Carbon Flux across Scales in a Changing Climate

by

Chao Song

(Under the Direction of Ford Ballantyne IV)

Abstract

Ecological patterns are scale dependent. Understanding how and why ecological patterns vary across scales is a central problem in ecology. Stream metabolism and soil respiration, two important processes in the global carbon cycle, are particularly scale dependent. In this dissertation, I employed a dynamic modeling approach to address multiple aspects related to the issue of scale in stream metabolism and soil respiration. Specifically, in chapter 2, I used a dynamic model of dissolved oxygen to quantify the temperature sensitivity of whole-stream metabolism in streams from six biomes, ranging from the tropics to the Arctic. I found that warming leads to convergence in stream metabolic balance, realized as reduced inter-site variability of GPP/ER. The GPP/ER ratio in streams with higher temperature and higher current GPP/ER is predicted to decrease in response to warming, whereas in streams with lower temperature and lower current GPP/ER it is expected to increase, although by a smaller magnitude. In chapter 3, I compared reach-scale metabolism quantified using open channel method and habitat-scale metabolism quantified using chamber incubations. I found that the reach-to-habitat ratio of GPP and ER, standardized to the same light and temperature conditions, decreased with the variance of habitat-scale metabolism within a reach. By combining theoretical analyses and numeric simulations, I showed that the heterogeneity of habitat-scale metabolism within a reach, the negative correlations between light and GPP

per light, and temperature used for habitat-scale incubations, could explain this pattern of mismatch between reach and habitat scale metabolism. In chapter 4, I demonstrated the importance of recognizing soil respiration as an aggregated process. I showed that aggregating over space influenced temperature sensitivity, but aggregation over time did no alter temperature sensitivity. I also demonstrated that recognizing soil respiration as the sum of contributions from distinct substrate pools could explain several often observed relationships between temperature sensitivity and temperature, and influenced interpretations of the mechanisms driving changes in temperature sensitivity of soil respiration. Collectively, these studies demonstrated scale dependency of soil respiration and stream metabolism, and highlighted the utility of dynamic modeling as a central approach to tackling the issue of scale.

INDEX WORDS: Scale, Stream metabolism, Gross primary production, Ecosystem respiration, Temperature sensitivity, Soil respiration, Global warming

Carbon Flux across Scales in a Changing Climate

by

Chao Song

B.S., Peking University, China, 2009

M.A., University of Kansas, 2011

M.S., University of Georgia, 2016

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2018

Carbon Flux across Scales in a Changing Climate

by

Chao Song

Major Professor:	Ford Ballantyne IV
Committee:	Sharon A. Billings
	John M. Drake
	Amy D. Rosemond

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2018

Acknowledgments

I express my deepest appreciation to my advisor Ford Ballantyne. He is a truly great advisor. He continually guided me to become a thoughtful scientist, and convincingly conveyed a spirit of excellent scholarship. This dissertation would not be possible without his guidance. I thank my committee members, Sharon Billings, John Drake, and Amy Rosemond, for their insightful advice in the dissertation research. I thank many collaborators, including Walter Dodds, Janine Rüegg, Alba Argerich, Christina Baker, Breck Bowden, Michael Douglas, Kait Farrell, Michael Flinn, Keith Gido, Erica Garcia, Ashley Helton, Tamara Harms, Shufang Jia, Jay Jones, Lauren Koenig, John Kominoski, Bill McDowell, Damien McMaster, Sam Parker, Claire Ruffing, Ken Sheehan, Matt Trentman, Matt Whiles, and Wil Wollheim, for their contributions. I benefited immensely from insightful discussions with colleagues from Odum School of Ecology and other departments at the University of Georgia, including Liuyang Zhang, Craig Osenberg, Jeff Minucci, Ashton Griffin, Carly Phillips, Reni Kaul, John Vinson, Nicole Lazar, Yanyan Tan, and Rui Xie. This dissertation research also received support from many staff members at Odum School of Ecology, including Katherine Adams, Brian Perkins, Tyler Ingram, Emily Schattler, and Shialoh Wilson. Part of the dissertation research is supported by the United States National Science Foundation (EF–1258994). Georgia Advanced Computing Resource Center provided computing facilities.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Ecological patterns are scale dependent. Understanding how and why ecological patterns differ across scales is often viewed as a central problem in ecology (Levin, 1992; Chave, 2013). which necessitates the consideration of scale in ecological studies. However, ecological studies may not always explicitly consider scale in the design and interpretation (Sandel & Smith, 2009). A lack of consideration of scale may lead to the incorrect application of ecological findings, and hinders our understanding of ecological problems of interest (Schindler, 1998). Although it is ideal to design studies directly on the scale of interest, this is sometimes infeasible in practice. For example, while many ecological problems of interest emerge on large scales, understanding these problems often requires manipulative experiments, which can only be done on a much smaller scale. Two promising general directions emerge to explicitly tackle the problem of scale. On the one hand, we may employ novel approaches to investigate ecological problems on the scales at which they are intended to be applied to. On the other hand, we may establish a link between findings from studies on different scales and understand how to translate ecological processes across scales.

Variability plays a central role in both approaches addressing the issue of scale. First, the natural variability in abiotic environments and biotic processes may serve as natural manipulative experiments on a relatively large scale where manipulative experiments are not feasible. Leveraging this variability allows us to make inference about ecological patterns of interest on large scales. For example, eddy flux data and daily temperature swings can be used to infer the temperature sensitivity of respiration at the whole ecosystem level (Mahecha *et al.*, 2010). Second, the existence of variability creates mismatches in quantities identified

on different spatial scales. The mismatch is not only an issue of representativeness of small-scale studies to problems on a large scale, but also can result from an inherent mismatch created by the interaction between nonlinearity and heterogeneity (Chesson, 2012; Melbourne & Chesson, 2006). We need to address both sources of mismatch to translate ecological processes across scales.

Dynamic modeling is the key method when addressing the issue of scale (Denny & Benedetti-Cecchi, 2012). To directly make inference on large scales, it is necessary to quantitively describe the dynamics of measurable quantities and subsequently make inference on the quantities not directly measurable. For example, formulating the dynamics of carbon flux is key to estimate temperature sensitivity of respiration at the ecosystem level with eddy flux data (Mahecha *et al.*, 2010). To establish the link between quantities identified on different scales, it is crucial to mechanistically characterize the dynamics of interests for two reasons. First, it allows us to correct differences in experimental conditions and artifacts between scales to resolve the representativeness issue. For example, measuring soil carbon efflux from respiration over a large spatial extent using eddy flux technique often requires a daily time span (Mahecha *et al.*, 2010). Measuring respiration over the same spatial extent using an *in situ* chamber requires measurements at multiple locations within the spatial extent, each taking less than an hour. As a result, the temperature conditions for the two measurements could be vastly different. While the large-scale measurement experiences daily temperature swings, each chamber measurement is made only under the temperature at the measurement time. If we want to compare the two respiration measurements under the same temperature condition, we need to standardize both measurements to the same temperature condition. Such standardization requires a mechanistic description of how temperature influences respiration. Second, the mechanistic description of dynamics enables us to explicitly incorporate the interaction between nonlinearity and heterogeneity to link measurements made on small scales to large scales. As Chesson (2012) pointed out in scale transition theory, incorporating the effects of spatial or temporal heterogeneity in linking measurements on different

scales requires an explicit formulation of how the spatial or temporal heterogeneous variables influence the response variables of interest.

In this dissertation, I used dynamic modeling approaches to address the issue of scale in ecosystem carbon flux, using stream metabolism and soil respiration as examples. Carbon flux in soil and stream ecosystems are both ecologically important in the global carbon cycle. While it is well appreciated that soil flux is a major contributor to the climate-carbon feedback (Schlesinger & Andrews, 2000; Cox *et al.*, 2000), the role of stream metabolism has also been increasingly recognized in recent years (Battin *et al.*, 2008, 2009; Hotchkiss *et al.*, 2015). In addition, the issue of scale is an important factor to consider in both systems. The dendritic structure of stream networks defines several spatial scales naturally: watershed, reach ($50 - 500$ m) and microhabitat (0.1 m)(Lowe *et al.*, 2006), that correspond to the common spatial scale of experiments and management practice. High spatial heterogeneity in soils means that measurements are unlikely to be directly transferrable across scales. Due to the practical usefulness and the natural emergence of specific scales of interests, soil respiration, and stream metabolism provide two ideal systems to explore the effects of scale.

This dissertation utilized data sets from the collaborative project "Scale, Consumer, and Lotic Ecosystem Rates". The data sets contain whole stream metabolism, chamber metabolism, and synoptic sampling of multiple streams along discharge gradients in six biomes. The data sets spanning multiple biomes provided a unique opportunity to examine multiple aspects of scale in ecosystem carbon flux in diverse ecological settings and contribute to a general understanding of the scale dependence of ecosystem carbon flux. Specifically, in the first chapter, I used a dynamic model of dissolved oxygen to estimate the temperature sensitivity of whole-stream metabolism, which is often difficult to obtain from manipulative experiments. The mechanistic characterization allowed me to estimate stream-specific temperature sensitivities of gross primary production and ecosystem respiration, predict the impact of warming on stream metabolic balance, and quantify the changes in carbon budget due to streams metabolism globally in a warming world. In the second chapter, I explicitly

linked stream metabolism quantified on different spatial scales. I demonstrated how mismatch in metabolism rate might arise as a result of the interaction between physiology and the environmental heterogeneity by dynamically modeling the dissolved oxygen concentration and estimating metabolism rate on both scales. In the third chapter, I demonstrated the importance of recognizing soil respiration as an aggregated process when estimating the temperature sensitivity of soil respiration. Soil respiration is inherently an aggregated process: respiration on a large spatial scale is the aggregation of smaller patches within the spatial extent; respiration on a longer time scale is the aggregation of respiration over all shorter time intervals; total respiration is the result of respiration from all substrate pools. I used simulation experiments and analyses of existing data sets to demonstrate how temperature sensitivity of soil respiration may depend on the spatial and temporal scale of consideration and thus provide a practical guide on how measurements made on one temporal and spatial scale can be translated to another scale. I also showed how viewing soil respiration as an aggregated process may influence our interpretation of the mechanisms driving the observed changes in soil carbon fluxes.

References

Battin, T.J., Kaplan, L.A., Findlay, S., Hopkinson, C.S., Marti, E., Packman, A.I., Newbold, J.D. & Sabater, F. (2008). Biophysical controls on organic carbon fluxes in fluvial networks. *Nature Geoscience*, 1, 95.

Battin, T.J., Luyssaert, S., Kaplan, L.A., Aufdenkampe, A.K., Richter, A. & Tranvik, L.J. (2009). The boundless carbon cycle. *Nature Geoscience*, 2, 598.

Chave, J. (2013). The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecology Letters*, 16, 4–16.

Chesson, P. (2012). Scale transition theory: its aims, motivations and predictions. *Ecological Complexity*, 10, 52–68.

Cox, P.M., Betts, R.A., Jones, C.D., Spall, S.A. & Totterdell, I.J. (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, 408, 184–187.

Denny, M. & Benedetti-Cecchi, L. (2012). Scaling up in ecology: mechanistic approaches. *Annual Review of Ecology, Evolution, and Systematics*, 43, 1–22.

Hotchkiss, E., Hall, R., Sponseller, R., Butman, D., Klaminder, J., Laudon, H., Rosvall, M. & Karlsson, J. (2015). Sources of and processes controlling CO_2 emissions change with the size of streams and rivers. *Nature Geoscience*, 8, 696.

Levin, S.A. (1992). The problem of pattern and scale in ecology: the robert h. macarthur award lecture. *Ecology*, 73, 1943–1967.

Lowe, W.H., Likens, G.E. & Power, M.E. (2006). Linking scales in stream ecology. *BioScience*, 56, 591–597.

Mahecha, M.D., Reichstein, M., Carvalhais, N., Lasslop, G., Lange, H., Seneviratne, S.I., Vargas, R., Ammann, C., Arain, M.A., Cescatti, A. *et al.* (2010). Global convergence in the temperature sensitivity of respiration at ecosystem level. *Science*, 329, 838–840.

Melbourne, B.A. & Chesson, P. (2006). The scale transition: scaling up population dynamics with field data. *Ecology*, 87, 1478–1488.

Sandel, B. & Smith, A.B. (2009). Scale as a lurking factor: incorporating scale-dependence in experimental ecology. *Oikos*, 118, 1284–1291.

Schindler, D.W. (1998). Whole-ecosystem experiments: replication versus realism: the need for ecosystem-scale experiments. *Ecosystems*, 1, 323–334.

Schlesinger, W.H. & Andrews, J.A. (2000). Soil respiration and the global carbon cycle. *Biogeochemistry*, 48, 7–20.

Chapter 2

Warming Induces Asymmetric Convergence of Stream Metabolic Balance

Streams play a significant role in the transport, storage, and transformation of organic carbon globally (Battin *et al.*, 2009; Butman *et al.*, 2016). Recent estimates suggest that 0.8–1.8 petagrams (Pg) of carbon evade from streams and rivers to the atmosphere annually (Cole *et al.*, 2007; Raymond *et al.*, 2013). This is comparable in size to the net annual terrestrial–atmosphere and net ocean–atmosphere carbon exchange (Ciais *et al.*, 2013). Stream metabolism, which is governed by gross primary production (GPP) and ecosystem respiration (ER), contributes substantially to the overall carbon flux out of streams. A recent study estimated that stream metabolism is responsible for up to 28% of the total carbon flux from streams to the atmosphere (Hotchkiss *et al.*, 2015), resulting in an estimated net flux of 0.12 Pg C per year (Battin *et al.*, 2008). As GPP and ER are both temperature dependent processes, sustained climate warming has the potential to profoundly alter the rates of carbon flux in and out of streams. Over the past century, mean water temperature in US rivers and streams increased at a rate of 0.009–0.077 °C per year (Kaushal *et al.*, 2010), and stream temperatures are predicted to increase by 1–3 °C with the doubling of atmospheric CO_2 concentration (Mohseni *et al.*, 1999). Consequently, understanding the feedback between stream metabolism and global warming is crucial when considering global or regional carbon cycles.

Although it is tempting to use well quantified temperature responses of photosynthesis and respiration at the cellular level to predict ecosystem level responses to warming, complex interactions among organisms and their abiotic environments can confound the temperature

responses of cellular processes at higher levels of organization. Taken at face value, the differential temperature sensitivities of photosynthesis and respiration at the cellular level defined by activation energy in the Arrhenius equations ($\approx$ 30.9 and 62.7 KJ mol^{-1} for photosynthesis and respiration respectively (Allen *et al.*, 2005)) prescribes a relatively faster increase in ER than GPP in response to warming. Consequently, we would predict that streams will become more heterotrophic (i.e. lower GPP/ER) as climate continues to warm. However, the implicit assumption of such a prediction, that the activation energies of photosynthesis and respiration at the cellular level are appropriate for describing the temperature sensitivities of GPP and ER in streams at the ecosystem level, may not hold.

Intrinsic variation in the temperature dependence of multiple processes that comprise aggregated ecosystem rates can cause the temperature sensitivities of whole ecosystem processes to deviate from the temperature dependence of cellular level responses. For example, variation in algal community composition can influence the temperature sensitivity of ecosystem level GPP because the activation energy of photosynthesis varies across phyla of algae (Galmes *et al.*, 2015; Chen & Laws, 2017). Similarly, the chemical structure of organic compounds influences the activation energy of decomposition reactions, and thus, variation in respiratory substrate composition can affect the temperature sensitivity of ER (Follstad Shah *et al.*, 2017). Alternatively, if ecosystem level GPP and ER are influenced by other temperature dependent processes, inferred temperature sensitivities of GPP and ER may reflect the influences of these processes and not necessarily the temperature sensitivities of cellular photosynthesis and respiration. For example, warming may accelerate the flux of nutrients and organic carbon from sediments to the water column (Duan & Kaushal, 2013) and transport of nutrients across cell membranes (Raven & Geider, 1988), both of which could result in amplified temperature sensitivities at the ecosystem level (Anderson-Teixeira *et al.*, 2008). Conversely, the temperature sensitivities of GPP or ER at the ecosystem level can be muted by nutrient limitation (Sand-Jensen *et al.*, 2007; López-Urrutia & Morán, 2007) or reflect the temperature sensitivity of a process that constrains GPP or ER, such

as nitrogen supply (Welter *et al.*, 2015). Finally, variation in the responses of different taxa to temperature variation can confound aggregate temperature sensitivity. For example, differential responses to warming across decomposer taxa have even been shown to cancel each other out, resulting in no net change in ecosystem carbon flux in response to warming (Boyero *et al.*, 2011).

In addition to the inherent complexity in ecosystem level temperature sensitivities of GPP and ER, the varied approaches employed to quantify them also have the potential to influence the inferred ecosystem level temperature dependence of GPP and ER. Incubations of stream substrata at different temperatures (Acuna *et al.*, 2008; Jankowski *et al.*, 2014) or mesocosm warming experiments(Yvon-Durocher *et al.*, 2010) do not include the entire focal ecosystem and may not encompass the processes key to determining the temperature sensitivities of GPP and ER at the ecosystem level. Comparisons among streams or within one stream over seasons (Sinsabaugh, 1997; Yvon-Durocher *et al.*, 2012; Huryn *et al.*, 2014; Demars *et al.*, 2011b; Perkins *et al.*, 2012; Welter *et al.*, 2015) yield ecosystem level estimates of temperature sensitivities, but temperature independent differences among streams or seasons due to hydrology (Demars *et al.*, 2011a), geomorphology (Jankowski *et al.*, 2014), nutrient availability (Cross *et al.*, 2015; Williamson *et al.*, 2016), and light availability (Huryn *et al.*, 2014) can easily confound the responses of GPP and ER to temperature. These confounding factors render the estimated temperature dependence not purely a response to temperature, but an integrated response to the suite of temperature dependent and independent differences across streams or seasons

Given the complexity of ecosystem level temperature sensitivities and the challenges associated with quantifying them, it is not surprising that various patterns have been reported. Some studies have found consistent temperature sensitivities of ER at the ecosystem and the cellular levels (Acuna *et al.*, 2008; Yvon-Durocher *et al.*, 2010; Demars *et al.*, 2011b; Perkins *et al.*, 2012), but others have demonstrated considerable deviation of ecosystem level activation energies of GPP (Demars *et al.*, 2011b; Yvon-Durocher *et al.*, 2010; Demars

et al., 2016) and ER (Yvon-Durocher *et al.*, 2012; Welter *et al.*, 2015) from the values of their cellular analogs. In studies that simultaneously examined the temperature dependence of GPP and ER in streams, a shift toward heterotrophy with warming has been observed in some instances (Yvon-Durocher *et al.*, 2010; Demars *et al.*, 2011b), but a recent synthesis based on geothermal streams concluded that warming increased GPP and ER to the same extent and resulted in no net change in metabolic balance (Demars *et al.*, 2016). To date, simultaneous quantification of the temperature dependence of GPP and ER have been constrained in mesocosm incubations or in geothermal streams. There is still considerable uncertainty about whether streams will become more heterotrophic (decreasing GPP/ER) or more autotrophic (increasing GPP/ER) in response to continued warming. Simultaneously quantifying the ecosystem level temperature sensitivities of GPP and ER in streams across broad bio-climatic regions is key to resolve such uncertainty.

When quantifying the ecosystem level temperature sensitivity of whole stream metabolism, every effort should be made to include the entire focal ecosystem and exclude the effects of temperature independent confounding factors. Estimating temperature dependence based on responses to short term temperature variation minimizes confounding effects, given that most confounding processes operate on a much longer time scale (Mahecha *et al.*, 2010). For whole stream metabolism, utilizing the response of dissolved oxygen (DO) concentration to diel temperature variation circumvents confounding factors to a large extent, and enables the temperature dependence of whole stream GPP and ER to be inferred (Holtgrieve *et al.*, 2016; Schindler *et al.*, 2017). By dynamically modeling GPP and ER, which drive diel DO dynamics, as explicit functions of temperature, and fitting the modeled DO trajectories to observed diel DO dynamics, the activation energies of GPP and ER can be estimated. This approach yields ecosystem level estimates of temperature dependence because DO dynamics integrate GPP and ER from all contributing components of the focal stream ecosystem. More importantly, using diel temperature variation to estimate temperature dependence of GPP and ER excludes temperature independent confounding factors that remain relatively

constant on the daily time scale. Combining this modeling approach with high resolution time series of light, temperature, and DO in streams across six biomes from the tropics to the arctic allows us to quantify the temperature dependence of stream metabolism across latitude and refine predictions of the feedback between stream metabolic balance and global warming.

Estimating activation energies of GPP and ER

We estimated the ecosystem level activation energies of GPP and ER in streams across six biomes by modeling diel changes in DO concentration. The six distinct biomes that span a wide range of latitude (13°S – 68 °N) include tropical forest (Luquillo Experimental Forest, Puerto Rico (LUQ)), tropical savanna (Litchfield National Park, North Territory, Australia (AUS)), tallgrass prairie (Konza Prairie, Kansas, USA (KNZ)), temperate rainforest (Andrews Experimental Forest, Oregon, USA (AND)), boreal forest (Caribou-Poker Creeks Research Watershed, Alaska, USA (CPC)), and arctic tundra (Toolik Lake Field Station, Alaska, USA (ARC)). In each biome, we measured DO concentration, photosynthetically active radiation, and water temperature at a 5 or 10 minute interval for 1–2 weeks in multiple stream reaches throughout a watershed. We modeled the response of DO concentration to diel temperature variation to estimate ecosystem level activation energies of GPP and ER. Specifically, we modeled the dynamics of DO concentration as:

$$\frac{d[\mathrm{O_2}]}{dt} = GPP - ER + K([\mathrm{O_2}]_{sat} - [\mathrm{O_2}]) \tag{2.1}$$

Here, $[\mathrm{O_2}]_{sat}$ is the saturated DO concentration and can be calculated from temperature and barometric pressure (American Public Health Association, 1995). GPP, ER, and K are instantaneous rates of primary production, respiration, and reaeration respectively. We modified previously published models of aquatic metabolism (Riley & Dodds, 2012; Jassby & Platt, 1976; Parkhill & Gulliver, 1999; Gulliver & Stefan, 1984; Elmore & West, 1961; Bott, 2006) by using the Arrhenius equation to describe the temperature dependence of GPP and

ER. Specifically, GPP, ER, and K were modeled as :

$$GPP = P_{max}\tanh(\frac{\alpha I}{P_{max}})e^{-\frac{Eap}{R}(\frac{1}{T}-\frac{1}{T_0})} \tag{2.2}$$

$$ER = R_{T_0}e^{-\frac{Ear}{R}(\frac{1}{T}-\frac{1}{T_0})} \tag{2.3}$$

$$K = K_{20} \times 1.024^{T-20} \tag{2.4}$$

Here, P_{max} (mg O_2 L^{-1} min^{-1}) is the maximum primary production rate, α (mg O_2 L^{-1} s m^{-2} μE^{-1} min^{-1}) is the slope of the light response curve of primary production at low light intensity, R_{T_0} (mg O_2 L^{-1} min^{-1}) is the respiration rate at reference temperature T_0 (Kelvin), which we set at the average daily water temperature across all days for each stream reach, K_{20} (min^{-1}) is the reaeration coefficient at 20 °C, I (μE m^{-2} s^{-1}) is photosynthetically active radiation, T (Kelvin) is water temperature, R (8.314 KJ mol^{-1} $Kelvin^{-1}$) is the ideal gas constant, E_{ap} (KJ mol^{-1}) and E_{ar} (KJ mol^{-1}) are the activation energies of GPP and ER respectively. We employed a Bayesian approach to estimate the parameters (P_{max}, α, R_{T_0}, K_{20}, E_{ap}, E_{ar}) in the model (Song *et al.*, 2016), and calculated daily GPP, ER, and GPP/ER using the estimated parameters and associated light and temperature profiles.

The estimated ecosystem level activation energies exhibited significant variability both within and across biomes (Fig. 2.1), and varied substantially from the activation energies of photosynthesis and respiration at the cellular level. Specifically, activation energies ranged from 0.5 to 839.2 KJ mol^{-1} for GPP and from 0.4 to 837.2 KJ mol^{-1} for ER. The median activation energies of GPP and ER were 68.2 KJ mol^{-1} and 67.5 KJ mol^{-1} respectively, which is consistent with a recent study quantifying the temperature sensitivity of GPP and ER in streams along a geothermal gradient (Demars *et al.*, 2016). However, this does not necessarily imply that warming will increase GPP and ER to the same extent. Due to the nonlinear nature of temperature dependence and substantial variability in the activation energies of GPP and ER, simply using the central tendency of the estimated activation

energies will not accurately describe the thermal response of stream metabolism within and across biomes. The inherent variation in activation energies underscores the importance of quantifying the thermal response of stream metabolism using activation energies of GPP and ER for individual streams rather than using the mean or median activation energies across all streams.

Activation energy of GPP/ER decreases with GPP/ER and temperature

The simultaneous quantification of the activation energies of GPP and ER allowed us to evaluate thermal response of stream metabolic balance across biomes. A common measure of metabolic balance in streams is the ratio of daily GPP to ER, which for our formulation of the instantaneous rates of GPP and ER is

$$\frac{GPP}{ER} = \frac{P_{max}\tanh(\frac{\alpha I}{P_{max}})}{R_{T_0}} e^{-\frac{E_{ap}-E_{ar}}{R}(\frac{1}{T}-\frac{1}{T_0})}. \tag{2.5}$$

The formulation of GPP/ER has the form of an Arrhenius equation, and thus, $E_{ap} - E_{ar}$ is the apparent activation energy of GPP/ER and determines how instantaneous metabolic balance changes with temperature. A positive $E_{ap} - E_{ar}$ means that GPP/ER will increase as temperature increases and a negative $E_{ap} - E_{ar}$ means GPP/ER will decrease as temperature increases.

Despite the significant variation in both E_{ap} and E_{ar} (Fig. 2.1), we observed that $E_{ap} - E_{ar}$ decreases significantly with daily GPP/ER (Fig. 2.2(a); linear mixed effects model, $F_{1,39.14} = 8.23$, $P = 0.0066$) and daily mean water temperature (Fig. 2.2(b); linear mixed effects model, $F_{1,44.28} = 8.4$, $P = 0.0058$). However, given that E_{ap} and E_{ar} were fitted parameters used to calculate daily GPP and ER, a relationship between $E_{ap} - E_{ar}$ and GPP/ER may be expected. To ensure that the observed relationship between $E_{ap} - E_{ar}$ and GPP/ER was not merely a statistical artifact, we performed simulations to establish the expected relationship between $E_{ap} - E_{ar}$ and GPP/ER under the assumption that parameters in equations 2.1–2.4 were chosen randomly and independently (see methods for details). Using sets of representative daily light and temperature measurements (Fig. 2.5), we observed a positive relationship

between $E_{ap} - E_{ar}$ and GPP/ER from simulations (Fig. 2.6). Consequently, the negative relationship between $E_{ap} - E_{ar}$ and GPP/ER (Fig. 2.2(a)) is not a statistical artifact, and rather an emergent property of stream ecosystems.

The negative correlation between $E_{ap} - E_{ar}$ and GPP/ER suggests a negative feedback between stream metabolic balance and its temperature dependence. Specifically, streams with a higher daily GPP/ER tended to have a lower, negative $E_{ap} - E_{ar}$ (Fig. 2.2(a)). Thus, we predict that streams with higher daily GPP/ER will shift towards greater heterotrophy when temperature increases. Conversely, streams with lower GPP/ER tended to have a higher, positive daily $E_{ap} - E_{ar}$, and are predicted to shift towards more autotrophy when temperature increases. If the observed pattern based on streams across the six biomes is generally applicable to stream ecosystems, we expect a convergence in stream metabolic balance, characterized by reduced inter-site variability of daily GPP/ER, in response to warming.

The significant decrease in $E_{ap} - E_{ar}$ with mean daily water temperature (Fig. 2.2(b)) also gives rise to a prediction for how GPP/ER will change in response to warming along the inter-biome temperature gradient. Specifically, streams experiencing higher mean daily temperature will shift toward greater heterotrophy (i.e. lower GPP/ER), and streams with lower temperature will shift toward greater autotrophy (i.e. higher GPP/ER) in response to temperature increases. Because temperature is correlated with latitude, the relationship between $E_{ap} - E_{ar}$ and mean daily temperature allows us to predict the influence of warming on stream metabolic balance across latitude. Since streams are mostly heterotrophic (Battin *et al.*, 2008), we expect streams in the tropics to be a relatively stronger carbon source and streams in the arctic to be a relatively weaker carbon source on average as temperature increases. However, given that daily GPP/ER also predicted the temperature sensitivity of stream metabolic balance (Fig. 2.2(a)), the exact latitudinal pattern of changes in stream metabolic balance will also depend on the effect size of GPP/ER and temperature, as well as the spatial distribution of daily GPP/ER at specific latitudes.

Warming induces asymmetric convergence in stream metabolic balance

A first step to assess warming induced changes in stream metabolic balance across biomes is to quantify how changes in stream metabolic balance, as described by GPP/ER, can be predicted by covariates that predict $E_{ap} - E_{ar}$. Since activation energy is proportional to the percentage change in reaction rate in an Arrhenius equation (Sierra, 2012), the fact that daily GPP/ER and temperature predict $E_{ap} - E_{ar}$ indicates that they also predict the percentage change in GPP/ER (ΔGPP/ER) as temperature increases. We performed a simulated warming experiment to calculate ΔGPP/ER, and established a relationship between ΔGPP/ER and predictors of $E_{ap} - E_{ar}$, namely daily GPP/ER and mean water temperature. Specifically, we added 1 °C to each recorded water temperature, which represents a 1 °C increase in mean daily water temperature while keeping daily temperature variability constant. The 1 °C increase in mean water temperature is a realistic estimate of stream temperature in the next century based on the current rate of warming in streams (Webb, 1996; Kaushal *et al.*, 2010). With light measurements, the elevated temperature trajectories, and parameters in the DO model (equation 2.1–2.4) estimated from field data, we calculated the daily GPP and ER under this warming scenario. The change in metabolic balance can then be calculated as

$$\Delta GPP/ER = \frac{GPP/ER_{warming} - GPP/ER_{current}}{GPP/ER_{current}} \tag{2.6}$$

where $GPP/ER_{current}$ and $GPP/ER_{warming}$ are daily GPP/ER currently and under 1 °C warming scenario respectively. We analyzed the effects of daily GPP/ER and mean water temperature on ΔGPP/ER in a linear mixed effects model. As expected, ΔGPP/ER had a significant negative relationship with both daily GPP/ER (Fig 2.3(a), $F_{1,39.29} = 12.50$, $P = 0.0011$) and temperature (Fig 2.3(b), $F_{1,42.41} = 7.60$, $P = 0.0086$). Quantitatively, ΔGPP/ER can be predicted based on the fixed effects in the linear mixed effect model as $\Delta\text{GPP/ER} = 0.46 - 0.45 \times \text{GPP/ER} - 0.019 \times \text{Temperature}$.

To establish how warming is likely to affect the metabolic balance in streams globally, we assembled a stream metabolism data set of daily GPP, ER, and mean water temperature based on two previous synthesis studies (Demars *et al.*, 2016; Hoellein *et al.*, 2013), and applied the linear model for ΔGPP/ER as a function of both GPP/ER and mean water temperature to the compiled data set. We selected data within the range of daily GPP/ER (0.016–0.978) and daily mean temperature (2.2–26.3 °C) found in our study for analyses, resulting in a total of 236 metabolism estimates (see supplementary materials). Using the predictive equation of ΔGPP/ER identified in the simulated warming experiment (i.e. $\Delta\text{GPP/ER} = 0.46 - 0.45 \times \text{GPP/ER} - 0.019 \times \text{Temperature}$), we quantified the ΔGPP/ER for each stream in the compiled data set and subsequently calculated the GPP/ER with the 1 °C increase in temperature. Two patterns of warming induced changes in stream metabolic balance emerged. First, the GPP/ER of streams converged under a 1 °C temperature increase, shown as a decrease in the inter-site variability of GPP/ER (Fig. 2.4(a)). The standard deviation of GPP/ER decreased from 0.26 currently to 0.20 with a 1 °C increase in temperature. Second, with a 1 °C increase in temperature, GPP/ER generally increased in streams with lower temperatures and decreased in streams with higher temperatures (Fig 2.4(b)). However, the direction and magnitude of changes in GPP/ER along the temperature gradient was modified by daily GPP/ER. The linear model for ΔGPP/ER establishes an isocline ($0.46 - 0.45 \times \text{GPP/ER} - 0.019 \times \text{Temperature} = 0$), along which daily GPP/ER is insensitive to changes in temperature. Thus the isocline represents the GPP/ER streams converge to in response to warming (Fig. 2.4(b)). The isocline separates the combinations of current daily GPP/ER and mean water temperature that give rise to an increase in GPP/ER from the combinations that give rise to a decrease in GPP/ER in response to warming. Notably, the magnitude of decrease in GPP/ER in streams with high temperatures and high daily GPP/ER was larger than the magnitude of increase in GPP/ER in streams with low temperatures and low daily GPP/ER. Such asymmetry suggests that warming will

influence the metabolic balance of streams with high temperature and daily GPP/ER more substantially, and change them into stronger carbon source (lower GPP/ER).

In conclusion, the temperature sensitivities of GPP and ER in streams quantified using a dynamic modeling approach is, to our knowledge, the first consistent ecosystem level quantification of the temperature sensitivities of stream metabolism across biomes. Based on the relationship between the temperature sensitivity of stream metabolic balance, daily GPP/ER, and the mean daily water temperature, we predict an asymmetric convergence in stream metabolic balance in response to warming, with a larger magnitude of decrease in GPP/ER in warmer, more autotrophic streams than the increase in GPP/ER in colder, more heterotrophic streams. Our findings suggest that the thermal response of stream metabolism cannot be simply extrapolated from the cellular analogs. Integrating the ecosystem level thermal response of stream metabolism is critical for an accurate projection of the feedback between stream metabolism and climatic warming.

