Stream metabolism and the open diel oxygen method: Principles, practice, and perspectives

Benoit O. L. Demars,\textsuperscript{*1} Joshua Thompson,\textsuperscript{2} J. Russell Manson\textsuperscript{3}

\textsuperscript{1}The James Hutton Institute, Ecological Sciences, Craigiebuckler, Aberdeen, Scotland, UK
\textsuperscript{2}Bangor University, School of Environment, Natural Resources & Geography, Deiniol Road, Bangor, Gwynedd, Wales, UK
\textsuperscript{3}The Richard Stockton College, Computational Science, 101 Vera King Farris Drive, Galloway, New Jersey

Abstract

Global quantitative estimations of ecosystem functions are vital. Among those, ecosystem respiration and photosynthesis contribute to carbon cycling and energy flow to food webs. These can be estimated in streams with the open channel diel oxygen method (single or two stations) essentially relying on a mass balance of oxygen over a defined reach. The method is generally perceived as low cost and easy to apply with new drift free optic sensors. Yet, it remains challenging on several key issues reviewed here: measurements of gas transfer at the air-water interface, appropriate mixing of tracers, uncertainty propagation in the calculations, spatial heterogeneity in oxygen concentrations, the derivation of net primary production (NPP) or autotrophic respiration, and the temperature dependence of photosynthesis and respiration. An extremely simple modeling tool is presented in an Excel workbook recommended for teaching the basic principles of the method. The only method able to deal with stream spatial heterogeneity is the method by Demars et al. Example data, Excel workbook, and R script are provided to run stream metabolism calculations. Direct gas exchange determination is essential in shallow turbulent streams, but modeling may be more accurate in large (deep) rivers. Lateral inflows should be avoided or well characterized. New methods have recently been developed to estimate NPP using multiple diel oxygen curves. The metabolic estimates should not be systematically temperature corrected to compare streams. Other recent advances have improved significantly the open channel diel oxygen method, notably the estimation of respiration during daylight hours.

Ecosystems may be metaphorically seen as superorganisms with organs and metabolic functions: forest as lung providing oxygen, river as blood bringing nutrients to floodplain, stream benthos as liver purifying water. The pace of activity is generally poetically expressed through breathing rate, pulse, or heartbeat (e.g., Baldocchi 2008; Palmer and Febria 2012). There is now a huge international effort trying to define river health (Sweeney et al. 2004; Young et al. 2008; Palmer and Febria 2012). Yet, quantitative estimation of many ecosystem functions remains challenging.

Aquatic ecosystem metabolism can be characterized by monitoring the diel change in dissolved oxygen concentration to estimate gross primary production (GPP), ecosystem respiration (ER), and the metabolic balance or net ecosystem production (NEP) as pioneered by Vinberg (1934, 1960) and Odum (1956). The estimates may be calculated directly as in Odum (1956) or indirectly via a model fitted on the diel oxygen curves (e.g., O’Connor and Di Toro 1970). There are also strong ecological incentives to go one step further and estimate autotrophic and heterotrophic respiration ($R_a$ and $R_h$, respectively) as well as net primary production (NPP) notably in studies of ecosystem ecological energetics (e.g., Le Cren and Lowe-McConnell 1980; Rosenfeld and Mackay 1987; Meyer 1989), nutrient cycling (Newbold et al. 1982; Hall and Tank 2003; Stutter et al. 2010), and metabolic balance under climate warming (López-Urrutia et al. 2006; Yvon-Durocher et al. 2010; Demars et al. 2011b). Despite the many methods available, no general consensus has yet emerged on the determination of $R_a$, $R_h$, or NPP in aquatic ecosystems (Robinson and Williams 2005; Kosten et al. 2014).

The present review on stream metabolism is limited to the open channel diel oxygen method as proposed by Odum (1956), and similarly applied to CO$_2$ diel change (e.g., Wright and Mills 1967; Hanson et al. 2003). Oxygen metabolism may be converted to carbon metabolism via a respiratory quotient (e.g., Hanson et al. 2003; Williams and del Giorgio 2005). The open diel oxygen change method is also applied to lakes (e.g., Staehr et al. 2010), oceans (e.g., del...
Giorgio and Duarte 2002), and with some modification to estuaries (e.g., Vallino et al. 2005). Recently, it has been integrated with isotopic approaches to better constrain parameter estimates (Venkiteswaran et al. 2008; Holtgrieve et al. 2010). Alternative in situ open methods have been developed for coastal systems which may be applied to some running water systems but they tend to be more technical and expensive (e.g., Berg et al. 2003; McGillis et al. 2011; Long et al. 2013). Those eddy and gradient flux methods also measure over a small footprint and are explicitly benthic. Other methods are available to quantify photosynthesis such as the microelectrode (Jørgensen et al. 1979), 14C (Steemann Nielsen 1946; Peterson 1980; Marra 2009), triple-isotope approach (Luz and Barkan 2000; Nicholson et al. 2012). There is still a debate however as to what exactly is being measured by the different methods (net vs. gross photosynthesis, Robinson and Williams 2005). There is also a wide range of closed systems (chamber methods) which may be used in combination to (or instead of) open in situ methods (e.g., Bott et al. 1997; Dodds and Brock 1998; Uzarski et al. 2001) but scaling issues remain problematic (Petersen et al. 2009; Hanson et al. 2011).

The open diel oxygen method has a set of assumptions. Oxygenic metabolism is generally the dominant pathway (Raven 2009) but respiration estimates will ineluctably include anaerobic metabolic pathways via oxidation of their respiration products in the upper part of the benthos (Canfield et al. 2005, p. 199; Trimmer et al. 2009). ER also includes chemical oxidation (nitrification) and photo-oxidation of organic matter, not to the same extent if estimated from O2 or CO2 diel changes (Estapa and Mayer 2010). ER is generally assumed constant throughout the day (but see Parkhill and Gulliver 1998, 1999) because until recently it has not been possible to separate respiration from photosynthesis during the daylight hours (Raven and Beadall 2005). Two techniques look promising in assessing respiration of oxygenic metabolism throughout the day: the diel change in δ18O of the dissolved oxygen (Tobias et al. 2007; Hotchkiss and Hall 2014) and the resazurin-resorufin approach (Haggerty et al. 2009; Argerich et al. 2011; González-Pinzón et al. 2012). Those studies suggest that assuming constant rates of ER throughout the day is unrealistic, bearing in mind large uncertainties in parameter estimation and underlying processes.

At first, the diel oxygen method appears to be a very simple oxygen mass balance, yet it can be deceptively complicated in practice. Recent efforts and new methodological developments have strived to improve the determination of gas exchange, quantify parameter sensitivity, propagate uncertainties, estimate autotrophic and heterotrophic respiration, and apply temperature corrections. Metabolic estimates should now be more accurate, but many studies fall short of understanding key principles or applying them in practice at the expense of additional (implicit) uncertainties. We aim to review recent improvements in understanding and practice of the art in rivers to complement a previous lake perspective (Staehr et al. 2010).

In the first section, the single station approach is presented with an Excel tool allowing the investigation of the sensitivity of diel oxygen to key parameters (recommended for teaching basic principles). Gas transfer velocity estimates from modeling approaches and empirical equations are compared to direct field measurements. In the second section, we critique the blind application of the Schmidt number dependence and discuss new developments in field gas tracer studies (mixing length, injection methods, and choice of tracers). The third section is devoted to spatial heterogeneity and propagation of uncertainties in calculations, a hot topic not confined to streams (e.g., van de Bogert et al. 2012; Antenucci et al. 2013; McNair et al. 2013). The method proposed by Reichert et al. (2009) is now put to the test devised by Demars et al. (2011b). New developments in transient storage metabolism are also considered. Finally, we comment on the temperature dependence and derivation of autotrophic and heterotrophic respiration in aquatic ecosystems, critical parameters in many ecological applications. We conclude with a set of recommendations and perspectives.

