Linking population and ecosystem processes through time and space in the highly dynamic Lake Mývatn, Iceland

by

Joseph S. Phillips

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Zoology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2019

Date of final oral examination: April 17th, 2019

The dissertation is approved by the following members of the Final Oral Committee:

Stephen R. Carpenter, Professor of Integrative Biology
Claudio Gratton, Professor of Entomology
Anthony R. Ives, Professor of Integrative Biology
Monica G. Turner, Professor of Integrative Biology
Jake Vander Zanden, Professor of Integrative Biology

Table of Contents

Acknowledgements

My time as a graduate student at UW-Madison, including the work reported in this dissertation, benefitted from the contributions and support of many people to whom I am sincerely grateful. This includes all of the members of my dissertation committee (Steve Carpenter, Claudio Gratton, Monica Turner, and Jake Vander Zanden), who have provided substantial feedback on my work and have always offered that feedback in the kindest spirit.

Of course, I am particularly grateful to my graduate advisor, Tony. I was originally interested in joining Tony's group because his work was a model for the kind of science I wanted to do[1]. Five years later, I can say that Tony is a model for the kind of scientist I want to be, which has far more to do with how you treat those around you than it does with the work itself. In my first summer working as part of Tony's long-term project at Mývatn, when I was the only graduate student in charge of the field crew upon Tony's return to the US, he told me that the only thing I could do wrong was to treat someone poorly. On a separate occasion, Tony said that mentoring students is far more important than publishing papers, which is one of those rare points that is obviously true (how much impact can a single paper or even ten papers really have?) but can be difficult to spot and is easy to forget. But mostly, Tony has led by example, with perhaps the most important being the way he pursues science by working with people that he likes and treating those people such that the feeling is mutual.

In this vein, I owe a tremendous debt to all of people I have been lucky enough to count as collaborators and friends on the Mývatn project. Cristina Herren and Kyle Webert helped

[1]I don't think Tony knows this, but I first became interested in his work after reading a paper by him and one of his postdocs on pea aphids the summer prior to my freshman year of college. I was sufficiently intrigued that I found his lab webpage. Based on the paper, I assumed he was a "real entomologist", but one who was interested in "real ecology". Tony would be the first to admit that he's far from a "real entomologist". But the fact that he could be confused for one in a "real ecology" paper says a lot about Tony's science and why I found it so appealing.

encouraged me to say “yes” to Tony’s invitation to work at Mývatn and played a big role in my orientation to the project. Caroline Daws was our field technician in 2015, whose support was absolutely essential to the success of my first summer as a graduate student. She was followed by Becca Blundell, an intern in 2015 who came back as our field tech in 2016 and whose sense of order and discipline was much appreciated and has been hard to replicate in the years sense. Amanda McCormick joined the project as a graduate student for the summer of 2016, and she and I jointly led the field crew (with others joining later—more on them in a moment) for the 2016, 2017, and 2018 field seasons. Those summers were great, and I can say with certainty that they would not have been nearly so successful had it not been for Amanda’s dual presence as the responsible adult and the ringleader of fun. Jamie Botsch joined the project as an intern in 2016 and decided to stick around forever (he’s now a graduate student on the project and will be the principle team leader this coming summer). He and I are perhaps a bit too similar in our obsessive tendencies (be it fawning about midges or constructing preposterous experimental apparatuses in the cold basement), but usually we arrive at a better place than we could have alone. Matt McCary joined us as a postdoc for the 2017 and 2018 seasons and is one of the most thoughtful scientists I’ve known, in addition to being a ringleader of fun in his own right. Claudio Gratton (another “real ecologist” masquerading as an entomologist[1]) as a co-PI on the project spent a few weeks at Mývatn during each of my field seasons, and even as my interests gravitated in a more aquatic direction, he continued to provide essential advice on how to function as a field scientist. Árni Einarsson was (and is) our host and collaborator at Mývatn, whose knowledge of the ecosystem and undying generosity were absolutely critical to every aspect of our time at Mývatn. I would also like to thank all of our field interns, without whom our field work simply wouldn’t have been possible: Aldo Arellano, Riley Book, Kelvin Chen,

Aliza Fassler, Shannon Grover, Jennifer Harris, Karen Jorgenson, Abby Lewis, Colleen Miller, Caroline Owens, Jose Sanchez-Ruiz, Natalie Schmer, Beth Smith, and Aspen Ward.

Back in Madison, many members of the Ives lab provided essential support, sometimes through their advice or expertise, but mostly through their friendship. Meghan Fitzgerald and Huan Fan were my first friends in Madison and played a central role to my introduction to life as a graduate student. Lucas Nell has exerted an inordinate influence on my habits and skills with computers, but more importantly has been a great collaborator on projects beyond Mývatn and in doing so has helped to remind me of my fondness for population biology beyond midges (he also puts up with my rambling rants and my general quirkiness, while contributing his own in equal measure). In a similar vein, Mike Bosch's work with ants and aphids has reminded me of my fondness for those critters in particular are for critters in general (and how important that is for doing biology properly). Several postdocs have spent time in the lab and have shaped me and my work both directly and by example: Brandon Barton, Rachel Penczykowski, and Tanjona Ramiadantsoa (a.k.a the pot-stirrer). Finally, I'd like to thank Claudio Gratton (again) for inviting me to spend time with the wonderful people in his lab, who have provided great feedback on my work and have taught me a lot about the more applied side of ecology.

My fellow graduate students beyond the Ives labs have been central to my time in Madison—in fact, I count them (along with my friends in the lab) as the primary determinant of my very happy time in Madison. Everyone talks about how great Madison is, but it's really the people that are great. They are far too numerous to name, but I am particularly indebted to Jen Chandler, Pete Guiden, Devin Merullo, and Jeremy Spool, for many hours of racquetball, birding, symphony concerts, games of Risk, and general hanging about since the beginning of my time in Madison.

Finally, I would like to thank my parents Jack and Jane, and my brother Jacob. Not only are they the most important people in my life, they are the most important dimension of my life. I am incredibly grateful that despite spending most of our time apart, not even the smallest distance has entered my relationships with any of them. May that always be so.

Abstract

Many populations vary substantially in abundance, which can have important consequences for the communities and ecosystems in which they exist. This is especially true for species with large per capita effects, such as ecosystem engineers that modify the structure of the physical environment experienced by themselves and other populations. Regardless of whether a population's ecosystem effects arise from high abundance, high per capita effects, or both, these effects are often mediated by their environmental context and therefore have the potential to vary substantially through space and time.

In this dissertation, I explore the connections between population dynamics, environmental variation, and ecosystem-consequences of single populations in the context of Lake Mývatn in northeastern Iceland. Mývatn is uniquely suited to exploring the intersection of these topics due to (1) its relatively simple ecological community dominated by a few taxa and (2) its high productivity that fuels high, but variable, abundance of those dominant taxa. Chief among these are midges that transform the benthic habitat in years of peak abundance through their building of silk-tubes, which in turn provides a substrate for algal growth. However, the dominant midge species has erratic fluctuations in abundance spanning several orders-of-magnitude, raising the question of how these dynamics arise and what their consequences are for the whole ecosystem. Furthermore, Mývatn is subject to episodic but occasionally very thick cyanobacterial blooms that have to potential to shift the majority of production from the benthic to the pelagic habitat, thereby altering the energy flow through the food web. Given that context, I discuss field experiments and modeling of ecological time series spanning multiple trophic levels and temporal scales at Mývatn. Together, this work illustrates the connections between population and ecosystem processes in the context of the highly dynamic Mývatn system.

Introduction

Many populations fluctuate dramatically in abundance, which can have important consequences for the communities and ecosystems in which they occur (Elton 1924, Elton and Nicholson 1942, Dwyer et al. 2004, Yang 2004, Hoekman et al. 2011). While ecologists have identified various causes of complex population dynamics, perhaps the most important are consumer-resource interactions and environmental variation (e.g. Elton 1924; Davidson & Andrewartha 1948; Krebs *et al.* 2001). Historically, these two mechanisms were considered in isolation, and there was even contentious debate over which was more important (Elton 1924; Nicholson & Bailey 1935; Andrewartha & Birch 1954; Turchin *et al.* 2000). Subsequent work revealed that not only are both mechanisms important, but understanding their interplay is essential for understanding population fluctuations (Ripa and Ives 2003, Benincà et al. 2011). For example, species interactions can amplify or dampen environmental stochasticity and can act as 'filters' that change the underlying structure of the stochasticity (Vasseur 2007). In turn, environmental variation can cause consumer-resource pairs that would otherwise be stable to fluctuate, sometimes altering the long-term persistence of the system (Turchin 2003). However, most investigations of the interplay between species interactions and environmental variation have been theoretical, and our understanding of how these mechanisms manifest in natural populations is rudimentary.

In addition to the interplay of species interactions and environmental variation in the context of complex population dynamics, there is the long-standing interest in how single populations that have large influences on their ecosystems (Paine 1966). For some populations, these large effects arise as a direct consequence of their high abundance. "Dominant species" are those that are consistently abundant through time and have been shown to be important for

enhancing primary production and limiting temporal variability of plant communities (Smith and Knapp 2003, Sasaki and Lauenroth 2011). Some populations are only transiently abundant but can nonetheless have large effects on their ecosystems that can potentially persist beyond the period of high abundance (e.g. insect outbreaks and phytoplankton blooms) (Dokulil and Teubner 2000, Yang and Gratton 2014). In contrast to populations that whose large effects arise primarily from their high abundance, "keystone" species have impacts that are disproportionate to their abundance (i.e. they have large per capita effects). The classic example comes from rocky intertidal communities, where high predation rates by sea stars promote species richness by suppressing mussels that would otherwise become dominant (Paine 1966). Other keystone species are "ecosystem engineers"; they directly alter the physical environment experienced by themselves and other populations, with cascading consequences for the whole system (Jones et al. 1994, Bertness and Leonard 1997, Wright and Jones 2006). For example, reefs built by corals and dams built by beavers create complex habitats that support high primary productivity and biodiversity (Bellwood 2001; Wright *et al.* 2002). Regardless of whether a population's ecosystem effects arise from high abundance, high per capita effects, or both, these effects are often mediated by the environmental context and therefore have the potential to vary substantially through space and time (Hastings et al. 2007).

In this dissertation, I explore the connections between population dynamics, environmental variation, and ecosystem-consequences of single populations in the context of Lake Mývatn in northeastern Iceland. Mývatn is uniquely suited to exploring the intersection of these topics for two main reasons: (1) a relatively simple community dominated by a few taxa and (2) high primary productivity fueling high, but variable, abundance of those dominant taxa (Einarsson et al. 2004). Mývatn's relative simplicity reflects the low species diversity of Iceland

as a whole, likely a function of it both being geographically isolated and at a high latitude. Mývatn's high productivity results from high nutrient inputs from the warm and cold springs that are the primary hydrological sources of the lake. The springs vary in their nutrient concentrations, but the overall external loading is estimated as 1.5 g m^{-2} y^{-1} of nitrogen (N), 1.4 g m^{-2} y^{-1} of phosphorus (P), and 340 g m^{-2} y^{-1} of silicon (Si) (Ólafsson 1979). The lake is quite shallow, with a maximum depth in the main (south) basin of approximately 3.3m and of around 5m in the secondary (north) basin due to sediment dredging; prior to dredging the north basin was much shallower (max depth ≈ 2m) (Einarsson et al. 2004). Mývatn's shallowness means that during clear periods much light reaches the bottom of the lake, which in conjunction with the high nutrient availability supports high benthic primary production, mainly by diatoms.

Mývatn's high benthic production supports large populations of midges (Diptera: Chironomidae), which spend most of their lives as aquatic larvae before emerging as flying adults that primarily function as the dispersal and reproductive stage. While Mývatn contains more than 30 midge species of differing ecologies, by far the most abundant are the related species *Tanytarsus graclientus* and *Chironomus islandicus* of the tribe Chironominae (Gardarsson et al. 2004). In peak years, the Chironomae can exceed densities of 200,000 m^{-2} and compose two-thirds of Mývatn's secondary biomass production (Einarsson et al. 2004). They in turn serve as a major food source for higher trophic levels, including fish (e.g. arctic char) and ducks. In addition to their numerical dominance, the Chironominae are ecosystem engineers, constructing silk tubes and a network of silk that stabilizes and provides 3-dimensional structure to the sediment (Hölker et al. 2015, Herren et al. 2017). Furthermore, the midge have the potential to alter sediment biogeochemistry through nutrient excretion and bioturbation that alters the redox conditions in the sediment (Hölker et al. 2015, Baranov et al. 2016, Baranov

Viktor et al. 2016). By virtue of being both abundant and having large per capita effects on the structure of the benthic environment, the Chironominae have the potential for large effects on function of the lake ecosystem.

While the Chironominae dominate the benthic environment in years of high abundance, both *Tanytarsus* and *Chirnomus* display high-amplitude fluctuations with episodic crashes that leave the sediment nearly absent of midge larvae (Lindegaard and Jónasson 1979). The fluctuations of *Tanytarsus* are particularly dramatic, spanning 5 orders-of-magnitude with crashes occurring every 4 to 7 years. Long-term data on adult *Tanytarsus* suggest that their fluctuations are tied to food limitation, as the individual wing-lengths (a proxy of overall condition) tend to decline immediately before the population crashes (Einarsson et al. 2002). Furthermore, the fluctuating abundance of *Tanytarsus* larvae and diatoms across nearly 30 years inferred from a sediment core resemble consumer-resource cycles, with *Tanytarsus* as the consumers and diatoms as the resource (Einarsson et al. 2016). However, the erratic nature of *Tanytarsus* population fluctuations is atypical of classic consumer-resource cycles. A model fit to a time series of *Tanytarsus* abundance from 1976-2002 suggests that the population dynamics are best explained by the existence of two alternative attractors: a stable equilibrium at moderately-high density and a high-amplitude cycle (Ives et al. 2008). The high-amplitude cycle is driven by depletion and recovery of the algal food source, with detritus buffering the decline in the midges that in turn prevents the immediate recovery of the algae. It is this decoupling of the midge-algae interactions that allows the existence of multiple attractors. In the presence of environmental stochasticity, the population is repeatedly pushed between these alternative dynamical states, resulting in the aperiodic high amplitude fluctuations. While this model provides a good phenomenological description of the complex dynamics of *Tanytarsus*

populations, the hypothesized processes that underlie the model have not been empirically demonstrated. Therefore, a mechanistic understanding of the midge fluctuations has yet to emerge.

The cyanobacterium *Dolichospermum spp.* is another species that has large but variable influences on the Mývatn ecosystem. Blooms of *Dolichospermum* are historically episodic, occurring in some years but not in others, and spatially variable with north basin typically having the highest concentrations (Einarsson et al. 2004). When the blooms due occur, they can be very thick; concentrations of phycocyanin (a cyanobacteria pigment) can exceed concentrations of >200 $\mu g\ L^{-1}$ and the bloom can block virtually all light 5cm below the water surface. Under these conditions, the lake is likely to shift from being dominated by benthic to pelagic primary production, which in turn could have consequences for the movement of energy throughout the food web (Elser 1999, Dokulil and Teubner 2000). Furthermore, *Dolichospermum* can fix nitrogen and heterocysts have been observed in their colonies which suggests that they actively fix nitrogen during summer blooms at Mývatn; this possibility has been supported by some preliminary attempts to measure nitrogen fixation rates in the field. While the ultimate fate of this nitrogen is unknown, eventually it probably becomes part of Mývatn's internal nitrogen loading and may be available for use to other organisms. In any case, the Mývatn's cyanobacterial blooms likely have major consequences for its ecosystem function, as has been observed in many other lakes.

The work described in this dissertation explores a variety of topics related to the causes and consequences of variation in the abundance of various populations in Mývatn. In Chapter 1, I describe two field experiments that (1) investigate the mechanisms through which midges (mainly *Tanytarsus*) influence benthic primary production and (2) illustrate the role of light in

mediating this effect. The results illustrate that ecosystem engineering by midges can have either positive or negative effects on net ecosystem production in the sediment, depending on variation in light conditions. To place these experimental results in context, I used measurements of the spatial variability in water column light attenuation due to a thick cyanobacterial bloom in 2015 to project the expected variation in midge effects on benthic production given variation in light. These results illustrate the role of environmental variation in mediating the ecological consequences of ecosystem engineering.

Chapter 2 extends Chapter 1's theme of light effects on primary production to both finer and longer temporal scales. Using hourly observations of dissolved oxygen and related variables from 6 summers at Mývatn between 2012-2018, I fit a model that characterized the time-varying dependence of lake metabolism on light and temperature. To do so, I extended a previous modeling approach by allowing the model parameters to change smoothly through time as a random walk, reflecting changes in the underlying ecological dynamics. When fit to data from Mývatn, the model showed substantial temporal variability in the response of lake metabolism to light. In particular, the maximum rate of gross primary production under light-saturating conditions made the dominant contribution to temporal variation in lake metabolism in Mývatn. This variation in turn was closely related to variation in cyanobacterial pigments, suggesting that Mývatn's episodic cyanobacterial blooms can have important consequences for ecosystem function.

Finally, Chapter 3 describes an extension of the statistical approach presented in Chapter 2 to a very different context: the problem of inferring temporal variation in demographic rates (e.g. survival and recruitment) from time series of observed population abundance. The focus of this chapter is partly methodological, including an exploration of the statistical performance of

the method. However, the method is illustrated through application to the surveyed abundance of Mývatn's population of arctic char (*Salvelinus alpinus*). The char were historically an important fishery prior to a decline that followed very high exploitation rates and other changes in the lake ecosystem (Gudbergsson 2004). After comparing the full model to reduced versions with different demographic rates fixed through time, I return to the fit of the model to the real data and describe how the estimated sequence of time-varying recruitment and survival can be used to partition their relative contributions to temporal variation in the per capita population growth rate. Per capita recruitment was the dominant contributor to variation in the arctic char population over the past 30 years. While the source of this variability is unknown, it is potentially connected to fluctuations in their food sources, including certain midges (mainly non-Chironominae) and zooplankton (many of which covary either positively or negatively with *Tanytarsus* abundance Einarsson and Örnólfsdóttir 2004, Webert et al. 2017).

References

Andrewartha, H. G., and L. C. Birch. 1954. The distribution and abundance of animals. University of Chicago Press.

Baranov, V., J. Lewandowski, P. Romeijn, G. Singer, and S. Krause. 2016. Effects of bioirrigation of non-biting midges (Diptera: Chironomidae) on lake sediment respiration. Scientific Reports 6.

Baranov Viktor, Lewandowski Jörg, and Krause Stefan. 2016. Bioturbation enhances the aerobic respiration of lake sediments in warming lakes. Biology Letters 12:20160448.

Bellwood, D. R. 2001. Regional-Scale Assembly Rules and Biodiversity of Coral Reefs. Science 292:1532–1535.

Benincà, E., V. Dakos, E. H. Van Nes, J. Huisman, M. Scheffer, A. E. U. Berger, and E. R. G. Shaw. 2011. Resonance of Plankton Communities with Temperature Fluctuations. The American Naturalist 178:E85–E95.

Bertness, M. D., and G. H. Leonard. 1997. The Role of Positive Interactions in Communities: Lessons from Intertidal Habitats. Ecology 78:1976–1989.

Davidson, J., and H. G. Andrewartha. 1948. The Influence of Rainfall, Evaporation and Atmospheric Temperature on Fluctuations in the Size of a Natural Population of Thrips imaginis (Thysanoptera). Journal of Animal Ecology 17:200–222.

Dokulil, M. T., and K. Teubner. 2000. Cyanobacterial dominance in lakes. Hydrobiologia 438:1–12.

Dwyer, G., J. Dushoff, and S. H. Yee. 2004. The combined effects of pathogens and predators on insect outbreaks. Nature 430:341–345.

Einarsson, Á., A. Gardarsson, G. M. Gíslason, and A. R. Ives. 2002. Consumer–resource interactions and cyclic population dynamics of Tanytarsus gracilentus (Diptera: Chironomidae). Journal of Animal Ecology 71:832–845.

Einarsson, Á., U. Hauptfleisch, P. R. Leavitt, and A. R. Ives. 2016. Identifying consumer–resource population dynamics using paleoecological data. Ecology 97:361–371.

Einarsson, A., and E. B. Örnólfsdóttir. 2004. Long-term changes in benthic Cladocera populations in Lake Myvatn, Iceland. Aquatic Ecology 38:253–262.

Einarsson, Á., G. Stefánsdóttir, H. Jóhannesson, J. S. Ólafsson, G. M. Gíslason, I. Wakana, G. Gudbergsson, and A. Gardarsson. 2004. The ecology of Lake Myvatn and the River Laxá: Variation in space and time. Aquatic Ecology 38:317–348.

Elser, J. J. 1999. The pathway to noxious cyanobacteria blooms in lakes: the food web as the final turn. Freshwater Biology 42:537–543.

Elton, C., and M. Nicholson. 1942. The Ten-Year Cycle in Numbers of the Lynx in Canada. The Journal of Animal Ecology 11:215.

Elton, C. S. 1924. Periodic fluctuations in the numbers of animals: their causes and effects. Journal of Experimental Biology 2:119–163.

Gardarsson, A., Á. Einarsson, G. M. Gíslason, T. Hrafnsdóttir, H. R. Ingvason, E. Jónsson, and J. S. Ólafsson. 2004. Population fluctuations of chironomid and simuliid Diptera at Myvatn in 1977–1996. Aquatic Ecology 38:209–217.

Gudbergsson, G. 2004. Arctic charr in Lake Myvatn: the centennial catch record in the light of recent stock estimates. Aquatic Ecology 38:271–285.

Gutiérrez, J. L., C. G. Jones, D. L. Strayer, and O. O. Iribarne. 2003. Mollusks as ecosystem engineers: the role of shell production in aquatic habitats. Oikos 101:79–90.

Hastings, A., J. E. Byers, J. A. Crooks, K. Cuddington, C. G. Jones, J. G. Lambrinos, T. S. Talley, and W. G. Wilson. 2007. Ecosystem engineering in space and time. Ecology Letters 10:153–164.

Herren, C. M., K. C. Webert, M. D. Drake, M. J. Vander Zanden, Á. Einarsson, A. R. Ives, and C. Gratton. 2017. Positive feedback between chironomids and algae creates net mutualism between benthic primary consumers and producers. Ecology 98:447–455.

Hoekman, D., J. Dreyer, R. D. Jackson, P. A. Townsend, and C. Gratton. 2011. Lake to land subsidies: experimental addition of aquatic insects increases terrestrial arthropod densities. Ecology 92:2063–2072.

Hölker, F., M. J. Vanni, J. J. Kuiper, C. Meile, H.-P. Grossart, P. Stief, R. Adrian, A. Lorke, O. Dellwig, A. Brand, M. Hupfer, W. M. Mooij, G. Nützmann, and J. Lewandowski. 2015. Tube-dwelling invertebrates: tiny ecosystem engineers have large effects in lake ecosystems. Ecological Monographs 85:333–351.

Ives, A. R., Á. Einarsson, V. A. A. Jansen, and A. Gardarsson. 2008. High-amplitude fluctuations and alternative dynamical states of midges in Lake Myvatn. Nature 452:84–87.

Jones, C. G., J. H. Lawton, and M. Shachak. 1994. Organisms as Ecosystem Engineers. Oikos 69:373–386.

Krebs, C. J., R. Boonstra, S. Boutin, and A. r. e. Sinclair. 2001. What Drives the 10-year Cycle of Snowshoe Hares? BioScience 51:25–35.

Lindegaard, C., and P. M. Jónasson. 1979. Abundance, Population Dynamics and Production of Zoobenthos in Lake Mývatn, Iceland. Oikos 32:202–227.

Nicholson, A. J., and V. A. Bailey. 1935. The Balance of Animal Populations.—Part I. Proceedings of the Zoological Society of London 105:551–598.

Ólafsson, J. 1979. The Chemistry of Lake Mývatn and River Laxá. Oikos 32:82–112.

Paine, R. T. 1966. Food Web Complexity and Species Diversity. The American Naturalist 100:65–75.

Ripa, J., and A. R. Ives. 2003. Food web dynamics in correlated and autocorrelated environments. Theoretical Population Biology 64:369–384.

Sasaki, T., and W. K. Lauenroth. 2011. Dominant species, rather than diversity, regulates temporal stability of plant communities. Oecologia 166:761–768.

Smith, M. D., and A. K. Knapp. 2003. Dominant species maintain ecosystem function with non-random species loss. Ecology Letters 6:509–517.

Turchin, P. 2003. Complex population dynamics: a theoretical/empirical synthesis. Princeton University Press.

Turchin, P., L. Oksanen, P. Ekerholm, T. Oksanen, and H. Henttonen. 2000. Are lemmings prey or predators? Nature 405:562–565.

Webert, K. C., C. M. Herren, Á. Einarsson, M. Bartrons, U. Hauptfleisch, and A. R. Ives. 2017. Midge-stabilized sediment drives the composition of benthic cladoceran communities in Lake Mývatn, Iceland. Ecosphere 8:e01659.

Wright, J. P., and C. G. Jones. 2006. The concept of organisms as ecosystem engineers ten years on: progress, limitations, and challenges. BioScience 56:203–209.

Wright, J. P., C. G. Jones, and A. S. Flecker. 2002. An Ecosystem Engineer, the Beaver, Increases Species Richness at the Landscape Scale. Oecologia 132:96–101.

Yang, L. H. 2004. Periodical Cicadas as Resource Pulses in North American Forests. Science 306:1565–1567.

Yang, L. H., and C. Gratton. 2014. Insects as drivers of ecosystem processes. Current Opinion in Insect Science 2:26–32.

Ziebarth, N. L., K. C. Abbott, and A. R. Ives. 2010. Weak population regulation in ecological time series. Ecology Letters 13:21–31.

Chapter 1

Spatiotemporal variation in the sign and magnitude of ecosystem engineer effects on lake ecosystem production

Joseph S. Phillips[1], Amanda R. McCormick[1], Árni Einarsson[2,3], Shannon N. Grover[1], and Anthony R. Ives[1]

[1]Department of Integrative Biology, University of Wisconsin-Madison, Madison, WI, 53706

[2]Mývatn Research Station, Skútustaðir, Iceland, IS-660

[3]Faculty of Life and Environmental Sciences, University of Iceland, Reykjavik, Iceland, IS-101

Status: In press at *Ecosphere*

Abstract

Ecosystem engineers can have diverse and conflicting effects their ecosystems, and the balance between these effects can depend on the physical environment. This context-dependence means that environmental variation can produce large differences in engineer effects through space and time. Here, we explore how local variability in environmental conditions can lead to large spatiotemporal variation in the effect of tube-building midges on benthic ecosystem metabolism in a shallow subarctic lake. Using field experiments, we found that midge engineering increases both gross primary production (GPP) and respiration (RESP) in the sediment. GPP and RESP have opposing influences on net ecosystem production (NEP), and the net effect of midges on the benthic ecosystem depends on the balance between their effects on GPP and RESP. Variation in light mediates this balance—under high light conditions primary producers are able to exploit the structural benefits provided by midges, while in the dark the elevation of respiration from midge engineering predominates. Benthic light levels vary spatially and temporally due to episodic cyanobacterial blooms that prevent almost all light from reaching the benthos. By quantifying the nonlinear relationship between midge engineering and light, we were able to project ecosystem-wide consequences of natural variation in light conditions across the lake. Our results illustrate how the sign and magnitude of ecosystem-wide effects of ecosystem engineers can vary through space and time.

Key Words: *Chironomus islandicus*; context-dependence; lake metabolism; light limitation; nutrient limitation; macroinvertebrates; *Tanytarsus gracilientus*

Introduction

Ecologists have long been interested in the potential for single populations to have disproportionate influences on their ecosystems (Paine 1966). Striking examples come from 'ecosystem engineering', whereby a single species alters the physical environment experienced by many others, with cascading consequences for the whole system (Jones et al. 1994, Bertness and Leonard 1997, Wright and Jones 2006). Different organisms in a community respond to environmental changes in different ways, leading to multiple effects of ecosystem engineering at the community and ecosystem scale (Jones et al. 1997, Hastings et al. 2007, Gribben et al. 2013). Ecosystem engineers also have non-engineering effects on other organisms within a community through direct interactions such as competition and herbivory (Jones et al. 1997, Bertness et al. 1999, Hastings et al. 2007, Largaespada et al. 2012). The net impact of ecosystem engineers depends on the balance between these diverse and potentially conflicting effects, which itself can depend on the environmental context (Bertness et al. 1999, Norkko et al. 2006, Daleo and Iribarne 2009, Brown and Lawson 2010, Lathlean and McQuaid 2017).

