



Wattage training

[List Info](#)
[Join This List](#)
[Previous Message](#)
[All Messages](#)
[Next Message](#)

TSS and IF - at last!



Andy Coggan



Mar 13, 2003 14:49 PST

"A watt is not always a watt" – Dave Harris

"Not all kilojoules are created equal" – Andy Coggan

Statement of the problem:

At least in theory, one of the advantages of training and racing with a power meter is that doing so enables you to more precisely control the overall training load. By continuously recording power output, the exact demands of each workout can be more accurately quantified, and the intensity or duration (or both) of subsequent training sessions can be modified as necessary to avoid either under- or overtraining. Successful application of this approach, however, requires that the athlete or coach be able to quickly make sense out of the huge amounts of data that are amassed when power output (along with other variables, e.g., heart rate) is recorded every second or so during multi-hour training rides. This task is made more difficult by the fact that power is highly variable when cycling outdoors, such that the overall average power may give little insight into the actual stress imposed by a given workout. This is especially true for races, since the fluctuations in power normally resulting from hills, wind, etc., are further exaggerated by tactical considerations, e.g., by the need to maintain one's position in a large field, or by the need to initiate or respond to attacks. The issue is therefore how to best summarize or condense power meter data while still adequately capturing or reflecting the actual demands of each race or training session.

One solution to the above problem is to calculate the frequency distribution of power output, i.e., the percentage of total ride time when power falls within a certain range (e.g., between 200 and 250 W) or level/zone (e.g., within level 4). Such frequency distribution analyses can be useful, but have two major limitations:

- 1) a relatively large number of numeric values is still needed to represent a single training session. Such data are therefore best presented graphically (e.g., as a bar chart), and are themselves not readily amenable to further analysis. Furthermore, while large differences in power distribution are readily detectable using this approach, more subtle differences are harder to detect.
- 2) more importantly, such analyses do not (and in fact readily cannot) take into account how long each "foray" into a given power range or level actually lasts. This has significant implications with respect to physiological responses, as will be discussed below.

Another means of expressing power meter data that is utilized by some is to simply record the total work (in kJ) performed during a race or training session. Expressing the data in this manner can be helpful in understanding the overall energy demands of training and e.g., how this compares to energy intake (useful, for example, when an athlete is trying to alter their body composition). However, like keeping track of miles or hours of training, total work only provides a measure of overall training volume, and says nothing about the actual intensity of that training.

The limitations of currently available methods for analyzing power meter data files led me to try to develop an alternative approach, which is the topic of this post.

Check It Out!



Topica Channels

[Best of Topica](#)
[Art & Design](#)
[Books, Movies & TV](#)
[Developers](#)
[Food & Drink](#)
[Health & Fitness](#)
[Internet](#)
[Music](#)
[News & Information](#)
[Personal Finance](#)
[Personal Technology](#)
[Small Business](#)
[Software](#)
[Sports](#)
[Travel & Leisure](#)
[Women & Family](#)

Start Your Own List!

Email lists are great for debating issues or publishing your views.

- [Start a List Today!](#)

is the topic of this post.

Proposed solution: TSS and IF:

Dr. Eric Bannister has previously described a way of quantifying training load in terms of a HR-based “training impulse”, or TRIMPS, score:

TRIMPS = exercise duration x average HR x an intensity-dependent weighting factor

Since HR is essentially linearly related to oxygen uptake (metabolic rate), the product of the first two factors in the above equation is proportional to the amount of energy expended, or (since efficiency is relatively constant), work performed. The third term then takes into account the intensity of the exercise, since many physiological responses (e.g., glycogen utilization, lactate accumulation) increase non-linearly with increasing intensity.

Reasoning by analogy, it seemed logical that data from a power meter could be used to derive what I have called a “training stress score”, or TSS:

TSS = exercise duration x average power x an intensity-dependent weighting factor

Similar to TRIMPS, the product of the first two factors in the above equation is equal to the total work performed, whereas the “intensity factor” (IF) serves to account for the fact that the physiological stress imposed by performing a given amount of work (e.g., 1000 kJ) depends in part on the rate at which that work is performed (i.e., on the power output itself).

Clearly, for such an approach to have merit, the IF must have some basis in reality, i.e., the relative weight given to higher vs. lower intensity exercise cannot be determined at random, but must be based on the actual physiological “costs”. Furthermore, since the physiological responses to exercise at a given power output depend in part on the duration for which that power is maintained, this fact must be recognized as well. The algorithm used to determine the IF is therefore the key to the whole approach, and so this is where developmental effort was focused.

