In-the-field drivetrain efficiency measurements using commercially available bike sensors: error analysis and recommendations

Bart Blockmans[1]

[1] Classified Cycling, Damplein 23, 2060 Antwerp, Belgium

Abstract. A popular approach to measure the relative efficiency of bicycle drivetrains in the field is the energy expenditure method, whereby in consecutive trials the average power and time required to cover a fixed distance are recorded and multiplied together to obtain the total mechanical energy expended. While the same experiment is typically repeated multiple times so as to reduce the standard error of the measurement, a careful error analysis is usually omitted, leaving the accuracy of the measurement undetermined. In this article a detailed error analysis of this measurement approach is presented, including a discussion of the accuracy of the involved measurement devices and of the influence of environmental conditions and pacing strategy. Practical measurement guidelines on how to achieve the required accuracy to draw valid conclusions regarding bicycle drivetrain efficiency are provided, as well as easy-to-use formulae to quantify the measurement error.

1. Introduction

With marginal gains winning the biggest cycling races and watts being the cyclist's primary valuta, cycling efficiency may have never been as highly pursued as today. Countless hours of wind tunnel testing are used to scrape off the last fraction of a watt from a cyclist's jersey, while sponsorship deals are put on the line to save 100 grams on a lightweight wheel set. Besides aerodynamic and gravitational efficiency, bicycle drivetrain efficiency has received increased interest in recent years. On the verge of what could possibly be a paradigm shift in bicycle drivetrain technology, the traditional front- and rear-derailleur-based *2x* systems are being challenged by "derailleur-killing" technologies such as *1x* configurations and drivetrains equipped with hub gears or mid-drive gearboxes. Whether or not said technology shift will actually happen, and if so in which of cycling's many disciplines, depends to a large extend on the proven efficiency of these new drivetrain concepts. The same holds true for less disruptive technologies such as oversized pulley wheels, ceramic bearings, low-friction chain wheels, and advanced lubricants.

Drivetrain efficiency is typically measured using one of two approaches: either in a controlled lab environment using dedicated test rig infrastructure, or in the field during an experiment that maximally represents actual cycling. Lab-based tests are primarily performed by drivetrain component manufacturers and are primarily concerned with the developed technology itself, with little to no reference to the systems interacting with it (see e.g. [1]). Field tests on the other hand consider the drivetrain as part of the bike-rider-world system and assess its efficiency under real-life conditions. These kinds of tests are particularly popular amongst technology journalists and enthusiasts, who rightfully feel the proof of the pudding is in the eating and typically have no access to dedicated test rig infrastructure. Equipped with little more than a bicycle power meter and a stop watch or GPS-based bicycle computer, they can – in theory – measure and compare the mechanical energies required to move from points A to B for different drivetrain configurations and in doing so establish a drivetrain efficiency hierarchy to share with their audiences.

The measurement technique just described is known as the energy expenditure approach, and involves recording the average power and time required to cover a fixed distance. By multiplying these quantities, the total expended mechanical energy is obtained, which will be higher for less efficient drivetrains and vice versa. As the comparatively large measurement error of this approach is intuitively clear to anybody performing this test, the same experiment is usually repeated multiple times so as to reduce the standard error of the measurement. Despite all good intentions, a detailed error analysis of the measurement is typically omitted, leaving the accuracy of the measurement – and hence the validity of the conclusions drawn from it – to be guessed at.

In this article a detailed error analysis of the energy expenditure method is presented. In Section 2 the mechanical power required to move a bike + rider at a given ground velocity is broken down into its main constituents, and a brief parameter sensitivity analysis is included to set the scene for the next section. In Section 3, an error analysis of the energy expenditure method is conducted based on the basic laws of error propagation. Section 4 investigates the influence of the selected pacing strategy on the outcome of the experiment, and Section 5 discusses the important topic of bicycle power meter accuracy (or rather: precision). Finally, Section 6 proposes an error-proof test protocol and summarizes the article's main findings.

2. Breakdown of mechanical power in cycling

The mechanical power P required the overcome the forces $F_{\leftarrow}$ that oppose the translational motion of the bike + rider moving at a ground speed of $V_{\rightarrow}$, can be decomposed as follows (see e.g. [2]):

$$\begin{aligned} P = F_{\leftarrow}V_{\rightarrow} = 1/2\,\rho C_d A(V_{\rightarrow} + w_{\leftarrow})^2 V_{\rightarrow} & \quad \text{air resistance} \\ + mg\sin\theta\, V_{\rightarrow} & \quad \text{change in elevation} \\ \pm maV_{\rightarrow} & \quad \text{change in speed} \\ + C_{rr}mg\cos\theta\, V_{\rightarrow} & \quad \text{rolling resistance} \\ + P_{loss,DT}(V_{\rightarrow}, P) & \quad \text{drivetrain losses} \end{aligned} \tag{1}$$

where the following symbols were used:

- ρ Air density [kg/m³]
- C_d Air-resistance coefficient [-]
- A Combined frontal surface area of the bicycle and cyclist [m²]
- $w_{\leftarrow}$ Wind component opposing the motion [m/s]
- m Combined mass of the bicycle and cyclist [kg]
- g Acceleration due to gravity [m/s²]
- θ Slope of the incline [rad]
- a Temporal rate of change of the ground speed $V_{\rightarrow}$ [m/s²]
- C_{rr} Rolling-resistance coefficient [-]
- $P_{loss,DT}$ Power losses in the transmission components of the bicycle drivetrain [W]

Note that given the focus of the current presentation, the power losses in the bicycle drivetrain, $P_{loss,DT}$, are explicitly included in Eq. 1, as opposed to through a drivetrain efficiency η as is common in literature (see e.g. [2]). The drivetrain power losses consist of the following three components:

$$P_{loss,DT} = P_{loss,H} + P_{loss,C} + P_{loss,B} \tag{2}$$

where the following symbols were used:

- $P_{loss,H}$ (hub) losses occurring between the rear-wheel sprocket or pulley, and the wheel [W]
- $P_{loss,C}$ (chain) losses occurring in the chain-sprocket and chain-ringwheel interfaces [W]
- $P_{loss,B}$ (bracket) losses occurring in the bicycle's bottom bracket [W]

In traditional 2x drivetrains, hub and bottom bracket losses are largely composed of bearing (seal) losses [3], and are relatively small compared to chain losses. The latter are due to frictional contact interactions between i) the bushings of the chain and the chain links, ii) the rollers of the chain and the sprocket teeth, and iii) the inner and outer side plates of the chain in the case of chain line offset [4]. In bicycles equipped with hub gears or mid-drive gearboxes, chain losses can be reduced at the expense of higher hub losses (in geared hubs) or bottom bracket losses (in mid-drive gearboxes) caused by frictional losses in the gear contacts (see [1] for a detailed analysis of geared bicycle hub losses). As a detailed analysis of mathematical drivetrain loss models is out of scope here, a first-order approximation of the models of [1], [3] and [4] is presented, exploiting the commonly applied decomposition of drivetrain power losses in terms of load-dependent and load-independent (or speed-dependent) losses:

$$P_{loss,DT} = C_P P + C_V V_{\rightarrow} \tag{3}$$

where the following symbols were used:

- C_P Load-dependent drivetrain loss coefficient [%]
- C_V Speed-dependent drivetrain loss coefficient [N]

Note that the load-dependent loss coefficients are percentages, whereas the speed-dependent loss coefficients have units of force [N]. Substitution of Eq. 3 into Eq. 1 yields the following expression:

$$\begin{aligned} P = F_{\leftarrow} V_{\rightarrow} &= \left[\frac{1}{2}\rho C_d A (V_{\rightarrow} + w_{\leftarrow})^2 + mg\sin\theta + ma + C_{rr} mg\cos\theta + C_V\right] V_{\rightarrow} + C_P P \\ &= \left[\frac{1}{2}\rho C_d A (V_{\rightarrow} + w_{\leftarrow})^2 + mg\sin\theta + ma + C_{rr} mg\cos\theta + C_V\right] \frac{V_{\rightarrow}}{1 - C_P} \end{aligned} \tag{4}$$

where the load-dependent drivetrain loss coefficient C_P takes the place of the overall drivetrain efficiency η in traditional power breakdown analysis. Under steady-state conditions (i.e. $V_{\rightarrow}$ is constant and $a = 0$), this equation can be used in two ways: either it returns the ground velocity $V_{\rightarrow}$ for a given input power P and the parameters of the cyclist $(m, C_d A)$, the bicycle (C_{rr}, C_P, C_V), and the environment $(\rho, w_{\leftarrow}, g, \theta)$, which requires solving a quadratic equation. Or it returns the required input power P for sustaining a given ground velocity $V_{\rightarrow}$, which can directly be obtained from of Eq. 4. Under unsteady conditions, the resulting ground velocity or required input power are time-dependent and require the piecewise application of the above equation along a discretized trajectory. In order to give the reader a feel for the relative contribution of the resistive forces, Figure 1 shows the mechanical power breakdown during cycling on a *false-flat* ($\theta = 0.5\%$) and an uphill ($\theta = 8\%$) section at 400 watts of input power, for the parameters given in Table 1. 250 seconds of cycling at these powers requires 100 kJ or approximately 24 kcal of mechanical energy; during that time the cyclist will cover 2.57 km on the false flat and 1.29 km on the uphill section, at average ground velocities of approximately 37 km/h and 18.5 km/h, respectively.

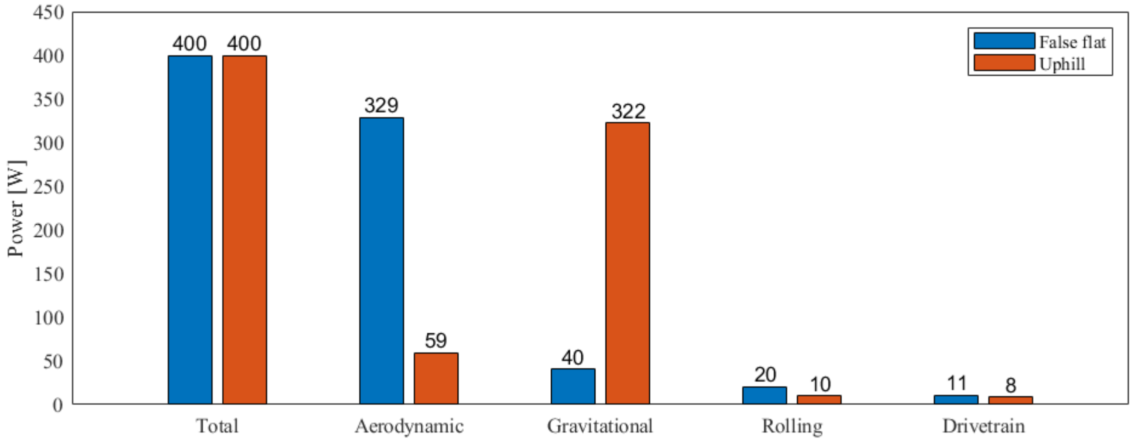


Figure 1. Breakdown of mechanical input power during false-flat ($\theta = 0.5\%$) and uphill ($\theta = 8\%$) cycling at 400 watts.

Parameter	Value	Parameter	Value
ρ	1.225 kg/m³	g	9.81 m/s²
$C_d A$	0.321 m²	C_{rr}	0.0025
$w_{\leftarrow}$	2 m/s	C_V	0.455 N
m	80 kg	C_P	0.015

Table 1. Reference parameters for mechanical power computations used in this article.

In view of the upcoming error analysis, it is interesting to have a look at the sensitivity of the mechanical power to parameters that are subject to change while assessing the efficiency of a given drivetrain configuration. As bicycle drivetrains tend to be very efficient, with efficiencies as high as $\eta = 98\%$, small changes in said parameters can have a significant influence on the outcome of the experiment. Assuming drivetrain efficiency degradation is avoided through proper lubrication of the drivetrain components in between trials, the bicycle parameters (C_{rr}, C_P, C_V) can be considered as constants for a given drivetrain technology. The same holds for the total mass m of the bicycle + cyclist, assuming the bike setup and rider (incl. kit) don't change inbetween experiments, and the test trajectory doesn't pass through a litter zone. Furthermore, assuming the location of the experiments is the same between all trials, the gravitational acceleration g is constant. The same cannot be said of the air density ρ due to its dependency on temperature T and the possibility of different trials being conducted at widely separated times. Finally, as gravity is a conservative force and the slope angle θ has a negligable influence on the rolling resistance ($\cos\theta \approx 1$), microscopic changes in θ due to minuscule trajectory deviations can be ignored. The mathematical expression of a sensitivity of a function f with respect to a variable x is given by the partial derivative $\partial f/\partial x$. Based on Eq. 4, the following sensitivities are obtained for the air resistance coefficient C_dA and wind speed $w_{\leftarrow}$:

$$\frac{\partial P}{\partial C_d A} = \frac{1}{2}\rho(V_{\rightarrow} + w_{\leftarrow})^2 V_{\rightarrow} \tag{5}$$

$$\frac{\partial P}{\partial w_{\leftarrow}} = \rho C_d A(V_{\rightarrow} + w_{\leftarrow})V_{\rightarrow} \tag{6}$$