The context-dependence of ecosystem engineering means that environmental variation can produce large differences in engineer effects through space and time (Wright et al. 2006, Hastings et al. 2007). For example, many engineers provide refuges from stressful environments (e.g. temperature or physical disturbance), and the response of other organisms in the communities varies across gradients in stress (e.g. tidal height or latitude) (Bertness et al. 1999, Crain and Bertness 2005, Arribas et al. 2014). While previous studies have demonstrated the existence of variation in the effects of ecosystem engineering across a range of scales (Wright et al. 2006, McAfee et al. 2016, Lathlean and McQuaid 2017), few studies have attempted to predict the magnitude of spatiotemporal variation in engineering effects at the ecosystem scale.

Scaling small-scale engineering to ecosystem-wide effects requires (1) modeling the functional response of engineering to continuous variation in the environment and (2) quantifying the spatial and temporal variation in the environment within the region of interest (Wright et al. 2006, Hastings et al. 2007). This is particularly important because responses of engineering to environmental conditions are likely nonlinear (as is true for many ecological processes) and the shape of this nonlinearity could have large consequences for the overall effects of engineering in variable environments (Ruel et al. 1999).

Here, we explore how local variability in environmental conditions can lead to large spatiotemporal variation in the effect of tube-building midges on benthic ecosystem metabolism in Lake Mývatn, Iceland. Midges (Diptera: Chironomidae) are widespread in lakes worldwide and exemplify other benthic invertebrates (e.g. annelids and mollusks) in their roles as ecosystem engineers (Armitage et al. 1995, Gutiérrez et al. 2003, Norkko et al. 2006, Arribas et al. 2014, Hölker et al. 2015). Larval midges construct silk tubes and a network of silk that stabilizes and provides 3-dimensional structure to the sediment (Olafsson and Paterson 2004, Hölker et al. 2015) (Fig 1). In Mývatn, midge populations are highly variable but can exceed densities of 200,000 m^{-2} and compose a majority of animal biomass in peak years (Lindegaard and Jónasson 1979, Einarsson et al. 2004). Therefore, Mývatn's midges have the potential to greatly alter benthic ecosystem function (Fig 2). We examined midge effects on net ecosystem production (NEP), which equals gross primary production (GPP) minus respiration (RESP) (Lovett et al. 2006, Chapin et al. 2006). NEP is a central component of ecosystem carbon budgets that influences the biomass available across trophic levels and the exchange of dissolved CO_2 with the atmosphere (Randerson et al. 2002, Cebrian and Lartigue 2004, Chapin et al. 2006, Davidson et al. 2015). There is much current interest in the biotic and abiotic factors determining the

balance between GPP and RESP in both aquatic and terrestrial systems, particularly in the context of global environmental change (Belshe et al. 2013, Holgerson and Raymond 2016, Demars et al. 2016).

Previous work showed that midges can increase GPP and producer biomass, possibly by providing a substrate for algal growth; chlorophyll concentrations can be twice as high in the tubes compared to the surrounding sediment (Pringle 1985, Olafsson and Paterson 2004, Hölker et al. 2015, Herren et al. 2017). However, midges may also decrease GPP through direct consumption of primary producers. Midges may ameliorate short-term nutrient limitation of GPP through mobilization (e.g. excretion or bioturbation) of N and P, but can also reduce P availability by oxidizing the sediment (Henriksen et al. 1983, Svensson 1997, Zhang et al. 2010, Hölker et al. 2015, Benelli et al. 2018). Furthermore, sediment oxygenation can increase microbial RESP by changing redox conditions and inducing shifts in bacterial community composition (Svensson and Leonardson 1996, Yeager et al. 2001, Nogaro et al. 2008, Hölker et al. 2015, Baranov et al. 2016a, 2016b). Therefore, midges may have opposing effects on NEP through their various effects on GPP and RESP (Fig 2). Determining how the physical environment regulates the balance of midge effects on production and respiration is important for understanding their role in the functioning of lake ecosystems (Hölker et al. 2015).

With two field experiments, we (1) explored the mechanisms by which midges stimulate benthic GPP and RESP and (2) quantified the role of light in mediating the balance between these effects and the consequences for NEP. In the first experiment, we experimentally mimicked two ways in which midges alter the benthic environment: modification of sediment structure and nutrient mobilization. We predicted that if midges stimulate ecosystem metabolism through these mechanisms, then experimental manipulations should (i) approximate the effects

of midges when midges are absent and (ii) be at least partially redundant with midges such that they have weaker effects when midges are present. We were interested in comparing the effects of midges through sediment structure and nutrient mobilization, because structural effects would be largely predictable from the presence of midges (Hölker et al. 2015). In contrast effects of nutrient mobilization would depend on the ambient nutrient content in the sediment pore water, which is highly dynamic in Mývatn (Gíslason et al. 2004).

In the second experiment, we quantified NEP across a light gradient in the presence and absence of midges. We expected the positive effect of midges on GPP to decline as photosynthesis became more limited by light, as the potential benefits provide by midges could only be exploited by the primary producers in the presence of sufficient light for photosynthesis. In contrast, we expected the positive effects of midge on RESP (at least from heterotrophs) to be present in both light and dark conditions. Therefore, a decline in light would tip the balance between positive midge effects on GPP and ER, making the midge effect on NEP less positive or more negative (Fig 2). In Mývatn, benthic light levels vary spatially and temporally due to variation in depth and episodic cyanobacterial blooms that prevent almost all light from reaching the benthos (Einarsson et al. 2004). To investigate how this spatiotemporal variation in light intensity can alter midge effects on NEP, we combined data on patterns of light transmission throughout the lake with a model parameterized from our field experiments. This allowed us to project the potential consequences of spatiotemporal variation in environmental conditions for determining the effect of midge engineering on NEP at the whole-lake scale.

Experimental Methods

Study System

Mývatn is located in northeastern Iceland (65°40'N 17°00'W) and has a tundra-subarctic climate (Björnsson and Jónsson 2004). The lake is shallow (mean depth ≈ 2.5m and max depth ≈ 4m in main basin) and is fed by cold and warm springs rich in phosphorus (N, P, and Si loading of 1.5, 1.4, and 340 g m^{-2} y^{-1} respectively) (Ólafsson 1979, Einarsson et al. 2004). Most of the primary production (roughly 288 g C m^{-2} y^{-1}) is benthic (Einarsson et al. 2004) as is likely true for many shallow lakes (Vadeboncoeur et al. 2002, Karlsson et al. 2009). Diatoms are the dominant benthic producers, with the genus *Fragilaria* composing most of the algal biomass (Jónasson 1979, Einarsson et al. 2004). Benthic primary production supports large populations of midges (>30 species), and the dominant tube-building genera *Tanytarsus* (tribe Tanytarsini) and *Chironomus* (tribe Chironomini) compose a majority of the lake's animal biomass in peak years. *Tanytarsus* generally has two generations per year accompanied by large emergences of adults that greatly reduce the number of larvae in the sediment. Furthermore, many of Mývatn's tube-building midge show large interannual fluctuations in abundance, with *Tanytarsus* spanning several orders of magnitude between high- and low-midge years (Ives et al. 2008). It is unknown how long the midge tubes persist following emergence, but they probably degrade within a few weeks.

Mesocosms

We performed two field experiments to explore (1) the mechanisms of positive midge effects and (2) light mediation of midge effects on benthic ecosystem metabolism. The experiments used mesocosms constructed from acrylic tubes (33cm height × 5cm diameter) with the bottom 15cm filled with sediment and the top 18cm with lake water. In both experiments we

manipulated the presence/absence of midges and their associated engineering to quantify their effects on ecosystem metabolism. A previous experiment at Mývatn directly demonstrated positive effects of midges on GPP and NEP by adding different midge densities to sediment that was sieved to remove any naturally occurring midges (Herren et al. 2017). In contrast, our goal was to quantify the magnitude of midge effects under the most realistic conditions possible, so we could place midge engineering in the context of natural environmental variability. Adding midges to sieved sediment does not accomplish this, because it takes time for the midges to recover from being relocated and for the impact of midge tube building to be fully realized in the sediment structure. Instead, we filled the mesocosms with either intact sediment cores ('midges present') or sediment sieved to remove midges and disrupt the structure created by midge tubes and silk ('midge absent'). Mývatn's sediment is composed of very fine particles (primarily diatom frustules and volcanic tephra; (Einarsson et al. 2004), and in the natural absence of midges the sediment has a very loose structure that was closely mimicked by the sieved sediment (Olafsson and Paterson 2004). Therefore, our midge treatments reflected the natural contrast in sediment structure between years of high and low midge abundance.

For the 'midges present' treatments, we extruded the top 15cm of sediment cores containing larval midges into each mesocosm without disturbing the sediment. For 'midges absent', we sieved the top 5cm (containing most of the active algal cells) and bottom 10cm of sediment cores through 63µm and 125µm mesh respectively, pooled the two sediment layers from all cores, and reassembled them as two layers in the mesocosms. The use of 63µm mesh for the top layer was necessary to remove first instar larvae that typically reside near the sediment surface. Sieved sediment was stored in a cool dark location for 2-3 days prior to establishing the mesocosms so that the sediment could settle, while cores for 'midges present' were collected on

the day each experiment was deployed. The sediment cores were collected from Mývatn using a Kajak corer, at times and locations of moderate larval midge density (15,000-25,000 m^{-2}). For both 'midges present' and 'midges absent' treatments the bottoms of the mesocosms were sealed with foam stoppers and tape. The bottom 15cm of each mesocosm was wrapped with four layers of black plastic to replicate the naturally dark conditions within the lake sediment.

Sieving and storing of sediment was clearly a large perturbation that could have altered algal abundance, productivity, or microbial activity. However, chlorophyll-a concentrations were similar between sieved and intact sediment (Appendix A). Furthermore, there were no indications of transient shifts in ecosystem metabolism of the sieved treatments that were not also present in the intact treatments; the sieved treatments were actually more consistent through time (*Results*; Appendix A). While this does not demonstrate that sieving had no artificial effects, it does suggest that the sieved treatments quickly reached an equilibrium state with their surrounding environment (consistent with the fast turnover times of diatoms that are the dominant benthic producers) (McCormick and Stevenson 1991, Sommer 1991) and provide a consistent baseline for evaluating the effects of midges engineering.

The mesocosms were deployed on the bottom of the lake for either 11 (experiment 1) or 15 (experiment 2) days at a depth of 1m. We left the tops of the mesocosms open to allow exchange with the ambient lake water, except during measurements of metabolism when they were sealed with rubber stoppers for several hours (see *Measurements of NEP, GPP, and RESP*). Mývatn is spring fed and has substantial lateral water flow even on days with low wind, so there was likely significant exchange between the mesocosms and overlying water (Bartrons et al. 2015).

Experiment 1: Mechanism of Midge Engineering

We used a factorial experiment to test two mechanisms by which midges may increase ecosystem metabolism rates: (1) enhanced substrate quality by building silk tubes and (2) promoting algal growth by mobilizing nutrients. The experiment had a 2 × 2 × 2 design crossing midge presence with artificial silk structures (to mimic midge tubes) and nutrient enrichment (to mimic nutrient mobilization), and we measured the responses of GPP, RESP, and chlorophyll-a. We predicted that if tube building by midges stimulated algal growth, then the presence of either silk or midges would increase GPP by a comparable amount. However, if the benefits of enhanced sediment structure (provided by either midges or silk) saturate, then the presence of *both* midges and silk should be lower than what would be expected from their separate effects. Similarly, we predicted that if nutrient mobilization by midges stimulated GPP, then midges and nutrient enrichment would both increase GPP when applied separately but that these effects would be at least partially redundant when applied together. We expected RESP to respond similarly to GPP, as is often the case in lake ecosystems with high *in situ* production. Since GPP and RESP have opposite effects on NEP, the effects of midges, silk, and nutrients would depend on the relative magnitude of their effects on GPP and RESP.

The presence of midges and associated sediment structure was manipulated as described in *Methods: Mesocosms*. To mimic midge engineering, we used thin sheets of natural silkworm silk (Undyed Silk Hankies, Yarn Designers Boutique, CA, USA) originally 25 × 25cm that we stretched into loops approximately 15cm in diameter. Each mesocosm received a single silk loop, which we loosely coiled into a 3-dimensional structure and positioned so that the bottom half of the loop sat below the sediment surface (Fig A1). While the silk coils were clearly not a perfect imitation of midge tubes, they are similar in that they provided 3-dimensional structure

that expanded the effective surface area available for algal growth. The silk could present a physical barrier reducing between the sediment and the water column, although this also is similar to the effect of the midge tubes and silk that can completely cover the sediment surface.

We manipulated nutrient concentration with agar rods (1.5cm diameter × 15cm length) either enriched with both N and P or not enriched (Tank et al. 2006) that were pushed vertically into the mesocosm sediment. While N and P are not the only nutrient that could be limited for benthic primary producers (particularly Si for diatoms) (Kilham 1971, Paasche 1973), N and P are the nutrients that are most obviously influenced by midges and are also the focus of most studies of nutrient limitation in freshwater systems (Svensson and Leonardson 1996, Hölker et al. 2015). We enriched the agar with NH_4Cl and KH_2PO_4 to produce N and P concentrations of 0.5M and 0.13M respectively, to roughly mimic the excretion ratios measured from *Tanytarsus* (Herren et al. 2017). Laboratory measurements of similarly enriched agar rods yielded average N and P release rates of 22.8 and 1.23 mg day^{-1}, respectively (Appendix A). The rods were constructed by pouring the hot agar into clean surgical tubing cut to the appropriate length, with a bamboo skewer inserted into the center for stability.

We secured the mesocosms to 2 racks with 12 mesocosms each. We intended to use a balanced design with 3 replicates in each of 8 treatment combinations (total n=24). However, due to a logistical error one 'silk present' and 'silk absent' treatment were switched for each nutrient and midge combination, so that each treatment had either 2 or 4 replicates. We deployed the experiment on the lake bottom near the southern shore from 2-13 August 2016 and measured GPP, RESP, and NEP on the 3rd and 11th days of deployment (see *Measurements of NEP, GPP, and RESP)*. On day 11 we collected 1mL of surface sediment from each mesocosm (after removing the silk in the silk treatment) and quantified chlorophyll-a concentration extracted in

methanol with a fluorometer (Turner Designs, Sunnyvale, CA, USA) (Welschmeyer 1994, Herren et al. 2017). We also measured the chlorophyll-a in the silk loops from each mesocosm in the 'silk present' treatment; chlorophyll-a was extracted from the entire silk loop using methanol, following the same procedure as for the sediment.

Experiment 2: Light Mediation of Midge Engineering Effects

We quantified the relationship between light and midge effects on benthic production by measuring NEP across a light gradient in the presence or absence of midges with their associated engineering. The experiment consisted of 48 mesocosms filled with sieved and intact sediment (*Methods: Mesocosms*). To create variation in light, we wrapped the sides of the upper 18cm of 40 mesocosms with 0, 1, 2, 3, or 4 layers of white mesh. We wrapped two layers of black plastic around the sides of the upper 18cm of the remaining 8 mesocosms. There was no algal growth or other fouling on the sides of the mesocosms during experiment, so the shading treatments remained consistent. Altogether we had 6 shading treatments with 8 replicates each. These replicates were evenly distributed across the two midge treatments, for a total of 12 treatment combinations.

We secured the mesocosms to 3 racks with 16 mesocosms each. Thirty-six replicates were distributed among racks in a complete block design, while the remaining 12 mesocosms were distributed haphazardly. We deployed the experiment in a small bay in the southeastern corner of Mývatn, free of the thick cyanobacterial bloom occurring across the rest of the lake that prevented most of the light from penetrating to a depth of 1m. *Tanytarsus* and *Chironomus* are absent from this portion of the lake (due to the hard substrate) and therefore could not colonize the mesocosms. This portion of the lake is near the springs that feed the main basin, and have substantially lower N:P ratios (~1:1) than the rest of the lake (~16:1 for the outlet on the opposite

side of the lake) . However, the nutrient content in the sediment pore-water of the mesocosms was likely much higher than in the water column, so differences in water column chemistry likely had minimal effects (Gíslason et al. 2004). The experiment was deployed from 28 July to 12 August 2015, and we measured NEP on 3rd and 15th days of deployment (see *Measurements of NEP, GPP, and RESP*).

To determine the effect of shading treatments on *in situ* light levels at the mesocosm sediment surface, we measured the photosynthetically active radiation (PAR) inside of a clear polycarbonate tube secured with the appropriate shading treatment. We secured a light meter (Li-192 Quantum Underwater Sensor, Li-COR, NE, USA) 18cm from the top of each tube and took the readings with the sensor just below the water's surface, at 0.5 below the surface, and 1m below the surface. For each set of readings, we also measured ambient light with the meter outside of the tube. This allowed us to calculate the effect of a given shading treatment as a fraction of surface irradiance at a given depth. We repeated each series of measurements 3 times so we could average across changes in ambient light intensity within a series of measurements at different depths. Throughout the incubation period for measuring NEP (see *Measurements of NEP, GPP, and RESP*), we measured PAR using the bare sensor at the locations above/below the water surface as described above. This allowed us to estimate the light level experienced by each mesocosm during the incubation, given its shading treatment and the ambient light level. Estimated light levels at the sediment surface ranged from 3.43 to 228 μmol-photons m^{-2} s^{-1}. From our long-term sampling location in May-August of 2013-2017, approximately 77% of light levels observed at midday from a depth of 2.5m (Mývatn's average depth) were within the experimental range. Day 15 was sunnier than day 3, so the most extreme mesocosm light levels

were greater for day 15, although with substantial overlap with day 3 due to the shading treatments.

To evaluate the efficacy of our midge (sieving) treatments and determine how much the larval midges grew, at the end of the experiment we sieved sediment from all cores in the full light and full dark shading treatments (n=16) and collected the larval midges (n=299). We identified the midges to either tribe (Tanytarsini or Chironomini) or subfamily (Appendix A). There was minimal colonization of the mesocosms by *Micropsectra* (tribe Tanytarsini); however, the 'midges absent' treatment remained largely midge-free. Tanytarsini dominated the 'midges present' treatment. The head capsule widths and body lengths of Tanytarsini larva were measured to determine whether the midges grew throughout the experiment. We used the distribution of head capsule widths from both the experiment and long-term data of the Tanytarsini population to define boundaries between instars (< 0.15mm for second instar, 0.15 - 0.24mm for third instar, and > 0.24mm for fourth instar). First instars were not present in the lake during the experimental period. We quantified midge growth by comparing the midge measurements from the mesocosms to those in sediment cores taken from the source location a few days before the beginning (n=133) and on the last day of the experiment (n=97).

Measurements of NEP, GPP, and RESP

Ecosystem metabolism rates were quantified in each mesocosm as the difference between the final and initial dissolved oxygen concentration per hour (g O_2 L^{-1} h^{-1}) over an incubation period of approximately 4 hours (ProODO, YSI Co., OH, USA) (Bott 2006, Herren et al. 2017). The incubations were performed on the lake bottom at a depth of 1m, with the mesocosms sealed with rubber stoppers for the duration (stoppers lead to an approximately 10% reduction in light at the mesocosm sediment surface). The mesocosm water columns were gently stirred to

homogenize the oxygen concentration before taking readings. Metabolism in the mesocosm water columns was negligible, as illustrated by experiments with similar mesocosms lacking sediment where changes in O_2 were not detectible. Therefore, we converted the differences in final and initial O_2 concentrations to fluxes of O_2 across the sediment surface (g O_2 m^{-2} h^{-1}) by multiplying by the depth of the mesocosm water column. We present ecosystem metabolism rates as fluxes of O_2, assuming that GPP and RESP correspond to roughly equal fluxes of C and that the net O_2 flux (NEP) is biologically meaningful (Bott 2006). We recognize that measuring metabolism through closed incubations has the potential to introduce some artifact (e.g. lag of water column mixing), but this approach has been used in previous studies (Bott 2006) and should provide a reasonable quantification of variation across treatments that is the major focus of our study.

For experiment 1, both NEP and RESP were quantified in sequential 4-hour incubations under light and dark conditions, respectively. Dark conditions were achieved by wrapping the tops of the mesocosms with four layers of black plastic, which blocked out essentially all light when the rubber stoppers were in place. GPP was calculated as the summed magnitudes of NEP and RESP (since NEP = GPP – RESP) (Lovett et al. 2006). The average water temperatures inside the mesocosms during the light and dark incubations were respectively 14.4°C and 15.9°C on day 3, and they were 14.7°C and 13.7°C on day 11. Average light levels at the deployment depth during the light incubation were 335 and 155 μmol-photons m^{-2} s^{-1} on days 3 and 11, respectively.

For experiment 2, we incubated each mesocosms once per measurement day with the shading treatments intact (as opposed to twice per measurement day, one light and one dark, as for experiment 1). This allowed us to quantify NEP across a light gradient, which in turn allowed

us to draw inferences about GPP and RESP across midge treatments (but not for each mesocosm separately; see *Statistical Methods*). Average water temperatures were 9.13°C on day 3 and 10.2°C on day 15 (see above for description of light levels during Experiment 2). We note that water temperatures were lower for Experiment 2 than Experiment 1, which was reflected in lower overall metabolism rates.

Statistical Methods

All analyses were conducted using R version 3.4.3 (R Core Team 2016). Code and data can be found at https://github.com/jsphillips2/midge_light.

Experiment 1: Mechanism of Midge Engineering

We fit separate linear mixed models (LMMs) for GPP, RESP, and NEP with two-way interactions between silk, midges, nutrients, and time. Mesocosm identity was included a random effect to account for repeated measures. A single extreme value was omitted for day 11, which did not alter the statistical conclusions. We calculated p-values using F-tests with the Kenward-Roger approximation, which can be applied to unbalanced designs (Luke 2017). We performed a parallel analysis to compare chlorophyll-a across treatments, although these data were only available for day 11 so there was no need to account for repeated measures. Unless otherwise noted we reported p-values for the full models, due to the risk of inflated Type-I errors following model selection (Freedman 1983). However, we also dropped non-significant interactions with backward selection to check for changes in the inference for corresponding fixed effects.

Experiment 2: Light Mediation of Midge Engineering Effects

Experiment 2 quantified NEP across a light gradient, with and without midges, and at two time points (days 3 and 15). We analyzed these data by fitting a modified productivity-irradiance curve with a Michaelis-Menten form (Jassby and Platt 1976):

$$\mathrm{NEP}_i = \alpha_i \times \mathrm{light}_i\,(1 + \beta \times \mathrm{light}_i) + \delta_i + \varepsilon_{\mathrm{rack/mescosm[i]}} \quad (1)$$

$$\alpha_i = \alpha_1 + \alpha_2 \times \mathrm{midge}_i + \alpha_3 \times \mathrm{time}_i + \alpha_4 \times \mathrm{midge}_i \times \mathrm{time}_i$$

$$\delta_i = \delta_1 + \delta_2 \times \mathrm{midge}_i + \delta_3 \times \mathrm{time}_i + \delta_4 \times \mathrm{midge}_i \times \mathrm{time}_i$$

where '*i*' is an index for observations, ε is the residual error nested within rack and mesocosm, 'light' is the *in situ* light level (PAR; μmol-photons m^{-2} s^{-1}), 'midge' and 'time' are binary indicators for midge treatment and time, and α_1, α_2, α_3, α_4, β, δ_1, δ_2, δ_3, and δ_4 are parameters to be estimated. The term α_i is the maximum rate of increase in NEP, and β is the rate at which NEP saturates with increasing light. The term δ_i captures light-independent effects on NEP. The hierarchical forms of α_i and δ_i allowed us to test the effects of midge presence, time, and their interactions on the light-dependent and independent components of NEP. In general, the light-dependent terms should be related to GPP and light-independent terms to RESP (Attard et al. 2014), although possible changes in algal biomass or microbial activity across light treatments potentially complicate this interpretation for day 15.

We evaluated the importance of the model parameters by fitting reduced versions of the full model that excluded different parameters and then ranked the models by their AIC scores (Burnham and Anderson 2002). This resulted in 25 models comprising all combinations of midge and time effects, with the restriction that models including midge × time interactions for either α or δ also included the corresponding main effects. The models were fit using maximum

likelihood, with normally distributed residuals and random effects to account for blocking (rack) and repeated measures (mesocosm). Following model selection, we refit the "optimal model" using restricted maximum likelihood for presentation of its results.

We quantified changes in midge stage structure by comparing the proportion of Tanytarsini individuals in second instar between the end of the experiment and (1) in-lake samples immediately prior to the experiment (to approximate the starting conditions for the experiment) and (2) in-lake samples at the end of the experiment with separate binomial generalized linear mixed models (GLMMs) including observation-level random effects to account for overdispersion. This allowed us to compare changes in midge stage structure in the mesocosms and the lake. Fewer than 5% of individuals were fourth instar, so this analysis essentially compared the proportions of second vs. third instars. To quantify midge growth, we performed a parallel analysis on body lengths using LMMs, with random effects for mesocosm or sediment core. We calculated p-values with either likelihood ratio tests (GLMMs) or F-tests using the Kenward-Roger approximation (LMMs) (Luke 2017).

We used the average densities of Tanytarsini (predominantly *Tanytarsus gracilentus*) and Chironomini (*Chironomus spp.*) observed in the 'midges present' mesocosms to estimate the potential contribution of respiration by midges themselves to NEP. We used these two taxa because Tanytarsini composed most of the individuals in the mesocosms and *Chironomus* are much more massive than any of the other taxa and therefore would contribute disproportionately to total respiration per individual. We used measurements of individual dry weights for these two taxa from (Herren et al. 2017) and literature values (Brodersen et al. 2004) for biomass-specific respiration measured at 10°C (the approximate summer average water temperature in Mývatn) to estimate total respiration for the midge densities observed in the mesocosms. We averaged

different values for *Tanytarsus gracilientus* and *Chironomus spp.* (*C. hyperboreus* and *C. riparius*—values for the main *Chironomus* species in Mývatn, *C. islandicus*, were not reported). This yielded the following estimates of total taxon-specific respirations:

Tanytarsini*:* 118.2 individuals × 0.0769 mg individual^{-1} × 7.7 μg O_2 h^{-1} mg^{-1} = 70.00 μg O_2 h^{-1}

Chironomini: 7.5 individuals × 0.6812 mg individual^{-1} × 1.862 μg O_2 h^{-1} mg^{-1} = 9.51 μg O_2 h^{-1}

Therefore, the combined respiration of Tanytarsini and Chironomini was 79.50 μg O_2 h^{-1}, or 0.0000795 g O_2 h^{-1}.

Lake-Scale Effects of Midges on NEP

To investigate how natural variation in benthic light intensity may alter midge effects on NEP, we combined the model fit to data from experiment 2 (equation 1) with field observations of light transmission to the lake bottom to project the potential spatiotemporal variation in midge effects on benthic production across the lake. The purpose of this projection was to place our experimental results in a broader context by illustrating how much variation one would expect in midge engineering effects across the lake given natural variation in light. This was in contrast to predicting the actual lake-wide effect of midges on NEP, which would require both characterizing the response of ecosystem metabolism to a gradient of midge densities and fine resolution of data on spatial variability in midge density, both of which were beyond the scope of this study.

We calculated the net midge effect as the partial derivative of equation 1 with respect to the binary midge index (identical to taking the difference between the equation evaluated with

and without midges). For the full version of the model, the net midge effect as a function of light level and time was calculated as

$$\frac{\partial \mathrm{NEP}}{\partial \mathrm{midge}} = (\alpha_2 + \alpha_4 \times \mathrm{time}) \times \mathrm{light} / (1 + \beta \times \mathrm{light}) + \delta 2 + \delta 4 \times \mathrm{time} \qquad (2)$$

with 'light' as continuous variation in PAR, 'time' as a binary indicator for the day of the experiment, and all parameters defined as above. The parameter values were based on the estimates associated with the final model as determined through the model selection procedure (terms that were not included in the final model were set to 0). Projections of midge effects at each time point were calculated by substituting light levels into this expression.