**Single station in homogeneous streams**

**Principles**

The net metabolism of an aquatic ecosystem can be computed from a single diel oxygen curve, the gas transfer velocity, and mixing depth (Odum 1956). The change in O2 concentration at a single station between two subsequent measurements can be approximated as:

\[ \frac{dC}{dt} = \frac{(C_t - C_{t-1})}{\Delta t} \]  

with \( C \) concentration of oxygen (mg O2 L\(^{-1} \) or g O2 m\(^{-3} \)) at time \( t \) and can be modeled as follows:

\[ \frac{dC}{dt} = \frac{K_L(C_{SAT} - C_t) + GPP - ER}{z} \]  

with \( K_L \), gas transfer velocity (m h\(^{-1} \)); \( C_{SAT} \), oxygen saturated concentration of O\(_2\) as a function of water temperature and atmospheric pressure (mg O\(_2\) L\(^{-1} \) or g O\(_2\) m\(^{-3} \)); \( C_t \), oxygen concentration at time \( t \) (mg O\(_2\) L\(^{-1} \) or g O\(_2\) m\(^{-3} \)); GPP, gross primary production (g O\(_2\) m\(^{-2} \) h\(^{-1} \)); ER, ecosystem respiration (g O\(_2\) m\(^{-2} \) h\(^{-1} \)); \( z \), mixing depth (m).

Models are often applied to contrasting datasets, but are seldom tested (Hanson et al. 2008). Holtgrieve et al. (2010) presented a set of scenarios using known \( K_L \), GPP, and ER to test how their Bayesian ecosystem metabolism model performed under known conditions. They were able to reproduce as in previous studies (e.g., Simonsen and Harremoes 1978; Kosinski 1984) known behavior such as the flattening of the diel oxygen curve with increasing \( K_L \) and the increasing time lag in oxygen peak relative to solar noon with
decreasing $K_L$ for daily GPP $\approx$ ER. They also reported uncertainties in model performance under a set of known daily GPP/ER ratios and $K_L$. Such test datasets with known parameters help understanding and visualizing the calculation methods and can readily be prepared as follows (see Excel workbook in Supporting Information).

First, a bell-shaped curve of photosynthetic radiation (PAR) was drawn with a peak at 1500 $\mu$mol Q m$^{-2}$ s$^{-1}$, and 10/14 h night-day cycle (this can easily be altered, e.g., Simonsen and Harremoës 1978). The scenarios were based on a 1-h time step (shorter time intervals are recommended in practice; see Kosinski 1984; Staehr et al. 2010). Then, GPP was calculated with a Michaelis-Menten type equation as follows:

$$GPP = \frac{GPP_{\text{MAX}} \cdot \text{PAR}}{k_{\text{PAR}} \cdot \text{PAR}}$$

(3)

with $GPP_{\text{MAX}}$, maximum GPP (set at 2 g O$_2$ m$^{-2}$ h$^{-1}$) and $k_{\text{PAR}}$, PAR at which half the $GPP_{\text{MAX}}$ is produced (set at 500 $\mu$mol Q m$^{-2}$ s$^{-1}$). Other similar equations are available for light saturation including a temperature correction for GPP (e.g., Eilers and Peeters 1988; Henley 1993; Parkhill and Gulliver 1999). ER was set as constant throughout the night-day cycle (but see Parkhill and Gulliver 1999). This allowed the calculation of NEP = GPP − ER at hourly time steps. Hourly ER was adjusted to ER = −0.54 g O$_2$ m$^{-2}$ h$^{-1}$ so that daily ER balanced daily GPP in the first instance (i.e., daily NEP = 0). It is interesting to play with the absolute metabolic rates of GPP and ER and GPP/ER to see how the oxygen curves change (e.g., Kosinski 1984; Holtgrieve et al. 2010; Staehr et al. 2010).

To simplify the calculations, normal atmospheric pressure (760 mm Hg) and 20 $^\circ$C constant temperature were applied (so $C_{\text{SAT}} = 9.59$ g O$_2$ m$^{-3}$). The diel curve of oxygen was therefore similar for O$_2$ concentration (g O$_2$ m$^{-3}$) and saturation (%). The mixing depth of the system was set to 1 m to include the oxygen curve change (e.g., Kosinski 1984; Holtgrieve et al. 2010; Staehr et al. 2010).

This simple tool allows understanding the basics of stream metabolism calculations, exploring uncertainties in parameter estimation and selecting appropriate equipment. For example, while river ecologists are used to conditions where daily ER $>$ GPP and moderate to high gas transfer velocities, it is not always immediately apparent, why oxygen saturation can be so highly variable and sustain saturation at night above 100% in deep slow flowing rivers and standing waters (Williams et al. 2000; Palmer-Felgate et al. 2011). This can easily happen during an algal bloom in systems with GPP $>$ ER and low $K_L$ (typically $<1$ m d$^{-1}$)–e.g., Staehr et al. 2010. Similarly at the end of the algal bloom, GPP $<$ ER and the system can quickly shift to very low saturation values down to anoxia (Williams et al. 2000; Palmer-Felgate et al. 2011).

The problem of oxygen sensor calibration (accuracy) is absolutely crucial and cannot be underestimated, especially in turbulent and/or low productivity streams. Calibration of the sensors, using water temperature, atmospheric pressure, and stream water, must be done before deployment, the oxygen sensors must be cross calibrated in the stream at the
start and end of deployment, and calibration checked again at the end. Long-term deployment requires additional regular calibration checks. Slight discrepancies between sensors can then be corrected for the two station method (see below).

Gas transfer velocity: Direct determination or modeling

It is clear from Fig. 1 that the gas transfer velocity (or gas exchange) greatly affects the diel oxygen curve. Note the results of the first day of the time series may be affected by the chosen initial concentration of oxygen, depending on the setting (another modeling property to explore). Large errors in stream metabolism estimates (especially respiration) will occur with high gas transfer velocity (McCutchan et al. 1998). Hence, models will perform best when $k$ is measured out in the field with typical uncertainties of 10% in small streams with discharge generally under a m$^3$ s$^{-1}$ (e.g., Thysen et al. 1987; Thene and Gulliver 1990; Genereux and Hemond 1992; Melching 1999) and 30% in larger (1–10 m$^3$ s$^{-1}$) wide and shallow heterogeneous streams (Demars, unpubl.). The gas transfer velocity may be derived from existing empirical or theoretical equations (both generally fitted on field or lab reaeration studies) but their precision is generally relatively poor (at best 40–125% error) and both tend to be restricted to relatively narrow conditions of applications (see, e.g., Cox 2003; Wallin et al. 2011; Raymond et al. 2012; Demars and Manson 2013; Palumbo and Brown 2014).

