International Journal of Computer Science in Sport

Volume 11/2012/Edition 2

TABLE OF CONTENTS

Editorial

Arnold Baca
Department of Biomechanics, Kinesiology and Applied Computer Science,
ZSU, University of Vienna

Dear readers:

Welcome to the summer 2012 issue of the **International Journal of Computer Science in Sport (IJCSS)**.

Two research papers, two scientific reports and one project report have been included within this issue.

Nicole Bandow, **Kerstin Witte** and **Steffen Masik** present a virtual test environment for performing reaction tasks sports. The evaluation of the developed procedure is accomplished on the basis of a comparative test.

The study by **Tade Souaiaia** and **Jonas Mureika** demonstrates a five-parameter model approximating the environmental effects on long jump performance of world class athletes.

Peter O'Donoghue compares two sets of two predictive models for the Rugby World Cup 2011 integrating multiple linear regression techniques.

Chueh-Wei Chang, **Yi-Po Wu** and **Hua-Wei Lin** discuss the design and implementation of an animation assisted sport simulation system for general baseball "cover, relay and cutoff play" training.

Jürgen Perl and **Stefan Endler** report on their current results regarding the PerPot simulation model, allowing the determination of the individual anaerobe threshold (IAT).

If you have any questions, comments, suggestions and points of criticism, please send them to me.

Enjoy the summer!

Arnold Baca, Editor in Chief
University of Vienna, arnold.baca@univie.ac.at

Development and Evaluation of a Virtual Test Environment for Performing Reaction Tasks

Nicole Bandow[1], Kerstin Witte[1] & Steffen Masik[2]

[1]*Department of Sport Science, Otto-von-Guericke University Magdeburg*

[2]*Fraunhofer Institute for Factory Operation and Automation Magdeburg*

Abstract

Virtual reality offers many advantages for standardized experimental setups as well as for manipulating selected parameters. This study describes the development of a three dimensional virtual environment for sports as well as the evaluation of its effectiveness by means of a comparative test. The test consists of measuring the participants' (n=33) mean simple reaction times ($\overline{RT}$) of an appearing ball in a real, a two dimensional and the developed virtual environment. To assess the participant's sensation of reality in the two dimensional and virtual environment, a short-questionnaire was used. Simple reaction times were measured by accelerometers fixed onto the participant's wrist. The results of the ANOVA and post-hoc analysis (Bonferroni) showed a significant difference (p<0.001) of $\overline{RT}$ between each environment. $\overline{RT}$ between real environment (188ms (±37ms)) and virtual environment (286 ms (±69 ms)) was 53% lower than between real and two dimensional environment (373 ms (±68 ms)). Results of the questionnaire showed that the majority of participants had a higher sensation of reality in the virtual environment than in the two dimensional environment. This leads to the conclusion that the virtual environment evokes a more realistic behavior than the two dimensional environment which is important for research and training in sports.

KEYWORDS: VIRTUAL ENVIRONMENT, CAVE, SPORTS, REACTION TIME

Introduction

This paper describes the development of a three dimensional virtual environment for research and training in sports. Although there is growing interest in using virtual reality in sports, only few studies use virtual reality technology for research and training of anticipation and perception (Tanaka, Hasegawa, Kataoka & Katz, 2010).

Virtual reality (VR) technology offers many advantages especially for research in sports. The main advantages of computer simulated environments are standardized conditions with high ecological validity, experimental controlling, feasibility of manipulation and repeatability (Armbrüster, 2007; Tanaka et al., 2010). The most important and differentiating advantage of VR is stereovision providing spatial information. Spatial information in sports is a necessary requirement to successfully assess game or combat situations (Panchuk & Vickers, 2009). Furthermore it is possible to create presence (the feeling of being in the virtual environment) by providing stereoscopic view jointly with a high level graphics VR. A high presence evokes the feeling of being able to act in a real environment. Moreover, this feeling elicits behavioral realism which is an important aspect for research in sports (Vignais, Bideau, Craig, Brault, Multon, Delamarche & Kulpa, 2009). Main advantages of VR for training are not being dependent on other athletes or on required spaces and environments (i.e. training establishment, soccer field) (Göbel, Geiger, Heinze & Marinos, 2010).

There have been very few studies with the aim of developing VR based training facilities for improving sports skills such as in archery, baseball, basketball, karate and table tennis (Göbel et al., 2010; Komura, Kuroda & Shingawa, 2002; Rusdorf, Brunnett, Lorenz & Winkler, 2007; Tanaka, 2009; Zhang & Wang, 2011).

The following studies show that the use of VR in research has only increased over the past decade. Dessing and Craig (2010) use VR to study the behavior of soccer goalkeepers. They examined whether goalkeepers misjudged the landing position of the ball when spin was added to it by using reproducible ball trajectories. They argue that presenting the same reproducible ball trajectories across several trials would not have been possible through video presentation or in reality (Craig, Berton, Rao, Fernandez & Bootsma, 2006). Bideau et al. (2010) and Vignais et al. (2009) use VR to analyze sport performance in handball. They examine whether the goalkeeper's choices, i. e. how to react to an approaching ball, are influenced by perception. The results show that good realistic virtual environments evoke good perception and realistic behaviors.

Up to now, most studies or training methods still use two dimensional presentations. Mori (2002) uses video presentations to examine relevant cues for anticipating karate attacks. Others use similar methods such as computer screens or video projections to analyze anticipation and perception in soccer, tennis, basketball, field hockey and squash. Interestingly, a few studies acknowledge the limitation of two dimensional presentations in regards to the lack of spatial information and propose using three dimensional presentations (Abernethy, 1990; Aglioti, Cesari, Romani & Urgesi, 2008; Clatworthy, Holder & Graydon, 1991; Hagemann, Schorer, Canal-Bruland, Lotz & Strauss, 2010; van der Savelsbergh, 2002; Williams, Kamp & Ward, 2002; Williams, Ward, Knowles & Smeeton, 2002).

The aim of the study at hand is twofold: (a) Developing a three dimensional virtual environment for research in sports. (b) Verifying it in terms of the degree of realism it evokes. In order to develop a VR environment for sports high demands need to be fulfilled. On the one hand, these are sports specific demands such as mobility, naturalness of the presented objects and characters, individual adjustable information (i. e. joint angles) as well as freedom of feedback. On the other hand, technological demands have to be taken into account which includes creating animated virtual characters based on natural motions (no latencies), adapting the virtual character to the environment and finally providing a good graphical quality of all visual components. Providing a good graphical quality depends on the graphical level of detail of the displayed virtual environment. It is important to note that a higher level of detail always leads to an increase in CPU run time. If all these demands are met, it is possible to evoke a high level of presence in the participant which is important for research in sport as mentioned before (Pronost, Multon, Li, Geng, Kulpa & Dumont, 2008).

The effectiveness of the developed virtual environment is tested by a comparative study measuring the athlete's reactions to a special event in different environments (Katz, Parker, Tyreman & Levy, 2008). In this study, the participant's simple reaction time (RT) to the appearance of an approaching ball is measured in a real, a two dimensional and the developed three dimensional virtual environment. Based on the fact that virtual environments cause a high level of presence and therefore evoke realistic behavior, it is assumed that the simple RTs in the three dimensional virtual environment are more similar to those in the real environment than the simple RTs in the two dimensional environment (Vignais, et al., 2009). To assess the participant's subjective impression of the degree of realism of the two dimensional and virtual environment, a short-questionnaire was used.

Methods

Development of a Three Dimensional Virtual Environment

To create the three dimensional virtual environment the following steps had to be accomplished: Selection of an appropriate VR technology, development of a virtual model and a three dimensional virtual environment, recording of motion capture data to animate the virtual model, merging of the motion capture data with the virtual model, and embedding the animated model in the three dimensional virtual environment (Figure 1). The following describes the procedure more precisely.

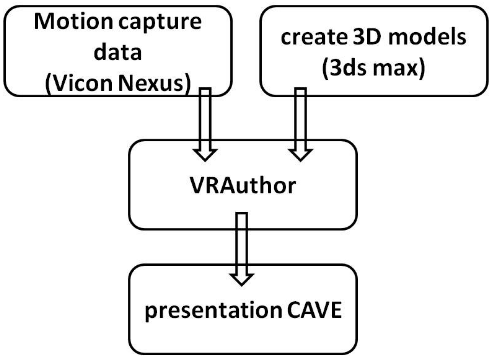


Figure 1. Schematic model of the development of the three dimensional virtual environment: creation of a virtual model and a three dimensional virtual environment, recording of motion capture data, merging of the motion capture data with the virtual model in the VRAuthor, and embedding the animated model in the three dimensional virtual environment for presentation in the CAVE.

The CAVE (Cave Automatic Virtual Environment) was selected as the most appropriate VR technology for our demands. Its four projection screens (left, right, front, floor), each 2.30m x 2.30m of size, offer a high level of immersion while giving sufficient space to move. Eight synchronized video projectors (JVC D-ILA - DLA-SX21S, 1400x1050px) driven by a PC-cluster of ten computers (two master PCs driving eight PCs controlling eight video projectors) were used to provide a stereoscopic view on all four projection screens. Polarization glasses were worn to provide stereovision. To adjust the virtual environment according to the viewer, an integrated optical tracking system with four cameras (ART, Weilheim, Germany) was used. Markers were fixed onto the polarization glasses to track the participant.

The three dimensional virtual environment and the virtual model, both based on computer graphics models, were created with the modeling software 3ds Max (Autodesk Inc., San Rafael, USA). To increase realism, the virtual ball was mapped with the texture of an official soccer ball. The virtual environment showed a black wall in front, bright walls on both left and right sides and a bright floor (Figure 2a).

In order to provide ball flights which were similar to the real and two dimensional environment, the virtual soccer ball was animated by real motion data. Motion data was collected by capturing the ball's flight trajectory by the optical motion capture system VICON (Oxford Metrics, Oxford, UK) with twelve MX13 cameras. To capture the ball six markers were fixed onto it, in a dice-like style.

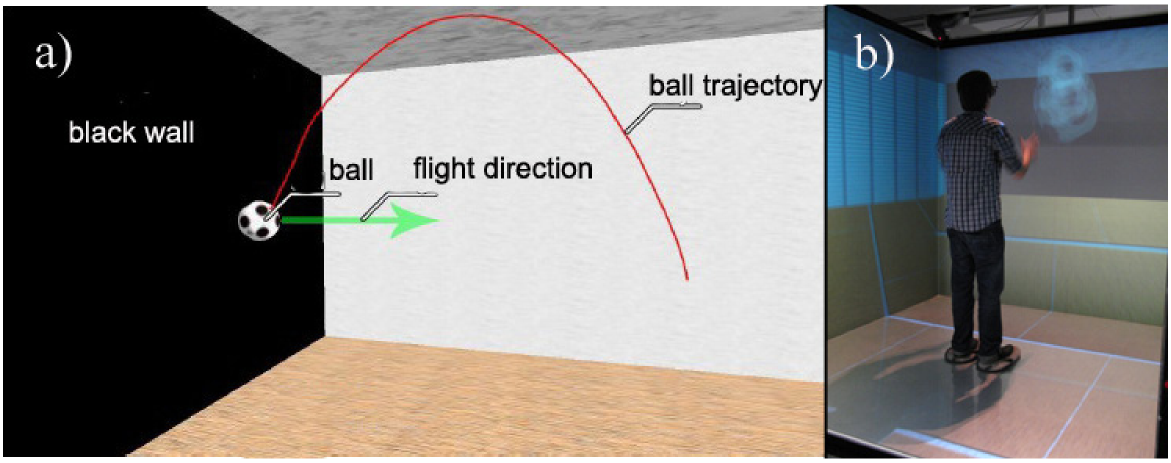


Figure 2. a) Schematic illustration of the virtual environment showing the virtual ball and its flight path. b) A participant in the CAVE reacting to an approaching ball.

The virtual environment and the animated virtual soccer ball were then transformed into the VRAuthor which is an author software based on the 3D graphics language OpenGL, and on OpenSceneGraph, an open source high performance 3D graphics toolkit that displays and simulates visual contents in virtual worlds. The VRAuthor offers the possibility to extent functions individually, and to present the visual contents through different VR technologies (Figure 2b).

Testing the Three Dimensional Virtual Environment

Research Designs

To assess the effectiveness of the developed virtual environment a comparative study was conducted. The study involved measuring participant's simple RT during a ball flight test in a real, a two and in a three dimensional virtual environment. At last the participants had to complete a short-questionnaire assessing their sensation of reality in the two and three dimensional virtual environment.

In the real test environment, participants had to react as fast as possible to a soccer ball flying towards them appearing out of a black wall 4 m in front of them. It was the same person throwing the ball throughout all tests in the real environment as well as for the video footage of the two dimensional environment, and the motion data for the three dimensional environment. The participants' movements as well as their right wrists were recorded by a high-speed camera (Basler Pilot 640-GC210; 640x412 px; 200Hz) positioned 1.15m to the right and 3m behind them (Figure 3a).

In the two dimensional environment participants had to react as fast as possible to approaching soccer balls that were projected onto a 2.25 m x 2.20 m screen. The video footage was back-projected (SAMSUNG SP-D400; 1024x768 px) from 6.50 m distance while the participants stood on the other side of the projection screen (Figure 3b). The video footage was created by filming ball throws in the real environment. The participants' movements as well as their right wrist were filmed by two high-speed cameras. The cameras were positioned 2.50 m behind the participants, one 1.15 m to the right and one 1.15 m to the left of them.

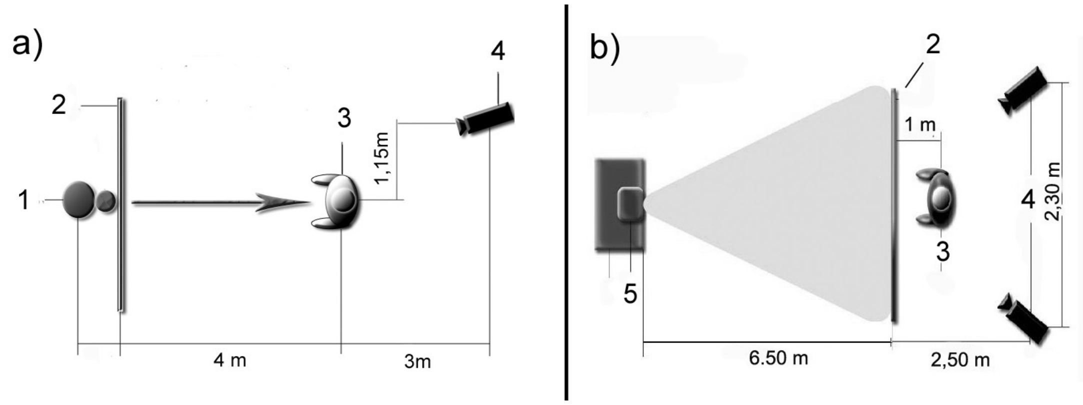


Figure 3. a) Real environment (1: ball; 2: black canvas (2.20 m x 2.25 m); 3: participant; 4: high-speed camera). b) Two dimensional environment (2: projection screen (2.20 m x 2.25 m); 3: participant; 4: high-speed cameras; 5: projector).

To conduct the test in the three dimensional virtual environment (i. e. CAVE), the tracking system had to be adjusted to each participant in the first place. The test procedure was the same as in the real and two dimensional environment. The participant had to react as fast as possible to approaching virtual soccer balls. Two high-speed cameras recorded the participants' movements and their right wrists. One camera was positioned 2.50 m behind the participants and 1.15 m to the left and the other camera 1,15 m in front of the participants in 2.55 m height (above the front projection screen) (Figure 4).

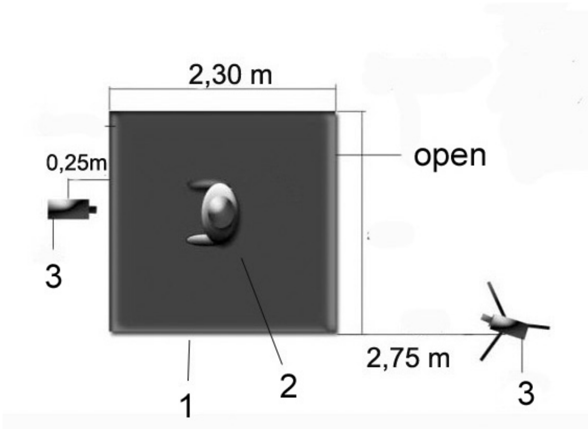


Figure 4. Three dimensional virtual environment (1: CAVE (2.30 m x 2.30 m); 2 participant; 3 high-speed cameras).

Thirty-three males were selected for the study. Participants were sports students, right handed and in average 24 (±3) years of age. All participants had to accomplish ten trials consecutively in each environment. Each execution was followed by a 5 second break. The first ten trials in each environment were conducted to familiarize the participants with the environment. The participants were instructed to react as fast as possible to the approaching ball. Since catching the ball was only possible in the real environment, but not in the two and three dimensional environment, it was not part of the analysis.

To assess the participants' subjective impression regarding their sensation of reality in the two dimensional and three dimensional virtual environment, they had to answer two questions in a short-questionnaire after the tests: '*How similar is the 2D environment compared to the real environment*' and *'How similar is the 3D 'virtual environment compared to the real environment'*. All participants were aware that the meaning of *similarity* here was related to the level of reality of each environment. A 6-point Likert scale was chosen as response mode. (Kirchhoff, Kuhnt, Lipp, & Schlawin, 2010) The answering options were: *not at all, hardly similar, little similar, quite similar, very similar, and identical.*

Determination and Analysis of Mean Reaction Times ($\overline{RT}$)

The reactions, defined as the simple RT (ms) from first visibility of the ball until the first physical reaction, were measured by an accelerometer (Myon; three dimensional; 1000Hz). The accelerometer was fixed one cm above the processus styloideus ulnae, onto the M. flexor digitorium superficialis of the right arm. Only the acceleration of the radial-ulna-plane was analyzed presenting the main movement direction of the reaction.

Acceleration and video data were collected synchronously for each test environment by means of the Nexus 1.3 software (VICON, Oxford Metrics, Oxford, UK). Acceleration data were smoothed by a moving average (20 samples) to minimize the noise before analysis.

To determine the simple RT, specifically the time differences between the ball becoming visible and the participants' first physical reactions, an analysis of the video footage and the acceleration data were necessary: Firstly, the ball's appearance time was determined by the bias of video footage. Secondly, the beginning of the participant's physical reaction was determined as the point when the acceleration exceeded the threshold of 20% of the amplitude's maximum of each deflection. The threshold was defined in a pre-analysis where the acceleration data was compared with the video footage. It was found that the beginning of the physical reaction based on the acceleration data was in average one or two frames (5-10 ms) before the visible reaction. The threshold was then set to 20% to ensure that smaller movements did not affect the identification of the beginning of the physical reaction (Figure 5).

