Eur J Appl Physiol (2014) 114:11–20
DOI 10.1007/s00421-013-2745-1

ORIGINAL ARTICLE

A comparison of methods for quantifying training load: relationships between modelled and actual training responses

L. K. Wallace · K. M. Slattery · Aaron J. Coutts

Received: 24 March 2013 / Accepted: 30 September 2013 / Published online: 9 October 2013

Abstract
Purpose To assess the validity of methods for quantifying training load, fitness and fatigue in endurance athletes using a mathematical model.
Methods Seven trained runners ($\dot{V}O_2$max: 51.7 ± 4.5 mL kg^{-1} min^{-1}, age: 38.6 ± 9.4 years, mean ± SD) completed 15 weeks of endurance running training. Training sessions were assessed using a heart rate (HR), running pace and rating of perceived exertion (RPE). Training dose was calculated using the session-RPE method, Banisters TRIMP and the running training stress score (rTSS). Weekly running performance (1,500-m time trial), fitness (submaximal HR, resting HR) and fatigue [profile of mood states, heart rate variability (HRV)] were measured. A mathematical model was applied to the training data from each runner to provide individual estimates of performance, fitness and fatigue. Correlations assessed the relationships between the modelled and actual weekly performance, fitness and fatigue measures within each runner.
Results Training resulted in 5.4 ± 2.6 % improvement in 1,500-m performance. Modelled performance was correlated with actual performance in each subject, with relationships being r = 0.70 ± 0.11, 0.60 ± 0.10 and 0.65 ± 0.13 for the rTSS, session-RPE and TRIMP input methods, respectively. There were moderate correlations between modelled and actual fitness (submaximal HR) for the session-RPE (−0.43 ± 0.37) and TRIMP (−0.48 ± 0.39) methods and moderate-to-large correlations between modelled and actual fatigue measured through HRV indices for both session-RPE (−0.48 ± 0.39) and TRIMP (−0.59 ± 0.31) methods.
Conclusions These findings showed that each of the training load methods investigated are appropriate for quantifying endurance training dose and that submaximal HR and HRV may be useful for monitoring fitness and fatigue, respectively.

Keywords Training dose · Fitness · Fatigue · Performance

Introduction

Physical training load is the dose of training completed by an athlete during an exercise bout. There are a variety of different methods used to quantify training loads undertaken by athletes (Borresen and Lambert 2009). These methods can be described as either external [i.e., the training completed by the athlete (e.g., distance, power)] or internal [i.e., the athlete's response to external training load (e.g., heart rate (HR), perception of effort)] loads. Typically, measures of internal training load [i.e., the HR-based training impulse (TRIMP) and session-RPE (sRPE)] have been reported to be more appropriate for monitoring the training process as these methods incorporate the relative physiological stress imposed on the athlete (Viru and Viru 2000). However, recent technological developments (e.g., power meters, GPS devices, accelerometers, etc.) and commercially available software (e.g., training peaks) have made external training load measures increasingly popular amongst athletes (Jobson et al. 2009). Indeed, the

Communicated by William J. Kraemer.

L. K. Wallace · K. M. Slattery · A. J. Coutts (✉)
Sport and Exercise Discipline Group, UTS: Health, University of Technology, Sydney (UTS), Kuring-gai Campus, PO Box 222, Lindfield, NSW 2070, Australia
e-mail: aaron.coutts@uts.edu.au

training stress score (TSS), which can be calculated from power meters, uses the concept of normalized power and an intensity factor based on an individual's lactate threshold of each training bout to provide a single estimate of the overall training load and physiological stress created by that training session. Whilst initially developed for cycling, the TSS concept has been modified for running using speed/distance measures and a unique algorithm based on the demands of running (Skiba 2006). These methods have recently attracted substantial interest from the coaching and athletic community; however, they are yet to receive critical scientific evaluation. Therefore, at present it is not known which method for quantifying training load is more appropriate for monitoring the training process or which relates best to the training outcomes (i.e., performance, fitness and fatigue).

Mathematical models can be used for describing and estimating the influence of physical training on athletic performance (Taha and Thomas 2003; Borresen and Lambert 2009). The first model proposed by Banister et al. (1975) considered the input dose effect that training has on the response elements of fitness and fatigue. The difference between these variables was suggested to reflect the performance of an athlete at a given time. This simplified model was shown as

$$\text{Performance} = \text{Fitness} - \text{Fatigue}$$

This basic model has been shown to significantly fit the training responses of athletes undertaking swimming, running, cycling, triathlon and hammer throwing [for review see (Taha and Thomas 2003)]. The accuracy of such models have since been refined (Busso 2003), with the introduction of time invariant parameters to take into account the accumulative effects of fatigue. This model showed a significantly improved fit compared with previous models described in the literature (Busso et al. 1991; Banister et al. 1975; Calvert et al. 1976).

Despite these refinements in modelling techniques, systems models have been unable to consistently predict performance on an individual basis in a real-world setting (Busso and Thomas 2006; Taha and Thomas 2003). This may be attributed to the lack consensus as to the most appropriate method quantifying training load, performance, fitness and fatigue. At present, the most commonly used methods for quantifying fitness and fatigue parameters include HR information at rest and during exercise, as well as blood lactate, biochemical markers and subjective questionnaires (Hooper and Mackinnon 1995; Lambert and Borresen 2006). More recently, heart rate variability (HRV) has been shown to reflect changes in the autonomic nervous system and has been suggested as a useful tool for measuring training adaptation and fatigue (Aubert et al. 2003; Achten and Jeukendrup 2003; Buchheit et al. 2010). Although HRV has been shown to be useful in guiding training in recreational athletes (Kiviniemi et al. 2009; Kiviniemi et al. 2007), the efficacy of HRV information in monitoring training in well-trained adult athletes remains unclear.