Hence, several techniques have been proposed to derive $K_L$ from the diel oxygen curve such as the night time method (Odum 1956; Hornberger and Kelly 1975; Thysen and Kelly 1985 implemented in RIVERMET©; Izagirre et al. 2007), day time method (Hornberger and Kelly 1975 corrected in Cox 2003; Chapra and Di Toro 1991), and iterative methods with simultaneous derivation of $K_L$, GPP, and ER by best fitting to the diel oxygen curve (Holtgrieve et al. 2010; Birkel et al. 2013; Riley and Dodds 2013). Those techniques are however restricted to sinusoidal diel curves (i.e., $K_L < 0.5$ m h$^{-1}$ and high productivity, Holtgrieve et al. 2010) where they perform similarly to suitable empirical equations ($≤40\%$ discrepancy against Owens et al. 1964; Williams et al. 2000) and tracer gas studies ($≤65\%$ median discrepancy, Riley and Dodds 2013)—note there is no point in comparing equations outside their range of applicability. The derivation of $K_L$ from the diel oxygen curve understandably fails under slow metabolism or high $K_L$ where $dC/dt$ varies little at night and the time lag in minimum oxygen deficit relative to solar noon is negligible (Hornberger and Kelly 1975; Kosinski 1984; Thysen et al. 1987; Cox 2003; Holtgrieve et al. 2010). Although Birkel et al. (2013) concluded that a modeling approach could simultaneously determine $K_L$, GPP, and ER reliably in “well-oxygenated upland channels with high hydraulic roughness,” their model has the same limits as Holtgrieve et al. (2010) and their results remained untested and partly based on spurious negative $K_L$ values (the direction of the reaeration flux must be determined by the oxygen deficit) and supersaturation at night (see Hall et al. 2015), both of which had no biophysical meaning. Moreover, in similar rivers, Young and Huryn (1999) compared the day and night time regression methods to gas tracer studies and found no correlation between the two with median discrepancy over 100% (calculated as the absolute difference between the two tracer studies divided by their average). This is questioning the findings of many studies where the gas exchange coefficient was poorly constrained (e.g., Iwata et al. 2007; Izagirre et al. 2008; Birkel et al. 2013).

In practice, the derivation of the gas exchange (time$^{-1}$) with long-term datasets can be further compromised by change in hydrology, pollution event, and equipment malfunction (e.g., Uehlinger 2000; Acuña et al. 2004; Izagirre et al. 2008; Stutter et al. 2010). To fill gaps, the gas exchange (time$^{-1}$) can often be related to discharge $Q$ (personal observations), but this is site specific (Genereux and Hemond 1992; Roberts et al. 2007; Wallin et al. 2011). Relatively large uncertainties in gas transfer velocity or gas exchange may also prevent to find strong correlations with $Q$ (e.g., Izagirre et al. 2008). Hence, gas tracer studies (or modeling in large rivers) and error propagations should be the norm in stream metabolism studies (e.g., McCutchan et al. 1998; Demars et al. 2011b; Hotchkiss and Hall 2014).

Reaeration with tracer gas: Theory and practice

The reaeration flux (mg O$_2$ m$^{-2}$ min$^{-1}$) is controlled by oxygen deficit $(C_{SAT} - C_0)$ and the gas exchange coefficient $k$ (min$^{-1}$), itself the product of gas transfer velocity $K_L$ (cm min$^{-1}$) and specific surface area (area/volume, cm$^{-1}$). In streams with smooth surface water, $K_L = k z$. Note that both surface area and water depth estimates should have uncertainties attached to them, they can be tricky to estimate due to surface water turbulence and bed roughness including macrophytes. Depth is best estimated from width, velocity, and discharge.

The friction velocity model and the Schmidt number dependence

One of the most popular equation to link gas transfer velocities of two gases was derived from the friction velocity model (e.g., Jähne et al. 1987b):

$$K_{L1}/K_{L2} = (D_{m1}/D_{m2})^n = (Sc_{L1}/Sc_{L2})^{-n}$$

with $n$ the Schmidt number dependence of the gas transfer velocity and $Sc$ the Schmidt number, ratio of the molecular transport properties $v/Dm$, where $v$ is the kinematic viscosity of water (cm$^2$ min$^{-1}$) and $Dm$ the molecular diffusivity (cm$^2$ min$^{-1}$).
The Schmidt number dependence has initially been shown theoretically to be constrained with $0.5 < n < 1$ (see, e.g., Dobbins 1956; Bennett and Rathbun 1972; Jähne et al. 1987b). Direct measurements were also derived from lab experiments (Jähne et al. 1987a,b) and a field study in a lake (Watson et al. 1991) as follows:

$$n = \ln \left( \frac{K_L}{K_L^2} \right) / \ln \left( \frac{Dm_1}{Dm_2} \right)$$  \hspace{1cm} (7)$$

The determination of $n$ requires gases with different diffusivity and accurate and precise estimates of $K_L$ and $Dm$. Unfortunately, there is no reliable theory to predict $Dm$ and the experimental measurements of $Dm$ vary widely for most gases among studies, hence it is best to use the $Dm$ values of two gases of interest from the same experimental set up (Bennett and Rathbun 1972; Jähne et al. 1987a), and correct adequately for temperature using preferably a theoretical equation (e.g., Stokes–Einstein equation, Edward 1970; Demars and Manson 2013). Empirical equations for Schmidt numbers (e.g., Wanninkhof 1992; Raymond et al. 2012) are not as precise because they were derived from different experimental set ups and the reader should refer to the original references.

Demars and Manson (2013) previously noted large discrepancies between studies in the application of dual gas tracer studies with, e.g., three different values for $n$ for similar small turbulent streams ($n = 0.7$ in Jones and Mulholland 1998; $n = -0.5$ in Hope et al. 2001; $n = 0.5$ in Wallin et al. 2011), hence a few important points are recalled.

Equation 6 was only derived for continuous surface water (i.e., no broken standing waves) and has been initially tested for a limited range of $K_L$, mostly low to moderate. The only laboratory experimental sets of data with breaking waves and bubbles initially showed $n$ down to about 0.3 (outside its theoretical range) for gas transfer velocity of about 0.3 m h$^{-1}$ (Wanninkhof et al. 1993). Asher et al. (1996) and Asher and Wanninkhof (1998) subsequently showed that under broken waves (high wind) bubble mediated transfer (a function of diffusivity and solubility) needed to be taken into account and developed a new model (Asher et al. 1997). There are many other theoretical models that can be used to relate $K_L$ values between two gases of interest, most, however, have limited hydraulic ranges and require additional parameters (see Demars and Manson 2013).

Dual tracer studies in turbulent streams have been published but the differences in diffusivity between the gases was too small to derive $n$ accurately (e.g., Genereux and Hemond 1992; Benson et al. 2014), although error propagation in Genereux and Hemond (1992) is questionable. Indeed, with $ln(x \pm \delta x) = ln(x) \pm \delta x/x$, the addition of the relative errors $(\delta x/x)$ of the ratio in quadrature and the data provided, $0.65 < n < 0.81$, rather than $n = 0.7 \pm 0.9$ for $113 < k_2 < 127$ day$^{-1}$ with $0.5 < n < 1$. Rathbun et al. (1978) simply related oxygen to propane gas exchange with a single coefficient $k_{C,H_4} = 1.39 \pm 0.03$ $k_{O_2}$, independently of temperature as expected from previous experimental data ($Dm_{O_2}/Dm_{C,H_4} = 1.28 \pm 0.09$ within 10–60 °C, Wise and Houghton 1966). Note from those results $1.25 < n < 1.30$, which is also outside the theoretical range, but routinely used in stream metabolism studies. The use of helium (or neon) in dual gas tracer study would have been more appropriate due to their higher molecular diffusivity than most gases of interest (e.g. King et al. 1995). But in the end, the quest for $n$ may be a red herring, as it would not be able to account for additional mechanisms not taken into account in the initial theory (e.g., broken waves, bubbles, renewal rate, see Vachon et al. 2010; Demars and Manson 2013) and in any case Eq. 6 is largely insensitive to $n$ in its predicted range ($0.5 < n < 1$). Hence, in practice if not sure what theoretical model should be used to link the two gases of interest in a given system, Eq. 6 may be used without $n$ as in early studies (Tsivoglou et al. 1965; Rathbun et al. 1978).

