Time-Series Forecasting: A Theoretical Model for Predicting Performance Potential

Joseph Walters[1] & S. Kyle Travis[1]

[1]Center of Excellence for Sport Science and Coach Education, Department of Exercise and Sport Sciences, East Tennessee State University, Johnson City, TN, USA

Introduction

Analyzing, interpreting, modifying, and understanding the training process to receive and repeat a specific performance outcome in sports has been the ultimate goal for sport scientist and coaches for decades. Mathematical modeling has gained elaborate uses in many research disciplines since the introduction of time series forecasting in the 1950s. Improving training-effect adaptability through analysis has traditionally oriented itself, like other areas of science such as medicine, biology and psychology, on the principle of reductionism (Pfeiffer, 2008). Using prior data trends over a series of time, forecasting future performance outcomes can be predictable to a level of certainty. Banister and Busso pioneered the use of mathematical modeling in sport to predict physiological changes that gauge possible performance outcomes for athletes (Bannister, Calvert, Savage, & Bach, 1975; Busso et al., 1992). Theoretical approaches for training adaptations in sport are based off of an antagonistic understanding of training effects (i.e., positive training effects, negative training effects) that present an expected result for actual performance. Classical reductionist approach investigating single-component and two-component factored models in sport should not be under simplified to capture superfluous performance later in the season (Pfeiffer, 2008). Having a grasp on the basic structure and building blocks of the training process, as well as understanding system functionality for mathematical modeling is vital. Therefore, the purpose of this paper is to identify and expound on basic models that are generally unknown to sport scientists and coaches that can be easily applied to training, practice, and game play capturing trends over time to repeat and predict desired performance outcomes.

Total Training Load

Utilizing antagonistic concepts between training effects (i.e., fitness and fatigue input) with actual performance allows the system to simulate upper and lower limit interactions between each variable. One of the most useful and simplest variables a sport scientist and coach can measure is total training load (TTL). Total TL can be measured internally and externally in various ways depending on the sport (i.e., volume load from strength training, conditioning, practice or game time, intensity, strength of opponent, heart rate, questionnaire, etc.) (Halson 2014). Total TL is one of the most important determinant factors for performance outcome and it should be monitored up to twice daily (e.g., strength training and practice time) over a period of weeks and months to observe training and performance trends (Gabbett, 2016). A simple method for obtaining training load is the use of session rate of perceived exertion (sRPE) multiplied by duration (i.e., minutes) of the training session, practice, competition, or game. Session RPE is a subjective measure provided by the athlete from a bout of training and can provide a valid measure of difficulty from the session (Wallace, Slattery, & Coutts, 2009). There are various scales that can be used for quantifying an athlete's subjective level of exertion, but the most used and reliable scale is the Borg Exertion Scale (Borg, 1998). After a training session has been completed, the sport scientist or coach acquires a number from 1 to 20 from the athlete that displays the level of difficulty the athlete felt during the session. Once this number is obtained then the duration of the session (e.g., 60 minutes) is multiplied towards that value. This simple method is the *input* information put into the system. Once the system processes this information, it can reveal an indirect measure of the output information (i.e., positive or negative performance outcomes).

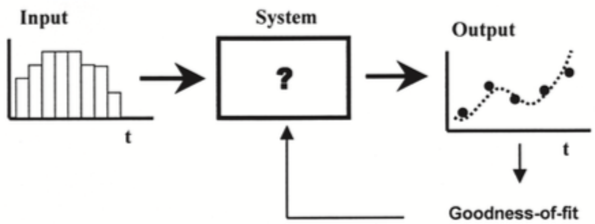

Figure 1. Adapted from Busso and Thomas, 2006. Schematic representation of the procedure for specific characterization of a dynamical process. The model parameters are estimated by fitting model output to actual previous data.

The use of measuring TTL can disclose a degree of subjectiveness from an athlete and provide appropriate planning or adjustments for the team or individual player. This model is dependent on the two-component input information (i.e., TTL x sRPE) to derive the mean potential output by fitting to actual previous data. Composed of two antagonistic first-order transfer functions, this model allows proper representation of the training responses derived from an average of the input in order to predict an accurate performance level. The quality of this prediction is measured by the mean differences between each factor from the previous weeks of training with a variation of 4.9% agreeing with the variations described by Torrents, Balagué, Perl and Schöllhorn (2007). Torrents et al. (2007) also state that actual performance variation can change up to 10%. The actual performance data is measured by team ranking, schedule strength, total points per game or practice, point attempts, etc. However, the authors suggest observing and inputting data over a full season, repeatedly, in order to reduce the variation within the output prediction to increase accuracy and goodness-of-fit.

Theoretical Model

The strategy data presented in this report is theoretical and based on training load variations that would be observed in team sports that tend to have extended seasons (i.e., basketball). The forecast method used is a **B**ox and Cox transformation, **A**RMA errors, **T**rend and **S**easonal components, also known as the BATS model. This approach is a combination of other models (e.g., moving average, Holt-Winters) and can capture seasonal trends in daily, weekly, and monthly sequences and improves forecast accuracy concerning daily analysis. Noticeably, the training load at the start of each week is low in comparison to the end of the week allowing the model to capture trends within training and coaching process. This can be particularly useful for sport scientist and coaches allowing accurate adjustments to daily workloads (e.g., practice, strength training sessions, conditioning sessions, etc.). By utilizing the prediction forecast, the upcoming weeks in the training plan can allow the team to compete at their highest level.