Since the influence of different training load inputs into systems models is not well understood, the purpose of the present investigation was to compare currently used methods for quantifying internal (i.e., TRIMP, sRPE) and external (i.e., rTSS) training loads using the time-invariant systems model previously described (Busso 2003). Furthermore, this investigation also compared the influence of a variety of fitness and fatigue markers on goodness-of-fit predictions within the model.

Methods

Participants

Seven trained endurance runners (mean ± SD, $\dot{V}O_2$max: 51.7 ± 4.5 mL kg^{-1} min^{-1}, age: 35.8 ± 9.1 years) volunteered to participate in this study. Each athlete had trained prior to the commencement of the study and completed between 5 and 10 sessions per week. All athletes were fully informed of the potential risks and benefits associated with participation. Written informed consent was obtained by each athlete and ethical approval was granted by the University Human Research Ethics Committee for all procedures.

Experimental design

A modelling post facto longitudinal research design was used to compare the performance, fitness and fatigue responses during a 15-week period. A 15-week period was chosen to complete sufficient performance fitness and fatigue measures for within-individual correlations between actual and modelled responses. Individual training dose was measured via a Polar RS 800 HR monitor with a Polar s3 foot pod stride sensor™ W.I.N.D. (Polar Oy, Polar Electro, Kempele, Finland). Foot pods were calibrated at the commencement of the study and at the mid-way testing point according to the recommendations of the manufacturer. Running speed, distance and heart rate (HR) were recorded during each session. In addition, perception of effort was also measured following each training session using a rating of perceived exertion (RPE) according to the category ratio scale (CR ten-scale) of Borg et al. (1985).

Throughout the investigation period, selected performance, physiological and psychological tests were completed by the athletes. Specifically, a 1,500-m running time-trial and a standardised submaximal HR test

(HR_{submax}) were performed weekly. These tests were completed at the same time of day (16:00) following a standardised warm-up. Upon waking on the morning of these tests, each athlete completed a HRV test and recorded resting HR (HR_{rest}) values. Directly following the HRV test, each athlete was required to complete the POMS psychological questionnaire to assess psychological state (McNair et al. 1971).

Testing procedures

Each athlete was tested at the same time of day following a full day of rest. Athletes were administered written guidelines on carbohydrate consumption prior to the study and were asked to standardise their diet for the 24 h prior to each testing session.

Physiological measures

Maximal oxygen uptake ($\dot{V}O_2$max) was measured using an incremental treadmill running test to exhaustion on a motorised treadmill (Startrac Unisen Inc., USA). Following a 5-min warm up at 8 km h^{-1}, the workload protocol commenced at a speed of 8.5 km h^{-1}. The workload was increased by 1.5 km h^{-1} every 4 min until volitional fatigue. The athletes received a 1-min rest period between workloads. Maximum oxygen uptake was measured using a gas analysis System (Physio-dyne® Fitness Instrument Technologies, Quogue, NY, USA) and was calibrated before and after each test with reference and calibration gases of known concentrations. The pneumotach was calibrated with ambient air using a 3-L syringe (Hans Rudolph Inc, Kansas City, USA). The reliability of $\dot{V}O_2$max measures for this laboratory were acceptable [coefficient of variation (CV %, ±90 % confidence intervals (CI) = 2.5 (1.8–4.3)].

Performance measures

Each athlete completed a 1,500-m time-trial once a week for the duration of the study. Maximal effort time trials have previously been suggested as ideal for evaluation of performance (Jeukendrup et al. 1996). Prior to the time-trial each athlete was required to complete a standardised warm up consisting of an 800-m jog, followed by the submaximal fitness test. Each athlete was then required to run 1,500 m on a tartan track in the shortest time possible. To minimise the effects of pacing, the athletes began the time trial in a staggered start with 10 s between each participant. The athletes were not informed of their lap splits and given equal verbal encouragement. The test–retest reliability of the 1,500-m time-trial was high [% CV (90 % CI) = 2.7 (2.1–3.5)].

Submaximal fitness test

The HR_{submax} required athletes to complete a 1,500-m circuit at a standardised pace of 210 m min^{-1}. This test was completed on a weekly basis prior to the 1,500-m time-trial. Pacing was achieved using instantaneous feedback from the Polar s3 stride sensor™ W.I.N.D. and RS800 running computer (Polar Oy, Polar Electro, Kempele, Finland). Heart rate information was collected during the entire exercise bout. Submaximal HR response was taken as the mean HR during the final 30 s of the exercise bout. A measure of RPE (6–20 scale) was also collected at the completion of the test (Borg 1973).

Training load quantification

A variety of methods were used to quantify the training load of the athletes during each exercise bout. The methods selected utilise a variety of training responses including HR-based information, perception of effort as well as external load measures. The TRIMP method proposed by Banister (1991) was used to quantify internal training load. This method was calculated using the following equation:

$$\mathrm{TRIMP} = D(\Delta\mathrm{HR\ ratio})\mathrm{e}^{b(\Delta\mathrm{HR\ ratio})}, \quad (1)$$

where D = duration of training session, b = 1.67 for females and 1.92 for males and (ΔHR ratio) is determined using the following equation:

$$\Delta\mathrm{HR\ ratio} = (\mathrm{HR_{ex}} - \mathrm{HR_{rest}})/(\mathrm{HR_{max}} - \mathrm{HR_{rest}}), \quad (2)$$

where HR_{rest} = the average heart rate during rest and HR_{ex} = the average HR during exercise (Morton et al. 1990).

The sRPE method proposed by Foster et al. (1995) was also used to quantifying internal training load. This method requires athletes to subjectively rate the intensity of the entire training session using a RPE according to the category ratio scale (CR-10) of Borg et al. (1985). The RPE value was then multiplied by the total duration (min) of the training session. To ensure the athletes reported a RPE for the entire training session, RPE measures were taken 30 min following the completion of the session. Standard instructions and anchoring procedures were explained during the familiarisation process (Noble and Robertson 1996).