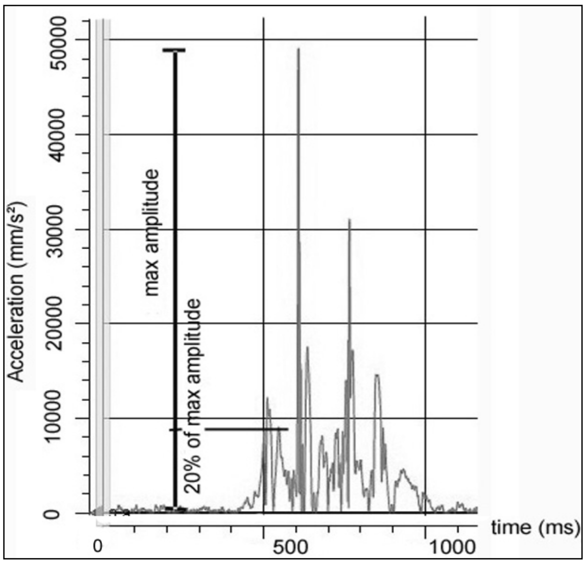


Figure 5. Generic acceleration waveform of a single reaction showing the 20% threshold of the amplitude's maximum.

For statistical analysis the $\overline{\mathrm{RT}}$ (ms) and standard deviations for each participant in each environment (3x10 trials) as well as the $\overline{\mathrm{RT}}$ over all participants for each environment (3x33 trials) were calculated. A one-way analysis of variance (ANOVA), with environment as the between-factor, was used to study the effects of the environment on the participants' reactions.

Results

Mean Simple Reaction Times

The findings of the ANOVA ($F_{(2,96)}$=79.9; p<0.001) showed significant differences in mean reaction time for each environment. $\overline{\mathrm{RT}}$ for in the real environment were 188 ms (±37 ms), for the two dimensional environment 373 ms (±68 ms), and for the three dimensional virtual environment 286 ms (±69 ms) (Figure 6). A post-hoc analysis (Bonferroni) revealed significant differences between *(1) real environment vs. 3D virtual environment (p<0.001), (2) real environment vs. 2D environment (p<0.001), and (3) 3D virtual environment vs. 2D* *environment (p<0.001)*. Remarkably, the difference between real and two dimensional environment (185 ms) was almost twice as high as between real and three dimensional environment (98 ms).

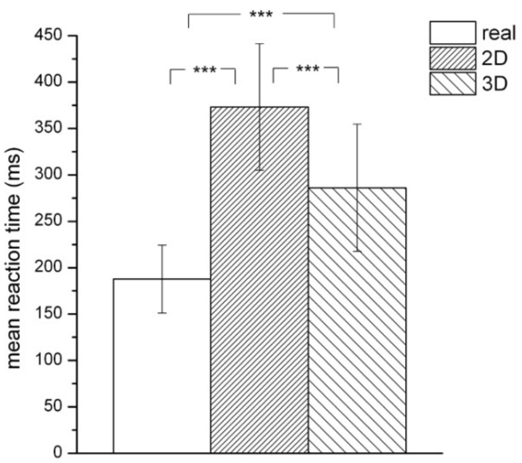


Figure 6. $\overline{\mathrm{RT}}$ (*n*=33) for each environment: (1) real environment vs. 3D virtual environment (p<0.001), (2) real environment vs. 2D environment (p<0.001), and (3) 3D virtual environment vs. 2D environment (p<0.001).

Sensation of Reality

The findings of the short-questionnaire showed that most participants had the impression that the three dimensional virtual environment rather than the two dimensional environment was more similar to the real environment in terms of sensation of reality.

Table 1. Frequency and percent of answers (n=33) to the question 'How similar is the 2D environment compared to the real environment?'

question	How similar is the 2D environment compared to the real environment?	
answers	frequency	%
not at all	2	6,1
hardly similar	19	57,6
little similar	6	18,2
quite similar	4	12,1
very similar	2	6,1
identical	-	-

Most participants (81%) responded to '*How similar is the 2D environment compared to the real environment?*' with choosing the answers representing the lowest sensation of reality ('not at all', 'hardly similar', 'little similar'). Only 21% answered with 'fairly similar' and 'very similar' (representing a higher sensation of reality). Nobody responded with 'identical' (Table

1).

Table 2. Frequency and percentage of answers (n=33) in regards to the question '*How similar is the 3D virtual environment compared to the real environment?*'

question	How similar is the 3D virtual environment compared to the real environment?	
answers	frequency	%
not at all	-	-
hardly similar	-	-
little similar	5	15,2
quite similar	16	48,5
very similar	12	36,4
identical	-	-

In contrast, only 15.2% of the participants answered the question '*How similar is the 3D environment compared to the real environment?*' with 'little similar' and none with 'not at all' or 'hardly similar'. Therefore 85% of the participants chose the answers representing a higher sensation of reality (Table 2).

Discussion

The results of the comparative test show significant differences in $\overline{RT}$ for each environment. The $\overline{RT}$ in real environment (188 ms) confirms the general statement from Wollny (2010) that visual simple RTs range between 100-350 ms depending on the participants' level of expertise. These findings can therefore be seen as a validated benchmark for further discussions. The results of the ANOVA prove a significant difference of $\overline{RT}$ between each environment. Remarkably is the difference of $\overline{RT}$ (98 ms) between real and three dimensional virtual environment compared to the difference between real and two dimensional environment (185 ms) which is 53% lower. This finding confirms the hypothesis that the simple RTs in the three dimensional virtual environment are more similar to those in the real environment than the simple RTs in the two dimensional environment. Since the similarity of reaction in the three dimensional virtual environment is closer to reality than in the two dimensional environment, it can be assumed that the presentation in the three dimensional environment is also more similar to reality. These findings strengthen the successful development of a three dimensional virtual environment that is able to evoke realistic behavior within the participants (Katz et al., 2008). Additionally, the findings of the short-questionnaire prove that the sensation of reality is higher in the three dimensional virtual environment than in the two dimensional environment. Moreover, these results show that the three dimensional environment can provoke more realistic reactions caused by more realistic presentations than the common two dimensional presentations. This finding is important for the application of three dimensional environments in research and training. Only if athletes react and behave authentically this technology can be used as research and training method in sports (Vignais et al., 2009).

Nonetheless, there is a lack of statistical proof of similarity between the real and three dimensional environment. The reasons therefore are considered consecutively. Firstly, the visual presentation of the virtual environment and models need improvement in regards to visual detail. Secondly, the adjustment of the distance between moving objects, i. e. the virtual ball, and the participants has to be enhanced. Some participants had the impression that the ball was flying to close towards them. Thirdly, the participant's individual perception has to be examined prior to the test. Studies have revealed that some people are incapable of seeing three dimensional virtual environments based on physical issues. Fourthly, the amount of trials to familiarize to stereovision was too small. A test to verify the participant's capability of stereovision as well as eye vision was not conducted but should be implemented in future examinations.

Conclusion

This study offers a first attempt at using VR technology for examining reactions. The findings show that the three dimensional virtual environment evokes a more realistic behavior within the participants than the two dimensional environment. It can be assumed that the reason therefore is due to spatial information and the feeling of presence arising thereby. Furthermore VR technology offers the possibility to create a reproducible and individually manipulative test environment. The developed three dimensional virtual environment can be seen as a start providing good opportunities for further research in anticipation and perception as well as for training in sports.

Acknowledgements

This study was financially supported by the German Federal Institute for Sport Sciences (Bundesinstitut für Sportwissenschaft – BISp, Bonn, Germany, project no. IIA1 – 071504/10).

References

Abernethy, B. (1990). Anticipation in squash: Differences in advance cue utilization between expert and novice players. *Journal of Sports Sciences*, *8*, 17–34.

Aglioti, S. M. Cesari P. Romani M. &. Urgesi C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, *11*(9), 1109–1116.

Armbrüster, C. (2007). *Virtuelle Realität in der experimentellen Psychologie: Forschungsmethode versus Forschungsgegenstand; Untersuchungen aus den Bereichen Wahrnehmung und Psychomotorik [Virtual reality in experimental psychology: research method versus research topic; Examinations of the field of perception and psycho-motorics]*. Hamburg: Kovač.

Bideau, B., Kulpa, R., Vignais, N., Brault, S., Multon, F., & Craig, C. (2010). Using Virtual Reality to Analyze Sports Performance. *IEEE Computer Graphics and Applications*, *30*(2), 14–21.

Clatworthy, R., Holder, T. &. Graydon, J. (1991). An investigation into the anticipation of field hockey penalty flicks using an ecologically valid temporal occlusion technique. *Journal of Sports Sciences*, *9*(4), 439–440.

Craig, C. M., Berton, E., Rao, G., Fernandez, L., & Bootsma, R. J. (2006). Judging where a ball will go: the case of curved free kicks in football. *Naturwissenschaften*, *93*(2), 97–101.

Dessing, J. C. &. Craig, C. M. (2010). Bending It Like Beckham: How to Visually Fool the Goalkeeper. *PLoS ONE*, *5*(10), 1–8.

Göbel, S., Geiger, C., Heinze, C., & Marinos, D. (2010). Creating a virtual archery experience. In *Proceedings of the International Conference on Advanced Visual Interfaces - AVI '10* (p. 337). ACM Press.

Hagemann, N., Schorer, J., Canal-Bruland, R., Lotz, S., & Strauss, B. (2010). Visual perception in fencing: Do the eye movements of fencers represent their information pickup? *Attention, Perception & Psychophysics*, *72*(8), 2204–2214.

Katz, L., Parker J., Tyreman H., &. Levy R. (2008). Virtual Reality. In P. Dabnichki & A. Baca (Ed.), *Computers in Sport* (pp. 3–39). Southampton [u.a.]: Witpress.

Kirchhoff, S., Kuhnt, S., Lipp, P., & Schlawin, S. (2010). *Der Fragebogen: Datenbasis, Konstruktion und Auswertung [The questionnaire: Data base, design and analysis]* (5. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Komura, T., Kuroda, A., & Shinagawa, Y. (Eds.) 2002. *NiceMeetVR: Facing Professional Baseball Pitchers in the Virtual Batting Cage*. New York, NY: ACM.

Panchuk, D., & Vickers, J. N. (2009). Using spatial occlusion to explore the control strategies used in rapid interceptive actions: Predictive or prospective control? *Journal of Sports Sciences*, *27*(12), 1249–1260.

Pronost, N., Multon, F., Li, Q., Geng, W., Kulpa, R., & Dumont, G. (2008). Interactive Animation of Virtual Characters: Application to Virtual Kung-Fu Fighting. In *Proceedings of the 2008 International Conference on Cyberworlds, CW 2008* (pp. 276–283).

Rusdorf, S., Brunnett, G., Lorenz, M., & Winkler, T. (2007). Real-Time Interaction with a Humanoid Avatar in an Immersive Table Tennis Simulation. *IEEE Transactions on Visualization and Computer Graphics*, *13*(1), 15–25.

Tanaka, K. (2009). Virtual Training System Using Visual Feedback for Sport Skill Learning. *International Journal of Computer Science in Sport*, *8*(2), 7–18.

Tanaka, K., Hasegawa, M., Kataoka, T., & Katz, L. (2010). The Effect of Self-Position and Posture Information on Reaction Time. *International Journal of Computer Science in Sport*, pp. 4–14.

van der Savelsbergh, G. J. P. Williams A. M., Kamp J., &. Ward P. (2002). Visual search, anticipation and expertise in soccer goalkeepers. *Journal of Sports Sciences*, *20*, 279–287.

Vignais, N. Bideau, B., Craig, C., Brault, S., Multon, F., Delamarche, P., &. Kulpa, R. (2009). Does the level of graphical detail of a virtual handball thrower influence a goalkeeper's motor response? *Journal of Sports Science and Medicine*, (8), 501–508.

Williams, A. M., Ward, P., Knowles, J. M., &. Smeeton, N. J.(2002). Anticipation Skill in a Real-World Task: Measurement, Training and Transfer in Tennis. *Journal of Experimental Psychology: Applied*, *8*(4), 259–270.

Wollny, R. (2010). *Bewegungswissenschaft: Ein Lehrbuch in 12 Lektionen [Kinesiology: A textbook in 12 lectures].* Aachen: Meyer und Meyer.

Zhang, L. &. Wang. L. (2011). VR-Based Basketball Movement Simulation. *Lecture notes in computer science*, (6530), 240–250.

A Mathematical Model of the Environmental Effects on Long Jump Performance of World Class Athletes

Tade Souaiaia[1], Jonas Mureika[2]

[1]*Department of Computational Biology and Bioinformatics, University of Southern California*

[2]*Department of Physics, Loyola Marymount University*

Abstract

A model is presented that combines experimental data on the relationship between approach velocity and takeoff angle with a five-parameter model to approximate environmental effects on approach velocity and in-flight travel. Results indicate that wind speed provides the greatest influence on jump distance, followed by air density which itself is a product of altitude, temperature, air pressure, and humidity. Local fluctuations in the Earth's surface gravitational field strength are shown to have a slight effect on performance. Previously, analysis attributed the majority of performance increase to faster approach speed and takeoff velocity. These new results suggest a diminishing return to performance from an increase in approach speed.

KEYWORDS: DRAG REDUCTION, WIND AND ALTITUDE ASSISTANCE, DENSITY ALTITUDE, MATHEMATICAL MODELING OF ATHLETIC PERFORMANCES

Introduction

Athletic performances in track and field can be manipulated both physiologically and environmentally. The testing for performance enhancing drugs is an attempt to normalize performances that were manipulated with drugs. Environmental factors are normalized by declaring performances with an excessive tailwind "illegal." The standard in sprints and long jump is to classify performances which benefit from a tailwind of greater than 2.0 m/s as ineligible for recognition as world records. The crude nature of this standard has led to studies that focus on instead adjusting performances with consideration to the effects of atmospheric drag. (see Ward-Smith (1984, 1986), Linthorne (1994b), Mureika (2001, 2003, 2006), Frolich (1984)) Most models examine sprint performances where the longer time interval allows atmospheric drag to provide greater influence than in jumps. However, historic data indicates that both the jumps and sprints may be measurably affected by atmospheric conditions. The 1968 Olympics – which took play in the high altitude of Mexico City -- saw records in both sprints and jumps broken by large margins. The most remarkable performance was recorded in the long jump. Bob Beamon of the United States jumped 8.90 meters, breaking the previous record by more than 0.5 meters. This jump was analyzed by Ward-Smith (1986), who examined primarily the effect of wind and altitude on the speed of the approach before the the jump rather than the flight through the air. Modeling the approach as a sprint provides inaccuracies because instead of approaching at maximum pace, long jumpers use a slower sprint to mitigate biomechanical difficulties of taking off at maximum pace (Seyfarth *et al.*, 2000).

We provide an accurate model for both the in-flight phase and slow sprint approach to the long jump. Analysis concerns differences in performance with respect to relative gravity, physical altitude, density altitude, and tailwind speed. This model can be used to normalize performance to an environmentally neutral locations and to provide information concerning takeoff velocity and the in-flight path of a jumper. The model is split into an aerial phase and an approach phase which provides an approximation for the effect of the environment on takeoff speed and angle.

The Aerial Phase

To accurately simulate atmospheric effects on the long jump, a model should incorporate and allow variation of all parameters of interest. The greatest environmental effect on the long jump is wind speed, which can decrease or increase the force of drag on an athlete. In the aerial portion of the model we will approximate the body of the long jumper as a rigid projectile, as shown in Figure 1.

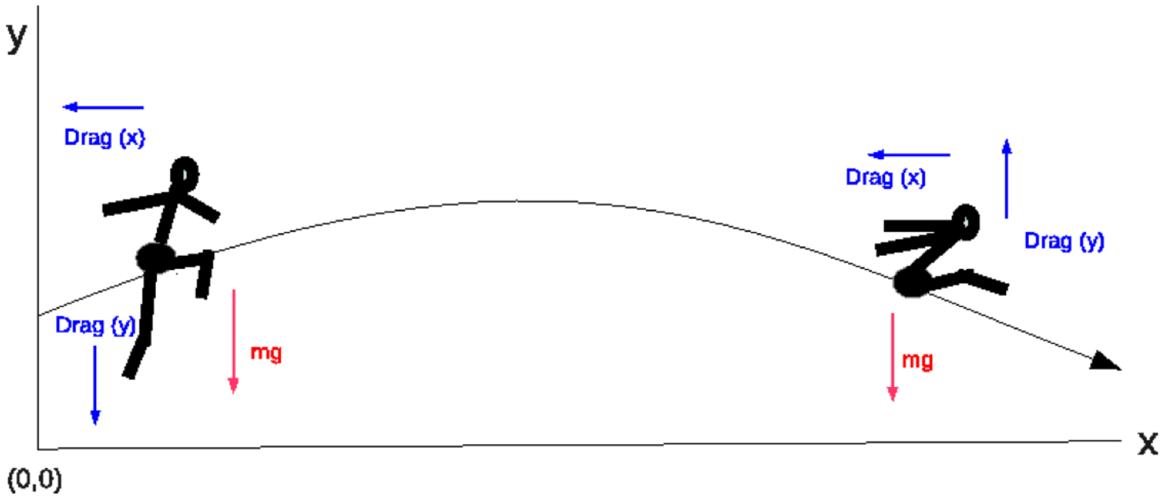


Figure 1. The free body diagram depicts the in-flight forces on the athlete. Specifically, the force of drag parameter (which is resistive in the horizontal direction but contributive in the vertical direction after the peak altitude has been reached) and the force of gravity which always acts in the negative vertical direction.

