Modern Techniques and Technologies Applied to Training and Performance Monitoring

William A. Sands, Ashley A. Kavanaugh, Steven R. Murray, Jeni R. McNeal, and Monèm Jemni

Athlete preparation and performance continue to increase in complexity and costs. Modern coaches are shifting from reliance on personal memory, experience, and opinion to evidence from collected training-load data. Training-load monitoring may hold vital information for developing systems of monitoring that follow the training process with such precision that both performance prediction and day-to-day management of training become adjuncts to preparation and performance. Time-series data collection and analyses in sport are still in their infancy, with considerable efforts being applied in “big data” analytics, models of the appropriate variables to monitor, and methods for doing so. Training monitoring has already garnered important applications but lacks a theoretical framework from which to develop further. As such, we propose a framework involving the following: analyses of individuals, trend analyses, rules-based analysis, and statistical process control.

Keywords: time-series analysis, dose-response, statistical process control

Training is a process and a historical sequence of events. As with all processes, whether biochemical, manufacturing, raising children, or others, these processes require some type of regulation and oversight.1 The level of oversight is related to the complexity of the process and the potential cost of losing process control.2 Complex processes tend to have more opportunities for error resulting in a deviation from the most desired path.3 For example, game sports are unique events that involve dynamic interactions between players, and as a result the observed behavior of an athlete or team is influenced by a situation or opponent.4 The “unstable” nature of game sports creates a challenge to quantify performance indicators. Statistical process control is one method used to identify stable performance traits.4 Statistical process control uses normative profiles or the averages of variables from several games, as well as tolerance limits expressed in confidence intervals based on a mean and variance estimates, to determine “reliable” or typical indicators of performance.5 Mathematical modeling, such as probability analysis, also has been used for estimating the impact of a single player or team on performance, predicting future behavior and identifying optimal decision-making strategies.6 Processes can operate along a continuum from deterministically controlled to utterly chaotic7–9 with both sharing a potential for unexpected problems and considerable expense in terms of money and threats to life and health.10

Historically, more effort has been expended in assessing the athlete as a system. However, the preparation and performance of an athlete can also be addressed as a system. When viewed from the perspective of dynamic systems, athletes are described as “non-ergodic, out of equilibrium systems, exploring larger and larger regions of the state space but eventually getting trapped within some relatively small set of the whole state space by the constraints of their sport discipline.”11,12 In short, athlete behavior is the result of systems within systems—physiology within biology, psychology within sociology, skills within events, and events within sports. The events of sport and the training process occur in a time series that are sequentially dependent in that the order of events matters.12

Training monitoring is receiving new and intensified attention. As sportsmen and women continue to search for new and better methods of performance enhancement, monitoring training and performance data are a natural step. Pursuit of an “edge” in training and competition has led coaches, athletes, and scientists along many paths to include increased training loads and durations,13 systematic attention to scheduling via periodization and planning,14 and nutritional supplements,15,16 among others.17–19 As pursuit of competitive advantages has slowed in some areas of application (ie, drugs, practical limits to training time and loads, and increased incidence of injuries, such as concussions); a return to the fundamentals of superior organization and incorporation of systems of training has brought a reemphasis to importance of training principles.20 One can assume that training processes that are better organized will likely outperform those that are less so.21 Organizing training and competition data may lead to hidden information that would not ordinarily be apparent. All of these concepts require data collection, archiving, and measurement. As such, there are three purposes for monitoring athlete training and competition:

- Providing feedback information on the effects of training
- Ensuring that the training plan is effective
- Assessing and determining the adaptation characteristics of the athlete

These 3 purposes have been augmented in recent years because of an increased interest in data mining, prediction of player potential, team and player earning potentials, and increasing development of “apps” that can be included and implemented on a smartphone. More recent monitoring goals include