Daily training load was also quantified from velocity data recorded with a Polar RS 800 running computer and a calibrated Polar s3 foot pod stride sensor™ W.I.N.D. (Polar Oy, Polar Electro, Kempele, Finland) and expressed as a rTSS. The rTSS is calculated using the gravity ordered velocity stress score (GOVSS) algorithm according to previously described methods (Skiba 2006). In general, this measure is a TRIMP measure derived from external load data and is calculated similar to other training impulse measures that combine exercise duration and intensity (Allen and Coggan 2006; McGregor et al. 2009). McGregor et al. (2009) have previously described this method for running using training data collected from training logs. The test–retest reliability of the foot pod units for measuring distance was determined in pilot testing was high [% CV (90 % CI) = 2.6 (2.1–3.5)].

Fitness

A variety of methods were used to measure fitness or adaptation to training. HR_{rest} was collected upon waking, each morning of the performance tests. The measurement was recorded as the minimum HR obtained during 5 min of lying in a supine position. Additionally HR_{submax} and submaximal RPE (RPE_{submax} 6–20) taken during the standardised submaximal warm-up were also used as a measure of fitness.

Fatigue

Several methods were used to assess the fatigue of the athletes in the study. Prior to each weekly performance test, each athlete completed mood states (POMS) questionnaire (McNair et al. 1971). This is a 65-item inventory of six subscales: tension-anxiety, depression-dejection, anger-hostility, vigour-activity, fatigue-inertia and confusion-bewilderment. The athlete used a five-point scale (0–'not at all' to 4–'extremely') to respond to each term according to the question, "How have you been feeling for the past week including today?" A specific sub-analysis of the POMS data using only the responses to the fatigue-inertia subset was completed and reported as a score out of 20.

Heart rate variability was recorded upon waking on each morning prior to the performance test. Polar RS 800 HR monitors (Polar Oy, Polar Electro, Kempele, Finland) were used to record R–R intervals at a timing accuracy of 2 ms. The measurement started with 5 min of lying supine followed by 5 min of standing. The R–R interval data were downloaded to a personal computer using Polar ProTrainer5 software (version 5.40.171, Finland) and analysed using Kubios HRV software (version 2.0, Biosignal Analysis and Medical Imaging Group, Finland). Occasional ectopic beats were automatically replaced with interpolated adjacent R–R interval values. Power spectral analysis was performed on the data using a traditional Fast Fourier Transform algorithm and a parametric method based on autoregressive time series modelling to establish power (ms^2) in distinct frequency bands: the high frequency (HF) range (HF = 0.15–0.40 Hz) and the low frequency (LF) range (LF = 0.04–0.15 Hz). The HRV ratio was determined as the HF/LF. In addition, the standard deviation of instantaneous beat-to-beat R–R interval variability measured from Poincaré plots (SD1) (Huikuri et al. 1996) was calculated during the last 3 min of the 5-min standing period as a vagal-related HRV index (Tulppo et al. 1996). Due to technical difficulty, only data were collected from six participants.

Fitting the model

The time-invariant systems model used in this study has been previously described (Busso 2003). This model assumes that the gain term of the fatigue effect is mathematically related to the training dose using a first-order filter. Performance output can be described as

$$\hat{p}^n = p^* + k_1 \sum_{i=1}^{n-1} w^i \mathrm{e}^{-(n-i)/\tau 1} - \sum_{i=1}^{n-1} k_2^i w^i \mathrm{e}^{-(n-i)/\tau 2}$$

in which the value of k_2 at day i is estimated by mathematical recursion using a first-order filter with a gain terms k_3 and a time constant τ^3:

$$k_2^i = k_3 \sum_{j=1}^{i} w^j \mathrm{e}^{-(i-j)/\tau 3}$$

The parameters for the model were determined by fitting the model performances with actual performances using the least squares method using the Solver function in Microsoft Excel (Microsoft, Redmond, USA). The set of model parameters was determined by minimizing the residual sum of squares between modeled performance and actual performances (RSS):

$$\mathrm{RSS} = \sum_{n=1}^{N} \left[p^n - \hat{p}^n\right]^2,$$

where n takes the N value corresponding to the days of measurement of the actual performance. Successive minimization of the RSS with a grid of values for each time constant gave the total set of model parameters.

Statistical analyses

Models were developed for each athlete and the goodness of fit values were used to determine the best fitting model from either the rTSS, HR- or RPE-based training load measures. The coefficient of determination (r^2), giving the variation explained by the model, was calculated to establish the goodness of fit for the model. Within-individual correlations between the various actual and predicted measures of performance, fitness and fatigue were analysed using the Pearson's correlation coefficient. The following criteria were adopted to interpret the magnitude of the correlation (r) between test measures: <0.1 trivial, 0.1–0.3 small, 0.3–0.5 moderate, 0.5–0.7 large, 0.7–0.9 very large, and 0.9–1.0 almost perfect. Differences between the mean within-individual correlations between each of the methods were assessed using a one-way ANOVA with Tukey HSD post hoc to locate differences. Statistical significance was set at $p < 0.05$. All data are presented as mean ± SD unless otherwise stated.

Results

Five hundred and forty-two individual training sessions were analysed during the investigation period. Individuals completed an average of 77 ± 20 individual training sessions with the weekly training duration of 389 ± 168 min and weekly training distance of 68 ± 36 km. This training resulted in a 5.4 ± 2.6 % improvement in 1,500-m time-trial performance (pre 5:33 ± 0:30, post: 5:08 ± 0:31 min:s). Modelled performance significantly correlated with actual performance in each athlete, with average correlations being 0.70 ± 0.11 (Table 1), 0.60 ± 0.10 (Table 2) and 0.65 ± 0.13 (Table 3) for the rTSS, sRPE and TRIMP input methods, respectively. The within-individual correlations between each of these methods were not significantly different between methods ($p = 0.33$).