It is an accurate approximation to assume drag forces are proportional to the squared velocity of the jumper. Analyzing the forces present on the athlete provide the most accurate model, because they can be decomposed into a horizontal and vertical components. It is understood from mechanics that for low wind speed, that only forces that act an in-flight athlete are gravitational acceleration (g) that acts only in the negative vertical direction, and the drag force, which acts in the direction opposing movement. We assume that wind (w) is present in only in the horizontal direction. These assumptions are consistent with the fact that modern athletic stadiums are relatively enclosed, which prevents both extreme horizontal wind and measurable wind in the vertical direction. We can write these equations as a system of force equations for each direction:

$$\frac{d^2x}{dt^2} = -k_x\left(\frac{dx}{dt} - w\right)^2 \qquad \frac{d^2y}{dt^2} = -g - k_y\left(\frac{dy}{dt}\right)\left|\frac{dy}{dt}\right| \tag{1}$$

where $g = 9.8$ m/s^2, w is the wind speed in the horizontal direction (vertical winds are ignored), and k represents the drag parameters in each direction as follows:

$$k_x = \frac{1}{2}\frac{C_d A_x \rho}{m} \qquad k_y = \frac{1}{2}\frac{C_d A_y \rho}{m} \tag{2}$$

where m is the average mass of a long jumper (75 kg), ρ is air density, and C_d is the dimensionless experimentally-determined drag coefficient of 0.6 (Brownlie *et al.* 2004), which is further described in Figure 2. Total drag is greatest when the air flow is neither laminar nor completely turbulent and both forms of drag are present. Material science in sport has provided different methods to reduce the time spent in this partially turbulent state. This can be accomplished by either quickly transitioning to turbulent flow (*e.g.* dimples on a golf ball) or by remaining in laminar flow (*e.g.* the smooth material of a speed skaters suit). Calculations on the material used by elite sprinters resulted in an effective result to the non-dimensional drag coefficient (C) of 0.40%, from 1.0 to 0.6.

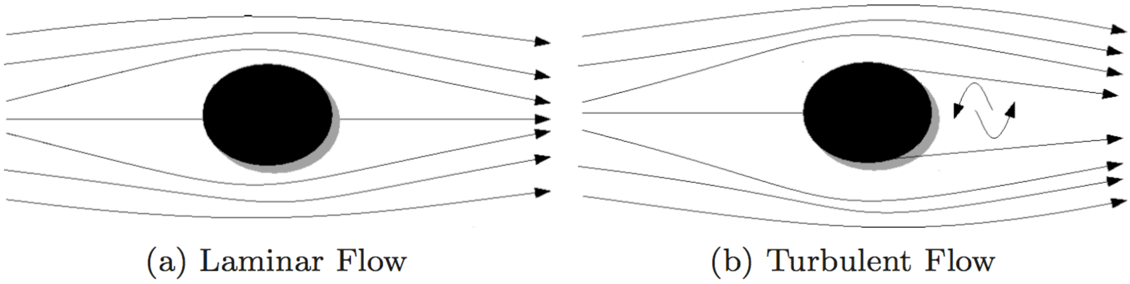


Figure 2. At low velocities (a) friction drag reduces an athletes ability to move throw the atmosphere. As velocity increases the point at which the boundary fluid layer separated farther from the front of the object resulting in a pressure differential which causes form drag. During turbulent flow (b) friction drag is reduced and form drag dominates.

To describe the cross sectional area of an athlete A_x, we first remember that the cross sectional area of a long jumper is time dependent in the horizontal direction. The jumper starts out in an outstretched position but finishes the jump in a crouch. This shrinking of cross sectional area will be approximated with the function $A = (t+2)^{-1}$, where the time (t) varies from zero to about one second, as A varies from 0.5 m^2 to 0.33 m^2. A_y is the horizontal cross sectional area of the athletes body and is approximated as 0.1 square meters.

The parameter most often studied in physical models is wind speed. Mathematically, the effect of the wind speed on the model is the simplest, but also makes the most difference in jump performance. In fact, the wind speed is the only aerodynamically-influential quantity measured at the Olympic Games. An analysis of the effect of wind speed will allow us to understand just how great a difference wind speed can have on an athlete's performance.

Positive wind speeds are tailwinds that decrease the ambient drag force acting on the sprinter. Conversely, negative wind speeds are headwinds into which the athlete runs, increasing the aerodynamic drag. It is important to remember that moderate wind speeds do not aid an athlete by providing a force to "push" the athlete through the air, but acts to counteract drag.

The parameter that is closest to being constant is the value of gravitational acceleration. The value of gravitational acceleration varies due to the distance between the earth's center and the athlete's center of mass. This distance varies with respect to physical altitude and latitude (because of the oblate shape of the earth). Differences in gravitational acceleration are not considered in most models because variation between most track and field venues is negligible. However, that the because the most famous long jump was performed at high altitude and close to the equator in Mexico City justifies the scrutiny of gravity. An approximation for measured gravity can be built up by combining the law of universal gravitation and the international gravity formula (Ceasure, 1987). When both are combined, the following approximation for measured gravity is obtained:

$$g = \frac{9.7803\left(\dfrac{(1+0.001913)\sin^2\lambda}{\sqrt{(1-0.006694)\sin^2\lambda}}\right)}{1+\dfrac{h^2}{R^2}} \tag{3}$$

where h is the physical elevation in meters, R is the Earth's average radius, and λ is the latitude. This equation should allow us to approximate the variation in measured gravity with respect to latitude λ and altitude h. To observe whether stadium conditions vary enough to affect the value of g, the measured gravity was calculated with accurate latitude and altitude for different stadiums. When six different major cities are used from around the world the data showed a maximum change in the gravitational acceleration of about .03 ms^{-2}, which is significant enough to affect an athlete's performance (at the centimetre scale).

The differences in measured gravity at six different track stadiums around can be observed in Table 1. The variations in g at different stadiums is significant enough to effect long jump performance, albeit to a small degree. However, to examine long jump performances differences while only varying the gravity is not necessary or realistic, because the factors that effect measured gravity, physical altitude and geographic location also cause variation air density which significantly effect the drag parameter k .

Table 1. Approximate value of the local gravitational acceleration [ms^{-2}] in different world class stadium locations (from Equation 3).

Venue	g [ms^{-2}]
Mexico City, Mexico	9.782
Colorado Springs, USA	9.790
Sydney, Australia	9.795
Los Angeles, USA	9.795
Oslo, Norway	9.801
Hammerfest, Norway	9.811

In past models air density (ρ) is often approximated as only a function of physical altitude. Physical altitude however does not provide as accurate a description of the air density as density altitude (H_ρ) that represents the effective altitude when considerations are made to barometric pressure, temperature and humidity. Mureika (2006c) has shown the following equation to be an accurate representation of density altitude,

$$H_\rho = \frac{T_0}{\Lambda}\left(1-\left(\frac{RT_o\rho(H)}{\mu P_o}\right)^{\frac{\Lambda R}{g\mu}-\Lambda R}\right) \tag{4}$$

where the parameters are defined as

$\Lambda = 6.5\times10^{-3}\mathrm{K}\cdot\mathrm{m}^{-1},\quad R = 8.314\ \mathrm{J}\cdot\mathrm{K}\cdot\mathrm{mol}^{-1}$
$T_o = 288.15\ K, P_o = 101.325\ kPa,\quad g = 9.80\ \mathrm{m}\cdot\mathrm{s}^{-2},\quad \mu = 2.89x10^{-2}\mathrm{kg}\cdot\mathrm{mol}^{-1}$

Although long jump performance could be effectively modeled by using density altitude as a parameter, the derivation of the drag parameter requires the air density (ρ) itself to be examined. Assuming ideal gases, air density can be described as

$$\rho = \frac{P-P_v}{R_aT}+\frac{P_v}{R_vT}\ , \tag{5}$$

where P is the total air pressure, the gas constants are R_a = 287.05, and R_v = 461.50, and the pressure of water vapor P_v is described by the Magnus-Teton equation (Murray, 1967)

$$P_v \approx \left(\frac{H_r}{100}\right)\cdot 10^{7.5T/(237.7+T)}\ , \tag{6}$$

where H_r is the relative humidity and T is the temperature in degrees Celsius.

This equation for the air density is crucial to our analysis, as it includes three of the parameters often overlooked in environmental analysis of athletic performance, barometric pressure, temperature and humidity. It is important to remember that the barometric pressure that is usually reported in weather forecasts is the corrected sea level pressure, similar to the altitude corrected for air density formula just derived. However, this sea level corrected barometric pressure (usually between 100-102 kPa) is not what we are examining for in the model. Thus, for forecast values, the correction for sea level barometric pressure is necessary, (Lawrence Livermore National Laboratory Meteorological Website, 1997)

$$P_{\text{corrected}} = P_{\text{SL}}\left(\frac{288 - 0.0065h}{288}\right)^{5.2561} . \tag{7}$$

This correction is used in the model to allow inputs to be sea level corrected barometric pressure. This correction allows reported values to be used for the inputs to the third and final parameter in the model air density. As is expected, differences in air density provide a greater effect on performance than gravity but less of an effect than wind speed.

The Approach Phase

The approximation of athletes as rigid bodies in motion during the aerial phase is a predictively-accurate but an incomplete way to model the environmental effects of the long jump. Although, aerial phase mechanics and the environment both affect athlete performance, the biggest influence on jump distance is the takeoff velocity and angle.

The environmental effect on take off velocity requires the approach sprint to be modeled with respect to wind speed and air density. Although sprint models are ubiquitous, modeling this phase in the same manner as a sprint model is inaccurate, because long jumpers do not take off at maximal sprint velocities. By modifying a model originally designed to simulate the 100 m sprint, the long jump preparatory phase can be obtained. Mureika (2001) discusses a quasi-physical model for the 100 m sprint using the following system of differential equations

$$\dot{x} = v(t)$$
$$\dot{v} = f_s + f_m - f_v - f_d \tag{8}$$

Here, (f_s) is the drive term, (f_m) is the maintenance term, (f_v) is the velocity limiting term, and (f_d) is the force of drag. Each term is defined as follows:

$f_s = f_0 \exp(-\sigma t^2)$, This term describes the drive force of the runner. It falls off very quickly with time to mimic the explosive start of a sprinter out of the blocks. This term is left out of the long jump model because long jumpers do not start from blocks, but rather from a "standing" position.

$f_m = f_1 \exp(-c\,t)$. This term describes the maintenance of force for the runner. The exponential being raised to a negative time coefficient means that this term will fall off as time passes (modeling fatigue).

$f_v = \alpha v(t)$ This term serves to limit top speed or leg turnover rate.

$f_d = \frac{1}{2}\rho C_d \frac{A}{m}(v(t) - w)^2$. This term represents drag. It is decreased by wind speed and lower air density.

This model can be used to simulate the approach or preparatory phase of the long jump if the drive term f_s removed (alternatively setting $f_0 = \sigma = 0$). It is assumed here that the main difference in the long jump run from the sprints is the lack of the driving out of the blocks. If this system is integrated numerically over a 30 meter distance (preparatory phase) it will provide initial velocity which can then be used along with take off angle as the inputs for the aerial phase. Following (Mureika 2001), the following values of the parameters are adopted: $f_1 = 5.15$, $\alpha = 0.323$, $c = 0.0385$, $m = 75$ kg, $C_d = 0.6$, and $A = 0.45$ m^2.

In a purely physical model it would seem advantageous for an athlete to jump at a 45 degree angle while maximizing velocity, since this angle would allow the athlete to maximize the horizontal the velocity necessary to travel forward as well as the vertical velocity necessary to allow the athlete time in the air before landing. However, an athlete approaching with a high velocity will be able to take off with a large horizontal take off velocity, but will not have enough ground contact time to create a large vertical force to allow hang-time. Athletes attempt to provide an increase the vertical force by planting the foot ahead of the body to allow more time to create the force. Unfortunately, the longer the athlete takes to create a vertical force the more horizontal breaking that that will take place. Thus, biomechanical compromises between vertical force and horizontal breaking force elite athletes to favor a take off angle between 20 and 25 degrees. Seyfarth *et al.* (2001) has experimentally determined the take off angle of elite long jumpers to be close to 21 degrees with a take off velocity of slightly less than 10 m/s.

More recently, Bridgett and Linthorne (2006) have conducted experimental research to determine how elite long jumpers take off angle changes with respect to initial velocity. Like Seyfarth *et al.* (2000), the authors found that elite long jumpers jump with take off angles close to 21 degrees, but that angle varies with respect to takeoff velocity. A linear fit experimental data that shows that as takeoff velocity increases takeoff angle decreases. Incorporating this relationship into the approach phase will provide an experimentally verified relationship between the environmentally affected takeoff velocities and take off angle.

Data and Analysis

The described model can generate data that yields information about long jump performance increase in both the aerial and approach phases with respect to the variation of the following inputs: wind speed, altitude, latitude, humidity, temperature, and air pressure. With the exception of wind speed all other inputs are heavily correlated to geographic location or in the case of elite athletes, specific stadiums during the summer and fall months. For this reason analysis will concern the effect of the stadiums most often implicated in changing jump performance. Three different locations with extremal realistic parameters [temperature (ºC), station pressure (kPa), relative humidity (%)] will be considered: Los Angeles, CA [25 ºC, 101.3 kPa, 10 %], Mexico City [35 ºC, 100.0 kPa, 60 %], and Hammerfest, Norway [15 ºC, 102.0 kPa, 2 %]. The cold and dry conditions in the northern location of Hammerfest should produce the greatest performance decreases, while the hot, humid conditions in Mexico should produce the highest performance increases. Los Angeles (a typical sea-level venue, similar to European arenas) will provide the baseline performance for different wind speeds, although it should be noted that aside from the lower relative humidity, these conditions are reflective of many sea level competition stadia.

Table 2. Effect on Performance (cm) for different venues as compared to Los Angeles at the same wind speed (m/s) at 25°C, 101.5 kPa, and 10% humidity. Values in parenthesis are temperature, barometric pressure, and relative humidity of indicated venues (°C, kPa, %).

	Mexico City (35,100,60)			Hammerfest (15,102,2)		
Wind (ms^{-1})	Total (cm)	Flight (cm)	Approach (cm)	Total (cm)	Flight (cm)	Approach (cm)
-4	17	4.3	12.7	-2	-0.4	-1.6
-2	15	5.6	9.4	-2	-0.5	-1.5
0	9.1	2.4	6.7	-1.9	-0.8	-1.1
2	6	1.6	4.4	-0.01	0.79	-0.8
4	6	3.2	2.8	-0.07	0.63	-0.7

The data generated by this model serves to illustrate two important characteristics of the environmental effects on long jump performance. The first effect to note is the actual performance increase without wind. Comparing performances without wind is important because all modern events have wind gauge readings reported, but rarely consider the air pressure or humidity during the event. The model suggests that without wind an athlete in Mexico City can expect about 9 cm increase in performance while the same athlete jumping in Norway should only see a performance loss of 2 cm. This suggests that while the performance increase in Mexico is real, the drop in performance often attributed to cold conditions may be the result of physiological difficultly rather than the environment.

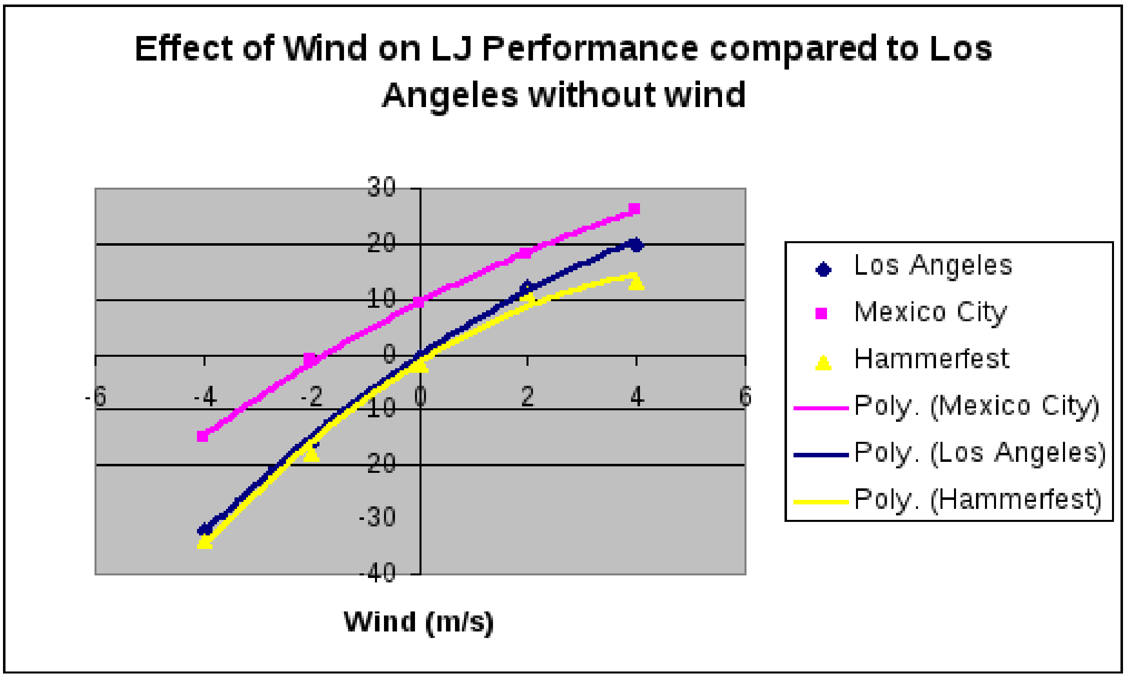


Figure 3. Altitude-based performance correction curves for various wind speeds, as compared to sea-level (Los Angeles) standard.

When the maximum legal wind in Mexico is considered (2 m/s) an elite athlete is expected to benefit from a performance increase of about 18 cm, far less than the value of 31 cm calculated by Ward-Smith (1986b) who attributed the majority of the performance increase to be an effect of the increased sprinting speed at altitude but did not take into account the different jumping angle the sprinting speed would create. Effectively, this model agrees with or is slightly higher than most aerial predictions (Ward Smith, 1986 and Frolich, 1985) but predicts significantly lower values for total performance increase because of the effect of sprint speed on take off angle. It is important to note that without the correction for take-off angle this model actually predicts a total performance of greater than 35 cm on a hot humid day in Mexico, 4 cm more than the difference found by Ward Smith. This difference is due to the effect of humidity, temperature, pressure and the locally-measured value of gravitational acceleration, all of which were not considered in the 1986 paper. It is only because of the confluence of speed and take off angle that the model's total performance increase is far less than what has been previously predicted.

Table 3. Total Performance corrections (cm) compared to performances in Los Angeles without wind.

	Mexico City		Hammerfest		Los Angeles	
Wind (ms^{-1})	**Total (cm)**	**Increase**	**Total (cm)**	**Increase**	**Total (cm)**	**Increase**
-4	-14.9		-33.9		-31.9	
-2	-1		-17.9		-15.9	
0	9.1		-1.9		0	
2	18.1		11.1		12.1	
4	26.1		13.1		20.1	

It is important that while this model suggests unfavorable conditions should not cause performance to decrease drastically, there is a very real (if overestimated) effect on long jump performance in favorable conditions that aided but did not invalidate the 8.90 meter jump of Bob Beamon. The model suggests that Beamon’s advantage of 18 cm (2 m/s of wind, compared to sea level without wind in Los Angeles) were responsible for less than a third of the margin between Beamon’s jump and the previous record. Even if one examines the possibility of a faulty wind reading and provides Beamon with what would have been a very unusual tailwind (4 m/s) the physical effects of drag would only have produced 26 centimeters of aid compared to jumping at sea level Los Angeles.