Sands is with the US Ski and Snowboard Association, Park City, UT. Kavanaugh is with the Dept of Exercise and Sport Science, East Tennessee State University, Johnson City, TN. Murray is with the Dept of Kinesiology, Colorado Mesa University, Grand Junction, CO. McNeal the Dept of Physical Education, Health and Recreation, Eastern Washington University, Cheney, WA. Jemni is with the Dept of Sport and Exercise Science, Qatar University, Doha, Qatar. Address author correspondence to William Sands at wmasands@hotmail.com.
Training and Competition Monitoring to Predict Future Performance

Professional sports have been leaders in collecting and archiving data involving everything from the number of times at bat to number of yards run to free throws attempted. Baseball, in particular, has a long historical tradition of extensive records from games. National, world, and Olympic records have been kept for over 100 years in various archives, libraries, museums, and publications. However, even among the sports with extensive competitive records, there has been a relative paucity of training records and associated analyses. In spite of a lack of long histories of training records, there have been numerous attempts to characterize and predict future performances via trend analyses and regression techniques for competition performances.29 Sports analytics have become a serious adjunct to many sports, with international conferences and publications.39-41 Although analytics hold considerable promise for performance prediction there have been few attempts to use training monitoring for prediction purposes. Most attempts at prediction have relied on competition information and data to use in trend analyses.

Trend Analyses

Attempts at predicting the future occur throughout our lives whether it be the sex of an unborn baby, weather for an outing, how a stock might perform, or who will be president. Predicting the future has many inherent problems. Nondeterministic processes, also called stochastic processes because they evolve and have some probabilistic components, vary within limits and are governed by contingencies and incumencies.7-9,42 Contingencies are described as "chancy" but are not random.7 Contingencies refer to the idea that any sequence of events in a process may vary based on principles from dynamic systems theory.43 Gould described the influence of contingencies in historical sequences using the metaphor of a "tape of time." Gould argued that if one could rewind the historical tape of time and then replay it—the historical sequence that would result would be different from the familiar sequence or process that had occurred previously. The deviation would be based on the inability to ensure identical conditions throughout the sequence and the strong reliance on initial conditions. For example, if one were to record the historical sequences of an athlete from the first day of training to a championship appearance, and then travel backward in time to observe what would happen if the process were repeated, it is likely that the athlete’s path would be dramatically different. Linked to the ideas of contingencies and incumencies is the area of nonlinear dynamics and firm reliance on sets of conditions that set the stage for the immediate and long-term future.46

Contingencies are the “branch points” in an historical sequence or the “bifurcation points” in dynamic systems.46 Contingencies are events that can cause the historical sequence to veer from its original direction to a new path. A key injury may serve to completely redirect a training process, team dynamics, and tactics. Fortunately, one can analyze the past, because it is already known, to attempt postdiction of the events that transpired. In so doing, the analysis can yield important information regarding how events were linked (ie, evolved) in the sequence.9 Of course, one cannot postdict without rich, thorough, relevant, and accessible data on the events that transpired. Contingencies are commonly portrayed in literature and film using time travel to a previous date and then altering the future by performing some act that alters the contingent path of the characters and events that follow.

Incumencies establish the limits within which a process, or time series of events, is constrained. Although process events may vary, the limits of the variance are completely governed by natural laws. For example, an athlete’s vertical jump is not likely to improve by a meter from one day to the next in spite of wishful thinking.
The vertical jump is constrained by the laws of physics, physiology, and psychology that place real practical limits on such change. Contingencies characterize the shifts in a process’s direction, incumbencies characterize the likely limits to any such direction change. Although the concepts of contingencies and incumbencies appear to border on randomness, trend analyses can still provide important information on the progress of training and training goals. Figure 1 shows the trend of Olympic gold-medal performances in the men’s 100-m sprint from 1896 to 2012. A linear trend is apparent and a least-squares regression line is included. The trend line appears to indicate that the gold medalist of the Games in 2020 should approach 9.50 seconds. This information can provide important training goals for those athletes who are considered contenders for the 100-m championship. Their training must embrace the need for producing speeds, stride rates, stride frequencies, and tactics that will lead to a performance of approximately 9.5 seconds. Each athlete has his or her own personal gifts of talent, motivation, and opportunity. Athletes who are not capable of performing at this level may need to reassess their goals and perhaps shift training emphases, establish new goals, or adjust their expectations.