In July and August 2015 Mývatn experienced a thick cyanobacterial bloom, during which we took light readings at 65 sites around the lake to determine the local rates of light attenuation (Appendix A). The sites were positioned on a 500×500m grid and at each site we recorded light levels at 0.5m intervals between the surface and 2m. We estimated the exponential light attenuation at each site, which characterized the spatial variation in the severity of the bloom. The spatial pattern of the bloom roughly matched those from 2011 (Bartrons et al. 2015) and 2016-2017 (unpublished data) based on measurement of cyanobacterial pigments and cell counts. For each site we calculated the light level on the lake bottom assuming moderate surface light of 320 μmol-photons m^{-2} s^{-1} and one of three water clarity scenarios: low clarity (100% measured light attenuation), medium clarity (40%), or high clarity (15%). Surface light was set to ensure that benthic light levels were within the range of observations over which the model was fit (which included approximately 77% of light levels observed at midday from a depth of 2.5m at our long-term sampling location in May-August of 2013-2017). Spatial variation in benthic light levels reflected both differences in depth and observed bloom severity. For each water clarity

scenario, we quantified the site-specific midge effect by evaluating equation (2) for the corresponding light level. We made projections using both time points in equation (2), corresponding to midge communities dominated by either second or third instar Tanytarsini (see RESULTS). This resulted in six scenarios of water clarity and midge stage.

Results

Experiment 1: Mechanism of Midge Engineering

Experiment 1 tested two mechanisms by which midges may stimulate GPP and RESP: enhanced substrate quality and nutrient mobilization. GPP was higher in the presence of midges ($F_{1,24.78}$=6.34, p=0.019) and silk ($F_{1,25.64}$=6.83, p=0.015) than in their absence, and these effects were of comparable magnitudes (Fig 3a,b). However, there was a significant negative interaction between these effects ($F_{1,15.33}$=16.31, p=0.001) (Fig 3a,b). While GPP was higher with both silk and midges than with neither, GPP with both was comparable to (or perhaps slightly lower than) GPP with only one or the other. This is consistent with the possibility that experimental silk and midges provide similar net benefits to GPP that saturate when both are present, although it could also be due to the silk directly altering the effects of the midges. Nutrient enrichment had a negative effect on GPP ($F_{1,27.04}$=6.65, p=0.016), although it was only visually obvious in the day 3 data and was only marginally significant when non-significant interactions including nutrients were dropped. GPP increased from day 3 to 11 across all treatments ($F_{1,17.90}$=5.0, p=0.038). This implies that nutrients were not limiting and therefore that nutrient mobilization could not explain the positive effect of midges on GPP in the experimental mesocosms.

Paralleling GPP, RESP increased with midges ($F_{1,24.78}$=5.85, p=0.023) and silk ($F_{1,25.64}$= 5.07, p=0.033), but with a negative interaction ($F_{1,16.31}$=9.58, p=0.007) (Fig 3c,d). There was a nutrient × day interaction ($F_{1,18.95}$=5.16, p=0.035), with the nutrient effect being relatively more

negative on day 3 than day 11. By definition GPP and RESP had opposite effects on NEP, but because the overall magnitude of GPP was greater than RESP, GPP had a disproportionate influence. Therefore, the response of NEP to the experimental treatments broadly paralleled GPP, with positive effects of midges ($F_{1,24.78}$=4.33, p=0.048) and silk ($F_{1,25.64}$=5.63, p=0.025), and negative midge × silk interaction ($F_{1,15.52}$=16.31, p=0.001), and negative effects of nutrients ($F_{1,27.04}$=11.90, p= 0.002) (Fig A2). NEP was higher on day 11 than day 3 ($F_{1,17.09}$=15.12, p=0.001), and there was a positive midge × day interaction ($F_{1,18.89}$=4.61, p= 0.045).

Sediment chlorophyll-a was higher with midges (mean±standard error = 21.4±1.6 mg/L) than without (16.2±0.8 mg/L) ($F_{1,20.0}$=9.87, p=0.005), although this effect was only significant when the nonsignificant interactions with silk and nutrients were dropped. There were no significant effects of silk ($F_{1,20}$=0.74, p=0.400) or nutrients ($F_{1,20}$=0.15, p=0.706) on chlorophyll-a in the sediment itself (i.e. with the silk removed) (p-values for the reduced model). On the silk itself, chlorophyll-a concentrations were >10 times (254±50 mg/L) than in the sediment (regardless of silk treatment).

Experiment 2: Light Mediation of Midge Engineering Effects

Experiment 2 quantified the net effect of midges on benthic NEP across a light gradient. We fit models (equation 1) of *in situ* NEP with different combinations of midge and time effects on light-dependent and independent processes. Of the 25 possible models, the three best were similarly well supported (ΔAIC < 2), and all remaining models had ΔAICs > 4 (Table 1). For simplicity, we present the results from the second-best model as it had the fewest parameters, noting that each of the three best models provide similar interpretations of the data. The second-

best model contained a light-dependent main effect of midges ('slopes' in Fig 4) and light-independent main and interaction effects of midges and time (y-intercepts in Fig 4).

Midges enhanced NEP at high light levels through their positive effect on GPP, but this effect declined as light decreased to levels too low to sustain algal photosynthesis (Fig 4). Furthermore, midges increased RESP, such that their net effect on NEP switched from positive to negative at low light levels. The direct contribution of midge respiration to total respiration was likely small (estimated as 0.0000795 g O_2 h^{-1}, while the difference in NEP between 'midges present' and 'midges absent' treatment in the dark was -0.044 g O_2 h^{-1}), indicating that midges stimulated respiration of other organisms in the sediment. NEP in the 'midges absent' mesocosms did not change between the beginning and end of the experiment (Appendix A; Fig A3), indicating an absence of transient artificial effects of the manipulation (i.e. sieving sediment) *per se*. In contrast, NEP declined substantially through time in the presence of midges, which resulted from an increase in midge stimulation of RESP (y-intercepts in Fig 4) while the midge effect on GPP ('slopes' in Fig 4) remained the same. Consequently, the amount of light required for the positive effects of midges on NEP to outweigh the negative effects was substantially higher at the end of the experiment (91 μmol-photons m^{-2} s^{-1}) than at the beginning (8 μmol-photons m^{-2} s^{-1}).

The change in midge effects on NEP through time coincided with a shift in midge stage structure (χ^2=13.35, df=1, p=0.0003). At the beginning of the experiment 26.3% of Tanytarsini larvae were third instar, while 72.2% were third instar by the end. Slightly more Tanytarsini were third instar in the lake (88.7%) than in the mesocosms at the end of the experiment, but this difference was not significant (χ^2=2.16, df=1, p=0.141). Tanytarsini body lengths exactly paralleled these results. The midges grew between the beginning (mean length=2.42mm) and end

(2.70mm) of the experiment ($F_{1,11.3}$=5.34, p=0.041), and were not significantly longer in the lake than in the mesocosms at the end (mean=3.10mm) ($F_{1,11.68}$=3.82, p=0.075). Approximating individual midges as cylinders and using the head capsule width as a proxy for diameter, the midges nearly doubled in volume (and therefore body mass) over the course of the experiment (Appendix A).

Lake-Scale Effects of Midges on NEP

We used equation 2 and field measurements of light attenuation during a cyanobacterial bloom to project potential midge effects on production across the lake. Based on the model selection described above, we employed the parameterization of the model that excluded the light-dependent midge × light interaction (α_4), yielding the following expression for the net midge effect:

$$\frac{\partial \text{NEP}}{\partial \text{midge}} = \alpha_2 \times \text{light} / (1 + \beta \times \text{light}) + \delta 2 + \delta 4 \times \text{time} \quad (3)$$

where α_2 = 0.00169 [(g O_2 m^{-2} h^{-1}/μmol-photons m^{-2} s^{-1})], β = 0.0202 [1/(μmol-photons m^{-2} s^{-1})], δ_2 = -0.0122 [g O_2 m^{-2} h^{-1}], and δ_4 = -0.0420 [(g O_2 m^{-2} h^{-1})/$time^{-1}$]. We made projections for low (100% observed light attenuation), medium (40%), and high (15%) water clarities crossed with two dominant midge stages (second and third instar, corresponding to days 3 and 15 of experiment 2). Across the six scenarios projected midge effects on benthic NEP ranged from strongly negative (-0.054 g O_2 m^{-2} h^{-1}) to strongly positive (0.053 g O_2 m^{-2} h^{-1}), thus spanning ~0.11 g O_2 m^{-2} h^{-1}. The estimated seasonal production by diatoms from Mývatn (200 g C m^{-2}) in a year of moderate midge abundance (Jónasson 1979) converted to the same timescale and units

is roughly 0.22 g O_2 m^{-2} h^{-1} (Appendix A), which means that the potential range of projected midge effects is substantial relative to the overall magnitude of production. In some cases, both moderately positive and moderately negative midge engineering effects occurred across the lake for a given water clarity scenario (Fig 5). Spatial variation was highest at medium water clarities, while midge effects at extreme water clarities tended to be either uniformly positive, negative, or neutral across the lake.

Discussion

Our results show how environmental conditions can alter the magnitude and sign of ecosystem engineer effects on their ecosystems: midges acting as both engineers and consumers increase benthic NEP when light is abundant but decrease NEP when light limits algal photosynthesis. Mývatn experiences patchy and episodic cyanobacterial blooms, which generate large variation in water clarity and benthic light levels across the lake, within a season, and between years. Therefore, the influence of midge ecosystem engineering on benthic production is spatiotemporally complex, and the positive effects of midges can be overshadowed by events in the pelagic zone. The majority of the world's lakes are shallow and likely to be limited by light (Karlsson et al. 2009, Cael et al. 2017), which is of increasing concern given increasing eutrophication and associated declines in water clarity (Vadeboncoeur et al. 2003, Taranu et al. 2015). Our results have implications for how such changes may produce spatiotemporal variation in the effects of benthic ecosystem engineers on the function of a variety of aquatic ecosystems.

We tested two hypothesized mechanisms for positive midge effects on benthic production at high light levels: (1) enhanced substrate quality by building silk tubes and (2) increased nutrient availability by mobilizing N and P. Experiment 1 supported the first hypothesis and

rejected the second. Artificial silk structures increased GPP in "midges absent" mesocosms to a level similar to the "midges present" mesocosms without silk. However, addition of silk to "midges present" mesocosms caused no further increase in GPP, suggesting that the potential benefit of artificial silk was already provided by the tubes and silk webs constructed by the midges (although this silk could also have had direct effects on the midges or their effects on GPP). Furthermore, chlorophyll-a was higher in the presence of midges and silk (on the silk itself), so it is likely that higher algal biomass supported by the enhanced substrate contributed to higher GPP (Pringle 1985). In contrast, addition of nitrogen and phosphorus slightly reduced GPP, implying that the algae were not strongly limited by those nutrients and hence would not benefit from nutrient mobilization by midges. The negative effect of nutrient enrichment could be due to a shift in the algal community (e.g. from diatom to green algae dominated) (Steinman et al. 2016) or from artifacts of the agar preparation such as release of hydrogen peroxide (Tanaka et al. 2014). An additional caveat is the potential for benthic production to be limited by nutrients other than N and P (e.g. Si for benthic diatoms) (Kilham 1971, Paasche 1973), although these two elements are the nutrients that are most obviously influenced by midges (Hölker et al. 2015) and are also the focus of many studies of nutrient limitation in freshwater systems (Elser et al. 2007).

The negative effect of midges on NEP at low light levels resulted from their effect on total respiration. In both experiments midges increased total respiration, yet respiration by midges themselves likely made a small contribution to this effect. Bioturbation by midges and other sediment-dwelling invertebrates can stimulate aerobic respiration by microbes through oxidizing the sediment and shifting bacterial community composition (Yeager et al. 2001, Hölker et al. 2015, Baranov et al. 2016a, 2016b). Furthermore, defecation by midges could provide a

carbon substrate for microbial respiration (Svensson and Leonardson 1996). Midge effects on RESP increased throughout experiment 2, causing the apparent effect of midges on NEP to become less positive or more negative. This change was accompanied by a large shift in size- and age-structure of the midge populations; the midges shifted from second to third instar dominated and nearly doubled in body mass. Larger midges may have increased bioturbation to meet their higher metabolic demands, further oxygenating the sediment and stimulating greater respiration (Hölker et al. 2015, Baranov et al. 2016a, 2016b).

By quantifying the functional dependence of the ecosystem engineering on continuous environmental variation, we were able to illustrate the expected spatiotemporal variation in the effects of midge engineering due to natural variation in light. These projections illustrate the potential for realistic variation in environmental conditions to produce large variation in engineer (e.g. midge) effects through space (across the lake) and time (between periods with and without cyanobacterial blooms). Cyanobacteria have many impacts on aquatic ecosystems that could affect benthic production beyond simply limiting light, such as nitrogen fixation and release of toxins (Elser 1999, Dokulil and Teubner 2000). Nonetheless, our projections illustrate one key respect in which cyanobacterial blooms could result in spatiotemporal variation in the effects of midge engineering on benthic production. While our analysis focused on the role of environmental variation in producing variation in engineer effects, the density of engineer populations is itself variable and likely to have large consequences for predicting engineer effects (Hastings et al. 2007). For example, *Tanytarsus* in Mývatn fluctuate across 5 orders of magnitude among years, with erratic population crashes that are difficult to anticipate (Ives et al. 2008). Given the large magnitude of their effects on benthic ecosystem processes, theses fluctuations in abundance likely have large consequences for the whole system. While Mývatn is

a fairly extreme example, many other ecosystem engineers show large spatiotemporal variation in abundance (Nalepa et al. 1993, Bos et al. 2007). Furthermore, the effects of ecosystem engineering can feed back to the engineer populations themselves (e.g. by affecting production of their food), and the strength and duration of these feedbacks could alter the qualitative nature of their dynamics (Hastings et al. 2007, Cuddington et al. 2009, Largaespada et al. 2012). Quantifying these feedbacks is essential to predicting the consequences of environmental variation through space and time, both in Mývatn and other cases of ecosystem engineering.

References

Armitage, P., P. S. Carnston, and L. C. V. Pinder. 1995. The Chironomidae. The biology and ecology of non-biting midges. Chapman and Hall, London.

Arribas, L. P., L. Donnarumma, M. G. Palomo, and R. A. Scrosati. 2014. Intertidal mussels as ecosystem engineers: their associated invertebrate biodiversity under contrasting wave exposures. Marine Biodiversity 44:203–211.

Attard, K. M., R. N. Glud, D. F. McGinnis, and S. Rysgaard. 2014. Seasonal rates of benthic primary production in a Greenland fjord measured by aquatic eddy correlation. Limnology and Oceanography 59:1555–1569.

Baranov, V., J. Lewandowski, and S. Krause. 2016a. Bioturbation enhances the aerobic respiration of lake sediments in warming lakes. Biology Letters 12:20160448.

Baranov, V., J. Lewandowski, P. Romeijn, G. Singer, and S. Krause. 2016b. Effects of bioirrigation of non-biting midges (Diptera: Chironomidae) on lake sediment respiration. Scientific Reports 6.

Bartrons, M., Á. Einarsson, R. L. G. Nobre, C. M. Herren, K. C. Webert, S. Brucet, S. R. Ólafsdóttir, and A. R. Ives. 2015. Spatial patterns reveal strong abiotic and biotic drivers of zooplankton community composition in Lake Mývatn, Iceland. Ecosphere 6:art105.

Belshe, E. F., E. a. G. Schuur, and B. M. Bolker. 2013. Tundra ecosystems observed to be CO2 sources due to differential amplification of the carbon cycle. Ecology Letters 16:1307–1315.

Benelli, S., M. Bartoli, M. Zilius, I. Vybernaite‐Lubiene, T. Ruginis, J. Petkuviene, and E. A. Fano. 2018. Microphytobenthos and chironomid larvae attenuate nutrient recycling in shallow-water sediments. Freshwater Biology 63:187–201.

Bertness, M. D., and G. H. Leonard. 1997. The Role of Positive Interactions in Communities: Lessons from Intertidal Habitats. Ecology 78:1976–1989.

Bertness, M. D., G. H. Leonard, J. M. Levine, P. R. Schmidt, and A. O. Ingraham. 1999. Testing the relative contribution of positive and negative interactions in rocky intertidal communities. Ecology 80:2711–2726.

Björnsson, H., and T. Jónsson. 2004. Climate and climatic variability at Lake Myvatn. Aquatic Ecology 38:129–144.

Bos, A. R., T. J. Bouma, G. L. J. de Kort, and M. M. van Katwijk. 2007. Ecosystem engineering by annual intertidal seagrass beds: Sediment accretion and modification. Estuarine, Coastal and Shelf Science 74:344–348.

Bott, T. L. 2006. Primary production and community respiration. Pages 663 – 690 *in* F. R. Hauer and G. A. Lamberti, editors. Methods in stream ecology. 2nd edition. Academic Press, Boston, Massachusetts, USA.

Brodersen, K., K. Peter, O. Pedersen, C. Lindegaard, and K. Hamburger. 2004. Chironomids (Diptera) and oxy-regulatory capacity: An experimental approach to paleolimnological interpretation. Limnology and Oceanography 49:1549–1559.

Brown, B. L., and R. L. Lawson. 2010. Habitat heterogeneity and activity of an omnivorous ecosystem engineer control stream community dynamics. Ecology 91:1799–1810.

Burnham, K. T., and D. R. Anderson. 2002. Model selection and inference: a practical information-theoretic approach. Second. Springer, New York, New York.

Cael, B. B., A. J. Heathcote, and D. A. Seekell. 2017. The volume and mean depth of Earth's lakes. Geophysical Research Letters 44:2016GL071378.

Cebrian, J., and J. Lartigue. 2004. Patterns of Herbivory and Decomposition in Aquatic and Terrestrial Ecosystems. Ecological Monographs 74:237–259.

Chapin, F. S., G. M. Woodwell, J. T. Randerson, E. B. Rastetter, G. M. Lovett, D. D. Baldocchi, D. A. Clark, M. E. Harmon, D. S. Schimel, R. Valentini, C. Wirth, J. D. Aber, J. J. Cole, M. L. Goulden, J. W. Harden, M. Heimann, R. W. Howarth, P. A. Matson, A. D. McGuire, J. M. Melillo, H. A. Mooney, J. C. Neff, R. A. Houghton, M. L. Pace, M. G. Ryan, S. W. Running, O. E. Sala, W. H. Schlesinger, and E.-D. Schulze. 2006. Reconciling Carbon-cycle Concepts, Terminology, and Methods. Ecosystems 9:1041–1050.

Crain, C. M., and M. D. Bertness. 2005. Community Impacts of a Tussock Sedge: Is Ecosystem Engineering Important in Benign Habitats? Ecology 86:2695–2704.

Cuddington, K., W. G. Wilson, and A. Hastings. 2009. Ecosystem Engineers: Feedback and Population Dynamics. The American Naturalist 173:488–498.

Daleo, P., and O. Iribarne. 2009. Beyond competition: the stress-gradient hypothesis tested in plant–herbivore interactions. Ecology 90:2368–2374.

Davidson, T. A., J. Audet, J.-C. Svenning, T. L. Lauridsen, M. Søndergaard, F. Landkildehus, S. E. Larsen, and E. Jeppesen. 2015. Eutrophication effects on greenhouse gas fluxes from shallow-lake mesocosms override those of climate warming. Global Change Biology 21:4449–4463.

Demars, B. O. L., G. M. Gíslason, J. S. Ólafsson, J. R. Manson, N. Friberg, J. M. Hood, J. J. D. Thompson, and T. E. Freitag. 2016. Impact of warming on CO2 emissions from streams countered by aquatic photosynthesis. Nature Geoscience 9:758–761.

Dokulil, M. T., and K. Teubner. 2000. Cyanobacterial dominance in lakes. Hydrobiologia 438:1–12.

Einarsson, Á., G. Stefánsdóttir, H. Jóhannesson, J. S. Ólafsson, G. M. Gíslason, I. Wakana, G. Gudbergsson, and A. Gardarsson. 2004. The ecology of Lake Myvatn and the River Laxá: Variation in space and time. Aquatic Ecology 38:317–348.

Elser, J. J. 1999. The pathway to noxious cyanobacteria blooms in lakes: the food web as the final turn. Freshwater Biology 42:537–543.

Elser, J. J., M. E. S. Bracken, E. E. Cleland, D. S. Gruner, W. S. Harpole, H. Hillebrand, J. T. Ngai, E. W. Seabloom, J. B. Shurin, and J. E. Smith. 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. Ecology Letters 10:1135–1142.

Freedman, D. A. 1983. A Note on Screening Regression Equations. The American Statistician 37:152.

Gíslason, S. R., E. S. Eiríksdóttir, and J. S. Ólafsson. 2004. Chemical composition of interstitial water and diffusive fluxes within the diatomaceous sediment in Lake Myvatn, Iceland. Aquatic Ecology 38:163–175.

Gribben, P. E., J. E. Byers, J. T. Wright, and T. M. Glasby. 2013. Positive versus negative effects of an invasive ecosystem engineer on different components of a marine ecosystem. Oikos 122:816–824.

Gutiérrez, J. L., C. G. Jones, D. L. Strayer, and O. O. Iribarne. 2003. Mollusks as ecosystem engineers: the role of shell production in aquatic habitats. Oikos 101:79–90.

Hastings, A., J. E. Byers, J. A. Crooks, K. Cuddington, C. G. Jones, J. G. Lambrinos, T. S. Talley, and W. G. Wilson. 2007. Ecosystem engineering in space and time. Ecology Letters 10:153–164.

Henriksen, K., M. B. Rasmussen, and A. Jensen. 1983. Effect of Bioturbation on Microbial Nitrogen Transformations in the Sediment and Fluxes of Ammonium and Nitrate to the Overlaying Water. Ecological Bulletins:193–205.

Herren, C. M., K. C. Webert, M. D. Drake, M. Jake Vander Zanden, Á. Einarsson, A. R. Ives, and C. Gratton. 2017. Positive feedback between chironomids and algae creates net mutualism between benthic primary consumers and producers. Ecology:447–455.

Holgerson, M. A., and P. A. Raymond. 2016. Large contribution to inland water CO2 and CH4 emissions from very small ponds. Nature Geoscience 9:222–226.

Hölker, F., M. J. Vanni, J. J. Kuiper, C. Meile, H.-P. Grossart, P. Stief, R. Adrian, A. Lorke, O. Dellwig, A. Brand, M. Hupfer, W. M. Mooij, G. Nützmann, and J. Lewandowski. 2015. Tube-dwelling invertebrates: tiny ecosystem engineers have large effects in lake ecosystems. Ecological Monographs 85:333–351.

Ives, A. R., Á. Einarsson, V. A. A. Jansen, and A. Gardarsson. 2008. High-amplitude fluctuations and alternative dynamical states of midges in Lake Myvatn. Nature 452:84–87.

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

Jónasson, P. M. 1979. The Lake Mývatn Ecosystem, Iceland. Oikos 32:289–305.

Jones, C. G., J. H. Lawton, and M. Shachak. 1994. Organisms as Ecosystem Engineers. Oikos 69:373–386.

Jones, C. G., J. H. Lawton, and M. Shachak. 1997. Positive and Negative Effects of Organisms as Physical Ecosystem Engineers. Ecology 78:1946–1957.

Karlsson, J., P. Byström, J. Ask, P. Ask, L. Persson, and M. Jansson. 2009. Light limitation of nutrient-poor lake ecosystems. Nature 460:506–509.

Kilham, P. 1971. A HYPOTHESIS CONCERNING SILICA AND THE FRESHWATER PLANKTONIC DIATOMS1: SILICA AND FRESHWATER PLANKTONIC DIATOMS. Limnology and Oceanography 16:10–18.

Largaespada, C., F. Guichard, and P. Archambault. 2012. Meta-ecosystem engineering: Nutrient fluxes reveal intraspecific and interspecific feedbacks in fragmented mussel beds. Ecology 93:324–333.

Lathlean, J. A., and C. D. McQuaid. 2017. Biogeographic variability in the value of mussel beds as ecosystem engineers on South African rocky shores. Ecosystems 20:568–582.

Lindegaard, C., and P. M. Jónasson. 1979. Abundance, Population Dynamics and Production of Zoobenthos in Lake Mývatn, Iceland. Oikos 32:202–227.

Lovett, G. M., J. J. Cole, and M. L. Pace. 2006. Is Net Ecosystem Production Equal to Ecosystem Carbon Accumulation? Ecosystems 9:152–155.

Luke, S. G. 2017. Evaluating significance in linear mixed-effects models in R. Behavior Research Methods 49:1494–1502.

McAfee, D., V. J. Cole, and M. J. Bishop. 2016. Latitudinal gradients in ecosystem engineering by oysters vary across habitats. Ecology 97:929–939.

McCormick, P. V., and R. J. Stevenson. 1991. Mechanisms of Benthic Algal Succession in Lotic Environments. Ecology 72:1835–1848.

Nalepa, T. F., J. F. Cavaletto, M. Ford, W. M. Gordon, and M. Wimmer. 1993. Seasonal and Annual Variation in Weight and Biochemical Content of the Zebra Mussel, Dreissena polymorpha, in Lake St. Clair. Journal of Great Lakes Research 19:541–552.

Nogaro, G., F. Mermillod-Blondin, B. Montuelle, J.-C. Boisson, and J. Gibert. 2008. Chironomid larvae stimulate biogeochemical and microbial processes in a riverbed covered with fine sediment. Aquatic Sciences 70:156–168.

Norkko, A., J. E. Hewitt, S. F. Thrush, and G. A. Funnell. 2006. Conditional Outcomes of Facilitation by a Habitat-Modifying Subtidal Bivalve. Ecology 87:226–234.

Ólafsson, J. 1979. The Chemistry of Lake Mývatn and River Laxá. Oikos 32:82–112.

Olafsson, J. S., and D. M. Paterson. 2004. Alteration of biogenic structure and physical properties by tube-building chironomid larvae in cohesive sediments. Aquatic Ecology 38:219–229.

Paasche, E. 1973. Silicon and the ecology of marine plankton diatoms. I. Thalassiosira pseudonana (Cyclotella nana) grown in a chemostat with silicate as limiting nutrient. Marine Biology 19:117–126.

Paine, R. T. 1966. Food Web Complexity and Species Diversity. The American Naturalist 100:65–75.

Pringle, C. M. 1985. Effects of Chironomid (insecta: Diptera) Tube-Building Activities on Stream Diatom Communities1,2. Journal of Phycology 21:185–194.

R Core Team. 2016. R version 3.3.1. The R Foundation for Statistical Computing Platform, Vienna, Austria.

Randerson, J. T., F. S. Chapin, J. W. Harden, J. C. Neff, and M. E. Harmon. 2002. Net Ecosystem Production: A Comprehensive Measure of Net Carbon Accumulation by Ecosystems. Ecological Applications 12:937–947.

Ruel, J. J., M. P. Ayres, J. J. Ruel, and M. P. Ayres. 1999. Jensen's inequality predicts effects of environmental variation. Trends in Ecology & Evolution 14:361–366.

Sommer, U. 1991. A Comparison of the Droop and the Monod Models of Nutrient Limited Growth Applied to Natural Populations of Phytoplankton. Functional Ecology 5:535–544.

Steinman, A., M. Abdimalik, M. E. Ogdahl, and M. Oudsema. 2016. Understanding planktonic vs. benthic algal response to manipulation of nutrients and light in a eutrophic lake. Lake and Reservoir Management 0:1–9.

Svensson, J., and L. Leonardson. 1996. Effects of bioturbation by tube-dwelling chironomid larvae on oxygen uptake and denitrification in eutrophic lake sediments. Freshwater Biology 35:289–300.

Svensson, J. M. 1997. Influence of Chironomus plumosus larvae on ammonium flux and denitrification (measured by the acetylene blockage- and the isotope pairing-technique) in eutrophic lake sediment. Hydrobiologia 346:157–168.

Tanaka, T., K. Kawasaki, S. Daimon, W. Kitagawa, K. Yamamoto, H. Tamaki, M. Tanaka, C. H. Nakatsu, and Y. Kamagata. 2014. A Hidden Pitfall in the Preparation of Agar Media Undermines Microorganism Cultivability. Appl. Environ. Microbiol. 80:7659–7666.

Tank, J. L., M. J. Bernot, and E. J. Rosi-Marshall. 2006. Nitrogen limitation and uptake. Pages 213–238 *in* F. R. Hauer and G. A. Lamberti, editors. Methods in stream ecology. 2nd edition. Academic Press, Boston, Massachusetts, USA.

Taranu, Z. E., I. Gregory-Eaves, P. R. Leavitt, L. Bunting, T. Buchaca, J. Catalan, I. Domaizon, P. Guilizzoni, A. Lami, S. McGowan, H. Moorhouse, G. Morabito, F. R. Pick, M. A. Stevenson, P. L. Thompson, and R. D. Vinebrooke. 2015. Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. Ecology Letters 18:375–384.