The second notable feature of the data generated by the model is the non-linear relationship between wind speed and increased performance. Two reasons serve to create this effect. First, the drag force depends on the square of the relative velocity of the athlete and the air which causes the performance decrease in running into a tailwind much larger than the increase in running in front of a headwind. Second, using Linthorne’s linear relationship between take-off speed and angle causes the positive effect of a tailwind to be lessened the faster the athlete is able to move. Although the relationship between speed and angle is approximated to be linear the trigonometric functions used to decompose the speed into separate vectors are non-linear.

Table 4. Top personal men's long jump performances per athlete and corrections (cm) with legal wind aid (w≤ 2 m/s). and illegal wind (w > 2.0 m/s). I = illegal wind, A = legal wind, U = unaided or negligible assitance (w < 0.5 m/s).

Rank	Name	Top Dist (m)	Wind (ms^{-1})	Legal (m)	Wind (ms^{-1})	Unaided (m)	Wind (ms^{-1})	Total Aid (m)
1 (I)	M Powell	8.99	4.4	8.95	0.3	8.95	0.3	4
2 (U)	M Powell	8.95	0.3	8.95	0.3	8.95	0.3	0
3 (A)	Lewis	8.91	3	8.79	1.9	8.72	-0.3	19
4 (A)	Beamon	8.9	2	8.9	2	8.3	0	50
5 (I)	M Powell	8.9	3.7	8.95	0.3	8.95	0.3	-5
6 (A)	Emmiyan	8.86	1.9	8.86	1.9	8.61	-0.3	25
7 (A)	Lewis	8.79	1.9	8.79	1.9	8.72	-0.2	7
8 (I)	Pedroso	8.79	3	8.7	1.6	8.66	0.3	13
9 (I)	Lewis	8.77	3.4	8.79	1.9	8.72	-0.3	5
10 (A)	Lewis	8.76	1	8.79	1.9	8.72	-0.3	4

Although this model predicts less performance increase than previous models, the experimental data generated by Linthorne is bounded by the data generated by the model. Experimentally, Linthorne found 2 m/s and 4 m/s tailwind to produce about 8 and 14 cm of aid, respectively while the model predicts 12 and 20 cm of aid. The computer model (Figure 3) also agrees with the effect of diminishing gains between take off speed and jump performance that Linthorne also observed.

The model can also be examined parameter by parameter to examine the predictive difference between past models. This model examined variations in gravitational acceleration rather than assuming a constant value, as well as the effect of temperature, humidity and pressure on air density. These considerations proved to slightly improve the environmental aid on a long jumper. However, this model also used a modified sprint model (drive term removed), parameterized the jumpers cross sectional area and used Linthorne's experimental data with respect to takeoff speed and angle. These considerations proved to drastically decrease the expected aid that the environment has on long jump distance. Perhaps most importantly in the world of track and field is the fact that such considerations show that less than half of the distance that Beamon broke the world record jump by can be attributed to Mexico City.

Finally, this model is anecdotally supported through historical data. Although the approach phase of the model was created using a pre-existing sprint model, the correction for speed and angle caused the model to predict significantly less performance aid for the long jump than the 100-meter sprint. If the top 10 performances (including those declared illegal for wind assistance) in the 100-meter sprint and the long jump are examined it is not surprising that every sprinter and eight of ten of the jumpers enjoyed greater than 1 m/s wind assistance. The top ten sprint times were bolstered by an average wind speed of 3.82 m/s which allowed the sprinters an to improve on their top performances with negligible wind speed (less than 0.5 m/s) by an average of 0.22 seconds, the same margin between the top legal 100m sprint and the 258th fastest recorded 100m sprint. In contrast, the long jump performances were achieved with an average tailwind of only 2.46 m/s which supplied the jumpers with an average boost of 12.2 cm (The model predicts 15 cm), which is approximately equivalent to the difference between the top legal long jump and the 4th farthest recorded long jump. Additionally, historical data suggests a generous tailwind provides a greater advantage in the sprints than the jumps. Five of the ten fastest men in any conditions joined the group with wind-assistance, while only one of the top ten longer jumpers of all time achieved his mark with wind assistance.

In closing, we comment on potential future applications of this research. Beyond normalization of performance, the model may be useful for training and technique implementation for different environments. For example, if the environment is very unfavorable to the aerial phase (*i.e.* High relative gravity, large headwind), the athlete may be better served to use a faster takeoff speed and a lower takeoff angle, which would reduce the length of time spent in the aerial phase but provide a faster initial horizontal velocity.

References

Lawrence Livermore National Laboratory Meterological Website. (1997). Retrieved from http://www-metdat.llnl.gov

Behncke, H. (1994). Small effects in running. *Journal of applied biomechanics*, *10*(3), 270–290.

Bridgett, L. A., & Linthorne, N. P. (2006). Changes in long jump take-off technique with increasing run-up speed. *Journal of sports sciences*, *24*(8), 889–897. doi:10.1080/02640410500298040

Brownlie, L. W., Kyle, C. R., Harber, E., MacDonald, R., & Shorten, M. R. (2004). Reducing the aerodynamic drag of sports apparel: development of the Nike Swift sprint running and SwiftSkin speed skating suits. In M. Hubbard, R. D. Mehta, & J. M. Pallis (Eds.), *The engineering of sport 5* (pp. 90–96). Sheffield: International Sports Engineering Association.

Dapena, J., & Feltner, M. E. (1987). Effects of wind and altitude on the times of 100-m sprint races: Die Auswirkungen von Wind und Hoehe auf die Zeiten im 100-m-Sprint. *International journal of sport biomechanics*, *3*(1), 6–39.

Davies, C. T. (1980). Effects of wind assistance and resistance on the forward motion of a runner. *Journal of applied physiology: respiratory, environmental and exercise physiology*, *48*(4), 702–709.

Emiliani, C. (1987). *Dictionary of the physical sciences: Terms, formulas, data*. New York: Oxford University Press.

Freitas, R. A. (1999). *Nanomedicine*. Austin, TX: Landes Bioscience.

Frohlich, C. (1984). Effect of wind and altitude on record performance in foot races, pole vault, and long jump. *Journal of applied physiology: respiratory, environmental and exercise physiology*, *48*, 702–709.
Greene, P. R. (1985). Running on flat turns: experiments, theory, and applications. *Journal of biomechanical engineering*, *107*(2), 96–103.
Linthorne, N. P. (1994). The effect of wind on 100-m-sprint times: Die Auswirkungen Von Wind auf die Zeiten beim 100-m-Sprint. *Journal of applied biomechanics*, *10*(2), 110–131.
Linthorne, N. P., Guzman, M. S., & Bridgett, L. A. (2005). Optimum take-off angle in the long jump. *Journal of sports sciences*, *23*(7), 703–712. doi:10.1080/02640410400022011
Linthorne, N. (1994). How does wind influence sprint times? Was Flojo's 100 m world record wind assisted?: Wie beeinflusst der Wind die Sprintzeiten? Lief Flojo ihren 100-m-Weltrekord mit Rueckenwind? *Modern athlete and coach*, *32*(1), 3–5.
Mureika, J. R. (2001). A realistic quasi-physical model of the 100 metre dash. *Canadian Journal of Physics*, *79*, 697–713.
Mureika, J. R. (2006). The Effects of Temperature, Humidity and Barometric Pressure on Short Sprint Race Times. *Canadian Journal of Physics*, *84*(4), 311–324.
Murray, F. W. (1967). On the computation of Saturation Vapor Pressure. *Journal of Applied Meteorology and Climatology*, *6*, 203–204.
Murrie, D. W. (1986). Determination of wind assistance in athletics: Are the measurements at Meadowbank stadium valid. In J. Watkins, T. Reilly, & L. Burwitz (Eds.), *Sports Science. Proceedings of the VII Commonwealth and International Conference on Sport, Physical Education, Recreation and Health* (pp. 387–392). London: E. & F. N. Spon.
Pritchard, W. G. (1993). Mathematical Models of Running. *SIAM Review*, *35*(3), 359. doi:10.1137/1035088
Quinn, M. D. (2003). The effects of wind and altitude in the 200-m sprint: Einfluss von Wind und Höhe im 200-m-Sprint. *Journal of applied biomechanics*, *19*(1), 49–59.
Seyfarth, A., Blickhan, R., & van Leeuwen, J. L. (2000). Optimum take-off techniques and muscle design for long jump. *The Journal of experimental biology*, *203*(Pt 4), 741–750.
Spiegel, J., & Mureika, J. R. (2003). A model of wind and altitude effects on 110-m hurdles: Ein Modell für Wind- und Höheneffekte beim 110 Meter Hürdenlauf. *Sportscience*, *7*, 12. Retrieved from http://www.sportsci.org/jour/03/jsjrm.pdf
Ward-Smith, A. J. (1984). Air resistance and its influence on the biomechanics and energetics of sprinting at sea level and at altitude. *Journal of biomechanics*, *17*(5), 339–347.
Ward-Smith, A. J. (1986). Altitude and wind effects on long jump performance with particular reference to the world record established by Bob Beamon. *Journal of sports sciences*, *4*(2), 89–99. doi:10.1080/02640418608732104
Ward-Smith, A. J. (1999). New insights into the effect of wind assistance on sprinting performance. *Journal of sports sciences*, *17*(4), 325–334. doi:10.1080/026404199366037

The Assumptions Strike Back! A Comparison of Prediction Models for the 2011 Rugby World Cup

Peter O'Donoghue

Cardiff School of Sport, Cardiff Metropolitan University, Cyncoed Campus, Cardiff, Wales, CF23 6XD, UK.

Abstract

Four studies out of a series of 6 previous studies have found that predictive models are more accurate at predicting actual match outcomes when the modelling assumptions are violated than when data are transformed to satisfy the assumptions. The current investigation produced two sets of two predictive models of the 2011 Rugby World Cup; one set of models used raw independent variables that violated the assumptions of the modelling techniques used and the other set used data that satisfied those assumptions. Each set of models contained a multiple linear regression based model where independent variables were included using the "Enter" method and a multiple linear regression based model where independent variables were included using the "Stepwise" method. The models created using data that satisfied the assumptions of linear regression were found to be more accurate at predicting actual match outcomes than when the data violated those assumptions. The assumptions strike back!

KEYWORDS: MULTIPLE LINEAR REGRESSION, FORECASTING

Introduction

Sports performance is difficult to predict as is evidenced by the fact that betting agencies take bets on the results of sports contests (Stefani, 1998). There are a variety of prediction techniques that have been investigated in sports performance including linear regression, discriminant function analysis, logistic regression, artificial neural networks, simulation and qualitative techniques (O'Donoghue et al., 2004). The data used in prediction techniques include factors that have an influence on game outcomes. The main measurable factors that have been shown to influence sports performance are the relative quality of the performers (O'Donoghue et al., 2008) and home advantage (Courneya and Carron, 1992; Nevill et al., 2002: Carron et al., 2005).

Statistical techniques used in predictive modelling involve two stages. Firstly the model of some dependent variable in terms of some independent variables is produced using previous case data. The second stage is to use the model to predict new cases where the values of independent variables are known but the values of the dependent variable are not known. These statistical modelling techniques have assumptions that should be satisfied by the data used to develop the models (Ntoumanis, 2001; Manly, 2005; Tabachnick and Fidell, 2007). However, a series of prediction studies in sport has shown that satisfying the assumptions of the modelling techniques does not always produce the most accurate forecasts of the actual outcomes of matches. Four of the 6 studies in the series have found that models where the data used satisfied the necessary have been less accurate than corresponding models where data have violated the assumptions (O'Donoghue and Williams, 2004; O'Donoghue, 2005; O'Donoghue, 2006; O'Donoghue, 2010). The one study in the series that found models where the assumptions were satisfied by the data to be more accurate was a study to predict the results of matches of the Euro 2008 soccer tournament (O'Donoghue, 2009). However, the difference in predictive accuracy between those models where the data satisfied the assumptions and those where data did not satisfy the assumptions might not be sufficient to justify the effort in transforming data variables. In each of the 6 studies, the independent variables used failed to satisfy the assumptions of the modelling techniques. It was, therefore, necessary to transform the variables and / or remove outliers in the previous case data in order for the data to satisfy the necessary assumption. The transformations could be logarithmic transformations, square root transformations or mapping functions that map variables onto a standard normal distribution. In one of the 6 studies, on the 2007 Rugby World Cup, it was not actually possible to produce a model where data satisfied the necessary assumptions (O'Donoghue, 2009). This was because as outliers were removed, cases that were not outliers originally became outliers as the variability in the data reduced. Eventually, so many outliers had been removed that there were no previous cases in the data that were upsets. This meant that the particular logistic regression model which was intended to predict matches to go to form or be upsets could not be produced. With the balance of evidence currently opposing the transformation data to satisfy the assumptions of statistical tests, the purpose of the current investigation was to continue the programme of research with a seventh study which was on the 2011 Rugby World Cup.

The specific modelling technique to be applied in the current investigation was multiple linear regression. Four models would be tested in two different ways producing 8 sets of predictions. Linear regression would be done using the "Enter" method and the "Stepwise" method for including independent variables. Each of these techniques would be applied to a set of raw untransformed independent variables and to a set of independent variables that had deliberately been transformed to ensure they satisfied the assumptions of multiple linear regression. Each model would produce an expected outcome (points difference) for each match. This expected points difference and the distribution of residual values would be used in 4 simulation models to predict the chances of matches being won, drawn and lost. The different predictions were compared with the results of actual matches played in the 2011 Rugby World Cup to assess their accuracy.

The Assumptions of Multiple Linear Regression

The following assumptions should be satisfied by data used to produce predictive models using multiple linear regression:

- There should be at least 20 cases for each independent variable.

- Linear regression assumes that the relationship between any independent variable and the dependent variable is linear (Newell *et al*., 2010: 140).
- There must be no outliers in individual independent variables, the dependent variable or residuals. As well as considering outliers within individual variables, we also need to check multivariate outliers. Distance measures such as Mahalanobis distances can be used to identify outliers within the multivariate space (Ntoumanis, 2001: 124-5).
- Multicollinearity should be avoided in the independent variables. This means that no pair of independent variables should be highly correlated (the absolute values of r should be less than 0.9).
- Residuals should be independent, homoscedastic and normally distributed. Rather than testing the distribution of the residuals for different subranges of each independent variable, the predicted value for the dependent variable is used. Therefore we test that there is little correlation between the predicted value of the dependent variable and the absolute residual values to show homoscedasticity. Independence can be checked using the correlation between the residuals and a variable representing the order of measurement of the cases. Normality of the residuals can be tested using z-scores for skewness and kurtosis which should both be between -1.96 and +1.96.

Methods

Data Sources

There are many factors that influence performance in sport; these factors have varying degrees of complexity and validity. The computer-based predictions modelled the relationship between the result of a match and three relevant factors in the 232 matches of the previous 6 Rugby World Cups (1987, 1991, 1995, 1999, 2003 and 2007). The three factors used were World ranking points, distance travelled to the tournament and recovery days between matches. These factors were chosen because reliable data was not available for physiological, technical and psychological aspects of squads.

Unlike soccer, where FIFA have provided world rankings and ranking points since 1993, there were no official world rankings for international Rugby Union during the first 4 World Cups. It was, therefore, necessary to devise a method of synthesising world ranking points for the teams that participated in the previous 4 World Cups. This was undertaken during a peer review exercise described by O'Donoghue and Williams (2003). The data for the 2003 and 2007 World Cups used actual World ranking points that were published by the International Rugby Board (IRB). This world ranking is based on previous results like the FIFA World ranking and should not be considered like the elo-number in chess.

The distance travelled to a tournament by a rugby team was deemed to be the giant circle distance between the country's capital city and the capital city of the host nation. This was obtained from an internet based distance calculator (Indonesia, 2006).

The 2003, 2007 and 2011 Rugby World Cups commenced with 4 pools of 5 teams operating a round robin competition involving 5 pairs of matches in each pool. This meant that each team would not play in one of the pairs of matches. This led to large differences in the recovery days from previous matches between the teams contesting some pool matches. For this reason, recovery advantage was included as a factor. Recovery advantage was the number of extra recovery days a team had since their previous match than the opponents had since their previous match. Where a team did not participate in the first pair of matches, they were assumed to have a recovery advantage of 6 recovery days over their opponents in the second pair of pool matches. This was justified by the 95th percentile for recovery differentials in the 2003, 2007 and 2011 Rugby World Cups being 6 days.

Models

Independent Variables

The models used 3 independent variables which were all determined with respect to the higher ranked of the two teams within matches according to the IRB World Rankings at the time matches were played:

- The difference in World Ranking Points, Rankδ: higher ranked team's value – lower'ranked team's value.
- The difference in distance travelled to the tournament, Distδ: higher ranked team's value – lower'ranked team's value.
- Difference in recovery days from previous match, Recδ: higher ranked team's value – lower'ranked team's value.

Dependent Variable

The dependent variable was the points difference, Pδ, between the higher ranked team in a match and the lower ranked team. If the higher ranked team won the match then this would be a positive value, if the match was an upset then this would be a negative value and if the match was a draw then the value would be 0.

Model A: "Enter" Method with Assumptions Violated

The first regression model was formed using all 232 previous Rugby World Cup matches with the three independent variables being entered in their raw form without any transformation or removal of outliers. This produced the model for Pd shown in equation (1).

$$P\delta = 0.619 + 2.285\ \mathrm{Rank}\delta - 0.0000266\ \mathrm{Dist}\delta + 0.892\ \mathrm{Rec}\delta \quad (1)$$

The assumption of at least 20 cases for each independent variable was satisfied and there were no high correlations between any pair of independent variables ($|r| \leq 0.128$). There was no order effect on residual values with no correlation between date of match and residual values (r

= +0.045). However, there were some assumptions that were violated by the data. The residual values were positively skewed (z_{Skew} = +5.75) and leptokurtic (z_{Kurt} = +6.58), Recδ contained outliers and the dependent value, Pδ, was positively skewed (z_{Skew} = +7.49) and contained outliers. There was also a worrying correlation between predicted values for Pδ and residual values (r = +0.255) meaning that the homoscedasticity of residuals could not be assumed. The residuals had a mean±SD of 0.000±19.613; this information would be used in the simulation model.