### Mixing length

There are a number of comprehensive primers on stream tracer studies (e.g. Kilpatrick et al. 1989). Although there are some empirical and theoretical formulae to predict solute and gas mixing length (Wallis and Manson 2004), in practice it is not always easy to define the optimal mixing length (e.g., Young and Huryn 1999 reported potential lack of mixing). Many studies have been content with qualitative observation using a dye; however, lateral and vertical mixing cannot be guaranteed, especially in deep slow flowing rivers where precise gas measurements are needed due to the low evaporation rate of the tracer gas. For this reason, in deep slow flowing rivers the extraction of the gas exchange coefficient directly from the oxygen curves using modeling is likely to provide a more accurate result. In small $(Q < 1$ m$^3$ s$^{-1})$ to medium $(Q < 10$ m$^3$ s$^{-1})$ size stream, the combination of a slug injection of a conservative tracer (e.g., Cl$^-$, Br$^-$, conservative dye) and continuous gas tracer study (e.g., C$_3$H$_8$, SF$_6$) can allow determining hydraulic parameters and mixing adequately. Mixing may be more accurately certified from continuous gas tracer studies, coupled with a conservative tracer, where gas samples are collected across the channel at different depth.

### Injection methods and choice of tracers

Bubbling gas across the channel with air stones has been successfully applied for decades (Kilpatrick et al. 1989). But it is rather wasteful in shallow streams (≈0.3 m deep) as more than 80–99% of the gas may escape in the atmosphere (Kilpatrick et al. 1989; Benson et al. 2014). This is of particular concern when SF6 is used as it is significantly more expansive than commercial propane and is a strong greenhouse gas. Regarding propane, icing of the gas regulator may occur under high gas flow ($\approx 1$ kg liquid gas h$^{-1}$, note this will vary between gas regulators from different companies) with subsequent drop in gas flow, so the use of a pressure gauge
on the outlet of the gas regulator is imperative. Two alternatives are available. One method is to diffuse the gas through silicone or Teflon tubing insuring more than 99% of the gas is solubilized in water (Sanford et al. 1996; Cook et al. 2006; Benson et al. 2014). Another method is to prepare a saturated solution of tracer gas in a collapsible bag and use a peristaltic pump for constant rate injection (Tobias et al. 2009; Jin et al. 2012). The latter would also ensure complete solubility of the tracer gas.

In large rivers (Q > 10 m$^3$ s$^{-1}$), a tracer gas solution is generally injected as a slug together with a conservative tracer. The use of propane may be limited at high discharge (Q > 100 m$^3$ s$^{-1}$) by its limit of quantification (0.5–1 mg m$^{-3}$, Kilpatrick et al. 1989; Theno and Gulliver 1990; Roldão 1991). In such case, it may be necessary to use tracer gas amenable to gas chromatography (GC) with electron capture detector for halogenated gases (e.g. Busenberg and Plummer 2010). GC mass spectrometry for noble gases (e.g. Wanninkhof et al. 1993; Mackinnon et al. 2002; Benson et al. 2014), or other isotopic methods (e.g. Yang et al. 2013). Alternatively, as the oxygen gas exchange coefficient tends to be low in large (deep) rivers, it may be easier and more accurate to estimate it directly from the oxygen curves using a modeling approach (e.g. Holtgrieve et al. 2010; Hotchkiss and Hall 2014).

The use of electric conductivity as a conservative tracer is pragmatic in small streams but the salt cation is generally not as conservative as the anion (e.g. Freeman et al. 1995). Fluorescent dyes are also used routinely in large systems although there are not always entirely conservative either in natural waters (Clark et al. 1996). Cl$^-$, Br$^-$, or $^3$He should be used for more accurate work (see Watson et al. 1991; Wanninkhof et al. 1993; Clark et al. 1994). The floating dome approach has known defects but is easy to use and inexpensive (see Richey et al. 2002; Vachon et al. 2010; Beaulieu et al. 2012).

Some old techniques have been superseded; one being the oxygen depletion method (Edwards et al. 1961) as it can interfere with the biota (Lavandier and Capblancq 1975), and another the use of radioactive gases $^{85}$Kr (Tsivoglou et al. 1968) due to strict controls on their use (Rathbun et al. 1978). The major shortcoming of current (dual) gas transfer studies (e.g., C$_3$H$_8$, SF$_6$, $^3$He) is their limited temporal scale of application despite known rapid changes in gas transfer velocity (Yotsukura et al. 1984; Tobias et al. 2009). Hence temporal scaling of $k$ (or $K_y$) rely on building rating curves with wind (Wanninkhof 1992), discharge (Roberts et al. 2007; Beaulieu et al. 2013), sound (Morse et al. 2007), or direct measurements of water turbulence (Vachon et al. 2010).

Two-station methods in heterogeneous streams

Spatial heterogeneity

The two station method has long been perceived as more accurate (Odum 1956; Marzolf et al. 1994; Hall and Tank 2005; McCutchan and Lewis 2006). Demars et al. (2011b) uncovered a paradox, however, where despite large saturation deficit of the oxygen curves at night (driven by intense respiration), the two station method can report positive net metabolism, i.e., photosynthesis at night. This happens, assuming constant discharge, when the difference (lagged by travel time) in dissolved oxygen concentration between the two stations exceeds reaeration $k_x(C_{SAT} - C_{AV})$ expressed in mg L$^{-1}$, with $\tau$ mean travel time and $C_{AV}$ average oxygen concentration. It is best visualized with turbulent streams ($k > 1$ h$^{-1}$) where the oxygen concentration is constant at night (see Fig. 1 above). In turbulent stream, the footprint of the oxygen sensors is also relatively limited and affected by fine scale spatial heterogeneity (e.g., shading, reaeration, lateral inflows, substrate). Our measurements of key parameters such as travel time ($\tau$), lateral inflows ($Q_{in}$), depth ($z$), reaeration coefficient ($k$), and saturation deficit ($C_{SAT} - C_{AV}$) are, however, estimated at coarse spatial scale (entire river reach). The 95% footprint (length scale $L$) of the oxygen sensor is generally calculated as $L \approx 3u/k$, but the 50% footprint $L \approx 0.7u/k$ is 4.3 times shorter (with $u$, velocity in m h$^{-1}$ and $k$, gas exchange coefficient in h$^{-1}$, see Supporting Information for derivation). When the oxygen sensors are precisely calibrated (see above), the difference (lagged by travel time) in dissolved oxygen concentration between the two stations is the result of spatial heterogeneity where the two oxygen sensors are affected differently by local conditions (see Fig. 1 in Demars et al. 2011b). This is the root of the problem because the two station method assumes spatial homogeneity. A simple solution to these problems is to average the dissolved oxygen concentration from the two (or multiple) stations to create a single diel oxygen curve. It allows complying with underlying assumptions (spatial homogeneity) and propagating uncertainties due to spatial heterogeneity (see Demars et al. 2011b). It also provides better results when oxygen sensors drift unpredictably (see Fig. S2 in Demars et al. 2011b).