Methods

Study sites and data collection

We conducted the study in six watersheds representing distinct biomes, including tropical forest (LUQ), tropical savanna (AUS), tallgrass prairie (KNZ), temperate rainforest (AND), boreal forest (CPC), and arctic tundra (ARC). Within each watershed, we selected 6–12 streams across a range of stream sizes to capture the physical gradients within the watershed. A detailed description of the study sites can be found in ref Rüegg *et al.* (2016). In each stream, we recorded DO concentration, water temperature, and barometric pressure using a YSI ProODO handheld optical DO meter (YSI Instruments, Yellow Springs, Ohio, USA), and photosynthetically active radiation using an Odyssey Irradiance logger (DataFlowSystems, Christchurch, New Zealand) at a single location in each stream. The DO meter was calibrated with water saturated air immediately before deployment. The readings from the irradiance logger were converted to photosynthetically active radiation by comparing to a calibrated

sensor. We recorded these data at an interval of 5 minutes (ARC) or 10 minutes (all other sites) for 1–14 days. We collected data during base flow periods (Feburary–March 2013 and March 2014 for LUQ, July–August 2013 for AUS, May–June 2013 and April–June 2014 for KNZ, July–August 2015 for AND, July–August 2013 and 2014 for CPC, July–August 2013 and 2014 for ARC). In total, we collected 709 daily DO trajectories from 69 stream reaches across the six biomes.

Estimating activation energies of GPP and ER

We modeled the dynamics of DO concentration with equations 2.1–2.4. We employed a Bayesian approach for parameter estimation (Song *et al.*, 2016). Specifically, for a given set of parameters, we used the Runge-Kuntta 4th order method implemented in R package `deSolve` (Soetaert *et al.*, 2010) with a step size of 2.5 minutes to numerically solve the differential equations describing DO dynamics (equations 2.1–2.4) and obtained a trajectory of modeled DO concentration. Numerically solving the differential equations with high accuracy requires the interpolation of discrete measurements of light and temperature. To this end, we used linear interpolation to approximate continuous trajectories of light and temperature from discrete measurements. We assumed that the differences between modeled and measured DO were independent and identically distributed normal random errors. Based on this assumption of error distribution, we computed the likelihood for any given set of parameters. We used uniform priors for all parameters in the model, setting the lower bound of the uniform priors at 0 and upper bound at values significantly larger than found in previous studies to ensure that the posterior inferences were not overly constrained by the prior distributions. In particular, we set the upper bound of the uniform prior for E_{ap} and E_{ar} at 1000 KJ mol^{-1}, which is significantly higher than found in existing literature (Acuna *et al.*, 2008; Yvon-Durocher *et al.*, 2010; Demars *et al.*, 2011b; Perkins *et al.*, 2012; Yvon-Durocher *et al.*, 2012; Jankowski *et al.*, 2014; Welter *et al.*, 2015). We used Markov Chain Monte Carlo to sample the posterior distributions of the parameters. Specifically, we

implemented the adaptive random walk Metropolis-Hasting algorithm (Haario *et al.*, 2001) with the function `metrop` in R package `mcmc` (Geyer & Johnson, 2014). We ran each Markov chain for half a million iterations and used a burn-in period of 300000 iterations to ensure stationarity. We performed visual inspection and Geweke diagnostic (Geweke, 1992) of the trace plots with R package `coda` (Plummer *et al.*, 2006) for proper mixing and convergence of the Markov chains. All parameters in the model (i.e. P_{max}, R_{T_0}, α, E_{ap}, E_{ar}, K_{20}) were simultaneously estimated. We used posterior means of the parameters for further statistical analyses.

We took two special considerations when estimating parameters. First, low diel variability in temperature in some streams prevented us from estimating E_{ap} and E_{ar} with confidence. Thus, we only used E_{ap} and E_{ar} estimates with 95% highest posterior density intervals narrower than 500 KJ mol^{-1} for further statistical analyses. This is to ensure that the estimated E_{ap} and E_{ar} are mainly determined by the data, not by the uniform priors. With this selection criteria, we obtained 292 estimated E_{ap} and E_{ar} from 48 reaches based on the 709 daily DO trajectories collected from 69 reaches. The choice of 500 KJ mol^{-1} as the threshold is arbitrary. Such an arbitrary choice influences the number of estimated E_{ap} and E_{ar} for further statistical analyses, but does not affect the findings of this study (Fig. 2.7). Second, when estimating parameters, we divided the data from the same stream into individual days, and estimated a unique set of parameters for each stream on each day after realizing the potential for significant day to day variation of the parameters for the same stream.

To obtain the posterior distributions of daily GPP and ER, we numerically integrated the instantaneous rates of GPP and ER over a day based on each iteration of parameters in the Markov Chain. We performed the same diagnostics of Markov chains to ensure stationarity, proper mixing, and convergence. We obtained the posterior distributions of GPP/ER by taking the ratio of the trace of daily GPP and ER. We reported the means of posterior

distributions as point estimates for daily GPP, ER, and GPP/ER. The estimated E_{ap}, E_{ar}, daily GPP, ER, and basic site information are included in the supplementary materials.

Expected relationship between $E_{ap} - E_{ar}$ and GPP/ER

A relationship between $E_{ap} - E_{ar}$ and daily GPP/ER may arise because E_{ap} and E_{ar} were used to calculate daily GPP and ER. To ensure that the observed relationship between $E_{ap} - E_{ar}$ and GPP/ER was not a statistical artifact, we performed simulations to establish the expected relationship between $E_{ap} - E_{ar}$ and daily GPP/ER when all parameters in equations 2.1–2.4 are independent. Specifically, we chose parameter values randomly and independently from uniform distributions. Without loss of generality, the ranges of the uniform distributions for all the parameters were 0.1–0.2 mg O_2 L^{-1} min^{-1} for R_{T_0}, 0.2–0.5 mg O_2 L^{-1} min^{-1} for P_{max}, 0.000035–0.000045 mg O_2 L^{-1} s m^{-2} μE^{-1} min^{-1} for α, 0.03–0.06 min^{-1} for K_{20}, 10–500 KJ mol^{-1} for E_{ap}, and 10–500 KJ mol^{-1} for E_{ar}. With each set of randomly and independently chosen parameters, we simulated a daily trajectory of DO concentration at a 10-minute interval using representative daily trajectories of light and temperature chosen from our field data (Fig. 2.5). We added normally distributed random observation errors to the simulated trajectory of DO. The standard deviation (σ) of the observation error was set ($\log(\sigma) = 2$) to be representative of the data. With each simulated DO concentration trajectory, we estimated GPP, ER, E_{ap}, and E_{ar} following the same methods outlined above. We repeated the simulation and estimation 1200 times. Given that all the parameters in the DO model (equations 2.1–2.4) were independent in the simulation study, any relationship between $E_{ap} - E_{ar}$ and daily GPP/ER based on the simulated data is purely driven by the fact that E_{ap} and E_{ar} were fitted parameters used to calculate daily GPP and ER. Comparing the expected relationship based on the simulations to the observed relationship allowed us to examine whether the findings based on the field data is an emergent ecosystem property or a statistical artifact.

Simulated warming experiment

With parameter estimates in the DO model (equations 2.1–2.4) for the 292 days of metabolism, we performed a simulated warming experiment to assess the response of stream metabolic balance to temperature increase. We added 1 °C to each individual measurement of water temperature. This warming scenario represents a 1 °C increase in daily mean temperature without changing the daily temperature variability. We calculated the daily GPP and ER under this warming scenario following the same procedure outlined above. We performed the same diagnostics of the trace plots of daily GPP and ER in the simulated warming experiment and excluded the estimates without proper convergence or mixing. In total, we successfully calculated 288 daily GPP and ER under the 1 °C warming scenario. The daily GPP and ER under the current temperature and the 1 °C warming scenario were used to calculate the proportional change in GPP/ER (ΔGPP/ER) as in equation 2.6.

Statistical analyses

We analyzed the pattern of $E_{ap} - E_{ar}$ as a function of current daily GPP/ER and daily mean temperature with a linear mixed effects model. Since we estimated a unique set of activation energies for each stream on each day, estimates of multiple days from the same stream could be correlated. Therefore, we included random effects of each stream nested in biome in the model to account for the repeated measurements. We treated the same streams measured in different years as different streams when specifying the random effects. Specifically, we started with a full model and performed backwards model selection to build the most parsimonious model. The fixed effects of the full model included daily GPP/ER, daily mean water temperature, and their interaction. The random effects of the full model included a random intercept and random slopes of both daily GPP/ER and mean water temperature for stream nested in biome. We first fit the full model using maximum likelihood and selected the adequate structure of random effects based on AIC. We found that eliminating the biome random

slopes and intercepts lead to a slight decrease in AIC (ΔAIC $= -0.86$), but eliminating the random intercept (ΔAIC $= 54.4$), random slope of daily GPP/ER (ΔAIC $= 22.7$), or random slope of daily mean water temperature (ΔAIC $= 10.8$) for each stream all resulted in substantial increases in AIC. Therefore, we specified the random effects with a random intercept and random slopes of GPP/ER and mean water temperature for each stream. We then refit the model with restricted maximum likelihood and used F-tests with Kenward-Roger approximation of degrees of freedom (Kenward & Roger, 1997) to select the fixed effects. We found no significant interaction between daily GPP/ER and mean water temperature ($F_{1,25.71} = 0.24$, $P = 0.63$). Thus, the most parsimonious model included daily GPP/ER and mean water temperature as fixed effects, and a random intercept and slopes of both daily GPP/ER and mean water temperature for each stream. We tested whether the fixed effects slopes of daily GPP/ER and mean water temperature were zero using F-test with Kenward-Roger approximation of degree of freedom to evaluate whether daily GPP/ER or mean water temperature had a significant effect on $E_{ap} - E_{ar}$.

Given that the percentage change in reaction rate is proportional to the activation energy in Arrhenius equation (Sierra, 2012), and that $E_{ap} - E_{ar}$ is the activation energy of GPP/ER (equation 2.5), it follows that predictors of $E_{ap} - E_{ar}$ should also be predictors of ΔGPP/ER. Therefore, we analyzed the effects of daily GPP/ER and mean water temperature on ΔGPP/ER using the same modeling structure as the most parsimonious model for $E_{ap} - E_{ar}$ without performing the model selection. We fit all the linear mixed effects models using function `lmer` in R package `lme4` (Bates *et al.*, 2015). F-test with Kenward-Roger approximation of degrees of freedom was implemented using R package `pbkrtest` (Halekoh & Højsgaard, 2014). All statistical analyses were performed in R 3.4.1 (R Core Team, 2017).

References

Acuna, V., Wolf, A., Uehlinger, U. & Tockner, K. (2008). Temperature dependence of stream benthic respiration in an alpine river network under global warming. *Freshwater Biology*, 53, 2076–2088.

Allen, A., Gillooly, J. & Brown, J. (2005). Linking the global carbon cycle to individual metabolism. *Functional Ecology*, 19, 202–213.

American Public Health Association (1995). *Standard methods for the examination of water and wastewater*. 19th edn. American Public Health Association, American Waterworkds Association and Water Environment Federation, Washington, DC.

Anderson-Teixeira, K.J., Vitousek, P.M. & Brown, J.H. (2008). Amplified temperature dependence in ecosystems developing on the lava flows of Mauna Loa, Hawai'i. *Proceedings of the National Academy of Sciences*, 105, 228–233.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.

Battin, T.J., Kaplan, L.A., Findlay, S., Hopkinson, C.S., Marti, E., Packman, A.I., Newbold, J.D. & Sabater, F. (2008). Biophysical controls on organic carbon fluxes in fluvial networks. *Nature Geoscience*, 1, 95–100.

Battin, T.J., Luyssaert, S., Kaplan, L.A., Aufdenkampe, A.K., Richter, A. & Tranvik, L.J. (2009). The boundless carbon cycle. *Nature Geoscience*, 2, 598–600.

Bott, T.L. (2006). Primary productivity and community respiration. In: *Methods in Stream Ecology* (eds. Hauer, F.R. & Lamberti, G.A.). Academic Press, San Diego, California, pp. 533–556.

Boyero, L., Pearson, R.G., Gessner, M.O., Barmuta, L.A., Ferreira, V., Graça, M.A., Dudgeon, D., Boulton, A.J., Callisto, M., Chauvet, E. *et al.* (2011). A global experiment

suggests climate warming will not accelerate litter decomposition in streams but might reduce carbon sequestration. *Ecology Letters*, 14, 289–294.

Butman, D., Stackpoole, S., Stets, E., McDonald, C.P., Clow, D.W. & Striegl, R.G. (2016). Aquatic carbon cycling in the conterminous United States and implications for terrestrial carbon accounting. *Proceedings of the National Academy of Sciences*, 113, 58–63.

Chen, B. & Laws, E.A. (2017). Is there a difference of temperature sensitivity between marine phytoplankton and heterotrophs? *Limnology and Oceanography*, 62, 806–817.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M. *et al.* (2013). Carbon and other biogeochemical cycles. In: *Climate Change 2013: the Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds. Stocker, T., Qin, D., Plattner, G.K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V. & Midgley, P.). Cambridge University Press, Cambridge, UK, pp. 465–570.

Cole, J.J., Prairie, Y.T., Caraco, N.F., McDowell, W.H., Tranvik, L.J., Striegl, R.G., Duarte, C.M., Kortelainen, P., Downing, J.A., Middelburg, J.J. *et al.* (2007). Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems*, 10, 172–185.

Cross, W.F., Hood, J.M., Benstead, J.P., Huryn, A.D. & Nelson, D. (2015). Interactions between temperature and nutrients across levels of ecological organization. *Global Change Biology*, 21, 1025–1040.

Demars, B., Manson, J., Olafsson, J., Gislason, G. & Friberg, N. (2011a). Stream hydraulics and temperature determine the metabolism of geothermal icelandic streams. *Knowledge and Management of Aquatic Ecosystems*, 402, 1–17.

Demars, B.O., Gíslason, G.M., Ólafsson, J.S., Manson, J.R., Friberg, N., Hood, J.M., Thompson, J.J. & Freitag, T.E. (2016). Impact of warming on CO_2 emissions from streams countered by aquatic photosynthesis. *Nature Geoscience*, 9, 758–761.

Demars, B.O., Russell Manson, J., Olafsson, J.S., Gislason, G.M., Gudmundsdottír, R., Woodward, G., Reiss, J., Pichler, D.E., Rasmussen, J.J. & Friberg, N. (2011b). Temperature and the metabolic balance of streams. *Freshwater Biology*, 56, 1106–1121.

Duan, S.W. & Kaushal, S. (2013). Warming increases carbon and nutrient fluxes from sediments in streams across land use. *Biogeosciences*, 10, 1193–1207.

Elmore, H. & West, W. (1961). Effect of water temperature on stream reaeration. *Journal of the Sanitary Engineering Division*, 87, 59–71.

Follstad Shah, J.J., Kominoski, J.S., Ardón, M., Dodds, W.K., Gessner, M.O., Griffiths, N.A., Hawkins, C.P., Johnson, S.L., Lecerf, A., LeRoy, C.J. *et al.* (2017). Global synthesis of the temperature sensitivity of leaf litter breakdown in streams and rivers. *Global change biology*.

Galmes, J., Kapralov, M., Copolovici, L., Hermida-Carrera, C. & Niinemets, Ü. (2015). Temperature responses of the Rubisco maximum carboxylase activity across domains of life: phylogenetic signals, trade-offs, and importance for carbon gain. *Photosynthesis research*, 123, 183–201.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: *Bayesian Statistics 4* (eds. Bernado, J., Berger, J., Dawid, A. & Smith, A.). Clarendon Press, Oxford, UK, pp. 169–193.

Geyer, C.J. & Johnson, L.T. (2014). *mcmc: Markov Chain Monte Carlo*. R package version 0.9–3.

Gulliver, J.S. & Stefan, H.G. (1984). Stream productivity analysis with dorm–I: Development of computational model. *Water Research*, 18, 1569–1576.

Haario, H., Saksman, E. & Tamminen, J. (2001). An adaptive metropolis algorithm. *Bernoulli*, pp. 223–242.

Halekoh, U. & Højsgaard, S. (2014). A kenward-roger approximation and parametric bootstrap methods for tests in linear mixed models – the R package pbkrtest. *Journal of Statistical Software*, 59, 1–30.

Hoellein, T.J., Bruesewitz, D.A. & Richardson, D.C. (2013). Revisiting odum (1956): a synthesis of aquatic ecosystem metabolism. *Limnology and Oceanography*, 58, 2089–2100.

Holtgrieve, G.W., Schindler, D.E. & Jankowski, K. (2016). Comment on demars et al. 2015, "stream metabolism and the open diel oxygen method: principles, practice, and perspectives". *Limnology and Oceanography: Methods*, 14, 110–113.

Hotchkiss, E., Hall Jr, R., Sponseller, R., Butman, D., Klaminder, J., Laudon, H., Rosvall, M. & Karlsson, J. (2015). Sources of and processes controlling CO_2 emissions change with the size of streams and rivers. *Nature Geoscience*, 8, 696–699.

Huryn, A.D., Benstead, J.P. & Parker, S.M. (2014). Seasonal changes in light availability modify the temperature dependence of ecosystem metabolism in an arctic stream. *Ecology*, 95, 2826–2839.

Jankowski, K., Schindler, D. & Lisi, P. (2014). Temperature sensitivity of community respiration rates in streams is associated with watershed geomorphic features. *Ecology*, 95, 2707–2714.

Jassby, A.D. & Platt, T. (1976). Mathematical formulation of the relationship between photosynthesis and light for phytoplankton. *Limnology and Oceanography*, 21, 540–547.

Kaushal, S.S., Likens, G.E., Jaworski, N.A., Pace, M.L., Sides, A.M., Seekell, D., Belt, K.T., Secor, D.H. & Wingate, R.L. (2010). Rising stream and river temperatures in the United States. *Frontiers in Ecology and the Environment*, 8, 461–466.

Kenward, M.G. & Roger, J.H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, pp. 983–997.

López-Urrutia, Á. & Morán, X.A.G. (2007). Resource limitation of bacterial production distorts the temperature dependence of oceanic carbon cycling. *Ecology*, 88, 817–822.

Mahecha, M.D., Reichstein, M., Carvalhais, N., Lasslop, G., Lange, H., Seneviratne, S.I., Vargas, R., Ammann, C., Arain, M.A., Cescatti, A. *et al.* (2010). Global convergence in the temperature sensitivity of respiration at ecosystem level. *Science*, 329, 838–840.

Mohseni, O., Erickson, T.R. & Stefan, H.G. (1999). Sensitivity of stream temperatures in the United States to air temperatures projected under a global warming scenario. *Water Resources Research*, 35, 3723–3733.

Parkhill, K.L. & Gulliver, J.S. (1999). Modeling the effect of light on whole-stream respiration. *Ecological Modelling*, 117, 333–342.

Perkins, D.M., Yvon-Durocher, G., Demars, B.O., Reiss, J., Pichler, D.E., Friberg, N., Trimmer, M. & Woodward, G. (2012). Consistent temperature dependence of respiration across ecosystems contrasting in thermal history. *Global Change Biology*, 18, 1300–1311.

Plummer, M., Best, N., Cowles, K. & Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. *R news*, 6, 7–11.

R Core Team (2017). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raven, J.A. & Geider, R.J. (1988). Temperature and algal growth. *New phytologist*, 110, 441–461.

Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C. *et al.* (2013). Global carbon dioxide emissions from inland waters. *Nature*, 503, 355–359.

Riley, A.J. & Dodds, W.K. (2012). Whole-stream metabolism: strategies for measuring and modeling diel trends of dissolved oxygen. *Freshwater Science*, 32, 56–69.

Rüegg, J., Dodds, W.K., Daniels, M.D., Sheehan, K.R., Baker, C.L., Bowden, W.B., Farrell, K.J., Flinn, M.B., Harms, T.K., Jones, J.B. *et al.* (2016). Baseflow physical characteristics differ at multiple spatial scales in stream networks across diverse biomes. *Landscape Ecology*, 31, 119–136.

Sand-Jensen, K., Pedersen, N.L. & Søndergaard, M. (2007). Bacterial metabolism in small temperate streams under contemporary and future climates. *Freshwater Biology*, 52, 2340–2353.

Schindler, D.E., Jankowski, K., A'mar, Z.T. & Holtgrieve, G.W. (2017). Two-stage metabolism inferred from diel oxygen dynamics in aquatic ecosystems. *Ecosphere*, 8.

Sierra, C.A. (2012). Temperature sensitivity of organic matter decomposition in the Arrhenius equation: some theoretical considerations. *Biogeochemistry*, 108, 1–15.

Sinsabaugh, R.L. (1997). Large-scale trends for stream benthic respiration. *Journal of the North American Benthological Society*, 16, 119–122.

Soetaert, K., Petzoldt, T. & Setzer, R.W. (2010). Solving differential equations in R: package desolve. *Journal of Statistical Software*, 33, 1–25.

Song, C., Dodds, W.K., Trentman, M.T., Rüegg, J. & Ballantyne, F. (2016). Methods of approximation influence aquatic ecosystem metabolism estimates. *Limnology and Oceanography: Methods*, 14, 557–569.

Webb, B. (1996). Trends in stream and river temperature. *Hydrological processes*, 10, 205–226.

Welter, J.R., Benstead, J.P., Cross, W.F., Hood, J.M., Huryn, A.D., Johnson, P.W. & Williamson, T.J. (2015). Does N_2 fixation amplify the temperature dependence of ecosystem metabolism? *Ecology*, 96, 603–610.

Williamson, T.J., Cross, W.F., Benstead, J.P., Gíslason, G.M., Hood, J.M., Huryn, A.D., Johnson, P.W. & Welter, J.R. (2016). Warming alters coupled carbon and nutrient cycles in experimental streams. *Global Change Biology*, 22, 2152–2164.

Yvon-Durocher, G., Caffrey, J.M., Cescatti, A., Dossena, M., del Giorgio, P., Gasol, J.M., Montoya, J.M., Pumpanen, J., Staehr, P.A., Trimmer, M. *et al.* (2012). Reconciling the temperature dependence of respiration across timescales and ecosystem types. *Nature*, 487, 472–476.

Yvon-Durocher, G., Jones, J.I., Trimmer, M., Woodward, G. & Montoya, J.M. (2010). Warming alters the metabolic balance of ecosystems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 2117–2126.

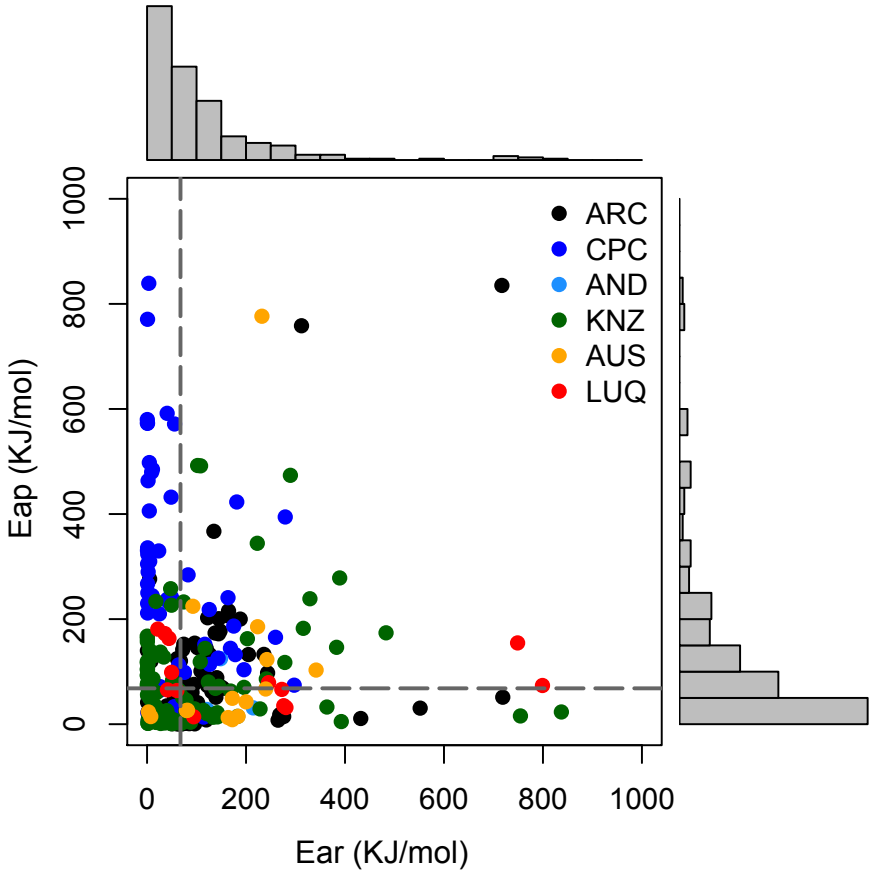

Figure 2.1: Intrinsic ecosystem level activation energies of GPP (E_{ap}) and ER (E_{ar}) in streams. Each point represents estimated E_{ap} and E_{ar} in a particular stream reach on one day. Histograms on the axes show the frequency distributions of E_{ap} and E_{ar}. Dashed lines are the medians of the frequency distributions.

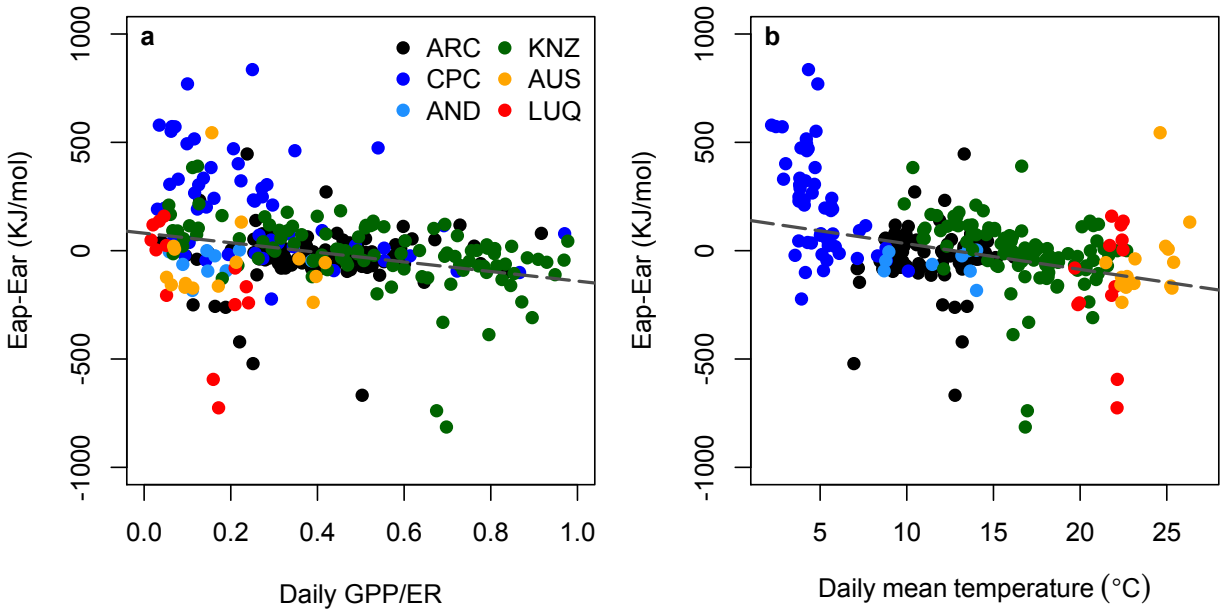

Figure 2.2: The empirical relationship between temperature sensitivity of GPP/ER, $E_{ap} - E_{ar}$, and (a) current GPP/ER and (b) mean daily temperature. Dashed lines are predictions based on fixed effects in the linear mixed effects model ($\mathrm{E_{ap}} - \mathrm{E_{ar}} = 236.92 - 221.20 \times \mathrm{GPP/ER} - 11.86 \times \mathrm{Temperature}$). In each panel, the prediction line is evaluated at the mean of the other covariate.

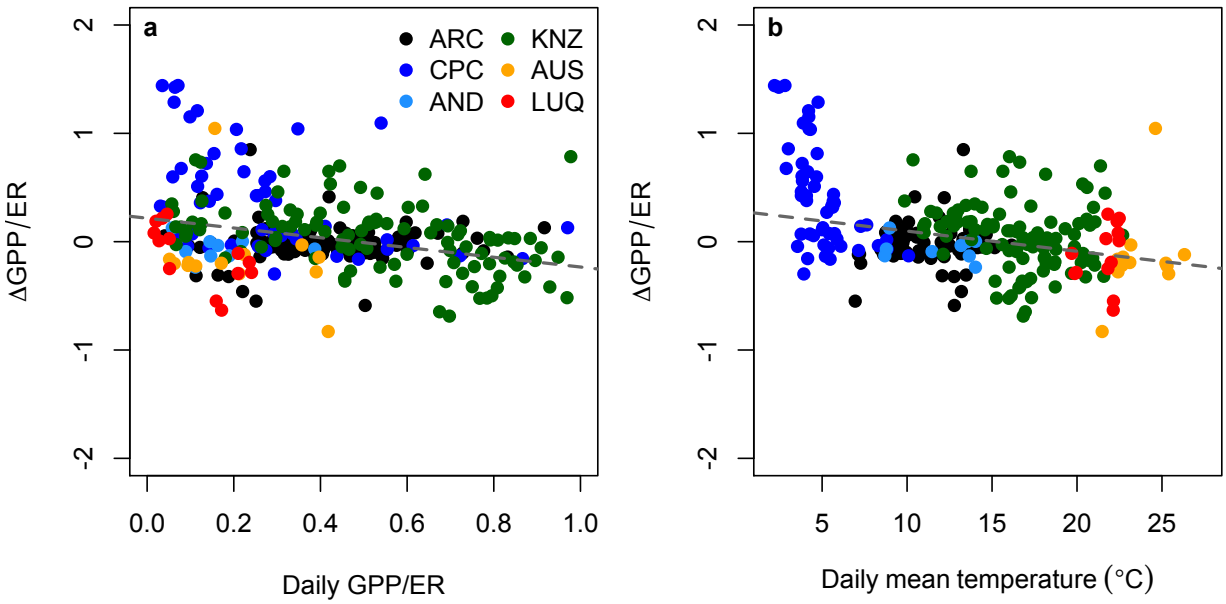

Figure 2.3: Proportional change in GPP/ER (ΔGPP/ER) with a 1 °C increase in temperature as a function of (a) current daily GPP/ER and (b) mean daily water temperature. Dashed lines are predictions based on fixed effects in the linear mixed effects model ($\Delta\text{GPP/ER} = 0.46 - 0.45 \times \text{GPP/ER} - 0.019 \times \text{Temperature}$). In each panel, the prediction line is evaluated at the mean of the other covariate.

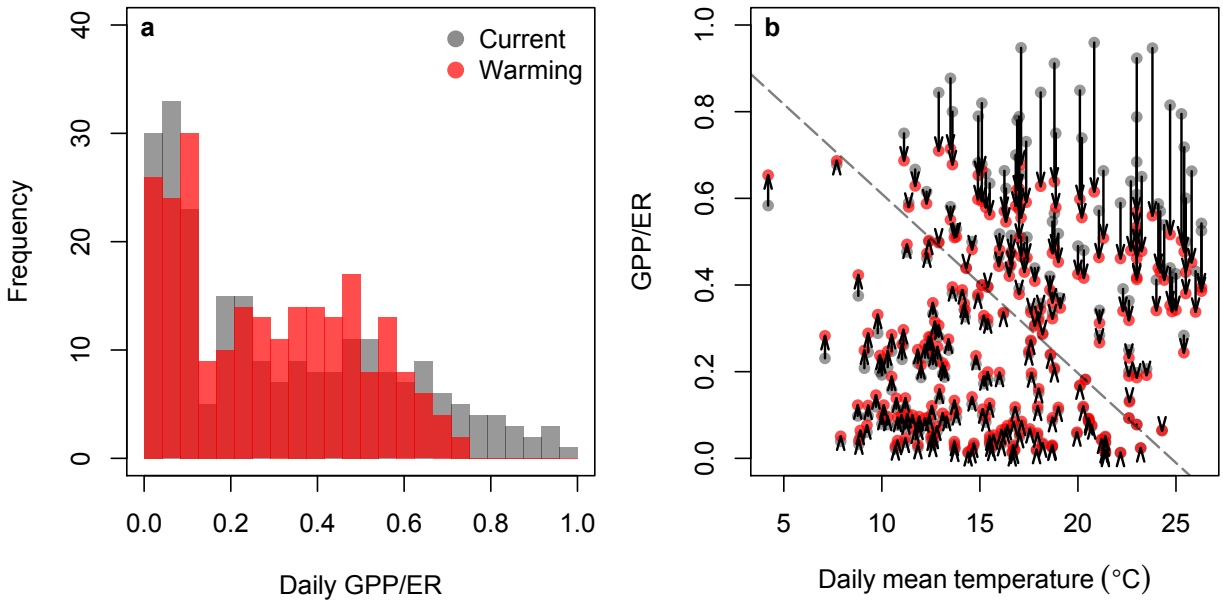

Figure 2.4: (a) Frequency distribution of GPP/ER currently and with a 1 °C increase in temperature. (b) Changes in GPP/ER with a 1 °C increase in temperature. Arrows indicate direction and magnitude of changes in GPP/ER from current temperature (gray dots) to 1 °C temperature increase (red dots). Dashed line is the isocline defined by the fixed effects from the linear mixed effects model ($0.46 - 0.45 \times \text{GPP/ER} - 0.019 \times \text{Temperature} = 0$), where GPP/ER is insensitive to temperature changes.