Model B: "Stepwise" Method with Assumptions Violated

When the raw untransformed variables were included in the stepwise regression analysis, only one independent variable, Rankδ, was included in the model produced as shown in equation (2).

$$P\delta = 0.650 + 2.315\ Rank\delta \qquad (2)$$

There was no association between date of match and residual value (r = +0.052). However, other assumptions were violated by the data used to create this model. The residuals were positively skewed (z_{Skew} = +5.68) and leptokurtic (z_{Kurt} = +6.33). The correlation between predicted values for Pδ and residual values (r = +0.268) meant that the homoscedasticity of residuals could not be assumed. The residuals had a mean±SD of 0.000±19.685.

Model C: "Enter" Method with Assumptions Satisfied

A number of steps needed to be taken in order for the previous case data to satisfy the assumptions of linear regression. The independent variable Recδ was excluded due to the high number of outliers; any Recδ value other than 0 was found to be a statistical outlier. The first attempt to produce a regression model with the remaining 2 independent variables revealed 7 outliers in the predicted values for Pδ. Three more outliers were removed during the second attempt to produce a regression model and the third attempt provided a model based on 222 previous cases with no outliers in the predicted values for Pδ. The model is shown in equation (3).

$$P\delta = 3.168 + 1.856\ Rank\delta - 0.00161\ Dist\delta \qquad (3)$$

The residual values were sufficiently normal (z_{Skew} = +1.73; z_{Kurt} = -1.67), homoscedastistic (r = +0.117) and independent of date of the match (r = +0.069). The residuals had a mean±SD of 0.000±16.302.

Model D: "Stepwise" Method with Assumptions Satisfied

Recδ was excluded from this model and the same 10 outlying cases were removed from the previous case data. The stepwise regression analysis only included on independent variable out of the remaining two, Rankδ, as shown in equation (4).

$$P\delta = 3.159 + 1.856\ Rank\delta \quad (4)$$

The residuals were sufficiently normal (z_{Skew} = +1.73; z_{Kurt} = -1.67), homoscedastistic (r = +0.116) and independent of date of the match (r = +0.068). The residuals had a mean$\pm$SD of 0.000$\pm$16.302.

Evaluation Process

Each of the 4 predictive models was evaluated in 2 ways. Firstly, the predicted values for the dependent variable Pδ were determined for all 48 matches including the 8 knockout stage matches. These were compared with the actual points difference values in the matches of the 2011 Rugby World Cup using 95% limits of agreement, mean absolute error and the 95th percentile for absolute error. Including the 8 knockout matches avoid errors made in the pool stages propagating into the knockout stage predictions.

The second way of evaluating the models was done using a simulation approach. The residual values were assumed to be normally distributed with a mean of 0.000 and the standard deviations reported previously. The NORMINV function in Excel allowed normally distributed random residuals to be generated with which to modify the expected Pδ value as shown in equation (5) where Expected Pδ is the value for points difference determined by the underlying regression model and $SD_{Residuals}$ is the standard deviation in the residuals used to create the underlying regression model.

$$P\delta = \text{NORMINV(RAND(), Expected } P\delta\text{, } SD_{Residuals}) \quad (5)$$

The simulator played the 2011 Rugby World Cup on 1000 occasions storing predicted results, calculating pool tables, establishing teams involved in quarter-finals, semi-finals, third place play-off and the final as well as the winners of the third place play-off and the final. For both pool and knockout matches, values of Pδ greater than +0.5 were deemed to be wins for the higher ranked team, values between -0.5 and +0.5 were deemed to be draws and values less than -0.5 were deemed to be wins for the lower ranked team. For each of the 1000 simulations of a model, a point was awarded for each pool match where the correct outcome was predicted giving a maximum of 40 points for the prediction of pool matches. No points were awarded for an incorrect outcome even if the prediction was a draw and one team won or vice versa. No points were awarded for quarter-finalists successfully predicted because this would have essentially given credit to the pool stage predictions twice. A point was awarded for each of the 4 semi-finalists correctly predicted, each of the 2 finalists correctly predicted, a point was awarded if the team finishing 3rd was correctly predicted and a point was awarded if the tournament winner was correctly predicted. This gave an overall evaluation score out of 48 points for each simulation of the Rugby World Cup. The mean and standard deviation of evaluation scores for the 1000 simulated tournaments was determined for each of the four underlying regression models.

Results

Expected Points Difference

Table 1 shows that all of the models under-estimated the points difference with respect to the higher ranked teams within matches. The two models that violated the assumptions had lower errors than the two corresponding models where steps were taken to ensure the models used data that satisfied the assumptions of linear regression.

Table 1. Errors between expected results from the predictive models and actual results of the 2011 Rugby World Cup.

Reliability statistic	Violating Assumptions		Satisfying Assumptions	
	Enter	Stepwise	Enter	Stepwise
Systematic Bias	-0.36	-0.37	-2.36	-2.41
Random Error	30.50	30.80	32.49	32.46
Mean Absolute Error	11.62	11.97	12.49	12.48
95th Percentile Abs Error	32.94	32.12	37.05	37.12

Simulation Results

Table 2 shows that the South Africa v Namibia, Australia v USA and Australia v Russia were the most accurately predicted matches while the drawn match between Japan and Canada was the least accurately predicted. Table 3 shows that the differences between the different predictions are minimal. The range of evaluation scores in the 1000 simulated tournaments was 29 to 42 for the "Enter" method where assumptions were violated, 28 to 42 for the "Stepwise" method where assumptions were violated and 28 to 43 for both methods where the assumptions were satisfied. These are very similar ranges of prediction accuracies.

Table 2. Correctness of predictions of pool matches during 1000 simulated tournaments (mean$\pm$SD).

Match	Violating Assumptions		Satisfying Assumptions	
	Enter	Stepwise	Enter	Stepwise
New Zealand 41-10 Tonga	0.99$\pm$0.12	0.99$\pm$0.12	0.99$\pm$0.10	0.98$\pm$0.13
France 47-21 Japan	0.91$\pm$0.29	0.92$\pm$0.28	0.93$\pm$0.26	0.94$\pm$0.24
Tonga 20-25 Canada	0.53$\pm$0.50	0.44$\pm$0.50	0.36$\pm$0.48	0.39$\pm$0.49

New Zealand 83-7 Japan	0.98±0.13	0.98±0.14	0.99±0.12	0.99±0.10
France 46-19 Canada	0.90±0.31	0.92±0.27	0.94±0.23	0.94±0.24
Tonga 31-18 Japan	0.59±0.49	0.51±0.50	0.57±0.50	0.59±0.49
New Zealand 37-17 France	0.82±0.39	0.79±0.41	0.85±0.36	0.83±0.38
Japan 23-23 Canada	0.01±0.12	0.02±0.15	0.02±0.15	0.02±0.15
France 14-19 Tonga	0.12±0.33	0.09±0.28	0.06±0.24	0.06±0.24
New Zealand 79-15 Canada	0.99±0.09	0.99±0.12	0.99±0.09	0.99±0.11
Scotland 34-24 Romania	0.95±0.22	0.93±0.25	0.94±0.23	0.94±0.23
England 13-9 Argentina	0.68±0.47	0.73±0.44	0.74±0.44	0.77±0.42
Scotland 15-6 Georgia	0.77±0.42	0.84±0.36	0.87±0.34	0.85±0.35
Argentina 43-8 Romania	0.94±0.24	0.93±0.26	0.95±0.22	0.96±0.21
England 41-10 Georgia	0.96±0.20	0.92±0.27	0.96±0.20	0.95±0.23
England 67-3 Romania	0.98±0.15	0.97±0.16	0.99±0.11	0.99±0.11
Scotland 12-13 Argentina	0.41±0.49	0.46±0.50	0.39±0.49	0.41±0.49
Georgia 25-7 Romania	0.78±0.42	0.71±0.45	0.73±0.45	0.73±0.44
England 16-12 Scotland	0.67±0.47	0.71±0.45	0.75±0.44	0.73±0.44
Argentina 25-7 Georgia	0.87±0.34	0.81±0.39	0.85±0.36	0.87±0.34
Australia 32-6 Italy	0.97±0.17	0.96±0.21	0.97±0.16	0.97±0.17
Ireland 22-10 USA	0.95±0.21	0.95±0.23	0.94±0.24	0.96±0.19
USA 13-6 Russia	0.53±0.50	0.65±0.48	0.68±0.47	0.72±0.45
Australia 6-15 Ireland	0.09±0.29	0.09±0.28	0.06±0.24	0.08±0.26
Italy 53-17 Russia	0.93±0.25	0.92±0.28	0.95±0.22	0.94±0.24
Australia 67-5 USA	1.00±0.05	1.00±0.04	1.00±0.03	1.00±0.03
Ireland 62-12 Russia	0.98±0.15	0.98±0.14	0.99±0.12	0.99±0.11
Italy 27-10 USA	0.88±0.32	0.86±0.35	0.88±0.32	0.88±0.33
Australia 68-22 Russia	1.00±0.00	1.00±0.06	1.00±0.00	1.00±0.03
Ireland 36-6 Italy	0.75±0.43	0.69±0.46	0.76±0.42	0.76±0.43
Fiji 49-25 Namibia	0.86±0.34	0.88±0.33	0.90±0.30	0.90±0.30
South Africa 17-16 Wales	0.71±0.45	0.72±0.45	0.78±0.42	0.76±0.42
Samoa 49-12 Namibia	0.97±0.18	0.93±0.25	0.96±0.21	0.95±0.21

South Africa 49-3 Fiji	0.95±0.22	0.96±0.20	0.96±0.19	0.97±0.18
Wales 17-10 Samoa	0.81±0.40	0.77±0.42	0.81±0.39	0.81±0.39
South Africa 87-0 Namibia	1.00±0.03	1.00±0.06	1.00±0.00	1.00±0.03
Samoa 27-7 Fiji	0.66±0.47	0.68±0.47	0.72±0.45	0.73±0.44
Wales 81-7 Namibia	0.99±0.10	0.99±0.10	0.99±0.08	0.99±0.11
South Africa 13-5 Samoa	0.92±0.27	0.92±0.27	0.93±0.25	0.92±0.28
Wales 66-0 Fiji	0.88±0.33	0.87±0.34	0.89±0.32	0.88±0.33

Table 3. Accuracy of prediction of different stages of the tournament based on 1000 simulations (mean±SD).

Stage of tournament	Violating Assumptions		Satisfying Assumptions	
	Enter	Stepwise	Enter	Stepwise
Pool stages / 40	31.67±2.07	31.43±2.05	32.03±1.93	32.10±1.93
Semi-finalists / 4	2.55±0.71	2.56±0.68	2.68±0.65	2.61±0.67
Finalists / 2	1.27±0.62	1.28±0.59	1.36±0.62	1.35±0.60
3rd place / 1	0.15±0.35	0.17±0.38	0.17±0.37	0.16±0.36
Winner / 1	0.46±0.50	0.50±0.50	0.52±0.50	0.50±0.50
Total / 48	36.09±2.37	35.94±2.36	36.76±2.30	36.72±2.27

Discussion

The actual performances of higher ranked teams within matches the 2011 Rugby World Cup was better than expected according to the four models. In particular, the models that were created using data that satisfied the assumptions of linear regression under-estimated points differences by 2.36 and 2.41 points. The performances of higher ranked teams was generally better than expected despite some notable upsets such as Tonga defeating France in a pool match. However, the 4 upsets that occurred in the 2011 World Cup was equal to the number that occurred in 1987, 1991 and 2003 and fewer than the 5 upsets that occurred in 1995, the 8 that occurred in 1999 and the 8 that occurred in 2007. The most likely explanation for the models under-estimating the performance of higher ranked teams is the removal of outliers from previous case data in order for the models to satisfy the assumptions of linear regression. There were two rounds of outlier removal which firstly removed 7 matches with points differences of 89 or greater and secondly removed 3 matches with points differences of 82 or greater. Outliers are typically removed from data in sports science studies because they are deemed to have arisen due to measurement error. Sports performance data is often nonparametric with skewed variables being common. The 10 matches won by 82 points or more were real matches and not the result of measurement error. Therefore, it may have been better to transform variables to satisfy assumptions rather than removing outliers.

There were 4000 simulated tournaments played altogether; 1000 for each of the 4 models. The overall accuracy in predicting outcomes was 75.8% which is a greater level of accuracy than has been seen in the prediction of soccer matches (O'Donoghue et al., 2004; O'Donoghue, 2005; O'Donoghue, 2006; O'Donoghue, 2009; O'Donoghue, 2010). The greater accuracy in rugby union is explained by a lower strength in depth in rugby union than would be seen in soccer. In the pool matches of the 2010 FIFA soccer World Cup there were 23 wins for the higher ranked teams, 11 draws and 14 upsets. This means that fewer than 50% of matches are won by the higher ranked soccer teams. In rugby union on the other hand, if the higher ranked team were predicted to win every match, an evaluation score of 40 out of 48 would have been achieved (35 pool matches, 3 semi-finalists, 1 finalist and the tournament winner being accurately predicted). In rugby union there are also fewer draws than would be observed in soccer meaning that with the exception of 4 drawn matches in Rugby World Cup history, the prediction task becomes a choice of 1 of 2 outcomes rather than 1 of 3.

Both of the models where the assumptions were violated predicted an average of 5.9 upsets within simulated tournaments and 0.4 draws while both models where the assumptions were satisfied predicted 4.8 upsets within tournaments and 0.4 draws. This meant that the models where the assumptions were violated had a greater chance of successfully predicting upsets as can be seen in Table 2. However, there were fewer upsets than matches won by higher ranked teams and although the number of upsets predicted was reasonably accurate, most simulated tournaments predicted the wrong matches to be upsets. The methods where the assumptions were satisfied achieved higher accuracy scores based on the larger number of matches won by higher ranked teams that were successfully predicted.

The slightly larger number of upsets predicted by the simulation models where assumptions were violated may be explained by outliers being retained within the data sets used to produce these models. The standard deviation for the residual values for the dependent variable, $P\delta$, averaged around 19 points where the full data set was included. The data sets that satisfied the assumptions of linear regression excluded 10 outliers which were very one-sided matches. This reduced the standard deviation of the residual values to around 16 points. This meant that there was a greater variability of predicted results in the simulation models where the assumptions were violated with more upsets being predicted where the actual results were wins for the higher ranked teams. A further issue which would have slightly reduced the accuracy of the two models where data failed to satisfy the assumptions was that the randomly generated points differences assumed a normal distribution which was not valid for these two models. The positive skew meant that there would be greater variance in outcomes favouring the higher ranked teams than the lower ranked teams.

The expected points difference values can be used to examine the variability of team performance taking into account the quality of the opponents faced. Table 4 shows the distribution points differences for the 4 teams who played 7 matches within the tournament. It also shows the distribution of the difference between the observed points difference and the expected points difference computed using the "Enter" model and satisfying the assumptions of linear regression. This shows that Wales performed much better than expected, with a greater points difference than expected in 6 out of their 7 matches. Australia's winning margins were greater than those of France. However, when the quality of opposition was taken into account, Australia's points difference over opponents was 8.56 less than expected while France's was 3.01 less than expected. Many observing the 2011 Rugby World Cup may have been left with an impression that the French performance was highly variable with a loss to Tonga, a single point victory over Wales who had a player sent off and then a narrow 1 point defeat to the World's highest ranked team New Zealand in the final. Table 4 shows that when opposition quality was taken into account, France's performance showed the lowest variability of the four teams who played 7 matches in the tournament. An issue here is that perceptions of variability are often based on match outcomes with highly variable performances of the same outcome not being recognised as variable. For example, Wales had some very high scoring wins which increased the variability of their performances. Another issue responsible for the variability results found is that a 10 point difference is considered the same by the linear model when that difference is the difference between a 5 point defeat and a 5 point victory and when the 10 point difference is the difference between an 80 point win and a 90 point win. Most rugby observers would not recognise the difference between an 80 point and 90 point win as being the same difference as the difference between an 80 point win and a 90 point win. Therefore, a logarithimic or square root transformation of the dependent variable might have resulted in a better reflection of variability for individual teams. Unfortunately this could not be done in the current investigation because the residuals were negatively skewed when a square root transformation was attempted (z_{Skew} = -3.24) and positively skewed when a logarithmic transformation was attempted (z_{Skew} = +12.75).

Table 4. Absolute points differences and relative points differences taking into account opposition quality and distance travelled (mean±SD) for the 4 teams who played 7 matches.

Outcome	New Zealand	France	Australia	Wales
Pδ	+32.71±25.26	+5.00±15.61	+16.57±26.74	+22.00±30.82
Pδ – Expected Pδ	+7.35±17.16	-3.01±12.41	-8.56±12.66	+15.88±16.70

Conclusions

The results of the current investigation challenge those of previous predictive modelling studies in sports performance. The balance of previous research has suggested more accurate predictions are achieved when the assumptions of statistical modelling techniques are violated. The current investigation has shown that where the data used to create models satisfy the assumptions of linear regression, the models are slightly more accurate than when the data violate these assumptions. Further research will be done on the Euro 2012 soccer tournament and the 2014 FIFA World Cup.

References

Carron, A.V., Loughhead, T.M. and Bray, S.R. (2005) ‘The home advantage in sport competitions: Courneya and Carron’s (1992) conceptual framework a decade later’, *Journal of Sports Sciences*, 23: 395–407.

Courneya, K.S. and Carron, A.V. (1992), The home advantage in sports competitions: a literature review. *Journal of Sport and Exercise Psychology, 14*, 13-27.

Indonesia (2006). www.indo.com/distance, accessed March 2006.

Manly, B.F.J. (1994), *Multivariate statistical methods: a primer, 2nd Edition*, London: Chapman Hall.

Manly, B.F.J. (2005) *Multivariate statistical methods: a primer*, 3rd edn, Boca Raton: Chapman Hall.

Nevill, A.M., Balmer, N.J. and Williams, A.M. (2002). Can crowd reactions influence decisions in favour of the home side? In *Science and Football IV,* Edited by Spinks, W., Reilly, T. and Murphy, A. (London: Routledge), pp. 308-319.

Ntoumanis, N. (2001), *A step-by-step guide to SPSS for sport and exercise studies*, London: Routledge.

O’Donoghue, P.G. (2005), Evaluation of Computer-based Predictions of the Euro2004 Soccer Tournament, 5th International Symposium of Computer Science in Sport, Hvar, Croatia, 25th-28th May, Book of abstracts, pp.36.