Note that the factor $(1 - C_P)^{-1}$ has been omitted from these expressions as the small value of $C_P \ll 1$ has only a minor influence on the sensitivities. For the sensitivity of the mechanical power with respect to the air density ρ, the ideal gas law is used:

$$\rho = \frac{p}{RT} \tag{7}$$

where p is the absolute pressure [Pa] and R is the specific gas constant for air [J/kgK]. When the air is humid, the ideal gas law should in theory be applied to a mixture of dry and humid air, where the mixture's composition depends on the relative humidity ϕ of the air. However, as the humidity of the air only has a considerable influence at high temperatures ($T > 30°$), it is not considered here; hence the specific gas constant for dry air ($R = R_0 = 287.058$ J/kgK) is used. Furthermore, assuming all trials are conducted at similar altitude and elevation gains of the selected trajectory are limited, changes in air pressure are neglected and the standard atmospheric pressure at sea level ($p = p_0 = 101.325$ kPa) is used. In doing so, the air density becomes a sole function of temperature. Using the chain rule of differentiation, the sensitivity of the power P with regard to temperature is obtained as

$$\frac{\partial P}{\partial T} = \frac{\partial P}{\partial \rho}\frac{\partial \rho}{\partial T} = -\frac{p_0}{R_0 T^2}\frac{\partial P}{\partial \rho} = -\frac{p_0}{2R_0 T^2}C_d A(V_{\rightarrow} + w_{\leftarrow})^2 V_{\rightarrow} \tag{8}$$

Finally, it is interesting to see how changes in mass and power loss coefficients across different drivetrain configurations influence the total mechanical power:

$$\frac{\partial P}{\partial m} = [g\sin\theta + a + C_{rr} g\cos\theta]V_{\rightarrow} \tag{9}$$

$$\frac{\partial P}{\partial C_V} = V_{\rightarrow} \tag{10}$$

$$\frac{\partial P}{\partial C_P} = \frac{P}{(1 - C_P)} \tag{11}$$

In Table 2 the six sensitivities derived above are expressed in terms of the numerical values in Table 1. The operating conditions of in-the-field efficiency tests are selected so as to eliminate the impact of certain sensitivities with high amplitude yet low controllability. For example, efficiency tests on an indoor cycling track are selected so as to minimize the influence of environmental conditions. Outdoor efficiency tests are typically performed on inclined trajectories so as to minimize the influence of changes in the air resistance coefficient C_dA. The ideal trajectory – an uphill indoor cycling track – unfortunately doesn't exist, so there will always be some influencing factors at work. In the next section the influence of these factors on the energy expenditure measurement approach will be investigated.

Sensitivity of P with respect to …	**Units**	**False flat**	**Uphill**
Air resistance coefficient C_dA	W/m²	1025	184.6
Wind speed $w_{\leftarrow}$	W/m/s	51.5	15.5
Temperature T	W/°C	-1.1	-0.2
Total mass bicycle + cyclist m	W/kg	0.75	4.15
Speed-dependent drivetrain loss C_V	W/N	10.26	5.15
Load-dependent drivetrain loss C_P	W/%	4.06	4.06

Table 2. Sensitivity of cycling power with respect to key influencing factors

3. Error analysis of the energy expenditure measurement approach

Energy expenditure is the sum of the basal metabolic rate (the energy expended while at complete rest), the thermic effect of food (the energy required to digest and absorb food), and the energy expended in physical activity. While physical activity amounts to no more than 20% of the total energy expended, the energy expenditure measurement is only concerned with the latter. Before discarding the remaining four fifths of the expended energy for the remainder of this article, it should be noted that bicycle drivetrain configurations – through the gear range and steps they offer – can have a significant effect on the cyclist's metabolic efficiency, and can as such have a bigger impact on the overall efficiency of the bicycle-cyclist system than would be apparent based on mechanical drivetrain efficiency alone.

Energy is the cyclist's ability to do work. Power is the rate at which work is done. In cycling, the instantaneous power applied to the pedals is measured using a power meter mounted in either the pedals, crank(s) or rear wheel of the bicycle. Most power meters function on the measurement of the local elastic deformation of a component under the influence of the applied pedaling force. This deformation is then converted to a torque M based on the power meter's calibration data and multiplied by the angular velocity ω of the crank or wheel to obtain the mechanical power $P = M\omega$. The average power $\bar{P}$ and the total mechanical energy E (expressed in Joules) required to cover a fixed distance between points A and B are related through the elapsed time $\Delta t = t_B - t_A$ as follows:

$$E = \int_{t_A}^{t_B} P(t)dt = \bar{P}\Delta t \tag{12}$$

This simple equation forms the basis of the energy expenditure measurement. In practice, $\bar{P}$ is obtained by averaging a time series of instantaneous power measurements that are each as close as possible to a pre-determined target power P_{target}. As long as the test trajectory stays the same across all experiments, small deviations with respect to P_{target} are thought to not substantially impact the outcome of the test. The elapsed time, Δt, is recorded using a timing device that records the times at which the cyclist crosses the trajectory's start and finish lines. By repeating these measurements for different drivetrain configurations, *ceteris paribus*, the relative efficiency of one drivetrain configuration with respect to another can be established.

Unfortunately, in general ceteris *non* paribus. As if it's not hard enough already to maintain constant power throughout repeated VO2max efforts, the test subject should make sure to keep his/her aerodynamic position fixed and to not deviate substantially from the trajectory that was mapped out for him/her. While these factors are difficult to control already, it's even harder to get a grip on environmental conditions such as changes in wind speed and/or direction, and temperature changes. While it might seem good practice to redo experiments with deviating operating conditions, a high number of valid experiments is critical to minimize the standard measurement error (see further) and there is a limit to what our poor test subject can endure. Another option is then to scale the respective energy measurements with regard to a chosen reference point based on the actual operating conditions during the respective trials. For example, if in one experiment the distance covered by the bicycle wheels, d, is one percent higher than the nominal trajectory distance d_0 (due to the test subject's inability to hold a line), this will amount to an energy surplus ΔE_d of approximately

$$\Delta E_d = \bar{P}\frac{\Delta d}{\bar{V}_{\rightarrow}} = \bar{P}\frac{0.01 d_0}{\bar{V}_{\rightarrow}} \tag{13}$$

For the experiments conducted in the previous section, this surplus amounts to 1 kJ, which represents 1% of the total energy expenditure and is thus of the same order as the amount of energy expended to overcome resistive drivetrain forces. To enable a fair comparison between different experiments, this energy surplus should hence be subtracted from the total energy measurement, to obtained the corrected energy:

$$E_{corr,d} = E - \Delta E_d = E\left(1 - \frac{\Delta d}{\bar{V}_{\rightarrow} t_{AB}}\right) = E\left(1 - \frac{\Delta d}{d}\right) = E\left(\frac{d_0}{d}\right) \tag{14}$$

