being analysed as different personal interpretations (technical definitions) could affect the tagging procedure (Liu et al., 2013).

The Digital.Stadium® VTS is a performance analysis device widely used at national (e.g. Italian Serie A) and international level (Euro 2016). The evaluation of its reliability could be paramount for its application at top-level competitions, as well as in research studies. Technical tagging reliability is a step necessary to validate the Digital.Stadium® VTS, and as reported, many of the contemporary commercial motion analysis systems do not report any validation process (Carling et al., 2008). The main purpose of this study was to assess the reliability (intra-day test-retest) of the VTS by evaluating several soccer-specific technical and tactical variables (i.e. shots, crosses, passes, corners, fouls, etc.) during an official match involving professional soccer players (ecological context).
Methods

Subjects
Thirty professional male soccer players (age 23 ± 5 years, body mass 78.3 ± 6.9 kg, body height 1.81 ± 0.06 m) were enrolled. Goalkeepers were considered in this study as well as substitutions. The experimental protocol was in accordance with the Declaration of Helsinki for the study on human subjects. The ethical committee of the University of Suffolk (Ipswich, UK) approved this study.

Design and research questions
The Digital.Stadium® VTS commercialised by SICS (Sports Information & Consultancy Services, Bassano, Italy) is currently used by a number of professional soccer teams and by the Italian national football federation. Digital.Stadium® is a semi-automatic technology, therefore the evaluation of reliability of such a system is paramount for its utilisation in elite contexts and for future research that will be able to use technical data recorded by this VTS. The aim of this study was to evaluate the reliability (intra-day test-retest) of technical and tactical events that had occurred during an official soccer match (reliability of physical data is not reported in this study). Intra-day reliability takes into account intra-operator reliability, and this was controlled by using the same 4 operators to perform the analyses of the two teams included in this study. This research can support the utilisation of SICS technology in a professional context. The Digital.Stadium® VTS can offer indications about the game style and strategies adopted (technical and tactical analysis) by professional teams during national and international competitions. Moreover, data collected by this VTS might be used for the development of original research in soccer performance analysis.

Experimental protocol and data analysis
Two professional teams were included in this study and the fixtures they played in occurred in 2016. Video analysis was done during a friendly match and data analysis was performed immediately after the game ended (Match) and replicated a week later (Replication). The soccer match analysed in this study took place on a natural grass field playing surface at the Giuseppe-Meazza stadium in Milan. None of the participating players were equipped with external detectors, devices or additional instruments.

SICS Digital.Stadium® version 4.3 is a portable hardware and software system that can be installed at a venue before a match. The installation and configuration procedure usually takes between 30 to 45 min. The main components of the Digital.Stadium® system are:

- Video Recording System: a set of 4 HD IP fixed cameras equipped with special no distortion lenses to record on a PC the match being played. Every part of the soccer field is covered by at least one camera, with some areas of overlay among the different cameras. The 4 cameras are installed on a horizontal bar mounted on a single tripod located in an elevated position in the stadium corresponding to the midfield line.
- File Transfer to Operational Centre module: to transfer the video files recorded by the cameras to the SICS operational offices;
- Automatic Tracking: software that detects the moving objects (players, ball, referees) on the field and calculates the position of each object for each frame with a frequency of 25 Hz;
- Manual Tracking: the movements of each object on the pitch are recorded by a group of operators (usually 4 per match) and the label with the identification of each object (players' names and numbers) is reassigned if needed;
- Technical Tagging Module: to tag each touch of the ball with its technical detail (pass, cross, shot, corner, etc.);
- Data Export Module: to obtain a full set of reports and raw data for each analysed player;
- VideoMatch for Digital.Stadium®: to analyse and present the data obtained by the aforementioned process both from a physical and technical point of view.

Before performing the Automatic Tracking step a process named “Masking” was performed that defined the areas like stairwells and seating areas for supporters; this ensured that movements in these areas were not detected by the software. Masking also consisted of a “Field Calibration” where some relevant points in the recorded videos were located to create a digital
map (e.g., corners, goal area lines, midfield line) and the x, y coordinates in the digital mapping of the field.

Video recordings were analysed by the computer system and then corrected by a human operator (semi-automatic system). Four operators were involved in the first analysis (Match) and the same four operators replicated the analysis (a week later) on the same match (Replication). The analysis took into account several technical variables deeply studied in soccer science such as passes, duels won, ball regain/lost, cross, shots, assists, key passes, number of dribbles, fouls and cards. A clear and deeply operational definition of such technical variables was previously reported by OPTA (Liu et al., 2013). Some physical parameters such as TD and total distance in possession of the ball were recorded in the study. Mean metabolic power (w·kg^{-1}) was also reported in order to offer comparison between our data and previous publications (Osognach et al., 2010). Physical importance of metabolic power, as well as of power actions in team sports had been previously reported (Osognach et al., 2010; Zamparo et al., 2015). The indirect estimation of the metabolic power using the rationale that accelerated running on a flat terrain is energetically analogous to uphill running at constant speed (Minetti et al., 2002):

\[
EC (J \text{ kg}^{-1} \cdot \text{m}^{-1}) = (155.4 \text{ ES}^5 - 30.4 \text{ ES}^4 - 43.3 \text{ ES}^3 + 46.3 \text{ ES}^2 + 19.5 \text{ ES} + 3.6) \text{ EM}
\]

where EC is the energy cost of accelerated running on grass, EM is the equivalent mass and ES is the equivalent slope (Osognach et al., 2010).