O’Donoghue, P.G. (2006), The effectiveness of satisfying the assumptions of predictive modeling techniques: an exercise in predicting the FIFA World Cup 2006, *International Journal of Computing Science in Sport(e),* 5(2), 5-16.

O’Donoghue, P.G. (2009), Predictions of the 2007 Rugby World Cup and Euro 2008, 3rd International Workshop of the International Society of Performance Analysis of Sport, Lincoln, UK, 6th-7th April 2009.

O’Donoghue, P.G. (2010), The effectiveness of satisfying the assumptions of predictive modeling techniques: an exercise in predicting the FIFA World Cup 2010, *International Journal of Computer Science in Sport*, 9(3), 15-27.

O’Donoghue, P.G. and Williams, J. (2004), An evaluation of human and computer-based predictions of the 2003 rugby union world cup, *International Journal of Computer Science in Sport (e)*, **3**(1), 5-22.

O'Donoghue, P.G., Dubitzky, W., Lopes, P., Berrar, D., Lagan, K., Hassan, D., Bairner, A. and Darby, P., (2004), An Evaluation of quantitative and qualitative methods of predicting the 2002 FIFA World Cup, *Journal of Sports Sciences*, **22**, 513-514.

O’Donoghue, P.G., Mayes, A., Edwards, K.M. and Garland, J. (2008), Performance norms for British National Super League Netball, *International Journal of Sports Science and Coaching*, 3, 501-511.

Stefani, R. (1998). Predicting Outcomes. In *Statistics in Sport*, (Edited by Bennett, J.), London: Arnold, pp. 249-275.

Tabachnick, B.G. and Fidell, L.S. (1996), *Using multivariate statistics, 3rd Edition*, New York: Harper Collins.

Tabachnick, B.G. and Fidell, L.S. (2007) *Using multivariate statistics*, 5th edn, New York: Harper Collins.

An Animation Assisted Training System for the Baseball Cover, Relay and Cutoff Play

Chueh-Wei Chang[1], Yi-Po Wu[1], Hua-Wei Lin[2]

[1]*Department of Computer Science and Information Engineering, National Taipei University of Technology, Taipei, Taiwan.*

[2]*National Taiwan University of Physical Education and Sport, Taichung, Taiwan.*

Abstract

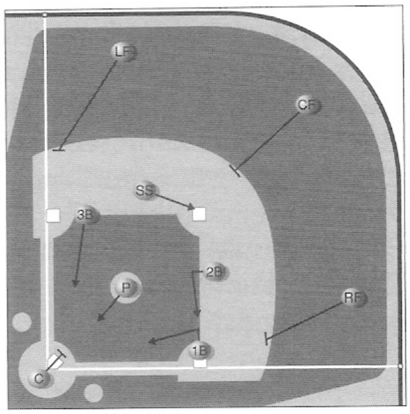
In baseball, the concepts of "cutoff", "relay", "cover" and "backup" are very important to the defensive aspect of the game. In this paper, we design and implement an animation assisted sport simulation system for general baseball "cover, relay and cutoff play" training. We treat the baseball and each player in the ball field as independent "objects" in that each object has its own objective. Each object in the field of play is subject to change once the ball is put into play by the pitcher. We follow a standard paradigm commonly used by the Taiwan national baseball team, and transfer the "cover, relay and cutoff play" into several structured program rules that we refer to as the Decision Making Method. We also design two mechanisms, Fielding Zone Map and the Critical Time Points Interval, for our Decision Making Method. This simulation system allows users to choose what happens once the baseball is put into play, and provides four different scenarios. The system simulates the chosen play through the use of graphic animation while continuously tracking the batted ball, batter, runner(s) on base, number of outs, hit/out batting result. By using this simulation system, players, especially those at the amateur junior level, can improve greatly on all aspects of play in a short period of time. Also, by using this hierarchical design, our system can be easily adapted to many other sports by adjusting the variables.

KEYWORDS: BASEBALL FIELDING TRAINING, ANIMATION PLATFORM, TACTICAL SIMULATION APPLICATION, IMAGERY PRACTICE

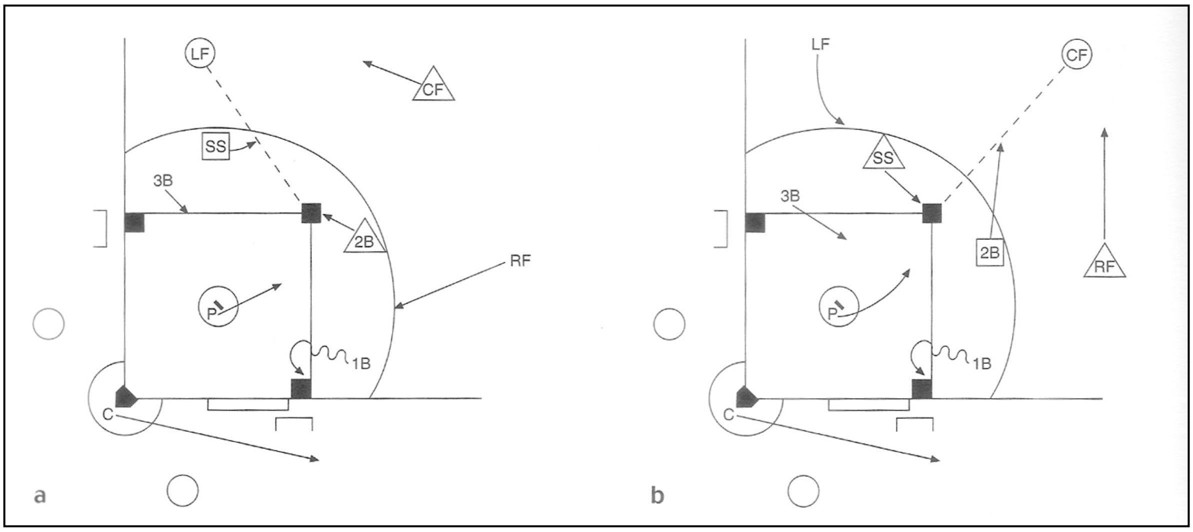
Introduction

A baseball game (Baseball Rules and Gameplay, 2011) relies on teamwork, and consists of two basic parts: offense and defense. Defense plays an integral role in the success of a team. The team that makes the fewest number of mistakes in fielding will likely be the team that wins the game. In defensive plays, every player has a specific task that needs to be completed. One of the most common reasons for failure in a defensive play is the fielders do not know how to work with teammates in a fielding sequence. Coaches need to instruct all the fielders involved in a play on defense strategies, i.e., how to properly execute a play quickly and instinctively. Fielders must comprehend their proper roles and the most efficient responses to various situations that can occur. Among the most widely instructed defense strategies are the “cover, relay and cutoff” methods designed for stopping the base runners from advancing bases (Kindall, & Winkin, 2000). These strategies require a great deal of practice and teamwork in order to consistently result in successful outcomes.

P	Holds runner close at first base, throws a mid-low fastball for a strike, and has front and left-side responsibilities
1B	Holds runner on until pitcher delivers the pitch, then charges the plate
2B	Cheats up and over and covers first base
SS	Covers second base
3B	Starts 15 feet in on grass and charges the plate, has priority over everyone on bunted ball
C	Covers in front of plate, calls where to throw the ball, has priority over pitcher and first baseman



a) Bunt defense: runner at first base, second baseman charges.



b) Fielding: a single to left field with no runner on bases and a single to center field with no runner on bases.

Figure 1. Standard paradigm illustrations for cover and cutoff play. (a) Bunt fielding with different runners on bases (Johnson, Leggett, & McMahon, 2001) (b) Fielding of different ball hit locations with no runner on bases (Stallings & Bennett, 2003). The abbreviation P stands for the Pitcher, and also 1B-First Baseman or First Base, 2B-Second Baseman or Second Base, SS-Shortstop, 3B-Third Baseman or Third Base, C-Catcher, LF-Left Fielder, CF-Center Fielder, RF-Right Fielder. These abbreviations will be used throughout.

There are many standard paradigms in the baseball textbooks for the general defensive cover, relay and cut-off play drills (Jiang, 2001; Johnson, Leggett, & McMahon, 2001; Stallings & Bennett, 2003). For example, Figure 1 describes and illustrates the most effective defensive strategies for specific game situations. Yet, these standard paradigms are planar and still images. One of the most commonly used baseball related software, Baseball Coach (All Stats Software, 2010), still uses static graph charts. These static graphs are unable to illustrate the movement and timing between fielders and runners.

The use of visualization imagery can improve performance and enhance a player's skills (Dalloway, 1992; Weinberg, 2008). With imagery training, players can learn to react to situations as efficiently and succinctly as possible. Mental practice can be just as important as physical practice for players to develop their defensive skills (Rushall, 1991; Morris, Spittle, & Watt, 2005). Dynamic moving illustrations are more effective in training than traditional material, such as book illustrations or verbal communication. By using a suitable visual simulation system, baseball players can quickly improve their "cover, relay and cutoff" play.

In recent years, with the help of information technology and computer software, it has become easier to use animation simulation in athletic training (Chang, Lin, & Chang, 2005; Leser, Uhlig, & Uhlig 2009) and competition management (Kao, Wu, & Chen, 2009; Vincent, Stergiou, & Katz, 2009). By using Microsoft Kinect sensors, electronic gyroscope, and computer graphics animation, training software have been widely utilized in baseball. These technologies also can provide an excellent foundation for simulation based training. Yet, it's still difficult to find a program that can accurately simulate "cover, relay and cutoff" play due to the complexity of the scenarios.

Therefore, we designed an economic, portable, reusable and flexible training tool that can help train players, especially amateur junior student players, to understand and remember each aspect of defense "cover, relay and cutoff play" situations. This software system allows players to repeat the simulation through the imagery training and retain their skills (Hoffler, & Leutner, 2007). We followed a standard paradigm commonly used by the Taiwan national baseball team, and present a methodology to design and implement an animation based simulation system that can help the general training of the "cover, relay and cutoff" play. Our system provides more information than the static version and allows players to get a feel for the timing of a play.

System Design Methods

Class Object Design

Before we started to design this simulation system, we needed to understand the actions that take place on a baseball field. Baseball is a team sport comprised of many actions such as throwing, catching, batting, and running. According to the baseball rules (Commissioner of Baseball, 2011), a pitcher throws a baseball toward home plate, a batter attempts to hit the ball with a bat into the field of play, a batter who hits the ball into the field must begin running toward 1B and beyond, a fielder tries to catch the ball cleanly and throw to the proper location in order to stop the runner from advancing on the bases, especially to home plate. There are nine defensive players on the field, and at minimum one, up to a maximum of four offensive players on the field for a given play. Therefore, we treat the baseball and each player as an independent "object" and each object has its own objective, such as catching a ball within their assigned defensive area, moving to cover a particular base, throwing the ball to the correct place, running to the next base, or just moving to a designated location, etc.

In order to design this system, we used a tree-structured scene description to define the spatial and temporal position of these objects and their movement in a given simulation. The system's compositor uses the scene description information, together with each objects data, to determine the final outcome. We use the 2.5 dimension (2.5D) view projection to create our view of the baseball field, and use Adobe Flash and Action Script language to create a Windows user interface, and generate a series of animations. The overall architecture of scene description is depicted in Figure 2. A full simulation consists of many animations of various objects which perform actions in adherence to the cutoff play and the base running strategy rules. For example, a fielder's action is displayed by a fielder moving (running) animation, a fielder catching animation and a fielder throwing animation accompanied with ball movement animation. By using such a hierarchy design method, our animation platform can be easily adapted into other sports by only changing the background image, object animations and decision rules.

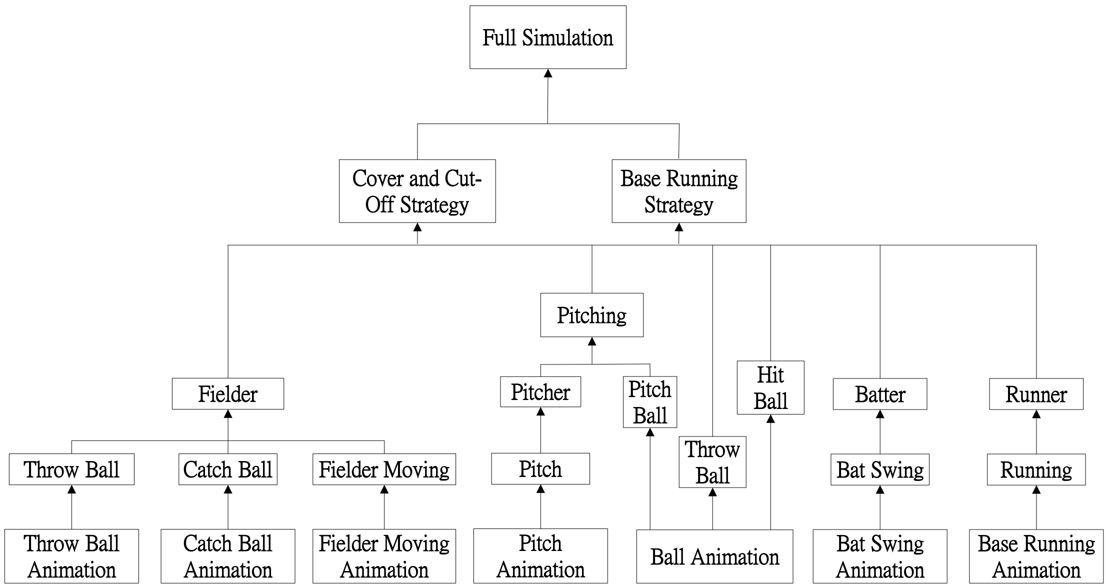


Figure 2. System scene description hierarchy.

System Operation Work Flow

At the beginning of each play, the nine defensive players are in their normal field positions. Also, we put the runner(s) of the batting team on the base(s) according to the user's selected parameters. Before we start the simulation, several factors must be clear: who will handle the ball first after the ball is hit, who will play the cutoff man and in which location, who will cover which base, how many bases the batter-runner and runner(s) can advance, and who will backup for an error and in which location. The system work flow we designed, is shown in

Figure 3, and consists of five steps.

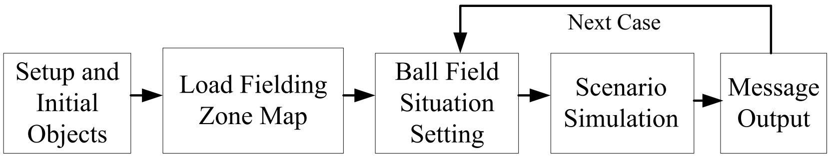


Figure 3. System operation block diagram.

(1) Setup and Initial Objects – This step will put all objects, including the ball field background bitmap, defensive and offensive players, and the baseball, onto the simulation scene by using Action Script AddChild functions. Each object has its own associated animation (running, batting, pitching, etc.) and pre-defined starting location according to standard paradigms.

(2) Load Fielding Zone Map – In order to define the area each of the nine fielders are responsible for, we designed a color fielding zone bitmap, as shown in Figure 4. The size of the fielding zone map is exactly the same as the ball field background image. We use the color of the ball landing location in the corresponding field zone map to assign which fielder will catch the ball without any ambiguity between fielders.

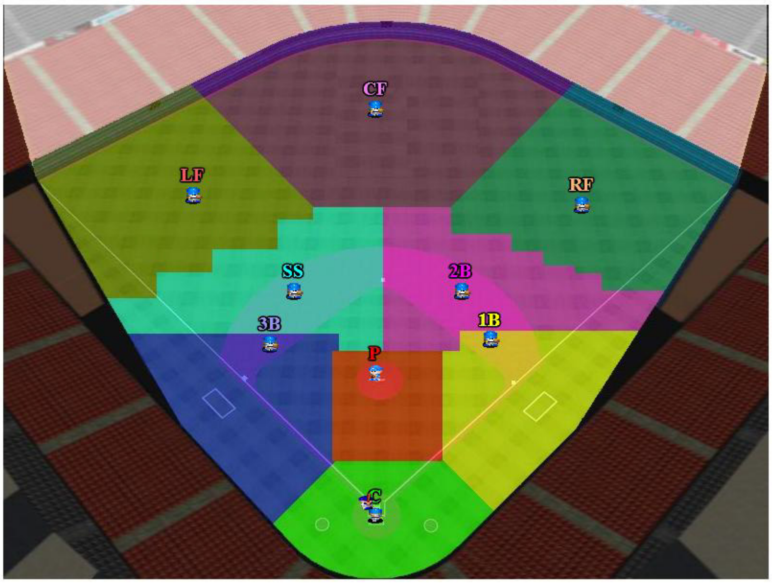


Figure 4. Fielding zone assignment for each fielder.

(3) Ball Field Situation Setting – Before we start to simulate a play, we need to set up four important factors: (a) the ball hit status, (b) the runner(s) on base(s), (c) the number of outs, (d) hit/out batting result. The ball hit status means the movement of the ball after it is hit. It could be a fly ball, line drive, or ground ball. The runner(s) on base(s) have eight different combinations from bases empty to bases loaded. The number of outs is from 0 to 2. There are five different options for the hit/out batting result we can simulate: single, double, triple, fly out, and bunt. Once these four parameters have been selected, the ball landing location also

needs to be specified. In order to simplify the case selection, we provide 18 commonly used training examples by using a push down list box in the bottom left corner of the system user interface.

(4) Scenario Simulation – After the scenario situation has been selected, the system can, just like players in a game, make decisions as to their subsequent actions from the time the pitcher delivers the ball until the play is completed. This includes many complicated physical actions that need to be performed in synchronization among the fielders.

(5) Message Output – After completing the simulation, the system will display the fielders' movement paths, ball hit and ball passed trajectories, and corresponding coaching advice that a fielder must know. These output messages provide sufficient information and details about what each player should do in such a situation.

Decisions Making

However, there are many diverse and complex scenarios in a baseball game. Our Decision Making Method only focuses on the most common cases and ensures there were no mistakes made by the fielders. Our Decision Making Method for each player is contingent upon where the ball lands and its movement. For example, on a single to left field with no one on base, the left fielder will move to field the ball and throw to the cutoff man (shortstop), the shortstop will move to align himself with the left fielder and second base to prepare for a potential cutoff throw, and the center fielder will move to back up the left fielder in case the ball eludes or deflects off his body.

Once the players have taken their appropriate positions on the field, we can divide the time of the play into several intervals and use the Critical Time Point method to evaluate and adjust, if necessary in accordance with the national team standard paradigm. In each critical time point interval, we need to determine which fielder will catch the ball and where the fielder will throw to next. We also need to decide what kind of action the other players will do in the situation. Therefore, we combine and apply the standard paradigms into the systems structure.