Another solution, mathematically grounded with useful recommendations, was presented earlier by Reichert et al. (2009). Their model provided an estimate of net metabolism in heterogeneous river reaches with an exponentially weighted average of the dissolved oxygen concentrations along the reach, with more weight on subreaches closer to the bottom station. Notably, assuming an homogeneous reach, they suggested that the two station method is most useful for a reach length $L$ within $0.4u/k < L < 3u/k$. Below $0.4u/k$, the bottom station is mostly (>2/3) affected by the top station (rather than within reach stream processes) and above $3u/k$ the two stations are essentially independent (and one-station method will provide the same result). The optimum is probably $0.4u/k < L < 1u/k$, where the bottom station is affected by 1/3 to 2/3 of the top station.

We calculated the net metabolism for the four reaches of the River Luteren using Demars et al. method (2011b, Fig. 2). There was a broad agreement between Demars and
Reichert methods using the Luteren data (Fig. 2), with differences in respiration estimates along the four reaches (1–4) of 14%, 33%, 3%, and 14% (based on absolute difference/average) similarly to a comparison between Odum (Bott 2007) and Reichert methods by Hondzo et al. (2013) in Minnehaha Creek. The apparent discrepancy in the second reach was further investigated. The net metabolism (NEP) in Luteren reach 2 at night was according to Odum (1956) \(0.47 \text{ g O}_2 \text{ m}^{-2} \text{ h}^{-1}\), Demars et al. (2011b) \(0.35 \text{ g O}_2 \text{ m}^{-2} \text{ h}^{-1}\), and Reichert et al. (2009) \(0.25 \text{ g O}_2 \text{ m}^{-2} \text{ h}^{-1}\). The method by Reichert et al. (2009) was then tested similarly to other methods as in Demars et al. (2011b). In Luteren reach 2, the difference (lagged by travel time) between stations in oxygen concentration was already close to reaeration (expressed in mg L\(^{-1}\) as above). The dissolved oxygen data of Luteren reach 2 were then slightly modified to increase the oxygen concentration difference between the two stations without changing the average expected oxygen saturation \((C_{\text{top}} - 0.2 \text{ mg L}^{-1} \text{ and } C_{\text{bot}} + 0.2 \text{ mg L}^{-1})\) in order to have the difference between the two stations exceeding the reaeration flux. The oxygen saturations at the top and bottom stations at night were around 87% and 92%, respectively (within 2% of the original oxygen saturation, similarly to the range of accuracy of many oxygen sondes)—Fig. 2. The resulting net metabolism were according to Odum (1956) \(+0.06 \text{ g O}_2 \text{ m}^{-2} \text{ h}^{-1}\), Demars et al. (2011b) unchanged at \(-0.35 \text{ g O}_2 \text{ m}^{-2} \text{ h}^{-1}\), and Reichert et al. (2009) \(+0.06 \text{ g O}_2 \text{ m}^{-2} \text{ h}^{-1}\) (Fig. 2). Both Odum and Reichert et al. methods indicated positive NEP, i.e., photosynthesis at night. This is physically impossible (oxygen was undersaturated) and suggests, as for previously tested methods, that some critical assumptions were violated in applying the two station diel oxygen method (see Demars et al. 2011b). Note that, under similar conditions, when \(C_{\text{top}} > C_{\text{bot}}\) respiration is over estimated and when the discrepancy between the two oxygen concentrations is not as extreme there can still be a significant bias (Demars et al. 2011b).

We also tried Riley and Dodds (2013) two station method using the gas exchange coefficient provided by Reichert et al. (2009) and light data from Säntis (Federal Office of Meteorology and Climatology, MeteoSwiss) adjusted to the ground measurements (Fig. 3 in Reichert et al. 2009). However, the model could not converge adequately on the Luteren data, probably due to the combination of moderate metabolic rates and high reaeration with average \(k = 2.6 \text{ h}^{-1}\) (cf. Fig. 1 above, Reichert et al. 2009; Holtgrieve et al. 2010).

Spatial heterogeneity is present at several scales from small macrophyte patches (Hondzo et al. 2013) to reach scale shading (Reichert et al. 2009). The recommendations of Reichert et al. (2009) on stream reach length assumed homogeneity within reach, and thus may not satisfy stream with high local heterogeneity such as in Minnehaha Creek (Hondzo et al. 2013). In practice, especially in small streams, it can be difficult to find an ideal spot for the oxygen sensor, and the question is now to define how many sensors are needed to reduce error due to spatial heterogeneity (Demars et al. 2011b).

![Fig. 2. Average O2 saturation (top panels) and net ecosystem production (NEP, bottom panels) of the Luteren reach 2 calculated according to Demars et al. (2011b) in black (thinner lines represent ±1 SD) and Reichert et al. (2009) in red. Left panels: as in original study, right panels: after slightly modifying the oxygen concentration of the top (−0.2 mg L\(^{-1}\)) and bottom (+0.2 mg L\(^{-1}\)) stations. Note unchanged NEP with larger uncertainties according to Demars et al. (2011b) and positive NEP at night according to Reichert et al. (2009) which is not logical with [O2] ≈ 90% saturation at night (top right panel).](image-url)
et al. 2011b; Hondzo et al. 2013), as done for lakes (van de Bogert et al. 2012).

It is good practice to check for spatial homogeneity in dissolved oxygen along and across the studied stream (e.g. Williams et al. 2000; Hondzo et al. 2013) including the zone of influence upstream the top station (see Demars et al. 2011b and below). Measurements at night may better point toward potential issues of lateral inflows (independently of photosynthetic activities), at least in turbulent streams with stable dissolved oxygen concentration at night.

Uncertainties

Most studies in stream metabolism estimates are blighted by unreported uncertainties and known unknowns (lateral inflows and spatial heterogeneity including the zone of influence above the top station). Yet, we know that uncertainties and associated biases can be in the same order of magnitude as the final metabolic estimates (McCutchan et al. 1998; Reichert et al. 2009; Demars et al. 2011b). More rigour is needed and the emergence of new modelling technique is promising (e.g., Holtgrieve et al. 2010; Solomon et al. 2013; Hotchkiss and Hall 2014).

First, it is wrong to think that unbiased estimates of GPP and (especially) ER can be derived from sites where the top station is differently affected by upstream processes compared to within reach processes, e.g., groundwater upwelling (Hall and Tank 2005; Riley and Dodds 2013), cascade or self-aerated flows (Birkel et al. 2013)—see Fig. 1 in Demars et al. (2011b). Riley and Dodds (2013, see their Fig. 2) presented change in dissolved oxygen saturation along two streams affected by low oxygen groundwater at the top of the reach. This type of data should not be used, however, for stream metabolism calculations without first correcting the bias in dissolved oxygen due to groundwater inflows upstream of the reach (see Demars et al. 2011b) or adding a transport model for dissolved oxygen where as it travels downstream, dissolved oxygen reaches equilibrium with the atmosphere according to some gas exchange coefficient $k$ all while having ER and GPP act on the dissolved oxygen concentration.

Second, when groundwater inflows are detected along a stream, it is important to assess the likelihood of bias. This will depend on both the relative size of the inflows and the difference in oxygen concentrations between inflows and main channel (Hall and Tank 2005; McCutchan and Lewis 2006). The statement by Izagirre et al. (2008) “We were unable to take groundwater inputs into account, but the relatively high metabolic rates measured and turbulent character of Basque streams suggest this problem is of minor importance in the overall picture” is not valid since high respiration rates could simply result from lateral inflows with low oxygen concentrations and reaeration rates do not affect the proportion of ER bias (it will simply amplify the bias in the same proportion as ER). Similarly Young and Huryn (1999) stated “[.] the two-station method appeared to give negative gross primary production estimates during the day. The most likely cause of this problem was an underestimate of the reaeration coefficient using the propane injection method” is also not valid because the reaeration rate only amplifies, i.e., it would not change the sign of photosynthesis. Negative GPP has been present in other studies (e.g. Marcarelli et al. 2010 as reported by Hall and Beaulieu 2013), most likely due to modeling artifacts.