Trend analyses also can serve training management in the day-to-day activities of athletes by providing information about apparent individual trends. For individual athletes, one often needs information more rapidly and may need to combine several trends to obtain a clear picture of training and performance changes. Individual athlete trends can be addressed by a procedure from single-subject-design research called a “split-middle” 47,48. Figure 2 shows data from monitoring female national team gymnasts in the three months leading to the 1988 Olympic trials. 25,27,30 The split-middle technique uses a calculated median slope. Using medians allows the slope calculations to be less sensitive to extreme scores. A split-middle is calculated by first dividing the period of interest in half and then calculating median values of the first and second halves. A line (ie, slope) is then graphed (or drawn) connecting

---

**Figure 1** — Trends of men’s 100-m gold-medal times in the Olympic Games. $R = .92$, $R^2 = .83$, SEE = 0.24, effect size = 0.92.

---

**Figure 2** — Split-middle trends for daily training time leading to the 1988 Olympic trials.
the 2 medians. The split-middle line is then extended to reach the beginning and end of the period of interest.

Figure 2 shows that training time was decreasing up until the first qualifying competition for Olympic Trials. In the next period, training time increased (height of the line) but stayed fairly constant. Finally, and probably indicative of poor training planning, the final weeks before the Olympic trials saw dramatic increases in training time. Of course, the increase in training time immediately before the second most important competition of the year is unlikely to create performance benefits. Sadly, the trend of the final weeks before the Olympic trials indicated that the athletes were likely to be overly fatigued. Of course, not all athletes were influenced in the same way, but the trend shown in Figure 2 usually results in unstable performances. Finally, one can observe that there were reductions in training time, probably as part of a taper before the qualifying competitions, but not before the Olympic trials—another failure in planning and in execution.

Figure 3 uses the split-middle technique to show the trend for increasing numbers of sleep disturbance symptoms in a gymnastics athlete as she approaches the National Collegiate Athletic Association national championships. This particular athlete shows little in the way of sleep disturbances during the preparatory period, but a later rather dramatic increase in sleep-disturbance symptoms. The lack of sleep disturbances during preparation and then the uptick in disturbances during the competitive period points away from a medical sleep disorder and toward the inability of the athlete to cope psychologically with the pressures of competition.

Coaches tend to use training data to perform informal predictions of which athletes are progressing according to plan. This type of prediction involves a mental model of where the athlete should be in training loads and performances during practices. Coaches and athletes rely on a model of preparedness to guide their development. Predicting far into the future is fraught with inaccuracies, guesswork, and wishful thinking. One way to gain some control over long-term development is a system of training management that records and describes dosage of training loads and the athletes’ responses to these loads.

Figure 4 shows the time series of resting heart rate and body mass for a single female college gymnast. The gymnast’s data in Figure 4 show a reduction in body mass, which is common during the preparatory period of this gymnastics team. However, as the training period continues, the gymnast’s resting heart rate first declines and then increases. The gymnast became ill coincidentally at the same time the resting heart rate increased. Decreasing body mass and increasing heart rate have been indicators of overtraining via autonomic response deviations and catabolic processes.

The nature of the prediction in this case should have been that overtraining was occurring and that training load should have been drastically reduced. The consequences of ignoring the information were dire, with the athlete suffering a career-ending injury.

**Rules-Based Training Management**

Management of training can certainly include trend analyses, and there may be little more than semantic distinctions between using identifiable trends in monitoring data as opposed to applying specific rules to identify important changes in training dosages and responses. Moreover, trends can be combined with rules (Figure 4) such that a trend direction or magnitude becomes the information needed to identify and change training loads. Training management usually involves identification of unusual data that indicate some shift from “normal,” or establishing confidence that training can proceed without hesitation or modification. Time series analyses can be invoked to address changes in data values as the training program progresses. Time series analysis involves trends, detection of outliers, and establishing normal data behavior via baselines or rules.