Maximum oxygen uptake was not significantly changed (51.7 ± 4.5 mL kg^{-1} min^{-1} vs. 51.3 ± 6.1 mL kg^{-1} min^{-1}, $p > 0.05$). Fitness measured by HR_{submax} decreased non-significantly from 77.9 ± 5.1 to 74.8 ± 5.5 % during the study ($p > 0.05$, Fig. 1a). Tables 2 and 3 show that there were moderate correlations between modelled and actual fitness measures (HR_{submax}) for the sRPE (−0.43 ± 0.37) and TRIMP (−0.48 ± 0.39) methods, respectively. Additionally, RPE_{submax} did not significantly change (11.7 ± 1.0 to 11.3 ± 1.4, from week 1 to week 15) during the test period and small correlations with modelled fitness using each of the training load input methods ($p > 0.05$, Fig. 1b). Finally, there were no significant group changes in HR_{rest} during the study. There were only trivial correlations between HR_{rest} and modelled fitness outcomes using any of the different training load measures (Tables 1, 2, 3).

Using the fatigue subset of the POMS questionnaire, the athletes reported substantial fluctuations in fatigue status

Table 1 Individual and mean ± SD correlations (r) between modelled and actual performance, fitness and fatigue using rTSS as the training load input method

Subject	Performance	Relationships with modelled fitness			Relationships with modelled fatigue			
		HR_{submax}	HR_{rest}	RPE_{submax}	POMS	Fatigue	HRV ratio	SD1
1	0.75	−0.14	−0.33	0.36	−0.10	0.06	0.06	0.02
2	0.55	0.02	−0.48	0.20	0.16	0.41	0.75	−0.59
3	0.60	−0.57	0.06	−0.57	−0.16	−0.33	0.13	−0.75
4	0.82	−0.41	0.13	−0.01	−0.02	0.16	−0.03	−0.67
5	0.63	0.29	−0.44	−0.62	−0.27	−0.27	0.00	−0.01
6	0.69	−0.74	0.99	−0.76	−0.19	−0.43	–	–
7	0.84	−0.72	−0.35	−0.23	0.65	0.43	0.44	−0.40
Mean ± SD	0.70 ± 0.11	−0.32 ± 0.39	−0.06 ± 0.52	−0.23 ± 0.43	0.01 ± 0.31	0.00 ± 0.35	0.23 ± 0.31	−0.40 ± 0.33

Table 2 Individual and mean ± SD correlations (r) between modelled and actual performance, fitness and fatigue using sRPE as the training load input method

Subject	Performance	Relationships with modelled fitness			Relationships with modelled fatigue			
		HR_{submax}	HR_{rest}	RPE_{submax}	POMS	Fatigue	HRV ratio	SD1
1	0.52	−0.02	−0.39	0.26	−0.20	−0.32	−0.05	0.20
2	0.62	0.10	−0.29	0.25	0.20	0.24	0.71	−0.76
3	0.69	−0.52	0.16	−0.53	0.09	0.05	0.07	−0.69
4	0.53	−0.34	0.14	−0.11	−0.07	0.15	−0.08	−0.64
5	0.53	−0.95	−0.15	−0.41	−0.01	0.42	0.16	−0.08
6	0.57	−0.69	0.94	−0.67	−0.19	−0.37	–	–
7	0.77	−0.58	−0.25	−0.21	−0.74	0.02	−0.74	−0.83
Mean ± SD	0.60 ± 0.10	−0.43 ± 0.37	0.02 ± 0.46	−0.20 ± 0.36	−0.13 ± 0.30	0.03 ± 0.29	0.01 ± 0.47	−0.47 ± 0.42

Table 3 Individual and mean ± SD correlations (*r*) between modelled and actual performance, fitness and fatigue using Banister's TRIMP as the training load input method

Subject	Performance	Relationships with modelled fitness			Relationships with modelled fatigue			
		HR_{submax}	HR_{rest}	RPE_{submax}	POMS	Fatigue	HRV ratio	SD1
1	0.78	−0.14	−0.37	0.29	−0.17	−0.02	0.09	−0.05
2	0.55	0.04	−0.47	0.17	0.18	0.48	0.76	−0.72
3	0.66	−0.56	0.09	−0.57	−0.11	−0.29	0.00	−0.65
4	0.67	−0.27	0.18	−0.17	0.21	0.35	0.08	−0.65
5	0.54	−0.95	−0.12	−0.40	0.05	−0.39	0.30	−0.29
6	0.51	−0.72	0.95	−0.76	−0.36	−0.28	–	–
7	0.86	−0.73	−0.38	−0.21	−0.67	−0.07	−0.68	−0.90
Mean ± SD	0.65 ± 0.13	−0.48 ± 0.36	−0.02 ± 0.49	−0.24 ± 0.38	−0.12 ± 0.31	−0.03 ± 0.33	0.09 ± 0.47	−0.54 ± 0.31