Vadeboncoeur, Y., E. Jeppesen, M. J. Vander Zanden, H.-H. Schierup, K. Christoffersen, and D. M. Lodge. 2003. From Greenland to green lakes: Cultural eutrophication and the loss of benthic pathways in lakes. Limnology and Oceanography 48:1408–1418.

Vadeboncoeur, Y., M. J. Vander Zanden, and D. M. Lodge. 2002. Putting the Lake Back Together: Reintegrating Benthic Pathways into Lake Food Web Models Lake ecologists tend to focus their research on pelagic energy pathways, but, from algae to fish, benthic organisms form an integral part of lake food webs. BioScience 52:44–54.

Welschmeyer, N. A. 1994. Fluorometric analysis of chlorophyll a in the presence of chlorophyll b and pheopigments. Limnology and Oceanography 39:1985–1992.

Wright, J. P., and C. G. Jones. 2006. The concept of organisms as ecosystem engineers ten years on: progress, limitations, and challenges. BioScience 56:203–209.

Wright, J. P., C. G. Jones, B. Boeken, and M. Shachak. 2006. Predictability of ecosystem engineering effects on species richness across environmental variability and spatial scales. Journal of Ecology 94:815–824.

Yeager, P. E., C. L. Foreman, and R. L. Sinsabaugh. 2001. Microbial community structure and function in response to larval chironomid feeding pressure in a microcosm experiment. Hydrobiologia 448:71–81.

Zhang, L., X. Gu, C. Fan, J. Shang, Q. Shen, Z. Wang, and J. Shen. 2010. Impact of different benthic animals on phosphorus dynamics across the sediment-water interface. Journal of Environmental Sciences 22:1674–1682.

Acknowledgements

We would like to thank A. Arellano, B. Blundell, J. Botsch, C. Daws, A. Fassler, C. Miller, C. Owens, J. Sanchez-Ruiz, and P. Uphues for assistance with data collection and the Jackson Lab at UW-Madison for performing nutrient analyses. Furthermore, we would like to thank members of the Ives, Vander Zanden, and Gratton labs for providing comments on the analysis and manuscript, and particularly L. Nell for assistance with preparing figures. Finally, we would like to thank C. Gratton, J. Ólafsson, K. Webert, and J. Welter for feedback on experimental methods. This work was supported by NSF LTREB DEB-1556208 to ARI and NSF Graduate Research Fellowship (DGE-1256259) supporting JSP.

Table 1: AIC values for models of NEP (equation 1) fit to data from experiment 2. The models vary by light-dependent (α) and light-independent (δ) effects of midge treatment and time. Models with interactions (×) also include corresponding main effects. All models include the saturation parameter β and intercepts for α and δ. For clarity only models with ΔAIC < 10 are shown (see Table A3 for details on all models).

Light-dependent (α) fixed effects	Light-independent (δ) fixed effects	Number of fixed effects parameters	AIC	ΔAIC
midge + time	midge × time	8	-474.67	0.00
midge	midge × time	7	-473.17	1.50
midge × time	midge × time	9	-472.69	1.99
midge × time	midge + time	8	-470.25	4.42
midge × time	time	7	-465.31	9.36

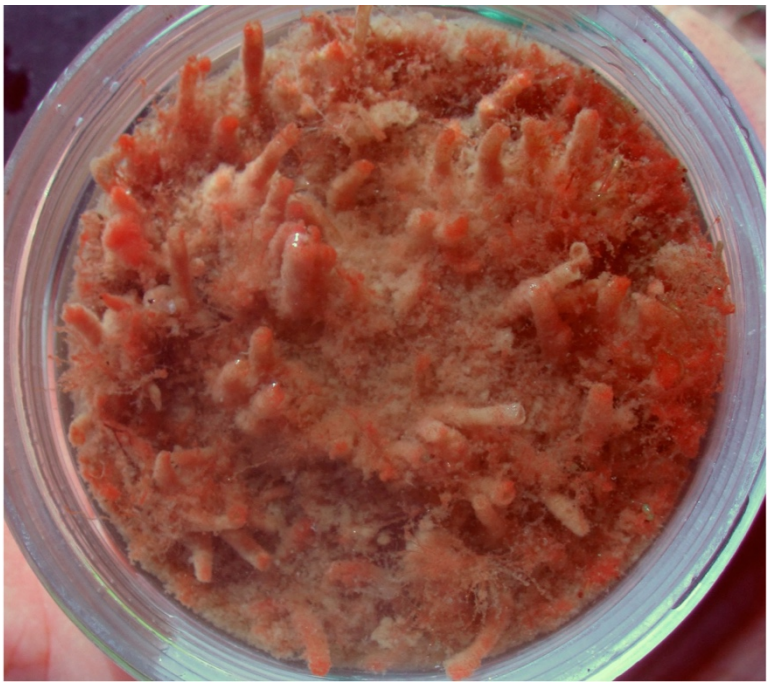

Figure 1. Midges build silk tubes that provide a substrate for algal growth. The sediment core in the photograph was taken from the bottom of Mývatn using a Kajak corer in a year of moderate midge abundance. The cylindrical structures projecting from the sediment surface are the silk tubes constructed by the midge larvae. Diatoms grow on the tubes and in the sediment. The image is in false color, replacing infrared with red to highlight the distribution of chlorophyll, which is more concentrated on the midge tubes than in the surrounding sediment. Photo credit: T. Ives.

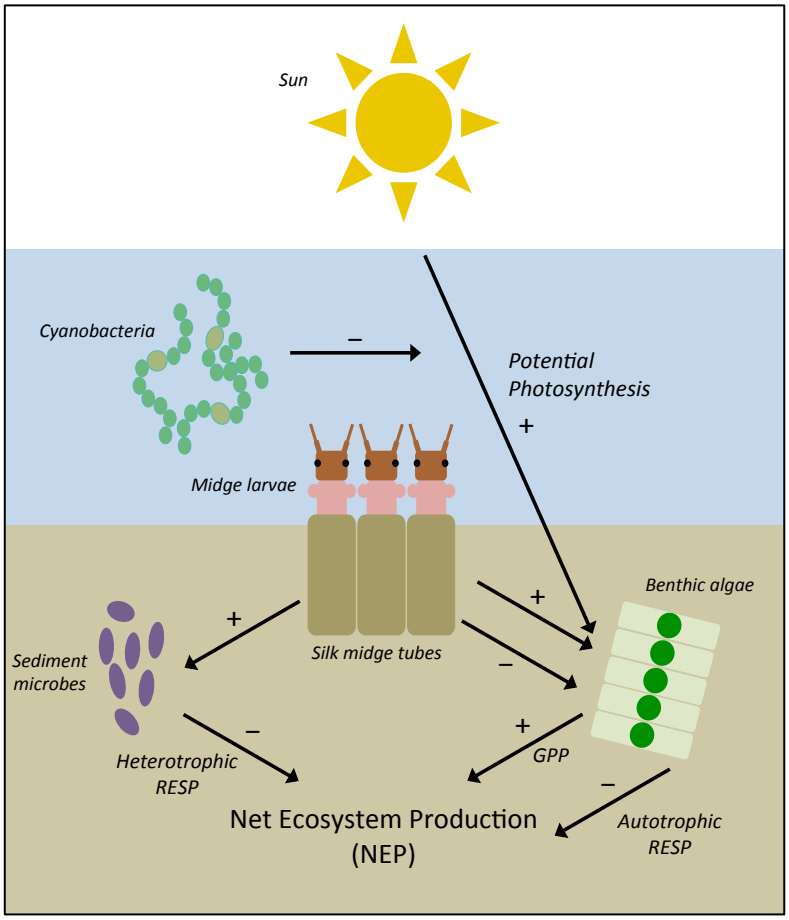

Figure 2. Midges can alter benthic ecosystem function. Larval midges build silk tubes that provide a substrate for algal growth and increase gross primary production (GPP) in the sediment. However, midges may inhibit GPP through consumption of algae. Furthermore, midges can stimulate microbial respiration (RESP) by oxygenating the sediment. GPP and RESP have opposite effects on net ecosystem production (NEP), so the effect of midges on NEP depends on the balance between their effects on GPP and RESP. We hypothesized that light mediates this balance, because the positive effects of midges on GPP would decline as photosynthesis became more limited by light. Episodic cyanobacterial blooms have a negative effect on benthic light levels, which could result in spatiotemporal variation in the net effects of midges on benthic production.

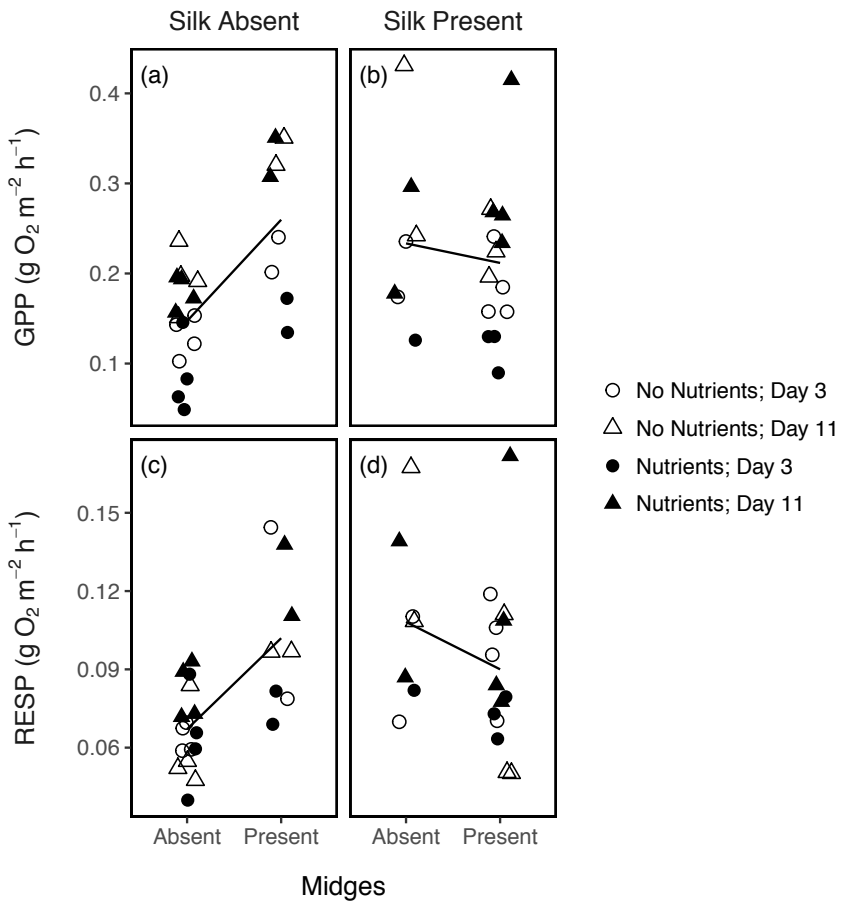

Figure 3. Midges and silk increase GPP and RESP. Each panel shows GPP (a,b) or RESP (c,d) for either absence (a,c) or presence (b,d) of silk. The data are grouped by midge treatment, nutrient treatments are represented by filled ("Nutrients") and open ("No Nutrients") points, and days are represented by circles ("Day 3") and triangles ("Day 11"). Lines show fits from LMMs. Using standard photosynthetic and respiratory quotients of 1.2 and 0.85, a 0.1 g O_2 m^{-2} h^{-1} O_2 difference equals 0.031 g C m^{-2} h^{-1} and 0.032 g C m^{-2} h^{-1}, respectively.

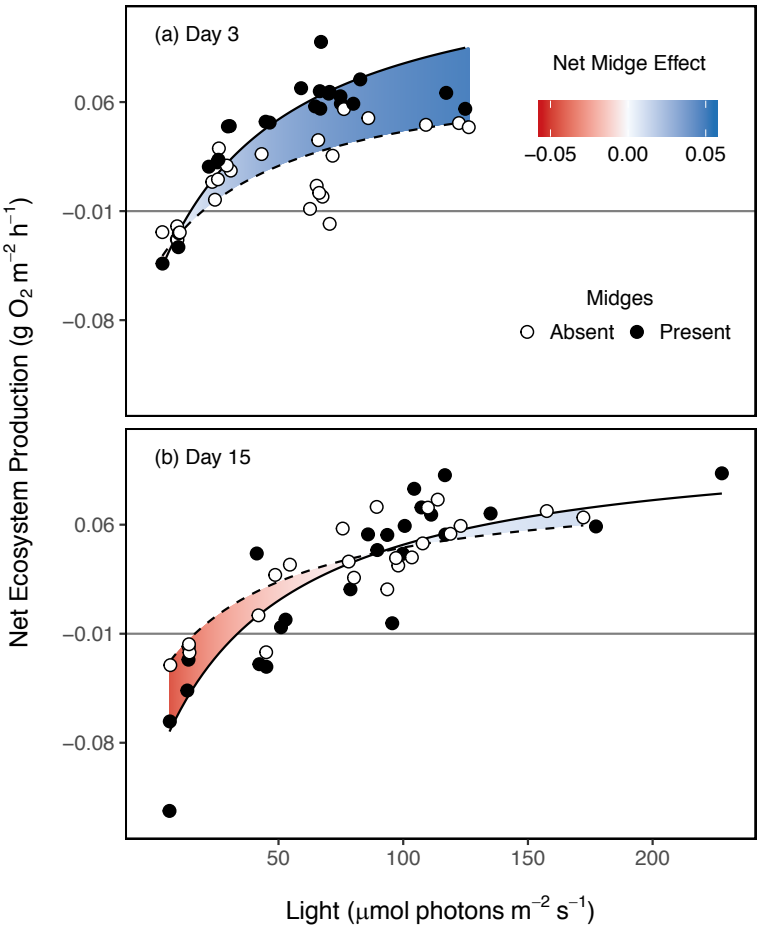

Figure 4. Light mediates midge effects on benthic net ecosystem production. NEP from experiment 2 is plotted against light for (a) day 3 and (b) day 15. The lines show model estimates evaluated with (solid) and without (dashed) midges. The colored shading indicates the net midge effect, calculated as the difference in the model estimates when evaluated with and without midges. The shift between days 3 and 15 is associated with a shift from second to third instar midges. Using a photosynthetic quotient of 1.2, a net midge effect of 0.1 g O_2 m^{-2} h^{-1} equals 0.031 g C m^{-2} h^{-1}.

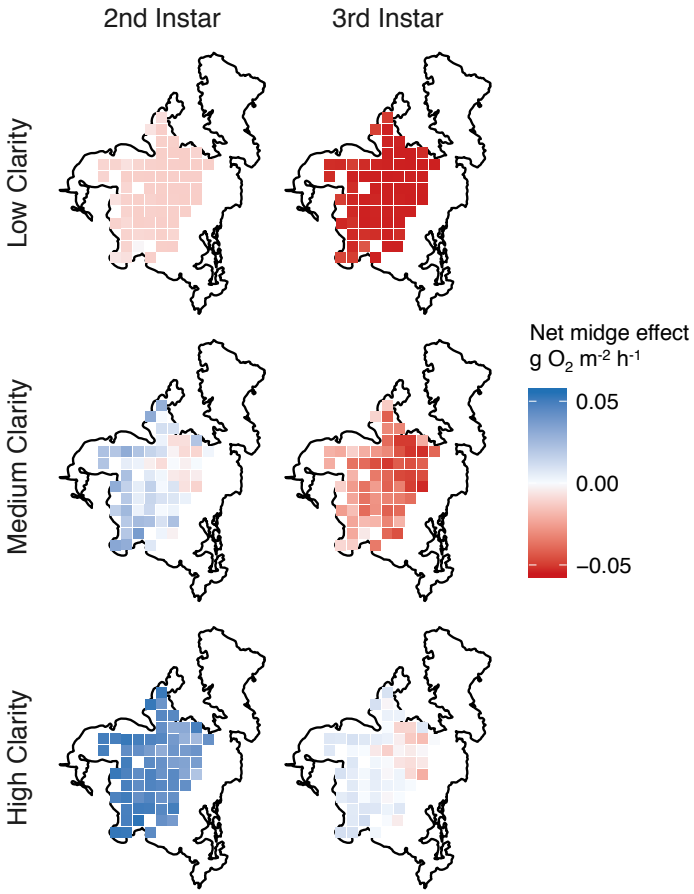

Figure 5. Projected midge effects on NEP vary through space and time. The maps show projected effects of midges on NEP across Mývatn. Midge effects were calculated as the difference in the model estimates when evaluated with and without midges. Different panels correspond to benthic light levels for the different water clarity scenarios (low, medium, high) and Tanytarsini stages (second and third instar-dominated). Using a photosynthetic quotient of 1.2, a net midge effect of 0.1 g O_2 m^{-2} h^{-1} equals 0.031 g C m^{-2} h^{-1}.

Appendices

Appendix A: Supplementary methods and results

Appendix A: Supplementary methods and results for Chapter 1

Supplementary methods

Experiment 1: Silk addition

Figure A1 shows a photograph of a silk coil in a mesocosm, typical of those used to mimic midge tubes.

Experiments 1 & 2: Comparison of Initial Chlorophyll-a between Sieving Treatments

For Experiment 1, we quantified chlorophyll-a concentrations (main text) from five samples of the homogenized sieved sediment (from the top 5cm) and samples of the top 5cm of five sediment cores collected from the same location and analyzed chlorophyll-a concentrations. Chlorophyll-a concentrations did not differ significantly between the sieved (mean ± standard error = 28.23 ± 3.51 mg L^{-1}) and intact sediment (36.01 ± 5.05 mg L^{-1}) (t-test on log-transformed values: t=-1.242, p=0.249).

For Experiment 2, we quantified chlorophyll-a concentrations in the same manner as for Experiment 1, except that we collected samples from both the top 5cm and bottom 10cm of sieved and intact sediment, with 3 samples for each layer × treatment combination. Chlorophyll-a concentration was substantially higher in the sediment from the top 5cm (13.07 ± 0.68 mg L^{-1}) than from the bottom 10cm (8.51 ± 0.86 mg L^{-1}). However, sieved vs. intact sediment did not differ within the 'top' (12.25 ± 0.55 vs. 13.89 ± 1.16 mg L^{-1}) (t-test on log-transformed values: t=1.321, p=0.257) and 'bottom' (7.30 ± 0.43 vs. 9.73 ± 1.42 mg L^{-1}) (t=1.559, p=0.194) sediment samples.

We note that for both Experiment 1 and Experiment 2, the low sample size provided limited power for detecting a difference in chlorophyll-a between sieved and intact sediment, and this approach does not provide information on variation in chlorophyll-a with sediment depth. Nonetheless, the average values are broadly similar, which is the central point. We also note that the chlorophyll-a concentrations were substantially higher for Experiment 1 than Experiment 2, likely due to the different sediment sources. This is consistent with the substantially higher primary production values in Experiment 1.

Experiment 2: Nutrient Release Rates from Agar

To determine the release rates of N and P from the agar rods, we constructed two enriched rods and two control rods containing agar (on a separate occasion from the main experiment) and placed them in separate 1L containers containing tap water. The enriched rods contained NH_4Cl and KH_2PO_4 to produce N and P concentrations of 0.5M and 0.13M respectively. We sampled 45mL of water from each container on 4 occasions over 6 days and analyzed the samples for N (NH_4) and P (SRP) concentration. Nutrient concentrations were analyzed using colorimetric methods on a segmented flow auto-analyzer (OI-A FS 3100, O-I-Analytical, TX, USA), following US-EPA protocols for ammonium (Unites States Environmental Protection Agency 1993a p. 36) and SRP(Unites States Environmental Protection Agency 1993b p. 36). Samples were also analyzed for NO_x following (Unites States Environmental Protection Agency 1993c), however their concentrations were negligible.

We estimated average daily release rates as the slope of total nutrient mass (accounting for the change in volume due to sampling) regressed against the days since the start of the release rate trial. N was released at an average rate of 22.8 mg day^{-1} in the enriched treatment, with

137.1 mg of the possible 185.5 mg being released over 6 days. P was released at a rate of 1.23 mg day^{-1} in the enriched treatment, with 7.42 mg of the possible 107 mg being released over 6 days. Neither N nor P increased through time in the control treatments, indicating that there was no release of nutrients in the absence of enrichment. However, the control treatment did have detectable levels of P (7.6mg), indicating some contamination in the release rate set-up. This contamination resulted from the fact that the agar rods for the release rate measurements were stored on a plastic tray that previously contained nutrient-enriched agar rods from the preparation for the main experiment. However, the surfaces used in the preparation of the rods for the main experiment were cleaned prior to use and so there is no reason to believe that there was contamination in the main experiment.

Supplementary results

Experiment 1: NEP

Figure A2 shows the results from Experiment 1 for NEP (GPP and RESP are shown in main text Figure 3). By definition GPP and RESP had opposite effects on NEP, but because the overall magnitude of GPP was greater than RESP, GPP had a disproportionate influence. Therefore, the response of NEP to the experimental treatments broadly paralleled GPP, with positive effects of midges ($F_{1,24.78}$=4.33, p=0.048) and silk ($F_{1,25.64}$=5.63, p=0.025), and negative midge × silk interaction ($F_{1,15.52}$=16.31, p=0.001), and negative effects of nutrients ($F_{1,27.04}$=11.90, p= 0.002) (Fig A2). NEP was higher on day 11 than day 3 ($F_{1,17.09}$=15.12, p=0.001), and there was a positive midge × day interaction ($F_{1,18.89}$=4.61, p= 0.045).

Experiment 2: Midge Counts

To evaluate the efficacy of our midge (sieving) treatments, we sieved the sediment from a subset of the mesocosms (all cores from the full light and full dark shading treatments) and counted the larval midges. The 'midges present' mesocosms had substantially more midges than the 'midges absent', indicating that the sieving treatments were effective (Table A1). Most of these were *Tanytarsus gracilentus*, the only member of the genus that occurs in Mývatn. There was minimal colonization of the mesocosms by *Micropsectra* (closely related to *Tanytarsus*), which are primarily restricted to the portion of the lake where the mesocosms were incubated. Two species of *Chironomus* (*C. islandicus* and *C. plumosus*) occur in Mývatn, however we did not distinguish between them. Members of Tanypodinae and Orthocladinae are quite diverse in Mývatn and we only identified them to subfamily.

Experiment 2: Midge Growth

We used the body lengths and head capsule widths of Tanytarsini midges collected from the lake near the beginning of the experiment and from a subset of the experimental cores at the end to estimate how much the midges grew. Mean body length of midges increased from 2.42mm to 2.70mm and head capsule widths increased from 0.12mm to 0.16mm. We estimated the change in midge biovolume by approximating individual midges as perfect cylinders, using head capsule with as the diameter and body length as the height of each cylinder (Smit et al. 1993). Using this approximation, mean biovolume increased by 91%, from 0.035mm^3 to 0.066mm^3. Body mass is generally proportional to biovolume; Smit *et al.* (1993) estimated wet specific gravity of 1.05 mg mm^{-3} for *Chironomus muratensis* (closely related to tribe Tanytarsini). Using this value, the average mass of Tanytarsini increased from 0.036mg to 0.069mg.

Experiment 2: Changes in NEP through Time by Midge Treatment

To evaluate whether the change in *in situ* NEP between days 3 and 15 was different across midge treatments, we fit a reduced version of equation (1) separately for the 'midges present' and 'midges absent' treatments (omitting the fixed effects of midges from α and δ). We evaluated the significance of the time parameters by examining the 95% confidence intervals based on REML estimates. For 'midges absent', the confidence intervals for both light-dependent and light-independent effects of time included 0, indicating that NEP did not change through time (Table A2). In contrast, while the confidence interval for light-dependent time effects for 'midges present' included 0, the light-independent effect excluded 0. The difference between temporal changes in the two midges treatments is especially striking when examined graphically (Fig A3)—changes through time are much more apparent when midges are present. These results indicate that the shift in apparent midge effects through time was driven by changes in the 'midges present' treatment.

Lake-Scale Effects of Midges on NEP

To quantify spatial variation in light attenuation as a result of a thick cyanobacterial bloom in summer of 2015, we measured PAR (μmol-photons m^{-2} s^{-1}) with a light sensor (Li-192 Quantum Underwater Sensor, Li-COR, NE, USA) at the water surface and at 0.5m, 1m, and 2m below the water surface for each of 65 sites around the lake. The sites were positioned on a 500×500m grid in the south basin of Mývatn. For each site, we calculated the light extinction coefficient by as the rate of exponential decay of PAR as a function of depth (using the lm() function in R3.3.1(R Core Team 2016)). The light readings were taken on 5 separate days between 30 July 2015 and 11 August 2015, and we acknowledge that substantial changes in the

spatial structure of the cyanobacterial bloom could have occurred over this time period. However, the spatial pattern of light attenuations (Fig A4) matched mapping of cyanobacterial blooms based on pigments over shorter windows of time in 2011(Bartrons et al. 2015) and 2016 (unpublished data). Therefore, the light attenuation data from 2015 likely provides a reasonable characterization of spatial patterns of light variation as a result of cyanobacterial blooms in general.

For each site we calculated the light level on the lake bottom assuming moderate surface light (320 μmol photons m^{-2} s^{-1}) and one of three water clarity scenarios: low clarity (100% measured light attenuation,) medium clarity (40% measured attenuation), or high clarity (15% measured attenuation). Benthic light levels varied from virtually 0 μmol photons m^{-2} s^{-1} to 180 μmol photons m^{-2} s^{-1}, which roughly matches the range of light levels used for the experiment and are typically of benthic light levels throughout much of the lake during the summer.

For each water clarity scenario we quantified the site-specific midge effect as the difference between equation (1) evaluated with and without midges. We made projections using both time points in equation (1), corresponding to midge communities dominated by either 2nd or 3rd instar *Tanytarsus*. This resulted in six scenarios of water clarity and midge stage (main text Fig 5). For comparison, we repeated this analysis using average light attenuations across the lake for each water clarity and midge scenario, such that the only spatial within scenarios was due to depth (Fig A5).

Projected midge effects on benthic production ranged from strongly negative (-0.054 g O_2 m^{-2} h^{-1}) to strongly positive (0.053 g O_2 m^{-2} h^{-1}) across the six scenarios. To facilitate comparison to these results, we converted the estimated benthic diatom production in Mývatn (200 g C m^{-2}) to a comparable scale. During much of the year Mývatn is frozen and likely has

negligible benthic production. We assumed a growing season of 4 months, which is probably somewhat shorter than the actual period and therefore provides a conservative comparison of the magnitude of midge effects relative to the overall benthic diatom production (by assuming more *in situ* production occurs in a shorter period of time). Using a photosynthetic quotient of 1.2 (Bott 2006), *in situ* diatom production was calculated as follows:

$200\ g\ C\ m^{-2} \times (12\ g\ C\ mol^{-1})^{-1} \times 1.2\ mol\ O_2\ mol^{-1}\ C \times 32\ g\ O_2\ mol^{-1} \times (4\ months \times 720\ h\ month^{-1})^{-1} = 0.222\ g\ O_2\ m^{-2}\ h^{-1}$

Table A1: Midge counts (mean ± standard error) by taxon from Experiment 2.

Midge Treatment	Tanytarsini	Chironomini	Tanypodinae	Orthocladinae
'Absent' (sieved)	17.00 ± 0.68	1.667 ± 0.494	4.500 ± 1.310	3.167 ± 1.376
'Present' (not sieved)	118.2 ± 33.71	7.500 ± 1.432	6.500 ± 2.825	8.200 ± 2.958

Table A2: Model coefficients (95% C.I.) for time effects estimated via REML separately for the 'midges absent' and 'midges present' treatments.

	Midges Absent	Midges Present
Light Dependent		
time ($\alpha 3$)	0.00025 (-0.00014, 0.00064)	0.0014 (-0.001, 0.0038)
Light Independent		
time ($\delta 3$)	-0.0045 (-0.017, 0.0077)	-0.060 (-0.099, -0.021)

Table A3: AIC values for models of NEP (eq. 1) fit to data from Experiment 2. The models vary by light-dependent (α) and light-independent (δ) effects of midge treatment and time. Models with interactions (×) also include corresponding main effects. All models include the saturation parameter β and intercepts for α and δ; '1' indicates intercept only.