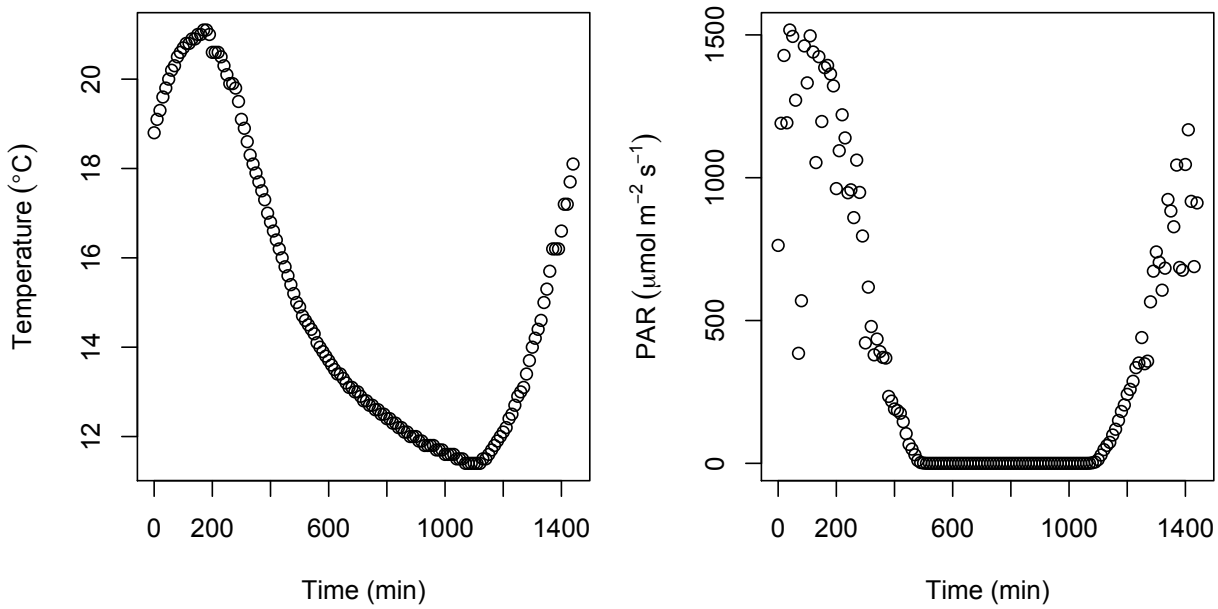

Figure 2.5: Trajectories of temperature and light used in the simulation study to establish the expected relationship between daily GPP/ER and $E_{ap} - E_{ar}$ as a result of statistical artifact.

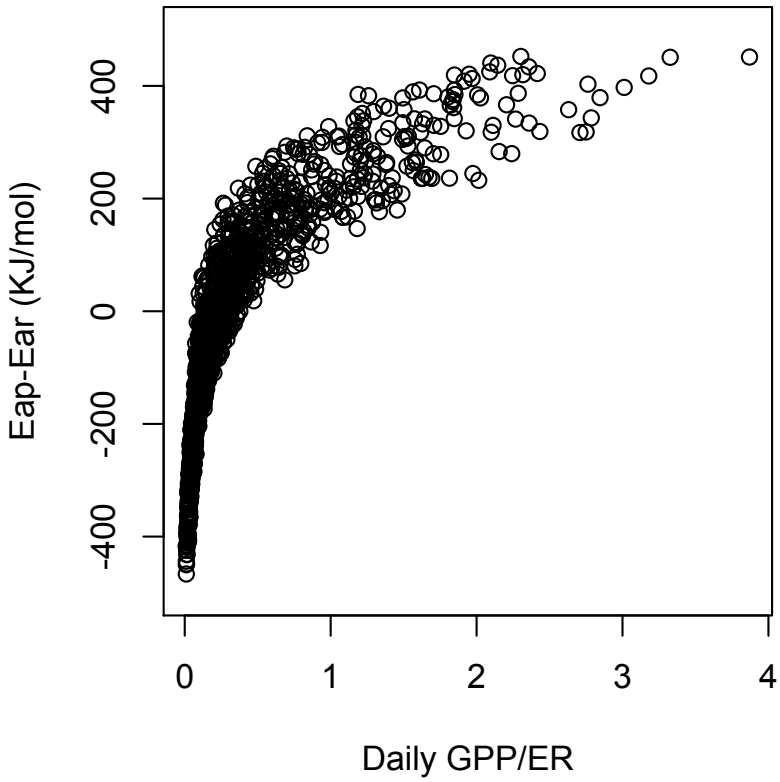

Figure 2.6: Expected relationship between GPP/ER and $E_{ap} - E_{ar}$ identified from simulated data. Since all parameters were drawn independently and randomly, the relationship is purely driven by the fact that E_{ap} and E_{ar} are fitted parameters used to calculated daily GPP and ER.

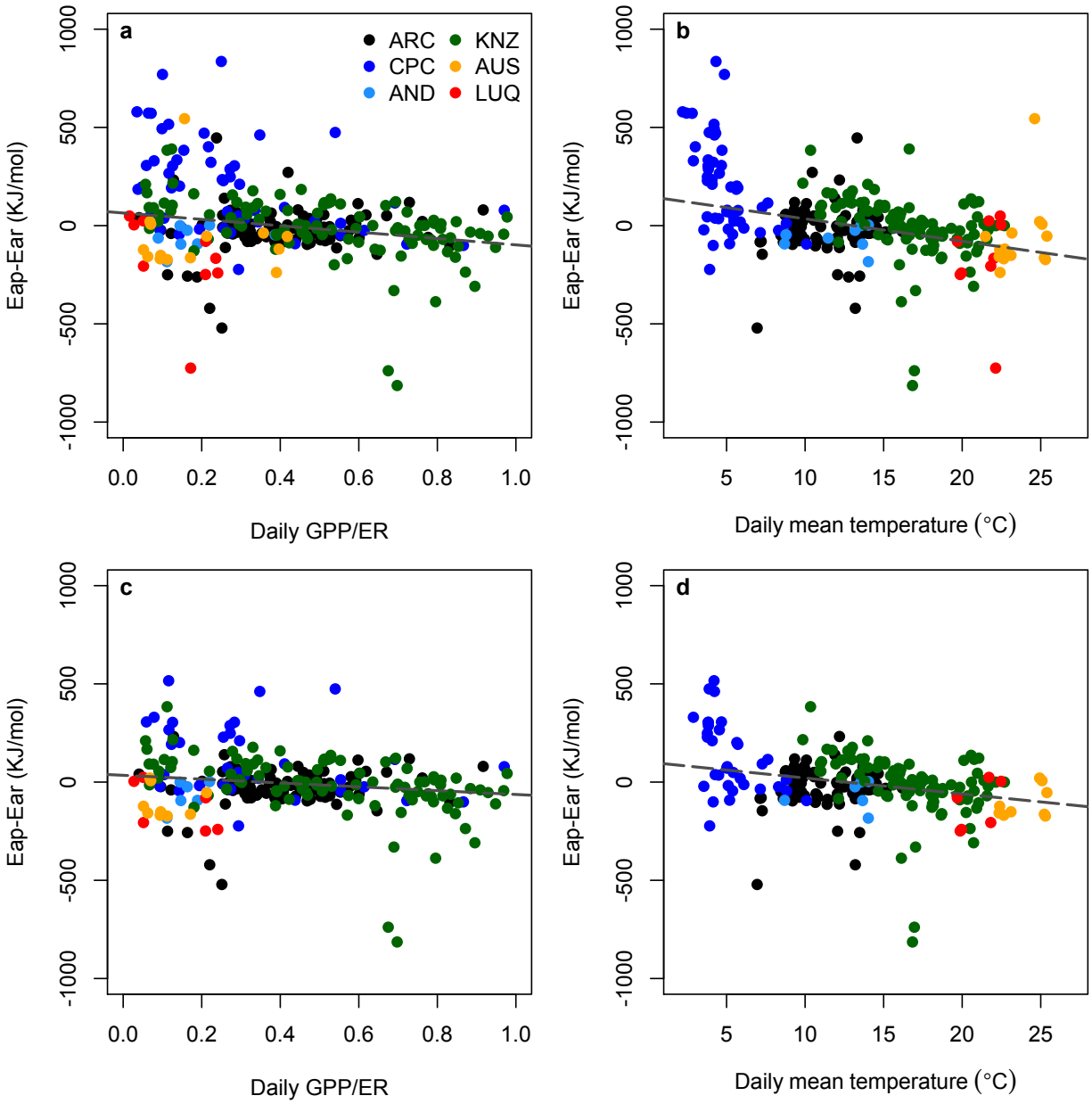

Figure 2.7: The empirical relationship between $E_{ap} - E_{ar}$, and (a, c) current GPP/ER and (b, d) mean daily temperature based on different thresholds of choosing estimable activation energies. Panel (a) and (b) use E_{ap} and E_{ar} estimates with 95% highest posterior density interval narrower than 400 KJ mol^{-1}; panel (c) and (d) use estimates with 95% highest posterior density interval narrower than 200 KJ mol^{-1}. Dashed lines are predictions based on fixed effects in the linear mixed effects model. In each panel, the prediction line is evaluated at the mean of the other covariate.

Chapter 3

Interaction Between Physiology and Environmental Heterogeneity Determines Discrepancies in Stream Metabolism across Spatial Scales

Introduction

Ecological processes operate on a variety of scales in ecosystems. The perceived ecological patterns are often strongly influenced by the scale of observation (Turner *et al.*, 1989; Wiens, 1989). In addition, many ecological problems require investigations on multiple scales (O'Neill *et al.*, 1986; Levin, 1992; Hewitt *et al.*, 2007). Therefore, the problem of scale is often acknowledged as a central problem in ecology (Levin, 1992; Schneider, 2001; Chave, 2013). Stream ecologists have long recognized the effects of spatial scales. Many general conceptual advances in stream ecology emphasized the explicit consideration of scales in understanding stream ecosystems (Vannote *et al.*, 1980; Poff, 1997; Habersack, 2000; Thorp *et al.*, 2006). In fact, the issue of scale is inherently embedded in stream ecosystems due to the dendritic structure of stream networks that naturally define distinct spatial scales. Stream habitats, characterized by distinct geomorphological and hydrological features, make up reaches, which connect together to form stream networks. These naturally defined spatial scales often coincide with characteristic scales on which ecological processes should be investigated. For example, while the transformation and transport of carbon in streams are best characterized on the reach or network scale (Battin *et al.*, 2008, 2009), investigation of the biological mechanism driving carbon flux in streams are best done on the habitat scale (Battin *et al.*, 2016). Given the multiple inherent scales nested in stream ecosystems, researchers have called for a multi-scale approach to understanding stream ecosystems and suggested that this could be a promising strategy for comprehending such complex systems (Cooper *et al.*, 1998; Lowe *et al.*, 2006).

A multi-scale approach is particularly important to study stream metabolism. Different applications of stream metabolism often require different spatial scales. In particular, reach and habitat scales are commonly used spatial scales to study stream metabolism. Reach, due to its relatively uniform hydrological characteristics, is usually considered as the basic unit for estimating network, regional, or global scale carbon flux (e.g. Battin *et al.* (2008); Raymond *et al.* (2013)). On the other hand, habitat scale manipulative experiments are often employed to investigate the mechanisms controlling metabolism. Stream metabolism has been extensively investigated on both reach and habitat scales. For example, the development of open channel method (Odum, 1956; Marzolf *et al.*, 1994; Young & Huryn, 1998) and related modeling approach (Holtgrieve *et al.*, 2010; Riley & Dodds, 2012; Grace *et al.*, 2015) have greatly facilitated studies of stream metabolism on the reach scale. At the same time, stream ecologists have a long history of measuring metabolism on the habitat scale using chamber incubations (Bott *et al.*, 1978; Dodds & Brock, 1998; Benstead *et al.*, 2009), leaf packs (Woodward *et al.*, 2012) or sediment cores (Grimm & Fisher, 1984; Jones *et al.*, 1995). These extensive investigations of stream metabolism on different spatial scales offer a great potential to combine the mechanistic understanding of the processes and the estimates on the desired spatial scale to comprehensively understand stream metabolism. However, a crucial missing component in realizing such potential is to understand how measurements made on different spatial scales relates to each other (Englund & Cooper, 2003). To date, empirical evidence comparing stream metabolism across spatial scales remains rare. Given the increasingly recognized importance of stream metabolism in global and regional carbon cycle (Battin *et al.*, 2008; Hotchkiss *et al.*, 2015), investigating the scale dependence of stream metabolism is imperative.

Ecological processes and patterns rarely translate across scales directly (Wiens, 1989; Duffy, 2009). Differences between metabolism measurements made on different scales could arise for multiple reasons. For example, abiotic conditions, such as light, temperature, nutrient concentration, and flow velocity, could differ among measurements made on different

spatial scales (Bott *et al.*, 1997). The spatial scale of measurements poses size constraints of sampling, causing over- or under-representation of particular types of substrata. Although the issue of representativeness across scales could be minimized by careful design of experimental equipment and control of experimental conditions, differences in metabolism across spatial scales can be inherent. Differences may persist even if the sampling on one scale is completely representative of another scale.

When biotic and abiotic conditions vary in space, large–scale dynamics differ from predictions of small–scale dynamics due to interactions between spatial heterogeneity and nonlinearity in small–scale dynamics. This phenomenon has been formalized in the scale transition theory (Chesson *et al.*, 2005; Chesson, 2012; Melbourne & Chesson, 2005; Baldocchi *et al.*, 2005; Melbourne & Chesson, 2006; Benedetti-Cecchi *et al.*, 2012). Mathematically, if function $f(x_i)$ describes metabolism rate as a function of abiotic and biotic conditions (x_i) on a small scale, $\overline{f(x_i)}$ is the metabolism rate on a larger scale because of the law of mass conservation. This is generally different from metabolism predicted based the average small scale biotic and abiotic conditions (i.e. $\overline{f(x_1, x_2, ...)} \neq f(\overline{x_1}, \overline{x_2}...)$) unless function $f(x_i)$ is linear. The differences across scales can be approximated with a second order Taylor expansion as

$$\overline{f(x_1, x_2...)} = f(\overline{x_1}, \overline{x_2}, ...) + \frac{1}{2} \sum f''_{x_i,x_j}(\overline{x_i}, \overline{x_j}) cov(x_i, x_j). \tag{3.1}$$

The differences across scales shown as $\frac{1}{2} \sum f''_{x_i,x_j}(\overline{x_i}, \overline{x_j}) cov(x_i, x_j)$, comes from the nonlinear ecological dynamics, represented in $f''_{x_i,x_j}(\overline{x_i}, \overline{x_j})$, and the existence of spatial variance of biotic and abiotic conditions, represented as $cov(x_i, x_j)$. The differences are the result of the fallacy of averaging and is an example of Jensen's inequality (Rastetter *et al.*, 1992). The existence of difference across scales only requires nonlinear dynamics ($f(x_i)$) and the spatial variation of biotic and abiotic conditions. The differences across scale can exist without any experimental artifacts or sampling constraints, and therefore can be considered as an inherent effect of scale and heterogeneity. The inherent differences across scales caused by the interaction between nonlinear dynamics and spatial heterogeneity are especially important for understanding stream metabolism across spatial scales. It is well known that metabolism

depends on abiotic conditions in a nonlinear fashion. For example, the temperature dependence of gross primary production (GPP) and ecosystem respiration (ER) and the light dependence of GPP (Jassby & Platt, 1976) in aquatic ecosystems are all nonlinear. In addition, streams are extremely heterogeneous in biotic and abiotic conditions (Palmer & Poff, 1997; Palmer *et al.*, 1997). The ubiquity of nonlinear dynamics and spatial heterogeneity suggest that stream metabolism is particularly scale dependent.

At the core of differences across scales is variability (Wiens, 2002). In an entirely homogeneous ecosystem, sampling bias or any inherent differences caused by nonlinear dynamics and heterogeneity would not exist. Therefore, proper incorporation and analysis of variability are crucial to understanding the effects of scale. One promising approach to investigate the scale dependence of stream metabolism is mechanistic modeling. On the one hand, mechanistic models for stream metabolism characterize the environmental dependence of metabolism, and thus allow for the standardization of metabolism measured on different spatial scales to the same environmental conditions. This partially resolves the issue of sampling representativeness. On the other hand, a mechanistic description of metabolism is required to analyze the direction and magnitude of differences caused by the interaction between heterogeneity and nonlinearity (equation 3.1). Combining the mechanistic modeling approach with metabolism measurements on multiple scales could be key in the multi-scale approach to stream ecology and help elucidate the processes responsible for the differences in metabolism across spatial scales.

Despite the widely recognized importance of scale in stream ecology and the extensive research in stream metabolism on multiple spatial scales, empirical measurements to directly test the effects of scale on metabolism are rare. A quantitative assessment of the inherent differences in metabolism rate across scales could help elucidate the mechanism responsible for the mismatch across scales and allow us to understand when and how metabolism measurements can translate across spatial scales. In this study, we estimated stream metabolism on the reach and habitat scales in six biomes to examine the scale dependence of stream

metabolism. We aimed to understand 1) whether metabolism, standardized to the same light and temperature conditions, differ across scales, and 2) whether the degree of differences changes with spatial heterogeneity in a consistent fashion as predicted by the interactions between spatial heterogeneity and nonlinearity. Specifically, we first analyze how differences in metabolism between reach and local habitat scale may change with the heterogeneity of habitat-scale metabolism based on scale transition theory. We then tested the expected pattern derived from the theoretical analyses against empirical data.

Methods and Materials

Study sites

We conducted this study in six biomes across two years. The six distinct biomes included tropical forest (Luquillo Experimental Forest, Puerto Rico (LUQ)), tropic savanna (Litchfield National Park, North Territory, Australia (AUS)), tallgrass prairie (Konza Prairie Long-term Ecological Research Station, Kansas, USA (KNZ)), temperate rainforest (Andrews Experimental Forest, Oregon, USA (AND)), boreal forest (Caribou-Poker Creeks Research Watershed, Alaska, USA (CPC)), and arctic tundra (Toolik Lake Field Station, Alaska, USA (ARC)). A detailed description of the study sites can be found in Rüegg *et al.* (2016). Within each biome, we chose three stream reaches for data collection each year, with some streams being sampled twice over the two years. The streams we chose in each biome were representative of the discharge gradients of the watershed. We collected data during base flow periods. Specifically, the sampling time for each biome was February–March 2013 and March 2014 for LUQ, July–August 2013 for AUS, May–June 2013 and April–June 2014 for KNZ, July–August 2015 for AND, July–August 2013 and 2014 for CPC, and July–August 2013 and 2014 for ARC.

Reach scale metabolism data collection

For reach scale metabolism, we recorded dissolved oxygen (DO) concentration, water temperature and barometric pressure using a YSI ProODO handheld optical DO meter (YSI Instruments, Yellow Springs, Ohio, USA) and photosynthetically active radiation (PAR) using an Odyssey Irradiance logger (DataFlowSystems, Christchurch, New Zealand) at a single location in the streams. The DO meters were calibrated to 100% saturation using water saturated air. The records from the irradiance logger were converted to PAR by comparison to a calibrated PAR sensor. Prior to deployment in each biome, all DO probes and irradiance loggers were run simultaneously at the same location to ensure consistency across probes. We recorded DO concentration, water temperature, and PAR at an interval of 5 minutes (ARC) or 10 minutes (all other sites) for two weeks.

Expressing the metabolism estimates as areal rates requires measurements of stream depth. We calculated the average depth of each stream reach based on discharge, velocity and width of the stream (depth = discharge/(width $\times$ velocity)). Specifically, we measured discharge and flow velocity of the streams by dilution gauging of short-term NaCl release (Kilpatrick & Cobb, 1985) once during the period when DO data were collected. We measured the wetted width of the streams at 10–20 transects along the streams. For streams reaches where discharge measurements were not made, we directly measured the depth of the reach in 10–20 transects, with 5 measurements of depth evenly spread along the width of the stream for each transect. Given that we collected reach scale metabolism data during the base flow periods without significant changes in flow conditions, the one–time measurement of depth was representative.

We measured the canopy cover above stream reaches in each cardinal directions using a spherical densiometer in the middle of the stream, where we collected metabolism data in CPC, KNZ, AUS, and LUQ. The average canopy cover of the four directions represented the canopy cover at each location within the focal stream reach. We performed the canopy cover measurements at 10–20 locations evenly distributed along each stream reach.

Habitat scale metabolism data collection

We incubated benthic substrata in recirculation chambers to measure metabolism on the habitat scale. We first filled plastic containers (10.0 cm×10.0 cm×6.6cm plastic containers with a mesh size of 1cm) with substrata representative of the streams. We distributed a set of five containers at each of the 8 to 10 locations in the reach. The containers were buried in the sediments with the top of the container flush with the stream bed for 30 days. After the 30-day incubation, we removed three containers at each location within a stream reach, and transported them to a sealed acrylic chamber. The chamber was equipped with an internal propeller driven circulation system (Rüegg *et al.*, 2015). Each chamber was filled with a known volume of stream water (about 10L) and sealed. We attached a YSI ProODO handheld optical DO meter (YSI Instruments, Yellow Springs, Ohio, USA) to a pre-made port on the chamber lid. The chamber was filled full and completely sealed to eliminate any air–water exchange of DO. PAR above each chamber was measured with an Odyssey Irradiance logger (DataFlowSystems, Christchurch, New Zealand). We recorded DO concentration, water temperature in the chamber, and PAR at a one minute or higher frequency for at least 30 minutes. During the first half of the incubation, we covered the chamber with a light impenetrable fabric. During the second half of the incubation, the chamber was exposed to ambient light. We maintained a constant circulation rate within the chamber by controlling the voltage supply to the propeller. All chamber metabolism measurements were performed during the period of reach scale metabolism data collection.

Metabolism modeling

We modeled the dynamics of DO concentration to estimate stream metabolism. In general, changes in DO concentration can be described as

$$\frac{d[\mathrm{O_2}]}{dt} = P(I,T) - R(T) + K(T)([\mathrm{O_2}]_{sat} - [\mathrm{O_2}]) \tag{3.2}$$

Here, $[O_2]$ is DO concentration ($mg\,L^{-1}$). $[O_2]_{sat}$ is the DO concentration at saturation, which can be calculated based on barometric pressure and temperature. P(I,T), R(T) and K(T) are GPP, ER, and reaeration coefficient, respectively. We modified the model formulae in Riley & Dodds (2012) to specify equation 3.2. Specifically, we calculated the $[O_2]_{sat}$ as (American Public Health Association, 1995)

$$[O_2]_{sat} = e^{-139.3441+\frac{157570}{T+273.15}-\frac{66423080}{(T+273.15)^2}+\frac{12438000000}{(T+273.15)^3}-\frac{862194900000}{(T+273.15)^4}} \times \frac{Pa \times 0.998}{101.3} \tag{3.3}$$

where T is temperature (°C) and Pa is barometric pressure (kPa). The reaeration coefficient K is temperature corrected (Elmore & West, 1961; Bott, 2006) as

$$K(T) = K_{20} \times 1.024^{(T-20)} \tag{3.4}$$

where K_{20} (min^{-1}) is the reaeration coefficient at 20°C. We modeled respiration rate as a temperature dependent process (Parkhill & Gulliver, 1999; Gulliver & Stefan, 1984) but use an Arrhenius equation to describe the temperature dependence

$$R(T) = R_{T_0} \times e^{-\frac{-E_{ar}}{R}(\frac{1}{T}-\frac{1}{T_0})} \tag{3.5}$$

where R_{T_0} ($mg\,L^{-1}\,min^{-1}$) is the respiration rate at reference temperature T_0, and E_{ar} ($KJ\,mol^{-1}$) is the activation energy of respiration. T_0 was set at the average daily temperature for reach scale metabolism and average water temperature during incubation for habitat scale metabolism to facilitate computational efficiency (Pinheiro *et al.*, 2015). We modeled GPP as a light saturating function (Jassby & Platt, 1976) with temperature dependence described by an Arrhenius equation.

$$P(I,T) = P_{max}\, tanh(\frac{\alpha I}{P_{max}}) \times e^{-\frac{-E_{ap}}{R}(\frac{1}{T}-\frac{1}{T_0})} \tag{3.6}$$

Here, I is PAR ($\mu E\ m^{-2}\ s^{-1}$), α ($mg\,L^{-1}\,min^{-1}\,\mu E^{-1}\,m^2\,s$) is the slope of the photosynthesis-irradiance curve at low light intensity at reference temperature T_0, and P_{max} is the photosynthesis rate at light saturation at reference temperature T_0. T_0 was specified the same way as in equation 3.5.

We employed a Bayesian approach to estimate parameters in equation 3.2–3.6 following Song *et al.* (2016). Specifically, for a given set of parameters, we used Runge-Kutta 4th order method with a step size of half of the DO probe logging interval to solve the differential equation describing DO dynamics (equation 3.2–3.6) and obtained a trajectory of modeled DO concentration. The Runge–Kutta 4th order method was implemented using function `rk4` in R package `deSolve` (Soetaert *et al.*, 2010). We used linear interpolation to approximate a continuous trajectory of light and temperature from discrete measurements. We assumed that the differences between modeled and measured DO were independent and identically distributed normal random errors. Based on such a distributional assumption of error, we computed the likelihood for any given set of parameters. We used uniform distributions as priors for all parameters in the model. We set the lower bound of the uniform priors at 0 and upper bound at values significantly larger than found in the literature for all parameters ($10\,\mathrm{mg\,L^{-1}\,min^{-1}}$ for R_{T_0}, $10\,\mathrm{mg\,L^{-1}\,\mu E^{-1}\,m^2\,s}$ for α, $20\,\mathrm{mg\,L^{-1}\,min^{-1}}$ for P_{max}, $10\,\mathrm{min^{-1}}$ for K_{20}, $1000\,\mathrm{KJ\,mol^{-1}}$ for both E_{ap} and E_{ar}). We used Markov Chain Monte Carlo to sample the posterior distribution of the parameters. Specifically, we implemented the adaptive random walk Metropolis–Hasting algorithm (Haario *et al.*, 2001) with the function `metrop` in R package `mcmc` (Geyer & Johnson, 2014). We ran each Markov chain for half a million iterations and used a burn-in period of 300000 iterations to ensure stationarity. We performed the visual inspection and Geweke diagnostic (Geweke, 1992) of the trace plot with R package `coda` (Plummer *et al.*, 2006) to ensure proper mixing and convergence of the Markov Chains. We calculated the 95% highest posterior density intervals for all parameters and only used estimates with the width of the interval less than half of the range of the prior distribution for further analyses to ensure that the estimated parameters in the DO model (equations 3.2–3.6) are mainly determined by data, not by the uniform prior distributions.

We applied the same model as prescribed by equation 3.2–3.6 to reach scale and habitat scale metabolism data. For reach scale metabolism, we simultaneously quantified reaeration, GPP and ER. Because the habitat level metabolism was measured

in sealed chambers without any air–water exchange of DO, we set K(T) in equation 3.2 to 0 for habitat scale metabolism. The parameter estimates for each scale was obtained following the same computational methods as outlined above. Parameter estimates for both reach and habitat scale metabolism (R_{T_0}, α, and P_{max}) were converted from volumetric rates to areal units. For reach scale parameters, they were converted as areal rate = volumetric rate/depth. For habitat scale parameters, they were converted as areal rate = volumetric rate × chamber volume/surface area, where surface area was 30 cm^2 for all chamber incubations.

Comparisons of metabolism across spatial scales

GPP and ER on the reach and habitat scale were not directly comparable because light and temperature were different between whole stream metabolism measurements and chamber incubations. Metabolism estimates on both scales should be standardized to the same light and temperature conditions for comparison. One way to correct the differences in temperature and light conditions is to compare parameters in equations 3.2–3.6 (R_{T_0}, α, and P_{max}) quantified on the reach and habitat scales. However, the chamber incubations for habitat scale metabolism measurements were not performed during full light. As a result, saturating photosynthetic rates P_{max} was usually not estimable on the habitat scale for most chambers. Thus, we were only able to compare R_{T_0} and α between reach and habitat scales. In addition, R_{T_0} and α were estimated at a particular reference temperature for each reach or chamber incubation. The reference temperature (T_0) was set at the mean temperature during the entire measurement period for reach scale metabolism and mean temperature during the chamber incubation for habitat scale metabolism to facilitate computational efficiency (Pinheiro *et al.*, 2015). To account for the difference in the reference temperature when comparing parameters across scales, we standardized reach level parameter estimates to the habitat scale reference temperature based on the estimated reach level activation energies. Specifically, to correct a reach level estimates made at temperature T_0 to the habitat scale

reference temperature T_1, we first sampled from the Markov Chains of reach scale parameters and used the following equations to obtain the posterior distribution of rates at T_1. We examined the width of the 95% highest posterior density intervals for R_{T_1} and α_{T_1}, and only used those estimates with the interval width less than half of the prior for further analyses. The mean of the posterior distribution of the research scale rates at T_1 is then compared to the corresponding rates on the habitat scale.

$$R_{T_1} = R_{T_0} \times e^{-\frac{E_{ar}}{R}(\frac{1}{T_1}-\frac{1}{T_0})} \tag{3.7}$$

$$\alpha_{T_1} = \alpha_{T_0} \times e^{-\frac{E_{ap}}{R}(\frac{1}{T_1}-\frac{1}{T_0})} \tag{3.8}$$

With such correction, the comparison between reach and habitat scale parameter estimates were made at the same temperature. For each stream reach, we measured reach scale metabolism over several days and habitat scale metabolism at multiple locations within the stream reach. Thus, we used the average over time to represent the reach scale metabolism and the average over space to represent the habitat scale metabolism. We calculated reach–to–habitat ratio of metabolism rates standardized to the same temperature to quantify the differences across scales.

Theoretical analysis of differences in metabolism across scales

The basic principle of analyzing the relationship between habitat and reach scale metabolism is the law of mass conservation. If measurements of metabolism are perfectly accurate on the habitat and reach scale, the sum of GPP and ER from all habitats within a focal reach should be equal to the GPP and ER in the reach. Therefore, the per area rate of GPP and ER on the reach scale should be equal to the average rate on the habitat scale. For GPP,

$$GPP_{reach} = \overline{GPP_{habitat}(P_{max,i}, \alpha_i, I_i, T)}. \tag{3.9}$$

Here, $P_{max,i}$, α_i, I_i and T are quantities for habitat i. We assume water temperature within a reach does not vary among habitats.

The habitat scale metabolism measurements were not made under full sunlight conditions, preventing us from estimating habitat scale P_{max} from most chamber incubations. Thus, we focused on comparing α between reach and habitat scales. Given that α describes how GPP changes with light at low light conditions, the estimation of α depends primarily on how GPP responds to light when light intensity is low (Jassby & Platt, 1976). Therefore, for comparison of α across scales, we can simplify the light response of GPP for theoretical analyses. At low light conditions, the saturating function we used to describe light response of GPP is approximately linear

$$P_{max}\tanh(\frac{\alpha I}{P_{max}}) \approx \alpha I. \tag{3.10}$$

With this approximation, we can simplify the relationship between reach and habitat scale GPP as

$$GPP_{reach} = \overline{\alpha_i I_i \times e^{-\frac{E_{ap,i}}{R}(\frac{1}{T}-\frac{1}{T_0})}} = \overline{\alpha_{T,i} I_i}. \tag{3.11}$$

The light intensity of individual habitat within a reach is not homogeneous. At any given average reach scale light $\overline{I}$, let a_i denote the light distribution within the reach such that $\overline{a_i} = 1$ and $I_i = a_i\overline{I}$ is the light intensity for habitat i. The reach scale GPP is

$$GPP_{reach} = \overline{\alpha_i a_i \overline{I}} = \overline{\alpha_i a_i}\overline{I}. \tag{3.12}$$

A second order Taylor expansion of the equation 3.12 led to:

$$GPP_{reach} \approx (\overline{\alpha_i}\,\overline{a_i} + \frac{1}{2}\mathrm{cov}(\alpha_i, a_i))\overline{I} = (\overline{\alpha_i} + \frac{1}{2}\rho_{\alpha,a}\sqrt{\mathrm{var}(\alpha_i)\mathrm{var}(a_i)})\overline{I}, \tag{3.13}$$

where $\rho_{\alpha,a}$ is the correlation coefficient between α_i and a_i

We derived the relationship between reach and habitat scale ER based on the same principle and analytical approach. Specifically, the reach and habitat scale ER can be linked based on the law of mass conservation as:

$$ER_{reach} = \overline{ER_{habitat}(R_{T_0,i}, T, E_{ar,i})} = \overline{R_{T_0,i} e^{(-\frac{E_{ar,i}}{R}(\frac{1}{T}-\frac{1}{T_0}))}}, \tag{3.14}$$

where $R_{T_0,i}$ and $E_{ar,i}$ are parameters for habitat i. After applying the second order Taylor expansion,

$$ER_{reach} \approx (1 + \frac{\text{var}(E_{ar,i})}{R^2}(\frac{1}{T} - \frac{1}{T_0})^2 - \frac{\text{cov}(E_{ar,i}, R_{T_0,i})}{\overline{R_{T_0,i}}R}(\frac{1}{T} - \frac{1}{T_0}))\overline{R_{T_0,i}}e^{-\frac{\overline{E_{ar,i}}}{R}(\frac{1}{T} - \frac{1}{T_0})}. \quad (3.15)$$

The theoretical analyses above assume that the ER and GPP on the habitat scale sum to the reach scale rates if measured accurately. The theoretically analyses demonstrate that when using average habitat scale biotic properties (i.e. $\overline{\alpha_i}$, $\overline{R_{T_0,i}}$) and reach scale average abiotic conditions (i.e. T and $\overline{I}$) to describe reach scale metabolism, the light and temperature dependence of GPP and ER on the reach scale, as approximated in equations 3.13 and 3.15, does not follow the same functional form describing the light and temperature dependence of GPP and ER on the habitat scale (equations 3.5 and 3.6). Consequently, if we apply the same function of light and temperature dependence of GPP and ER to reach and habitat scale, the estimated α and R_{T_0} will differ across scales. Such difference exists even if measurements on each scale are accurate and representative.

Specifically, equation 3.13 suggests that differences in α between reach and habitat scales should depend on the variance of habitat scale α and light within a reach, and the direction of the difference across scale as a function of habitat scale heterogeneity should depend on the correlation between habitat scale α and light. The theoretical analyses motivated us to examine how the differences in α between reach and habitat scales change with the variance of habitat scale α and light. Since we did not directly quantify the spatial heterogeneity of light within a stream reach, we used the variance of canopy cover as an indicator of the light variability within a reach. Similarly, equation 3.15 shows that the variance of habitat scale E_{ar}, the covariance between habitat scale E_{ar} and R_{T_0}, and the temperature used for comparing R_{T_0} across scale are the key elements determining the direction and magnitude of differences in R_{T_0} across scales. Given that we only directly quantified habitat scale R_{T_0}, the theoretical analyses led us to examine how differences in R_{T_0} between reach and habitat scales change with the variance of habitat scale R_{T_0}.