Similar corrections should be applied to the other influencing factors if the outcomes of the various trials are to be compared with each other. The energy surplus due to a change in wind speed $\Delta w_{\leftarrow}$ equals

$$\Delta E_w = \frac{\partial E}{\partial w_{\leftarrow}}\Delta w_{\leftarrow} = \frac{\partial P}{\partial w_{\leftarrow}}\Delta w_{\leftarrow}\Delta t = \rho C_d A (V_{\rightarrow} + w_{\leftarrow}) V_{\rightarrow} \Delta w_{\leftarrow} \Delta t \tag{15}$$

For the experiments of the previous section, a wind speed change of $\Delta w_{\leftarrow} = 0.5$ m/s results in an energy surplus of more than 6 kJ (= 6%) for false-flat cycling and approximately 2 kJ (= 2%) for uphill cycling. Subtraction of this energy surplus from the total energy measurement yields:

$$E_{corr,w} = E - \Delta E_w = E\left(1 - \phi_{aero}\frac{2\Delta w_{\leftarrow}}{V_{\rightarrow} + w_{\leftarrow}}\right) \tag{16}$$

where ϕ_{aero} is the ratio of energy expended to overcome aerodynamic drag to the total energy expenditure. For the experiments conducted in the previous section, $\phi_{aero} \approx 0.8$ for false-flat cycling and $\phi_{aero} \approx 0.15$ for uphill cycling. Similarly, the energy deficit due to a temperature change is approximately equal to:

$$\Delta E_T = \frac{\partial E}{\partial T}\Delta T = \frac{\partial P}{\partial T}\Delta T \Delta t = -\frac{1}{2}\rho_0 \frac{\Delta T}{T} C_d A (V_{\rightarrow} + w_{\leftarrow})^2 V_{\rightarrow} \Delta t \tag{17}$$

which for a temperature change of $\Delta T = 2K$ amounts to energy deficits of 0.5 kJ (= 0.5%) and 0.1 kJ (= 0.1%) for false-flat and uphill cycling, respectively. Correction of the total energy measurement for the recorded change in temperature yields:

$$E_{corr,T} = E - \Delta E_T = E\left(1 + \phi_{aero}\frac{\Delta T}{T}\right) \tag{18}$$

Finally, a positive change in the air resistance coefficient $C_d A$ due to the test subject's inability to maintain a constant aerodynamic position throughout all experiments results in an energy surplus of:

$$\Delta E_{CdA} = \frac{\partial E}{\partial C_d A}\Delta C_d A = \frac{\partial P}{\partial C_d A}\Delta C_d A \Delta t = \frac{1}{2}\rho (V_{\rightarrow} + w_{\leftarrow})^2 V_{\rightarrow} \Delta C_d A \Delta t \tag{19}$$

For a 10% increase in air resistance coefficient, the energy surplus equals approximately 8 kJ (= 8%) in case of false-flat cycling and 1.5 kJ (= 1.5%) in case of uphill cycling. The corresponding corrected energy is:

$$E_{corr,CdA} = E - \Delta E_{CdA} = E(1 - \phi_{aero}\Delta C_d A) \tag{20}$$

Putting everything together, the corrected mechanical energy expended during an experiment subject to changes in distance travelled by the bicycle wheels (Δd), wind speed ($\Delta w_{\leftarrow}$), temperature (ΔT) and aerodynamic position ($\Delta C_d A$) is equal to:

$$E_{corr} = \bar{P}\Delta t\left(1 - \frac{\Delta d}{d}\right)\left[1 - \phi_{aero}\left(\frac{2\Delta w_{\leftarrow}}{V_{\rightarrow} + w_{\leftarrow}} + \Delta C_d A - \frac{\Delta T}{T}\right)\right] = \bar{P}\Delta t\left(\frac{d_0}{d}\right)[1 - \varphi_{aero}] \tag{21}$$

which brings us far from the simple energy equation of Eq. 12. Usage of this equation requires measuring the distance travelled by the bicycle wheels, the wind speed, the air resistance coefficient, and the temperature, which suddenly makes the simple energy expenditure measurement quite a complicated affair. If such measurements are not available, the above equation can be used to estimate (part of) the measurement error. For example, if our (wishfully thinking) experimenter believes the error in following the designed trajectory is no more than 0.25%, the wind speed changes with no more than 2%, the temperature changes with no more than 0.5 K, and the prescribed aerodynamic positions declines with no more than 2%, the energy measurement – assuming perfect average power and time registration (we'll get to that soon) – will have a maximal possible error of ±2.7 kJ (= 2.7%) in the case of false-flat cycling and ±0.76 kJ (= 0.76%) in the case of uphill cycling. When these changes are assumed to be random rather than systematic and follow a normal distribution, the error in the averaged energy measurement can be reduced by repeating the same experiment n times, which will reduce the standard error by a factor of approximately $1/\sqrt{n}$. This means that in order to halve the error on the estimated mean energy expenditure, four times as many experiments will be required; reducing the error by 75% requires 16 times as many experiments. Clearly a nearly flat test trajectory is a test subject killer.

Thus far all measurements were considered to be flawless. Obviously, this is not the case in reality, and the above analysis needs to be complemented with a measurement error analysis. To make life a bit easier, let's assume no measurement error is involved in the φ_{aero} term in Eq. 21. Then, the corrected energy E_{corr} is a function of the average power $\bar{P}$, the elapsed time Δt, and the travelled distance d. While the former two are independent variables, there is a correlation between Δt and d which will be neglected in the upcoming analysis. According to the law of error propagation (see e.g. [3]), the error in the measurement of the energy ϵ_E can be written in terms of the power, time and distance measurement errors ϵ_P, ϵ_t, and ϵ_d as follows:

$$\epsilon_E = \sqrt{\left(\frac{\partial E}{\partial \bar{P}}\right)^2 \epsilon_P^2 + \left(\frac{\partial E}{\partial \Delta t}\right)^2 \epsilon_t^2 + \left(\frac{\partial E}{\partial d}\right)^2 \epsilon_d^2} \tag{22}$$

where the subscript *'corr'* is omitted for brevity. Substitution of Eq. 21 (with omission of the factor φ_{aero}) into the above expression yields