Light-dependent (α) fixed effects	Light-independent (δ) fixed effects	Number of fixed effects parameters	AIC	ΔAIC
midge + time	midge × time	8	-474.67	0.00
midge	midge × time	7	-473.17	1.50
midge × time	midge × time	9	-472.69	1.99
midge × time	midge + time	8	-470.25	4.42
midge × time	time	7	-465.31	9.36
time	midge × time	7	-461.62	13.05
midge × time	midge	7	-460.84	13.83
1	midge × time	6	-460.16	14.51
midge × time	1	6	-455.69	18.99
midge + time	midge + time	7	-443.62	31.06
midge + time	time	6	-443.02	31.65
midge	midge + time	6	-442.98	31.69
midge	time	5	-442.48	32.19
time	midge + time	6	-439.87	34.80
1	midge + time	5	-439.27	35.40
midge + time	midge	6	-438.51	36.16
time	time	5	-437.79	36.88

midge + time	1	5	-437.45	37.22
1	time	4	-437.29	37.38
time	midge	5	-434.09	40.58
midge	midge	5	-432.58	42.09
time	1	4	-432.12	42.56
midge	1	4	-430.90	43.77
1	midge	4	-427.43	47.24
1	1	3	-425.35	49.33

Figure A1. Top-down view of experimental mesocosm with a silk coil partially submerged in sediment.

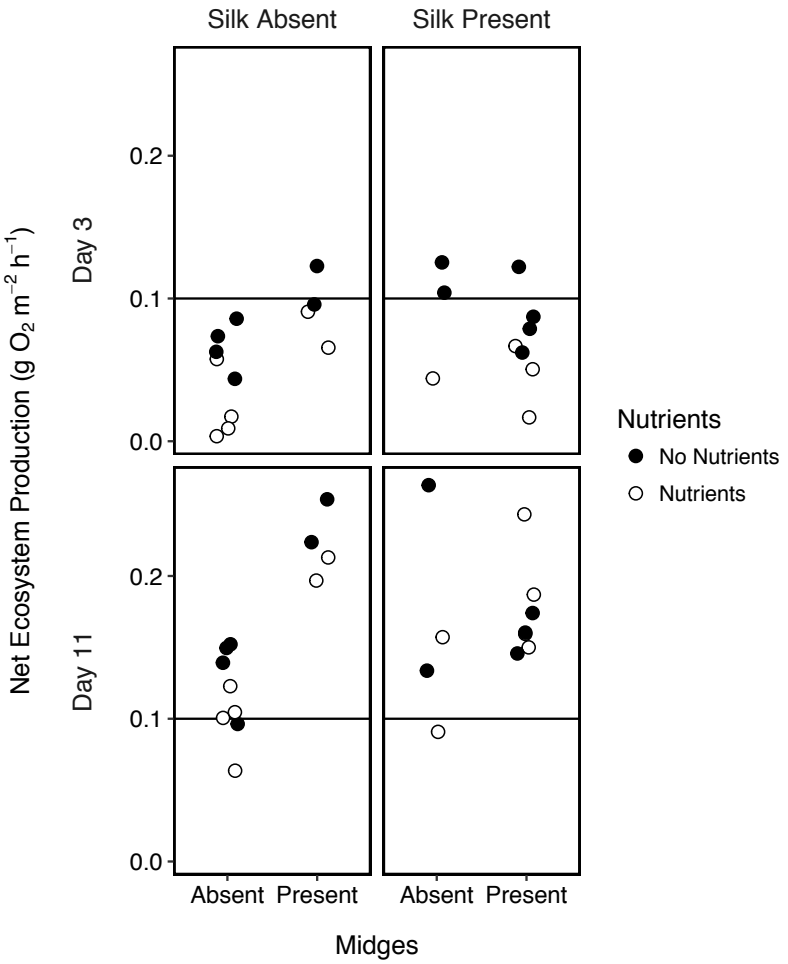

Figure A2. Midge, silk, and nutrient effects on NEP. These effects mirrored those for GPP and RESP, with midges and silk having positive effects but with negative interactions. The effect of nutrients was negative, although this was only obvious in the day 3 data. Using a standard photosynthetic of 1.2, a 0.1 g O_2 m^{-2} h^{-1} O_2 difference equals 0.031 g C m^{-2} h^{-1}.

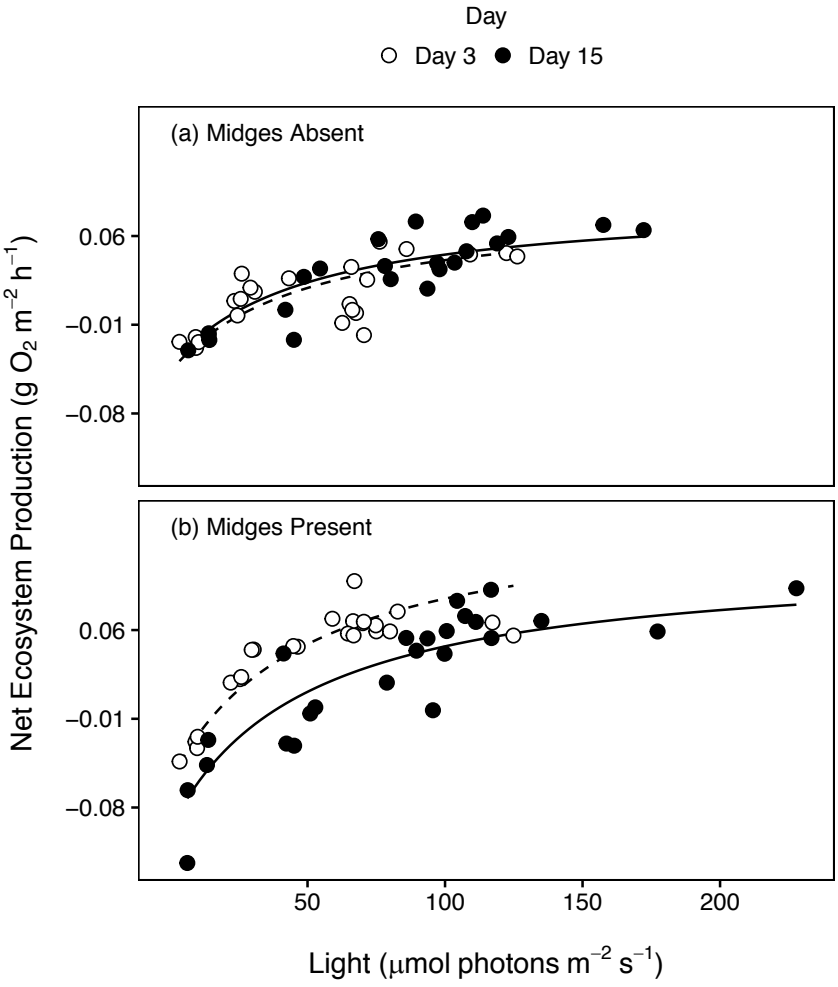

Figure A3. NEP changed through time, but only when midges were present. Each panel shows NEP from Experiment 2 for either the (a) midges absent or (b) midges present treatments, with open circles (dashed lines) indicating day 3 and filled circles (solid lines) indicating day 15. These data are the same as presented in Main Text Fig 4, but organized to highlight the changes in NEP through time within each midge treatment.

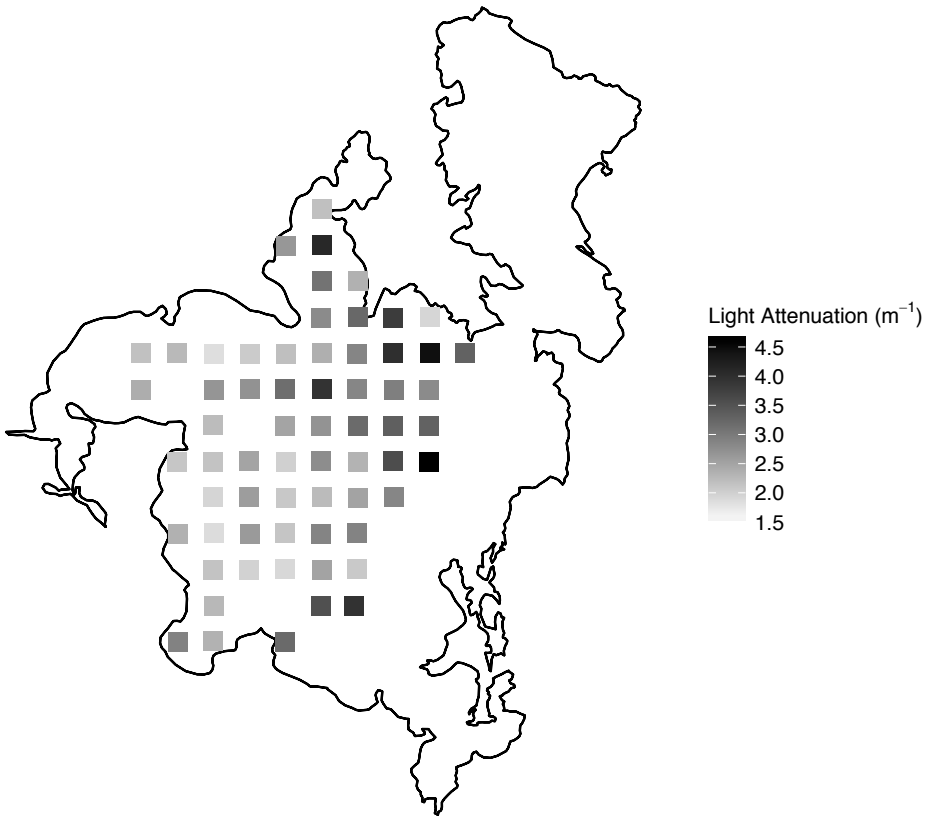

Figure A4. Spatial variation in light attenuation across Mývatn. Attenuation was calculated from readings of PAR at multiple depths at each site between 30 July and 11 August 2015.

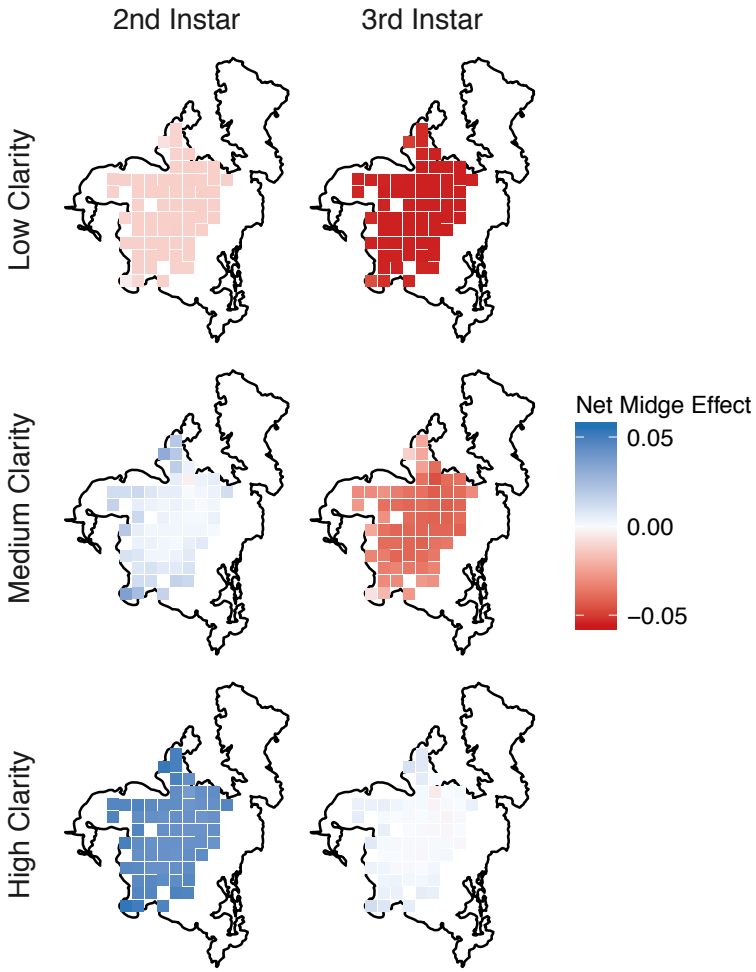

Figure A5. Projected midge effects on NEP through space and time, based on the model (Main Text Fig 4) and different water clarify scenarios (low, medium, high). Midge effects were calculated as the difference in the model predictions when evaluated with and without midges. In contrast to Main Text Fig 5, this figure uses a single water clarity value for each scenario (rather than different values for each site), such that the only variation in benthic light levels is due to variation in depth.

Chapter 2

Time-varying responses of lake metabolism to light and temperature

Joseph S. Phillips[1]

[1]Department of Integrative Biology, University of Wisconsin-Madison, Madison, WI, 53706

Status: In revision for *Limnology & Oceanography*

Abstract

Light is a primary driver of lake ecosystem metabolism, and the dependence of primary production on light is often quantified as a photosynthesis-irradiance or "P-I" curve. The parameters of the P-I curve (e.g. the maximum primary production when light is in excess) can change through time due to a variety of biological factors (e.g. changes in biomass or community composition), which themselves are subject to external drivers (e.g. herbivory or nutrient availability). However, the relative contribution of variation in the P-I curve to overall ecosystem metabolism is largely unknown. I developed a statistical model of ecosystem metabolism with time-varying parameters governing the P-I curve, while also accounting for the influence of temperature. I parameterized the model with dissolved oxygen time-series spanning six summers from Lake Mývatn, a shallow eutrophic lake in northern Iceland with large temporal variability in ecosystem metabolism. All of the estimated parameters of the P-I curve varied substantially through time. The sensitivity of primary production to light under light-limiting conditions was particularly variable (>15-fold). However, the 3.5-fold variation in the maximum potential primary production made the largest contribution to variation in ecosystem metabolism, accounting for around 90% of the variance in net ecosystem production. Much of the variation in maximum primary production was attributable to cyanobacterial blooms, which occur in some but not all years in Mývatn. Overall, these results illustrate how changes in the P-I curve contribute substantially to temporal variation in lake ecosystem metabolism.

Keywords: benthic production; cyanobacteria; ecosystem metabolism; pelagic production; photosynthesis-irradiance curve; state-space model

Introduction

Ecosystem metabolism, the biological assimilation and release of carbon, is central to the function of aquatic ecosystems (Hanson et al. 2003, Staehr et al. 2010b, 2010a). Photosynthesis stores light energy in the form of organic compounds, and gross carbon assimilation (known as gross primary production or GPP) provides metabolic energy and structural material to primary producers, higher trophic levels, and detrital food webs (Del Giorgio et al. 1999, Carpenter et al. 2001, Chapin et al. 2006). In ecosystems without substantial external carbon inputs, in situ GPP constrains total biological productivity, which in turn influences ecological dynamics across organizational scales (Del Giorgio et al. 1999, Fussmann 2000, Staehr et al. 2010b). The counterpart to GPP is ecosystem respiration (ER), which includes the oxidation of organic carbon and subsequent energy utilization by all organisms in the ecosystem (Chapin et al. 2006, Solomon et al. 2013). The balance between GPP and ER determines the net assimilation or release of carbon due to biological processes (net ecosystem production or NEP), which plays a key role in ecosystem carbon budgets (Randerson et al. 2002, Chapin et al. 2006, Demars et al. 2016). Freshwater ecosystems can be either sources or sinks of CO_2 and make substantial contributions to regional carbon cycling, especially in landscapes with high densities of freshwater bodies (Cole et al. 2007, Raymond et al. 2013, Holgerson and Raymond 2016, Demars et al. 2016). Therefore, characterizing the physical and biological controls of ecosystem metabolism is important for predicting both the local function and regional impacts of freshwater systems.

Light is a fundamental requirement of photosynthesis and therefore is a key driver of ecosystem metabolism. Many studies have characterized photosynthesis-irradiance (P-I) curves (Jassby and Platt 1976, Behrenfeld and Falkowski 1997) at organizational scales spanning laboratory monocultures (Myers and Graham 1971) to whole ecosystems (Holtgrieve et al. 2010,

Schindler et al. 2017). For ecosystem metabolism, the P-I curve can be defined as the response of NEP (i.e. GPP minus ER) to light, with a y-intercept equal to –ER (since GPP is 0 is the absence of light) and GPP as the light-dependent increase above this level (Figure 1) (Holtgrieve et al. 2010). Defining ER as the intercept of the P-I curve neglects possible light-dependent changes in ER, but nonetheless is a common assumption for studies of ecosystem metabolism (Schindler et al. 2017). P-I curves are modeled in various ways (Jassby and Platt 1976, Henley 1993, Behrenfeld and Falkowski 1997) but often take the form of a saturating relationship with parameters such as the maximum rate of GPP attained at high light levels ("maximum GPP") and the rate at which GPP increases with light at low light levels ("initial slope").

Variation in NEP can arise from two sources: variation in light levels and variation in the biological processes governing the P-I curve. Aquatic ecosystems experience substantial variability in light due to changes in solar irradiance (e.g. diel and seasonal cycles; changes in cloud cover) and water clarity (e.g. resuspension of sediment; phytoplankton blooms). Changes in the P-I curve result from biological processes across organizational scales (Falkowski 1984). At the cellular level, photosynthetic rates vary with temperature, availability of nutrients, and physiological regulation (Falkowski 1984, Davison 1991, Edwards et al. 2011). For example, aquatic primary producers can increase production of photosynthetic pigments in response to chronic light limitation, thereby increasing their sensitivity to a given amount of light (Falkowski 1984, Moore et al. 2006, Dubinsky and Stambler 2009). At the population and community scale, photosynthetic rates may change due to variation in primary producer biomass, rapid evolution, and species composition (Carpenter and Kitchell 1984, Vrede et al. 2009, Schwaderer et al. 2011, Edwards et al. 2015), which can either occur in response to the physical environment or through interactions with other organisms (e.g. competition with other producers; herbivory by

zooplankton) (Carpenter and Kitchell 1984, Bergquist and Carpenter 1986, Yoshiyama et al. 2009, Hashioka et al. 2013).

ER comprises respiration by both autotrophs (i.e. primary producers) and heterotrophs and is determined in part by their respective biomasses (Del Giorgio et al. 1999). The extent to which heterotrophic production (and therefore respiration) is coupled with GPP is expected to depend on the dominant source of carbon in the system (Del Giorgio et al. 1999, Solomon et al. 2013). Many freshwater ecosystems are dominated by allochthonous (external) sources of carbon, especially those with extensive boundaries along terrestrial landscapes (e.g. streams) (Thorp and Delong 2002, Staehr et al. 2010b, Solomon et al. 2013). In these ecosystems, heterotrophic respiration (and total ER) can exceed GPP and changes in response to variation in the allochthonous resource (Hanson et al. 2003, Cole et al. 2006, Ask et al. 2009, Staehr et al. 2010b). Conversely, in ecosystems with relatively low allochthonous inputs, ER should be more strongly coupled to GPP, due to both respiration by primary producers and respiration by heterotrophs whose primary source of carbon is in situ GPP (Solomon et al. 2013). On shorter times scales, ER is driven by variation in temperature (Solomon et al. 2013, Demars et al. 2016, Song et al. 2018), due to the thermal dependence of enzyme kinetics. GPP may also be influenced by changes in temperature, and while ER is assumed to be more sensitive to temperature (but see Demars et al. 2016), little is known about how these temperature-metabolism relationships compare to the coupling of ER and GPP through other biological processes.

Despite the large potential for temporal variation in the P-I curve, the relative contribution of this variation to ecosystem metabolism is largely unknown. In this study, I present a statistical model of ecosystem metabolism with time-varying parameters governing the

P-I curve, and I fit the model to observations of dissolved oxygen (DO) from Lake Mývatn in northern Iceland. The model includes explicit dependence of GPP and ER on temperature, as temperature is a well-documented driver of the physiological processes contributing to ecosystem metabolism (Demars et al. 2016). Beyond the effects of light and temperature, the parameters of the P-I curve (e.g. "maximum GPP" and "initial slope") vary smoothly through time as stochastic processes and thereby capture temporal variation in ecosystem metabolism due changes in the P-I curve itself. These changes could be due to a variety of biological factors (e.g. organismal traits, biomass, or community composition) which themselves could be influenced by changes in other variables (e.g. herbivory, nutrient availability, or lake mixing).

This approach builds on a large body of work utilizing time series of DO to infer ecosystem metabolism in aquatic systems, relying on the fact that photosynthesis releases oxygen while aerobic respiration consumes oxygen and the respective conversions between CO_2 and O_2 fluxes assumed to be approximately equal (Odum 1956, Hanson et al. 2008, Holtgrieve et al. 2010, Staehr et al. 2010a). However, common methods either do not provide an explicit characterization of the P-I relationship (i.e. model-free "bookkeeping") (Staehr et al. 2010a) or assume that the relationship between ecosystem metabolism and light or other covariates (e.g. temperature; chlorophyll) is fixed over some time interval (usually one to several days) (Holtgrieve et al. 2010, Winslow et al. 2016, Richardson et al. 2017). In both cases, temporal variation in responses to light is investigated by applying the method iteratively to separate sequences of data that are assumed to be independent. In contrast, the present method explicitly models the temporal changes in the P-I curve utilizing many time periods simultaneously. This takes advantage of the fact that the physical and biological processes governing ecosystem metabolism and other aspects of DO dynamics are correlated through time, which means that this

shared information can be used to inform the parameter estimates across all time points. Furthermore, this method is statistically unified since it uses all of the data to fit a single model, which facilitates characterizing the uncertainty in the estimates of temporally variable parameters of the P-I curve.

I applied this method to DO time series from Lake Mývatn spanning six summers to (a) characterize temporal variation in Mývatn's P-I relationship, (b) evaluate the contribution of this variation to changes in ecosystem metabolism, and (c) evaluate the contribution of different variables to the coupling of GPP and ER. Mývatn is dynamic, with much of its ecology being driven by midges (Diptera: Chironomidae) that fluctuate across four orders-of-magnitude and constitute the majority of the lake's secondary biomass production (Einarsson et al. 2004). The midges are ecosystem engineers and locally stimulate both GPP and ER, such that their population fluctuations may contribute to large variation in Mývatn's whole-lake metabolism (Herren et al. 2017). Furthermore, while Mývatn is generally dominated by benthic primary production during clear-water phases, it is also subject to cyanobacterial blooms that are occasionally very thick and may shift the majority of production from benthic to pelagic (Einarsson et al. 2004). The dynamic nature of Mývatn's ecology likely results in large variation in its responses to light and temperature, making it a useful test case for partitioning the relative contribution of different sources of variation in ecosystem metabolism.

Methods

Study System

Mývatn is located in northeast Iceland (65°40'N 17°00'W), approximately 100km south of the Arctic Circle and has a tundra-subarctic climate (Björnsson and Jónsson 2004). The lake has two basins (north and south) connected by a narrow channel; the south basin is larger and is the central focus of this study. The south basin is shallow (mean depth = 2.3m) and does not stratify in the summer, with water temperature tracking the air temperature and ice cover for 190 days per year on average (Einarsson et al. 2004). Phosphorous- and silicon-rich cold springs (inputs of N, P, and Si = 1.4, 1.5, and 340 g m-2 y-1) feed the south basin (Ólafsson 1979). Consequently, Mývatn is naturally eutrophic with high primary production. According the previous estimates, the majority of primary production is benthic and attributable to benthic diatoms (chiefly Fragilaria) and mats of filamentous green algae (Cladophora glomerata and Aegagropila linnai) (Ólafsson 1979, Einarsson et al. 2004). However, there are occasional blooms of green algae, diatoms, and cyanobacteria which may make substantial contributions to primary production. The cyanobacterial blooms (chiefly Dolichospermum spp.) in particular can be quite thick (peak phycocyanin concentration >200 μg L-1), blocking almost all light below 5cm of the water surface (JS Phillips unpublished data). Benthic primary production supports large populations of midges (>30 species) that in peak years compose over two-thirds of the lake's secondary production and show dramatic fluctuations in abundance spanning 4 orders of magnitude (Lindegaard and Jónasson 1979, Einarsson et al. 2004, Ives et al. 2008). The midges in turn are an important food source for ducks and fish (sticklebacks, brown trout, and arctic char). The two dominant midge genera (Tanytarsus and Chironomus) are tube-building ecosystem engineers that locally stimulate benthic GPP by providing a stable 3-dimensional substrate for diatom growth (Hölker et al. 2015, Herren et al. 2017).

Data Collection and Processing

A sonde multiprobe (Hydrolab DS5X, Hach Co., CO, USA) was deployed near the center of Mývatn's south basin during the summer months (late May through late August) from 2012 to 2018. This location corresponded with a routine monitoring station that was manually sampled approximately weekly during the periods of sonde deployment for standard water column variables (e.g. DO, light, and temperature) as well as benthic variables such as larval midge density and sediment chlorophyll. The sonde was attached to a buoy so that the sensors sat at a depth of 0.5m. This location is quite shallow (3.3m) and well-mixed, such that the sensor measurements represent integrated values throughout the water column. The sonde measured temperature, DO, turbidity, and phycocyanin (a cyanobacterial pigment) every 30 minutes. However, DO data were unavailable from 2014 due to probe failures, so only 2012-2013 and 2015-2018 data were used in this analysis. The sensor data had occasional anomalies, which I removed by first excluding values that exceeded thresholds of plausibility (< 5 mg L-1 for DO and > 18°C for temperature) based on routine monitoring with handheld DO probes (ProODO, YSI Co., OH, USA) or visual inspection (> 200 volts for turbidity). I then excluded values that deviated from the mean on a given day by more than two standard deviations. This resulted in the omission of 5.6% of the observations. There were no indications of sensor drift (Figures B3 and B4), and therefore no drift correction was applied. I averaged the half-hourly measurements to an hourly scale to match the resolution of the local weather data (solar irradiance and wind speed), which resulted in minimal loss of information due to the high temporal autocorrelation in the measurements (Figure B7).

Light readings were taken either every 15 or 30 minutes (depending on the year) with a light/temperature logger (HOBO Pendant, Onset Computer Corporation, MA, USA) attached to

the top of the buoy deploying the sonde; I averaged these data to an hourly scale for further analysis. HOBO loggers measure visible light intensity in lux, which is closely related to the spectral range (400-700nm) used to calculate photosynthetically active radiation (PAR; measured in μmol-photons m-2 s-1) (Thimijan and Heins 1983). I also obtained downwelling solar irradiance from the local whether station (in Watts). When converted to PAR with standard scalar conversions (Thimijan and Heins 1983), the solar irradiance and HOBO logger data were strongly correlated ($r = 0.79$; Figures B1 and B2). However, solar irradiance generally showed higher peaks at midday than the HOBO logger data, likely due to the wider spectral range used for calculating solar irradiance (Thimijan and Heins 1983). When available I used HOBO logger data. However, HOBO logger data were missing in 2012, so I used solar irradiance instead (solar irradiance was not available in 2018, so it would not have been possible to use the same metric for all 6 years of data). Both HOBO logger and solar irradiance data were scalarly converted to PAR (Thimijan and Heins 1983); for simplicity I henceforth refer to them both as "PARout" To estimate light below the water surface, I regressed PAR measurements taken approximately 5cm below the water surface during routine monitoring (using a Li-192 Quantum Underwater Sensor, Li-COR, NE, USA) against PARout at the corresponding time with the y-intercept forced through 0, since light beneath the water surface must be 0 when PARout is 0. I used resulting regression equation ($y = 0.44 \times PARout$; $df = 54$; $R2 = 0.80$) to translate PARout into PAR below the water surface for each time point in the sonde data.

Primary production occurs throughout the water column and in the benthos. Therefore, I estimated the mean light level throughout the water column given observed variation in surface light and water clarity as

$$\overline{L_{d,h}(z)} = \frac{1}{z_{\max}} \int_0^{z_{\max}} L_{d,h}^0 \; e^{-c_{d,h}\, z} \; dz = \frac{L_{d,h}^0 (1 - e^{-c_{d,h}\, z_{\max}})}{z_{\max}\, c_{d,h}} \quad (1)$$

where z [m] is the vertical position in the water column, z_{max} is the water column depth (3.3m), $L^0_{d,h}$ is the PAR [μmol-photons m^{-2} s^{-1}] at the water surface ($z = 0$) and $c_{d,h}$ is the rate at which light attenuates through the water column. For simplicity, I henceforth denote $\overline{L_{d,h}(z)}$ as $L_{d,h}$. To estimate $c_{d,h}$, I regressed observe light attenuation calculated from weekly measurements of PAR at 0.5m depth intervals against turbidity measured simultaneously by the sonde. I used this regression (y = 0.43 + 0.059 × turbidity; df = 34; R^2 = 0.66) to predict $c_{d,h}$ from turbidity for each time point in the sonde data.

Model Structure

I separated net ecosystem production ($NEP_{d,h}$) in hour h on day d into contributions from $GPP_{d,h}$ and $ER_{d,h}$ as

$$NEP_{d,h} = GPP_{d,h} - ER_{d,h} \quad (2)$$

with all three components quantified as total change in oxygen throughout the water column, resulting a flux per unit area of the lake [mg O_2 m^{-2} h^{-1}]. The following equations govern the dependence of $GPP_{d,h}$ and $ER_{d,h}$ on hourly measurements of light and temperature; these equations collectively specify the P-I curve for a given set of parameters that vary between days. In other words, a separate P-I curve is estimated for each day (although informed by data from all days), factoring out the influence of temperature.