Expected differences in metabolism across scales based on simulations

The theoretical analyses identified variability of habitat scale R_{T_0} and α as potential key elements for differences in α and R_{T_0} across scales, but a quantitative description of these relationships is difficult to see directly from the theoretical analyses. Therefore, we performed two sets of numerical simulations to establish the expected relationship between differences in metabolism across scales and habitat scale variance. The overarching goal of the simulation is to examine whether metabolism still differs across scales when GPP or ER from the habitat sum to the rate at the reach scale.

We used the reach-to-habitat ratio of α and R_{T_0} to quantify the differences in metabolism across scales. As guided by the theoretical analyses, for comparison of α across scales, we simulated how reach-to-habitat ratio of α changes with the variance of habitat scale α in three scenarios: positive, negative, or no correlation between habitat scale α and light (i.e. a_i in equation 3.12). For comparison of R_{T_0} across scales, we performed simulations with positive, negative, and no correlation between habitat scale R_{T_0} and E_{ar}. For each scenario of correlation, we chose three temperatures (i.e. T_0) to which reach and habitat scale ER were standardized for comparison, lower bound, upper bound and the mean of reach scale daily temperature. We performed the first set of simulations using a hypothetical range of light and temperature that encompass the typical range of light and temperature observed in the streams we studied. We also carried out a second set of simulation using the actual measurements of light and temperature. The details of the simulations can be found in the appendix.

Statistical analyses of field data

As motivated by the theoretical analyses, we analyzed how reach–to–habitat ratio of α changes with the variance of canopy cover and habitat scale α separately, and how reach-to-habitat ratio of R_{T_0} changes with the variance of habitat scale R_{T_0} in linear mixed effects

models. We log–transformed both dependent and independent variables for analyses to conform to the assumptions of linearity and homoscedasticity of variance. Specifically, we used habitat scale variance of α and R_{T_0} or variance of canopy cover as continuous predictors for reach–to–habitat ratios of α and R_{T_0} respectively. The most complex random effects structure included a random intercept and a random slope for each biome. We first fit the model with maximum likelihood and performed backward model selection using AIC scores to select an adequate structure of the random effects (Bolker *et al.*, 2009). With a properly specified random effects structure, we tested the effects of habitat scale variance on reach–to–habitat ratio of α or R_{T_0} using F tests. All statistical analyses were performed in R 3.4.1 (R Core Team, 2017)

Results

Biome specific direction of differences in metabolism across scales

Metabolism quantified on different scales did not match in general. When directly comparing GPP and ER between reach and habitat scale without accounting for the light and temperature differences, habitat-scale GPP and ER were generally higher (Fig. 3.1). When reach and habitat scale metabolism was standardized to the same light and temperature conditions, GPP and ER differed across scales as well (Fig. 3.2). We observed no consistent over- or under- estimation of metabolism rates across scales. However, although the magnitude of differences in metabolism across scales within each biome varied substantially, the direction of differences between reach and habitat scale ER (i.e. R_{T_0}) and GPP (i.e. α) were generally consistent within a biome (Fig. 3.2). Notably, In LUQ, the reach scale R_{T_0} and α were generally larger than the habitat scale estimates. In contrast, the reach scale R_{T_0} and α were generally smaller than the habitat scale estimates in KNZ (Fig. 3.2). In ARC, we observed larger reach scale α but smaller reach scale R_{T_0} compared to the habitat scale estimates (Fig. 3.2).

Reach–to–habitat ratio of metabolism decreased with within–reach heterogeneity

The reach to habitat ratio of R_{T_0} and α decreased significantly with the habitat scale variance of R_{T_0} (Fig. 3.3a, $F_{1,30} = 20.40$, $P = 9.09 \times 10^{-5}$) and α (Fig. 3.3b, $F_{1,31} = 17.46$, $P = 0.00022$) within the reach respectively. In addition, the reach–to–habitat ratio of α decreased significantly with the variance of canopy cover within a reach (Fig. 3.3c, $F_{1,17} = 4.49$, $P = 0.049$). We did not find evidence for variation in how the reach-to-habitat ratio of metabolism changes with the variance of habitat scale metabolism within a reach or variance of canopy cover as the models without random slope or intercept had the smallest AIC scores (Table 3.1).

Correlation between habitat scale metabolism and abiotic conditions

We quantified the correlations between habitat scale R_{T_0} and E_{ar}. As outlined in the metabolism modeling section, T_0 was set at the mean temperature during chamber incubations for computational efficiency, and thus differed from reach to reach. Therefore, we first standardized R_{T_0} to the same temperature (20 °C, the mean temperature of all chamber incubations) according to equation 3.7 following the same procedure of standardizing reach scale R_{T_0} to chamber incubation temperatures. Due to the limited range of temperature during chamber incubations, we were only able to estimate habitat scale E_{ar} and standardize R_{T_0} to 20 °C for 13 chamber incubations from 3 reaches in KNZ. We observed a negative correlation between habitat scale R_{20} and E_{ar} (Pearson correlation coefficient, -0.48, $n = 13$, Fig. 3.4b).

We also quantified the correlation between habitat scale α and light. Ideally, the correlation between habitat scale α and light should be quantified using α and light measured on a specific habitat, and α should be standardized to the same temperature. However, we only measured light intensity at one location of the stream reach and did not directly measure the light intensity in each habitat we sampled for incubation. Moreover, due to the limited range

of temperature in chamber incubations and more parameters governing the GPP function (equation 3.6), the habitat scale E_{ap} was not estimable for almost all chambers. Thus, we used the average habitat scale α within a reach without temperature correction and the mean daily light intensity measured from the single location to approximate the correlation. Given the large range of habitat scale α, we log–transformed habitat scale α for calculation of correlation coefficient.We found a negative correlation between habitat scale α and light (Pearson correlation coefficient, -0.42, $n = 41$, Fig 3.4a).

Expected pattern of reach–to–habitat ratio of metabolism along gradients of habitat scale variance

The direction of correlation between habitat scale α and light determined how reach-to-habitat ratio of α changed with the variance of habitat scale α. If habitat scale α and light were uncorrelated, the reach and habitat scale α matched. The reach-to-habitat ratio of α did not change with the variance of habitat scale α (Fig. 3.5e). When habitat scale α and light were positively correlated, the reach-to-habitat ratio of α increased with the variance of habitat scale α (Fig. 3.5f). In contrast, a negative correlation between the reach-to-habitat ratio of α and the variance of habitat scale α arose if habitat scale α and light were negatively correlated (Fig. 3.5d).

The temperature T_0 used to compare reach and habitat scale R_{T_0} dictated the relationship between the reach-to-habitat ratio of R_{T_0} and the variance of habitat scale R_{T_0}. If T_0 was at the lower or upper bound of the daily temperature range, the reach-to-habitat ratio of R_{T_0} decreased with the variance of habitat scale R_{T_0} (Fig 3.5a, c). When T_0 was at the mean daily temperature, the reach-to-habitat ratio of R_{T_0} had a positive relationship with the variance of habitat scale R_{T_0} (Fig. 3.5b). The direction of correlation between R_{T_0} and E_{ar} did not influence the direction of the how difference in metabolism across scales changes with habitat scale heterogeneity for R_{T_0} (Fig. 3.9).

We used negative correlations between habitat scale α and light, and between R_{T_0} and E_{ar} as identified in the field data (Fig. 3.4) when simulating the difference in R_{T_0} across scales based on field light and temperature conditions and the empirical mean and variance of α and R_{T_0}. We found that the reach-to-habitat ratio of α and R_{T_0} both decreased with the habitat scale variance under realistic field conditions (Fig. 3.6).

Discussions

Biome specific directions of differences in metabolism across scales

When not accounting for differences in temperature and light between reach and habitat scale measurements, GPP and ER quantified on the habitat scale were generally higher than estimates on the reach scale (Fig. 3.1), primarily due to the high temperature of chamber incubation (Fig. 3.8). When standardized to the same light and temperature conditions, GPP and ER quantified on reach and habitat scales did not match in general (Fig 3.2). Differences between metabolism measurements made on different scales could arise for multiple reasons. Samples used to measure metabolism on a small scale may not be representative. For example, metabolism quantified with chamber incubations cannot fully include the contributions from the hyporheic zone, which could be a substantial part of total stream metabolism (Grimm & Fisher, 1984; Naegeli & Uehlinger, 1997; Fellows *et al.*, 2001). Metabolism often differs in substrata of different types (Hoellein *et al.*, 2009; Tromboni *et al.*, 2017) or sizes (Cardinale *et al.*, 2002), which indicates that over– or under–representing substrata of a particular type or size in the sample could influence metabolism measurements. The differences in metabolism across spatial scales could also result from inaccurate estimates of metabolism on different scales associated with inherent shortcomings of the methods used to quantify metabolism. Small scale measurements, such as chamber incubation or benthic cores, are prone to experimental artifacts such as altering the flow velocity, turbulence, nutrient concentration, dissolved oxygen concentration, and various other ambient conditions that could influence metabolic activities (Carpenter, 1996). On the other hand, although reach scale

metabolism model includes key processes driving DO dynamics, it may not include all processes affecting DO dynamics, such as groundwater input (McCutchan *et al.*, 2002; Hall & Tank, 2005), spatial heterogeneity of metabolism (Reichert *et al.*, 2009), influence of riparian vegetation (Dodds *et al.*, 2017), and different sources of respiratory substrates (Schindler *et al.*, 2017), all of which could cause significant differences in estimated GPP and ER in certain streams. Although we have standardized GPP and ER to the same light and temperature conditions to compare across scales, the existence of differences after standardization suggests that the differences and constraints associated with the methodology used on different spatial scales are still important.

The differences in metabolism rates between reach and habitat scales were biome specific (Fig. 3.2). The direction of differences in metabolism across scales was generally consistent within each biome. The biome specific direction of differences in metabolism across scales suggests that the particular biotic and abiotic conditions of the streams in a biome may be critical in determining the differences in metabolism measured on different scales. Although we did not directly test specific mechanisms responsible for the differences in metabolism across scales in each biome, previous studies on the drivers of metabolism could offer plausible explanations. For example, Chestnut & McDowell (2000) found a substantial contribution of carbon and nitrogen from riparian and hyporheic subsurface flow in LUQ streams. Therefore, chamber incubation of stream substrata may cause underestimation of metabolism because chamber incubation eliminates such supply of carbon and nitrogen, offering a plausible mechanism consistent with the observed high reach to habitat ratio of GPP and ER in LUQ. In ARC, bryophytes are responsible for the majority of photosynthesis and nutrient uptake in streams (Arscott *et al.*, 1998; Stream Bryophyte Group, 1999). Since the benthic substrata sampled for chamber incubations often under-represent bryophytes, we expect lower photosynthetic rates and higher nutrients availability in chamber incubations. As a result, chamber incubations likely underestimate GPP but over-estimate ER, consistent with the observation in ARC. Conversely, the measured depth to anoxia in KNZ (Wilson & Dodds, 2009) is

shallower than the depth of substrata baskets we used to quantify habitat scale metabolism. Thus, taking the substrata basket out of the stream for chamber incubation may create a more aerobic condition, increase nitrification, and decrease denitrification. As a result, the higher oxygen concentration and nitrogen availability in chamber incubation may lead to an overestimation of GPP and ER on the habitat scale. Overall, the biome-specific direction of differences in metabolism across scales despite consistent methodology across biomes indicates the importance of considering the particular conditions of the focal ecosystem when relating metabolism across scales.

Interactions between physiology and heterogeneity lead to differences in metabolism across scales

The reach–to–habitat ratio of α and R_{T_0} both decreased significantly with the variance of habitat scale α and R_{T_0} (Fig. 3.3). Such a trend of difference across scales cannot be explained by methodological differences used for measurements on different spatial scales, but can be partially explained by the interactions between nonlinear dynamics and spatial heterogeneity. The theoretical analyses showed that the functional form of the light and temperature dependence of GPP and ER describing habitat scale metabolism does not apply directly to the reach scale (equations 3.13, 3.15). As a result, if we fit the same function form of light and temperature dependence of GPP and ER to reach and habitat scale data, the estimated R_{T_0} and α will differ across scales.

Specifically, if GPP responds to light linearly at low light conditions on the habitat scale (i.e. $GPP_{habitat} \approx \alpha_i I_i$), the slope of the reach scale GPP–light relationship is not equal to the average on the habitat scale $\overline{\alpha_i}$ (Fig. 3.7b). The reach scale α can be approximated as $\overline{\alpha_i} + \frac{1}{2}\rho_{\alpha,a}\sqrt{\text{var}(\alpha_i)\text{var}(a_i)}$ (Equation 3.13). Thus, If the correlation between habitat scale α and light is negative (i.e. $\rho_{\alpha,a} < 0$), we expect a decreasing reach-to-habitat ratio of α with the variance of habitat scale α and variance of habitat scale light. This was confirmed in our

simulation under both hypothetical (Fig. 3.5d) and field light conditions (Fig. 3.6b). Meanwhile, based on the field data, we observed a negative correlation between habitat scale α and light (Fig 3.4a). Thus, the decreasing reach-to-habitat ratio of α with the variance of habitat scale α (Fig. 3.3b) was consistent with the theoretical prediction (Fig 3.5d, equation 3.13). Additionally, although we did not directly quantify the variance of habitat scale light within a reach, the variance of canopy cover determine the heterogeneity of light to a large extent (Stovall *et al.*, 2009; Warren *et al.*, 2013, 2016). Therefore, the negative relationship between reach-to-habitat ratio of α and variance of canopy (Fig. 3.3c) cover was also consistent with the theoretical prediction.

Similarly, the form of the temperature dependence of ER on the habitat scale does not directly apply to the reach scale (equation 3.15). If we describe the temperature dependence of ER on the habitat scale with an Arrhenius equation, the temperature dependence of ER on the reach scale, calculated by aggregating the habitat scale ER, does not follow the form of an Arrhenius equation (equation 3.15). As a result, fitting an Arrhenius equation to the reach scale data, as we did in modeling reach scale metabolism, will not recover the average habitat scale R_{T_0} (Fig. 3.7a).

To illustrate how differences in respiration standardized to the same temperature across scales occur schematically, we plotted the logarithm of ER against the inverse of temperature, known as the Arrhenius plot. If the temperature dependence of ER is described by Arrhenius equation, we expect a linear relationship between the logarithm of ER and the inverse of temperature. However, in the presence of spatial heterogeneity of E_{ar} within a stream reach, the relationship between the logarithm of ER and inverse of temperature on the reach scale is convex (Fig. 3.7a). The curvature increases with the variance of habitat scale E_{ar}. If a higher variance of E_{ar} corresponds to a higher variance of R_{T_0}, a reasonable assumption supported by a close association between R_{T_0} and E_{ar} observed in this study and studies in terrestrial ecosystems (Fierer *et al.*, 2005; Craine *et al.*, 2010; Lehmeier *et al.*, 2013), the curvature also increases with the variance of habitat scale R_{T_0}. Consequently, if T_0 is near

the lower or upper bound of the daily temperature range, the fitted reach scale R_{T_0} will under–estimate the average habitat scale R_{T_0}, and the reach-to-habitat ratio of R_{T_0} will decrease with the variance of habitat scale R_{T_0} (Fig. 3.7a). Conversely, if T_0 is near the mean daily temperature, the fitted reach scale R_{T_0} will over–estimate the average habitat scale R_{T_0}, and the reach-to-habitat ratio of R_{T_0} will increase with the variance of habitat scale R_{T_0} (Fig. 3.7a). Such expectation based on the analysis of the Arrhenius plot is confirmed in the simulation (Fig. 3.5a–c). Given that the chamber incubation temperature is often near the upper end of the daily temperature range (Fig. 3.8), we expect a negative relationship between the reach-to-habitat ratio of R_{T_0} and the variance of habitat scale R_{T_0}. This is consistent with the field data (Fig. 3.3a) and simulation under realistic field conditions (Fig. 3.6a). The consistency among field data (Fig. 3.3a), simulation (Fig 3.5c) and theoretical analyses (Fig 3.7a) indicates that the spatial heterogeneity in habitat scale ER is a plausible explanation for the decreasing reach-to-habitat ratio of R_{T_0} with the variance of habitat scale R_{T_0}.

The theoretical analyses above show two essential elements that gave rise to the observed relationship between differences in metabolism across scale and habitat scale variance in metabolism. The first key element is the existence of spatial variance of habitat scale GPP and ER within a reach. It is well recognized that spatial variances of biotic properties are ubiquitous in stream ecosystems (Palmer & Poff, 1997; Palmer *et al.*, 1997). For example, spatial variance of GPP and ER could arise from heterogeneity of abiotic environments, such as flow velocity (Sobczak & Burton, 1996), nutrients concentrations (Dent & Grimm, 1999), light availability caused by canopy cover (Keeton *et al.*, 2007; Stovall *et al.*, 2009; Warren *et al.*, 2013) or variation in substrata topography (Murdock & Dodds, 2007). Spatial variance in GPP and ER can also be reinforced by self–organization of benthic communities along heterogeneous abiotic environments (Dong *et al.*, 2017; Warren *et al.*, 2017). The second key component for the decreasing reach-to-habitat ratio of α with habitat scale variance of α is the negative correlation between habitat scale α and light. We consider this correlation

a general phenomenon as it has well established physiological basis. Studies on the algal photosynthesis–irradiance curve have shown that algae adapt to the low light environment by increasing chlorophyll concentration per cell, resulting in a negative correlation between α and light (Falkowski & Owens, 1980; Richardson *et al.*, 1983; Neale & Melis, 1986). For stream algae assemblages, shading experiments (Hill *et al.*, 1995; Rier *et al.*, 2006) or comparisons among streams with different light availability (Guasch & Subater, 1995; Roberts *et al.*, 2004) showed a negative relationship between α and light availability. α has even been shown to increase along the depth of the biofilms due to decreasing light intensity by self–shading (Dodds *et al.*, 1999). The ubiquity of spatial variance and the generality of physiology suggest that the decreasing reach–to–habitat ratio of GPP and ER with habitat scale variance of metabolism observed in this study is a general property of stream ecosystems.

Our findings showed that habitat scale heterogeneity in GPP and ER affect the differences between reach and habitat scale metabolism (Fig. 3.3). We also observed that the habitat scale heterogeneity in GPP and ER differ from biome to biome. For example, habitat scale heterogeneity in GPP and ER is generally lower in LUQ but higher in CPC. The differences in the heterogeneity in different biomes may stem from the differences in abiotic environment and the biotic communities of the streams. For example, metabolism in streams with permafrost soils, such as streams in CPC, are strongly driven by dissolved organic carbon and nutrient input (Petrone *et al.*, 2007). As a result, the heterogeneity of hyporheic flow from permafrost soils (Balcarczyk *et al.*, 2009) could lead to high spatial heterogeneity in local habitat scale respiration. On the other hand, ER in LUQ streams is primary driven by allochthonous input from adjacent riparian forest (Wantzen *et al.*, 2008). The relative low habitat scale heterogeneity in ER may be the result of the relative similar allochthonous input from the adjacent riparian forest. Together, this suggests that the magnitude of heterogeneity induced difference in metabolism across scales are likely dependent on biome.

In conclusion, we illustrated a predictable pattern of differences across scales along gradients of spatial heterogeneity in stream metabolism. In essence, our findings highlight that

mechanistic response functions, such as the GPP–light function and the ER–temperature function, do not directly translate across scales. The interaction among nonlinear dynamics, physiology, and spatial heterogeneity generates differences in metabolism across scales. Given their generality, we argue that such interaction is a universal source of differences for stream metabolism across scales, and thus should be considered when scaling up small–scale metabolism measurements. Our approach combining scale transition theory (Chesson, 2012; Melbourne & Chesson, 2005, 2006), mechanistic modeling (Denny & Benedetti-Cecchi, 2012), and numerical simulations explicitly quantified the contribution of such interaction to the differences in metabolism across scales, and provide a means to analytically incorporate the effects of such interaction when scaling tractable, small–scale observations to predictions of large–scale patterns.

Appendix

Simulating differences in metabolism across scales

We used numerical simulation to establish how differences in metabolism across scales change with habitat scale metabolism. We used the reach-to-habitat ratio of α and R_{T_0} to quantify the differences across scales. As guided by the theoretical analyses, for comparison of α across scales, we simulated how reach-to-habitat ratio of α changes with the variance of habitat scale α in three scenarios: positive, negative, or no correlation between habitat scale α and light (i.e. a_i in equation 3.12).Specifically, we assumed a stream reach comprised of 50 habitats. For each round of simulation, we randomly chose α_i and a_i in equation 3.12 for all habitats from gamma distributions. The mean and standard deviation of the gamma distribution for a_i was 1 and 0.2. The gamma distribution for α_i had a mean of 0.02 ($\mathrm{mg\,min^{-1}\,\mu E^{-1}\,s}$) and standard deviation chosen from a uniform distribution ranging from 0.005 to 0.05. The randomly chosen standard deviation allowed us to obtain a gradient of the variability of habitat scale α in different rounds of simulations. The mean and variance of α_i represented the realistic range of habitat scale α observed in the field data. When simulating under the assumption of a positive correlation between habitat scale α and light, we order the randomly chosen α_i and a_i such that the largest value of α_i corresponds to the largest value of a_i. When simulating under the assumption of negative correlation between habitat scale α and light, we order the randomly chosen α_i and a_i such that the largest value of α_i corresponds to the smallest value of a_i. At any average reach light intensity, we calculated the GPP of habitat i as $\alpha_i a_i \overline{I}$. The reach scale GPP was then calculated as the average of all habitat scale GPP. We evaluated reach scale GPP over a range of average reach scale light (from 0 to 300 $\mu\mathrm{mol\,m^{-2}\,s^{-1}}$). We regressed GPP on the average reach scale light and obtained the slope as the reach scale α.

For comparison of R_{T_0} across scales, we performed simulations with positive, negative, and no correlation between habitat scale R_{T_0} and E_{ar}. For each scenario of correlation, we

chose three temperatures (i.e. T_0) to which reach and habitat scale ER were standardized for comparison, lower bound, upper bound and the mean of reach scale daily temperature. Specifically, we randomly chose R_{T_0} for the habitats from a gamma distribution with a mean of 4 ($\mathrm{mg\,m^{-2}\,min^{-1}}$), and E_{ar} for the habitats from a gamma distribution with a mean of 70 ($\mathrm{KJ\,mol^{-1}}$). We used the same method as in the simulations for α to simulate positive or negative correlation between habitat scale R_{T_0} and E_{ar}. We varied the standard of the gamma distributions for R_{T_0} (from 0.5 to 3) and E_{ar} (from 40 to 45) in different rounds of simulations to obtain a range of variance of habitat scale R_{T_0}. We further assumed that a stream reach with a higher variance of habitat scale R_{T_0} also had a higher variance of habitat scale E_{ar}. To implement this assumption in the simulations, we ensured that a higher variance of R_{T_0} always corresponded to a higher variance of E_{ar} in different rounds of simulations. With habitat scale R_{T_0} and E_{ar} chosen, we calculated ER of habitat i at a particular temperature T as $R_{T_0} e^{-\frac{E_{ar}}{R}(\frac{1}{T}-\frac{1}{T_0})}$. The reach scale ER at temperature T was calculated as the average of habitat scale ER. We assumed daily temperature ranged from 18 to 22 °C, and calculated the reach scale ER over the hypothetical time series of daily temperature. We then fit a nonlinear regression between reach scale ER and temperature according to equation 3.5 to obtain the estimates of reach scale R_{T_0}. The nonlinear regression was fit using Levenberg–Marquardt algorithm (Levenberg, 1944; Marquardt, 1963) implemented in function `nlsLM` of R package `minpack.lm` (Elzhov *et al.*, 2016). As in the simulations for comparison of α, we quantified how reach-to-habitat ratio of R_{T_0} changes with the variance of habitat scale variance of R_{T_0} in each of the nine scenarios of simulations. We performed the simulations 1000 times. All simulations were done in R 3.4.1 (R Core Team, 2017).

We also performed simulations to examine how reach-to-habitat ratios of α and R_{T_0} change with the variance of habitat scale α and R_{T_0} under field light and temperature conditions. To this end, we followed the same method above for simulations but made a few modifications. For each stream reach, we randomly sampled habitat scale R_{T_0} and α based on the mean and variance estimated from chamber incubations in that reach, and evaluated

reach scale GPP and ER over the daily light and temperature measurements in that stream reach. The temperature to which ER was standardized for comparison was chosen at the average temperature of chamber incubations for the reach. This was to mimic the fact that we standardized the reach scale R_{T_0} to the average chamber incubation temperature for comparison in the analyses of the field data. We chose the direction of correlation between habitat scale α and light, R_{T_0} and E_{ar} in the simulations based on the identified direction of correlation from field data.

References

American Public Health Association (1995). *Standard methods for the examination of water and wastewater*. 19th edn. American Public Health Association, American Waterworkds Association and Water Environment Federation, Washington, DC.

Arscott, D.B., Bowden, W.B. & Finlay, J.C. (1998). Comparison of epilithic algal and bryophyte metabolism in an arctic tundra stream, alaska. *Journal of the North American Benthological Society*, 17, 210–227.

Balcarczyk, K.L., Jones, J.B., Jaffé, R. & Maie, N. (2009). Stream dissolved organic matter bioavailability and composition in watersheds underlain with discontinuous permafrost. *Biogeochemistry*, 94, 255–270.

Baldocchi, D.D., Krebs, T. & Leclerc, M.Y. (2005). "Wet/dry Daisyworld": a conceptual tool for quantifying the spatial scaling of heterogeneous landscapes and its impact on the subgrid variability of energy fluxes. *Tellus B*, 57, 175–188.

Battin, T.J., Besemer, K., Bengtsson, M.M., Romani, A.M. & Packmann, A.I. (2016). The ecology and biogeochemistry of stream biofilms. *Nature Reviews Microbiology*, 14, 251–263.

Battin, T.J., Kaplan, L.A., Findlay, S., Hopkinson, C.S., Marti, E., Packman, A.I., Newbold, J.D. & Sabater, F. (2008). Biophysical controls on organic carbon fluxes in fluvial networks. *Nature Geoscience*, 1, 95–100.

Battin, T.J., Luyssaert, S., Kaplan, L.A., Aufdenkampe, A.K., Richter, A. & Tranvik, L.J. (2009). The boundless carbon cycle. *Nature Geoscience*, 2, 598–600.

Benedetti-Cecchi, L., Tamburello, L., Bulleri, F., Maggi, E., Gennusa, V. & Miller, M. (2012). Linking patterns and processes across scales: the application of scale-transition theory to algal dynamics on rocky shores. *Journal of Experimental Biology*, 215, 977–985.

Benstead, J.P., Rosemond, A.D., Cross, W.F., Wallace, J.B., Eggert, S.L., Suberkropp, K., Gulis, V., Greenwood, J.L. & Tant, C.J. (2009). Nutrient enrichment alters storage and fluxes of detritus in a headwater stream ecosystem. *Ecology*, 90, 2556–2566.

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H. & White, J.S.S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24, 127–135.

Bott, T., Brock, J., Baatrup-Pedersen, A., Chambers, P., Dodds, W., Himbeault, K., Lawrence, J., Planas, D., Snyder, E. & Wolfaardt, G. (1997). An evaluation of techniques for measuring periphyton metabolism in chambers. *Canadian Journal of Fisheries and Aquatic Sciences*, 54, 715–725.

Bott, T., Brock, J., Cushing, C., Gregory, S., King, D. & Petersen, R. (1978). A comparison of methods for measuring primary productivity and community respiration in streams. *Hydrobiologia*, 60, 3–12.

Bott, T.L. (2006). Primary productivity and community respiration. In: *Methods in Stream Ecology* (eds. Hauer, F.R. & Lamberti, G.A.). Academic Press, San Diego, California, pp. 533–556.

Cardinale, B.J., Palmer, M.A., Swan, C.M., Brooks, S. & Poff, N.L. (2002). The influence of substrate heterogeneity on biofilm metabolism in a stream ecosystem. *Ecology*, 83, 412–422.

Carpenter, S.R. (1996). Microcosm experiments have limited relevance for community and ecosystem ecology. *Ecology*, 77, 677–680.

Chave, J. (2013). The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecology Letters*, 16, 4–16.

Chesson, P. (2012). Scale transition theory: its aims, motivations and predictions. *Ecological Complexity*, 10, 52–68.

Chesson, P., Donahue, M.J., Melbourne, B.A. & Sears, A.L. (2005). Scale transition theory for understanding mechanisms in metacommunities. In: *Metacommunities: spatial dynamics and ecological communities* (eds. Holyoak, M., Leibold, M.A. & Holt, R.D.). The University of Chicago Press, Chicago, IL, pp. 279–306.

Chestnut, T.J. & McDowell, W.H. (2000). C and n dynamics in the riparian and hyporheic zones of a tropical stream, luquillo mountains, puerto rico. *Journal of the North American Benthological Society*, 19, 199–214.

Cooper, S.D., Diehl, S., Kratz, K. & Sarnelle, O. (1998). Implications of scale for patterns and processes in stream ecology. *Australian Journal of Ecology*, 23, 27–40.

Craine, J.M., Fierer, N. & McLauchlan, K.K. (2010). Widespread coupling between the rate and temperature sensitivity of organic matter decay. *Nature Geoscience*, 3, 854–857.

Denny, M. & Benedetti-Cecchi, L. (2012). Scaling up in ecology: mechanistic approaches. *Annual Review of Ecology, Evolution, and Systematics*, 43, 1–22.

Dent, C.L. & Grimm, N.B. (1999). Spatial heterogeneity of stream water nutrient concentrations over successional time. *Ecology*, 80, 2283–2298.

Dodds, W. & Brock, J. (1998). A portable flow chamber for *in situ* determination of benthic metabolism. *Freshwater Biology*, 39, 49–59.

Dodds, W.K., Biggs, B.J. & Lowe, R.L. (1999). Photosynthesis-irradiance patterns in benthic microalgae: variations as a function of assemblage thickness and community structure. *Journal of Phycology*, 35, 42–53.

Dodds, W.K., Tromboni, F., Aparecido Saltarelli, W., Cunha, F. & Gasparini, D. (2017). The root of the problem: Direct influence of riparian vegetation on estimation of stream ecosystem metabolic rates. *Limnology and Oceanography Letters*, 2, 9–17.

Dong, X., Ruhí, A. & Grimm, N.B. (2017). Evidence for self-organization in determining spatial patterns of stream nutrients, despite primacy of the geomorphic template. *Proceedings of the National Academy of Sciences*, p. 201617571.

Duffy, J.E. (2009). Why biodiversity is important to the functioning of real-world ecosystems. *Frontiers in Ecology and the Environment*, 7, 437–444.

Elmore, H. & West, W. (1961). Effect of water temperature on stream reaeration. *Journal of the Sanitary Engineering Division*, 87, 59–71.

Elzhov, T.V., Mullen, K.M., Spiess, A.N. & Bolker, B. (2016). *minpack.lm: R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds*. R package version 1.2–1.

Englund, G. & Cooper, S.D. (2003). Scale effects and extrapolation in ecological experiments. *Advances in ecological research*, 33, 161–213.

Falkowski, P.G. & Owens, T.G. (1980). Ligh–shade adaptation two strategies in marine phytoplankton. *Plant Physiology*, 66, 592–595.

Fellows, C.S., Valett, M.H. & Dahm, C.N. (2001). Whole-stream metabolism in two montaine streams: contribution of the hyporheic zone. *Limnology and Oceanography*, 46, 523–531.

Fierer, N., Craine, J.M., McLauchlan, K. & Schimel, J.P. (2005). Litter quality and the temperature sensitivity of decomposition. *Ecology*, 86, 320–326.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: *Bayesian Statistics 4* (eds. Bernado, J., Berger, J., Dawid, A. & Smith, A.). Clarendon Press, Oxford, UK, pp. 169–193.

Geyer, C.J. & Johnson, L.T. (2014). *mcmc: Markov Chain Monte Carlo*. R package version 0.9–3.

Grace, M.R., Giling, D.P., Hladyz, S., Caron, V., Thompson, R.M. & Mac Nally, R. (2015). Fast processing of diel oxygen curves: Estimating stream metabolism with BASE (BAyesian Single-station Estimation). *Limnology and Oceanography: Methods*, 13, 103–114.

Grimm, N.B. & Fisher, S.G. (1984). Exchange between interstitial and surface water: implications for stream metabolism and nutrient cycling. *Hydrobiologia*, 111, 219–228.

Guasch, H. & Subater, S. (1995). Seasonal variations in photosynthesis-irradiance responses by biofilms in Mediterranean streams. *Journal of Phycology*, 31, 727–735.

Gulliver, J.S. & Stefan, H.G. (1984). Stream productivity analysis with dorm–I: Development of computational model. *Water Research*, 18, 1569–1576.

Haario, H., Saksman, E. & Tamminen, J. (2001). An adaptive metropolis algorithm. *Bernoulli*, 7, 223–242.

Habersack, H.M. (2000). The river-scaling concept (RSC): a basis for ecological assessments. *Hydrobiologia*, 422, 49–60.

Hall, R.O. & Tank, J.L. (2005). Correcting whole-stream estimates of metabolism for groundwater input. *Limnology and Oceanography: Methods*, 3, 222–229.

Hewitt, J.E., Thrush, S.F., Dayton, P.K. & Bonsdorff, E. (2007). The effect of spatial and temporal heterogeneity on the design and analysis of empirical studies of scale-dependent systems. *The American Naturalist*, 169, 398–408.

Hill, W.R., Ryon, M.G. & Schilling, E.M. (1995). Light limitation in a stream ecosystem: responses by primary producers and consumers. *Ecology*, 76, 1297–1309.

Hoellein, T.J., Tank, J.L., Rosi-Marshall, E.J. & Entrekin, S.A. (2009). Temporal variation in substratum-specific rates of n uptake and metabolism and their contribution at the stream-reach scale. *Journal of the North American Benthological Society*, 28, 305–318.

Holtgrieve, G.W., Schindler, D.E., Branch, T.A. & A'mar, Z.T. (2010). Simultaneous quantification of aquatic ecosystem metabolism and reaeration using a bayesian statistical model of oxygen dynamics. *Limnology and Oceanography*, 55, 1047–1063.