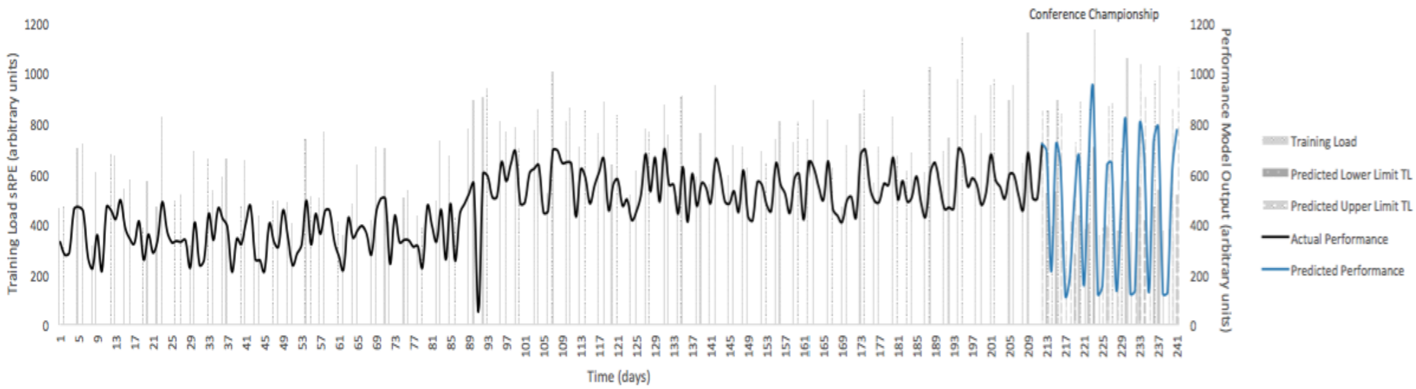

Figure 2. Daily TL variations and actual performance trending towards conference championship.

Discussion

Because computational power has magnified in recent decades, time series forecasting can be used in sport with basic fundamental knowledge of data science. Sports scientist and coaches can use the above model to aid them in daily and weekly forecasting and could provide useful in determining exercise selection, duration, intensity, etc. However, sport scientist and coaches that are considering using time series forecast should understand that certain models have limitations in prediction accuracy, and choosing the correct model may require trial and error to determine the best fit. As previously suggested, the more data points collected will increase the accuracy of the model and therefore correct any outstanding errors. By utilizing mathematical models, sport scientists and coaches are provided with an additional tool for developing training plans, minimize fatigue, predict performance outcomes, and optimize the team and athlete's ability to compete at their highest level.

Practical Applications

Mathematical modeling using the system provided within this report (i.e., sRPE x duration with actual performance changes) can aid in enhancing clarity and confidence regarding possible reasons for changes in performance and reduce the level of uncertainty associated with any changes outside of the normal trend. By examining trends, positive or negative, this could be helpful for determining which athletes are ready for demands from competition at important times during the season (i.e., conference championship). Overall, using this model can help the sport scientist and coach, or athlete understand the importance of intensity, time, adaptability, and recovery during the season.

Resources

Hyndman, R. J., & Athanasopoulos, G. (2017). *Forecasting: principles and practice* (2nd ed.). Retrieved from https://www.otexts.org/fpp

Lewis, N. D. (2017). *Automated Time Series Forecasting Made Easy with R: An intuitive Step by Step Introduction for Data Science*. www.AusCov.com.

References

Bannister, E., Calvert, T., Savage, M., & Bach, T. (1975). A systems model of training for athletic performance. *Aust J Sports Med*, *7*, 57–61.

Borg, G. (1998). *Borg's Perceived Exertion And Pain Scales*. Human Kinetics.

Busso, T., Häkkinen, K., Pakarinen, A., Kauhanen, H., Komi, P. V., & Lacour, J. R. (1992). Hormonal adaptations and modelled responses in elite weightlifters during 6 weeks of training. *European Journal of Applied Physiology and Occupational Physiology*, *64*(4), 381–386.

Busso, Thierry, & Thomas, L. (2006). Using mathematical modeling in training planning. *International Journal of Sports Physiology and Performance*, *1*(4), 400–405.

Gabbett, T. J. (2016). The training—injury prevention paradox: should athletes be training smarter and harder? *Br J Sports Med*, bjsports-2015-095788.

Halson, S. L. (2014). Monitoring Training Load to Understand Fatigue in Athletes. *Sports Medicine (Auckland, N.z.)*, *44*(Suppl 2), 139–147.

Hyndman, R. J., & Athanasopoulos, G. (2017). *Forecasting: principles and practice* (2nd ed.). Retrieved from https://www.otexts.org/fpp

Jones, A. M., Vanhatalo, A., Burnley, M., Morton, R. H., & Poole, D. C. (2010). Critical power: implications for determination of V˙O2max and exercise tolerance. *Medicine and Science in Sports and Exercise*, *42*(10), 1876–1890.

Lewis, N. D. (2017). *Automated Time Series Forecasting Made Easy with R: An intuitive Step by Step Introduction for Data Science*. www.AusCov.com.

Pfeiffer, M. (2005). Modeling the Relationship between Training and Performance - A Comparison of Two Antagonistic Concepts. *International Journal of Computer Science in Sport*, *Volume 7*(2), 13–32.

Schöllhorn, W. I., & Torents, C. (2007). Linear and Nonlinear Analysis of the Traditional and Differential Strength Training. *UGDYMAS• KŪNO …*. Retrieved from http://www.academia.edu/27215733/Linear_and_Nonlinear_Analysis_of_the_Traditional_and_Differential_Strength_Training

Siff, M. (2005). *Supertraining* (5th ed.). Denver.

Wallace, L. K., Slattery, K. M., & Coutts, A. J. (2009). The ecological validity and application of the session-RPE method for quantifying training loads in swimming. *Journal of Strength and Conditioning Research*, *23*(1), 33–38.