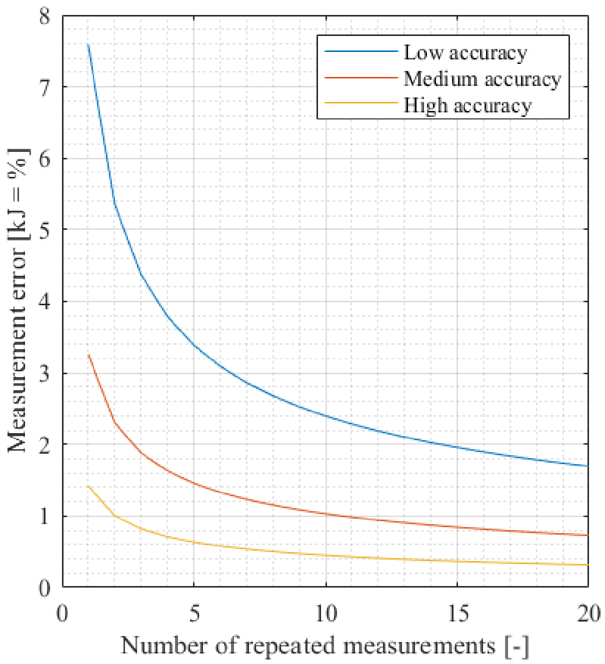
$$\epsilon_E = \sqrt{\left[\Delta t\left(\frac{d_0}{d}\right)\right]^2 \epsilon_P^2 + \left[\bar{P}\left(\frac{d_0}{d}\right)\right]^2 \epsilon_t^2 + \left[-\bar{P}\Delta t\left(\frac{d_0}{d^2}\right)\right]^2 \epsilon_d^2} \approx \sqrt{\Delta t^2 \epsilon_P^2 + \bar{P}^2 \epsilon_t^2 + \left[-\frac{\bar{P}\Delta t}{d}\right]^2 \epsilon_d^2} \tag{23}$$

where use was made of the assumption that $d \approx d_0$ and thus only marginally influences the measurement error. Say the experiment is conducted with a commercial power meter that has a claimed accuracy of ±1%, time is registered by the experimenter using a stop watch with an error of ±0.5 s (i.e. twice the average human reaction time to visual stimuli, as the stop watch needs to be pressed at both start and end of the trajectory), and the distance travelled is measured by an encoder mounted on the front wheel axle with an accuracy of one quarter of the circumference of the wheel (i.e. approximately ±0.5 m), then for the 400W experiments conducted in the previous section, the measurement error would be approximately equal to ± 1 kJ (= 1%). Slightly higher errors are obtained in case of GPS-based measurements, where ϵ_d can easily reach ±5 m, resulting in a timing error ϵ_t of ±0.5-1.0 s based on the cyclist's ground velocity. Assuming again these errors are random rather than systematic (which, again, is far from obvious), the standard measurement error can (and will have to) be reduced by repeated measurements, thereby reducing the standard error by a factor of approximately $1/\sqrt{n}$.

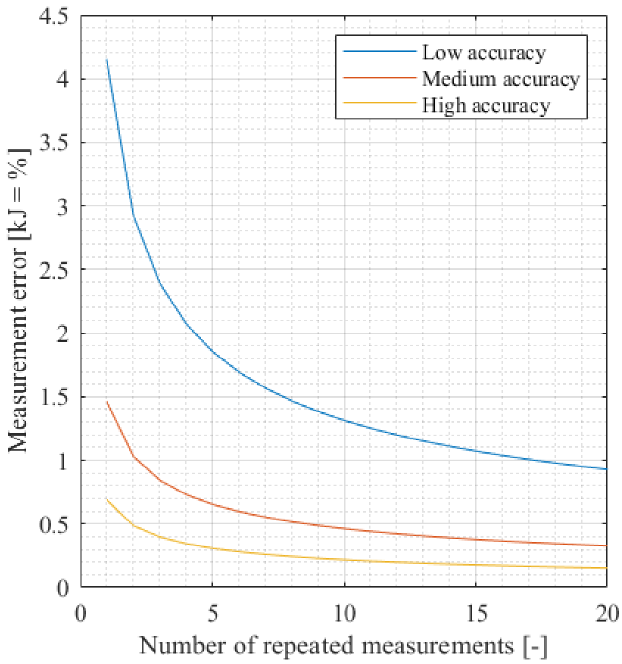
Summarizing our findings thus far, the energy expenditure measurement requires correcting for the actual distance travelled by the wheels (see Eq. 14), and involves two kinds of errors: a measurement error, ϵ_E, due to inaccuracies in the measurement devices (see Eq. 23) and an "experiment error", Δ_E, due to difficult-to-control experiment parameters (see Eq. 21). The overall error of the energy expenditure measurement, $\mathcal{E}_E$, can thus be expressed as:

$$\mathcal{E}_E = \frac{1}{\sqrt{n}}(\epsilon_E + \Delta_E) = \frac{1}{\sqrt{n}}\left[\sqrt{\Delta t^2 \epsilon_P^2 + \bar{P}^2 \epsilon_t^2 + \left(\frac{\bar{P}\Delta t}{d}\right)^2 \epsilon_d^2} + E\phi_{aero}\left(\frac{2\Delta w_{\leftarrow}}{V_{\rightarrow} + w_{\leftarrow}} + \Delta C_d A - \frac{\Delta T}{T}\right)\right] \quad (24)$$

In Figure 2, the overall error is shown in function of the number of repeated experiments n for the two experiments studied in the previous sections and for 3 levels of measurement accuracy, the data of which are shown in Table 3. In this table, decreased instrument measurement errors are typically associated with increased resources, whereas decreased experiment deviations imply either improved control of the experimental conditions, or actual measurement of parameter deviations and application of the correction formulas given in Eq. 21. From Figure 2 it is immediately clear that uphill trajectories have a clear advantage in terms of measurement error due to the small contribution of aerodynamic drag (see also Figure 1). In order to obtain a maximal measurement error of 0.5% (which is not an unnecessary luxury when studying bicycle drivetrain efficiency), the low accuracy measurement approaches are ruled out a priori, the medium accuracy approach needs approximately 9 measurements per study case, and the highly accurate measurement can suffice with 2 repetitions. In Section 5 we'll have a look into how realistic such a high measurement accuracy really is given the currently available bicycle power meters. First we'll have a look at the influence of power pacing strategy.



(a)



(b)

Figure 2. Total measurement error in function of number of repeated 100 kJ experiments for false-flat (a) and uphill (b) trajectories.

Measurement accuracy	ϵ_P	ϵ_t	ϵ_d	Δd	$\Delta w_{\leftarrow}$	$\Delta C_d A$	ΔT
Low (excl. distance correction)	2%	1 s	n.a.	1%	2%	5%	1 K
Medium (incl. distance correction)	1%	0.5 s	1 m	0.5%	1%	2.5%	0.5 K
High (incl. distance correction)	0.5%	0.2 s	0.25 m	0.25%	0.5%	1%	0.25 K

Table 3. Measurement errors and deviations involved in three levels of measurement accuracy (see Figure 2)