Hotchkiss, E., Hall Jr, R., Sponseller, R., Butman, D., Klaminder, J., Laudon, H., Rosvall, M. & Karlsson, J. (2015). Sources of and processes controlling CO_2 emissions change with the size of streams and rivers. *Nature Geoscience*, 8, 696–699.

Jassby, A.D. & Platt, T. (1976). Mathematical formulation of the relationship between photosynthesis and light for phytoplankton. *Limnology and Oceanography*, 21, 540–547.

Jones, J.B., Fisher, S.G. & Grimm, N.B. (1995). Vertical hydrologic exchange and ecosystem metabolism in a Sonoran Desert stream. *Ecology*, 76, 942–952.

Keeton, W.S., Kraft, C.E. & Warren, D.R. (2007). Mature and old-growth riparian forests: structure, dynamics, and effects on Adirondack stream habitats. *Ecological Applications*, 17, 852–868.

Kilpatrick, F.A. & Cobb, E.D. (1985). Measurement of discharge using tracers. In: *U.S. Geological Survey Techniques of Water-Resources Investigations, Book 3*. U.S. Geological Survey, chap. A16.

Lehmeier, C.A., Min, K., Niehues, N.D., Ballantyne, F. & Billings, S.A. (2013). Temperature-mediated changes of exoenzyme-substrate reaction rates and their consequences for the carbon to nitrogen flow ratio of liberated resources. *Soil Biology and Biochemistry*, 57, 374–382.

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics*, 2, 164–168.

Levin, S.A. (1992). The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology*, 73, 1943–1967.

Lowe, W.H., Likens, G.E. & Power, M.E. (2006). Linking scales in stream ecology. *BioScience*, 56, 591–597.

Marquardt, D.W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11, 431–441.

Marzolf, E.R., Mulholland, P.J. & Steinman, A.D. (1994). Improvements to the diurnal upstream-downstream dissolved oxygen change technique for determining whole-stream metabolism in small streams. *Canadian Journal of Fisheries and Aquatic Sciences*, 51, 1591–1599.

McCutchan, J., Saunders, J., Lewis, W., Hayden, M.G. *et al.* (2002). Effects of groundwater flux on open–channel estimates of stream metabolism. *Limnology and Oceanography*, 47, 321–324.

Melbourne, B.A. & Chesson, P. (2005). Scaling up population dynamics: integrating theory and data. *Oecologia*, 145, 178–186.

Melbourne, B.A. & Chesson, P. (2006). The scale transition: scaling up population dynamics with field data. *Ecology*, 87, 1478–1488.

Murdock, J.N. & Dodds, W.K. (2007). Linking benthic algal biomass to stream substratum topography. *Journal of Phycology*, 43, 449–460.

Naegeli, M.W. & Uehlinger, U. (1997). Contribution of the hyporheic zone to ecosystem metabolism in a prealpine gravel-bed-river. *Freshwater Science*, 16, 794–804.

Neale, P.J. & Melis, A. (1986). Algal photosynthetic membrane complexes and the photosynthesis-irradiance curve: a comparison of light-adaptation responses in *chlamydomonas reinhardtii (chlorophyta). Journal of Phycology*, 22, 531–538.

Odum, H.T. (1956). Primary production in flowing waters. *Limnology and Oceanography*, 1, 102–117.

O'Neill, R.V., Deangelis, D.L., Waide, J. & Allen, T.F. (1986). *A hierarchical concept of ecosystems*. Princeton University Press.

Palmer, M.A., Hakenkamp, C.C. & Nelson-Baker, K. (1997). Ecological heterogeneity in streams: why variance matters. *Journal of the North American Benthological Society*, 16, 189–202.

Palmer, M.A. & Poff, N.L. (1997). The influence of environmental heterogeneity on patterns and processes in streams. *Journal of the North American Benthological Society*, 16, 169–173.

Parkhill, K.L. & Gulliver, J.S. (1999). Modeling the effect of light on whole-stream respiration. *Ecological Modelling*, 117, 333–342.

Petrone, K., Hinzman, L., Shibata, H., Jones, J. & Boone, R. (2007). The influence of fire and permafrost on sub-arctic stream chemistry during storms. *Hydrological Processes*, 21, 423–434.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team (2015). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-122.

Plummer, M., Best, N., Cowles, K. & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R news*, 6, 7–11.

Poff, N.L. (1997). Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society*, 16, 391–409.

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rastetter, E.B., King, A.W., Cosby, B.J., Hornberger, G.M., O'Neill, R.V. & Hobbie, J.E. (1992). Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. *Ecological Applications*, 2, 55–70.

Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C. *et al.* (2013). Global carbon dioxide emissions from inland waters. *Nature*, 503, 355–359.

Reichert, P., Uehlinger, U. & Acuña, V. (2009). Estimating stream metabolism from oxygen concentrations: effect of spatial heterogeneity. *Journal of Geophysical Research: Biogeosciences*, 114.

Richardson, K., Beardall, J. & Raven, J. (1983). Adaptation of unicellular algae to irradiance: an analysis of strategies. *New Phytologist*, 93, 157–191.

Rier, S.T., Stevenson, R.J. & LaLiberte, G.D. (2006). Photo-acclimation response of benthic stream algae across experimentally manipulated light gradients: a comparison of growth rates and net primary productivity. *Journal of phycology*, 42, 560–567.

Riley, A.J. & Dodds, W.K. (2012). Whole-stream metabolism: strategies for measuring and modeling diel trends of dissolved oxygen. *Freshwater Science*, 32, 56–69.

Roberts, S., Sabater, S. & Beardall, J. (2004). Benthic microalgal colonization in streams of differing riparian cover and light availability. *Journal of Phycology*, 40, 1004–1012.

Rüegg, J., Brant, J.D., Larson, D.M., Trentman, M.T. & Dodds, W.K. (2015). A portable, modular, self-contained recirculating chamber to measure benthic processes under controlled water velocity. *Freshwater Science*, 34, 831–844.

Rüegg, J., Dodds, W.K., Daniels, M.D., Sheehan, K.R., Baker, C.L., Bowden, W.B., Farrell, K.J., Flinn, M.B., Harms, T.K., Jones, J.B. *et al.* (2016). Baseflow physical characteristics differ at multiple spatial scales in stream networks across diverse biomes. *Landscape Ecology*, 31, 119–136.

Schindler, D.E., Jankowski, K., A'mar, Z.T. & Holtgrieve, G.W. (2017). Two-stage metabolism inferred from diel oxygen dynamics in aquatic ecosystems. *Ecosphere*, 8.

Schneider, D.C. (2001). The rise of the concept of scale in ecology the concept of scale is evolving from verbal expression to quantitative expression. *BioScience*, 51, 545–553.

Sobczak, W.V. & Burton, T.M. (1996). Epilithic bacterial and algal colonization in a stream run, riffle, and pool: a test of biomass covariation. *Hydrobiologia*, 332, 159–166.

Soetaert, K., Petzoldt, T. & Setzer, R.W. (2010). Solving differential equations in R: Package desolve. *Journal of Statistical Software*, 33, 1–25.

Song, C., Dodds, W.K., Trentman, M.T., Rüegg, J. & Ballantyne, F. (2016). Methods of approximation influence aquatic ecosystem metabolism estimates. *Limnology and Oceanography: Methods*, 14, 557–569.

Stovall, J.P., Keeton, W.S. & Kraft, C.E. (2009). Late-successional riparian forest structure results in heterogeneous periphyton distributions in low-order streams. *Canadian Journal of Forest Research*, 39, 2343–2354.

Stream Bryophyte Group (1999). Roles of bryophytes in stream ecosystems. *Journal of the North American Benthological Society*, 18, 151–184.

Thorp, J.H., Thoms, M.C. & Delong, M.D. (2006). The riverine ecosystem synthesis: biocomplexity in river networks across space and time. *River Research and Applications*, 22, 123–147.

Tromboni, F., Dodds, W.K., Neres-Lima, V., Zandonà, E. & Moulton, T.P. (2017). Heterogeneity and scaling of photosynthesis, respiration, and nitrogen uptake in three Atlantic Rainforest streams. *Ecosphere*, 8, e01959.

Turner, M.G., O'Neill, R.V., Gardner, R.H. & Milne, B.T. (1989). Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology*, 3, 153–162.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. & Cushing, C.E. (1980). The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences*, 37, 130–137.

Wantzen, K.M., Yule, C.M., Mathooko, J.M. & Pringle, C.M. (2008). Organic matter processing in tropical streams. In: *Tropical stream ecology*. Elsevier, pp. 43–64.

Warren, D.R., Collins, S.M., Purvis, E.M., Kaylor, M.J. & Bechtold, H.A. (2017). Spatial variability in light yields colimitation of primary production by both light and nutrients in a forested stream ecosystem. *Ecosystems*, 20, 198–210.

Warren, D.R., Keeton, W.S., Bechtold, H.A. & Rosi-Marshall, E.J. (2013). Comparing streambed light availability and canopy cover in streams with old-growth versus early-mature riparian forests in western oregon. *Aquatic Sciences*, 75, 547–558.

Warren, D.R., Keeton, W.S., Kiffney, P.M., Kaylor, M.J., Bechtold, H.A. & Magee, J. (2016). Changing forest–changing streams: riparian forest stand development and ecosystem function in temperate headwaters. *Ecosphere*, 7.

Wiens, J.A. (1989). Spatial scaling in ecology. *Functional ecology*, 3, 385–397.

Wiens, J.A. (2002). Riverine landscapes: taking landscape ecology into the water. *Freshwater biology*, 47, 501–515.

Wilson, K.C. & Dodds, W.K. (2009). Centimeter-scale stream substratum heterogeneity and metabolic rates. *Hydrobiologia*, 623, 53–62.

Woodward, G., Gessner, M.O., Giller, P.S., Gulis, V., Hladyz, S., Lecerf, A., Malmqvist, B., McKie, B.G., Tiegs, S.D., Cariss, H. *et al.* (2012). Continental-scale effects of nutrient pollution on stream ecosystem functioning. *Science*, 336, 1438–1440.

Young, R.G. & Huryn, A.D. (1998). Comment: Improvements to the diurnal upstream-downstream dissolved oxygen change technique for determining whole-stream metabolism in small streams. *Canadian Journal of Fisheries and Aquatic Sciences*, 55, 1784–1785.

Table 3.1: Backwards model selection using AIC for proper structure of the random effects. β_0 and β_1 are fixed effects intercept and slope. b_{0i} and b_{1i} are random effect intercept and slope for biome i. ε is random error.

Model	AIC
$\log(\frac{R_{T_0,reach}}{R_{T_0,habitat}}) = \beta_0 + \beta_1 \log(\text{var}(R_{T_0,habitat})) + b_{0i} + b_{1i} \log(\text{var}(R_{T_0,habitat})) + \varepsilon$	109.99
$\log(\frac{R_{T_0,reach}}{R_{T_0,habitat}}) = \beta_0 + \beta_1 \log(\text{var}(R_{T_0,habitat})) + b_{0i} + \varepsilon$	105.99
$\log(\frac{R_{T_0,reach}}{R_{T_0,habitat}}) = \beta_0 + \beta_1 \log(\text{var}(R_{T_0,habitat})) + \varepsilon$	103.99
$\log(\frac{\alpha_{reach}}{\alpha_{habitat}}) = \beta_0 + \beta_1 \log(\text{var}(\alpha_{habitat})) + b_{0i} + b_{1i} \log(\text{var}(\alpha_{habitat})) + \varepsilon$	150.39
$\log(\frac{\alpha_{reach}}{\alpha_{habitat}}) = \beta_0 + \beta_1 \log(\text{var}(\alpha_{habitat})) + b_{0i} + \varepsilon$	147.98
$\log(\frac{\alpha_{reach}}{\alpha_{habitat}}) = \beta_0 + \beta_1 \log(\text{var}(\alpha_{habitat})) + \varepsilon$	145.98
$\log(\frac{\alpha_{reach}}{\alpha_{habitat}}) = \beta_0 + \beta_1 \log(\text{var}(\text{canopy})) + b_{0i} + b_{1i} \log(\text{var}(\text{canopy})) + \varepsilon$	92.75
$\log(\frac{\alpha_{reach}}{\alpha_{habitat}}) = \beta_0 + \beta_1 \log(\text{var}(\text{canopy})) + b_{0i} + \varepsilon$	89.07
$\log(\frac{\alpha_{reach}}{\alpha_{habitat}}) = \beta_0 + \beta_1 \log(\text{var}(\text{canopy})) + \varepsilon$	87.07

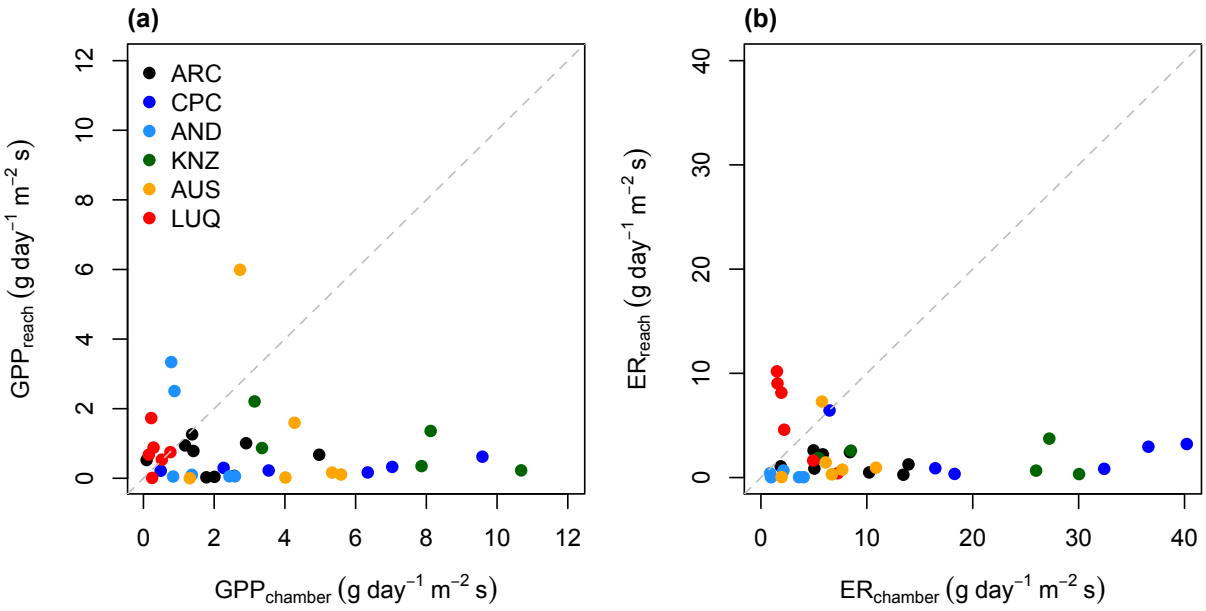

Figure 3.1: Direct comparison of GPP and ER quantified on the reach and habitat scale. GPP and ER shown here are not corrected for different light and temperature conditions between reach and habitat scale measurements.

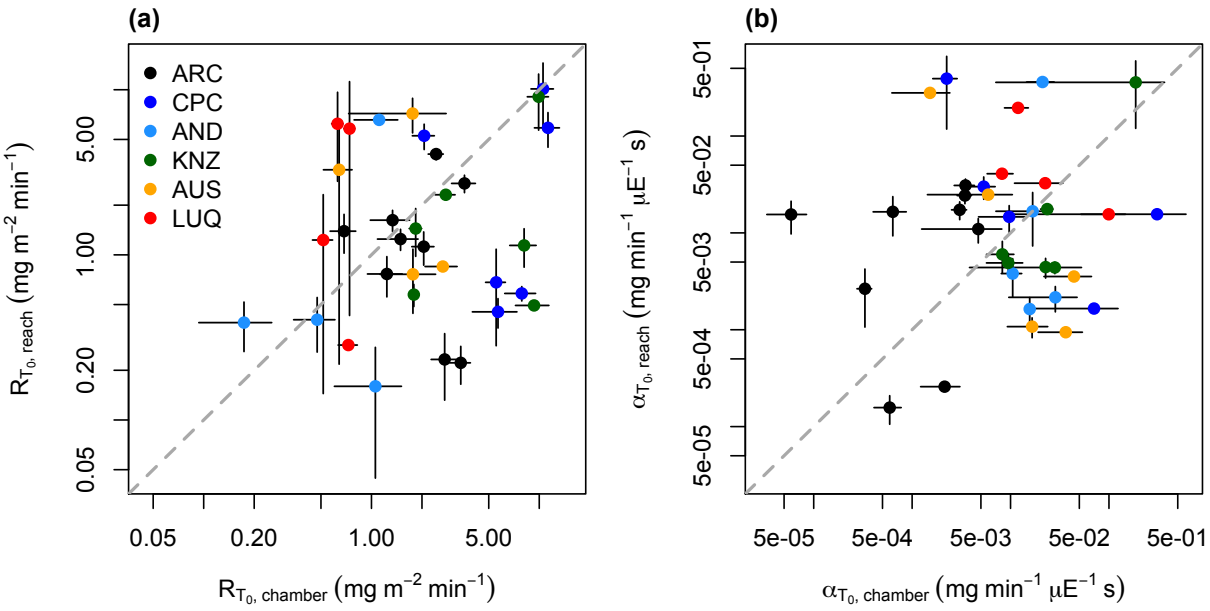

Figure 3.2: Reach and habitat scale (a) R_{T_0} and (b) α standardized to the mean temperature of chamber incubations. Axes are in logarithm scale. Error bars represent one standard error of mean. The dashed line is the 1:1 line.

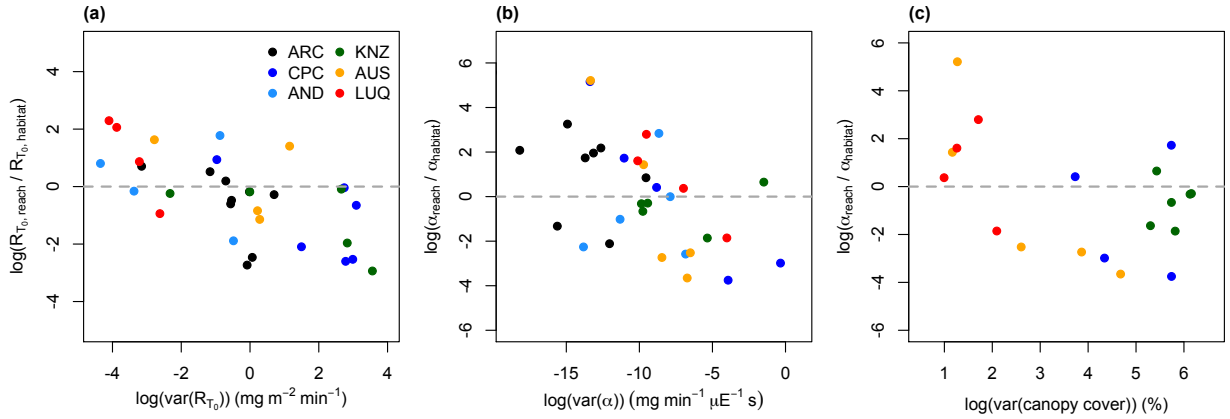

Figure 3.3: Reach–to–habitat ratios of (a) R_{T_0} and (b) α decrease with the habitat scale variance within a reach. (c) Reach–to–habitat ratio of α decreases with the variance of canopy cover.

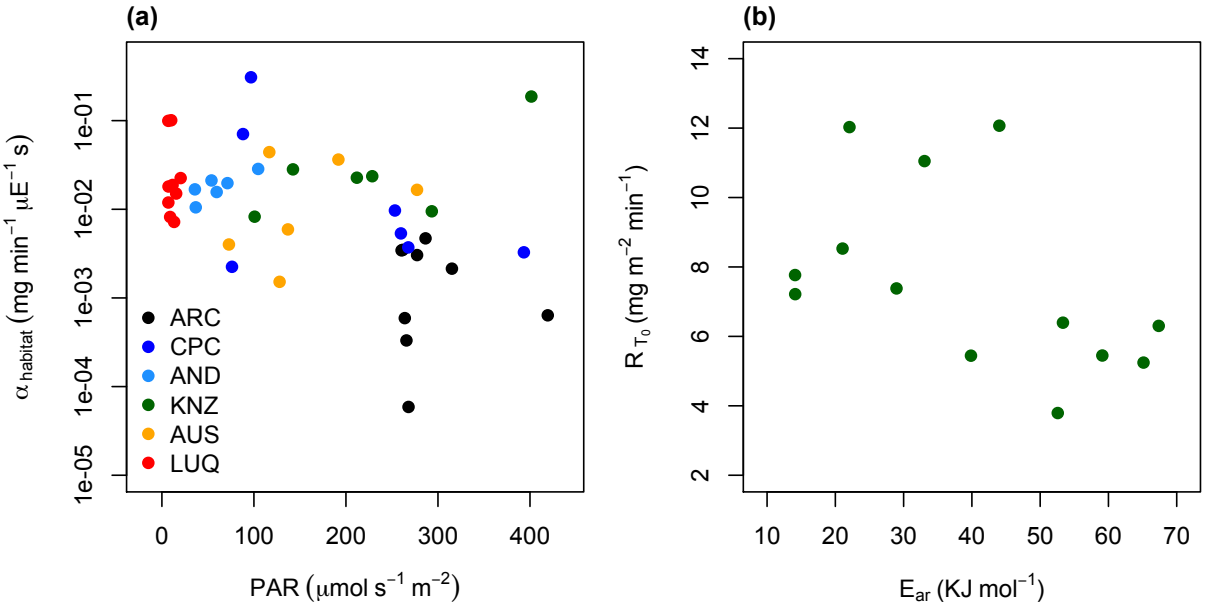

Figure 3.4: (a) Negative correlation between habitat scale α and light. Each point is the average daily light of a reach measured at one location and average habitat scale α within the reach. (b) Negative correlation between habitat scale E_{ar} and R_{T_0} . Each point in the figure is an estimate from one chamber incubation.

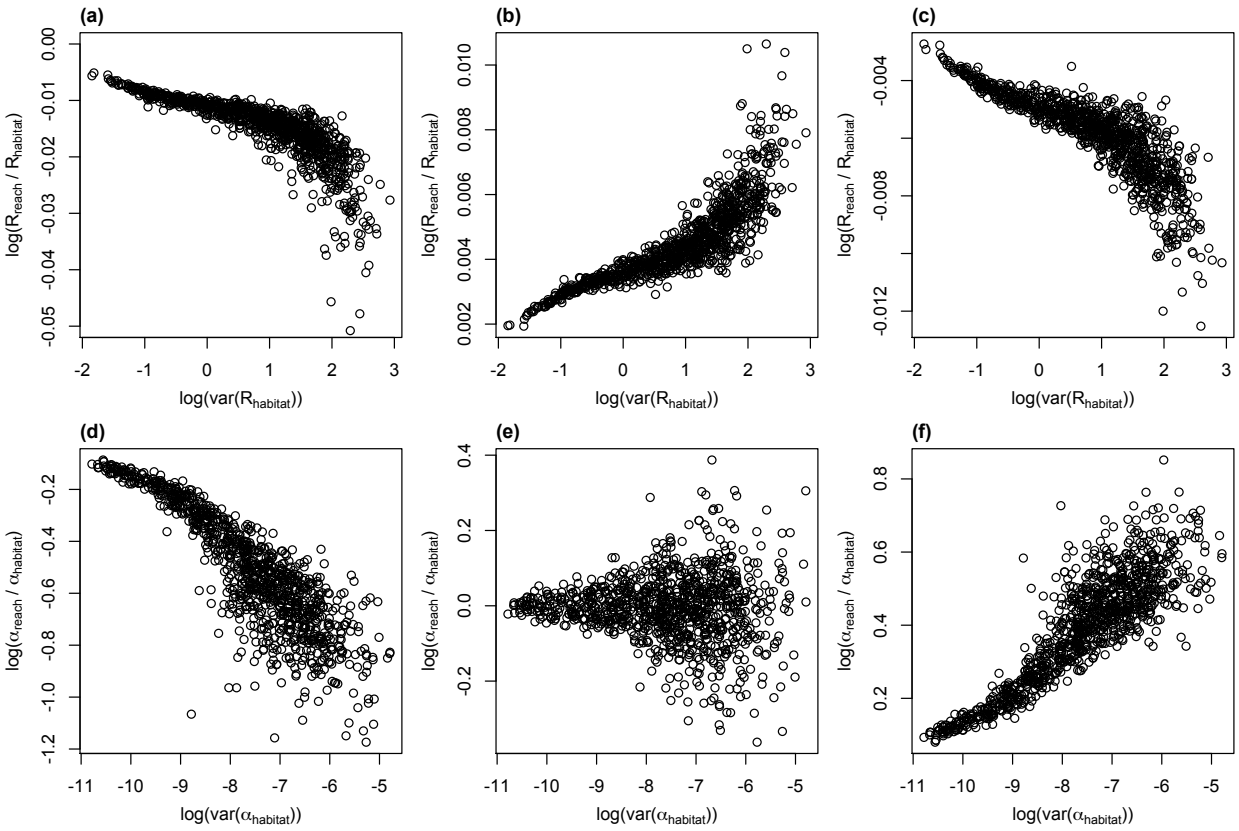

Figure 3.5: Simulation demonstrating expected relationship between habitat-to-reach ratio of R_{T_0} and the variance of habitat scale R_{T_0} when temperature for comparison (T_0) was at (a) lower bound, (b) mean, and (c) upper bound of daily temperature. Habitat scale R_{T_0} and E_{ar} were chosen independently. Expected relationship between habitat-to-reach ratio of α assuming a (d)negative, (e)no, or (f) positive correlation between habitat scale α and light.

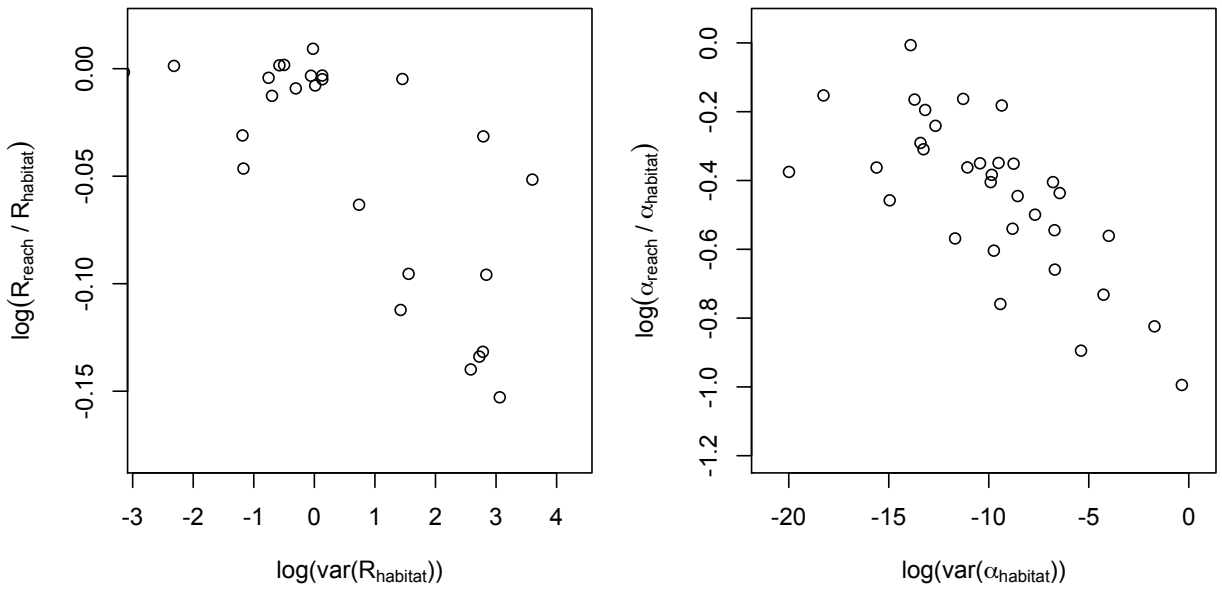

Figure 3.6: Simulation of the expected relationship between habitat-to-reach ratio of (a) R_{T_0} and (b) α based on field conditions. Reach scale light and temperature measurements were used in the simulation. Habitat scale R_{T_0} and α were randomly chosen based on the mean and variance estimated from chamber incubations. We assumed negative correlations between habitat R_{T_0} and E_{ar}, and between habitat scale α and light in the simulation.

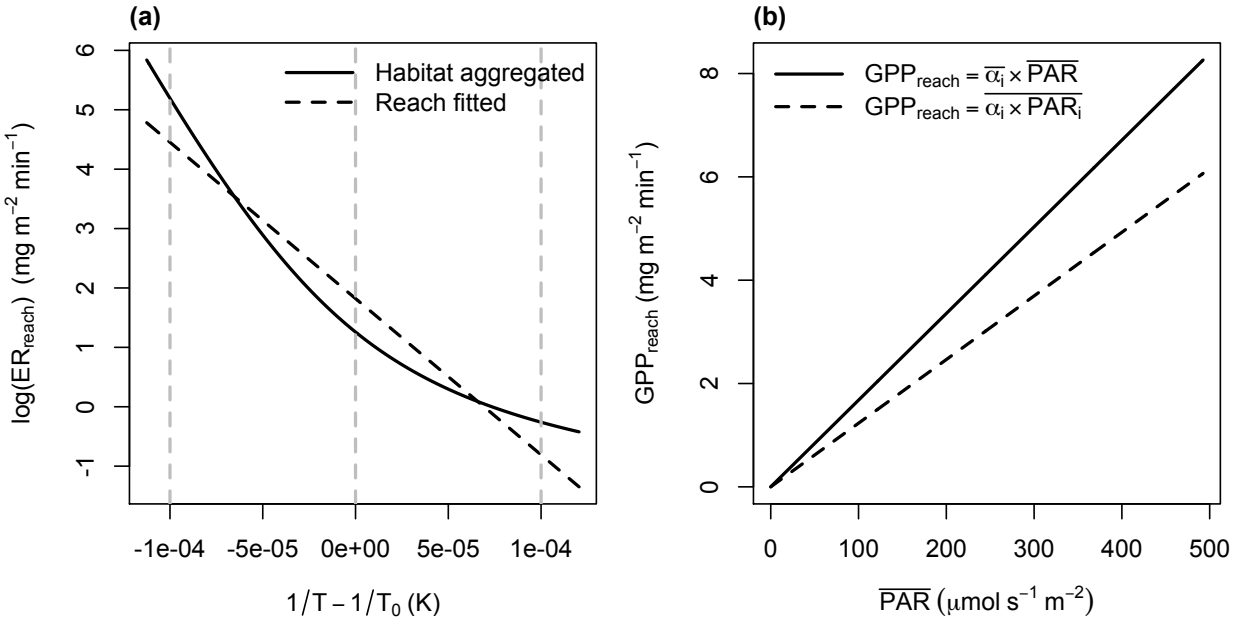

Figure 3.7: Diagram illustrating how differences in R_{T_0} and α across scales occur. (a) Reach scale ER calculated by aggregating habitat scale ER is a convex curve in Arrhenius plot. Fitting an Arrhenius equation to the reach scale ER thus cause a discrepancy between reach and habitat scale R_{T_0}. The magnitude of discrepancy depends on the curvature, which increases with the variance of habitat scale ER. (b) Reach scale GPP–light response curve calculated by aggregating habitat scale GPP differs from prediction based on $\overline{\alpha_i}$, the average habitat scale α. The diagram assumes a negative correlation between habitat scale α and light.

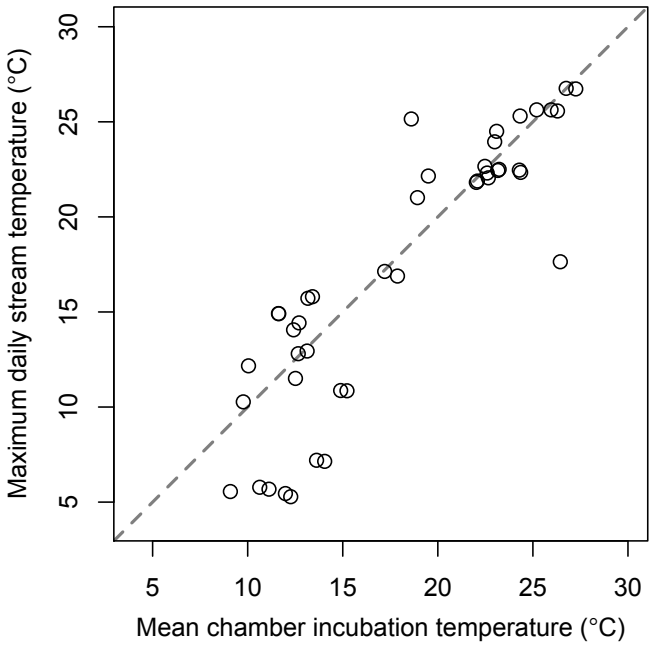

Figure 3.8: The mean temperature during chamber incubation and the maximum daily water temperature in the corresponding stream. The dashed line is the 1:1 line.