I modeled $GPP_{d,h}$ as a saturating function of average light throughout the water column $L_{d,h}$ [μmol-photons m^{-2} s^{-1}]:

$$GPP_{d,h} = \beta_{d,h} \tanh\left(\frac{\alpha_{d,h}}{\beta_{d,h}} L_{d,h}\right) \quad (3)$$

where $\beta_{d,h}$ [mg O_2 m^{-2} h^{-1}] is the maximum potential GPP ("max GPP"; Fig 1) attained at high light levels and $\alpha_{d,h}$ [mg O_2 s μmol-photons^{-1} h^{-1}] is the rate at which GPP increases with $L_{d,h}$ when $L_{d,h}$ is near zero ("initial slope") (Jassby and Platt 1976, Holtgrieve et al. 2010).

While many studies of lake metabolism assumed that GPP is insensitive to temperature, primary production tends to increase with temperature particularly when light is saturating (Edwards et al. 2016, Demars et al. 2016, Richardson et al. 2017). Therefore, I modeled the maximum potential GPP as

$$\beta_{d,h} = \beta_d^0 \, \gamma_\beta^{\frac{T_{d,h} - T_0}{\Delta T}} \quad (4)$$

where β_d^0 is the maximum potential GPP on a given day at the mean water temperature for the entire time series T_0 [12°C], $T_{d,h}$ is the observed water temperature, and ΔT = 1°C to make the exponent dimensionless. The parameter γ_β [dimensionless] describes the scaling of $\beta_{d,h}$ with changes in temperature. For example, if γ_β = 1.2, then a 1°C increase in $T_{d,h}$ above the mean temperature would result in a 20% increase in $\beta_{d,h}$. This is phenomenologically similar to the Arrhenius equation that has previously been used to model the relationship between ecosystem

metabolism and temperature based on enzyme kinetics (Demars et al. 2016, Schindler et al. 2017). I opted for this alternative parameterization as it is more concise.

Daily changes in max GPP not due to temperature (e.g. changes in primary producer biomass or biomass-specific photosynthetic rates) were described by changes in β_d^0, which I modeled as a multiplicative stochastic process analogous to exponential population growth:

$$\begin{aligned} \beta_d^0 &= \beta_{d-1}^0 \, e^{\delta_{d-1}^{\beta}} \\ \delta_d^{\beta} &\sim \mathcal{N}\left(0, \sigma_{\beta}\right) \end{aligned} \tag{5}$$

where δ_d^{β} is daily the rate of change in β_d^0 with standard deviation σ_{β}. This is equivalent to modeling β_d^0 as a random walk on a log scale; this ensures that β_d^0 remains positive (negative values are biologically unintelligible). The stochastic process implies temporal autocorrelation in the estimates of β_d^0, which has the consequence of smoothing the estimates through time with the degree of smoothing being influenced by the observed data (Zeng et al. 1998, Ives and Dakos 2012). Analogously, I modeled daily variation in the initial slope of the P-I curve α_d as

$$\begin{aligned} \alpha_d &= \alpha_{d-1} \, e^{\delta_{d-1}^{\alpha}} \\ \delta_d^{\alpha} &\sim \mathcal{N}(0, \sigma_{\alpha}) \end{aligned} \tag{6}$$

with rate of change δ_d^{α} and standard deviation σ_{α}.

Daily and hourly variation in ER was modeled as a function of temperature:

$$ER_{d,h} = \rho_d \, \gamma_{\rho}^{\frac{T_{d,h} - T_0}{\Delta T}} \tag{7}$$

where ρ_d [mg O_2 m^{-2} h^{-1}] is the baseline respiration rate at T_0 and γ_ρ [dimensionless] is the temperature scaling of ER_d (Holtgrieve et al. 2010). Similar to β_d^0 and α_d, daily variation in ρ_d (e.g. changes in biomass or biomass-specific respiration of all aerobic organisms) was modeled as

$$\begin{aligned} \rho_d &= \rho_{d-1} e^{\delta_{d-1}^{\rho}} \\ \delta_d^{\rho} &\sim \mathcal{N}\left(0, \sigma_\rho\right) \end{aligned} \tag{8}$$

with rate of change δ_d^{ρ} and standard deviation σ_ρ. Various studies have documented diel variation in ER in response to factors other than temperature (Sadro et al. 2011), and these have occasionally been incorporated into models of lake ecosystem metabolism (e.g. Schindler et al. 2017). However, the primary goal of the present analysis was to estimate variation in ecosystem metabolism rates between days, and it is likely that incorporating diel variation in respiration not due to temperature would have a minimal impact on the estimates of daily variation (Hanson et al. 2008). Therefore, to reduce the complexity of the model I did not include diel variation in ER driven by factors other than temperature.

I estimated parameters for equations 2-8 from the observed DO time-series by modeling the hourly concentration of DO [mg O_2 m^{-3}] as

$$DO_{d,h} = DO_{d,h-1} + \frac{\Delta h\left(NEP_{d,h-1} + EXC_{d,h-1}\right)}{z_{mix}} + \epsilon_{d,h-1}^{proc} \tag{9}$$

where Δh = 1 hour, $EXC_{d,h}$ [$mg\ O_2\ m^{-2}\ h^{-1}$] is the exchange of oxygen between the water and the air, z_{mix} is the mixing depth (3.3m), and $\epsilon_{d,h-1}^{proc}$ is the processes error, defined as

$$\epsilon_{d,h}^{proc} \sim \mathcal{N}\left(0,\ \sigma_{proc}\right) \tag{10}$$

where σ_{proc} is the standard deviation of the stochastic changes in DO (see below for description of the observation model linking equation 9 to the observed data).

$EXC_{d,h}$ is the oxygen flux across the air-water boundary, defined with respect to the water (i.e. a negative flux corresponds to a loss of O_2 from the water):

$$EXC_{d,h} = k_{d,h}\left(DO_{d,h}^{eq} - DO_{d,h}\right) \tag{11}$$

where $DO_{d,h}^{eq}$ is the atmospheric equilibrium (i.e. saturation) concentration of oxygen based on temperature, pressure, and salinity (Staehr et al. 2010a). The parameter $k_{d,h}$ is known as the "piston velocity" and scales the rate of gas exchange across the air-water boundary. While some studies estimate $k_{d,h}$ directly in the field, in most studies it is modeled as a function of wind speed with the following equation:

$$k_{d,h} = \left(k_0^{600} + k_1^{600}\omega_{d,h}^{k_2^{600}}\right)\left(\frac{\mathrm{Sc}_{\mathrm{O}_2}\left(T_{d,h}, S\right)}{600}\right)^{-0.5} \tag{12}$$

where ω is the wind speed [$m\ s^{-1}$], and k_0^{600}, k_1^{600}, and k_2^{600} are constants defined for a standard Schmidt number of 600, which is converted for oxygen at the observed temperature and salinity

based on its corresponding Schmidt number $\mathrm{Sc}_{\mathrm{O}_2}(T_{d,h}, S)$ (Wanninkhof 1992, Staehr et al. 2010a). While in principle it is possible to estimate k_0^{600}, k_1^{600}, and k_2^{600} from DO time series (similar to Holtgrieve et al. 2010), empirical estimates for these constants exist in the literature (Cole and Caraco 1998, Staehr et al. 2010a) and are widely utilized in studies of ecosystem metabolism. I opted for using the existing values for statistical simplicity and consistency with previous studies.

It is important to note that equation 9 does not strictly follow mass-balance; because $ER_{d,h}$ is not a function of oxygen concentration, theoretically it can exceed the amount of available oxygen. In most applications (including here), oxygen concentrations are sufficiently high that respiration is not oxygen limited. However, it is important to recognize the potential for negative values to arise during model fitting, as this influences the choice of appropriate descriptions of process and observation error. For example, using a lognormal error model might fail because it could require taking the log of negative predicted values. I used the most straightforward approach (as in other studies) of Gaussian process and observation error, relying on the data to guide the model to physically meaningful values. This yielded the following observation equation:

$$DO_{d,h}^{obs} = DO_{d,h} + \epsilon_{d,h}^{obs} \tag{13}$$

where $\epsilon_{d,h}^{obs}$ is the observation error, modeled as

$$\epsilon_{d,h}^{obs} \sim \mathcal{N}(0, \sigma_{obs}) \tag{14}$$

where σ_{obs} is the standard deviation of the observation process. The central statistical difference between these types of error is that process errors propagate through time and manifest as relatively smooth variations, while observation errors do not propagate and appear "noisy" (Box et al. 1994).

Model Fitting

Observed DO time series were fit to all six years (2012-2013 and 2015-2018) simultaneously. Continuous series of hourly observations (i.e. without missing data or spanning years) were initialized as separate time series, with starting values modeled with observation error according to equations 13 and 14 with the mean of the Gaussian distribution set to the observed value for that time step. The presence of missing values meant that some continuous series had a small number of observations. While in principle the model could accommodate these short series because the parameters estimates would be informed by the longer series, the short series likely contributed little information. Therefore, I removed any continuous series with fewer than 24 observations (i.e. one day's worth of observations). A single continuous stochastic process was used to model β_d^0, α_d, and ρ_d for each year, with the model "smoothing" across unobserved days.

I fit the model with a Bayesian approach via Stan 2.17.0 run from R 3.5.1 (R Core Team 2018) using the 'rstan' package (Stan Development Team 2018). Stan utilizes Hamiltonian Monte Carlo to generate a multivariate stochastic process that converges to a stationary distribution approximating the joint posterior distribution for the unobserved variables (Carpenter et al. 2017). I centered the observed DO on its mean and divided by its standard deviation across all time points and then scaled the model parameters accordingly (Appendix B).

Scaling the model stabilized the fitting process and also facilitated the use of normal distributions on unit scale as weakly informative priors (standard deviations of 1 and truncated at 0 or 1 for those parameters with corresponding lower bounds; Appendix B). I used four independent chains with diffuse initial values and 2000 iterations including a 1000-iteration "warm-up". Convergence was determined by examining trace plots for parameter estimates and the potential scale reduction factor ($\hat{R}$), which quantified the relative variance within and between chains to determine whether stationarity was achieved (values near 1 indicate convergence, with 1.1 as a generally accepted upper threshold). See Appendix B for further details for the fitting routine and additional diagnostics.

Attempting to estimate both σ_{obs} and σ_{proc} independently resulted in very inefficient sampling of the posterior distribution as σ_{obs} slowly approached 0 (Figure B5). Therefore, I fit the model with σ_{obs} set to 1% of the observed standard deviation in DO (which was computationally more convenient than setting it to exactly 0). This resulted in much more efficient sampling and estimates for model parameters that were similar to the model with σ_{obs} estimated independently (Figure B6). It was not surprising that σ_{obs} was low, as the observed time series of DO (filtered for anomalous values) lacked the appearance of high frequency noise that would be indicative of observation error (Figure B7). This is partially due to the fact that I used hourly data that were averaged from half-hourly data; observation error would likely have been more substantial for higher frequency data.

The posteriors for all parameters were unimodal and at most only modestly skewed, justifying the use of quantiles to provide summaries for parameter estimates. Throughout I use posterior medians for point estimates and 16% and 84% quantiles for the bounds of 68% uncertainty intervals ('$ui_{68\%}$') to match the coverage of standard errors. I expressed values for

total daily values for GPP, ER, and NEP in units g O_2 m^{-2} day^{-1} rather than mg O_2 m^{-2} h^{-1} to facilitate a more natural comparison to other reported values for ecosystem metabolism in both aquatic and terrestrial systems.

Model Assessment

The main objective of this analysis was to infer temporal variation in the biological processes contributing to the P-I curve and ecosystem metabolism. Therefore, I assessed the performance of the model by addressing two questions: (1) how much of the variation in observed DO was attributable to explicitly modeled biological processes (i.e. the strength of the biological signal) and (2) what was the risk of inferring temporal variation in the P-I curve when it actually did not vary (i.e. the risk of "false positives" for inferred changes in the P-I curve). For the former, I rearranged equation 9 to yield

$$\frac{\Delta DO_{d,h-1}}{\Delta h} = \frac{DO_{d,h} - DO_{d,h-1}}{\Delta h} = \frac{NEP_{d,h-1}}{z_{mix}} + \frac{EXC_{d,h-1}}{z_{mix}} + \frac{\epsilon_{d,h-1}^{proc}}{\Delta h} \quad (15)$$

where $\frac{\Delta DO_{d,h-1}}{\Delta h}$ was the observed hourly change in DO (since observation error was negligible). Because the primary goal of this assessment was to relate the biological processes to observed DO, I subtracted the atmospheric exchange term $\frac{EXC_{d,h-1}}{z_{mix}}$ from each side to obtain a "corrected" change in DO:

$$\frac{\Delta DO_{d,h-1}^{corrected}}{\Delta h} = \frac{\Delta DO_{d,h-1}}{\Delta h} - \frac{EXC_{d,h-1}}{z_{mix}} = \frac{NEP_{d,h-1}}{z_{mix}} + \frac{\epsilon_{d,h-1}^{proc}}{\Delta h} \quad (16)$$

The corrected change in DO was therefore due to either NEP (expressed volumetrically) or processes error (expressed per hour). By comparing the time-series of corrected DO change and NEP, I was able to visually assess the relative contributions of NEP and process error, providing a measure of the strength of the biological signal. For clarity, I present the summed hourly values of corrected change in DO and NEP for each day (i.e. the net daily fluxes) (Figure 2).

To determine the risk of false positives in inferred changes of the P-I curve, I fit the model to simulated data with the time-varying parameters of the P-I curve (β_d^0, α_d, and ρ_d) fixed to their mean values according to the model fit to the real data. The simulated data implied that standard deviations associated with the changes in the parameters of the P-I curve (σ_β, σ_α, and σ_ρ) had "true" values of 0. Therefore, the model fit to these data should produce estimates shifted toward 0 relative to the estimates for the actual data. Since the goal was to assess the model as fit to the particular data used in this analysis, I used the observed environmental data and the particular sequence of process errors as inferred from the original model fit to construct the simulated data. Therefore, the simulated data approximated what would have actually been observed if the parameters of the P-I curve had been fixed through time (see Appendix B for full simulation details). The values of σ_β, σ_α, and σ_ρ for the simulated data also provide a benchmark to which to compare the corresponding estimates for the real data, which are dimensionless.

Variance Partitioning

I partitioned variance in GPP, ER, and NEP into contributions from light, temperature, and the time-varying parameters using the "delta method" (Clark 2007). This approach differs from a more conventional partitioning (e.g. ANOVA) by explicitly accounting for nonlinearities and covariances between variables. The delta method is a locally linear approximation of the

variance relative to the mean values of the input variables (see Appendix B for full mathematical details). In general, the contribution of a parameter (e.g. temperature) to the variance in a function of that parameter (e.g. GPP) is determined by (i) the sensitivity of the function to changes in the parameter and (ii) the overall variability of that parameter. For example, temperature could explain a large portion of the variance in GPP either because (i) GPP is very sensitive to temperature or (ii) because temperature itself has a high variance. For each variable, I report the squared coefficient of variation ("CV^2") and scaled sensitivity, with the latter including the effect of covariances between variables. The product of CV^2 and scaled sensitivity equals the relative contribution of a given variable to the variance in the response.

Code and data availability

Data and code can be found at https://github.com/jsphillips2/sonde_oxygen.

Results

Model Assessment

To assess the extent to which biological fluxes of O_2 as inferred by the model accounted for variation in the observed data, I compared the corrected DO flux (equation 15) to NEP (Figure 2). The corrected DO flux excluded the atmospheric exchange and was therefore due to either modeled biological processes (i.e. NEP) or to processes (physical or biological) not explicitly included in the model (i.e. process error). Much of the variation in the observed DO flux was attributable to modeled NEP, although with substantial variation due to process error (Figure 2). Nonetheless, this demonstrates that the ecosystem metabolism inferred by the model reflects real variation in the observed data.

To assess the risk of false positives in inferred temporal changes in the P-I curve I fit the model to simulated data with the parameters of the P-I curve (β_d^0, α_d, and ρ_d) fixed to their means for a given year, implying that the corresponding standard deviations for their daily variations (σ_β, σ_α, and σ_ρ) equaled 0. When fit to the real data, the estimates of σ_β, σ_α, and σ_ρ were all clearly different from 0, indicating substantial temporal variation in the corresponding metabolism parameters (Table I). In contrast, the variance estimates for the simulated data were all strongly right skewed, with posterior densities concentrated near 0 (Figure B9; Table I). This suggests that the temporal variation in β_d^0, α_d, and ρ_d inferred by the model reflected real changes in the underlying P-I curve.

Parameter Estimates

Table I presents the estimates and 68% uncertainty intervals for the model parameters. ER was very sensitive to temperature, with γ_ρ = 1.15 implying that 1°C increase (or decrease) from the reference temperature would result in a 15% increase (or 8.7% decrease) in ER. GPP was also sensitive to temperature (γ_β = 1.06), although substantially less so than ER. The metabolism parameters β_d^0, α_d, and ρ_d all varied through time, with the standard deviations of their associated stochastic processes clearly different from 0 (Table I; Figure B9). This resulted in substantial temporal variation in the resulting P-I curves (Figure 1).

The initial slope of the P-I curve (α_d) was by far the most temporally variable parameter relative to its overall magnitude (σ_α = 0.21), indicating that the response of GPP to variation in light under light-limiting conditions was very dynamic (Figure 3). Several years showed similar seasonal patterns, with α_d starting low and increasing to a peak mid-summer, followed by a steep decline. However, in 2012 and 2016 α_d declined gradually throughout the season. The

estimate for α_d in 2012 was substantially lower than in other years, which could be due to the use of solar irradiance rather than HOBO logger light data.

Maximum GPP at high light (β_d^0) was slightly more variable relative to overall magnitude (σ_β = 0.06) than the baseline ER (ρ_d) (σ_ρ = 0.05). However, this difference was exaggerated by the fact that magnitude of β_d^0 greatly exceeded the ρ_d, making the maximum GPP substantially more variable on the absolute scale (Figure 4). Maximum GPP showed fairly consistent seasonal trends, with steady increases from early to mid-summer followed by a plateau or a modest decline. Temporal variation in the baseline ER loosely mirrored variation in maximum GPP, with the exception of 2015 when maximum GPP steadily increased while baseline ER decreased throughout the season.

The parameters of the P-I curve were estimated using the average light throughout the water column, which varied due to changes in surface light and to water clarity. However, variation in light attenuation was estimated from turbidity and was therefore subject to some uncertainty. Furthermore, other studies typically use PAR at the surface as the measure of light for fitting P-I curves. To explore this further, I fit the model using surface light and compared the resulting parameter estimates to those using the average water column light (as above). The estimates of baseline ER and maximum GPP were almost identical, as they should be given that they characterize metabolism when light is 0 (i.e. when surface and water column light are the same) and when light is saturating (i.e. when differences between surface and water column later are irrelevant) (Figure B10). The average value of the initial slope of the P-I curve was substantially lower for the model fit using surface light, which again was as expected because the surface light values were always higher than the water column values (Figure B11). However, the temporal patterns were very similar, with the largest differences being relatively high values

in 2015 and 2018 when using the water column light. These two years had substantial declines in water column light due to reduced water clarity associated with cyanobacterial blooms, which suggests that the initial slope of the P-I curve partially compensated for the declines in water clarity.

Ecosystem Metabolism

Mean daily estimates of GPP (5.34 g O_2 m^{-2} day^{-1}; $ui_{68\%}$ = [5.20, 5.49]) and ER (4.63 g O_2 m^{-2} day^{-1}; $ui_{68\%}$ = [4.46, 4.79]) were large overall, reflecting Mývatn's eutrophic state. Both showed substantial variation through time, both within and between years (Figure 5). However, because GPP and ER were correlated (r = 0.73; Figure 6), NEP remained comparatively stable and slightly above zero (0.72 g O_2 m^{-2} day^{-1}; $ui_{68\%}$ = [0.64, 0.80]). This indicates that Mývatn was net autotrophic during the summer months over this time period. Correlation between GPP and ER is expected in ecosystems dominated by autochthonous production due to respiration by autotrophs and heterotrophic respiration of organic carbon directly resulting from endogenous GPP (Del Giorgio et al. 1999, Solomon et al. 2013). However, the correlation between maximum GPP (β_d^0) and baseline ER (ρ_d) estimated from the model was only modest (r = 0.51; Figure 7) and the remainder of the correlation between overall GPP and ER was due to their joint responses to temperature.

Most of the variation in GPP was due to variation in the maximum GPP, with temperature as the second most important contributor (Table II). In contrast, the contributions of variation in light and the initial slope of the P-I curve were quite small, despite the fact that these two variables had comparatively high coefficients of variation. This is because GPP was relatively insensitive to changes in these variables, due the average light level (144 μmol-

photons m^{-2} s^{-1}) generally being in the saturated part of the P-I curve. Variation in baseline ER and temperature made comparable contributions to variance in ER, owing to their similar variability and sensitivities. Note that the variance partitioning is an approximation around the means of the input variables (i.e. light, temperature, and time-varying parameters) and so reflects the contribution of variation in different variables centered around the average conditions.

Since GPP and ER had opposite effects on NEP, variables with correlated effects on GPP and ER had reduced contributions to the variance in NEP. Consequently, both temperature (positively related to both GPP and ER) and the baseline ER (correlated with max GPP) had negligible or even slightly negative contributions. The contribution of variation in light was slightly higher for NEP than for GPP, while the contribution of variation in the initial slope was slightly lower. Maximum GPP was by far the largest contributor to variation in NEP, with a relative contribution of 0.89 despite being correlated with baseline ER. This is because the sensitivity of NEP to maximum GPP was substantially higher than its sensitivity to baseline ER.

Drivers of Maximum GPP

Maximum GPP at high light (β_d^0) was quite variable, and it was the dominant contributor to variation in overall GPP and NEP. To explore the possible drivers of variation in maximum GPP, I fit linear models regressing β_d^0 against daily mean phycocyanin concentration (as a measure of cyanobacterial abundance) and larval midge abundance. Phycocyanin data were available for all days for which there were estimates of β_d^0, while midge data were only available approximately weekly. Therefore, I fit two models: (1) including phycocyanin for all days, and (2) including both phycocyanin and midge abundance for those days when midge samples were collected. For midge abundance, I used total counts of the two dominant midge taxa previously

shown to have positive effects on benthic primary production (*Tanytarsus gracilentus* and *Chironomus islandicus*) (Herren et al. 2017) averaged across replicated samples for a given date. I z-transformed (subtracted the mean and divided by the standard deviation) all variables and included an autoregressive correlation structure grouped by year to account for temporal autocorrelation in estimates of β_d^0 (Pinheiro et al. 2018). Maximum GPP was positively related to phycocyanin in both the model for all observations (df = 343; t = 3.65; p = 0.0003) and the model for days with midge samples (df = 39; t = 2.87; p = 0.0066) (Figure 8). In contrast, there was no relationship between maximum GPP and midge abundance (df = 39; t = –0.20; p = 0.84) (Figure B12). This suggests that increases in cyanobacterial abundance during blooms increased the maximum GPP, while midges had no effect.

Discussion

In this study I used time-series of DO from Lake Mývatn to estimate temporal variation in P-I curves and ecosystem metabolism across six summers. To do so, I employed a new approach that extended previous methods by explicitly modeling temporal variation in the parameters governing ecosystem metabolism, taking advantage of the shared information across the full set of data. All of the parameters of the P-I curve (maximum GPP, initial slope of the P-I curve, and baseline ER) varied substantially through time, with the initial slope being particularly variable. However, variation in the maximum GPP made the largest contribution to variation in estimated NEP with respect to average conditions, with this variation being associated with increases in the abundance of cyanobacteria during blooms. Overall, these results illustrate how changes in the P-I curve contributes substantially to temporal variation in lake ecosystem metabolism.

The method is quite flexible and could be readily applied to DO time-series from other systems to quantify temporal variation in components of ecosystem metabolism. While the present analysis places emphasis on the P-I curve as something of interest in itself, this method has utility for the broader goal of identifying trends and drivers of metabolism by removing short-term fluctuations and thereby clarifying longer-term changes. Diel variation in light and temperature major contributors to short-term variability in ecosystem metabolism, potentially obscuring other drivers that may be of primary interest (Coloso et al. 2011, Richardson et al. 2017). In the present model, all variation in ecosystem metabolism due to biological and physical variables other than light and temperature (e.g. biomass, nutrient concentration, etc.) manifest entirely in the time-varying parameters of the P-I curve. For example, the maximum GPP can be understood as the ecosystem-wide photosynthetic potential (Jassby and Platt 1976, Behrenfeld and Falkowski 1997) that could be directly linked to drivers such as primary producer biomass or nutrient limitation. Furthermore, the model utilizes information across all time points to directly estimate temporal variation in the parameters of the P-I, which essentially smooths across short-term fluctuations, thereby clarifying the longer-term variations that are often of central interest.

Many studies have considered the coupling of GPP and ER, because the balance between these two processes determines whether a given ecosystem is a net carbon source ("heterotrophic") or sink ("autotrophic"). It is generally argued that systems dominated by autochthonous carbon should have a strong coupling between GPP and ER, due both to respiration by photosynthetic organisms themselves and heterotrophic respiration of the organic carbon fixed through primary production (Del Giorgio et al. 1999, Solomon et al. 2013, Richardson et al. 2017, Schindler et al. 2017). Recent work has also suggested that correlated responses to temperature can couple GPP and ER (Solomon et al. 2013, Demars et al. 2016,

Richardson et al. 2017), and distinguishing between these two mechanisms is important for predicting the response of GPP-ER coupling to long-term change (e.g. climate). In Mývatn GPP and ER were strongly correlated, and this was due to both temperature (to which both GPP and ER were sensitive) and correlations between the time-varying ecosystem metabolism rates (maximum GPP and baseline ER, controlling for temperature). Even though ER was more sensitive to temperature than GPP relative to its overall scale (i.e. $\gamma_\rho > \gamma_\beta$), the magnitude of GPP was much greater; consequently, the temperature-driven variations in ER and GPP largely negated each other in the resulting NEP. This suggests that net carbon flux from Mývatn is relatively insensitive changes in temperature. In contrast, the coupling between maximum potential GPP and baseline ER was only modestly positively (R = 0.51), allowing variation in maximum GPP to have a large effect on NEP. Given that Mývatn likely has low allochthonous inputs (being surrounded by low-productivity tundra; Gratton et al. 2008), it is surprising the temperature-independent coupling of GPP and ER is not stronger. Solomon et al. (2013) argue that ER-GPP coupling should be weaker in eutrophic lakes when ER is substantially lower than GPP. However, Mývatn does not fit this pattern as overall ER during the most productive summer months was 86% of GPP. The unexcepted decoupling of ER and GPP could be due to the high organic content of the lake sediment (itself the source of past primary production; Einarsson et al. 2004), which could fuel high respiration during periods of low GPP.

The relative contribution of benthic and pelagic organisms to primary production is of growing interest given the role of eutrophication in stimulating phytoplankton blooms, often at the expense of benthic production (Vadeboncoeur et al. 2002, 2003, Karlsson et al. 2009, Taranu et al. 2015). Mývatn is generally regarded as being dominated by benthic production (Jónasson 1979, Einarsson et al. 2004), which is typical of shallow lakes that have large amounts light

reaching the benthos (Karlsson et al. 2009). Furthermore, Mývatn's high benthic production sustains very large midge populations (peak larval densities exceeding 500000 m^{-2}) that are an important subsidy to the surrounding landscape (Gratton et al. 2008, 2017, Dreyer et al. 2015). However, water column phycocyanin concentrations were a strong predictor of the maximum potential GPP at high light which suggests that much of the temporal variation in GPP was due pelagic production. This could explain the absence of a relationship between maximum GPP and larval midge density, despite previous experiments showing that midges have large effects on benthic GPP through their engineering of the sediment. Nonetheless, GPP early in each summer during the "clear water phase" is dominated by benthic production (as illustrated by in situ incubations of sediment and water; JS Phillips unpublished data), and this is reflected by early summer NEP being positive despite the low abundance of primary producers in the water column. Together, these results suggest that the relative contributions of benthic and pelagic production vary substantially both within and between years.