4. Influence of pacing strategy

Thus far we have only been concerned with the average power $\bar{P}$ applied to the pedals along the test trajectory, assuming thus that the energy expenditure is not influenced by the distribution of power $P(t)$ along the trajectory. And indeed, when the environmental parameters (like inclination or headwind) don't change along the trajectory, this is a reasonable assumption. Say, for example, that part of the trajectory was covered at an average power of P_1 and another part at an average power of P_2, where $P_1 \neq P_2$. The time-averaged power is then equal to

$$\bar{P} = \frac{P_1 \Delta t_1 + P_2 \Delta t_2}{\Delta t_1 + \Delta t_2} = \frac{P_1(d_1/V_1) + P_2(d_2/V_2)}{(d_1/V_1) + (d_2/V_2)} = \frac{P_1(\alpha d/V_1) + P_2((1-\alpha)d/V_2)}{(\alpha d/V_1) + ((1-\alpha)d/V_2)} \tag{25}$$

where α is the fraction of the total trajectory distance d covered at the average power P_1 such that the total average power equals $\bar{P}$. Solving this equation for α yields:

$$\alpha = \frac{P_2 V_1 - \bar{P} V_1}{\bar{P}(V_2 - V_1) - P_1 V_2 + P_2 V_1} \tag{26}$$

In Table 4 the differences in total energy expenditure are compared for the false-flat and uphill test trajectories, where two different pacing strategies have been applied: one where the average power is maintained throughout the whole course, resulting in the total energy expenditure $\bar{E}$, and a second where two power levels – the first 10% lower and the second 10% higher than the total average power – are maintained throughout different parts of the course, resulting in a total energy expenditure E_{12}. While the average powers over the entire trajectory are the same in both cases, the time required to cover the trajectory increases for the second pacing strategy, resulting in a small surplus in energy expenditure. However, given the large power variation (10%) and the small energy surplus (less than 0.2%), one may conclude that pacing strategy has no important influence on the energy expenditure measurement method in the case of a constant trajectory.

Trajectory	$\bar{P}$	P_1	P_2	α	$\bar{E}$	E_{12}	ΔE	$\Delta(\Delta t)$
False flat (2.57 km)	400 W	440 W	360 W	0.5214	100.00 kJ	100.15 kJ	147 J	0.3680 s
Uphill (1.29 km)	400 W	440 W	360 W	0.5417	100.00 kJ	100.16 kJ	156 J	0.3897 s

Table 4. Influence of pacing strategy on a constant trajectory

Things are a bit different when the environmental parameters of the trajectory change along the trajectory's course. For example, consider the case where the false-flat trajectory has a headwind ($w_{\leftarrow}$ = 5 m/s) and a tailwind ($w_{\leftarrow}$ = -5 m/s) section, each 1397.5 m long. With the remaining parameters as in Table 1, completing the course using a constant-power strategy ($P(t) = \bar{P} = 400$ W) amounts to a total energy expenditure of 100 kJ. Old cycling wisdom however claims that in order to complete such a course in the smallest amount of time, it is better to push harder during the headwind section. If that is true, there should be a pacing strategy with equal overall average power $\bar{P}$ but smaller elapsed time Δt. Finding this optimal pacing strategy and the corresponding minimum of Δt requires maximizing the average ground speed:

$$\bar{V}_{\rightarrow} = \frac{d}{\Delta t} = \frac{d}{\Delta t_{hw} + \Delta t_{tw}} = \frac{d}{d/2V_{hw} + d/2V_{tw}} = \frac{2V_{hw}V_{tw}}{V_{hw} + V_{hw}} \tag{27}$$

where the subscripts 'hw' and 'tw' refer to headwind and tailwind quantities, respectively. Obviously the resulting time-averaged power should remain equal to $\bar{P}$, which can be expressed using the following constraint equation (cfr. Eq. 25):

$$C(V_{hw}, V_{tw}) = \frac{P_{hw}(d_{hw}/V_{hw}) + P_{tw}(d_{tw}/V_{tw})}{(d_{hw}/V_{hw}) + (d_{tw}/V_{tw})} - \bar{P} = 0 \tag{28}$$

In order to solve this constrained optimization problem, the method of Lagrange multipliers can be used, which involves the introduction of the following Lagrangian $L(V_{hw}, V_{tw})$:

$$L(V_{hw}, V_{tw}) = \bar{V}_{\rightarrow}(V_{hw}, V_{tw}) + \lambda C(V_{hw}, V_{tw}) \quad (29)$$

where λ is a Lagrange multiplier. The solution to this optimization problem is obtained as the stationary points of the Lagrangian:

$$\frac{\partial L}{\partial V_{hw}} = 0 = \frac{2V_{tw}^2}{(V_{hw} + V_{hw})^2} + \lambda \frac{\partial C(V_{hw}, V_{tw})}{\partial V_{hw}} \quad (30)$$

$$\frac{\partial L}{\partial V_{tw}} = 0 = \frac{2V_{hw}^2}{(V_{hw} + V_{hw})^2} + \lambda \frac{\partial C(V_{hw}, V_{tw})}{\partial V_{tw}} \quad (31)$$

$$\frac{\partial L}{\partial \lambda} = 0 = C(V_{hw}, V_{tw}) \quad (32)$$

These equations represent a system of three coupled equations in the unknowns V_{hw}, V_{tw} and λ, which can be solved numerically. As it turns out, the optimal pacing strategy consists of completing the headwind section of the course at approximately 500 W and the tailwind section at 266 W, resulting in a 3.1s time bonus and a total energy expenditure of 98.76 kJ, i.e. 1.24 kJ or 1.24% less than the constant-power strategy. Within reasonable bounds of $P_{hw} = 400 \pm 100$ W, the difference between the optimal and the worst pacing strategy is as big as 6 kJ, or a 6% difference in energy expenditure – all for the same time-averaged power of $\bar{P}$ = 400W.

A similar analysis can be conducted for an uphill trajectory consisting of two inclined sections of different grades. For example, if we replace our earlier uphill trajectory of 8% (see Section 2) by an uphill trajectory consisting of a 650.5 m long section of 6% followed by an equally long section of 10%, then completing this course using a constant-power strategy at $\bar{P}$ = 400W requires 100 kJ of mechanical energy, whereas the optimal pacing strategy yields a 1.5s time bonus and requires approximately 99.4 kJ of energy, i.e. a reduction of roughly 0.6%. Within reasonable bounds of P_{hw}, the difference between the optimal and worst pacing strategy is approximately 2.37 kJ or 2.37% in energy expenditure. While the effect is less pronounced than in the false-flat experiment, it is evident that on trajectories with variable resistances, pacing strategy can have a detrimental impact on the outcome of the experiment.

5. Accuracy and precision of bicycle power meters

In Section 3 of this article, the overall error of the energy expenditure measurement is expressed in terms of the accuracy of the selected power meter (see Eq. 24). Given the definition of energy as the product of power and time, the measurement error of the power meter obviously has a large influence on the overall measurement error. When using power meters to assess the relative efficiency of drivetrain technologies, absolute accuracy of the power meter (also called *trueness*) is not our main concern (it rarely is, really). Instead, we want its precision to be high, meaning that the deviation in power meter reading between identical tests is sufficiently small to allow for meaningful comparisons to be made, irrespective of the trueness of the measurement. Hence the power meter measurement error ϵ_P in Eq. 24 is a precision error, not an accuracy error. Now how realistic are the precision errors listed in Table 3 really?