Spatial heterogeneity and lateral groundwater inflows generally remain known unknowns at the likely cost of large discrepancies and potential lack of correlations in ER between one and two station methods (e.g. Uzarski et al. 2001; Reichert et al. 2009; Hondzo et al. 2013). Both can be taken into account, however, and their uncertainties propagated into the final estimates as, e.g., in Demars et al. (2011b). The concentration of dissolved oxygen in the lateral inflows can be spatially very heterogeneous and extremely difficult to estimate, and so wherever possible sites without lateral inflows should be chosen. The simple direct calculation method of Demars et al. (2011b) gives better metabolic estimates than any other methods in spatially heterogeneous streams. Here, we provide example data, an Excel workbook and R script to run the calculations (see Supporting Information). The R script does not have the correction for lateral inflows, but this could be easily implemented. The propagation of uncertainties in Demars et al. (2011b) could be improved, however, using bootstrapping.

Hyporheic vs. channel metabolism

Spatial heterogeneity has been extended to estimating how stream metabolism depends on the hyporheic or transient storage zone of the stream (Haggerty et al. 2009; Gonzalez-Pinzón et al. 2012). The transient storage zone ($A_s$) is characterized by the stream bed interstitials and other near stagnant water regions that occur in natural channels and by the rate at which material is transferred in and out of these regions. Direct measurement of $A_s$ is not possible for natural geometries and so recourse is to an indirect measurement wherein the temporal residence time of a conservative tracer within a stream reach is measured. A conservative tracer solution is injected into a stream and then recorded over time downstream the mixing zone at two locations under constant discharge. The upstream tracer curve is used as an inflow boundary condition for the solution of the partial differential equations developed by Bencala and Walters (1983) that describe solute transport in streams,

$$\frac{\partial C}{\partial t} = -u \frac{\partial C}{\partial x} + D \frac{\partial^2 C}{\partial x^2} + \alpha (s-C) - \frac{q}{A} C$$

$$\frac{ds}{dt} = -s \frac{A}{A_s} (s-C)$$

where $C$ is the concentration of tracer in the main channel, $u$ is average water velocity (m s$^{-1}$), $s$ is the concentration of tracer in the transient storage zone, $A$ is the stream channel geometry and so recourse is to an indirect measurement wherein the temporal residence time of a conservative tracer within a stream reach is measured. A conservative tracer solution is injected into a stream and then recorded over time downstream the mixing zone at two locations under constant discharge. The upstream tracer curve is used as an inflow boundary condition for the solution of the partial differential equations developed by Bencala and Walters (1983) that describe solute transport in streams,

$$\frac{\partial C}{\partial t} = -u \frac{\partial C}{\partial x} + D \frac{\partial^2 C}{\partial x^2} + \alpha (s-C) - \frac{q}{A} C$$

$$\frac{ds}{dt} = -s \frac{A}{A_s} (s-C)$$

where $C$ is the concentration of tracer in the main channel, $u$ is average water velocity (m s$^{-1}$), $s$ is the concentration of tracer in the transient storage zone, $A$ is the stream channel geometry and so recourse is to an indirect measurement wherein the temporal residence time of a conservative tracer within a stream reach is measured. A conservative tracer solution is injected into a stream and then recorded over time downstream the mixing zone at two locations under constant discharge. The upstream tracer curve is used as an inflow boundary condition for the solution of the partial differential equations developed by Bencala and Walters (1983) that describe solute transport in streams,
cross-sectional area ($m^2$), $A_s$ is the transient storage cross-sectional area ($m^2$ m$^{-1}$ or $m^3$), $D$ is the longitudinal dispersion coefficient, $x$ is the solute exchange parameter between the main channel and the storage zones ($s^{-1}$), $q$ is the lateral volumetric inflow rate per unit length ($m^3$ s$^{-1}$ m$^{-1}$), $x$ is the longitudinal spatial co-ordinate (m), and $t$ is time (s).

Current best practice suggest that these equations should be solved numerically (Cox and Runkel 2008; Demars et al. 2011a) and the resulting downstream prediction is fitted to the observed downstream tracer curve by optimal choice of the parameters ($u, D, x, A_s$); $q$ is calculated beforehand from the curves as in a dilution gauging calculation. These model parameters may be recovered from tracer data using this technique so long as the Damkohler number is roughly within an order of magnitude of unity (Wagner and Harvey 1997). Although these hydraulic parameters may be recovered, large uncertainties remain for the parameters $A_s$ and $x$ (Wallis et al. 2013; Ward et al. 2013).

Where these hydraulic parameters can be estimated with reasonable accuracy, they may be used in the estimation of stream metabolism by another method. The following equations describe the fate and transport of dissolved oxygen in a stream with transient storage zones (under constant discharges, without groundwater inflows),

$$\frac{\partial C}{\partial t} = -u\frac{\partial C}{\partial x} + D \frac{\partial^2 C}{\partial x^2} + x(C_s - C) + k(C_{\text{SAT}} - C) + P - R$$

(10)

$$\frac{dC}{dt} = -\frac{x}{A_s} \frac{A}{C_s} (C_s - C)$$

(11)

where $C$ is dissolved oxygen concentration in the main channel ($g$ m$^{-3}$), $C_s$ is dissolved oxygen concentration in the transient storage zone ($g$ m$^{-3}$), $C_{\text{SAT}}$ is saturated dissolved oxygen concentration ($g$ m$^{-3}$); $P$ and $R$ represent the daily average rate of dissolved oxygen production and depletion by photosynthesis (GPP) and respiration (ER), respectively, in g m$^{-3}$ s$^{-1}$. Groundwater inflows can be taken into account with an additional term $\frac{x}{A_s} (C_s - C)$ in Eq. 10, with $C_s$ oxygen concentration of the groundwater inflows.

Using the known hydraulic parameters, $u, D, x, A_s,$ and $q$, and taking the upstream diel dissolved oxygen curve as the inflow boundary condition, a model prediction based on these equations (and the same numerical method as above) may be fitted to the downstream diel dissolved oxygen curve by optimizing for $P$ and $R$ which, in the simplest analysis, may be taken as constants. Alternatively, more complex models of photosynthesis and respiration with additional parameters can be employed. For example, a model for photosynthetic rate which incorporates the effect of varying light levels could be employed,

$$P(I) = P_{\text{max}} \left(1 - e^{-\gamma I}\right)$$

(12)

where $P_{\text{max}}$ is the specific photosynthesis rate at optimum light conditions, $I$ is the light intensity in W m$^{-2}$ (observed), and $\gamma$ is the initial slope of the light-saturation curve (Jassby and Platt 1976). $P_{\text{max}}$ and $\gamma$ now become fitting parameters rather than simply a constant.

While more complex, the transient storage model may be tested experimentally in stream with the use of resazurin, a smart tracer changing colour in the presence of oxygenic metabolism (Haggerty et al. 2009; Argerich et al. 2011; González-Pinzón et al. 2012).