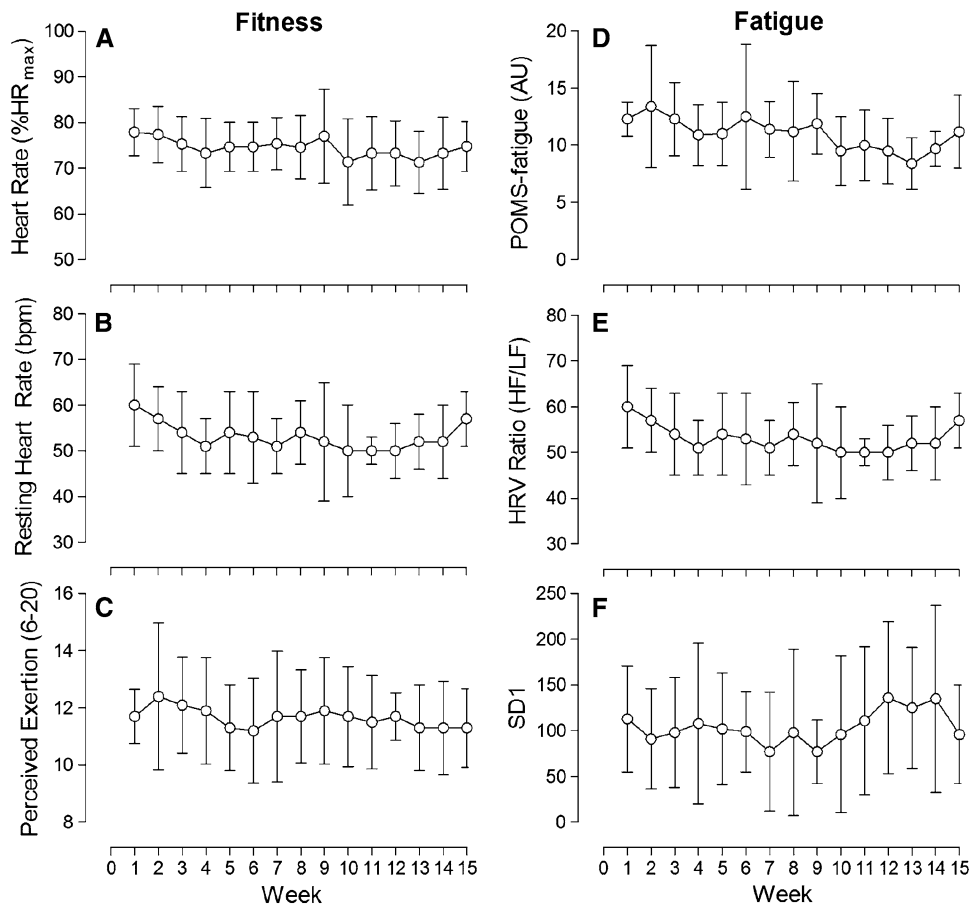

Fig. 1 Mean ± SD fitness (**a** HR_{submax}, **b** HR_{rest} and **c** RPE_{submax}) and fatigue (**d** POMS-fatigue, **e** HRV ratio and **f** SD1) measures during the 15-week study

across the 15-weeks training period (range 7–20). However, there were no group changes in the POMS fatigue scores during the study period (Fig. 1d). Notably, however, these measures showed trivial correlations with modelled fatigue using each of the various training load input methods. Similarly, the HRV ratio was unrelated to the modelled fatigue in this study (Fig. 1e). In contrast, there were moderate-to-large correlations between modelled and actual fatigue measures using SD1 when both sRPE (−0.48 ± 0.39, Table 2) and TRIMP (−0.59 ± 0.31, Table 3) methods were used as the input methods. There were no significant group changes in SD1 during the study (Fig. 1f).

Discussion

The purpose of this study was to compare the validity of three common methods for quantifying training load (i.e., rTSS, sRPE & TRIMP) using a systems model approach. This study also examined the influence of the various training load inputs on actual training outcomes (i.e., performance, fitness and fatigue) compared with those predicted by the model. The main results demonstrated large correlations between each of the different methods for quantifying training loads and modelled running performance. Notably, the relationships between rTSS and performance were slightly larger than both methods for quantifying internal training load (i.e., sRPE and TRIMP). These results contrast with common training theory which suggests that it is the internal stimulus that determines training adaptation (Booth and Thomasson 1991; Viru and Viru 2000; Impellizzeri et al. 2005). A possible explanation for the reduced relationship between the internal training load methods and performance may be attributed to the relatively poor measurement reliability of the CR10 Borg scale for estimating exercise intensity and/or the inability of the generic HR methods to adjust for individual fitness/performance characteristics.

We have recently observed high measurement error of the session-RPE using the CR-10 scale (CV 28.1 %) compared with HR (CV 3.9 %) during endurance cycling (Wallace et al. 2013) and with intermittent running protocols (CV 31.9 %) (Scott et al. 2013). More specifically, we have observed a poor sensitivity to small changes in intensity during moderate-to-hard exercise with this scale. In the present study, 63 % of all training sessions were rated as 'moderate-to-hard', which may have influenced the fit of the between modelled and actual performance. Therefore, the combination of the higher measurement error of the CR-10 RPE scale coupled with the scales reduced sensitivity at moderate-to-high intensities may explain the reduced relationship between sRPE and predicted running performance.

Relatively poor levels of test–retest reliability for the Banisters' generic HR TRIMP (15.6 % CV) have also been reported in healthy individuals undertaking steady-state cycle training in a laboratory setting (Wallace et al. 2013). Notably, however, the reliability of the HR TRIMP method improved when adjusted to account for individual differences in ventilatory thresholds [i.e., Banister's TRIMP (15.6 % CV) versus Lucia's TRIMP (10.7 % CV)] (Wallace et al. 2013). Moreover, it has also been shown that training loads calculated from individual HR-lactate relationships (i.e., the individual HR-based TRIMP) rather than the Banister's TRIMP were related to both changes in fitness (running speed and 2 and 4 mmol L^{-1} blood lactate) and running performance (5,000 and 10,000-m time trials) in eight long-distance runners during a 9-week training period (Manzi et al. 2009). Taken collectively with these previous observations, the present findings suggest that the relationship between HR TRIMP methods and training outcomes may be improved if the TRIMP calculation is modified to account for individual physiological thresholds. The absence of this individualised process in the present study may explain the reduced relationship between the TRIMP method and endurance running performance.

Running speed was also collected which was then expressed as an arbitrary measure of external training load using the rTSS. Similar to the individualised TRIMP methods, the rTSS calculates training dose by multiplying training duration with training intensity. However, unlike TRIMP and sRPE, the rTSS calculates training intensity using an intensity factor calculated from a percentage of an individuals' 'threshold' running pace. The rTSS is based on the TSS, where exercise intensity is calculated from normalised power measures (Allen and Coggan 2006). The TSS was originally adapted from Banister's TRIMP and has been reported to be appropriate for monitoring individual training (Allen and Coggan 2006). Whilst the present study is the first to examine the relationship between rTSS and running performance, previous studies have successfully used the TSS to quantify training load in cycling (MacLeod and Sunderland 2009) and running (McGregor et al. 2009). In this study, the relationship between rTSS and modelled performance was the strongest of each of the training load methods. The improved associations between the rTSS compared with the sRPE and HR TRIMP methods are likely explained by the ability of the rTSS to adjust for differences in individual performance characteristics and the improved measurement reliability of the foot pods compared with the other training load quantification methods.