(1) Critical Time Points Selection – In each play, pitching, batting, catching and running all have their own particular actions and methods. We divide each play into four time intervals, as shown in Figure 5, for each time the ball changes direction. The time interval t_0 to t_1 represents the time it takes for the ball to travel from the pitcher to home plate. The time interval t_1 to t_2 represents the time it takes once the ball is hit by the batter until it is fielded. The time interval t_2 to t_3 represents the action of the receiving fielder till he throws to the cutoff fielder. The time interval t_3 to t_4 represents the interval for passing the ball from the cutoff man to a specific baseman.

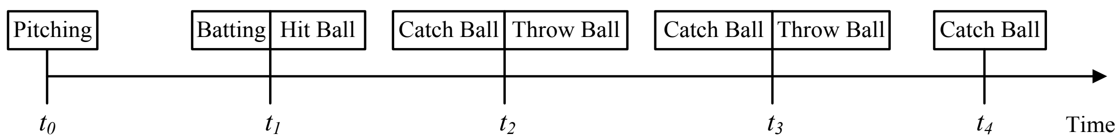


Figure 5. The critical time point and interval diagram.

(2) Fielding Player Selection – After the pitcher delivers and the batter hits the ball, we must determine which fielder will handle the ball first. By using the color-coordinated fielding zone function the correct fielder is easily identified. Once identified this fielder needs to move to where the ball lands and field the ball. The remaining eight fielders move to their specified places to cover bases, cutoff and relay, or backup in the event the ball gets passed the determined fielder.

(3) Cutoff Player Selection – Next, we need to decide who will be the cutoff man according to the standard paradigm. Sometimes, we need to select more than one cutoff man if there is more than one runner, or another complicating factor. The cutoff man needs to move into a proper position, which is most likely the midpoint between two fielders, the one fielding and throwing ball and the one covering the base.

(4) The Second Pass Ball Decision – If the ball is hit deep into the outfield, the fielders need to pass the ball twice to the proper destination. The system can simulate this situation if necessary.

(5) Base Runner Moving Method – At most, there are four base runners running at the same time. For each runner, how many bases he can advance will heavily depend on where ball lands, the number of outs, and whether its' a ground ball or a fly ball. For example, if the ball is hit deep into right field and a runner is on third base, regardless of whether the ball is caught the runner on third base has a chance to cross home plate and score.

Our system's animation will start from when the pitcher delivers the ball and continue until the play is completed and all the players are stable, with no more action required. We provide several bitmapped animation sequences for each player, including pitching, catching, running, throwing in order to make the simulation as realistic and easy to follow as possible.

Implementation Results and Discussion

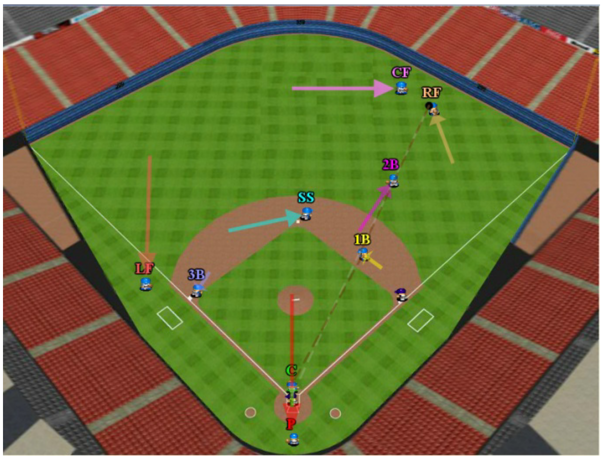
We use an object-oriented programming language Action Script associated with Flash Player API to implement this animation based system. The Flash Player API is made up of classes that represent and provide access to object animation function. Figure 6 is our system user interface. The size of the working area is 1024×768, and the ball field area resolution is 800×600. A user can simply start a simulation by selecting a location where the ball lands with his mouse in the ball field area and pushing the "Start Demo" button on the bottom right corner of the window. The functions in the Scenario area are the ball field situations we mentioned in the previous section. Those commonly used 18 typical training examples we provide, such as all bases are empty for single hit to left field, are put in the "Options" area. Several radio buttons are designed for toggling between showing or hiding the trajectory lines of the ball and the fielders. There is also a description for each player in the upper right "Description" area, such as, the shortstop lines up a throw to 2B and possibly cuts and relays that throw to 2B, the second baseman covers 2B, the right fielder moves to backup 2B in line with the throw, the third baseman covers 3B, the catcher trails the runner to 1B and backs up 1B, the first baseman makes sure the runner touches the base and then covers 1B. The user can easily repeat each scenario and learn where he should go, what he should do, and how to collaborate with his teammates in each such situation.

Figure 6. System user interface.

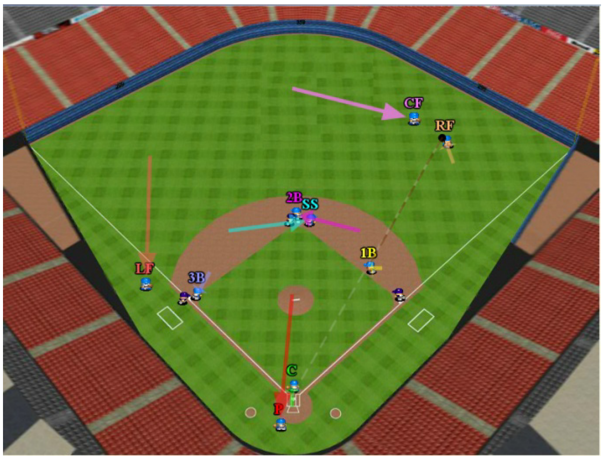
We can see four other examples in Figure 7 of how our system works. Figure 7a illustrates the bunt defense with a runner at first base and indicates that the pitcher, catcher and first baseman will cover the area in front of home plate at the same time, while the second baseman will cover 1B. Figure 7b illustrates a fly ball to deep right field with a runner on third base. The runner on third base has the opportunity to score, and depending on the number of outs, will likely do so. Figure 7c illustrates a single base hit to right field with runners on first and second base. The first baseman becomes the cutoff man, and the right fielder initiates a double relay, first throwing to the second basemen, who, in turn, throws to the first basemen who throws the ball to home plate to prevent the runner from scoring. Figure 7d illustrates a single base hit to left field with runners on first and second base. Due to the depth at which the ball is fielded, the runner on second base is able to cross home plate safely.

We have already tested and provided this training software to many amateur baseball teams in Taiwan. A compressed Chinese version of this cutoff training system software can be downloaded and installed from the following web page: http://csie.ntut.edu.tw/~labmit/cwchang/CutOffTrainingSystem(Merge).rar

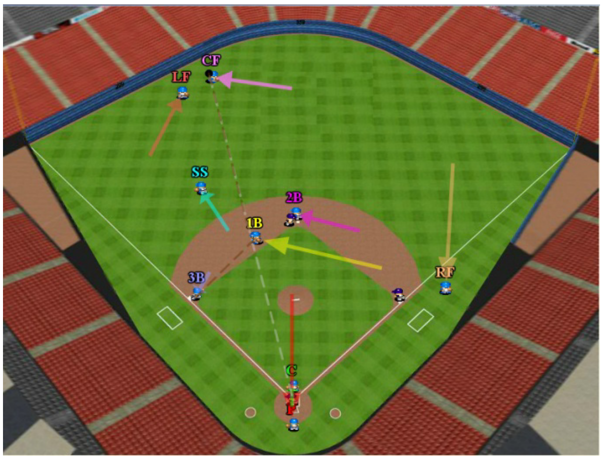
a) Bunt fielding with runner on first base.



b) Sacrifice fly with runner on third base and scores.



c) Single base hit with runner on first and second bases.



d) Single base hit with runners on first and second bases and scores.

Figure 7. System simulation examples.

We realize that not every user will agree with the national team standard paradigms of fielding that we have used as reference, because every team has its own characteristics. However, coaches and players can discuss and change the ideas from the output of this system. We hope that this simulation software will not only provide a tool that can show the method most teams incorporate, but also establish a basis for communication between coaches and players.

Conclusions

Within this paper, we discuss the design and implementation of an animation assisted sport simulation system for "cover, relay and cutoff play" training. This simulation system is an economic, portable, reusable and flexible training tool that can help baseball players better understand and retain each aspect of various defensive situations, and apply them as efficiently as possible. We treat the baseball and each player on the ball field as an independent "object" that has its own objective to complete. Each object in the ball field is activated once the ball is thrown by the pitcher. We follow and transfer the “cover, relay and cutoff play” standard paradigms and program them into the structure of the system. We also designed the Fielding Zone Map and the Critical Time Points Interval, to create our Decision Making Method. This simulation system allows users to arbitrarily choose where the ball lands in the ball field, simulates the play through a series of animations, and allows for the selection of four different scenario options: the ball hit status, the runner(s) on base, the number of outs, hit/out batting result. This system’s animation starts from the pitch ball and continues until the play is completed. By using this simulation system, baseball players can more easily study and understand the complexities of the “cover, relay and cutoff play” and easily adapt them to their play.

We also know that by using such a hierarchy design method, our animation platform can be easily applied to other sports. This can be accomplished by changing the background image, object animations and associated domain parameters. We plan to design another two tactic simulation system for volleyball and basketball in the near future. Furthermore, we also plan to design an editable version with a database for coaches to create their own tactical paradigms. This editable version will be able to record and replay the movement of each object when the user changes an object’s position.

We believe we can still improve many functions of the simulation system such as, providing more realistic scenes by using 3D animation, multiple view directions, more in-depth case studies and special cases in real games.

References

All Stats Software (2010). *Baseball Coach Software*. Web site: http://www.baseball-software.com/index.html

Baseball Rules and Gameplay (2011). *Wikipedia, the free encyclopedia*. Web site: http://en.wikipedia.org/wiki/Baseball

Chang, Chueh-Wei, Lin, Hua-Wei, & Chang, Chen-Kang (2005). The construction and applications of a baseball strategy database with animation playback. *Sports Coaching Science*, 5, 147-158.

Commissioner of Baseball (2011). *Official baseball rules*. Major League Baseball Office of the Commissioner.

Dalloway, M. (1992). *Visualization: The master skill in mental training*. Phoenix, AZ: Optimal Performance Institute.

Hoffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, 17, 722-738.

Jiang, M. H. (2001). *Basic baseball drills*. Taipei, Taiwan: Da Kun Book Store (In Chinese).

Johnson, M., Leggett, J., & McMahon, P. (2001). *Baseball skills & drills*. Champaign, IL: Human Kinetics.

Kao, S. J., Wu, C. C., & Chen, C. F. (2009). Computer-assisted evaluation system for volleyball referee's executive judgment. *International Journal of Computer Science in Sports*, 8(2), 40-49.

Kindall, J., & Winkin, J. (2000). *The baseball coaching bible*. Champaign, IL: Human Kinetics.

Leser, R., Uhlig, J., & Uhlig, M. (2009). Development of an application for learning and teaching soccer tactics. *International Journal of Computer Science in Sport*, 8(2), 19-31.

Morris, T., Spittle, M., & Watt, A. P. (2005). *Imagery in sport*. Champaign, IL: Human Kinetics.

Rushall, B. S. (1991). *Imagery training in sports: A handbook for athletes, coaches, and sport psychologists*. Spring Valley, CA: Sports Science Associates.

Stallings, J., & Bennett, B. (2003). *Baseball strategies*. Champaign, IL: Human Kinetics.

Vincent, J., Stergiou, P., & Katz, L. (2009). The role of databases in sport science: Current practice and future potential. *International Journal of Computer Science in Sports*, 8(2), 50-66.

Weinberg, R. (2008). Does imagery work? Effects on performance and mental skills. *Journal of Imagery Research in Sport and Physical Activity*, 3(1), 1-21.

PerPot
Individual Anaerobe Threshold
Marathon Scheduling

Jürgen Perl[1], Stefan Endler[2]

[1]Mainz, Germany, D-55124

[2]University of Mainz, Germany, D-55099

Abstract

Today an optimization of workout units and the competitions is more important than ever. That applies especially for amateur athletes, who practice beside their normal work. An optimization with the modern lactate analysis is too expensive and not practical for this group of athletes. In this paper we will present a model based alternative, called PerPot. By means of simulation, workout units and competitions can be optimized using only the heart rate profile and the speed profile of the athletes. Furthermore, the individual anaerobic threshold (IAT) can be simulated. Our results show a high correspondence between the athletes' actual (half-) marathon finishing times and the PerPot-simulated results.

KEYWORDS: ENDURANCE SPORTS, HEART RATE, PERFORMANCE, SIMULATION

Introduction

Originally, the antagonistic meta-model PerPot was thought to qualitatively analyze phenomena like delayed reaction on load, collapse effecting overload, or optimizing load profiles in order to approximate given performance profiles. Applying the model to data from practice it turned out, however, that PerPot was able to even provide quantitative results and to predict load-based performance development very precisely (see Perl & Endler (2006), Pfeiffer & Perl (2009), Perl (2010), Endler & Perl (2011)).

Based on those results extensions of PerPot have been developed, which are now able to determine the individual anaerobe threshold (IAT) by simulation, giving heart rate-oriented load scheduling a new quality.

The following contribution therefore focuses especially on running speed as load and heart rate as performance in order to demonstrate exemplarily how load scheduling in endurance sports by means of PerPot works.

PerPot Basics

In the following the dynamics of PerPot is described only briefly with a focus on overflow and reserve, which leads to the question of predicting and avoiding overload as well as underperforming in endurance sports like marathon.

Antagonistic Dynamics, Prediction and the Role of Reserve

As is presented in Perl (2008-2), the meta-model PerPot describes physiological adaptation on an abstract level as an antagonistic process, as is shown in Figure 1: A load input flow is feeding identically a strain potential as well as a response potential. From the response potential the performance potential is increased by a positive flow, while the strain potential reduces it by a negative flow. All flows show specific delays modelling the time that components of the modelled system need to react. In particular in marathon or marathon-like running sports delays play an important role for the process of tiring and recovering.

A typical situation in marathon is a temporary unobserved overload, which is much later followed by an unexpected break down. As can be simulated by the model (see Figure 1) the reason is a delayed reduction of the reserve (grey) of the fatigue potential, which then causes a sudden overflow of strain together with a significant loss of performance. Therefore reserve is the central aspect of predicting future performance development in order to optimize current load management to avoid overload and underperforming. Moreover, the reserve dynamics can be used to determine IAT by simulation, which is helpful not only for scheduling runs.

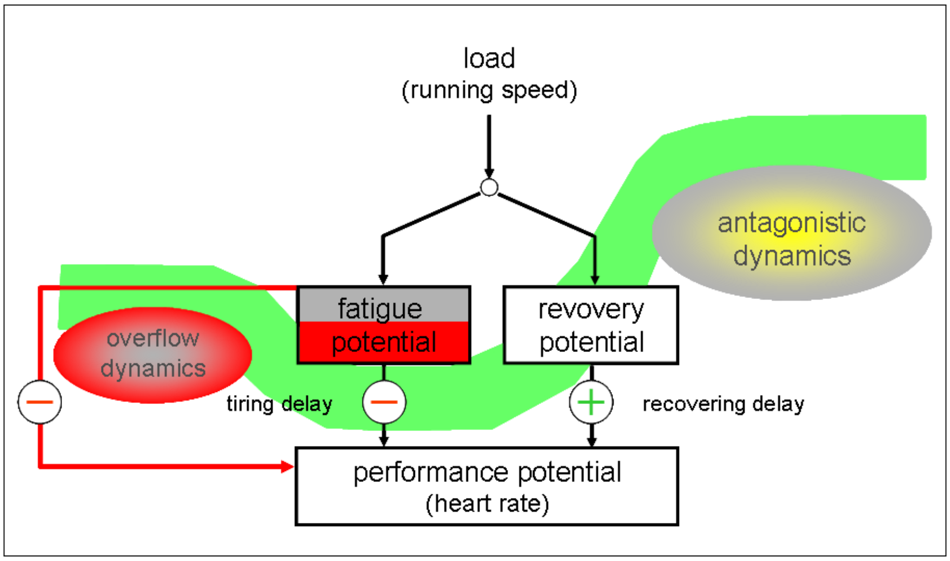


Figure 1. Basic PerPot structure with the highlighted area of reserve and overflow: The red part of the fatigue potential represents the amount of accumulated fatigue, while the grey part means the still available reserve. If the reserve is reduced to zero an overflow of fatigue reduces the performance potential with small delay, possibly causing a sudden break down, which cannot compensate by the slow recovery flow. (The green pattern in the background just symbolizes the super compensation effect.)

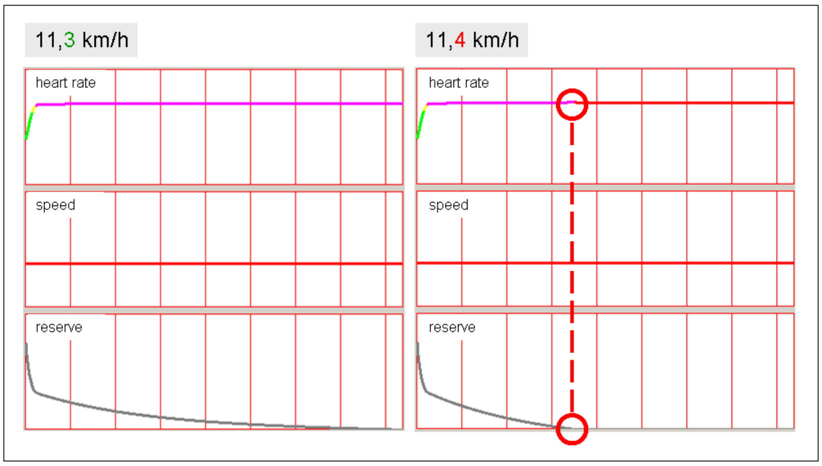


Figure 2. Break down caused by a minimally increased speed. (Colours of the speed curves are explained in the text.)

More details about additional components and effects like overflow, reserve and atrophy can be found in Perl (2005), (2008-1) and (2008-2).

Figure 2 demonstrates how sensible the dependency between small overload and break down can be: Both graphics show the simulated results of a run with constant speed of 11.3 km/h (left) resp. 11.4 km/h (right). The corresponding heart rate curves have changing colours, symbolizing low (green), medium (yellow), and high (violet) rates. The red colour means 'beyond IAT'. The simulation shows that even an only very little increase of speed can cause a reduction of reserve below zero and, correspondingly, crossing the IAT-line.