Lake ecosystem metabolism is a topic of growing interest, with extensive effort devoted to developing new estimation methods and identifying drivers of spatiotemporal variation. In this study, I developed a new method for modeling temporal variation in ecosystem metabolism responses to two key physical drivers: light and temperature. In Lake Mývatn, the maximum GPP at high light showed 3.5-fold variation through time and was the most important contributor to variation in overall ecosystem metabolism. Variation in the response of metabolism to light was associated with cyanobacterial blooms that shifted relative contributions of benthic and pelagic production. These results illustrate how modeling temporal variation in P-I curves can provide insights into the physical and biological drivers of lake ecosystem metabolism.

References

Ask, J., J. Karlsson, L. Persson, P. Ask, P. Byström, and M. Jansson. 2009. Terrestrial organic matter and light penetration: Effects on bacterial and primary production in lakes. Limnology and Oceanography 54:2034–2040.

Behrenfeld, M. J., and P. G. Falkowski. 1997. A consumer's guide to phytoplankton primary productivity models. Limnology and Oceanography 42:1479–1491.

Bergquist, A. M., and S. R. Carpenter. 1986. Limnetic Herbivory: Effects on Phytoplankton Populations and Primary Production. Ecology 67:1351–1360.

Björnsson, H., and T. Jónsson. 2004. Climate and climatic variability at Lake Myvatn. Aquatic Ecology 38:129–144.

Box, G. E., G. M. Jenkins, and G. C. Reinsel. 1994. Time series analysis: forecasting and control. Third. Prentice Hall, Englewood Cliffs, New Jersey, USA.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell. 2017. Stan: A Probabilistic Programming Language. Journal of Statistical Software 76.

Carpenter, S. R., J. J. Cole, J. R. Hodgson, J. F. Kitchell, M. L. Pace, D. Bade, K. L. Cottingham, T. E. Essington, J. N. Houser, and D. E. Schindler. 2001. Trophic Cascades, Nutrients, and Lake Productivity: Whole-Lake Experiments. Ecological Monographs 71:163–186.

Carpenter, S. R., and J. F. Kitchell. 1984. Plankton Community Structure and Limnetic Primary Production. The American Naturalist 124:159–172.

Chapin, F. S., G. M. Woodwell, J. T. Randerson, E. B. Rastetter, G. M. Lovett, D. D. Baldocchi, D. A. Clark, M. E. Harmon, D. S. Schimel, R. Valentini, C. Wirth, J. D. Aber, J. J. Cole, M. L. Goulden, J. W. Harden, M. Heimann, R. W. Howarth, P. A. Matson, A. D.

McGuire, J. M. Melillo, H. A. Mooney, J. C. Neff, R. A. Houghton, M. L. Pace, M. G. Ryan, S. W. Running, O. E. Sala, W. H. Schlesinger, and E.-D. Schulze. 2006. Reconciling Carbon-cycle Concepts, Terminology, and Methods. Ecosystems 9:1041–1050.

Clark, J. S. 2007. Models for Ecological Data: An Introduction. Princeton University Press, Princeton, New Jersey, USA.

Cole, J. J., and N. F. Caraco. 1998. Atmospheric exchange of carbon dioxide in a low-wind oligotrophic lake measured by the addition of SF6. Limnology and Oceanography 43:647–656.

Cole, J. J., S. R. Carpenter, M. L. Pace, M. C. V. de Bogert, J. L. Kitchell, and J. R. Hodgson. 2006. Differential support of lake food webs by three types of terrestrial organic carbon. Ecology Letters 9:558–568.

Cole, J. J., Y. T. Prairie, N. F. Caraco, W. H. McDowell, L. J. Tranvik, R. G. Striegl, C. M. Duarte, P. Kortelainen, J. A. Downing, J. J. Middelburg, and J. Melack. 2007. Plumbing the Global Carbon Cycle: Integrating Inland Waters into the Terrestrial Carbon Budget. Ecosystems 10:172–185.

Coloso, J. J., J. J. Cole, and M. L. Pace. 2011. Difficulty in Discerning Drivers of Lake Ecosystem Metabolism with High-Frequency Data. Ecosystems 14:935.

Davison, I. R. 1991. Environmental Effects on Algal Photosynthesis: Temperature. Journal of Phycology 27:2–8.

Del Giorgio, P. A., J. J. Cole, N. F. Caraco, and R. H. Peters. 1999. Linking Planktonic Biomass and Metabolism to Net Gas Fluxes in Northern Temperate Lakes. Ecology 80:1422–1431.

Demars, B. O. L., G. M. Gíslason, J. S. Ólafsson, J. R. Manson, N. Friberg, J. M. Hood, J. J. D. Thompson, and T. E. Freitag. 2016. Impact of warming on CO2 emissions from streams countered by aquatic photosynthesis. Nature Geoscience 9:758–761.

Dreyer, J., P. A. Townsend, J. C. Hook III, D. Hoekman, M. J. Vander Zanden, and C. Gratton. 2015. Quantifying aquatic insect deposition from lake to land. Ecology 96:499–509.

Dubinsky, Z., and N. Stambler. 2009. Photoacclimation processes in phytoplankton: mechanisms, consequences, and applications. Aquatic Microbial Ecology 56:163–176.

Edwards, K. F., C. A. Klausmeier, and E. Litchman. 2011. Evidence for a three-way trade-off between nitrogen and phosphorus competitive abilities and cell size in phytoplankton. Ecology 92:2085–2095.

Edwards, K. F., M. K. Thomas, C. A. Klausmeier, and E. Litchman. 2015. Light and growth in marine phytoplankton: allometric, taxonomic, and environmental variation: Light and growth in marine phytoplankton. Limnology and Oceanography 60:540–552.

Edwards, K. F., M. K. Thomas, C. A. Klausmeier, and E. Litchman. 2016. Phytoplankton growth and the interaction of light and temperature: A synthesis at the species and community level: Light-Temperature Interactions. Limnology and Oceanography 61:1232–1244.

Einarsson, Á., G. Stefánsdóttir, H. Jóhannesson, J. S. Ólafsson, G. M. Gíslason, I. Wakana, G. Gudbergsson, and A. Gardarsson. 2004. The ecology of Lake Myvatn and the River Laxá: Variation in space and time. Aquatic Ecology 38:317–348.

Falkowski, P. G. 1984. Physiological responses of phytoplankton to natural light regimes. Journal of Plankton Research 6:295–307.

Fussmann, G. F. 2000. Crossing the Hopf Bifurcation in a Live Predator-Prey System. Science 290:1358–1360.

Gratton, C., J. Donaldson, and M. J. Vander Zanden. 2008. Ecosystem Linkages Between Lakes and the Surrounding Terrestrial Landscape in Northeast Iceland. Ecosystems 11:764–774.

Gratton, C., D. Hoekman, J. Dreyer, and R. D. Jackson. 2017. Increased duration of aquatic resource pulse alters community and ecosystem responses in a subarctic plant community. Ecology 98:2860–2872.

Hanson, P. C., D. L. Bade, S. R. Carpenter, and T. K. Kratz. 2003. Lake metabolism: Relationships with dissolved organic carbon and phosphorus. Limnology and Oceanography 48:1112–1119.

Hanson, P. C., S. R. Carpenter, N. Kimura, C. Wu, S. P. Cornelius, and T. K. Kratz. 2008. Evaluation of metabolism models for free-water dissolved oxygen methods in lakes. Limnology and Oceanography: Methods 6:454–465.

Hashioka, T., M. Vogt, Y. Yamanaka, C. L. Quéré, E. T. Buitenhuis, M. N. Aita, S. Alvain, L. Bopp, T. Hirata, I. Lima, S. Sailley, and S. C. Doney. 2013. Phytoplankton competition during the spring bloom in four plankton functional type models:18.

Henley, W. J. 1993. Measurement and Interpretation of Photosynthetic Light-Response Curves in Algae in the Context of Photoinhibition and Diel Changes. Journal of Phycology 29:729–739.

Herren, C. M., K. C. Webert, M. D. Drake, M. Jake Vander Zanden, Á. Einarsson, A. R. Ives, and C. Gratton. 2017. Positive feedback between chironomids and algae creates net mutualism between benthic primary consumers and producers. Ecology:447–455.

Holgerson, M. A., and P. A. Raymond. 2016. Large contribution to inland water CO2 and CH4 emissions from very small ponds. Nature Geoscience 9:222–226.

Hölker, F., M. J. Vanni, J. J. Kuiper, C. Meile, H.-P. Grossart, P. Stief, R. Adrian, A. Lorke, O. Dellwig, A. Brand, M. Hupfer, W. M. Mooij, G. Nützmann, and J. Lewandowski. 2015. Tube-dwelling invertebrates: tiny ecosystem engineers have large effects in lake ecosystems. Ecological Monographs 85:333–351.

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

Ives, A. R., and V. Dakos. 2012. Detecting dynamical changes in nonlinear time series using locally linear state-space models. Ecosphere 3:art58.

Ives, A. R., Á. Einarsson, V. A. A. Jansen, and A. Gardarsson. 2008. High-amplitude fluctuations and alternative dynamical states of midges in Lake Myvatn. Nature 452:84–87.

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

Jónasson, P. M. 1979. The Lake Mývatn Ecosystem, Iceland. Oikos 32:289–305.

Karlsson, J., P. Byström, J. Ask, P. Ask, L. Persson, and M. Jansson. 2009. Light limitation of nutrient-poor lake ecosystems. Nature 460:506–509.

Lindegaard, C., and P. M. Jónasson. 1979. Abundance, Population Dynamics and Production of Zoobenthos in Lake Mývatn, Iceland. Oikos 32:202–227.

Moore, C. M., D. J. Suggett, A. E. Hickman, Y.-N. Kim, J. F. Tweddle, J. Sharples, R. J. Geider, and P. M. Holligan. 2006. Phytoplankton photoacclimation and photoadaptation in response to environmental gradients in a shelf sea. Limnology and Oceanography 51:936–949.

Myers, J., and J.-R. Graham. 1971. The Photosynthetic Unit in Chlorella Measured by Repetitive Short Flashes. PLANT PHYSIOLOGY 48:282–286.

Odum, H. T. 1956. Primary Production in Flowing Waters. Limnology and Oceanography 1:102–117.

Ólafsson, J. 1979. The Chemistry of Lake Mývatn and River Laxá. Oikos 32:82–112.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and R Core Team. 2018. nlme: Linear and Nonlinear Mixed Effects Models.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Randerson, J. T., F. S. Chapin, J. W. Harden, J. C. Neff, and M. E. Harmon. 2002. Net Ecosystem Production: A Comprehensive Measure of Net Carbon Accumulation by Ecosystems. Ecological Applications 12:937–947.

Raymond, P. A., J. Hartmann, R. Lauerwald, S. Sobek, C. McDonald, M. Hoover, D. Butman, R. Striegl, E. Mayorga, C. Humborg, P. Kortelainen, H. Dürr, M. Meybeck, P. Ciais, and P. Guth. 2013. Global carbon dioxide emissions from inland waters. Nature 503:355–359.

Richardson, D. C., C. C. Carey, D. A. Bruesewitz, and K. C. Weathers. 2017. Intra- and inter-annual variability in metabolism in an oligotrophic lake. Aquatic Sciences 79:319–333.

Sadro, S., C. E. Nelson, and J. M. Melack. 2011. Linking diel patterns in community respiration to bacterioplankton in an oligotrophic high-elevation lake. Limnology and Oceanography 56:540–550.

Schindler, D. E., K. Jankowski, Z. T. A'mar, and G. W. Holtgrieve. 2017. Two-stage metabolism inferred from diel oxygen dynamics in aquatic ecosystems. Ecosphere 8:n/a-n/a.

Schwaderer, A. S., K. Yoshiyama, P. de Tezanos Pinto, N. G. Swenson, C. A. Klausmeier, and E. Litchman. 2011. Eco-evolutionary differences in light utilization traits and distributions of freshwater phytoplankton. Limnology and Oceanography 56:589–598.

Solomon, C. T., D. A. Bruesewitz, D. C. Richardson, K. C. Rose, M. C. Van de Bogert, P. C. Hanson, T. K. Kratz, B. Larget, R. Adrian, B. L. Babin, C.-Y. Chiu, D. P. Hamilton, E. E. Gaiser, S. Hendricks, V. Istvànovics, A. Laas, D. M. O'Donnell, M. L. Pace, E. Ryder, P. A. Staehr, T. Torgersen, M. J. Vanni, K. C. Weathers, and G. Zhu. 2013. Ecosystem respiration: Drivers of daily variability and background respiration in lakes around the globe. Limnology and Oceanography 58:849–866.

Song, C., W. K. Dodds, J. Rüegg, A. Argerich, C. L. Baker, W. B. Bowden, M. M. Douglas, K. J. Farrell, M. B. Flinn, E. A. Garcia, A. M. Helton, T. K. Harms, S. Jia, J. B. Jones, L. E. Koenig, J. S. Kominoski, W. H. McDowell, D. McMaster, S. P. Parker, A. D. Rosemond, C. M. Ruffing, K. R. Sheehan, M. T. Trentman, M. R. Whiles, W. M. Wollheim, and F. Ballantyne. 2018. Continental-scale decrease in net primary productivity in streams due to climate warming. Nature Geoscience 11:415.

Staehr, P. A., D. Bade, M. C. Van de Bogert, G. R. Koch, C. Williamson, P. Hanson, J. J. Cole, and T. Kratz. 2010a. Lake metabolism and the diel oxygen technique: State of the science: Guideline for lake metabolism studies. Limnology and Oceanography: Methods 8:628–644.

Staehr, P. A., K. Sand-Jensen, A. L. Raun, B. Nilsson, and J. Kidmose. 2010b. Drivers of metabolism and net heterotrophy in contrasting lakes. Limnology and Oceanography 55:817–830.

Stan Development Team. 2018. RStan: the R interface to Stan.

Taranu, Z. E., I. Gregory-Eaves, P. R. Leavitt, L. Bunting, T. Buchaca, J. Catalan, I. Domaizon, P. Guilizzoni, A. Lami, S. McGowan, H. Moorhouse, G. Morabito, F. R. Pick, M. A. Stevenson, P. L. Thompson, and R. D. Vinebrooke. 2015. Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. Ecology Letters 18:375–384.

Thimijan, R. W., and R. D. Heins. 1983. Photometric, Radiometric, and Quantum Light Units of Measure A Review of Procedures for Interconversion. HortScience 18:5.

Thorp, J. H., and M. D. Delong. 2002. Dominance of autochthonous autotrophic carbon in food webs of heterotrophic rivers. Oikos 96:543–550.

Vadeboncoeur, Y., E. Jeppesen, M. J. Vander Zanden, H.-H. Schierup, K. Christoffersen, and D. M. Lodge. 2003. From Greenland to green lakes: Cultural eutrophication and the loss of benthic pathways in lakes. Limnology and Oceanography 48:1408–1418.

Vadeboncoeur, Y., M. J. Vander Zanden, and D. M. Lodge. 2002. Putting the Lake Back Together: Reintegrating Benthic Pathways into Lake Food Web Models Lake ecologists tend to focus their research on pelagic energy pathways, but, from algae to fish, benthic organisms form an integral part of lake food webs. BioScience 52:44–54.

Vrede, T., A. Ballantyne, C. Mille-Lindblom, G. Algesten, C. Gudasz, S. Lindahl, and A. K. Brunberg. 2009. Effects of N : P loading ratios on phytoplankton community composition, primary production and N fixation in a eutrophic lake. Freshwater Biology 54:331–344.

Wanninkhof, R. 1992. Relationship between wind speed and gas exchange over the ocean. Journal of Geophysical Research: Oceans 97:7373–7382.

Winslow, L. A., J. A. Zwart, R. D. Batt, H. A. Dugan, R. I. Woolway, J. R. Corman, P. C. Hanson, and J. S. Read. 2016. LakeMetabolizer: an R package for estimating lake metabolism from free-water oxygen using diverse statistical models. Inland Waters 6:622–636.

Yoshiyama, K., J. P. Mellard, E. Litchman, and C. A. Klausmeier. 2009. Phytoplankton Competition for Nutrients and Light in a Stratified Water Column. The American Naturalist 174:190–203.

Zeng, Z., R. M. Nowierski, M. L. Taper, B. Dennis, and W. P. Kemp. 1998. Complex population dynamics in the real world: modeling the influence of time-varying parameters and time lags. Ecology 79:2193–2209.

Acknowledgements

The data used in this manuscript were collected during routine sampling of Mývatn supported by NSF LTREB DEB-1556208 to Anthony R. Ives and the Mývatn Research Station directed by Árni Einarsson. This work was also funded by NSF Graduate Research Fellowship (DGE-1256259) supporting JSP. I would like to thank Jamie Botsch, Stephen Carpenter, Árni Einarsson, Anthony Ives, and Amanda McCormick for feedback on the analyses and manuscript.

Table I: Posterior summaries of fixed (non-time varying) parameters, using posterior medians as point estimates and 16% and 84% quantiles as bounds of 68% uncertainty intervals ('$ui_{68\%}$') to match the coverage of standard errors. The "Fixed" column shows point estimates for the model fit to simulated data, where the metabolism parameters were fixed to their means, implying corresponding standard deviations (σ_β, σ_α, and σ_ρ) of 0.

Parameter	Description	Estimate† [$ui_{68\%}$]	Fixed
γ_β	Scaling of max GPP with temperature	1.06 [1.05, 1.08]	
γ_ρ	Scaling of ER with temperature	1.15 [1.13, 1.18]	
σ_α	Standard deviation of changes in initial slope	0.21 [0.18, 0.25]	0.01
σ_β	Standard deviation of changes in max GPP	0.06 [0.05, 0.07]	0.004
σ_ρ	Standard deviation of changes in baseline ER	0.05 [0.04, 0.08]	0.004
σ_{proc}	Standard deviation of process error	0.099 [0.098, 0.010]	

†The estimates for σ_{proc} are expressed in terms of the standard deviation of the observed DO data. All other parameters of dimensionless.

Table II: Variance partitioning of GPP, ER, and NEP into contributions from light, temperature, and the time-varying parameters of the P-I curve. The CV^2 and scaled sensitivity are dimensionless and can therefore be compared for different variables. The relative contribution to the variance equals the product of the CV^2 and scaled sensitivity.

		CV^2 [$ui_{68\%}$]	Scaled Sensitivity [$ui_{68\%}$]	Relative Contribution to Variance [$ui_{68\%}$]
GPP				
	initial slope	0.38 [0.30, 0.48]	0.09 [0.04, 0.15]	0.04 [0.02, 0.06]
	max GPP	0.09 [0.07, 0.11]	7.95 [6.69, 9.63]	0.71 [0.63, 0.78]
	temperature	0.03†	5.81 [3.72, 7.97]	0.17 [0.11, 0.23]
	light	1.67†	0.05 [0.02, 0.8]	0.08 [0.04, 0.13]
ER				
	baseline ER	0.08 [0.05, 0.11]	5.56 [3.96, 7.44]	0.44 [0.32, 0.58]
	temperature	0.03†	19.10 [14.7, 23.2]	0.56 [0.42, 0.68]
NEP				
	initial slope	0.38 [0.30, 0.48]	0.06 [0.00, 0.13]	0.02 [0.00, 0.05]
	max GPP	0.09 [0.07, 0.11]	10.10 [8.44, 12.2]	0.90 [0.79, 0.99]
	baseline ER	0.08 [0.05, 0.11]	–0.41 [-1.40, 0.373]	-0.03§ [-0.09, 0.04]
	temperature	0.03†	-0.79 [-1.70, 0.13]	-0.02§ [-0.05, 0.00]
	light	1.67†	0.08 [0.04, 0.13]	0.13 [0.07, 0.22]

†The environmental variables are inputs to the model and therefore have no modeled uncertainty in their coefficients of variation.

§Negative contributions to variance in metabolism arise due to covariances with other parameters, resulting in compensatory effects on metabolism.

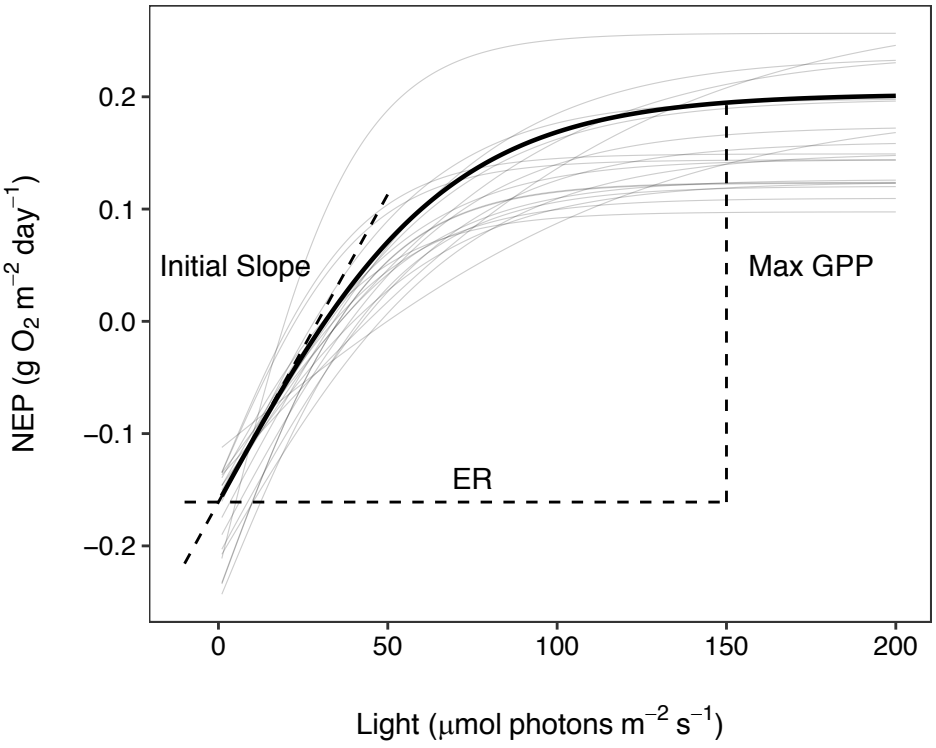

Figure 1: Photosynthesis-irradiance (P-I) curves for ecosystem metabolism. The dashed lines illustrate the contribution of each parameter of the P-I curve. The thin lines show P-I curves for a random subset of days at Lake Mývatn as inferred by the model described in the main text, while the thick line shows the P-I curve using the mean parameter values across this subset of days.

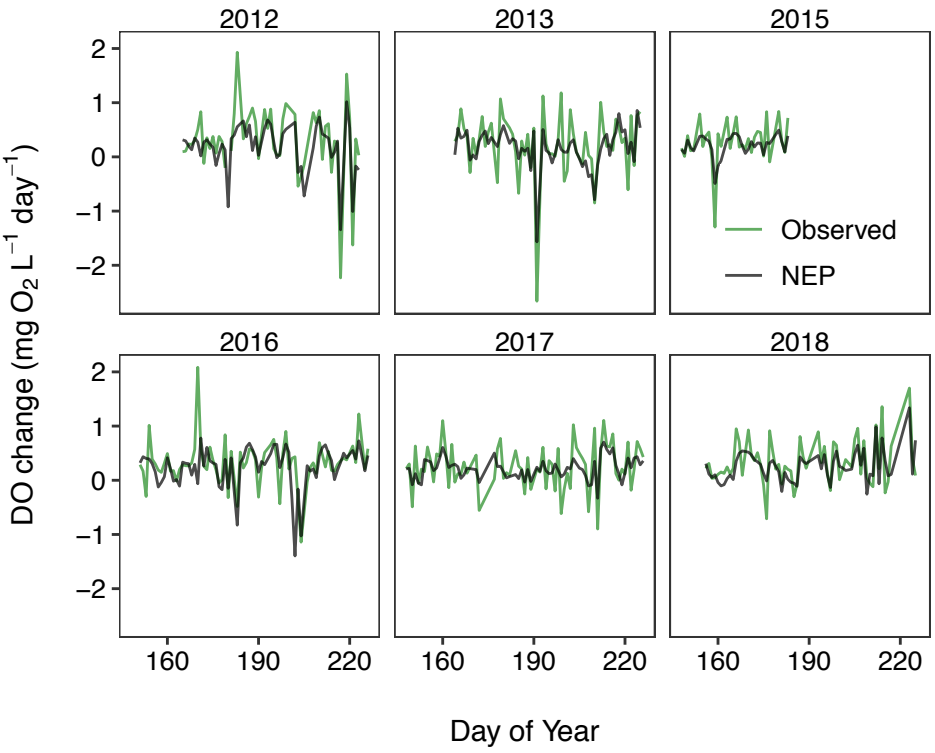

Figure 2: Observed daily change in DO (with flux to atmosphere removed) and the modeled daily NEP, plotted through time. After accounting for the atmospheric flux, the change in DO is either attributable to the explicitly modeled biological flux (i.e. NEP) or process error.

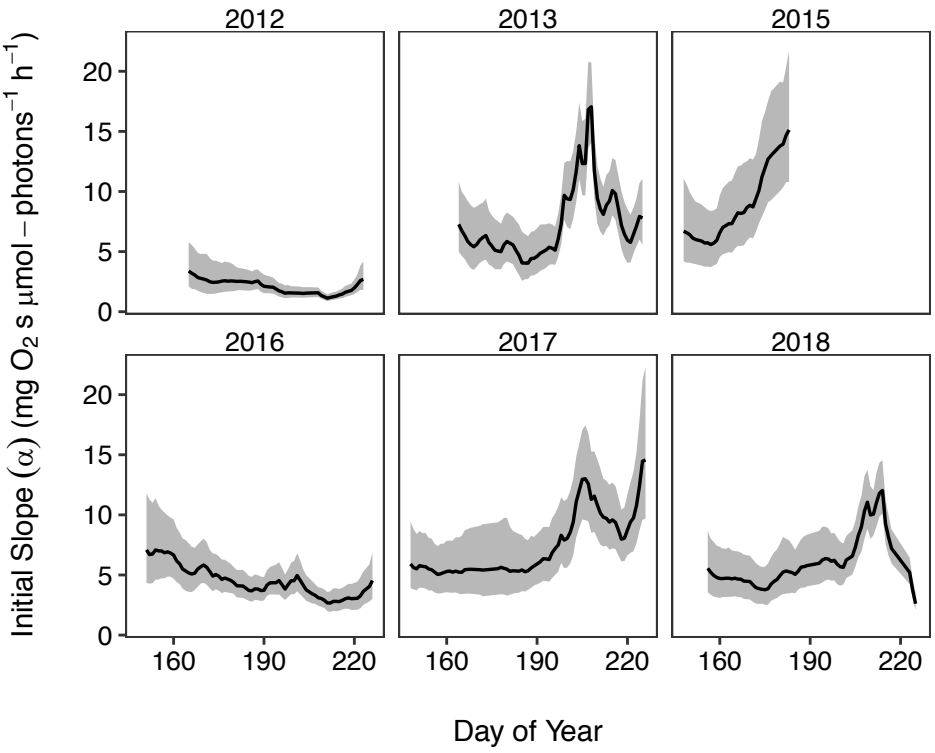

Figure 3: Initial slope of the P-I curve as inferred from the model, plotted through time. The lines are the posterior medians, and the shaded regions are the 68% uncertainty intervals matching the coverage of standard errors.

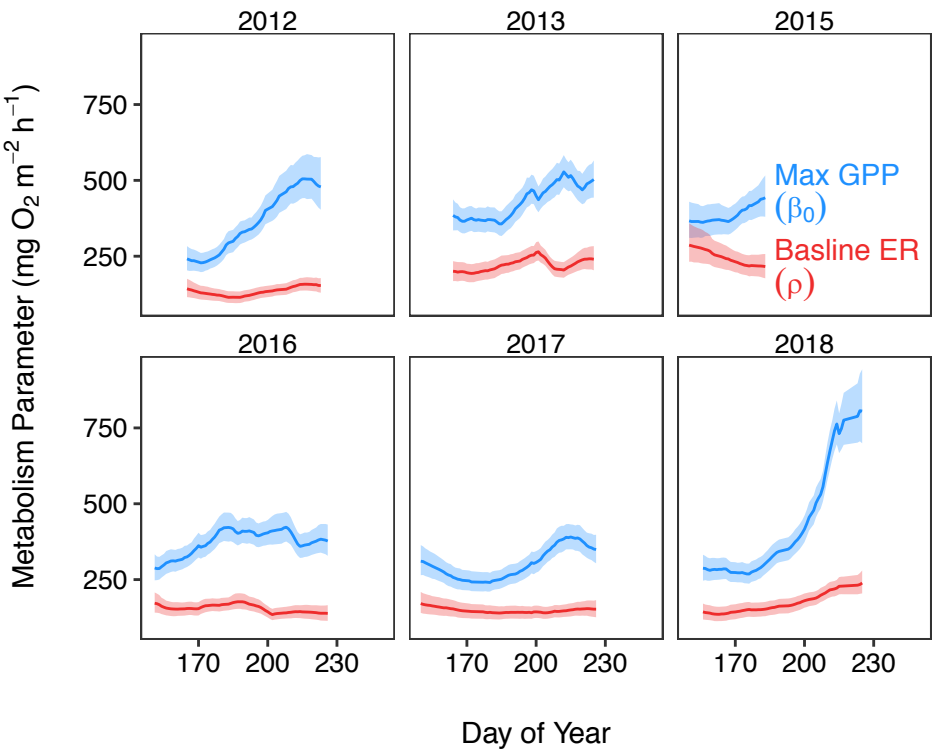

Figure 4: Maximum GPP and baseline ER (at the reference temperature of 12°C) as inferred from the model, plotted through time. The lines are the posterior medians, and the shaded regions are the 68% uncertainty intervals matching the coverage of standard errors.