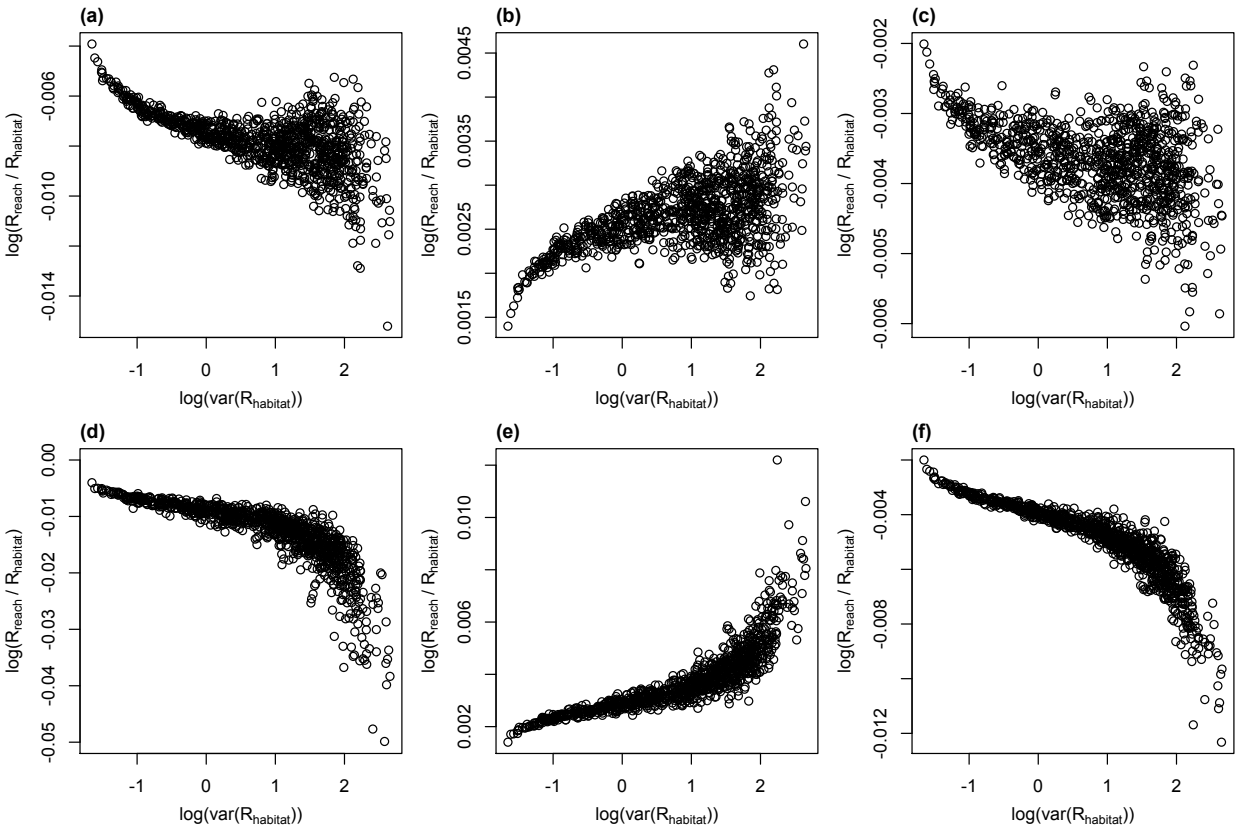

Figure 3.9: Simulation of the expected relationship between habitat-to-reach ratio of R_{T_0} and the variance of habitat scale R_{T_0} when temperature for comparison (T_0) was at (a, d) lower bound, (b, e) mean, and (c, f) upper bound of daily temperature. Panels a, b, and c show simulations assuming a positive correlation between habitat scale R_{T_0} and E_{ar}. Panels d, e, and f show simulations assuming a negative correlation between habitat scale R_{T_0} and E_{ar}.

Chapter 4

Soil respiration as an aggregated process: implications for scaling up and data interpretation

Introduction

Soil respiration releases about 80 petagrams of carbon to the atmosphere annually (Schlesinger, 1977; Raich & Schlesinger, 1992; Raich & Potter, 1995; Raich *et al.*, 2002). As the largest source of CO_2 flux from terrestrial ecosystems, soil respiration exerts a strong influence on global climate (Schlesinger & Andrews, 2000). At the same time, global warming is expected to accelerate soil respiration (Trumbore *et al.*, 1996; Rustad *et al.*, 2001; Bond-Lamberty & Thomson, 2010), creating a positive feedback between global warming and carbon release via soil respiration (Jenkinson *et al.*, 1991; Cox *et al.*, 2000). However, considerable uncertainty remains in the predicted strength of the climate-carbon feedback (Friedlingstein *et al.*, 2006; Bradford *et al.*, 2016). A key quantity to reduce such uncertainty is the temperature sensitivity of soil respiration (Jones *et al.*, 2003; Luo, 2007; Todd-Brown *et al.*, 2013; Exbrayat *et al.*, 2014). Practically, predicting the climate-carbon feedback on the global or regional scale requires a relatively simple prescription of the temperature dependence of soil respiration. However, it is well recognized that the temperature sensitivity of soil respiration is driven by a myriad of biotic and abiotic processes (Subke & Bahn, 2010; Conant *et al.*, 2011; Davidson & Janssens, 2006; Davidson *et al.*, 2006). In particular, the varied temperature sensitivities of decomposition of heterogeneous soil organic carbon pools have been considered key to an accurate description of soil carbon dynamics (Davidson *et al.*, 2000; Knorr *et al.*, 2005; Luo & Weng, 2011). Since predicting the feedback between

soil carbon dynamics and climate globally often requires simple description of the temperature dependence, a fundamental question is how the temperature sensitivity of bulk soil respiration relates to the temperature sensitivities of the decomposition of heterogeneous soil organic carbon pools.

When linking temperature sensitivities of the bulk soil respiration and the decomposition of different substrate pools, we aimed at using the temperature dependence of pools to describe the aggregate temperature dependence of bulk soil respiration. This is essentially a problem of aggregation because the bulk soil respiration is the sum of respiration from all contributing substrate pools. However, using the mean temperature sensitivity of the decomposition of all contributing substrate pools may not appropriately represents the temperature sensitivity of bulk soil respiration, due to the nonlinear nature of temperature dependence of respiration (Rastetter *et al.*, 1992). Early theoretical work extensively investigated the conditions for appropriate aggregation (Iwasa *et al.*, 1987) and the error associated with it (O'Neill & Rust, 1979; Gardner *et al.*, 1982; Iwasa *et al.*, 1989). Gardner *et al.* (1982) showed that aggregation error could be within 10% when turnover times of components to be aggregated differ up to three times. However, the turnover times of different soil carbon substrate pools differ substantially, ranging from years to hundreds of years (Trumbore *et al.*, 1996). This suggests that error introduced by aggregating respiration from multiple substrate pools could not be ignored. Moreover, these theoretical studies showed that error associated with aggregation also depends on the specific functional forms and the heterogeneity among components to be aggregated. Therefore, although the general theoretical grounds of aggregation has been extensively investigated, it is still necessary to analyze the specific form of temperature dependence and the heterogeneity among substrate pools to understand the relationship between temperature sensitivities of bulk soil respiration and decomposition of contributing substrate pools.

More generally, insights into the relationship between temperature sensitivities of bulk soil respiration and decomposition of multiple substrate pools may extend to aggregating soil

respiration over space. The fact that bulk soil respiration is the aggregation of multiple substrate pools is analogous to the fact that the total soil respiration over a large spatial extent is the aggregation of respiration from all patches within the area. Since temperature sensitivity of soil respiration is influenced by multiple biotic and abiotic factors that are spatially variable (e.g., Boone *et al.* (1998); Davidson & Janssens (2006); Davidson *et al.* (2006)), the temperature sensitivity of soil respiration is spatially heterogeneous (Qi *et al.*, 2002; Reichstein *et al.*, 2003; Scott-Denton *et al.*, 2003). As a result, the temperature sensitivity of respiration on a small scale does not directly translate to a large spatial scale.

Understanding the temperature sensitivity of aggregated soil respiration over space is practically useful. Temperature sensitivity of soil respiration is commonly measured using controlled warming experiments on a smaller spatial scale, such as measurements made with automatic chambers or incubations, but understanding the role of soil respiration in regional or global carbon cycles often requires characterization of the temperature sensitivity on a relatively large spatial extent. Even though the temperature sensitivity of soil respiration over large spatial extent can be directly quantified using eddy covariance data (Baldocchi *et al.*, 2001) and novel statistical approaches (Sanderman *et al.*, 2003; Mahecha *et al.*, 2010), such quantification is usually limited to observational studies. Small-scale manipulative experiments are necessary to elucidate the biological mechanisms underlying the observed temperature sensitivity, which are crucial for accurately predict how soil carbon dynamics will respond to a changing climate (Allison *et al.*, 2010; Wieder *et al.*, 2013). As a result, investigating how temperature sensitivity of soil respiration quantified on a small spatial scale can translate to a large spatial scale is still practically useful.

Concurrent with aggregation over space often comes aggregation over time. Practically, temporal scale usually increases with the spatial scale of measurements (Delcourt & Delcourt, 1988; Levin, 1992). For example, small spatial scale measurements of temperature sensitivity are often based on automatic chambers or incubations that span minutes to hours while earth system models usually simulate soil respiration on a daily or monthly time step (e.g., Anav

et al. (2013); Arora *et al.* (2013); Todd-Brown *et al.* (2014)). The temperature sensitivity that describes the thermal response of soil respiration on the scale of minutes or hours may not adequately describe how soil respiration over a month or year responds to the mean monthly or annual temperatures. Studies using spatial temperature gradients to estimate temperature sensitivity showed that temperature sensitivity of annual respiration was lower than the temperature sensitivity of soil respiration measured on a sub-daily time scale (Kirschbaum, 2010; Yvon-Durocher *et al.*, 2012). However, studies based on comparisons across sites may not represent the response of a single site (Lauenroth & Sala, 1992; Bradford *et al.*, 2017). For example, monthly or daily temperature in the arctic are more variable within a year compared to the tropics. Thus, when estimating temperature sensitivity from the latitudinal temperature gradient (Kirschbaum, 2010), the inferred the temperature sensitivity reflects not only the effects of differences in mean temperature, but also the effects of difference in temperature variability, which may not exist at a single site. It remains unclear how temperature sensitivity of soil respiration at a single site might depend on the temporal scale of aggregation.

Recognizing the fact that soil respiration is an aggregated process is also crucial for interpreting the observed changes in the temperature sensitivity of bulk soil respiration. Much debate in the literature on the temperature sensitivity of soil respiration stems from whether soil respiration is viewed as an aggregated process or not. For example, the constant residence time of soil organic matter or respiration rate along a temperature gradient could be interpreted as the insensitivity of soil respiration to temperature when viewing soil respiration as derived from a single pool (Liski *et al.*, 1999; Giardina & Ryan, 2000), but could also be explained by changes in the composition of carbon substrate pools, each with a unique temperature sensitivity (Trumbore *et al.*, 1996; Davidson *et al.*, 2000; Knorr *et al.*, 2005) or changes in substrates quality (Ågren, 2000; Ågren & Bosatta, 2002). Furthermore, inference on the mechanisms driving changes in temperature sensitivity may depend on how we conceptualize soil respiration. For example, changes in the temperature sensitivity of soil respiration

in response to warming could result from adaptation or acclimation of microbial community (Luo *et al.*, 2001; Karhu *et al.*, 2014), but quick depletion of labile carbon substrate may offer an equally plausible explanation (Melillo *et al.*, 2002; Kirschbaum, 2004, 2006). Different mechanisms for observed changes in temperature sensitivity may lead to vastly different predictions of soil carbon dynamics in the future. For example, modeling studies showed that explicitly incorporating the microbial processes, such as enzymatic kinetics (Allison *et al.*, 2010) or microbial carbon use efficiency (Wieder *et al.*, 2013), and their responses to changes in climate result in substantially different estimates of soil carbon stock in the future. Therefore, when making inferences on the mechanisms driving observed changes in temperature sensitivity of soil respiration, it is crucial to consider how aggregating multiple substrate pools could lead to apparent changes in temperature sensitivity, and view it as a possible mechanism to ensure a comprehensive conclusion from the data.

Although the view that soil respiration is an aggregated process is well recognized in the literature (Davidson *et al.*, 2000; Kirschbaum, 2004, 2006; Knorr *et al.*, 2005), a comprehensive analysis of how temperature sensitivity of bulk soil respiration may arise by aggregating over space, time, or heterogeneous carbon substrate pools is still lacking. Therefore in this study, we utilized numerical simulations and re-analyzed existing data sets to explore the effects of aggregation over space, time, and multiple substrate pools on temperature sensitivity. Specifically, we first used theoretical analyses and numerical simulations to identify key factors determining the effects of aggregation over space on temperature sensitivity. We then used continuously measured soil respiration data in Harvard Forest and Tibet alpine grasslands to examine how activation energy of soil respiration depends on the time scale of aggregation. Finally, we used simulations to demonstrate how temperature sensitivity of bulk soil respiration may appear to be dependent on temperature, and used a published soil incubation data set to demonstrate how viewing soil respiration as an aggregated process may provide a possible explanation for observed changes in temperature sensitivity.

Methods

Aggregation over space

We first explored how aggregation over space influenced temperature sensitivity of soil respiration analytically. The goal was to quantify the temperature sensitivity of total soil respiration over a particular spatial extent consisting of multiple patches, each with a unique respiration rate and temperature dependence. The temperature dependence of respiration for each patch can be described by an Arrhenius equation:

$$R_i = A_i e^{-\frac{E_{ai}}{RT}}, \tag{4.1}$$

where i indicates patch number, R_i is the respiration rate per area of patch i, A_i is the pre-exponential factor, E_{ai} is the activation energy, R is the ideal gas constant, and T is temperature in Kelvin. Similarly, we may use the Arrhenius equation to describing the temperature dependence of soil respiration over a large spatial extent (R_{agg}):

$$R_{agg} = A_{agg} e^{-\frac{E_{a,agg}}{RT}}, \tag{4.2}$$

where R_{agg} is the total respiration over a large spatial extent, A_{agg} and $E_{a,agg}$ are pre-exponential factor and activation energy for respiration in the spatial extent of interests respectively. After taking the logarithm of both sides in equation 4.2, we obtained:

$$\log(R_{agg}) = \log(A_{agg}) - \frac{E_{a,agg}}{R}\frac{1}{T}. \tag{4.3}$$

Equation 4.3 suggests that $E_{a,agg}$ can be estimated with a linear regression between $\log(R_{agg})$ and $1/T$. In a simple linear regression $y = a + bx$, the slope b is estimated as $\mathrm{cov}(y, x)/\mathrm{var}(x)$. Thus, $E_{a,agg}$ is estimated as

$$E_{a,agg} = \frac{\mathrm{cov}(\log(R_{agg}), \frac{1}{T})}{\mathrm{var}(\frac{1}{T})}(-R) \tag{4.4}$$

When total respiration over a large spatial extent is not directly quantified, we use small scale measurements of respiration to estimate R_{agg}. On a per area basis, the aggregated soil

respiration is the average of patch respiration. Thus, the estimate of R_{agg} from small patch respiration is

$$\widehat{R_{agg}} = \overline{A_i e^{-\frac{E_{ai}}{RT}}}. \tag{4.5}$$

Combining equation 4.4 and 4.5, $E_{a,agg}$ can be estimated as

$$E_{a,agg} = \frac{\mathrm{cov}(\log(\overline{A_i e^{-\frac{E_{ai}}{RT}}}), \frac{1}{T})}{\mathrm{var}(\frac{1}{T})}(-R). \tag{4.6}$$

When A_i and E_{ai} follow a joint log-normal distribution, we can derive an analytical expression of $E_{a,agg}$ (see appendix for derivation) as

$$E_{a,agg} = \overline{E_{ai}} + \frac{1}{2}\sigma_e \left[\rho\sqrt{\log(\frac{\sigma_A^2}{\overline{A_i}^2} + 1)} - \frac{\sigma_e}{R}\frac{\mathrm{cov}(\frac{1}{T^2}, \frac{1}{T})}{\mathrm{var}(\frac{1}{T})}\right]. \tag{4.7}$$

Here, σ_e is the standard deviation of E_{ai}, σ_A is the standard deviation of A_i, and ρ is the correlation coefficient of A_i and E_{ai}. Equation 4.7 relates the activation energy that describes the temperature dependence ($E_{a,agg}$) of large spatial scale respiration to the naive estimate if we direct extrapolate activation energy quantified on small scale ($\overline{E_{ai}}$) to the large scale.

We also performed numeric simulations to complement the theoretical analysis. To that end, we assumed a spatial extent consisting of 50 patches. We randomly chose parameter values (i.e. A_i and $E_{a,i}$) in equation 4.1 for each patch and calculated the total respiration R_{agg} as the sum of respiration from all patches (equation 4.5). We calculated the total respiration R_{agg} over a temperature range of 10 to 20 °C, and estimated $E_{a,agg}$ using a linear regression between $\log(R_{agg})$ and $\frac{1}{T}$. Specifically, for each round of simulation, A_i was chosen from a gamma distribution with a mean chosen randomly from a uniform distribution between 0.5×10^{14} and 1.5×10^{14} (μmol C m^{-2} s^{-1}) and a standard deviation uniformly chosen between 1×10^{13} and 1×10^{14}. E_{ai} was chosen randomly from a gamma distribution with mean chosen randomly from a uniform distribution between 40 and 100 (KJ/mol) and a standard deviation uniformly chosen between 10 and 50. To simulate different degrees of correlation between A_i and E_{ai}, we matched the largest A_i with the highest E_{ai} for a positive correlation or lowest E_{ai} for a negative correlation. We performed 200 rounds of simulations with a positive correlation

between A_i and E_{ai} and 200 rounds of simulations with a negative correlation between A_i and E_{ai}. In the simulations, the distributions of A_i and E_{ai} were intentionally chosen to be different from the distributional assumptions used in the theoretical derivation (equation 4.7) to demonstrate that the qualitative prediction from the theoretical derivation was not driven by the particular form of distributions of these parameters.

Aggregation over time

We used two long-term *in situ* soil respiration measurements data sets to examine the effect of aggregation over time on the temperature sensitivity of soil respiration. Both data sets contain continuous *in situ* soil respiration measurements over 3–5 years. Briefly, the Harvard Forest respiration data set includes soil temperature recorded at a depth of 5 cm and soil respiration measurements using automatic chambers at the Environmental Measurement Site in Harvard Forest (42°20'N, 72°11'W) from 2003 to 2006. The soil respiration rates and soil temperature were measured at eight fixed locations at this site. Respiration rate and soil temperature were logged every 4–5 hours from April to December. The Tibet grassland respiration data set contains respiration measurements from two locations, a mesic grassland location, and a meadow location, within the Haibei Alpine Grassland Ecosystem Research Station (37°30'N, 101°12'E) located in the northeastern part of the Tibetan Plateau, China. Respiration and soil temperature at 5 cm were recorded hourly with four automatic chambers from 2008 to 2012 at the mesic grassland location and from 2010 to 2012 in the meadow location. Respiration rates were only measured in growing season (May to September) due to the below-freezing soil temperature during the non-growing season. Only daily mean respiration for each site are available, resulting in two time series of daily respiration and temperature for analyses. More detailed descriptions of the sites and measurement protocols for Harvard Forest (Phillips *et al.*, 2010; Giasson *et al.*, 2013) and Tibet grasslands (Wang *et al.*, 2018) can be found in previously published studies.

To examine the effects of the time scale of aggregation on the estimated temperature sensitivity, we calculated the mean soil respiration rate and the mean soil temperature by averaging all soil respiration and soil temperature measurements within the period of aggregation. Specifically, to estimate activation energy describing the temperature dependence of average respiration over n days, we calculated the mean respiration and mean soil temperature over each consecutive n days, and regressed the logarithm of mean respiration rate on the inverse of mean temperature in Kelvin to obtain the activation energy. If days of measurements within a year is not a multiple of the days averaged, we excluded the remainder from the analyses. For example, if we calculated monthly mean respiration based on 182 days of soil respiration measurements within a year, we obtained 6 monthly respiration rates and excluded the remaining two days of data from the analyses. We explored how the activation energy of soil respiration changed with the number of days averaged ranging from 1 to 30 days, which correspond to the typical time steps used in earth system models to simulate carbon dynamics (Anav *et al.*, 2013; Arora *et al.*, 2013; Todd-Brown *et al.*, 2014). Because sub-daily records of soil respiration are available in the Harvard Forest data set, we also calculated temperature sensitivity based on respiration measured every 4–5 hours for the Harvard Forest data set.

We also examined the effects of aggregation over time using simulated time series of respiration based on a constant activation energy and the actual temperature data in Harvard Forest and Tibet grasslands. Because seasonal variation in temperature sensitivity (Xu & Qi, 2001; Janssens & Pilegaard, 2003) may cause deviation in temperature sensitivity when aggregating over time, the simulated data sets, where soil respiration is purely driven by temperature variation, allowed us to isolate the effects of aggregation over time. Without loss of generality, we simulated soil respiration rate using the Arrhenius equation with $R_{T_0} = 4$ μmol s^{-2} s^{-1}, $T_0 = 12$ °C and $E_a = 75$ KJ mol^{-1} based on the soil temperature measurements in the Harvard Forest and Tibet grasslands. The chosen values of the parameters represented realistic values for *in situ* soil respiration rates at these two sites.

We followed the same methods outlined above to calculate activation energy over a range of days averaged.

Aggregation over multiple substrate pools

We quantified temperature sensitivity of bulk soil respiration when the total respiration is the sum of respiration from heterogeneous substrate pools. We used first-order kinetics to model the decomposition of each carbon substrate pool as commonly used in earth system models. For a particular carbon substrate pool, the respiration rate from this pool is

$$\frac{dC_i(t)}{dt} = k_i C_i(t), \tag{4.8}$$

where $C_i(t)$ is the pool size of carbon substrate i at time t and k_i is its decomposition rate. The decomposition of all substrate pools can be represented in a matrix form:

$$\frac{d\boldsymbol{C}(t)}{dt} = \boldsymbol{A}\boldsymbol{K}\boldsymbol{C}(t). \tag{4.9}$$

Here, $\boldsymbol{C}(t)$ is a column vector describing the size of each carbon substrate pool at time t. Matrices $\boldsymbol{A}$ allows the model to account for transfers among substrate pools. For example, matrices $\boldsymbol{A}$ and $\boldsymbol{C}$ for a three–pool with transfer model are given by

$$\boldsymbol{A} = \begin{bmatrix} -1 & 0 & 0 \\ f_{21} & -1 & 0 \\ 0 & f_{32} & -1 \end{bmatrix}, \ \boldsymbol{K} = \begin{bmatrix} k_1(T) & 0 & 0 \\ 0 & k_2(T) & 0 \\ 0 & 0 & k_3(T) \end{bmatrix}. \tag{4.10}$$

The elements f_{ij} in matrix $\boldsymbol{A}$ denote the transfer from pool j to i. Matrix $\boldsymbol{K}$ represents the first order decomposition rates. T is temperature. We modeled the temperature dependence of k_i with Arrhenius equations

$$k_i(T) = k_i(T_0) e^{-\frac{E_{ai}}{R}(\frac{1}{T} - \frac{1}{T_0})}. \tag{4.11}$$

We can modify matrices $\boldsymbol{A}$ and $\boldsymbol{K}$ to represent other models. For example, a two pool model can be represented by using the first two rows and columns of matrices $\boldsymbol{A}$ and $\boldsymbol{K}$. A

model without transfer among substrate pools can be represented by setting all f_{ij} to 0 in matrix $\boldsymbol{A}$. The system of differential equations represented by equation 4.9 can be solved as

$$\boldsymbol{C}(t) = e^{\boldsymbol{A}\boldsymbol{K}(T)t}\boldsymbol{C}(0), \tag{4.12}$$

and the total respiration rate at time t can be calculated as

$$\mathbf{1}^T \boldsymbol{R} e^{\boldsymbol{A}\boldsymbol{K}(T)t}\boldsymbol{C}(0), \tag{4.13}$$

where $\mathbf{1}$ is a vector with all elements 1, and $\boldsymbol{R}$ is a matrix for the respiration rate. For a three pool with transfer model, $\boldsymbol{R}$ is

$$\boldsymbol{R} = \begin{bmatrix} k_1(T)(1-f_{21}) & 0 & 0 \\ 0 & k_2(T)(1-f_{32}) & 0 \\ 0 & 0 & k_3(T) \end{bmatrix}. \tag{4.14}$$

After establishing the mathematical representation of the multiple substrate models (equations 4.8–4.12), we performed numeric simulations to explore how temperature sensitivity of bulk soil respiration depends on temperature. Briefly, we assumed 50 distinct carbon substrate pools and randomly chose values for $k_i(T)$, f_{ij}, and E_{ai} in equations 4.10 and 4.11 to describe the decomposition dynamics of each substrate pool. Without loss of generality, we randomly chose f_{ij} from a uniform distribution (0, 0.5), E_{ai} from a uniform distribution (20, 150) (KJ/mol), and k_i at 10 °C from a uniform distribution (1/3650, 1/365) (day^{-1}). We then simulated the pool size over 730 days according to equation 4.12 and calculated respiration rate over time as in equation 4.13. We performed such simulation under four temperatures (10, 12, 14 and 16 °C). At any particular time, we calculated three activation energies for total respiration based on different range of temperature: one by comparing 10 and 12 °C, one by comparing 12 and 14 °C, and one by comparing 14 and 16 °C. We quantified the activation energy over time, resulting in 3 time trajectories of activation energies for total respiration, each quantified based on specific range of temperature. Because soil respiration can be expressed in different units, we quantified activation energy based on

respiration expressed in the unit of total respiration from all substrate pools and respiration per unit carbon.

Integration of aggregation over space and substrate pools

The theoretical analysis (equation 4.7) showed that spatial variability of the temperature dependence of respiration for each patch is critical for determining the difference between patch average activation energy and the aggregated activation energy. When respiration of each patch is the results of respiration from multiple substrate pools, changes in the composition of substrate pools, which may occur with increases in temperature (Kirschbaum, 2004), could alter the spatial variability of the temperature dependence of patches and subsequently influence aggregation over space. We performed a simulation study here to investigate how warming may influence the difference in activation energy when aggregating over space by changing the substrate composition of the patches.

We assumed a space consisting of 50 patches. The respiration from each patch is derived from 10 substrates with distinct pool size, decomposition rate, and activation energy. Further, we assumed that the input to each substrate pool at the current temperature (10 °C) balanced the respiratory loss so that the carbon substrate pools were at steady state at the current temperature. The inclusion of an input in the model allowed us to simulate the long-term dynamics of soil carbon. Without input, all carbon in the soil will be eventually respired. We then increased the temperature to 15 °C and simulated the carbon pools dynamics for each patch. Mathematically, the dynamics of carbon pools within each patch can be described as

$$\frac{d\boldsymbol{C}(t)}{dt} = \boldsymbol{AK}\boldsymbol{C}(t) + \boldsymbol{I}, \tag{4.15}$$

where $\boldsymbol{A}$, $\boldsymbol{K}$, and $\boldsymbol{C}(t)$ are defined the same way as in equation 4.9. $\boldsymbol{I}$ is a vector of input rates to each carbon substrate pool and is assumed to be constant over time in the simulation. Equation 4.15 can be solved analytically as

$$\boldsymbol{C}(t) = e^{\boldsymbol{AK}t} + e^{\boldsymbol{AK}t} * \boldsymbol{I}, \tag{4.16}$$

where $e^{\boldsymbol{AK}t} * \boldsymbol{I}$ denotes the convolution

$$e^{\boldsymbol{AK}t} * \boldsymbol{I} = \int_0^t e^{\boldsymbol{AK}(t-u)} \boldsymbol{I} du. \tag{4.17}$$

For a constant $\boldsymbol{I}$, equation 4.17 evaluates to $\boldsymbol{U}\boldsymbol{\kappa}$, where $\boldsymbol{U}$ is a matrix with eigenvectors of $\boldsymbol{AK}$ as the columns and $\boldsymbol{\kappa}$ is a column vector with element i

$$[(\boldsymbol{U}^{-1}\boldsymbol{I})_i]\frac{e^{-k_i t} - 1}{-k_i}, \tag{4.18}$$

When specifying equation 4.15 for each patch, we randomly chose decomposition rate k_i, transfer coefficient between substrate pools f_{ij}, and activation energy E_{ai} for carbon pools within each patch. Specifically, we chose k_i from a uniform distribution between 1/36500 and 1/365 day^{-1}. The choice of decomposition rates reflects the range of residence time of organic carbon typically observed in soils (Trumbore *et al.*, 1996). We chose transfer coefficient from a uniform distribution between 0 and 0.5. The initial substrate pool size was chosen from a uniform distribution (1, 20) mg g^{-1}. E_{ai} was chosen from a gamma distribution, where the mean was randomly chosen from a uniform distribution (80, 120) KJ mol^{-1} and the standard deviation was randomly chosen from a uniform distribution (45, 75) KJ mol^{-1} for each iteration of simulation. This is to generate the distribution for E_{ai} with different skewness in different iterations of simulation. We simulated the carbon pool dynamics for 500 years. At each time point during the simulation, we calculated activation energy for each patch following the same methods outlined in the aggregation over multiple substrate pools and used the same procedure as outlined in the aggregation over space to calculate aggregated activation energy over space. We performed the simulation for 100 iterations.

Fitting multi-pool models to soil incubation data

We demonstrate how viewing soil respiration as an aggregated process could influence the inferred mechanisms driving changes in temperature sensitivity by analyzing a data set from a soil incubation experiment. Karhu *et al.* (2014) used a novel approach, cooling and rewarming soils over the course of incubation, to infer the mechanism behind warming-driven changes

in temperature sensitivity. Briefly, they sampled soils around the globe. Twenty replicates of soil sampled from each particular location were incubated at a control temperature set at 3 °C above the mean annual temperature of the sampling location. After 84 days, ten replicates were cooled by 6 °C from the control temperature. After another 60 days, five of the ten cooled soil replicates were rewarmed to the control temperature. Soil respiration rates were measured weekly initially and biweekly later in the experiments. They compared the respiration rate at the same amount of cumulative respiratory carbon loss in control, cooled and rewarmed soils to infer mechanisms driving the changes in temperature sensitivity.

Karhu *et al.* (2014) established two criteria for defining an enhancing microbial community response to warming: 1) a faster proportional decrease of respiration rate with cumulative carbon loss for cooled soils, and 2) lower respiration rates of rewarmed soils compared to the control at the same amount cumulative carbon loss. The opposite is defined as a compensatory response, and no difference between cooled, rewarmed, and control soils regarding the two criteria were defined as no response. Although the majority of the soils exhibited enhancing responses, we selected data from 9 soils randomly from each category of response for analyses.

To explore whether the observed changes in temperature sensitivity of bulk soil respiration in this experiment can be explained by viewing soil respiration as an aggregation of respiration from multiple substrate pools, we fit multi-pool models described by equations 4.9 and 4.10 to the data. Because the incubation experiment only had two temperature treatments, we only considered two candidate models for the soil incubation data: two–pool and two–pool with transfer (Liang *et al.*, 2015). Each model can be specified by varying matrices $\boldsymbol{A}$ and $\boldsymbol{K}$ in equation 4.9. For the two–pool without transfer model, we specify the transfer coefficients f_{ij} in matrix $\boldsymbol{A}$ as 0.

We employed a Bayesian approach to estimate parameters in the models with various structures of carbon substrate pools for its convenience in dealing with the identifiability issues and constraining ranges of parameters in model fitting. For a particular set of param-

eters in each model, we first calculated the size of each carbon substrate pool over time according to equation 4.12. Specifically, for soils in control temperature throughout the experiment, we specified matrix $\boldsymbol{A}$ and $\boldsymbol{K}$ as in equation 4.10 to calculate the carbon pool size over time. For soils that were cooled after 84 days in the control temperature, we calculated carbon pool size at the end of 84 days. The carbon pool size on day 84 was the initial carbon pool size for the cooling treatment. We then used the decomposition rate in the cooled temperature (equation 4.11) to calculate the carbon pool sizes over time at the cooled temperature. For soils that were rewarmed following the cooling treatment, we calculated the carbon pool size at the end of the cooling treatment as the initial carbon pool size for the rewarming treatment. Since the soils were rewarmed to the control temperature, we calculated the carbon pool size following the rewarming of soils using the same k_i as the control soils. After calculating the carbon pool sizes over time, we obtained the respiration rate over time as in equation 4.13. We assumed that the differences between modeled and measured soil respiration rates were independent and identically distributed random errors following a normal distribution. With such an assumption of error distribution, we calculated the likelihood for a particular set of parameters. We used uniformly distributed improper priors for all parameters. We sampled the posterior distributions of the parameters using Markov Chain Monte Carlo and reported the mean of the posterior distribution as the point estimates for all the parameters. We implemented an adaptive Metropolis Hasting algorithm for efficient mixing (Haario *et al.*, 2001). We examined the trace plots visually and performed Geweke diagnostic to ensure convergence. The adaptive Metropolis–Hasting algorithm was implemented using function `metrop` in R 3.4.2 (R Core Team, 2017)

Results

Aggregation over space or time

The theoretical analysis (equation 4.7) showed that aggregation over space changed the temperature sensitivity. Specifically, the difference between activation energy on a larger spatial

extent ($R_{a,agg}$) and the average patch activation energy ($\overline{E_{ai}}$) depends on the variability of A_i, E_{ai} and the correlation between A_i and E_{ai} (ρ). When ρ is negative, the activation energy over large spatial extent is lower than the average patch activation energy. When ρ is positive, and the coefficient of variation of A_i is relatively big compared to the variation of E_{ai}, the activation energy over large spatial extent could be higher than the average patch activation energy. The numeric simulation confirmed the results from the theoretical analysis (Fig. 4.1). Depending on the particular values of ρ, σ_A, σ_e, the difference between average patch activation energy and the activation energy over large spatial extent could be substantial (Fig. 4.1).

We found that aggregation over time span ranging from hours to 30 days did not systematically change the activation energy in both the Harvard forest (Fig. 4.2a) and Tibet grasslands data sets (Fig. 4.2b). Although activation energy appeared to decrease slightly with the number of days averaged based on the simulated data sets (Fig. 4.2c, d), such decreasing trend was not significant considering the minor magnitude of decrease and the uncertainty in the estimated activation energy.

Aggregation over multiple substrate pools

The simulation showed that when soil respiration is the result of respiration from multiple substrate pools, the activation energy of total soil respiration depends on the range of temperature used to quantify the temperature sensitivity (Fig. 4.3). But the direction of the temperature–activation energy relationship depends on the unit in which soil respiration is expressed. When respiration is expressed as total respiration from all substrate pools, activation energy increases with the temperature used to quantify it. In contrast, when respiration is in the unit of respired carbon per unit soil carbon, activation energy is likely to decrease with the temperature used to quantify it. This simulation showed that an apparent temperature dependence of activation energy could arise when soil respiration is the result of aggregation over multiple substrate pools.

Aggregation over space and multiple substrate pools

When examining how aggregation over multiple substrate pools could influence the aggregation over space, we found that differences in the activation energy and the average patch activation energy when aggregating over space showed a similar pattern as when the substrate pool dynamics in each patch is not explicitly considered; the aggregated activation energy was mostly lower than the patch average activation energy. Both the aggregated activation energy and the patch average activation energy increased with warming, but the difference between aggregated activation energy and patch average activation energy remained roughly the same over time (Fig. 4.4).