Most bicycle power meters are essentially force sensors that convert the measured deformation of a reference component (e.g. a bicycle crank) into force or torque. By sampling the deformation sensors (mostly strain gauges) at high sampling frequencies, the amplitude of the force can be recorded at multiple points along the pedal stroke. As power is force times linear velocity or torque times angular velocity, a second measurement is required. The majority of commercial power meters are equipped with a cadence sensor that measures the averaged angular velocity over one complete crank revolution. The effect of this comparatively low sampling frequency is illustrated in Figure 3. The top figure (based on measurement data reported in [5]) shows the amount of torque generated by the cyclist's right ('R', dashed line) and left ('L', dotted line) legs. Clearly the majority of pedaling force is generated during the downstroke, which extends from 0 to 180 degrees of crank angle for the right leg and from 180 to 360 degrees for the left leg, with a small amount of negative force generated during the upstroke. The middle figure in Figure 3 (based on the same measurement data reported in [5]) shows the variation in the crank's angular velocity as it completes one full revolution, with a dashed line showing the average velocity. The true power P generated is equal to the instantaneous torque $M(\theta)$ multiplied by the instantaneous angular velocity $\omega(\theta)$. Instead, the power displayed by the majority of commercial bicycle power meters is the instantaneous torque multiplied by the angular velocity $\bar{\omega}$ averaged over one crank revolution, thus assuming the following approximation holds:

$$\bar{P} = \frac{1}{2\pi}\int_0^{2\pi} M(\theta)\omega(\theta)\, d\theta \cong \frac{\bar{\omega}}{2\pi}\int_0^{2\pi} M(\theta)\, d\theta \tag{33}$$

where θ is the crank angle. In the bottom figure in Figure 3, the mechanical power computed based on the instantaneous angular velocity (solid line in middle figure) and averaged angular velocity (dashed line) are shown. The average power over one crank revolution is 335.3 W in the former and 331 W in the latter case, which represents a 1.3% difference in power reading. Different pedaling patterns having their angular velocity peaks at different crank angles may show the opposite behavior, where the power measurement based on averaged angular velocity overestimates the true power (see e.g. [6]). This implies that the error of the power meter reading depends on the cyclist's pedaling pattern, and hence becomes as much a matter of precision as it is of accuracy. This is especially true when the investigated drivetrain technology affects the pedaling pattern, e.g. in the case of oval chain rings or technologies that favor different gear ratios. The advice is therefore to use a power meter that includes higher-frequency sampling of the angular velocity.

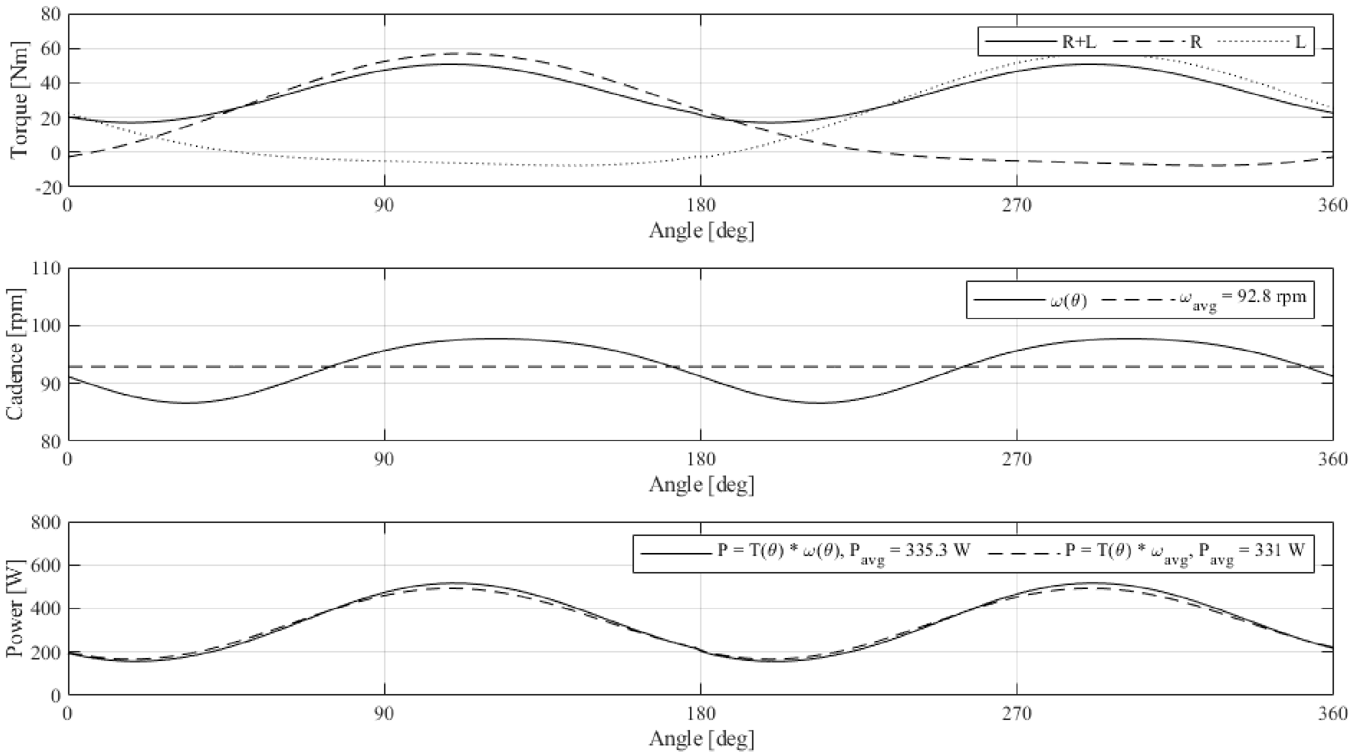


Figure 3. Torque, cadence and power in function of crank angle

Angular velocity sampling aside, commercial power meters are typically advertised with claimed accuracies of ±2% or better. Probably the most extensive investigation into the accuracy of commercial power meters has been conducted in 2017 by Thomas Maier and colleagues [7], who compared the performance of 54 power meters from 9 different manufacturers in terms of both accuracy and precision. Particularly interesting to the current study is that Maier and co-workers assessed the accuracy of time-averaged power measurements over repeated intervals of 1 minute, which compares well to the kinds of efforts involved in drivetrain efficiency measurements. The results of their analysis are summarized in Figure 4, which shows both the trueness (expressed in terms of % deviation of the ground truth) and the precision (expressed in terms of the coefficient of variation) of the investigated power meters (54 dots in total in each graph). A first observation is that only a minority of the power meter manufacturers see their advertised accuracy claims confirmed in the experiments of Maier et al. More importantly, the prescribed 1% precision required for the qualification of "medium accuracy" (see Table 2) seems difficult to attain, with only 4 out of 9 of the manufacturers achieving said precision in the majority of their power meters. A precision of 0.5% or more, as required to qualify as a highly accurate measurement according to the ranking of Table 2, is observed in less than 20% of the investigated power meters. Careful selection of the power meter used in the efficiency experiments is thus of utmost importance for the accuracy of the measurement – that, and a bit of luck.