### Net primary production (NPP), autotrophic ($R_s$) and heterotrophic ($R_h$) respiration

#### Methods

The determination of autotrophic and heterotrophic respiration requires additional assumptions and careful considerations of ecosystem functioning (Roxburgh et al. 2005; Lovett et al. 2006), because it is not possible using a single diel oxygen curve to derive $R_s$, $R_h$, and NPP from the diel oxygen method (Kosten et al. 2014). This said, using multiple diel oxygen curves collected across sites or year round at one site, Solomon et al. (2013) and Hall and Beaulieu (2013) suggested a regression approach based on ER-GPP coupling (at standard temperature, see below) to determine basal respiration ($\gamma$ axis intercept) and proportion of $R_s$ to GPP (slope $\gamma$ of ER on GPP). It assumes that autotrophic respiration includes heterotrophic respiration by heterotrophs of autochthonously produced labile organic matter. Hence both methods assume heterotrophic respiration tends towards zero ($R_h \rightarrow 0$) in streams with negligible allochthonous carbon inputs (e.g., desert or spring fed streams). Although $R_h$ could result from autochthonously produced $C$ from an earlier time, generally ER correlates closely to GPP (e.g., Stockner 1968; Huryn et al. 2014). The method requires at least 50 days of daily metabolism estimates at a given site and thus the approach is limited to available long term data (Hall and Beaulieu 2013). Most of the results from Hall and Beaulieu (2013) were derived from poorly constrained metabolic estimates (data from Izaguirre et al. 2008, see above). Despite their high similarity Solomon et al. (2013) and Hall and Beaulieu (2013) may give very different answers. Solomon et al. (2013) approach will work best for systems where GPP and ER are strongly correlated. This is, however, not always the case (e.g. Robinson and Williams 2005; Hall and Beaulieu 2013).

A wide range of other studies have reported estimates of $R_s$, $R_h$ or NPP/GPP using a wide range of alternative methods (see Table 1). None of these studies and methods are either bullet-proof or generally applicable (Robinson and Williams 2005), and congruency in results may not warrant better accuracy. The fraction $\gamma = R_h / GPP$ has also generally been assumed constant in a given study, although this fraction likely changes along the stream continuum (Meyer 1989), grazing activity (McIntire et al. 1996) and environmental gradients such as light or nutrient (Solomon et al. 2013).
One alternative is to use a range of carbon use efficiencies $(1 - \alpha)$ values and see what are the ecological implications (e.g., Meyer 1989; Demars et al. 2011b). Just as Meyer (1989) concluded 25 years ago, new complementary or alternative methods to quantify $\alpha$ in aquatic ecosystems are still needed.

### Temperature dependence

The general consensus is that more comparable metabolic estimates are obtained when standardized for a reference temperature, generally 20 °C (e.g., Wilcock et al. 1998; Riley and Dodds 2013). There are three related ways of standardizing metabolic estimates and the gas exchange coefficient: von Hoff $Q_{10}$, Arrhenius activation energy $E_a$ (eV), and temperature coefficient $\theta$ derived from a simplification of Arrhenius equation (e.g., Yongsiri et al. 2004; Hernández-León and Ikeda 2005). The mathematically most accurate is the activation energy (Gillooly et al. 2001; Yongsiri et al. 2004). The three constants are related as follows:

$$E_a = \ln \left( \frac{Q_{10}}{E_a} \right) R T_1 T_2$$

(13)

$$\theta = \exp \left[ E_a \left( \frac{R T_1 T_2}{2} \right) \right]$$

(14)

with Boltzmann gas constant $R = 8.62 \times 10^{-5}$ eV K$^{-1}$ and $T$ temperature in Kelvin.

The major issue is that there is not a universal $Q_{10}$, $E_a$, or $\theta$ and that many studies have simply used a constant without any justification. There is in fact a wide range of $\theta$ values for the gas transfer velocity derived from theoretical modeling, also changing with temperature and turbulence (Demars and Manson 2013). Many studies have used the same constant to correct for ER and GPP estimates, but recent studies have suggested higher activation energy for ER than GPP in the ocean (López-Urrutia et al. 2006), experimental ponds (Yvon-Durocher et al. 2010), and streams (Demars et al. 2011b). The search for the temperature dependence of GPP and ER is in fact an active field of research.

Three approaches have been explored. One is to review the literature on annual metabolic estimates across sites (steady state) or seasonal metabolic estimate within site at nonsteady state (DeNicola 1996; Morin et al. 1999; Yvon-Durocher et al. 2012). A major caveat is that other factors may be responsible for the correlation, such as organic matter availability (e.g., Sinsabaugh 1997; Acula et al. 2004), nutrient supply (Cross et al. 2015; Welter et al. 2015), hydrology (e.g., Uehlinger et al. 2003; Acula and Tockner 2010), geomorphology (Jankowski et al. 2014), or simply light availability (e.g., Kelly et al. 1983; Dodds et al. 2013; Griffiths et al. 2013). The use of mixed model analyses on large datasets is probably the best approach (Yvon-Durocher et al. 2012) although while it will be effective at removing random effects, systematic bias are likely. In general, all review studies were poorly constrained and unsurprisingly many studies (not cited in the literature reviews) fell far away from the prediction of the metabolic theory of ecology (e.g., Sinsabaugh 1997: $E_p = 1.56$ and $E_n = 1.53$ or $Q_{10} \approx 9$; Huryn et al. 2014), with even significant negative correlation (e.g., Acula et al. 2004; Marcarelli et al. 2010) or no real effects (e.g., Roberts et al. 2007; Griffiths et al. 2013). Similar, there was no correlation between temperature and GPP or ER measured during the summer at 62 sites selected across biomes and anthropogenic impacts (Bernot et al. 2010; see also Hoellein et al. 2013). Stream metabolism is difficult to

---

**Table 1.** Fraction $\alpha$ of autotrophic respiration ($R_a$) relative to GPP in aquatic ecosystems (note that $1 - \alpha = \epsilon$ carbon use efficiency). MTE, metabolic theory of ecology.

<table>
<thead>
<tr>
<th>Reference</th>
<th>$\alpha$ = $R_a$/GPP</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duarte and Cebrian (1996)</td>
<td>0.35</td>
<td>Literature review of marine ecosystems</td>
</tr>
<tr>
<td></td>
<td>0.26</td>
<td>Phytoplankton</td>
</tr>
<tr>
<td></td>
<td>0.52</td>
<td>Microphytobenthos</td>
</tr>
<tr>
<td></td>
<td>0.57</td>
<td>Macroalgae</td>
</tr>
<tr>
<td>McIntire et al. (1996)</td>
<td>0.14/0.42</td>
<td>Derived from stream modeling, with/without grazing</td>
</tr>
<tr>
<td>Webster and Meyer (1997)</td>
<td>0.5</td>
<td>Arbitrary, stream ecosystems across biomes</td>
</tr>
<tr>
<td>Young and Huryn (1999)</td>
<td>0.2</td>
<td>Arbitrary, to reflect partially grazed stream</td>
</tr>
<tr>
<td>Allen et al. (2005)</td>
<td>0.6</td>
<td>Taken from literature, fixed value used in subsequent aquatic ecosystem MTE papers</td>
</tr>
<tr>
<td>Thomas et al. (2005)</td>
<td>0.39</td>
<td>Estimated average weighted mean daily R/P of the various stream autotrophs</td>
</tr>
<tr>
<td>Nicholson et al. (2012)</td>
<td>0.33–0.71</td>
<td>Field estimation using modeling and triple oxygen isotope/14C method, marine phytoplankton</td>
</tr>
<tr>
<td>Hall and Beaulieu (2013)</td>
<td>0.34–0.69</td>
<td>Range derived from quantile regression of GPP and ER of stream data (excluding values derived from Izagirre et al. 2008)</td>
</tr>
<tr>
<td>Solomon et al. (2013)</td>
<td>0.62 (0.02–5.95)</td>
<td>Median (range) derived from regression of temperature corrected GPP and ER of lake data (excluding lakes with $r^2 &lt; 0.4$)</td>
</tr>
</tbody>
</table>
measure year round due to lateral inflows which can prevent reliable estimation of ER (O’Gorman et al. 2012). Poorly constrained studies with no assessment of lateral inflows should not be used (e.g. Izagirre et al. 2008; Birkel et al. 2013).