Fitting multi–pool models to incubation data

We fit a two-pool model and a two–pool without transfer model to the soil incubation data in Karhu *et al.* (2014). While the two–pool model can be fit with uniquely identifiable parameters, the transfer coefficients (i.e. f_{21}) in the two-pool with transfer model cannot be estimated practically, primary due to the limited changes in soil carbon stock from a relatively short term incubation (Subke & Bahn, 2010). Thus, we only presented the results from the two-pool model (Table 4.1).

The two-pool model provided a visually good fit to the incubation data (Fig. 4.5). Unique patterns of parameters in the two-pool model emerged for soils exhibiting different types of responses. Specifically, soils exhibiting the enhancing responses typically had the fast decomposing pool (i.e. high k_i) with lower E_{ai}. Soils consisting of two pools with roughly the same activation energy often exhibited no response. In contrast, soils exhibiting the compensatory response typically had the fast decomposing pool with high E_{ai}.

Discussion

Effects of aggregating over space

We showed through theoretical analysis and numerical simulations that the temperature sensitivity of bulk soil respiration cannot always be represented by the average temperature sensitivity of contributing patches when aggregating over space. When the pre-exponential factor and temperature sensitivity of patches are spatially variable, the activation energy of total respiration could be either higher or lower than the average activation energy from all patches, depending primarily on the direction of correlation between the pre-exponential factor and activation energy for patches (Fig. 4.1). Although the theoretical analyses (equation 4.7) relied on an additional assumption about form of the joint distribution of A_i and E_{ai} (see appendix), the numeric simulation that did not made such assumption showed the same qualitative results about the effects of aggregation over space as the theoretical analysis (Fig. 4.1). The consistency between the numeric simulation and theoretical analyses suggests that the additional distribution assumption is not responsible for the effects of aggregation over space.

We suggest that the temperature sensitivity of soil respiration over a large spatial extent should be predominantly lower than that in the corresponding contributing patches. The theoretical analysis showed that when A_i and E_{ai} are negatively correlated and spatially variable, aggregation over space leads to smaller activation energy over a large spatial extent compared to the average patch activation energy. Both A_i and E_{ai} were known to be spatially heterogeneous at the plot scale (Xu & Qi, 2001; Qi *et al.*, 2002; Scott-Denton *et al.*, 2003). In addition, in incubation studies on temperature sensitivity of substrates with varying qualities, respiration rate was often negatively correlated with activation energy (Fierer *et al.*, 2005; Hartley & Ineson, 2008; Conant *et al.*, 2008; Craine *et al.*, 2010). If this pattern extends to the spatial context, where patches with higher A_i have lower activation energy, we would

expect lower activation energy on a large spatial extent than the average activation energy in all contributing patches.

Lower activation energy as a result of aggregation over space may explain why activation energy estimated for a larger spatial extent tended to be smaller. On a relatively large spatial extent, Mahecha *et al.* (2010) reported an average Q_{10} of 1.4 based on eddy covariance data. Zhou *et al.* (2009) used a model inversion approach and estimated an average Q_{10} of 1.72 for all the 1 °by 1 °grids globally. At the global mean temperature around 15 °C, these values of Q_{10} are equivalent to activation energies of 23.5 and 37.9 KJ/mol respectively. In contrast, on a relatively small spatial scale, Yvon-Durocher *et al.* (2012) synthesized field warming and incubation studies and estimated an average activation energy of 0.65 eV (62.7 KJ/mol) for soil respiration. Although differences associated with methodologies, such as using soil or air temperature (Xu & Qi, 2001; Graf *et al.*, 2008; Kirschbaum, 2010; Phillips *et al.*, 2011) and inclusion of aboveground respiration in eddy flux measurements (Graf *et al.*, 2011), may contribute to the different estimates of activation energy in these studies, our study demonstrated that aggregation over space may also be a source of differences in activation energy on different spatial scales. The simulation procedure we used in this study demonstrated one possible way to account for the difference caused by aggregation over space and calculate activation energy on a large spatial scale based on multiple small-scale measurements.

Effects of aggregation over time

We showed that aggregation over time did not significantly change the temperature sensitivity of soil respiration. We did not detect a systematic trend of estimated temperature sensitivity over the time span of aggregation in either field observations (Fig. 4.2a, b) and simulated data sets (Fig. 4.2c, d). Our findings offered evidence that temperature sensitivity quantified on the scale of hours-to-days might be directly applied to describe the temperature dependence of soil respiration on the scale of weeks to months, which are typically the

time steps for simulating carbon dynamics in the earth system models (Anav *et al.*, 2013; Arora *et al.*, 2013; Todd-Brown *et al.*, 2014).

Our findings are in contrast to previous studies that showed a lower temperature sensitivity on longer time scales (Kirschbaum, 2010; Yvon-Durocher *et al.*, 2012). The contrasting findings may arise from the different methods used to quantify temperature sensitivity. The activation energy of soil respiration in these studies was derived from comparing respiration rates over a spatial temperature gradient. Over space, locations with higher mean temperature have lower temperature variability (Mearns *et al.*, 1984; Kirschbaum, 2010). This means that colder locations experience relatively longer warm periods compared to warmer locations. Consequently, when comparing the average respiration over a period of time, the difference in respiration between colder and warmer locations is reduced by the larger variability of temperature at the colder locations (Savage, 2004). However, the negative correlation between mean and variability of temperature does not exist for a single location. At a single location, warmer weeks or months do not necessarily have lower temperature variability. As a result, when using single site data to examine the dependence of activation energy on time scale of aggregation, we did not observe a changing activation energy with the time span of averaging. More generally, our findings raise questions about the use of comparative analyses across sites to parameterize ecosystem models and make predictions about single sites. Sites often differ in multiple biotic and abiotic conditions, making it practically difficult to isolate the effect of any one particular factor. Growing evidence suggests that equating cross-site patterns to single-site patterns may lead to erroneous predictions (Lauenroth & Sala, 1992; Bradford *et al.*, 2014; Waring *et al.*, 2016; Bradford *et al.*, 2017).

Our findings demonstrated that aggregation over time itself does not systematically change the temperature sensitivity. However, temperature sensitivity of soil respiration could depend on the time span of aggregation for other reasons, especially on long time scales. For example, temperature sensitivity of annual total respiration may reflect the temperature sensitivity of primary production because long-term respiration could be constrained by the

total amount of substrates produced from photosynthesis in terrestrial ecosystems (Vargas *et al.*, 2010; Yvon-Durocher *et al.*, 2012; Giardina *et al.*, 2014). Respiration on a longer time scale may also exhibit stronger temperature dependence due to the faster accumulation of resources necessary for respiration over the relatively longer time span (Anderson-Teixeira *et al.*, 2008). Although we showed no effects of aggregating over time spans ranging from weeks to months on temperature sensitivity, cautions should be taken to examine whether other biological processes may constrain respiration and subsequently influence temperature sensitivity, especially on annual or longer time scales.

Effects of aggregating over multiple substrate pools

In the simulations demonstrating the effects of aggregating over multiple substrate pools, we quantified activation energy by comparing soils that had been exposed to different temperatures for some time. This mimics the common way by which temperature sensitivity of soil respiration is quantified in the literature (e.g., Lloyd & Taylor (1994); Yvon-Durocher *et al.* (2012)). We showed through simulations that activation energy of total soil respiration depends on the range of temperature used to quantify it (Fig. 4.3), even if the parameters governing the decomposition kinetics of each substrate pool remain unaffected by temperature. This occurs because warming not only alters the relative proportion of different substrate pools, but also changes the pool size for each substrate. As a result, expressing respiration in different units leads to different activation energy–temperature relationship (Fig. 4.3). The different units we used in the simulation to express temperature sensitivity corresponds to commonly used units for soil respiration. For example, respiration per unit soil carbon is often used in incubation studies (Karhu *et al.*, 2014). Respiration per area (Lloyd & Taylor, 1994) or respiration per unit weight of soil (Hartley & Ineson, 2008), which are equivalent to total respiration from all substrate pools, are also common in field measurements and incubation studies.

Our finding that the temperature–activation energy relationship could depend on the unit of soil respiration may reconcile the apparently contradictory findings from previous studies. For example, several synthesis studies showed that temperature sensitivity of soil respiration decreased with temperature (Lloyd & Taylor, 1994; Chen & Tian, 2005). But increasing temperature sensitivity with warming has also been observed in a global scale incubation experiment (Karhu *et al.*, 2014). These studies differed in the units of soil respiration. Lloyd & Taylor (1994); Chen & Tian (2005) used respiration per area to quantify temperature sensitivity while Karhu *et al.* (2014) estimated temperature sensitivity from respiration per unit soil carbon. Our simulation study demonstrated that the opposite activation energy–temperature relationship might arise simply due to the different units of soil respiration used in the study and are the results of aggregation over multiple substrate pools.

Integration of aggregation over space and substrate pools

We found that considering aggregation over multiple substrate pools for each patch does not change how aggregation over space influences the activation energy of respiration. The patch activation energy, which is aggregated over multiple substrate pools, increased during warming (Fig. 4.4). This is to be expected based on the kinetic property of soil respiration prescribed by the Arrhenius equation (Sierra, 2012). Warming often leads to the depletion of carbon substrates with low activation energy (Kirschbaum, 2004) and therefore result in a higher average activation energy (Craine *et al.*, 2010). But because warming causes similar shift in the composition of substrate pools and thus the activation energy for all patches, it does not fundamentally alter the spatial variability of activation energy and the correlation between activation energy and pre-exponential factors in the Arrhenius equation. As a result, how aggregation over space changes the activation energy of soil respiration is not influenced by warming. This result suggests that the differences in the activation energy of soil respiration caused by aggregation over space will likely remain in a warming world.

Data interpretation: the importance of considering soil respiration as an aggregated process

We demonstrated that recognizing soil respiration as an aggregated process is important for interpreting observed changes in the temperature sensitivity of soil respiration. While Karhu *et al.* (2014) suggested microbial community changes as the mechanisms driving the enhanced temperature sensitivity under warming, we showed that such observation might occur without invoking any microbial mechanisms. A model with two carbon substrate pools, each with a distinct temperature sensitivity, provided an equally adequate fit to the observed respiration rates (Fig. 4.5) and thus offered an equally plausible explanation for the enhanced temperature sensitivity of bulk soil respiration under warming.

The fundamental difference between the two interpretations stems from viewing soil respiration as derived from a single pool or multiple pools. Karhu *et al.* (2014) established two criteria for defining an enhancing microbial community response to warming: 1) a faster proportional decrease of respiration rate with cumulative carbon loss for cooled soils, and 2) lower respiration rates of rewarmed soils compared to the control at the same amount cumulative carbon loss. They concluded that microbial community responses predominantly enhanced temperature sensitivity under warming. Such conclusion is extricably tied to the dynamic relationship between available carbon pool size and respiration rate prescribed by the Q model (Ågren & Bosatta, 1996). The Q model is a single carbon substrate pool model that includes the effects of substrate quality on decomposition kinetics. For the Q model to exhibit the enhancing response, a change in one or more parameters that were synonymous with microbial community responses must occur.

We showed that the pattern of soil respiration following the cooling and rewarming treatments could also be generated by a two-pool model without transfer among pools (Fig. 4.5). Specifically, the enhancing responses, which were observed for the majority of soils in the experiment, could occur if the substrate with higher activation energy has lower respiration rate (i.e. $E_{a1} > E_{a2}$ and $k_1(T_{control}) < k_2(T_{control})$) (Table 4.1). The reason such

responses arose is that substrate with higher activation energy experienced a larger proportional decrease in decay rate when cooled (Davidson & Janssens, 2006; Sierra, 2012). As a result, the proportion of respiration derived from substrates with higher activation energy and slower decomposition rate was lower, and the proportion of these substrates in total soil organic matter was greater in the cooled soils than the control. Because respired soil carbon was derived more from the fast decomposing substrates under cooling treatment, soils in the cooling treatment experienced a more rapid proportional decrease in total respiration rates as soil carbon got respired and lower respiration rates following rewarming. Evidently, the negative correlation between activation energy and reaction rate is key for generating the enhancing response in the two-pool model. Such negative correlation is consistent with previous studies based on enzyme assays (Lehmeier *et al.*, 2013) and soil incubations (Fierer *et al.*, 2005; Karhu *et al.*, 2010; Craine *et al.*, 2010). Thus, our findings showed that the two pool model, with realistic parameters consistent with previous research, offered a plausible explanation for the observed enhancing responses in this incubation experiment.

Whether viewing soil respiration as an aggregated process or not has generated a lot of debate in the literature on what drives the observed changes in temperature sensitivity of soil respiration. When soil respiration is viewed as derived from a single pool, changes in temperature sensitivity is often interpreted as an indication of changes in the decomposer community. In contrast, when viewing soil respiration as an aggregated process, changes in the relative proportion of substrate pools may explain variation in the temperature sensitivity of respiration. For example, Luo *et al.* (2001) attributed the warming-induced decrease in temperature sensitivity of bulk soil respiration to acclimation while Kirschbaum (2004) suggested that fast depletion of labile substrates in response to warming was an equally plausible mechanism. Constant decomposition rate over a range of temperatures may be interpreted as the insensitivity of organic matter decomposition to temperature (Giardina & Ryan, 2000), but could also occur when multiple substrates with distinct temperature sensitivities simultaneously decompose (Davidson *et al.*, 2000; Knorr *et al.*, 2005).

Although viewing soil respiration as derived from one pool or multiple pools may both reasonably explain the same pattern of temperature sensitivity observed in an experiment, the different inferred mechanisms could lead to considerable differences in the predicted soil carbon dynamics under future climate scenarios. For example, modeling studies showed that explicitly incorporating how enzymatic kinetics and microbial carbon use efficiency respond to warming leads to substantial differences in predicted soil carbon stocks compared to the prediction purely based on the kinetic properties of soil carbon substrates (Allison *et al.*, 2010; Wieder *et al.*, 2013, 2015; Luo *et al.*, 2016). Therefore, it is important to make comprehensive and robust inferences on what processes drive the variations in temperature sensitivity. As our study demonstrated, it is at least necessary to consider a multi-pool model as a candidate model describing the temperature sensitivity of soil respiration to adequately access the confidence and uncertainty in the inferred mechanisms. In addition, our findings suggest that respiration data along are unlikely to be sufficient to uniquely identify whether changes in temperature sensitivity of soil respiration is driven by microbial mechanisms or changes in substrates composition, even in a specifically designed experiment as in Karhu *et al.* (2014). Experimental approaches to quantifying soil microbial dynamics and the kinetic properties of biochemical processes involved in soil respiration could be key to elucidating mechanisms driving the temperature dependence of soil respiration (Billings *et al.*, 2015).

Conclusions

In conclusion, we demonstrated through numerical simulations and analyses of existing data sets the importance of viewing soil respiration as an aggregated process. We showed that aggregation over space influences activation energy. In particular, the spatial variability in A_i and E_{ai} for the Arrhenius equation describing small patch scale respiration, and the correlation between A_i and E_{ai} determine the differences between aggregated activation energy and the small scale average activation energy. Practically, this suggests that quantifying the

spatial variability of patch scale parameters (A_i and E_{ai}) in the Arrhenius equation is critical when trying to use small scale measurements to infer the temperature sensitivity on a large scale. The theoretical analyses provide a quantitative method to relate small scale measurements to large scale temperature sensitivity. We further showed that incorporating the dynamics of multiple substrate pools does not influence the effects of aggregation over space, suggesting that the difference in activation energy when aggregating over space will likely remain the same in a warming world. We also showed that aggregation over time does not create a consistent trend of deviation in activation energy. This means that activation energy quantified based on short term measurements could be directly applied to describing the temperature dependence on longer time scales. Finally, explicitly considering the kinetics of heterogeneous soil carbon substrate pools may not only explain commonly observed relationship between temperature sensitivity and temperature, but also influence the interpretation of observed changes in temperature sensitivity in experiments. Collectively, these findings suggest that soil respiration is an inherently aggregated process. Recognizing such fact as a candidate conceptual model when designing experiments and interpreting experimental data is important.

Appendix

Derivation of activation energy when aggregating over space

We first state the following properties that establish the relationship of mean or variance between normal and log-normal distributions. These properties were used in the derivation of activation energy when aggregating over space. First, if random variable x follows a normal distribution $N(\mu_x, \sigma_x^2)$, e^x follows a log-normal distribution with mean $e^{\mu_x+\frac{\sigma_x^2}{2}}$ and variance $(e^{\sigma_x^2}-1)e^{\mu_x+\frac{\sigma_x^2}{2}}$. Second, if random variable y follows a log-normal distribution with mean μ_y and variance σ_y^2, $\log(y)$ follows a normal distribution with mean $\log(\mu_y) - \frac{1}{2}\log(\frac{\sigma_y^2}{\mu_y^2}+1)$ and variance $\log(\frac{\sigma_y^2}{\mu_y^2}+1)$. Third, if variable x and y follow a joint normal distribution, any linear combination of x and y, $ax+by$, follows a normal distribution $N(a\mu_x + b\mu_y,\ a^2\sigma_x^2 + b^2\sigma_y^2 + ab\rho_{xy}\sigma_x\sigma_y)$, where μ_x is the mean of x, μ_y is the mean of y, σ_x is the standard deviation of x, σ_y is the standard deviation of y, and ρ_{xy} is the correlation coefficient between x and y.

Following equation 4.6, $E_{a,agg}$ is calculated as

$$E_{a,agg} = \frac{\mathrm{cov}(\log(\overline{A_i e^{-\frac{E_{ai}}{RT}}}), \frac{1}{T})}{\mathrm{var}(\frac{1}{T})}(-R).$$

Thus, quantifying $E_{a,agg}$ requires an explicit expression of $\overline{A_i e^{-\frac{E_{ai}}{RT}}}$. Due to the nonlinearity of Arrhenius equation, using the Delta methods to approximate respiration when activation energy is heterogeneous is inaccurate. Thus, we took an alternative approach. We made assumptions about the distributions of A_i and E_{ai} that allow us to derive $\overline{A_i e^{-\frac{E_{ai}}{RT}}}$ analytically. We assumed:

$$E_{ai} \sim N(\overline{E_{ai}},\ \sigma_e^2); \tag{4.19}$$

$$A_i \sim logN(\overline{A_i},\ \sigma_A^2), \tag{4.20}$$

where N stands for normal distribution, logN stands for log-normal distribution, σ_e and σ_A are standard deviations of E_{ai} and A_i respectively. Based on the definition of log-normal distribution, the logarithm of a random variable following a log-normal distribution is normally

distributed:

$$\log(A_i) \sim N\Big(\log(\overline{A_i}) - \frac{1}{2}\log(\frac{\sigma_A^2}{\overline{A_i}^2}+1),\ \log(\frac{\sigma_A^2}{\overline{A_i}^2}+1)\Big). \tag{4.21}$$

We further assumed that $\log(A_i)$ and E_{ai} follow a joint normal distribution. As a property of multivariate normal distribution, $\log(A_i) - \frac{E_{ai}}{RT}$, a linear combination of $\log(A_i)$ and E_{ai}, follows a normal distribution:

$$\log(A_i) - \frac{E_{ai}}{RT} \sim N\Big(\log(\overline{A_i}) - \frac{1}{2}\log(\frac{\sigma_A^2}{\overline{A_i}^2}+1) - \frac{\overline{E_{ai}}}{RT}, \log(\frac{\sigma_A^2}{\overline{A_i}^2}+1) + \frac{\sigma_e^2}{R^2T^2} - \rho\frac{\sigma_e}{RT}\sqrt{\log(\frac{\sigma_A^2}{\overline{A_i}^2}+1)}\Big) \tag{4.22}$$

Because the exponential of a normally distributed random variable follows a log-normal distribution, $A_i e^{-\frac{E_{ai}}{RT}} = e^{\log(A_i)-\frac{E_{ai}}{RT}}$ follows a log-normal distribution, and the mean of the log-normal distribution is

$$\overline{A_i e^{-\frac{E_{ai}}{RT}}} = \exp\Big(\log(\overline{A_i}) - \frac{\overline{E_{ai}}}{RT} + \frac{1}{2}\frac{\sigma_e^2}{R^2T^2} - \frac{1}{2}\rho\frac{\sigma_e}{RT}\sqrt{\log(\frac{\sigma_A^2}{\overline{A_i}^2}+1)}\Big) \tag{4.23}$$

Plug equation 4.23 in equation 4.4, we obtain the expression for $E_{a,agg}$ shown in equation 4.7:

$$E_{a,agg} = \overline{E_{ai}} + \frac{1}{2}\sigma_e\left[\rho\sqrt{\log(\frac{\sigma_A^2}{\overline{A_i}^2}+1)} - \frac{\sigma_e}{R}\frac{\mathrm{cov}(\frac{1}{T^2},\frac{1}{T})}{\mathrm{var}(\frac{1}{T})}\right].$$

Bayesian parameter estimates for the multiple pool models

We demonstrate the details of the Bayesian parameter estimate procedure for the two pool without transfer model. Let $\boldsymbol{\theta}$ be the vector of parameters. For a two pool without transfer model, $\boldsymbol{\theta}$= $(k_1, k_2, E_{a1}, E_{a2}, \sigma)^T$. For a given set of observed respiration rate over time $\boldsymbol{y}$ = $(y_{t_1}, y_{t_2}, ..., y_{t_n})^T$, we assumed independent and identically distributed normal observational errors. Thus, the likelihood for any set of parameter is

$$\ell(\boldsymbol{\theta}) = \prod_{t=t_1}^{t_n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_t - \mathbf{1}^T e^{\boldsymbol{A}\boldsymbol{K}(T)t)^2}\boldsymbol{C}(0)}{2\sigma^2}}. \tag{4.24}$$

Let $p(\boldsymbol{\theta})$ be the prior probability density of parameter $\boldsymbol{\theta}$. Based on Bayes' theorem, the posterior probability density $q(\boldsymbol{\theta})$ is

$$q(\boldsymbol{\theta}) \propto \ell(\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{4.25}$$

Let $g(\boldsymbol{\theta}'|\boldsymbol{\theta})$ be the proposal density in the Metropolis–Hasting algorithm. We used a normal distribution centered at $\boldsymbol{\theta}$ for $g(\boldsymbol{\theta}'|\boldsymbol{\theta})$. Given that a normal proposal density is symmetric, $g(\boldsymbol{\theta}'|\boldsymbol{\theta}) = g(\boldsymbol{\theta}|\boldsymbol{\theta}')$. Thus, the acceptance ratio is

$$\alpha(\boldsymbol{\theta}'|\boldsymbol{\theta}) = \min\big(1, \frac{q(\boldsymbol{\theta}')}{q(\boldsymbol{\theta})}\big). \tag{4.26}$$

Then, sampling the posterior distribution of parameter $\boldsymbol{\theta}$ using the Metropolis–Hasting algorithm can be done in the following steps. First, start with an arbitrarily chosen parameter $\boldsymbol{\theta}$ and propose a new value of $\boldsymbol{\theta}'$ based on the proposal density $g(\boldsymbol{\theta}'|\boldsymbol{\theta})$. Based on the adaptive Metropolis–Hasting algorithm by Haario *et al.* (2001), we set the variance–covariance matrix of the proposal density as $(2.38)^2\boldsymbol{\Sigma}/d$, where $\boldsymbol{\Sigma}$ is the variance–covariance matrix of $\boldsymbol{\theta}$ based on the sampled values in the Markov Chain and d is the number of parameters. Second, calculate the acceptance ratio α (equation 4.26). Third, move to the proposed value $\boldsymbol{\theta}'$ with probability α. These steps are repeated for enough iterations until we reached stationarity.

References

Ågren, G.I. (2000). Temperature dependence of old soil organic matter. *Ambio*, 29, 55–55.

Ågren, G.I. & Bosatta, E. (1996). Quality: a bridge between theory and experiment in soil organic matter studies. *Oikos*, 76, 522–528.

Ågren, G.I. & Bosatta, E. (2002). Reconciling differences in predictions of temperature response of soil organic matter. *Soil Biology and Biochemistry*, 34, 129–132.

Allison, S.D., Wallenstein, M.D. & Bradford, M.A. (2010). Soil-carbon response to warming dependent on microbial physiology. *Nature Geoscience*, 3, 336–340.

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., Jones, C., Jung, M., Myneni, R. & Zhu, Z. (2013). Evaluating the land and ocean components of the global carbon cycle in the CMIP5 earth system models. *Journal of Climate*, 26, 6801–6843.

Anderson-Teixeira, K.J., Vitousek, P.M. & Brown, J.H. (2008). Amplified temperature dependence in ecosystems developing on the lava flows of Mauna Loa, Hawai'i. *Proceedings of the National Academy of Sciences*, 105, 228–233.

Arora, V.K., Boer, G.J., Friedlingstein, P., Eby, M., Jones, C.D., Christian, J.R., Bonan, G., Bopp, L., Brovkin, V., Cadule, P. *et al.* (2013). Carbon–concentration and carbon–climate feedbacks in CMIP5 earth system models. *Journal of Climate*, 26, 5289–5314.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R. *et al.* (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, 82, 2415–2434.

Billings, S.A., Tiemann, L.K., Ballantyne IV, F., Lehmeier, C. & Min, K. (2015). Investigating microbial transformations of soil organic matter: synthesizing knowledge from disparate fields to guide new experimentation. *Soil*, 1, 313.

Bond-Lamberty, B. & Thomson, A. (2010). Temperature-associated increases in the global soil respiration record. *Nature*, 464, 579–582.

Boone, R.D., Nadelhoffer, K.J., Canary, J.D. & Kaye, J.P. (1998). Roots exert a strong influence on the temperature sensitivity of soil respiration. *Nature*, 396, 570.

Bradford, M.A., Veen, G.C., Bonis, A., Bradford, E.M., Classen, A.T., Cornelissen, J.H.C., Crowther, T.W., Jonathan, R., Freschet, G.T., Kardol, P. *et al.* (2017). A test of the hierarchical model of litter decomposition. *Nature Ecology & Evolution*, 1, 1836–1845.

Bradford, M.A., Warren II, R.J., Baldrian, P., Crowther, T.W., Maynard, D.S., Oldfield, E.E., Wieder, W.R., Wood, S.A. & King, J.R. (2014). Climate fails to predict wood decomposition at regional scales. *Nature Climate Change*, 4, 625.

Bradford, M.A., Wieder, W.R., Bonan, G.B., Fierer, N., Raymond, P.A. & Crowther, T.W. (2016). Managing uncertainty in soil carbon feedbacks to climate change. *Nature Climate Change*, 6, 751–758.

Chen, H. & Tian, H.Q. (2005). Does a general temperature-dependent Q_10 model of soil respiration exist at biome and global scale? *Journal of integrative plant biology*, 47, 1288–1302.

Conant, R.T., Drijber, R.A., Haddix, M.L., Parton, W.J., Paul, E.A., Plante, A.F., Six, J. & Steinweg, J.M. (2008). Sensitivity of organic matter decomposition to warming varies with its quality. *Global Change Biology*, 14, 868–877.

Conant, R.T., Ryan, M.G., Ågren, G.I., Birge, H.E., Davidson, E.A., Eliasson, P.E., Evans, S.E., Frey, S.D., Giardina, C.P., Hopkins, F.M. *et al.* (2011). Temperature and soil organic matter decomposition rates–synthesis of current knowledge and a way forward. *Global Change Biology*, 17, 3392–3404.

Cox, P.M., Betts, R.A., Jones, C.D., Spall, S.A. & Totterdell, I.J. (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, 408, 184–187.

Craine, J.M., Fierer, N. & McLauchlan, K.K. (2010). Widespread coupling between the rate and temperature sensitivity of organic matter decay. *Nature Geoscience*, 3, 854–857.

Davidson, E.A. & Janssens, I.A. (2006). Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature*, 440, 165–173.

Davidson, E.A., Janssens, I.A. & Luo, Y. (2006). On the variability of respiration in terrestrial ecosystems: moving beyond Q10. *Global Change Biology*, 12, 154–164.

Davidson, E.A., Trumbore, S.E. & Amundson, R. (2000). Biogeochemistry: soil warming and organic carbon content. *Nature*, 408, 789–790.

Delcourt, H.R. & Delcourt, P.A. (1988). Quaternary landscape ecology: relevant scales in space and time. *Landscape Ecology*, 2, 23–44.

Exbrayat, J.F., Pitman, A. & Abramowitz, G. (2014). Response of microbial decomposition to spin-up explains CMIP5 soil carbon range until 2100. *Geoscientific Model Development*, 7, 2683–2692.

Fierer, N., Craine, J.M., McLauchlan, K. & Schimel, J.P. (2005). Litter quality and the temperature sensitivity of decomposition. *Ecology*, 86, 320–326.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I. *et al.* (2006). Climate–carbon cycle feedback analysis: results from the C^4MIP model intercomparison. *Journal of Climate*, 19, 3337–3353.

Gardner, R., Cale, W. & O'Neill, R. (1982). Robust analysis of aggregation error. *Ecology*, 63, 1771–1779.

Giardina, C.P., Litton, C.M., Crow, S.E. & Asner, G.P. (2014). Warming-related increases in soil CO_2 efflux are explained by increased below-ground carbon flux. *Nature Climate Change*, 4, 822–827.

Giardina, C.P. & Ryan, M.G. (2000). Evidence that decomposition rates of organic carbon in mineral soil do not vary with temperature. *Nature*, 404, 858–861.

Giasson, M.A., Ellison, A.M., Bowden, R.D., Crill, P.M., Davidson, E.A., Drake, J.E., Frey, S.D., Hadley, J.L., Lavine, M., Melillo, J.M. *et al.* (2013). Soil respiration in a northeastern us temperate forest: a 22-year synthesis. *Ecosphere*, 4, 1–28.

Graf, A., Weihermüller, L., Huisman, J., Herbst, M., Bauer, J. & Vereecken, H. (2008). Measurement depth effects on the apparent temperature sensitivity of soil respiration in field studies. *Biogeosciences*, 5, 1175–1188.

Graf, A., Weihermüller, L., Huisman, J.A., Herbst, M. & Vereecken, H. (2011). Comment on "global convergence in the temperature sensitivity of respiration at ecosystem level". *Science*, 331, 1265–1265.

Haario, H., Saksman, E., Tamminen, J. *et al.* (2001). An adaptive Metropolis algorithm. *Bernoulli*, 7, 223–242.

Hartley, I.P. & Ineson, P. (2008). Substrate quality and the temperature sensitivity of soil organic matter decomposition. *Soil Biology and Biochemistry*, 40, 1567–1574.

Iwasa, Y., Andreasen, V. & Levin, S. (1987). Aggregation in model ecosystems. i. perfect aggregation. *Ecological Modelling*, 37, 287–302.

Iwasa, Y., Levin, S.A. & Andreasen, V. (1989). Aggregation in model ecosystems ii. approximate aggregation. *Mathematical Medicine and Biology*, 6, 1–23.

Janssens, I.A. & Pilegaard, K. (2003). Large seasonal changes in q10 of soil respiration in a beech forest. *Global Change Biology*, 9, 911–918.

Jenkinson, D. (1990). The turnover of organic carbon and nitrogen in soil. *Philosophical Transaction of Royal Society of London B*, 329, 361–368.

Jenkinson, D.S., Adams, D. & Wild, A. (1991). Model estimates of CO_2 emissions from soil in response to global warming. *Nature*, 351, 304–306.

Jones, C.D., Cox, P. & Huntingford, C. (2003). Uncertainty in climate–carbon-cycle projections associated with the sensitivity of soil respiration to temperature. *Tellus B*, 55, 642–648.

Karhu, K., Auffret, M.D., Dungait, J.A., Hopkins, D.W., Prosser, J.I., Singh, B.K., Subke, J.A., Wookey, P.A., Ågren, G.I., Sebastià, M.T. *et al.* (2014). Temperature sensitivity of soil respiration rates enhanced by microbial community response. *Nature*, 513, 81–84.

Karhu, K., Fritze, H., Hämäläinen, K., Vanhala, P., Jungner, H., Oinonen, M., Sonninen, E., Tuomi, M., Spetz, P., Kitunen, V. *et al.* (2010). Temperature sensitivity of soil carbon fractions in boreal forest soil. *Ecology*, 91, 370–376.

Kirschbaum, M.U. (2004). Soil respiration under prolonged soil warming: are rate reductions caused by acclimation or substrate loss? *Global Change Biology*, 10, 1870–1877.

Kirschbaum, M.U.F. (2006). The temperature dependence of organic-matter decomposition—still a topic of debate. *Soil Biology and Biochemistry*, 38, 2510–2518.

Kirschbaum, M.U.F. (2010). The temperature dependence of organic matter decomposition: seasonal temperature variations turn a sharp short-term temperature response into a more moderate annually averaged response. *Global Change Biology*, 16, 2117–2129.

Knorr, W., Prentice, I.C., House, J. & Holland, E. (2005). Long-term sensitivity of soil carbon turnover to warming. *Nature*, 433, 298–301.

Lauenroth, W. & Sala, O.E. (1992). Long-term forage production of north american shortgrass steppe. *Ecological Applications*, 2, 397–403.

Lehmeier, C.A., Min, K., Niehues, N.D., Ballantyne, F. & Billings, S.A. (2013). Temperature-mediated changes of exoenzyme-substrate reaction rates and their consequences for the carbon to nitrogen flow ratio of liberated resources. *Soil Biology and Biochemistry*, 57, 374–382.

Levin, S.A. (1992). The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology*, 73, 1943–1967.

Liang, J., Li, D., Shi, Z., Tiedje, J.M., Zhou, J., Schuur, E.A., Konstantinidis, K.T. & Luo, Y. (2015). Methods for estimating temperature sensitivity of soil organic matter based on incubation data: A comparative evaluation. *Soil Biology and Biochemistry*, 80, 127–135.

Liski, J., Ilvesniemi, H., Maekelae, A. & Westman, C. (1999). CO_2 emissions from soil in response to climatic warming are overestimated. the decomposition of old soil organic matter is tolerant of temperature. *Ambio*, 28, 171–174.