The intra-day test-retest reliability of metabolic power was not reported in this study, because a systematic analysis of all the physical performance variables (e.g. TD, high intensity running activity, metabolic variables) recorded by this VTS was published previously (Beato and Jamil, 2017).

Statistical analysis

Data are presented as means ± SD. A Shapiro-Wilk test was performed for the evaluation of normality (assumption) for statistical distribution. Log transformation was done for non-normal data. The Pearson Correlation Coefficient reported in this study as the variance explained (R^2) was used to determine the relationship between matches’ data and thus the validity of the variable (Hopkins, 2000). Statistical significance was set at p < 0.05. Absolute and relative reliability of data (between Match and Replication) was assessed using the typical error of measurement expressed as a percentage coefficient of variation (CV) and an interclass correlation (ICC). Differences between Match and Replication were reported as a mean of change with confidence intervals (CI 90%) (Hopkins, 2000). Statistical analysis was performed using SPSS (SPSS Statistics 20.0) for Mac OS X Yosemite.

Results

Reliability of Match and Replication is reported in Table 1 as CV, ICC and mean of change with CI 90%. This study reported a near perfect relationship between Match and Replication variables. R^2 coefficients were highly significant for each of the performance variables (p < 0.001).

In this study over 3500 events were recorded during the technical tagging process. Furthermore, we reported a single value for each variable that did not show any differences between Match and Replication, contrariwise, we reported both Match and Replication data for the variables that showed differences. More specifically, the following variables were recorded: shot = 1.35 ± 0.6, dribbles = 1.9 ± 1.0, crosses = 2.19 ± 1.27 (Match) and 2.11 ± 1.26 (Replication), key passes = 3.00 ± 2.26 (Match) and 3.05 ± 2.29 (Replication), passes = 35.21 ± 20.24 (Match) and 35.62 ± 20.18 (Replication), assists = 1.4 ± 0.70, corners = 0.6 ± 0.4, free kicks = 0.9 ± 0.6, ball recovery = 5.7 ± 3.4, lost balls = 6.3 ± 3.5, duels won 4.9 ± 3.5, fouls done 1.9 ± 1, and fouls received 1.5 ± 1.2 (Match) and 1.7 ± 1.3, cards = 0.2 ± 0.3.

Mean TD was 8304 ± 3760 m and 8298 ± 3754 m in Match and Replication, respectively (TD, as well as every other physical variable, took in consideration GKS and substitutions). Total distance covered in possession was 193.5 ± 112.2 m and 196.6 ± 112.5 m in Match and Replication, respectively. Metabolic power was 10.64 ± 2.13 and 10.62 ± 2.11 in Match and Replication, respectively.
Table 1

Reliability data recorded during Match (30 players) and after its Replication (intra-day test retest).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean of change (CI 90%)</th>
<th>Typical Error as CV (%)</th>
<th>ICC</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shots</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dribbles</td>
<td>0.18</td>
<td>0.2</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td>Crosses</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Key passes</td>
<td>0.05 (-0.03; 0.13)</td>
<td>0.2</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>Passes</td>
<td>0.39 (0.02; 0.77)</td>
<td>0.8</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>Assists</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Corners</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Free kicks</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Balls recovered</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Lost balls</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Duels won</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Fouls done</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Fouls received</td>
<td>0.19 (0.01; 0.36)</td>
<td>0.3</td>
<td>0.96</td>
<td>0.89</td>
</tr>
<tr>
<td>Cards</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>TD (m)</td>
<td>-6.4 (-18.0; 5.2)</td>
<td>0.26</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>Total distance with the ball (m)</td>
<td>2.7 (-0.03; 5.8)</td>
<td>0.7</td>
<td>0.99</td>
<td>0.98</td>
</tr>
</tbody>
</table>

$ICC = \text{interclass correlation}$, $CV = \text{coefficient of variation}$, $CI = \text{confidence intervals}$, $R^2 = \text{variance}$, $TD = \text{Total distance}$. 

Figure 1

*Digital.Stadium® VTS.*
Discussion

In this study, we analysed Digital.Stadium® VTS reliability (intra-day test-retest) in the measurement of soccer specific technical tagging (e.g. cross, passes, number of dribbling) and external load variables during an official match with professional soccer players. Digital.Stadium® is a technology commonly used in professional soccer, and its reliability is crucial for scientific acknowledgment and credibility (Carling et al., 2008). Operators recorded over 3500 events, and this value is in line with the number of events reported previously by OPTA (Liu et al., 2013). This study reports high levels of absolute and relative reliability (CV and ICC), as well as a small mean of change (between Match and Replication) in every variable (Table 1). During our protocol, this VTS was tested in an ecological context, where reliability could be affected by human intervention (Barris and Button, 2008). Generally, VTS technologies report common limitations such as difficulty in tracking players in congested areas of the pitch (Barros et al., 2007). Furthermore, high intensity, high speed and power data are associated with a low level of reliability (Buchheit et al., 2014; Stevens et al., 2014). Our results support the reliability of this VTS for both technical and tactical variables that were analysed, as close to perfect reliability values were reported.