Whilst internal training load methods appear to be theoretically robust for quantifying training load (Impellizzeri et al. 2005; Viru and Viru 2000), more work is required to determine appropriate weighting factors for HR methods and to increase measurement reliability when using perceptual measures to assess exercise intensity. It is possible that the substitution of the CR-10 scale with the CR100 or 6–20 RPE scales may, in part, address this issue. Indeed, whilst the present study shows that the rTSS relates best to predicted performance, further research is required to assess the efficacy of this tool in a different cohort of athletes with varying training goals.

The second purpose of this investigation was to examine the relationships between actual and predicted measures of fitness and fatigue using a systems modelling approach for assessing training responses. Moderate correlations were observed between HR_{submax} and predicted fitness when TRIMP and sRPE were used as training inputs. It is widely recognised that HR_{submax} decreases with endurance training in adult populations with these changes being largely attributed to decreases in sympathetic activity of the heart (Carter et al. 2003) and increased plasma volume (Covertino 1991). Despite this, several previous studies have only shown small to moderate changes in HR_{submax} in trained endurance athletes following intensive training periods (Buchheit et al. 2010; Uusitalo et al. 1998; Swaine et al. 1994). It was suggested that training elicits differing effects on indices of fitness which limit the efficacy of HR_{submax} as a marker of cardiovascular fitness. These previous findings may explain the moderate relationship between HR_{submax} and predicted fitness using HR and sRPE load input methods in the present study. Collectively, the current findings support the use of HR_{submax} as a valid simple fitness test for assessing fitness changes in endurance runners; however, the moderate strength of the correlations indicate that other measures may be required to accurately monitor how an athlete is responding to training.

Many studies have shown HR_{rest} to decrease slightly following endurance training (Uusitalo et al. 1998; Buchheit et al. 2010; Wilmore et al. 2001). This phenomenon has been attributed to decreases in intrinsic rhythmicity of the heart and an increase in the predominance of parasympathetic control (Smith et al. 1989). Despite this, no relationships were observed between actual HR_{rest} and modelled fitness in the runners in the present study. These findings are in accordance with previous research showing no decreases in HR_{rest} following periods of intensified training (Melanson and Freedson 2001; Fry et al. 1992; Zavorsky 2000). The lack of agreement between the findings between these studies may be due to the differences in training undertaken by the participants and/or the inter-individual differences of the training status of the athletes. Furthermore since HR_{rest} can also be influenced by factors such as age, hydration and environmental conditions, the present study does not support the use of HR_{rest} as an idiosyncratic marker of cardiovascular fitness.

Heart rate variability represents the beat-to-beat variation in R–R intervals and is widely used as a non-invasive measurement of autonomic nervous system activity (Achten and Jeukendrup 2003). Recent research has focussed on the effectiveness of HRV for assessing training adaptation at the level of the individual athletes (Hautala et al. 2010; Buchheit et al. 2011; Buchheit et al. 2010) and guiding training on an individual basis (Kiviniemi et al. 2009, 2007). We observed moderate-to-large correlations between instantaneous beat-to-beat variability (SD1) and predicted fatigue when sRPE and TRIMP were used as training load inputs. However, in contrast, there were no significant relationships between predicted fatigue and most other HRV indices (i.e., HF, LF, rMSSD, SD2, data not reported) within time and frequency domains. The lack of association between these measures with modelled fatigue may be due to the level of accumulated fatigue which may not have been sufficient to alter changes in cardiac autonomic function or the large noise from these measure when assessed in a field setting. Whilst the SD1 results indicate that this measure may provide a non-invasive and objective method for assessing short-term fatigue in endurance athletes, the majority of HRV measures have low practical efficacy when used to monitor training in the field setting.

Psychological tools such as the profile of mood states (POMS) questionnaire have been used to assess mood states in athletic populations (Martin et al. 2000; Hooper et al. 1997). From these investigations, several links have been made between changes in POMS fatigue scores in athlete undertaking intensified training periods (Martin et al. 2000; Liederbach et al. 1992) or exhibiting symptoms of overtraining or staleness (Hooper et al. 1997). However, in the present study, a poor relationship was observed between the fatigue subset of POMS and predicted fatigue. These results are similar to one earlier study examining the relationship between the fatigue subset of POMS and predictions of fatigue using a modelling approach (Wood et al. 2005). The previous authors reported a moderate correlation between the fatigue subset of POMS and predicted fatigue, but only provided data from a single runner where ten fatigue (POMS) measures were taken over a 12-week training period. The reduced strength in this correlation was attributed to the inability of the fatigue subset to detect the source of fatigue (i.e., global fatigue vs. training induced fatigue). The poorer correlation in the present study compared with Wood et al. (2005) may be explained by the increased number of POMS measures taken from the runners in this study (10 versus 15). Importantly, some individuals in this study did exhibit large-to-moderate correlations between predicted fatigue for the POMS-subset with each of the training load inputs. Collectively, however, the lack of relationship between the POMS fatigue subsets and predicted fatigue suggests that these measures are insensitive to small changes in cumulative fatigue in trained athletes. It may be that other psychometric tools that assess sport-related fatigue, rather than mood states, such as the RESTQ-sport (Kellmann and Kallus 1993), the daily analysis for life demands of athletes (DALDA) (Rushall 1990) or the training distress questionnaire (Main and Grove 2009) are more appropriate for assessing training-related fatigue in endurance athletes. In particular, the RESTQ-sport may be more appropriate for monitoring training as it assesses components of both stress and recovery. It also possible that a limitation of the model used in this study is that it does not account for the influence of recovery practices on performance, fitness and fatigue.