The second approach is to use a comparative survey where the sites differ widely in their temperature everything else being similar and when the streams can be assumed at steady state under stable flow conditions when standing biomass and nutrient concentrations remain stable over time (Brookshire et al. 2009; Demars et al. 2011b). Temperature was tentatively shown to alter the metabolic balance of stream due to the higher activation energy for ER ($E_a = 0.67$) than GPP ($E_p = 0.54$). The temperature dependence of GPP was however not significant after taking into account the effect of transient storage (Demars et al. 2011a) and was not present at the end of the winter when the streams were clearly not under steady state (O’Gorman et al. 2012)—a reminder that other factors do affect stream metabolism across seasons (see above).

The results from Demars et al. (2011b) were confirmed for respiration, however, using an experimental approach (Perrins et al. 2012), the third approach. Other recent experiments in stream ecosystems have also investigated the respiratory response of suspended particles (Sand-Jensen et al. 2007) and periphyton (Acuña et al. 2008). Very few experiments have investigated the simultaneous temperature dependence of GPP and ER in channels (e.g., Beyers 1962; Kevern and Ball 1965; Phinney and McIntire 1965) and pond mesocosms (Moss 2010; Yvon-Durocher et al. 2010; Liboriussen et al. 2011). Surprisingly, the temperature treatment in the pond studies resulted in completely different metabolic responses, the reasons of which are unclear. Microcosm and mesocosm experiments are also not always easily transferred to natural ecosystems due to scaling issues (Petersen et al. 2009; Hanson et al. 2011).

As it stands, the temperature dependence of aquatic ecosystem metabolism remains poorly constrained (Le Cren and Lowe-McConnell 1980; Demars et al. 2011b; Yvon-Durocher et al. 2012). Moreover, the underlying mechanisms of the MTE (see Gillooly et al. 2001; Allen et al. 2005) are questionable, notably the intrinsic (cellular) activation energy of photosynthesis derived for C3 plants ($E_p \approx 0.32$) is inadequate for aquatic photosynthesizers with carbon concentrating mechanisms suppressing photorespiration (Williams and del Giorgio 2005; Raven et al. 2012). The activation energy of RUBISCO carboxylase activity is higher than C3 plants in algae ($E_p = 0.59 \pm 0.13$; Raven and Geider 1988). The apparent activation energy at the ecosystem level will also result from the balance of other dominant factors.

**Comments and recommendations**

In light of the present review, it is recommended for students (or novices) to play with the very simple single station modeling tool provided in Supporting Information to better grasp some of the fundamental principles, notably the sensitivity of key parameters. Extreme care in calibration of oxygen sensors is required in turbulent streams. The determination of the gas exchange coefficient (or gas transfer velocity) from the diel curve or other empirical and theoretical equations will carry high uncertainties (at best 40–125%) affecting the metabolic estimates. These high uncertainties become inevitable at global or continental scale modeling (Butman and Raymond 2011), but in detailed studies one should expect direct measurements of the gas exchange coefficient with appropriate tracers and field technique (e.g., Kilpatrick et al. 1989) or in large (deep) rivers with low reaeration rate adequate modeling technique to extract the gas exchange coefficient directly from the oxygen curves (e.g., Holtgrieve et al. 2010; Hotchkiss and Hall 2014). Tracer and oxygen gas exchange at the air-water interface can be simply related by the ratio of their molecular diffusivity, without the Schmidt number dependence. The associated cost of running the tracer studies may not be higher than buying the oxygen sensors. Gas samples can generally be kept for a month (Demars et al. 2011b) allowing transport by mail to adequate facilities for analysis (abroad if necessary). The cost of using very inaccurate data must also be considered. The health and safety concerns (e.g., Birkel et al. 2013) amount generally to little more than daily use of gas cylinders for barbecues or kitchen hobs. In large streams, the floating dome may be used with appropriate corrections but it would be worth trying out more widely the use of silicon or Teflon tubing to diffuse the tracer gas more effectively in the water column than with the traditional ceramic diffusers. It also becomes apparent that it is crucial to propagate uncertainties in all the calculations. There is no excuse not to with available methods and software.

It is wrong to think that the traditional two station approach can deal with spatial heterogeneity, including in the zone of influence upstream of the top station. As it stands, the most accurate method for heterogeneous streams is Demars et al. (2011b) and we provide an Excel workbook with example data as well as R script to run the method. The method by Demars et al. suffers from a lack of mathematical derivation, but this is compensated by recognizing explicitly critical assumptions not taken into account in other methods, including Reichert et al. (2009). The averaging of the two (multiple) oxygen curves could be used as inputs to more sophisticated modeling methods.

More work on new methods should be pursued such as including transient storage (Haggerty et al. 2009), the use of stable isotopes to better constrain parameters (Holtgrieve et al. 2010), and the estimation of respiration during daylight hours (Argerich et al. 2011; Hotchkiss and Hall 2014). New methods to quantify autotrophic respiration have been presented based on a large number of GPP and ER estimates (>50) for either a given system or across sites (notably Hall
and Beaulieu 2013; Solomon et al. 2013). Heterotrophic respiration (necessary for C cycling studies) and NPP (the base of the food chain) may be calculated more readily. However, none of the current methods are generally applicable, and the good work must continue.

While it is appreciated that temperature corrected gas exchange coefficient and metabolic estimates have been better related to other drivers, the intrinsic (cellular) temperature dependence of ER and GPP (Q10 ≈ 2 or E ≈ 0.6 in aquatic ecosystems) is affected by many other dominant processes at the ecosystem level and the apparent temperature dependence (or activation energy) can vary considerably from one stream to another. Therefore, ER and GPP should not be systematically temperature corrected to compare streams.

References

Brookshire, E. N., H. M. Valett, and S. Gerber. 2009. Mainte-
Busenberg, E., and L. N. Plummer. 2010. A rapid method for the measurement of sulfur hexafluoride (SF6), trifluoro-
methyl sulfur pentafluoride (SF5CF3), and Halon 1211 (CF2ClBr) in hydrologic tracer studies. Geochem. Geo-


Morse, N., W. B. Bowden, A. Hackman, C. Pruden, E. Steiner, and E. Berger. 2007. Using sound pressure to


Acknowledgments
The authors thank Marc Stutter and an anonymous referee for their comments on the manuscript. Bob Hall merit a special thank for his detailed and engaging review. JRM acknowledges the support of the Richard Stockton College of New Jersey. This article was supported by the Scottish Government Rural and Environment Science and Analytical Services (RESAS) and additional funding to work under the United Kingdom Natural Environment Research Council (NERC) Coupled Macronutrient Cycling Program. JT was funded by NERC grant NE/J011967/1, The Multi-Scale Response of Water Quality, Biodiversity and C Sequestration to Coupled Macronutrient Cycling from Source to Sea.

Submitted 26 November 2014
Revised 22 March 2015
Accepted 6 April 2015

Associate editor: Paul Kemp