Lloyd, J. & Taylor, J. (1994). On the temperature dependence of soil respiration. *Functional ecology*, 8, 315–323.

Luo, Y. (2007). Terrestrial carbon-cycle feedback to climate warming. *Annual Review of Ecology, Evolution, and Systematics*, pp. 683–712.

Luo, Y., Ahlström, A., Allison, S.D., Batjes, N.H., Brovkin, V., Carvalhais, N., Chappell, A., Ciais, P., Davidson, E.A., Finzi, A. *et al.* (2016). Toward more realistic projections of soil carbon dynamics by Earth system models. *Global Biogeochemical Cycles*, 30, 40–56.

Luo, Y., Wan, S., Hui, D. & Wallace, L.L. (2001). Acclimatization of soil respiration to warming in a tall grass prairie. *Nature*, 413, 622–625.

Luo, Y. & Weng, E. (2011). Dynamic disequilibrium of the terrestrial carbon cycle under global change. *Trends in Ecology & Evolution*, 26, 96–104.

Luo, Y., White, L.W., Canadell, J.G., DeLucia, E.H., Ellsworth, D.S., Finzi, A., Lichter, J. & Schlesinger, W.H. (2003). Sustainability of terrestrial carbon sequestration: a case study in Duke Forest with inversion approach. *Global biogeochemical cycles*, 17.

Mahecha, M.D., Reichstein, M., Carvalhais, N., Lasslop, G., Lange, H., Seneviratne, S.I., Vargas, R., Ammann, C., Arain, M.A., Cescatti, A. *et al.* (2010). Global convergence in the temperature sensitivity of respiration at ecosystem level. *Science*, 329, 838–840.

Mearns, L.O., Katz, R.W. & Schneider, S.H. (1984). Extreme high-temperature events: changes in their probabilities with changes in mean temperature. *Journal of Climate and Applied Meteorology*, 23, 1601–1613.

Melillo, J., Steudler, P., Aber, J., Newkirk, K., Lux, H., Bowles, F., Catricala, C., Magill, A., Ahrens, T. & Morrisseau, S. (2002). Soil warming and carbon-cycle feedbacks to the climate system. *Science*, 298, 2173–2176.

O'Neill, R. & Rust, B. (1979). Aggregation error in ecological models. *Ecological Modelling*, 7, 91–105.

Parton, W., Schimel, D.S., Cole, C. & Ojima, D. (1987). Analysis of factors controlling soil organic matter levels in Great Plains Grasslands. *Soil Science Society of America Journal*, 51, 1173–1179.

Phillips, C.L., Nickerson, N., Risk, D. & Bond, B.J. (2011). Interpreting diel hysteresis between soil respiration and temperature. *Global Change Biology*, 17, 515–527.

Phillips, S.C., Varner, R.K., Frolking, S., Munger, J.W., Bubier, J.L., Wofsy, S.C. & Crill, P.M. (2010). Interannual, seasonal, and diel variation in soil respiration relative to ecosystem respiration at a wetland to upland slope at harvard forest. *Journal of Geophysical Research: Biogeosciences*, 115.

Qi, Y., Xu, M. & Wu, J. (2002). Temperature sensitivity of soil respiration and its effects on ecosystem carbon budget: nonlinearity begets surprises. *Ecological Modelling*, 153, 131–142.

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raich, J.W. & Potter, C.S. (1995). Global patterns of carbon dioxide emissions from soils. *Global Biogeochemical Cycles*, 9, 23–36.

Raich, J.W., Potter, C.S. & Bhagawati, D. (2002). Interannual variability in global soil respiration, 1980–94. *Global Change Biology*, 8, 800–812.

Raich, J.W. & Schlesinger, W.H. (1992). The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate. *Tellus B*, 44, 81–99.

Rastetter, E.B., King, A.W., Cosby, B.J., Hornberger, G.M., O'Neill, R.V. & Hobbie, J.E. (1992). Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. *Ecological Applications*, 2, 55–70.

Reichstein, M., Rey, A., Freibauer, A., Tenhunen, J., Valentini, R., Banza, J., Casals, P., Cheng, Y., Grünzweig, J.M., Irvine, J. *et al.* (2003). Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices. *Global Biogeochemical Cycles*, 17.

Rustad, L., Campbell, J., Marion, G., Norby, R., Mitchell, M., Hartley, A., Cornelissen, J. & Gurevitch, J. (2001). A meta-analysis of the response of soil respiration, net nitrogen mineralization, and aboveground plant growth to experimental ecosystem warming. *Oecologia*, 126, 543–562.

Sanderman, J., Amundson, R.G. & Baldocchi, D.D. (2003). Application of eddy covariance measurements to the temperature dependence of soil organic matter mean residence time. *Global Biogeochemical Cycles*, 17, 1061.

Savage, V.M. (2004). Improved approximations to scaling relationships for species, populations, and ecosystems across latitudinal and elevational gradients. *Journal of Theoretical Biology*, 227, 525–534.

Schlesinger, W.H. (1977). Carbon balance in terrestrial detritus. *Annual Review of Ecology and Systematics*, 8, 51–81.

Schlesinger, W.H. & Andrews, J.A. (2000). Soil respiration and the global carbon cycle. *Biogeochemistry*, 48, 7–20.

Scott-Denton, L.E., Sparks, K.L. & Monson, R.K. (2003). Spatial and temporal controls of soil respiration rate in a high-elevation, subalpine forest. *Soil Biology and Biochemistry*, 35, 525–534.

Sierra, C.A. (2012). Temperature sensitivity of organic matter decomposition in the arrhenius equation: some theoretical considerations. *Biogeochemistry*, 108, 1–15.

Subke, J.A. & Bahn, M. (2010). On the "temperature sensitivity" of soil respiration: Can we use the immeasurable to predict the unknown? *Soil Biology and Biochemistry*, 42, 1653–1656.

Todd-Brown, K., Randerson, J., Hopkins, F., Arora, V., Hajima, T., Jones, C., Shevliakova, E., Tjiputra, J., Volodin, E., Wu, T. *et al.* (2014). Changes in soil organic carbon storage predicted by earth system models during the 21st century. *Biogeosciences*, 11, 2341–2356.

Todd-Brown, K., Randerson, J., Post, W., Hoffman, F., Tarnocai, C., Schuur, E. & Allison, S. (2013). Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences*, 10, 1717–1736.

Trumbore, S.E., Chadwick, O.A. & Amundson, R. (1996). Rapid exchange between soil carbon and atmospheric carbon dioxide driven by temperature change. *Science*, 272, 393–395.

Vargas, R., Baldocchi, D.D., Allen, M.F., Bahn, M., Black, T.A., Collins, S.L., Yuste, J.C., Hirano, T., Jassal, R.S., Pumpanen, J. *et al.* (2010). Looking deeper into the soil: biophysical controls and seasonal lags of soil co2 production and efflux. *Ecological Applications*, 20, 1569–1582.

Wang, Y., Song, C., Yu, L., Mi, Z., Wang, S., Zeng, H., Fang, C., Li, J. & He, J.s. (2018). Convergence in temperature sensitivity of soil respiration: evidence from the tibetan alphine grasslands. *Soil Biology and Biochemistry*, 122, 50–59.

Waring, B.G., Adams, R., Branco, S. & Powers, J.S. (2016). Scale-dependent variation in nitrogen cycling and soil fungal communities along gradients of forest composition and age in regenerating tropical dry forests. *New Phytologist*, 209, 845–854.

Wieder, W.R., Allison, S.D., Davidson, E.A., Georgiou, K., Hararuk, O., He, Y., Hopkins, F., Luo, Y., Smith, M.J., Sulman, B. *et al.* (2015). Explicitly representing soil microbial processes in earth system models. *Global Biogeochemical Cycles*, 29, 1782–1800.

Wieder, W.R., Bonan, G.B. & Allison, S.D. (2013). Global soil carbon projections are improved by modelling microbial processes. *Nature Climate Change*, 3, 909–912.

Xu, M. & Qi, Y. (2001). Spatial and seasonal variations of q 10 determined by soil respiration measurements at a sierra nevadan forest. *Global Biogeochemical Cycles*, 15, 687–696.

Yvon-Durocher, G., Caffrey, J.M., Cescatti, A., Dossena, M., del Giorgio, P., Gasol, J.M., Montoya, J.M., Pumpanen, J., Staehr, P.A., Trimmer, M. *et al.* (2012). Reconciling the temperature dependence of respiration across timescales and ecosystem types. *Nature*, 487, 472–476.

Zhou, T., Shi, P., Hui, D. & Luo, Y. (2009). Global pattern of temperature sensitivity of soil heterotrophic respiration (Q10) and its implications for carbon-climate feedback. *Journal of Geophysical Research: Biogeosciences*, 114.

Table 4.1: Estimates of decomposition rates and activation energies for the two pool model without transfer. Type of response following the cooling and rewarming treatments are defined as Karhu *et al.* (2014). Unit for the decomposition rates k_1 and k_2 is day^{-1}. Unit for the activation energy E_{a1} and E_{a2} is KJ mol^{-1}.

Soil ID	k_1	k_2	E_{a1}	E_{a2}	Type of response
5E2	1.68×10^{-3}	6.74×10^{-6}	66.85	100.01	Enhancing
2H	2.60×10^{-4}	1.34×10^{-6}	35.32	211.97	Enhancing
1D	3.04×10^{-4}	5.82×10^{-6}	14.46	170.89	Enhancing
1H	2.07×10^{-4}	4.87×10^{-6}	19.77	230.39	Enhancing
1C	2.82×10^{-4}	3.45×10^{-6}	19.85	207.06	Enhancing
1A	7.32×10^{-4}	0.58×10^{-6}	156.59	154.53	No response
3A	3.18×10^{-4}	1.79×10^{-6}	261.20	19.96	Compensatory
2D	1.12×10^{-3}	3.36×10^{-6}	173.69	82.59	Compensatory
4G	1.92×10^{-3}	1.01×10^{-5}	100.96	80.06	Compensatory

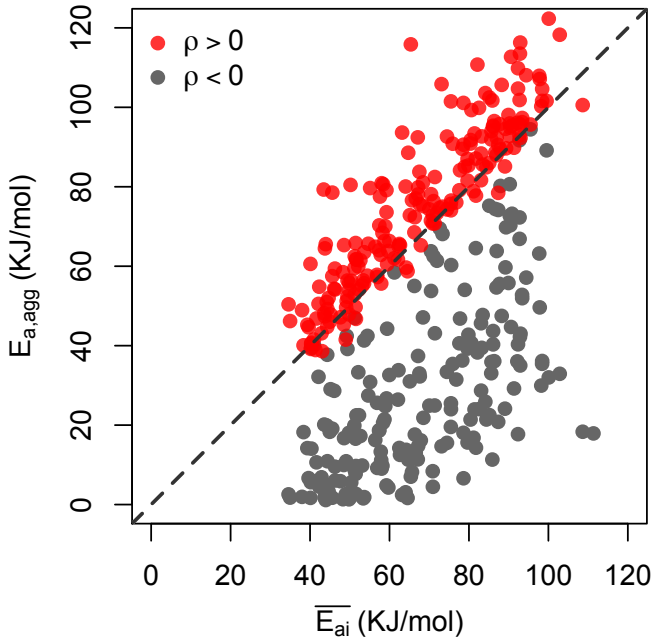

Figure 4.1: The temperature sensitivity of average soil respiration within a spatial extent and the average patch activation energy when the correlation between A_i and E_{ai} (ρ) was positive or negative. The dashed line is the 1:1 line.

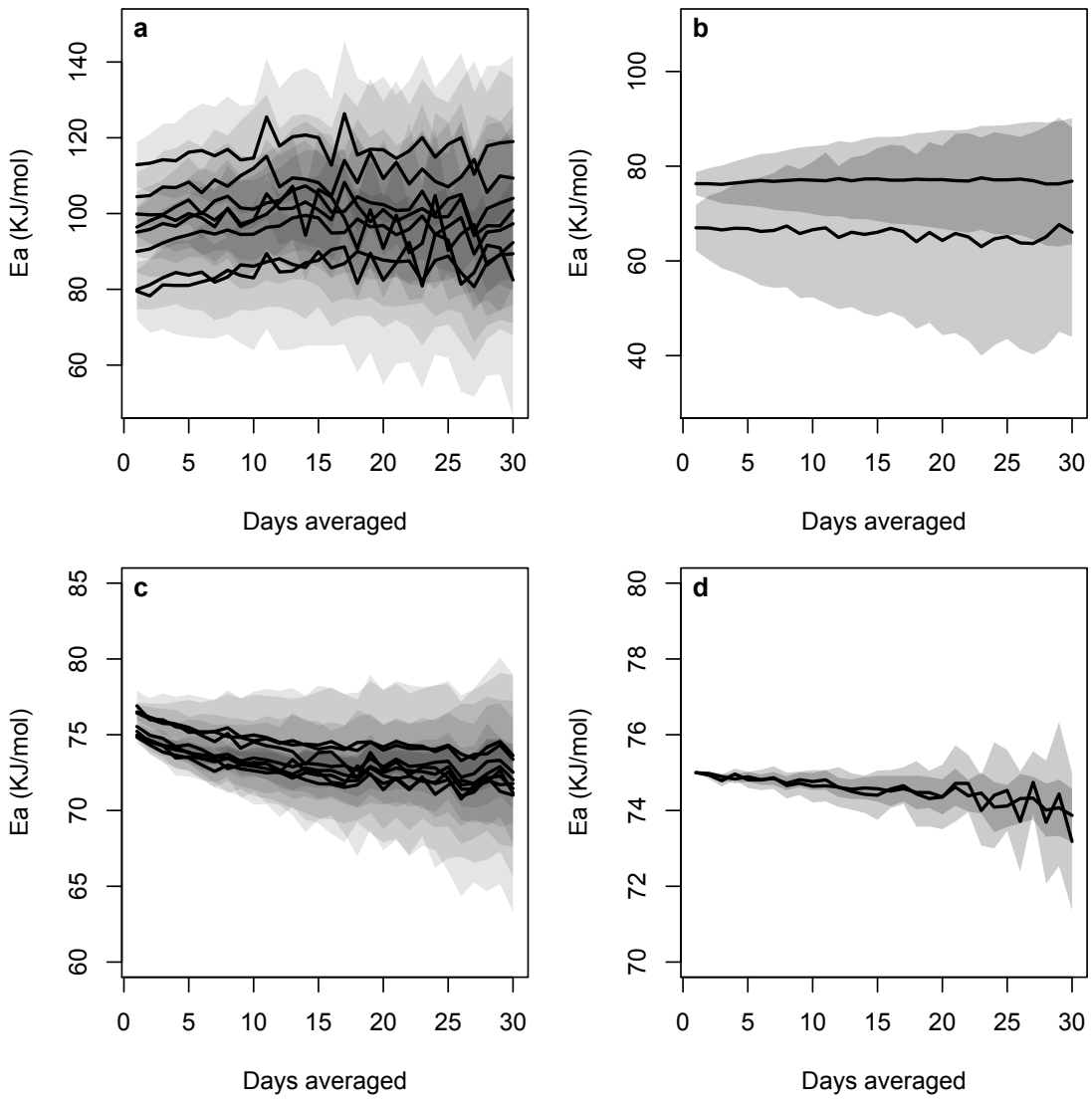

Figure 4.2: Estimated activation energy based on the average respiration rate and temperature over different number of days using data from (a) Harvard forest, (b) Tibet alpine grassland, (c) simulated respiration using Harvard Forest temperature, and (d) simulated respiration using Tibet grassland temperature. Shaded areas are 95% confidence interval for the estimated activation energy. In each panel, each line corresponds to data from one chamber

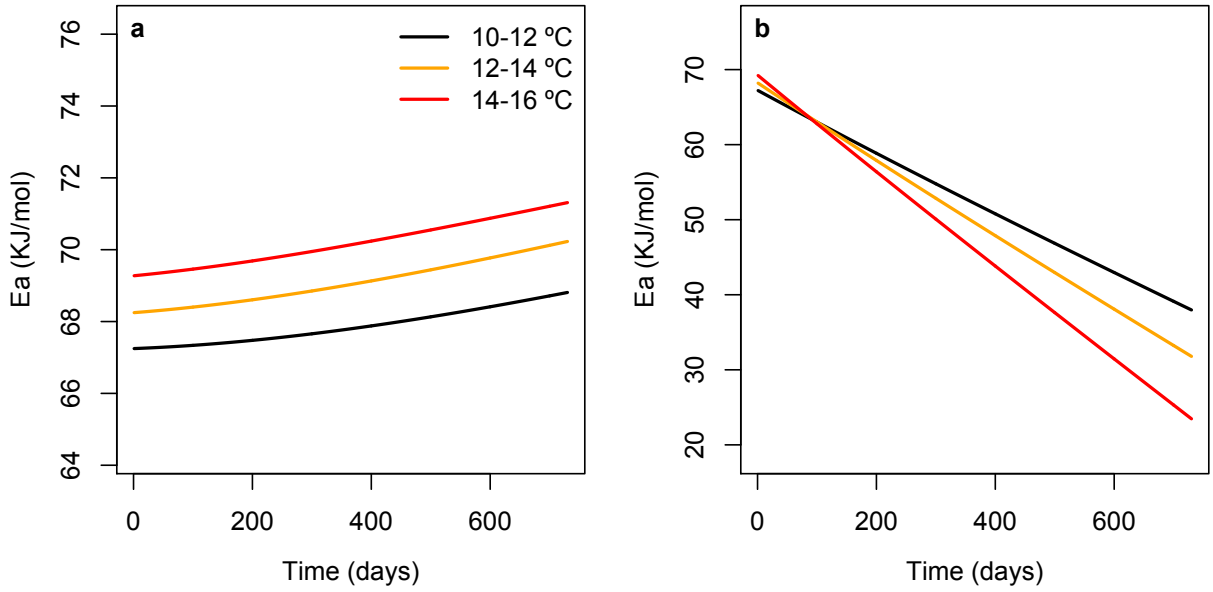

Figure 4.3: Relationship between activation energy and temperature based on respiration data expressed in the unit of (a) respiration per unit soil carbon and (b) total respiration from all substrate pools.

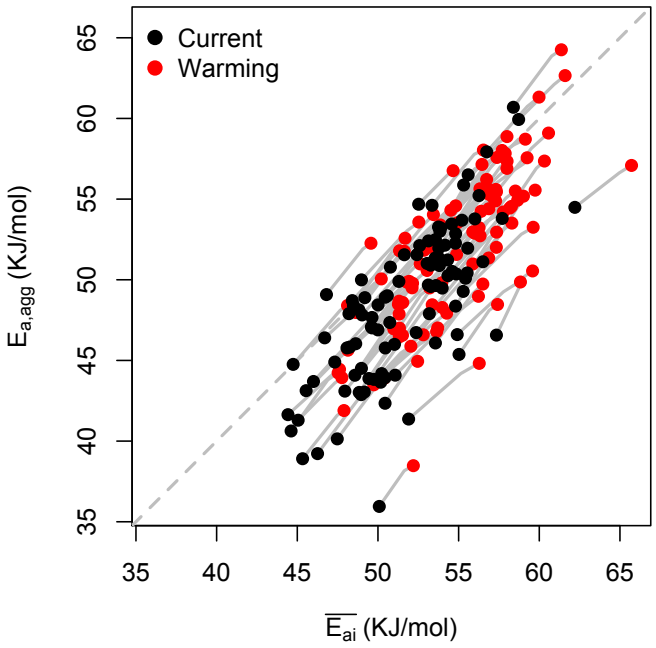

Figure 4.4: Effects of aggregation over space on temperature sensitivity when respiration from each patch is consisted of respiration from multiple substrate pools. Black dots and red dots represent the aggregated activation energy and patch average activation energy at the current steady state and steady state under a 5 °C increase in temperature. Gray lines connecting the dots represent the trajectory over time

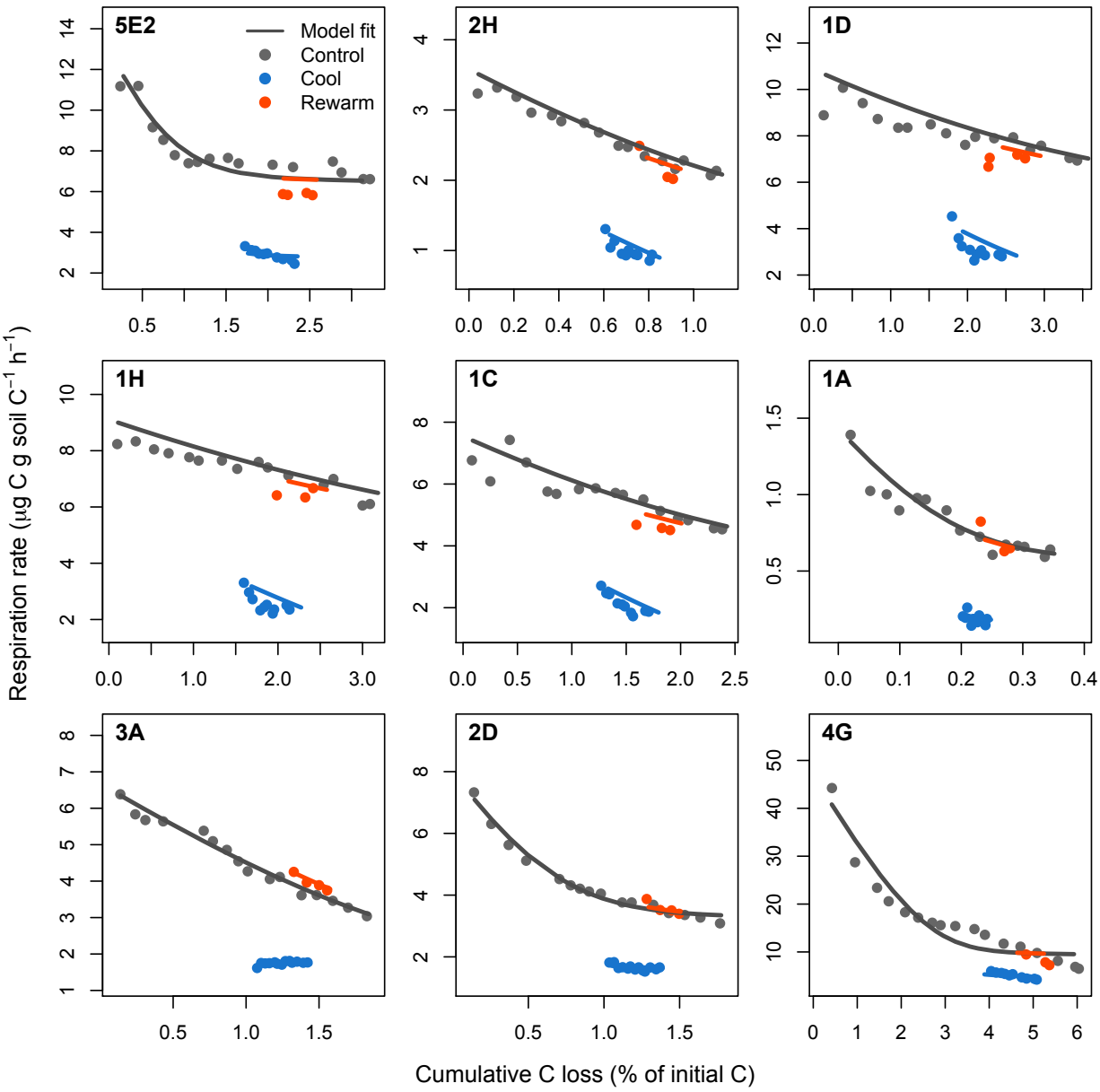

Figure 4.5: Two pool model fit to the incubation data by Karhu *et al.* (2014). Each panel shows data from incubation of one soil sample. Sample number for each soil matches the numbers in Karhu *et al.* (2014)

Chapter 5

Conclusions

Examining patterns and mechanisms that explicitly consider scale is key to developing a robust predictive theory for ecology, as almost all ecological problems are in the context of a particular spatial and temporal scale (Wiens, 1989; Levin, 1992). Ideally, the scale of study should match the scale of intended application (Hewitt *et al.*, 2007). When this is infeasible, understanding how patterns and mechanisms observed on one scale translate to the scale of interests becomes critical. Throughout this dissertation, I strived to address the issue of scale from these two approaches. In chapter 2, I used a dynamic model of dissolved oxygen to directly estimate the temperature sensitivity of whole-stream metabolism. The modeling approach allowed me to quantify the temperature sensitivity on the intended spatial scale. In chapter 3, I investigated the role of spatial heterogeneity in driving differences in stream metabolism across scale. In chapter 4, I explored how aggregation over space, time, and heterogeneous carbon substrate pools influenced the perceived temperature sensitivity of soil respiration. Collectively, these chapters share a common theme: ecosystem carbon flux is the sum of its parts, but the property of carbon flux, such as the temperature sensitivity, is not directly transferrable from the parts to the whole system.

Central to the explicit consideration of scale is a dynamic and mechanistic view of ecosystem processes (Denny & Benedetti-Cecchi, 2012). The mechanistic view is necessary for developing the modeling approaches to quantify temperature sensitivity of stream metabolism on reach scale, which is difficult for direct experimental work. It also served as the foundation to explicitly link metabolism across spatial scales, or link temperature sensitivity of soil respiration across space, time, and ecological organizations. Based on the

mechanistic view of ecological processes, this dissertation research followed two guiding principles: first, inference and prediction were done in a consistent and rigorous way based on a mechanistic understanding of the ecological processes (Clark *et al.*, 2001; Dietze *et al.*, 2018). For example, the estimates of the temperature sensitivities of GPP and ER, and the prediction of warming effects on stream metabolism were both based on the mechanistic description of DO dynamics. Second, ecological patterns should be tested against predictions derived explicitly from the mechanistic descriptions of the processes involved instead of testing bivariate correlation empirically. For example, the difference in metabolism across scales and its expected relationship with spatial heterogeneity is first derived from theoretical analyses and then tested against empirical data. Such approach allows us to examine the hypothesized pattern with multiple data sets and reduce the risk of data dredging.

Finally, this dissertation research provided several unique contributions to the research on carbon dynamics. In particular, to my best knowledge, chapter 2 provided the first borad scale quantification of the temperature dependence of stream metabolic balance, and presented the pattern of warming-induced changes in metabolic balance that has not been discovered before. Chapter 3 is the first cross-biome study that demonstrated how the physiology of primary production and respiration determined the difference in metabolism across scales. Chapter 4 made an conceptual synthesis and provided a potential unifying framework to explain commonly observed pattern of the temperature sensitivity of soil respiration. The progresses made in this dissertation shed lights to several future research directions.

One area for future investigation is to integrate stream and adjacent terrestrial ecosystems when evaluating the feedback between carbon cycle and climatic warming. This dissertation work predicted how warming is likely to influence metabolic balance in streams at the current state of the stream ecosystems. However, changes in streams and adjacent terrestrial ecosystems concurrent with warming may complicate this prediction. For example, warming is expected to change the quantity and quality of allochthonous carbon inputs by stimulating soil organic matter decomposition (Freeman *et al.*, 2001) and altering riparian communities

(Kominoski *et al.*, 2013). Thermal adaptation of benthic communities (Padfield *et al.*, 2016, 2017) and changes in hydrology or nutrient availability (Cross *et al.*, 2015; Demars *et al.*, 2015) may further amplify or damp the predicted convergence of metabolic balance. Developing more integrative models that incorporate the warming response of stream metabolic balance identified in this study will improve our ability to quantify the feedback between carbon dynamics and future climate changes.

Another future directions is to link reach scale and watershed scale measurements. While this dissertation research quantified how habitat scale metabolism was linked to reach scale metabolism, it remains a challenge to scale up reach scale measurements to the watershed or landscape scale. To date, most watershed scale models are based on simple empirical relationships that relate easily measurable abiotic quantities to the biological processes of interests. Realistic parameterization of biological processes throughout the watersheds remains rare. However, the more extensive application of automated sensors and development of computational approaches to data–model integration approach provide a promising way to more realistically parameterize reach scale data over sufficient spatial extent and resolution, and validate such parameterization against observations made on the watershed or landscape scale (Luo *et al.*, 2011; Peng *et al.*, 2011). For example, our modeling approach to quantify reach scale temperature sensitivity of GPP and ER, coupled with data from spatially distributed DO sensors and watershed models, may allow for the quantification of watershed scale carbon budget in a warming climate.

Finally, modeling soil carbon dynamics at regional or global scales in a warming climate remains a grand challenge. Different mechanistic descriptions of soil carbon dynamics lead to substantially different predictions of soil carbon flux and storage in a warming climate (Allison *et al.*, 2010; Wieder *et al.*, 2013). Debates on what mechanisms drive the temperature dependence of soil organic carbon decomposition have persisted in the literature for years, yet we have not reached a consensus on the best way to conceptualize and model soil carbon dynamics (Wieder *et al.*, 2015; Luo *et al.*, 2016). This dissertation research demonstrated

that shifts in the composition of carbon substrate pools as a result of the inherent differences in the temperature sensitivity among substrates is a potential explanation for many observed temperature-activation energy relationship. Thus, pool sizes and flux data alone are unlikely to be sufficient to reveal the mechanisms driving the soil carbon dynamics uniquely. Studies that directly quantify microbial properties, either in the field (Frey *et al.*, 2013) or in highly controlled experiments (Lehmeier *et al.*, 2016), may be key to advance our understanding in this area.

All together, this dissertation examined the patterns, causes, and consequences of the scale dependence of ecosystem carbon flux and highlighted the scale dependence of ecosystem carbon flux. The results provided novel insights into the differential impacts of climate changes on ecosystem carbon flux on different spatial and temporal scales, and have also critically allowed me to establish the link between carbon flux on different scales. The findings suggest that properties of carbon flux, such as the temperature sensitivity, changes across scales in the presence of heterogeneity and nonlinearity. Accounting for variance is key to correctly link and translate ecological patterns across scales, and explicitly formalizing the dynamics of the ecological processes of interests is a powerful and versatile tool to address the issue of scale in ecosystem ecology.

References

Allison, S.D., Wallenstein, M.D. & Bradford, M.A. (2010). Soil-carbon response to warming dependent on microbial physiology. *Nature Geoscience*, 3, 336–340.

Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.A., Lodge, D.M., Pascual, M., Pielke, R., Pizer, W. *et al.* (2001). Ecological forecasts: an emerging imperative. *science*, 293, 657–660.

Cross, W.F., Hood, J.M., Benstead, J.P., Huryn, A.D. & Nelson, D. (2015). Interactions between temperature and nutrients across levels of ecological organization. *Global Change Biology*, 21, 1025–1040.

Demars, B.O., Thompson, J. & Manson, J.R. (2015). Stream metabolism and the open diel oxygen method: Principles, practice, and perspectives. *Limnology and Oceanography: Methods*, 13, 356–374.

Denny, M. & Benedetti-Cecchi, L. (2012). Scaling up in ecology: mechanistic approaches. *Annual Review of Ecology, Evolution, and Systematics*, 43, 1–22.

Dietze, M.C., Fox, A., Beck-Johnson, L.M., Betancourt, J.L., Hooten, M.B., Jarnevich, C.S., Keitt, T.H., Kenney, M.A., Laney, C.M., Larsen, L.G. *et al.* (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences*.

Freeman, C., Evans, C., Monteith, D., Reynolds, B. & Fenner, N. (2001). Export of organic carbon from peat soils. *Nature*, 412, 785–785.

Frey, S.D., Lee, J., Melillo, J.M. & Six, J. (2013). The temperature response of soil microbial efficiency and its feedback to climate. *Nature Climate Change*, 3, 395.

Hewitt, J.E., Thrush, S.F., Dayton, P.K. & Bonsdorff, E. (2007). The effect of spatial and temporal heterogeneity on the design and analysis of empirical studies of scale-dependent systems. *The American Naturalist*, 169, 398–408.

Kominoski, J.S., Follstad Shah, J.J., Canhoto, C., Fischer, D.G., Giling, D.P., González, E., Griffiths, N.A., Larrañaga, A., LeRoy, C.J., Mineau, M.M. *et al.* (2013). Forecasting functional implications of global changes in riparian plant communities. *Frontiers in Ecology and the Environment*, 11, 423–432.

Lehmeier, C.A., Ballantyne, F., Min, K. & Billings, S.A. (2016). Temperature-mediated changes in microbial carbon use efficiency and 13 c discrimination. *Biogeosciences*, 13, 3319–3329.

Levin, S.A. (1992). The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology*, 73, 1943–1967.

Luo, Y., Ahlström, A., Allison, S.D., Batjes, N.H., Brovkin, V., Carvalhais, N., Chappell, A., Ciais, P., Davidson, E.A., Finzi, A. *et al.* (2016). Toward more realistic projections of soil carbon dynamics by Earth system models. *Global Biogeochemical Cycles*, 30, 40–56.

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J.S. & Schimel, D.S. (2011). Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*, 21, 1429–1442.

Padfield, D., Lowe, C., Buckling, A., Ffrench-Constant, R., Jennings, S., Shelley, F., Ólafsson, J.S. & Yvon-Durocher, G. (2017). Metabolic compensation constrains the temperature dependence of gross primary production. *Ecology letters*, 20, 1250–1260.

Padfield, D., Yvon-Durocher, G., Buckling, A., Jennings, S. & Yvon-Durocher, G. (2016). Rapid evolution of metabolic traits explains thermal adaptation in phytoplankton. *Ecology Letters*, 19, 133–142.

Peng, C., Guiot, J., Wu, H., Jiang, H. & Luo, Y. (2011). Integrating models with data in ecology and palaeoecology: advances towards a model–data fusion approach. *Ecology Letters*, 14, 522–536.

Wieder, W.R., Allison, S.D., Davidson, E.A., Georgiou, K., Hararuk, O., He, Y., Hopkins, F., Luo, Y., Smith, M.J., Sulman, B. *et al.* (2015). Explicitly representing soil microbial processes in earth system models. *Global Biogeochemical Cycles*, 29, 1782–1800.

Wieder, W.R., Bonan, G.B. & Allison, S.D. (2013). Global soil carbon projections are improved by modelling microbial processes. *Nature Climate Change*, 3, 909–912.

Wiens, J.A. (1989). Spatial scaling in ecology. *Functional ecology*, 3, 385–397.