Training monitoring should involve characterization of both dosage (what the athlete does) and response (how the athlete changes behavior or perceives the dosage). For example,
sets, repetitions, resistance, distance run, meters swum, speeds, accelerations, and others compose dosage. Heart rate, variations in heart rate, ratings of perceived exertion, fatigue ratings, adaptation, incidence of injuries and illness, and others compose response. Both dimensions of training are necessary. “Integral to a systematic approach to the planning and periodization of training and competition is the need to document carefully the quality and quantity of training and competition undertaken.”

Characteristic and Idiosyncratic

Training monitoring experience has shown that athletes tend to respond to the stressors of training in individually characteristic ways. For example, an athlete who tends to respond to stress with sleep disturbances will usually respond to stress with sleep disturbances. Athletes who respond to a stressor by becoming ill will often become ill after other stressful situations. This tendency helps define the nature of data collection and the selection and use of markers of stress and adaptation. Thus, it is important to collect a range of variables so that each individual athlete’s responses can be captured. Individual athletes are the unit of analysis. Group data are too often of little use because of the loss of individual dosage and response information. Mean values tend to obscure the most important information for training monitoring—what happened with each athlete?

In an effort to characterize the most common indications of overtraining, Sands used an artificial intelligence approach and computer language (ie, Prolog) to codify over 200 rules extracted from the literature on overtraining and training planning and periodization. The rules for training monitoring were coded in “if-then”-type algorithms such that data were used to activate rules indicating whether data were an indicator of overtraining or stability. Dosage information was managed by recording and displaying individual athlete data using means and trends (Figure 5), followed by a 1-page summary (Figure 6) based on the Prolog rules which identified potential trouble areas. Figure 5 shows an example weekly report provided for a single gymnastics athlete. The dosage variables are grouped from right to left in terms of averages and recency, on the left of the diagram beginning with the last training day and regressing to the seventh to last training day. This is followed further to the left by the average of the eighth-to-last training day to the fourteenth-to-last training day. The left-most bar graph shows the average of the training dosage from the fifteenth-to-last training day to the beginning of the training records. The right side of the diagram shows trends by displaying a sloped line indicating whether the dosages are increasing, staying the same, or decreasing within the same time periods. Response data are shown on the lower portion of the diagram. These data are depicted in two bar graphs and two slopes for each variable. The rightmost response bar graph and slope line shows the last 10 training days. The leftmost response bar graph and slope line show the remaining training days to the start of training records. The durations of the phases being graphed was determined by years of monitoring systems experimentation and experience with these athletes.

In addition to the graphic information presented to the coach and athlete, an expert system was used to translate the numerical and graphic information to a simple list of potential concerns. Figure 6 shows an example of 2 athletes whose reports indicate an athlete with few overtraining symptoms (ID: 602) and an athlete (ID: 582)
Figure 5 — Example output from analyses of daily training data from a single gymnastics athlete.
Monitoring Methods

S2-69

USGF Athlete Tracking Program – Expert System

ID Number = 620
Name =

Dosage Information
Average volume has decreased.
Recent volume is stable.
Average intensity has decreased.
Recent intensity is decreasing.
Average routines have increased.
Recent routines have increased.

Response Information
Average resting heart rate has increased. ***This may indicate maladaptation/overtraining.
Recent resting heart rate has increased. ***This may indicate maladaptation/overtraining.
Average weight is stable.
Recent weight is stable.
Average sleep time has decreased. ***This may indicate maladaptation/overtraining.
Recent sleep time is stable.
Average mood states are stable.
Recent mood states are stable.
Average illnesses are stable.
Recent illnesses are stable.
Average injury reports are stable.
Recent illnesses are stable.

ID Number = 582
Name =

Dosage Information
Average volume has increased.
Recent volume has increased.
Average intensity has increased.
Recent intensity has increased. ***Check training load for overstress.
Average routines have increased. ***Check training load for overstress.
Recent routines have increased. ***Increasing too many load factors may cause overtraining.