As previously stated human error can affect reliability data, particularly as performance analysis and technical counts are associated with an analyst’s background and previous experience (Liu et al., 2013). Physical performance analysis is largely performed by an automatic system, where only limited corrections associated with player tracking in congestion are conducted manually. Previous studies have reported a 95% automatic tracking rate (DVideo) (Barros et al., 2007; Carling et al., 2008). Previous literature has also reported contrasting results during curvilinear and high-intensity movements (Barris and Button, 2008; Carling et al., 2008), contrariwise, other studies reveal good test-retest reliability with a CV equivalent to 1% (Mohr et al., 2003). Physical data reported in this study mirrors the CV and reliability scores reported in previous literature (CV = 1%). Furthermore, TD and distance covered in possession showed a close to perfect ICC (0.99). Systematic analysis about reliability scores of kinematic variables has been recently published and it supports our results (Beato and Jamil, 2017). During physical performance analysis only limited human interventions are necessary, contrariwise to the technical tagging process that is mainly manual and thus subject to major human errors. After test-retest reliability of technical tagging analysis, it was possible to state that meaningless differences were found in all the match actions (ICC = 0.98-099). Our results relating to Digital.Stadium® VTS reliability are in line with reliability scores previously reported by OPTA Sportsdata (Liu et al., 2013).

Performance analysis offers an abundance of information about team and players’ position workloads, transient fatigue, technical counts and tactical principles (Carling, 2013). Much previous scientific literature focused on physical performance analysis, while much less attention has been afforded to technical and tactical actions even though these factors can be better predictors of success in soccer compared to purely physical variables (Liu et al., 2016). Previous studies reported that the player work rate was influenced by opponents’ quality and tactical strategies (Carling et al., 2008; Carling, 2013). Moreover, it was reported that a greater amount of high intensity distance was covered against better opponents than lower level teams (Rampinini et al., 2007). This evidence is supported by a recent article that reports “Tier A clubs” (first quartile) as being more technically and tactically efficient in possession of the ball than their lower tier counterparts, with a resultant reduction in non-productive high-intensity efforts (Bradley et al., 2016). A recent study analysing matches played at the international tournament Euro 2012® reported results that supported the importance of two different tactical situations for scoring goals: transition play after losing possession of the ball and the general efficiency in attacking play (Winter and Pfeiffer, 2016). As reported in the same article, technical tagging and tactical analysis can be used to evaluate different game styles in order to produce goal scoring opportunities (Winter and Pfeiffer, 2016).

This study has some limitations. A larger sample size may be necessary to better evaluate the Digital.Stadium® VTS reliability. Moreover, as reported above, technical tagging reliability is associated with operator experience and this
study was performed involving SICS analysts. Different reliability results could be found when involving less experienced staff or using other VTS, therefore we recommend using the results of this study as evidence of the reliability of the Digital.Stadium® VTS and to avoid cross comparison with other VTS brands.

Technical tagging reliability is a necessary step when validating the Digital.Stadium® VTS, and as reported, many of the contemporary commercial motion analysis systems do not report any validation process (Carling et al., 2008). The evaluation of VTS reliability is very important to its application at a top-level context, as well as in research studies where a high standard of quality is required (Liu et al., 2013). Future studies will also be able to analyse the validity of such technology alongside other gold standard measures such as time gates and laser and radar guns, during specific soccer circuits (Akenhead et al., 2014; Rampinini et al., 2015). In addition, future studies could evaluate the technical tagging process in different climatic conditions (variations in light) that can affect data reliability and accuracy (Carling et al., 2008; Carling, 2013).

Conclusions

The findings reported in this study have an important practical application since they underline that all data reported by the Digital.Stadium® VTS show high levels of absolute and relative reliability. This study reports meaningless differences between match and its replication (intra-day reliability). We may conclude that the semi-automatic process behind the Digital.Stadium® VTS is more than capable of recording technical tagging data accurately. Sport scientists should consider the accuracy of technical tagging and performance variables of this VTS, as well as other similar technologies for future scientific analyses and studies. In conclusion, Digital.Stadium® VTS can offer accurate indications about the game style and strategies adopted (technical and tactical analysis) by professional teams during national and international competitions. Moreover, data collected by this VTS might be used for the development of new research in performance analysis.

Acknowledgements

The authors thank SICS and its performance analysis staff for their precious help in data collection.

References

Beato M, Coratella G, Schena F. Brief review of the state of art in futsal. J Sport Med Phys Fit, 2016b; 56: 428–432
Carling C. Interpreting physical performance in professional soccer match-play: Should we be more

Corresponding author:

Dr Marco Beato,
Department of Science and Technology,
University of Suffolk, Ipswich, UK
E-mail: M.Beato@uos.ac.uk