There are some limitations of this study that must be acknowledged. First, the efficacy of the mathematical model itself consists of several limitations that may reduce its adequacy. It is generally reported that a large number of performance tests are required to gain a stable fit in the model. Indeed, it has been suggested that between 20 and

200 performance tests are required within a short time to obtain a robust model (Taha and Thomas 2003). Whilst we measured performance 15 times, which is high in comparison with most other modelling studies, the stability of the model maybe inadequate to truly describe the relationships reliably. Since it would be practically unrealistic to substantially increase the number of maximal performance tests with athletes in a normal training environment, this approach may only be limited to laboratory studies. Second, although reasonable attempts were made to control the performance test environment, not all factors could be controlled. It is possible that factors such as the climate (i.e., temperature, wind, etc.) and athlete motivation may have affected 1,500-m time-trial performance independent of the other factors assessed in this study and, therefore, also influences the relationships between the modelled and actual performance. Finally, the athletes in this study were endurance athletes who do not regularly compete in relatively short events such as 1,500 m-time-trial or train for these events. It is, therefore, possible that the lack of specificity in the performance test may have reduced the strength of the training impulse-performance outcome for these athletes.

Conclusions

The main findings of this study are that there were large relationships between each of the different methods for quantifying training loads and modelled running performance. However, notably, the relationships between rTSS and modelled 1,500-m time-trial performance were slightly larger than both methods for quantifying internal training load (i.e., sRPE and TRIMP). From a practical point of view, these results suggest that it is important to select a reliable measure of training load and that methods for quantifying load should be adjusted to account for individual athlete characteristics. However, other factors such as practical usefulness need to be considered when monitoring athletes (particularly large groups). Therefore, the HR and in particular, the session-RPE method may be suitable practical choices for monitoring load in a training environment. The moderate relationships between some of the fitness (i.e., %HRmax) and fatigue (i.e., SD1) variables indicate that these markers may be useful for monitoring athletes to better understand their response to the training process. Taken together, these results suggest that each of the methods used in this study is appropriate for monitoring training dose in endurance athletes. However, coaches and scientists should be aware that ideally, the training load measures should be reliable and account for differences in individual physiological/performance characteristics of athletes.

Acknowledgments The authors would like to thank Dr. Chris Barnes from the Australian Institute of Sport for his expertise in developing the customised modelling spreadsheets used in this study.

References

Achten J, Jeukendrup AE (2003) Heart rate monitoring: applications and limitations. Sports Med 33(7):517–538

Allen H, Coggan A (2006) Training and racing with a power meter. Velo Press, Boulder

Aubert AE, Seps B, Beckers F (2003) Heart rate variability in athletes. Sports Med 33(12):889–919 (pii 33123)

Banister EW (1991) Modeling elite athletic performance. In: Green HJ, McDougal JD, Wenger HA (eds) Physiological testing of elite athletes. Human Kinetics, Champaign, pp 403–424

Banister EW, Calvert TW, Savage MV, Bach T (1975) A systems model of training for athletic performance. Aust J Sports Med Exerc Sci 7:57–61

Booth FW, Thomasson DB (1991) Molecular and cellular adaptations of muscle in response to exercise: perspectives of various models. Physiol Rev 71(2):541–585

Borg G (1973) Perceived exertion: a note on “history” and methods. Med Sci Sports 5(2):90–93

Borg GAV, Hassmen P, Langerstrom M (1985) Perceived exertion in relation to heart rate and blood lactate during arm and leg exercise. Eur J Appl Physiol 65:679–685

Borresen J, Lambert MI (2009) The quantification of training load, the training response and the effect on performance. Sports Med 39(9):779–795. doi:10.2165/11317780-000000000-00000

Buchheit M, Chivot A, Parouty J, Mercier D, Al Haddad H, Laursen PB, Ahmaidi S (2010) Monitoring endurance running performance using cardiac parasympathetic function. Eur J Appl Physiol 108(6):1153–1167. doi:10.1007/s00421-009-1317-x

Buchheit M, Simpson MB, Al Haddad H, Bourdon PC, Mendez-Villanueva A (2011) Monitoring changes in physical performance with heart rate measures in young soccer players. Eur J Appl Physiol. doi:10.1007/s00421-011-2014-0

Busso T (2003) Variable dose-response relationship between exercise training and performance. Med Sci Sports Exerc 35(7):1188–1195

Busso T, Thomas L (2006) Using mathematical modeling in training planning. Int J Sports Physiol Perform 1(4):400–405

Busso T, Carasso C, Lacour JR (1991) Adequacy of a systems structure in the modeling of training effects on performance. J Appl Physiol 71(5):2044–2049

Calvert TW, Banister EW, Savage MV, Bach T (1976) A systems model of the effects of training on physical performance. IEEE Trans Syst Man Cybern 6:94–102

Carter JB, Banister EW, Blaber AP (2003) Effect of endurance exercise on autonomic control of heart rate. Sports Med 33(1):33–46

Covertino VA (1991) Blood volume: its adaption to endurance training. Med Sci Sports Exerc 23:1338–1448

Foster C, Hector LL, Welsh R, Schrager M, Green MA, Snyder AC (1995) Effects of specific versus cross-training on running performance. Eur J Appl Physiol 70(4):367–372

Fry RW, Morton AR, Garcia-Webb P, Crawford GP, Keast D (1992) Biological responses to overload training in endurance sports. Eur J Appl Physiol 64(4):335–344

Hautala AJ, Karjalainen J, Kiviniemi AM, Kinnunen H, Makikallio TH, Huikuri HV, Tulppo MP (2010) Physical activity and heart rate variability measured simultaneously during waking hours. Am J Physiol Heart Circ Physiol 298(3):H874–H880. doi:10.1152/ajpheart.00856.2009

Hooper SL, Mackinnon LT (1995) Monitoring overtraining in athletes: recommendations. Sports Med 20(5):321–327

Hooper SL, Mackinnon LT, Hanrahan S (1997) Mood states as an indication of staleness and recovery. Int J Sports Psychol 28(1):1–12