This correspondence between reserve and IAT, together with some adjustments, can be used to calculate IAT by simulating, using speed and heart rate data of the athlete only, which simply can be taken from a standard step test like that in Figure 3.

Simulative Determination of the Individual Anaerobe Threshold

Figure 3 shows one of the interfaces of a special PerPot-derivation used for IAT calculation and marathon simulation. On the left hand side the result of a step test is presented, where speed and heart rate are measured using appropriate devices, while the reserve curve is simulated by PerPot. Together with the reserve PerPot calculates the IAT, which in this example has a value of 180, as well as some derived heart rate values as orientations for regeneration, base endurance 1 and base endurance 2. Finally, from the internal delays adaptation times can be derived, which in this example say that the athlete reacts in less than 4 minutes on increasing speed, but needs more than 5 minutes to recover in case of decreasing speed.

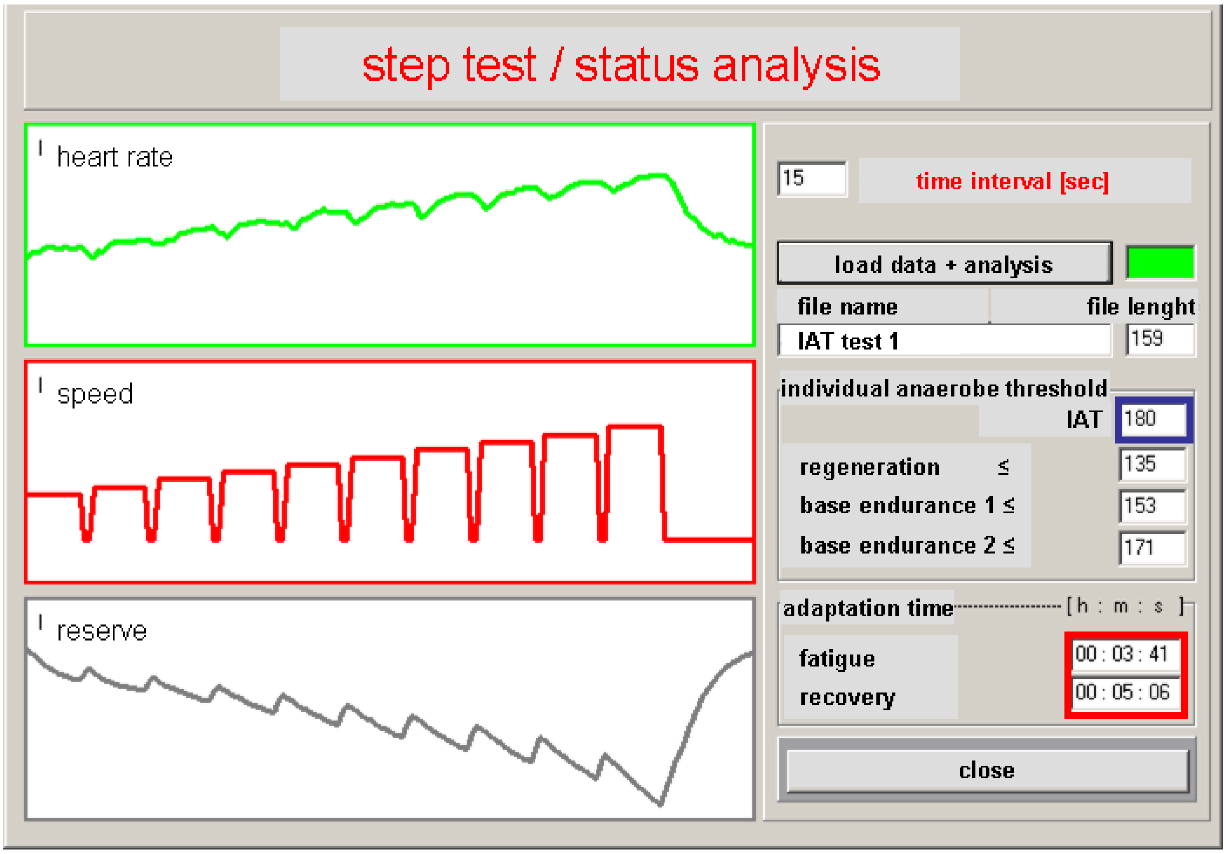


Figure 3. Interface of a PerPot derivation used for IAT calculation and marathon simulation.

Another interface of that tool can be used for segment-oriented simulation of the run itself, providing automatic optimization as well as stepwise manual modification. Figures 2 and 5 show parts of that interface, where the vertical red lines mark the segments, in which the speed can be manipulated or optimized as a whole.

Systematic tests were run exemplarily in cooperation with Mark Pfeiffer from the Institute of Sport Science, University of Bayreuth, Germany. In a double blind test data of 14 athletes were analysed. The athletes completed a step test on a treadmill (starting speed: 6km/h, step length: 3 Min., step raise: 1 km/h, intermission: variable, until blood test). The calculated IAT-values were sent back and compared to the results of the common lactate-tests. Figure 4 presents the results of the 3 most common lactate procedures compared to the PerPot simulation.

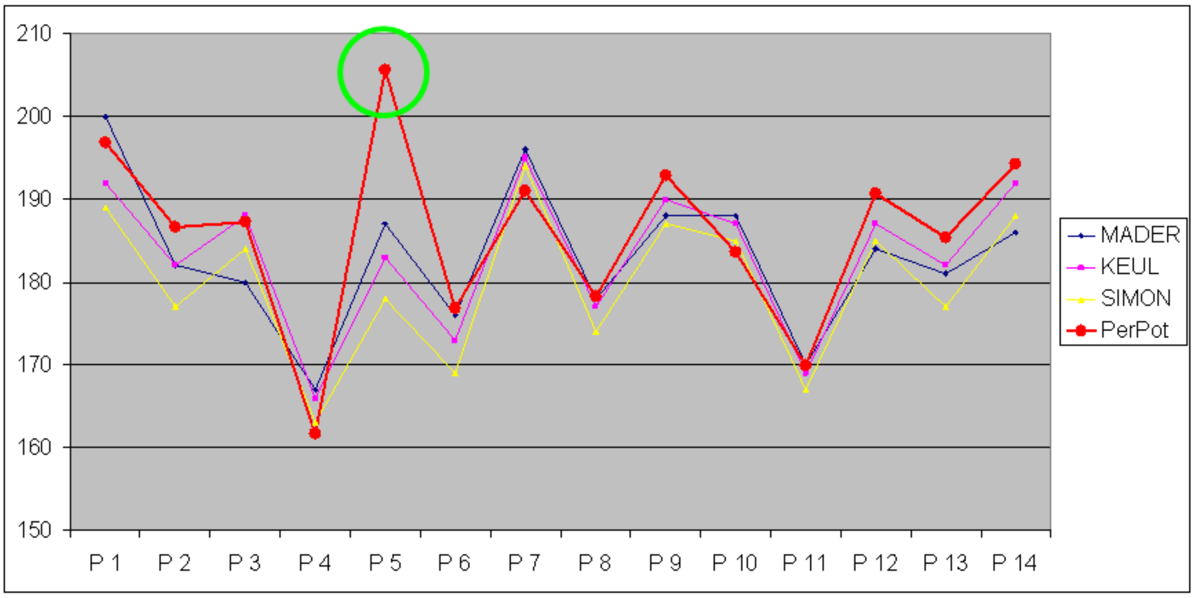


Figure 4. Three IAT lactate tests of 14 athletes compared to PerPot tests. (Note that the connection of the discrete values was done only to clarify the range of variation.)

It turns out that the PerPot results are perfectly in the range – despite athlete P5, where PerPot shows a quite different value. The reason is that P5 has a pathologic high level of heart rate, which could be recognized by PerPot simulation but could not by lactate-analysis.

Marathon Scheduling and Controlling

Using simulation and IAT features the marathon can easily be simulated and optimized under the aspect of avoiding overload and underperforming. Figure 5 shows a result where, depending on a short step test of about 20 minutes, a constant speed is calculated, which optimizes reserve and heart rate: The heart rate reaches the IAT exactly on the last 1000-meter-segment, and exactly then the reserve drops into the overload area.

Of course, individual optimization can be done as well with non-constant speed and depending on course profiles (see Endler & Perl (2006)).

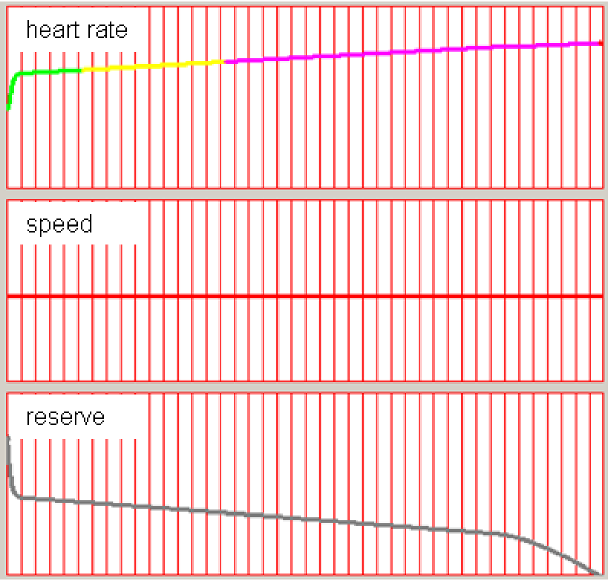


Figure 5. Example of an optimized marathon speed with reaching IAT and overload in the last 1000-meter-segment.

The only problem is to transfer the results of analysis into run control. One very simple solution consists in taking the calculated heart rate profile as an indicator, as has been done in the following tests and is demonstrated in Figure 6: Depending on the (flat) course, 8 time segments of 21 minutes each plus one segment of 9 minutes of specific constant heart rates are prepared and can be used as controlling indicators during the run by means of heart rate measuring.

The particular example presented in Figure 6 stems from the Munich Marathon 2009, where one of the authors (Stefan Endler) scheduled a finishing time of 2:57:00 and in fact finished in 2:56:54.

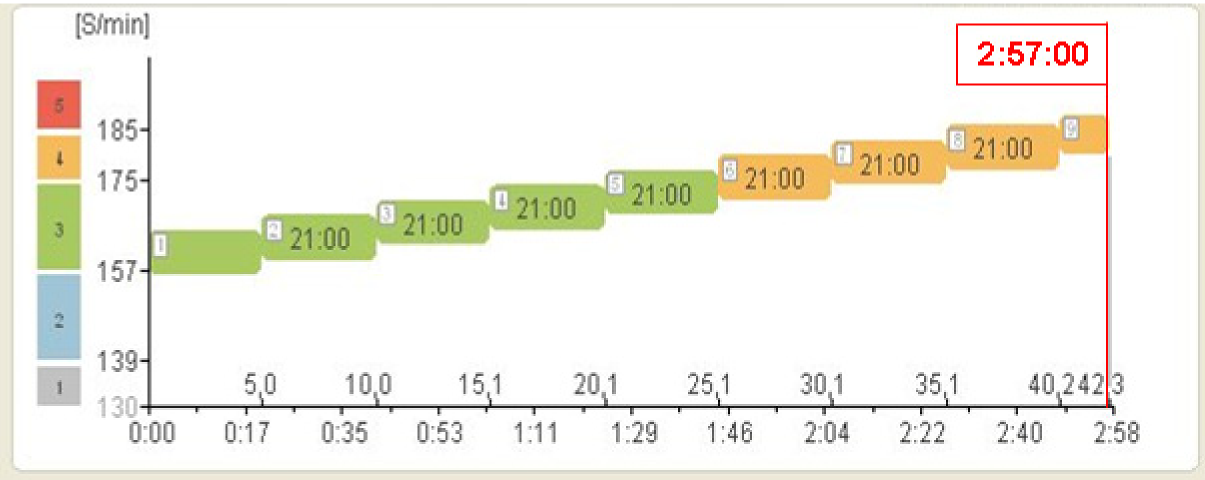


Figure 6. Munich Marathon 2009, optimized by simulation and controlled by heart rate profile: Calculated time: 2:57:00. Finishing time: 2:56:54.

Even more precisely run control can be done using speed instead of heart rate data, if a speed sensor or a GPS-device is available (also see *Conclusion and Outlook*).

Results

Of course, not always the results are that perfect. But during the 18 month since Munich there have been a lot of tests which show that simulation-based run optimization works surprisingly well (also see Endler & Perl (2011)):

Person / Year of birth	Marathon / Half marathon	Year	PerPot time	Finishing time	Deviation
Stefan Endler / 1983					
	Marathon München	2009	2:57:00	2:56:54	-0,06%
weather	Marathon Rom	2010	2:59:00	3:10:53	6,64%
illness	Marathon Mainz	2010	2:55:00	3:13:57	10,83%
illness	Half marathon Worms	2010	1:22:45	1:31:31	10,59%
	Half marathon Frankfurt	2011	1:28:32	1:30:11	1,12%
	Half marathon Mittelrhein	2011	1:30:25	1:29:24	-1,14%
Marion Endler / 1958					
weather	Marathon Rom	2010	4:25:00	4:39:38	5,52%
	Half marathon Mainz	2010	1:57:00	2:00:22	2,88%
	Half marathon Frankfurt	2011	1:59:26	2:01:29	1,72%
	Half marathon Mainz	2011	2:10:31	2:10:38	0,09%
Peter Kossok / 1939					
	Marathon Mainz	2010	4:10:45	2:18:44 (HM)	
	Half marathon Mainz	2011	2:06:36	2:07:38	0,82%
Daniel Roth / 1983					
	Half marathon Frankfurt	2010	1:23:30	1:27:30	4,79%
	Marathon Mainz	2010	2:58:15	3:01:25	1,78%
	Half marathon Worms	2010	1:25:00	1:24:13	-0,93%
unkown	Marathon Köln	2010	2:58:20	3:14:46	9,21%
	Half marathon Frankfurt	2011	1:26:43	1:25:35	-1,32%
	Marathon Mainz	2011	3:06:11	3:04:14	-1,06%
Egor Dranischnikow / 1981					
	Half marathon Mittelrhein	2010	1:45:00	1:47:30	2,38%
Ulrich Heil / 1980					
	Half marathon Frankfurt	2011	1:43:46	1:46:42	2,83%
	Half marathon Mainz	2011	1:41:17	1:42:05	0,79%
Christoph Morrison / 1980					
	Half marathon Mittelrhein	2011	2:27:13	2:24:16	-2,04%

Figure 7. Marathon and half marathon results of the last two years. Highlighted are results with negative (green) or large positive (orange, red) deviations between finishing time and simulated time. In case of positive deviations the reasons were added.

The recruited persons have different sex, age and running pre-condition to cover a wide range of conditional states. Figure 7 shows all long distance results (marathon and half marathon). Moreover, the simulation was used successfully for shorter distances (see Endler (2011)).

The normal difference between the simulated PerPot time and the finishing time is about ±2%. There are only a few larger deviations, which normally are caused by specific reasons or conditions:

The simulation is based on the last step test before the competition. If the time difference between that calibration day and the competition day is too long, running conditions like weather as well as personal state and parameters can change meanwhile.

One example was the Rom marathon 2010. The calibration process was run in Germany at a temperature of about zero degrees Celsius, while the temperature in Rom during the marathon was about twenty degrees Celsius.

Another reason can be an illness after calibration, which causes a significant reduction of the runner's maximum performance.

Conclusion and Outlook

As has been demonstrated, simulation-based calculation of IAT, which is not meant to replace medical lactate tests, in practice can help a lot for better estimating the current status of the athlete in order to optimize his speed load during the run. This seems not only important for well-trained athletes to improve their performance but also for leisure time runners or rehabilitation under the aspect of health care.

Coming projects will deal with ways of on-line controlling and adjusting the running speed by means of on-line data recording and look ahead-simulation. A promising co-operation currently is projected with the working group of Arnold Baca, University of Vienna, Austria, who works with the concept of internet-based remote data acquisition and analysis and therefore in combination with PerPot simulation would enable a perfect on-line run control.

First tests, which where run by one of us (Stefan Endler) in Mainz using the Vienna technology, proved that internet-based on-line coaching also works on long distance.

Finally, the DoMo-Version of PerPot (Perl & Pfeiffer, 2011) enables the handling of course profiles as load components like running speed, which improves precision of prediction together with on-line coaching significantly.

References

Perl, J. (2004). PerPot – a meta-model and software tool for analysis and optimisation of load-performance-interaction. In International Journal of Performance Analysis of Sport-e, Volume 4, Number 2, (pp. 61-73).

Perl, J. (2005). Dynamic Simulation of Performance Development: Prediction and Optimal Scheduling. In International Journal of Computer Science in Sport, 4, 2, (pp. 28-37).

Perl, J. & Endler, S. (2006). Training- and Contest-scheduling in Endurance Sports by Means of Course Profiles and PerPot-based Analysis. In International Journal of Computer Science in Sport, 5, 2, (pp. 42-46).

Perl, J. (2008-1). Modelling. In P. Dabnichki & A. Baca (Eds.), Computers in Sport, (S. 121-160). Wit Press.

Perl, J. (2008-2). Physiologic Adaptation by Means of Antagonistic Dynamics. In M. Khosrow-Pour (Ed.), Encyclopaedia of Information Science and Technology (2nd ed.), VI, (3086-3092).

Pfeiffer, M. & Perl, J. (2009). Simulative Trainingswirkungsanalyse bei einem Fahrradergometertraining mittels antagonistischer Modelle [Simulative training effect analysis in ergometer biking by means of antagonistic models]. In Lames, Augste, Cordes, Dreckmann, Görsdorf & Siegle (Hrsg.), Schriften der Deutschen Vereinigung für Sportwissenschaft, Band 189, (S. 41-51).

Perl, J. (2010). Trainingswirkungsanalyse: Planung und Optimierung mithilfe des antagonistischen Metamodells PerPot [Training effect analysis: Scheduling and optimization by means of the antagonistic meta-model PerPot]. Zeitschrift für Angewandte Trainingswissenschaft, 2/09, 117-127.

Endler, S. & Perl, J. (2011). Leistungsoptimierung beim Marathon mit sportinformatischen Modellen [Performance optimization in marathon by means of models from computer science in sport]. In Link and Wiemeyer (Hrsg.), Sportinformatik trifft Sporttechnologie, 221-225.

Endler, S. (2011). Marathonprojekt. www.informatik.uni-mainz.de/marathon.php (last check: 22.07.2011).

Perl, J. & Pfeiffer, M. (2011). PerPot DoMo: Antagonistic meta-model processing two concurrent load flows. Submitted to International Journal of Computer Science in Sport.