To derive an appropriate algorithm, I relied on blood lactate data collected from a large number of trained cyclists exercising at intensities both below and above their LT. This choice was made because many physiological responses (e.g., muscle glycogen and blood glucose utilization, catecholamine levels, ventilation) tend to parallel changes in blood lactate during exercise – in this context, then, blood lactate levels can be viewed as an overall index of physiological stress. To reduce variability between individuals, the data were normalized by expressing both the power output and the corresponding blood lactate level as a percentage of that measured at LT. The normalized data were then used to derive a best-fit curve. Perhaps not surprisingly, an exponential function provided the best fit, but a power function of the following form proved to be nearly as good:

blood lactate (% of lactate at LT) = power (% of power at LT)^{3.90};
R²=0.806, n=76

Based on these data, a 4th-order function was used in the algorithm for determining the IF (the exponent was rounded from 3.90 to 4.00 for simplicity's sake).

The other physiological knowledge that seemed necessary to incorporate into the algorithm for calculating IF was the fact that physiological responses to changes in exercise intensity are not instantaneous, but followed a characteristic time course. Because of this, exercise in which the intensity alternates every 15 seconds between a high and a low level (e.g., 400 and 0 W) results in physiological, metabolic, and perceptual responses nearly identical to steady-state exercise performed at the average intensity (e.g., 200 W). The specific reasons for this are beyond the scope of this discussion, but the important facts are 1) the half-lives (50% response time) of many physiological responses are directly or indirectly related to metabolic events in exercising muscle, and 2) such half-lives are typically on the order of 30 seconds. Thus, the decision was made to smooth power data using a 30 second rolling

the decision was made to smooth power data using a 30 second rolling average before applying the 4th order weighting as described above. Finally, the decision was made to 1) express the IF as a ratio of the "corrected" power obtained by smoothing/weighting to the individual's power at LT, and 2) normalize the TSS to the amount of work that could be performed during one hour of cycling at threshold power (=100 TSS "points"). While these last two steps are not necessary for comparisons within a given individual, they should make it easier for coaches or anyone dealing with multiple athletes to more quickly grasp the significance of a given value.

The steps required to calculate IF and TSS then become:

- 1) starting at 30 seconds, calculate a 30 second rolling average for power (data point by data point)
- 2) raise the values obtained in step 1 to the 4th power
- 3) take the average of all the values obtained in step 2
- 4) take the 4th root of the number obtained in step 3
- 5) divide the "corrected" power obtained in step 4 by the individual's power at LT – this decimal value is the IF
- 6) multiply the average power (uncorrected) for the workout by the duration (in seconds) to obtain the total work performed (in J)
- 7) multiply the total work by the IF (step) to derive the "raw" TSS
- 8) divide the "raw" TSS by the amount of work performed in one hour at LT (LT power x 3600 seconds) and multiply by 100 to obtain the final TSS

(These calculations are obviously too cumbersome to routinely perform on every power meter file, or part thereof, even when e.g., using a macro in Excel – hopefully, in the near future software will be available to automate the process.)

Applications:

The most obvious application of this method is to quantify the overall training load, in terms of the number of TSS points accumulated during a given period of time. (Indeed, this was the original purpose of developing it.) For example, by keeping track of the total TSS per week or per month, it may be possible to identify an individual's "breaking point", i.e., the maximum quantity and quality of training that still leads to improvements, rather than overtraining. As well, a very high TSS resulting from a single race or training session may be an indicator that additional recovery on subsequent days is required. Until additional experience is gained with the method, it is difficult to say exactly what a "high TSS score" exactly is – however, the table below gives some rough guidelines:

| | |
|---------|---|
| <100 | low (easy to recover by following day) |
| 100-200 | medium (some residual fatigue may be present the next day, but gone by 2nd day) |
| 200-300 | high (some residual fatigue may be present even after 2 days) |
| 300 | epic (residual fatigue lasting several days likely) |

Note that while the TSS score is normalized to an individual's LT, such that comparison across individuals is possible, there could still be differences between athletes in how they respond to a given "dose" of training. Such difference may be due to natural ability, or may be the result of specific training (i.e., the more you do the more you can do). This isn't really a problem, however, since comparison within a given individual is the primary interest.