Huikuri HV, Seppanen T, Koistinen MJ, Airaksinen J, Ikaheimo MJ, Castellanos A, Myerburg RJ (1996) Abnormalities in beat-to-beat dynamics of heart rate before the spontaneous onset of life-threatening ventricular tachyarrhythmias in patients with prior myocardial infarction. Circulation 93(10):1836–1844

Impellizzeri FM, Rampinini E, Marcora SM (2005) Physiological assessment of aerobic training in soccer. J Sports Sci 23(6):583–592

Jeukendrup AE, Saris WHM, Brouns F, Kester ADM (1996) A new validated endurance performance test. Med Sci Sports Exerc 28(2):266–270

Jobson SA, Passfield L, Atkinson G, Barton G, Scarf P (2009) The analysis and utilization of cycling training data. Sports Med 39(10):833–844. doi:10.2165/11317840-000000000-00000

Kellmann M, Kallus KW (1993) The recovery-stress-questionnaire: a potential tool to predict performance in sports. In: Nitsch JR, Seiler R (eds) Movement and sport: psychological foundations and effects. Academia, Sankt Augustin, pp 242–247

Kiviniemi AM, Hautala AJ, Kinnunen H, Tulppo MP (2007) Endurance training guided individually by daily heart rate variability measurements. Eur J Appl Physiol 101(6):743–751. doi:10.1007/s00421-007-0552-2

Kiviniemi AM, Hautala AJ, Kinnunen H, Nissila J, Virtanen P, Karjalainen J, Tulppo MP (2009) Daily exercise prescription based on heart rate variability among men and women. Med Sci Sports Exerc. doi:10.1249/MSS.0b013e3181cd5f39

Lambert MI, Borresen J (2006) A theoretical basis of monitoring fatigue: a practical approach for coaches. Int J Sports Sci Coach 1(4):371–387

Liederbach M, Gleim GW, Nicholas JA (1992) Monitoring training status in professional ballet dancers. J Sports Med Phys Fit 32(2):187–195

MacLeod H, Sunderland C (2009) Fluid balance and hydration habits of elite female field hockey players during consecutive international matches. J Strength Cond Res 23(4):1245–1251. doi:10.1519/JSC.0b013e318192b77a

Main LC, Grove JR (2009) A multi-component assessment model for monitoring training distress among athletes. Eur J Sport Sci 9(4):195–202

Manzi V, Iellamo F, Impellizzeri F, D'Ottavio S, Castagna C (2009) Relation between individualized training impulses and performance in distance runners. Med Sci Sports Exerc 41(11):2090–2096. doi:10.1249/MSS.0b013e3181a6a959

Martin DT, Andersen MB, Gates W (2000) Using profile of mood states (POMS) to monitor high-intensity training in cyclists: group versus case studies. Sport Psychol 14:138–156

McGregor SJ, Weese RK, Ratz IK (2009) Performance modeling in an Olympic 1500-m finalist: a practical approach. J Strength Cond Res 23(9):2515–2523. doi:10.1519/JSC.0b013e3181bf88be

McNair DM, Lorr M, Droppleman LF (1971) EITS profile for mood states. Educational and Industrial Testing Service, San Diego

Melanson EL, Freedson PS (2001) The effect of endurance training on resting heart rate variability in sedentary adult males. Eur J Appl Physiol 85(5):442–449

Morton RH, Fitz-Clarke JR, Banister EW (1990) Modeling human performance in running. J Appl Physiol 69(3):1171–1177

Noble BJ, Robertson RJ (1996) Perceived exertion. Human Kinetics, Champaign

Rushall BS (1990) A tool for measuring stress tolerance in elite athletes. J Appl Sport Psychol 2(1):51–66

Scott TJ, Black CR, Quinn J, Coutts AJ (2013) Validity and reliability of the session-RPE method for quantifying training in Australian football: a comparison of the CR10 and CR100 scales. J Strength Cond Res 27(1):270–276. doi:10.1519/JSC.0b013e3182541d2e

Skiba PF (2006) Calculation of power output and quantification of training stress in distance runners: the development of the GOVSS algorithm. http://www.physfarm.com/govss.pdf. Accessed 5 May 2011

Smith ML, Hudson DL, Graitzer HM, Raven PB (1989) Exercise training bradycardia: the role of autonomic balance. Med Sci Sports Exerc 21(1):40–44

Swaine IL, Linden RJ, Mary DA (1994) Loss of exercise training-induced bradycardia with continued improvement in fitness. J Sports Sci 12(5):477–481

Taha T, Thomas SG (2003) Systems modelling of the relationship between training and performance. Sports Med 33(14):1061–1073

Tulppo MP, Makikallio TH, Takala TE, Seppanen T, Huikuri HV (1996) Quantitative beat-to-beat analysis of heart rate dynamics during exercise. Am J Physiol 271(1 Pt 2):H244–H252

Uusitalo AL, Uusitalo AJ, Rusko HK (1998) Exhaustive endurance training for 6–9 weeks did not induce changes in intrinsic heart rate and cardiac autonomic modulation in female athletes. Int J Sports Med 19(8):532–540

Viru A, Viru M (2000) Nature of training effects. In: Garret WE Jr, Kirkendall DT (eds) Exercise and sport science. Lippincott Williams Wilkins, Philadelphia, pp 67–95

Wallace LK, Impellizzeri FM, Slattery KM, Coutts AJ (2013) Establishing the criterion validity and reliability of common methods for quantifying training load. J Strength Cond Res (in press)

Wilmore JH, Stanforth PR, Gagnon J, Rice T, Mandel S, Leon AS, Rao DC, Skinner JS, Bouchard C (2001) Heart rate and blood pressure changes with endurance training: the HERITAGE Family Study. Med Sci Sports Exerc 33(1):107–116

Wood RE, Hayter S, Rowbottom D, Stewart I (2005) Applying a mathematical model to training adaptation in a distance runner. Eur J Appl Physiol 94:310–316

Zavorsky GS (2000) Evidence and possible mechanisms of altered maximum heart rate with endurance training and tapering. Sports Med 29(1):13–26