Response Information
Average resting heart rate has increased. ***This may indicate maladaptation/overtraining.
Recent resting heart rate has increased. ***This may indicate maladaptation/overtraining.
Average weight has decreased. ***This may indicate maladaptation/overtraining.
Recent weight has decreased. ***This may indicate maladaptation/overtraining.
Average sleep time has decreased. ***This may indicate maladaptation/overtraining.
Recent sleep time is stable.
Average sleep disturbances are stable.
Recent sleep disturbances are stable.
Average mood states have decreased. ***This may indicate maladaptation/overtraining.
Recent mood states have decreased. ***This may indicate maladaptation/overtraining.
Average illnesses have increased. ***This may indicate maladaptation/overtraining.
Recent illnesses have increased. ***This may indicate maladaptation/overtraining.
Average injury reports have increased. ***This may indicate maladaptation/overtraining.
Recent injury reports are stable.

Figure 6 — Expert system examples of 2 athletes. Athlete 602 shows few symptoms of overtraining, while athlete 582 shows many symptoms of concern. This information is for illustration only.

Statistical Process Control

Trend analyses and training monitoring via rules have problems of scalability and face an important question. At what point are trends or data values (means or slopes) deviant enough to demand further concern? When and how to intervene most effectively remains a seat-of-the-pants decision. Statistical process control uses various properties of probability, characteristics of data distributions, and simple calculations to describe when outlier or unusual data merit concern and intervention. The premise behind using statistical process control lies in something called a control chart. \(^2,10,74,75\) Figure 7 shows the basic idea. Performance data proceed from left to right in time. These data form a distribution that is better pictured as a normal distribution lying on its side. The normal distribution has well-known properties such as a mean and standard deviation. The standard deviation is a useful tool for measuring the probabilities of data values about the mean. As training data are recorded, one can use the mean and standard deviation to determine if the values are common (ie, close to the mean) or unusual (ie, far from the mean). Moreover, each data value can be assigned a probability value. Values that are more than 1 or 2 SDs from the mean are correspondingly rare and may demand explanation. Those values that lie close to the mean, say, less than one standard deviation away from the mean, are frequent and may not warrant explanation. However, values that lie several standard deviations from the mean are rare and demand explanation.

An example of data from a college gymnastics athlete is shown in Figure 8. These data show that whenever the athlete’s resting heart rate exceeded the upper control limit (2 SD), there was a corresponding illness at the same time. There was 1 exception after the winter break in late December that may have reflected the sudden return to training after the break. However, the remaining 4 examples of extreme resting heart rates correspond to greater sums of illness symptoms. Heart-rate values beyond the 2 SDs result in a rough-chance probability of 2.5%. A probability value of 2.5% is obviously rare and demands explanation. However, requiring a value greater than 2 SDs to merit explanation may be too stringent for detection of important outliers that should be investigated. Future efforts using statistical process control may require shifting the control limit values, as suggested in the statistical process control literature.\(^10,74\)

Statistical process control procedures can be programmed in simple algorithms and are easily scaled for identification of problem data that could help automate the coach’s ability to gain insight into which athletes may need interventions to gain or regain control of training adaptations. Moreover, these methods are also modifiable as the training process proceeds. Rather than using somewhat arbitrarily derived rules, intervention decisions are based on repeatable probabilities. Although shown in Figure 8 involving response variables, the same procedures can be followed for dosage information.
Conclusion

Recent experience has shown that many investigators are seeking to implement training monitoring with varying success. Sadly, many investigators have admitted that while data are plentiful, quality information usable for training decisions has been much more difficult to obtain. Monitoring for prediction is particularly attractive for administrators and Las Vegas point spreads. 76 Prediction can also help coaches and athletes in establishing information about what can and will happen if events proceed down a particular path. The management of training using monitoring tools, in our opinion, holds more potential benefits for athletes and coaches than prediction. The ability to help athletes train within a window of adaptation is a laudable goal and more achievable than long-term prediction.

Figure 7 — A simple model-control chart showing contrived data to illustrate how data may be distributed over a period of time.

Figure 8 — Applied statistical process control methods on prepractice resting heart rate and illnesses for an entire training year.
References

52. Anderson O. Monitoring your recovery: should you check your blood, or do your moods provide enough information? Run Res. 2002;18(1):1, 6–8.