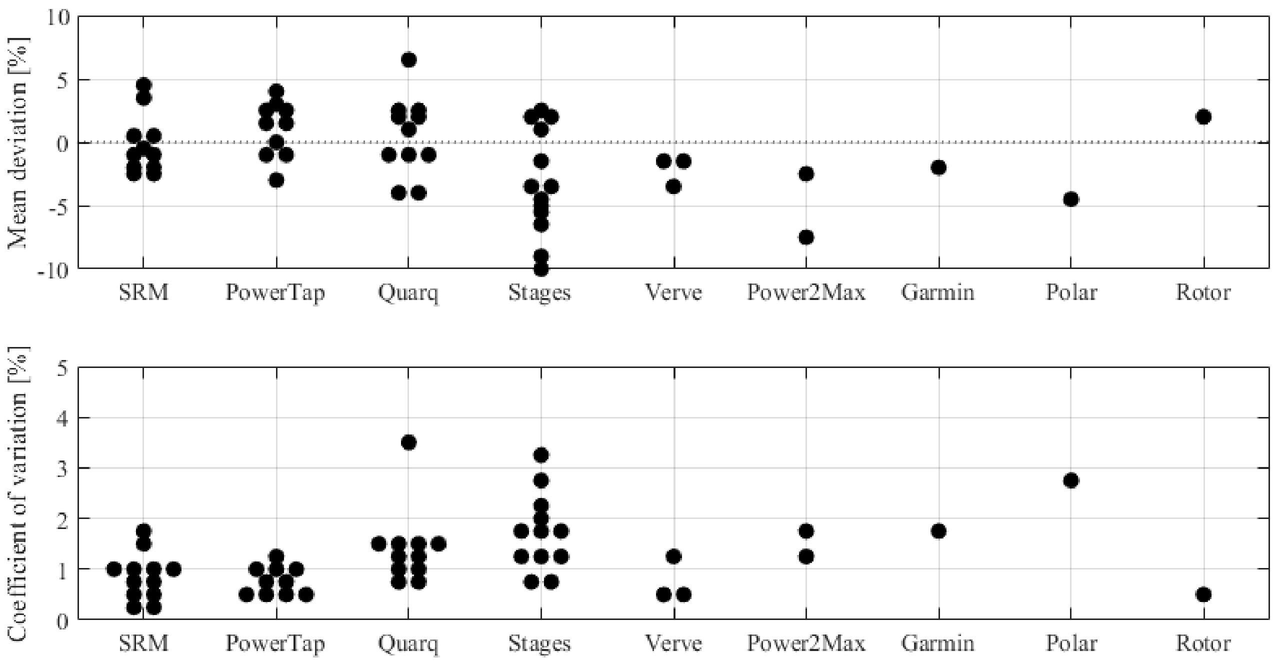


Figure 4. Trueness and precision of individual commercial power meters sorted per manufacturer (reproduced from [7]).

6. Conclusions and recommendations

This article presented the findings of an error analysis of the energy expenditure approach, which is a measurement technique often used to assess the relative efficiency of different bicycle drivetrain configurations in the field. An analytical formula for the overall error of the energy expenditure measurement was derived that consists of a measurement error and an experiment error. The former accounts for imperfections in the measurement devices whereas the latter takes the influence of difficult-to-control experiment parameters into account. In addition, the non-negligible influence of the selected pacing strategy on test trajectories with variable resistive forces was demonstrated, and a brief analysis on the accuracy and precision of commercial power meters was included to highlight the importance of proper power meter selection. Based on this analysis, the following recommendations for accurate in-the-field efficiency testing of bicycle drivetrain technology can be made:

- With regard to the selected trajectory:
    - choose an uphill trajectory that favors moderate velocities at high powers, so as to minimize the influence of environmental parameters and aerodynamic position;
    - choose a trajectory with minimal change in resistive forces (wind & inclination), so as to minimize the influence of pacing strategy; if not possible, decide on a fixed pacing strategy across all trials.

  - choose a trajectory that is sufficiently long to let the law of large numbers do its job, but sufficiently short to allow for multiple high-quality repetitions, so as to reduce the standard error of the measurement;
  - choose a trajectory with a minimum of curves and if possible draw a line on the asphalt to aid the test subject in covering a similar distance in each trial.

- With regard to the measurement devices:
  - choose a trusted power meter with high precision and preferably high-frequency sampling of both torque and angular velocity;
  - add an accurate distance sensor to the bicycle based on wheel rotation rather than GPS measurement; increase the resolution by conducting multiple measurements per wheel rotation, and correct the measured energy according to the distance travelled;
  - set up a mechanical or optical timing device with a reaction time that is faster than the mean human reaction time of 0.25 seconds;
  - if possible, add a temperature, wind speed or even CdA sensor (yes, they exist!) to the bike to increase the accuracy of the measurement.

Failure to incorporate the majority of the above recommendations is prone to lead to measurement errors in excess of the comparatively small power losses in modern bicycle drivetrains.

References

[1] Blockmans, B. Ketelslegers, D., Valero, T., Van Druten, R. (2023) Numerical and experimental investigation of the efficiency of Classified Cycling's Powershift hub. Classified Cycling, White paper.
[2] Martin, J. C., Milliken, D. L., Cobb, J. E., McFadden, K. L., & Coggan, A. R. (1998). Validation of a mathematical model for road cycling power. Journal of applied biomechanics, 14(3), 276-291.
[3] SKF. Rolling Bearings, PUB BU/P1 10000/2 EN; 2013.
[4] Spicer, J. B., Richardson, C. J., Ehrlich, M. J., Bernstein, J. R., Fukuda, M., & Terada, M. (2001). Effects of frictional loss on bicycle chain drive efficiency. J. Mech. Des., 123(4), 598-605.
[5] Favero Electronisc, S. R. L. Influence of Angular Velocity of Pedaling on the Accuracy of the Measurement of Cyclist Power. Research article. Available online: https://cycling. favero. com/(accessed on 18 January 2022).
[6] Danek, T., Slawinski, M. A., & Stanoev, T. (2021). On modelling bicycle power-meter measurements. arXiv preprint arXiv:2103.09806.
[7] Maier, T., Schmid, L., Müller, B., Steiner, T., & Wehrlin, J. P. (2017). Accuracy of cycling power meters against a mathematical model of treadmill cycling. International Journal of Sports Medicine, 38(06), 456-461.