While the goal at the outset was to develop a method of quantifying the overall training load (duration x intensity) via TSS, the IF score may actually prove to be even more useful. For example, it can be used to compare the intensity of even markedly dissimilar training sessions or races, either within (most valid/relevant) or across (to assess tactical or drafting skill, or just for plain old "bragging rights" <g>) individuals (see below):

Typical IF values for different events or training sessions:

| | |
|-----------|--|
| <0.75 | level 1 recovery rides |
| 0.75-0.85 | level 2 endurance training sessions |
| 0.85-0.95 | level 3 tempo rides, aerobic and anaerobic interval workouts |

(work and rest periods combined), longer (>2.5 h) road races
 0.95-1.05 level 4 intervals, shorter (<2.5 h) road races, criteriums,
 circuit races, 40k TT (by definition)
 1.05-1.15 shorter (e.g., 15 km) TTs, track points race
 1.15 prologue TT, track pursuit, track miss-and-out

Perhaps even more importantly, *for the first time ever* the algorithm used to derive IF makes it possible to estimate steady-state power at LT from highly variable power data! That is, if sustainable power (either constant or non-constant) is essentially “capped” by power at LT, and if the 30 second smoothing/4th order weighting algorithm appropriately corrects the variable power data, then the power estimated at step 4 in the calculation of TSS/IF (see above) provides an estimate of the equivalent steady power that could be produced for the same physiological stress.* Stated another way, the correction algorithm simply provides a means of expressing highly variable power data in physiologically-relevant “language”. Consequently, if an individual pushes themselves just as hard in a ~1 hour mass start race (or time trial in very hilly terrain) as they might in a flat time trial, then corrected power provides an estimate (generally to w/in 5-10 W) of their power at LT. This observation reduces, or perhaps even completely eliminates, the need to perform a time trial to determine power at LT. Instead, the results of mass start races can be used for this purpose, for example for beginning power meter users who have never done a time trial using such a tool. Even for riders whose power at LT is well established, the IF score can be used to detect significant changes in fitness – for example, if a rider’s IF score for a ~1 h race is greater than 1.05, then their LT power should be reassessed (ideally using the same means used to establish it originally) to determine whether it has truly changed.

*Astute readers will have already picked up on the fact that the IF values given in the table above are the fraction or percentage of power at LT that was equivalently maintained. Indeed, it was suggested to me that the IF should be multiplied by 100 to express it as a percentage, since decimal values less than 1 can be more difficult to immediately grasp. I resisted this quite valid suggestion, however, because I was afraid that scaling IF this way might result in people confusing IF values with TSS scores. As well, expressing IF as a percentage rather than a decimal could result in individuals confusing these values with the percentages limits of the training levels I laid out previously. A really astute reader will realize that they are in fact essentially measures of the same thing, i.e., power output relative to the individual’s power at LT – the absolute values differ, however, because deriving the IF score corrects for the effects of variations in power on physiological responses, whereas the training levels have simply been offset to lower power levels to account for this fact (e.g., level 1, recovery, is defined as an average power of <55% of power at LT, but the IF value of <0.75 corresponds to <75% of power at LT).

Finally, yet another application of the IF algorithm/score is as a teaching tool, as it helps demonstrate why, even when power is highly variable, it is still an individual’s “metabolic fitness” (i.e., power at LT) that is important in determining performance. That is, by illustrating (via a 4th order relationship – greater even than the 3rd order relationship between power and wind resistance!) how physiologically “costly” every sustained burst above LT proves to be, the IF algorithm may 1) help less experienced riders understand why it is important to learn how to modulate their effort during mass start races, so that they don’t fatigue themselves unnecessarily, and 2) help even experienced riders understand how appropriate training aimed at raising LT can improve performance even in events seemingly much different than a time trial (a point Amit has already picked up on).

Limitations and concluding remarks:

As mentioned previously, the key to everything I’ve written about above is the weighting algorithm, and thus the validity/robustness of the TSS and IF scores/values depend entirely on it. I believe that it is based on sound physiological reasoning, and in my experience so far it seems to work quite well (better than I could have hoped, actually). I have

not, however, had the chance to evaluate thousands, much less hundreds, of data files, so the possibility of the occasional “outlier” still exists. A greater limitation to the entire concept, though, is that the basic premise – i.e., that you can adequately describe the training load and the stress it imposes on an individual based on just one number (TSS), completely ignoring how that “score” is achieved and other factors (e.g., diet, rest) – is, on its face, ridiculous. In particular, it must be recognized that just because, e.g., two different training programs produce the same weekly TSS total, doesn’t mean that an individual will respond in exactly the same way. Nonetheless, I believe that TSS (and IF) should prove useful to coaches and athletes for evaluating/managing training.

Finally, I am releasing this idea onto the list because I strongly believe that knowledge is to be shared, not hoarded, and I hope that others will benefit from my efforts. To that end, I encourage people to try calculating TSS and IF for some of their own files, and share any interesting observations or questions that arise as a result – somebody might even want to try writing an Excel macro to speed up the calculations a bit. However, I would be very disappointed if anyone tried to capitalize on these ideas by producing or incorporating them into a commercial program without my permission.

[Join This List](#)[List Info](#)[◀ Previous Message](#)[All Messages](#)[Next Message ▶](#)

© 2001 Topica Inc. TFMB

Concerned about privacy? Topica is TrustE certified.
See our [Privacy Policy](#).