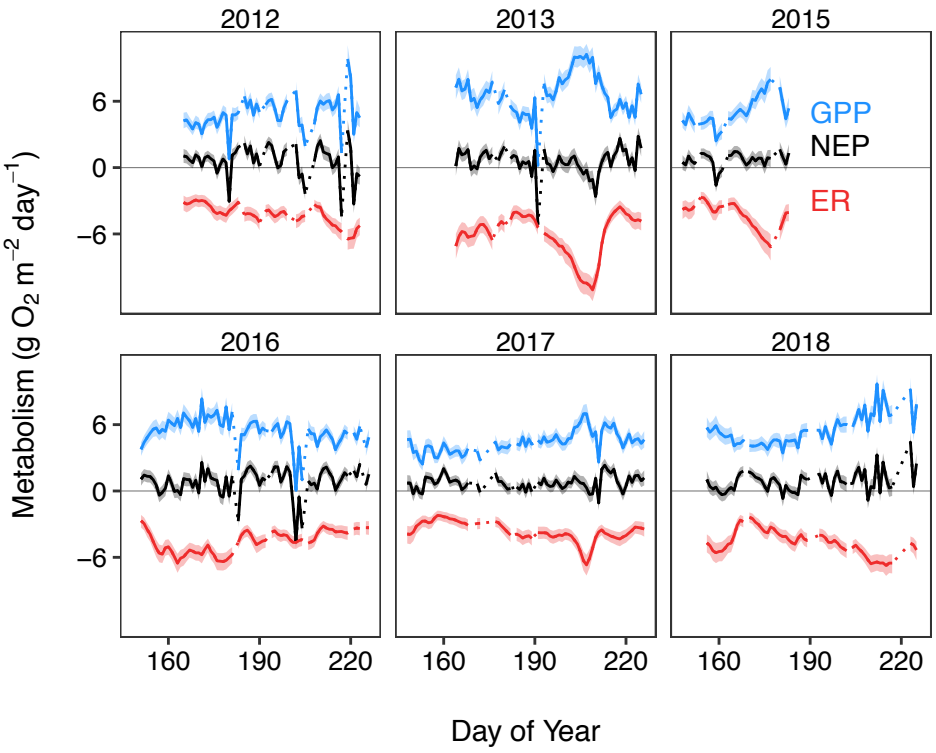

Figure 5: Daily estimates of GPP, ER, and NEP, plotted through time. The lines are the posterior medians, and the shaded regions are the 68% uncertainty intervals matching the coverage of standard errors.

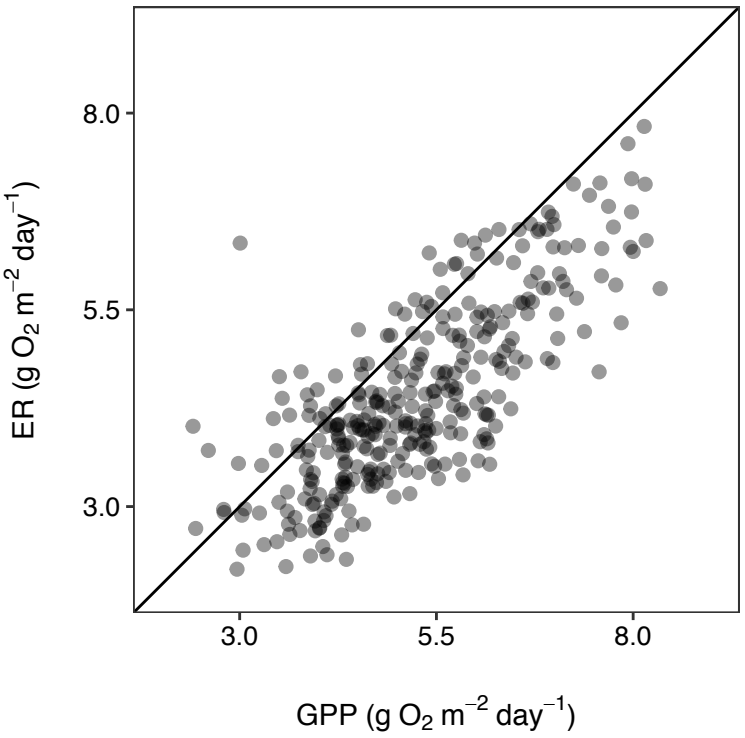

Figure 6: Estimated daily ER plotted against GPP, with the 1-to-1 line. The estimates of ER and GPP include the influence of temperature, which contributes to their correlation ($R = 0.73$)

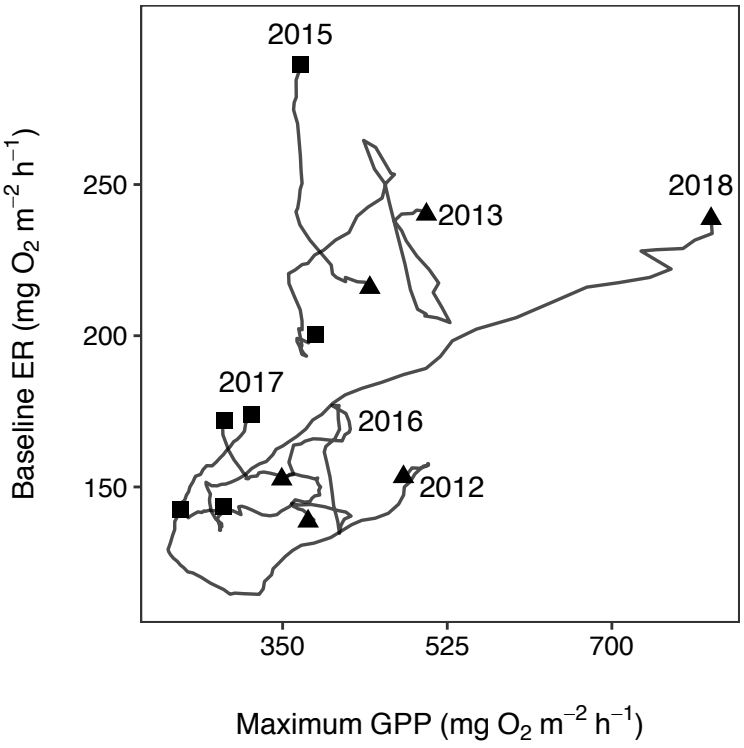

Figure 7: Baseline ER plotted against maximum GPP, with both at the reference temperature of 12°C. Each line represents a separate year, with squares indicating the beginning and triangles the end of each series.

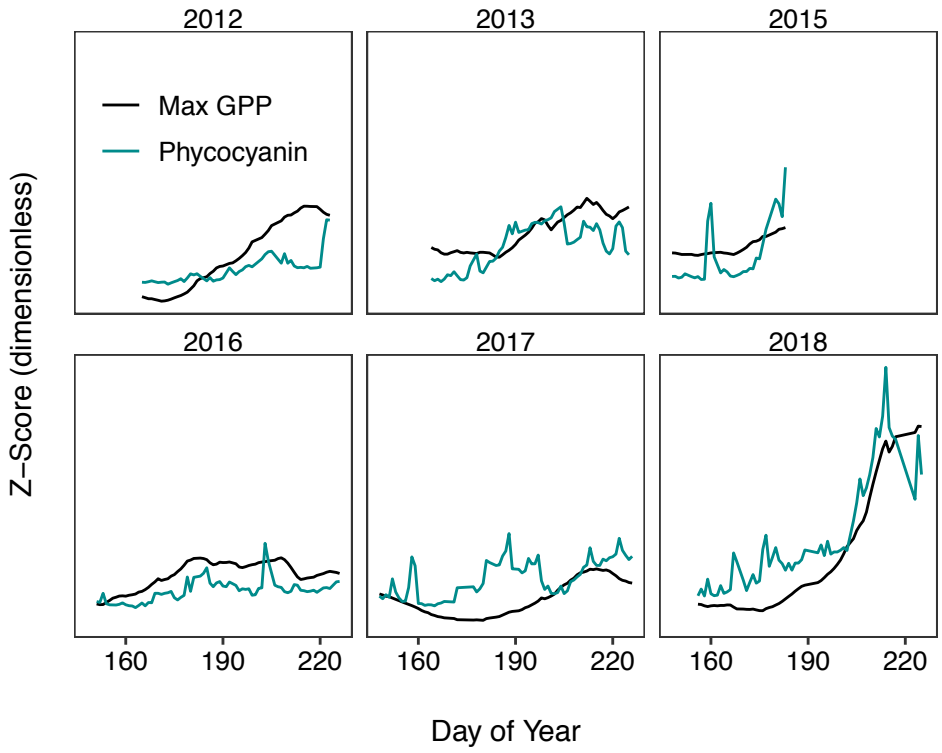

Figure 8: Maximum GPP inferred from the model and phycocyanin concentration (a cyanobacterial pigment), plotted through time. Both variables are shown as z-scores (centered on mean and divided by standard deviation).

Appendices

Appendix B: Supplemental methods and results for Chapter 2

Appendix B: Supplementary methods and results for Chapter 2

Code and data availability

Data and code for data processing and model fitting, diagonstics, and analysis can be found at https://github.com/jsphillips2/sonde_oxygen.

Data preparation

Light

Light readings were taken either every 15 or 30 minutes (depending on the year) with a light/temperature logger (HOBO Pendant, Onset Computer Corporation, MA, USA) attached to the top of the buoy deploying the sonde; I averaged these data to an hourly scale for further analysis. HOBO loggers measure visible light intensity in lux, which is closely related to the spectral range (400-700nm) used to calculate photosynthetically active radiation (PAR; measured in μmol-photons m-2 s-1) (Thimijan and Heins 1983). I also obtained downwelling solar irradiance from the local whether station (in Watts). Following Thimijan and Heins 1983, visible light intensity and downwelling irradiance can be scalarly converted to photon flux density based on an approximate standard spectrum for sunlight. Specifically, the visible intensity is divided by 54 and downwelling irradiance is multiplied by 4.57.

When approximately converted to PAR, visible light intensity and downwelling irradiance were strongly correlated (r=0.79 on natural scale and r=0.89 on log-log scale; Figure B1). However, when hourly values are plotted through time it appears that downwelling irradiance has higher peaks than the visible intensity (Figure B2), which makes sense as the downwelling irradiance is calculated using a wider spectral range that visible intensity.

Sensor drift

To evaluate whether the sonde DO readings drifted through time, I compared them to measurements and the same depth (0.5m) and location made with handheld probes (ProODO, YSI Co., OH, USA) during routine sampling of the central monitoring station. The sonde and handheld probe measurements were strongly correlated (r=0.92; Figure B3). Furthermore, the DO time-series measured by the sonde failed to display linear trends that were consistent between years (Figure B4). Therefore, there did not appear to be any major drift in the sensor readings.

Model fitting

Scaling the model

For fitting the model, I z-scored (centered on the mean and divided by the standard deviation) the observed DO concentrations and scaled the model accordingly. I then nondimensionalized the scaled model to remove the scale parameters (mixing depth and mean/standard deviation for the observed DO). This put the observed values and model parameters on unit scale, which improved the computational efficiency and allowed for the use for weakly informative priors on unit scale. I then back-transformed the scaled parameters to recover their values on the original scale.

The model for DO dynamics (equations 9 and 10 in the Chapter 2) is:

$$DO_{t+1} = DO_t + \frac{\Delta h}{z_{mix}}\left(NEP_t + k_t(DO_t^{eq} - DO_t)\right) + \epsilon_t^{proc}$$

where DO_t is the DO concentration [$mg\ O_2\ m^{-3}$], DO_t^{eq} is the DO concentration at saturation [$mg\ O_2\ m^{-3}$], Δh = 1 hour, NEP_t is the net ecosystem production [$mg\ O_2\ m^{-2}\ h^{-1}$], k [$m\ h^{-1}$] is the rate of oxygen exchange, z_{mix} is the mixing depth [$3.3m$], and ϵ_{proc} is the processes error [$mg\ O_2\ m^{-3}$]. For simplicity, only a single time index t is used, rather than indexing by hour and day as in the main text. For clarity, DO is substituted with y and NEP with ϕ, yielding

$$y_{t+1} = y_t + \frac{\Delta h}{z_{mix}}\left(\phi_t + k_t(y_t^{eq} - y_t)\right) + \epsilon_t^{proc}.$$

Z-scoring the observed DO values by the mean (μ) and standard deviation (τ) of the full set of DO observations implies the following scaling of the model:

$$\frac{y_{t+1} - \mu}{\tau} = \frac{1}{\tau}\left(y_t - \mu + \frac{\Delta h}{z_{mix}}\left(\phi_t + k_t(y_t^{eq} - y_t)\right) + \epsilon_t^{proc}\right)$$

which can be expanded as

$$\frac{y_{t+1} - \mu}{\tau} = \frac{1}{\tau}y_t - \frac{1}{\tau}\mu + \frac{1}{\tau}\frac{\Delta h}{z_{mix}}\phi_t + \frac{1}{\tau}\frac{\Delta h}{z_{mix}}k_t(y_t^{eq} - y_t) + \frac{1}{\tau}\epsilon_t^{proc}.$$

Rearranging to express all "y" terms to resemble the left-hand side yields

$$\frac{y_{t+1} - \mu}{\tau} = \frac{y_t - \mu}{\tau} + \frac{1}{\tau}\frac{\Delta h}{z_{mix}}\phi_t + \frac{\Delta h}{z_{mix}}k_t\left(\left(\frac{y_t^{eq} - \mu}{\tau} + \frac{\mu}{\tau}\right) - \left(\frac{y_t - \mu}{\tau} + \frac{\mu}{\tau}\right)\right) + \frac{1}{\tau}\epsilon_t^{proc}$$

which simplifies to

$$\frac{y_{t+1}-\mu}{\tau}=\frac{y_t-\mu}{\tau}+\frac{1}{\tau}\frac{\Delta h}{z_{mix}}\phi_t+\frac{\Delta h}{z_{mix}}k_t\left(\frac{y_t^{eq}-\mu}{\tau}-\frac{y_t-\mu}{\tau}\right)+\frac{1}{\tau}\epsilon_t^{proc}.$$

Substituting with scaled variables yields

$$x_{t+1}=x_t+\hat{\phi}_t+\hat{k}_t\,(x_t^{eq}-x_t)+\hat{\epsilon}_t^{proc}$$

where

$$x_t=\frac{y_t-\mu}{\tau}$$

$$x_t^{eq}=\frac{y_t^{eq}-\mu}{\tau}$$

$$\hat{\phi}_t=\frac{1}{\tau}\frac{\Delta h}{z_{mix}}\phi_t$$

$$\hat{k}_t=\frac{\Delta h}{z_{mix}}k_t$$

$$\hat{\epsilon}_t^{proc}=\frac{1}{\tau}\epsilon_t^{proc}$$

I fit the nondimensionalized version of the model to the data, resulting in estimates for the scaled parameters (e.g. $\hat{\phi}_t$). The parameter values on the original scale could be obtained by back-transforming the scaled values. For example, $\phi_t = \tau \frac{z_{mix}}{\Delta h} \hat{\phi}_t$.

NEP (ϕ_t) is the difference between GPP and ER. Therefore, the scaling of NEP implies that GPP (denoted χ) and ER (denoted κ) are scaled as

The original equation for χ (main text equation 3) is

$$\chi_t = \beta_d \tanh\left(\frac{\alpha_t}{\beta_t} L_t\right)$$

where β_t is the maximum rate of GPP [$mg\ O_2\ m^{-3}\ h^{-1}$] on a particular day and α [$mg\ O_2\ s\ \mu mol\ photons^{-1}\ m^{-1}\ h^{-1}$] is the initial slope of the P-I at low light. This is scaled as

$$\hat{\chi}_t = \frac{1}{\tau}\frac{\Delta h}{z_{mix}}\beta_t \tanh\left(\frac{\frac{\lambda}{\tau}\frac{\Delta h}{z_{mix}}\alpha_t \frac{L_t}{\lambda}}{\frac{1}{\tau}\frac{\Delta h}{z_{mix}}\beta_t}\right)$$

where λ is the mean light level (so that the initial slope can be expressed on unit scale). This simplifies to

$$\hat{\chi}_t = \hat{\beta}_t \tanh\left(\frac{\hat{\alpha}_t}{\hat{\beta}_t}\hat{L}_t\right)$$

where

$$\hat{\beta}_t = \frac{1}{\tau}\frac{\Delta h}{z_{mix}}\beta_t$$

$$\hat{\alpha}_t = \frac{\lambda}{\tau}\frac{\Delta h}{z_{mix}}\ \alpha_t$$

$$\hat{L}_t = \frac{L_t}{\lambda}.$$

This scaling can be applied by direct analogy to ER and to the time varying components of β_t, α_t, and ρ_t (main text eqns 5, 6, and 7).

Priors

I fit the model using weakly informative priors based on normal distributions (see http://mc-stan.org/users/documentation/case-studies/weakly_informative_shapes.html). Because the parameters were fit on unit scale, I generally used priors with standard deviations of 1 and means of 0. However, because the parameters describing the scaling of GPP and ER with temperature (γ_β and γ_ρ) had lower bounds of 1, I used corresponding priors with means of 1. For those parameters with lower bounds (1 for γ_β and γ_ρ and 0 for all standard deviations), I used normal priors truncated at their respective lower bounds.

Estimating observation error

Attempting to estimate the standard deviations for observation error (σ_{obs}) and process error (σ_{proc}) independently resulted in very inefficient sampling of the posterior distribution as

σ_{obs} slowly converged to 0. This can be seen by examining the parameter trace plots from the MCMC (Figure B5). While σ_{proc} showed good mixing across all four chains, the chains for σ_{proc} did not mix, indicating a lack of convergence. However, the overall magnitude of σ_{obs} was close to 0 and the lack of convergence is likely due to difficulties associated with estimating the parameter when near its boundary. Therefore, I refit the model fixing σ_{obs} to near 0 (1% of the observed standard deviation in obseved DO, so that I could use the same model specification in Stan). This resulted in nearly identical estimates for most parameters except for σ_{proc}, which was nontheless very similar (Figure B6). It is not surprising that observation error was so low, given the smooth appearance of the DO time series (Figure B7).

Variance partitioning

I partitioned variation in NEP, GPP, and ER resulting from variation in physical and biological parameters by applying the delta method to Chapter 2 equations 2, 3, and 7, respectively. This method gives the first-order approximation of the variance in a function F that depends on a vector $\mathbf{\Theta_{d,h}}$ of time-varying parameters $\theta^1_{d,h}, \theta^2_{d,h}, \ldots \theta^K_{d,h}$ as

$$\mathrm{Var}\left[F\left(\mathbf{\Theta_{d,h}}\right)\right] \approx \sum_i \left(\frac{\partial \hat{F}}{\partial \hat{\theta}^i}\right)^2 \sigma^2_{\theta^i} + \sum_{i|i\neq j} \sum_{j|j\neq i} \frac{\partial \hat{F}}{\partial \hat{\theta}^i} \frac{\partial \hat{F}}{\partial \hat{\theta}^j} \sigma_{\theta^i,\theta^j}$$

where $\hat{\theta}^i$ is the mean value of $\theta^i_{d,h}$ across all d and h, $\hat{F}$ is the function evaluated at the vector of mean values for each parameter vector $\widehat{\mathbf{\Theta}}$, $\frac{\partial \hat{F}}{\partial \hat{\theta}^i}$ is the partial derivative of $\hat{F}$ with respect to $\hat{\theta}^i$, $\sigma^2_{\theta^i}$ is the variance in $\theta^i_{d,h}$, and $\sigma_{\theta^i,\theta^j}$ is the covariance between $\theta^i_{d,h}$ and $\theta^j_{d,h}$. The relative

contribution of a given parameter $\theta_{d,h}^{i=I}$ is calculated as the sum of all terms in equation where $i = I$, divided by the total variance $\mathrm{Var}[F(\boldsymbol{\Theta}_{\boldsymbol{d,h}})]$. Note that the contribution of a parameter will be reduced by its negative covariances with other parameters and can even be negative if the contribution of its negative covariance terms exceeds the contribution of its (necessarily positive) variance term. The advantage of an explicit partitioning based on the model equations, as opposed to a more conventional "statistical" variance partitioning (e.g. ANOVA), is that the explicit partitioning accounts for nonlinearities and interactions between explanatory variables through both derivatives and covariances.

Generically, the contribution of a variable θ^i to the variance in F is determined by (1) the sensitivity of F to variation in $\theta_{d,h}^{i}$ (i.e. the partial derivatives) and (2) the variance of $\theta_{d,h}^{i}$. This can be illustrated by considering the contribution of a single parameter $\theta_{d,h}^{i=I}$ relative to the total variance:

$$\left[\left(\frac{\partial \hat{F}}{\partial \hat{\theta}^I}\right)^2 \sigma_{\theta^I}^2 + \sum_{j|j\neq I} \frac{\partial \hat{F}}{\partial \hat{\theta}^I}\frac{\partial \hat{F}}{\partial \hat{\theta}^j} r_{\theta^I,\theta^j}\sigma_{\theta^I}\sigma_{\theta^j}\right] \Big/ \mathrm{Var}\left[F\left(\boldsymbol{\Theta}_{\boldsymbol{d,h}}\right)\right]$$

where the covariance between a given pair of variables θ^I and θ^j is expressed in terms of their correlation (r_{θ^I,θ^j}) and respective standard deviations (σ_{θ^I} and σ_{θ^j}). Factoring $\sigma_{\theta^I}^2$ and expressing it in terms of the coefficient of variation (CV_{θ^I}) yields

$$\mathrm{CV}_{\theta^I}^2\left(\hat{\theta}^I\right)^2 \left[\left(\frac{\partial \hat{F}}{\partial \hat{\theta}^I}\right)^2 + \sum_{j|j\neq I} \frac{\partial \hat{F}}{\partial \hat{\theta}^I}\frac{\partial \hat{F}}{\partial \hat{\theta}^j} r_{\theta^I,\theta^j}\frac{\sigma_{\theta^j}}{\sigma_{\theta^I}}\right] \Big/ \mathrm{Var}\left[F\left(\boldsymbol{\Theta}_{\boldsymbol{d,h}}\right)\right]$$

which can be simplified by expressing it as the product of $\mathrm{CV}_{\theta^I}^2$ and the "scaled sensitivity" ψ_{θ^I}:

$$\mathrm{CV}_{\theta^I}^2 \psi_{\theta^I}$$

Both $\mathrm{CV}_{\theta^I}^2$ and ψ_{θ^I} are dimensionless and can be compared for different variables. $\mathrm{CV}_{\theta^I}^2$ quantifies the overall variation in $\theta_{d,h}^I$, while ψ_{θ^I} quantifies how sensitive F is to variation in $\theta_{d,h}^I$ given its correlations with other variables.

Simulations

Methods

A main objective of the analysis was to infer the temporal variation in the parameters of the P-I curve (β_d^0, α_d, and rho_d), with the magnitude of variation characterized by the standard deviations of the respective stochastic processes (σ_β, σ_α, and σ_ρ). To explore the capacity of the model to distinguish real changes in these parameters from noise, I fit the model to simulated data with β_d^0, α_d, and rho_d fixed to their mean values for each year as inferred by the fit of the model to real data. The simulated data implied that σ_β, σ_α, and σ_ρ had "true" values of 0; therefore, the model fit to these data should produce estimates shifted toward 0 relative to the estimates for the actual data.

The goal was to validate the model as fit to the particular data used in this analysis. Therefore, I constructed the simulated data using the observed environmental data and the particular sequence of process errors (i.e. ϵ_d^{proc}) as inferred from the model fit to the original data. The observed DO for each year had gaps due to missing data, which resulted in many contiguous time-series within each year. For the simulated data, I retained this structure as it

could have implications for the efficacy of the model fitting. Initial values for the first time-series in each year were set to the corresponding value in the observed data. However, for the initial values of all subsequent time-series within a year I projected the model as if there were no gap in the data (such that each year would have a single contiguous stretch of observations) and then inserted gaps into the simulated time-series to match the structure of the original data. This is was necessary as using the observed values to initial each time-series would lead to substantial differences between the initial value and the equilibrium determined by metabolism rates when fixed to their means. The real and simulated data appear quite different, indicating that the temporal variation in the ecosystem metabolism parameters has large effects on the DO dynamics (Figure B8).

Results

As discussed in the main text, σ_β, σ_α, and σ_ρ were all clearly different from 0 when estimated from the real data. In contrast, the variance estimates for the simulated data were all strongly right skewed, with posterior densities concentrated near 0 (Figure B9). This suggests that the temporal variation in β_d^0, α_d, and rho_d inferred by the model reflected real changes in the underlying P-I curve.

Surface vs. average water column PAR

Methods

Because primary production occurs throughout the water column and in the benthos, I fit the model using the average water column light, calculated as

$$\overline{L_{d,h}(z)} = \frac{1}{z_{\max}} \int_0^{z_{\max}} L_{d,h}^0 \; e^{-c_{d,h}\, z} \; dz = \frac{L_{d,h}^0 (1 - e^{-c_d, h\, z_{\max}})}{z_{\max}\, c_{d,h}}$$

where z $[m]$ is the vertical position in the water column, z_{max} is the water column depth, $L_{d,h}^0$ is the PAR at the water surface ($z = 0$), and $c_{d,h}$ is the light attenuation coefficient (we used c rather than the standard k_d to avoid notational confusion with gas exhange constants used below). For notational simplicity, we henceforth denote $\overline{L_{d,h}(z)}$ as $L_{d,h}$. To estimate $c_{d,h}$, I regressed observe light attenuation calculated from weekly measurements of PAR at 0.5m depth intervals against turbidity and used this regression (y = 0.43 + 0.059 × turbidity; df = 34; R^2 = 0.66) to predict $c_{d,h}$ from turbidity for each time point in the sonde data.

The estimates $\overline{L_{d,h}(z)}$ have some uncertainty to do the inference of water clarity from turbidity. Furthermore, it is common in other studies of ecosystem metabolism to use light as measured near the water surface. To evaluate the consequences of using $\overline{L_{d,h}(z)}$, I refit the model to surface light ($L_{d,h}^0$) and compared the estimates of the time-varying parameters to those for the original model fit.

Results

The estimates of baseline ER and maximum GPP were almost identical for both surface and water column light, as they should be given that they characterize metabolism when light is 0 (i.e. when surface and water column light are the same) and when light is saturating (i.e. when differences between surface and water column later are irrelevant) (Figure B10). The average value of the initial slope of the P-I curve was substantially lower for the model fit using surface light, which again was as expected since the surface light values were always higher than the

water column values (Figure B11). However, the temporal patterns were very similar, with the largest differences being relatively high values in 2015 and 2018 when using the water column light. These two years had substantial declines in water column light due to reduced water clarity associated with cyanobacterial blooms.

Drivers of Maximum GPP

Maximum GPP at high light β_d^0 was quite variable, and it was the dominant contributor to variation in overall GPP and NEP. In the main text, I report results from linear models regressing β_d^0 against daily mean phycocyanin (as a measure of cyanobacterial abundance) and larval midge abundance. Phycocyanin data were available for all days for which there were estimates of β_d^0, while midge data were only available approximately weekly. Therefore, I fit two models: (1) including phycocyanin for all days, and (2) including both phycocyanin and midge abundance for those days when midge samples were collected. In the main text I show a figure relating β_d^0 against phycocyanin for all days. Here I show the corresponding figure for those days where midges were samples (Figure B12).

References

Thimijan, R. W., and R. D. Heins. 1983. Photometric, Radiometric, and Quantum Light Units of Measure A Review of Procedures for Interconversion. HortScience 18:5.

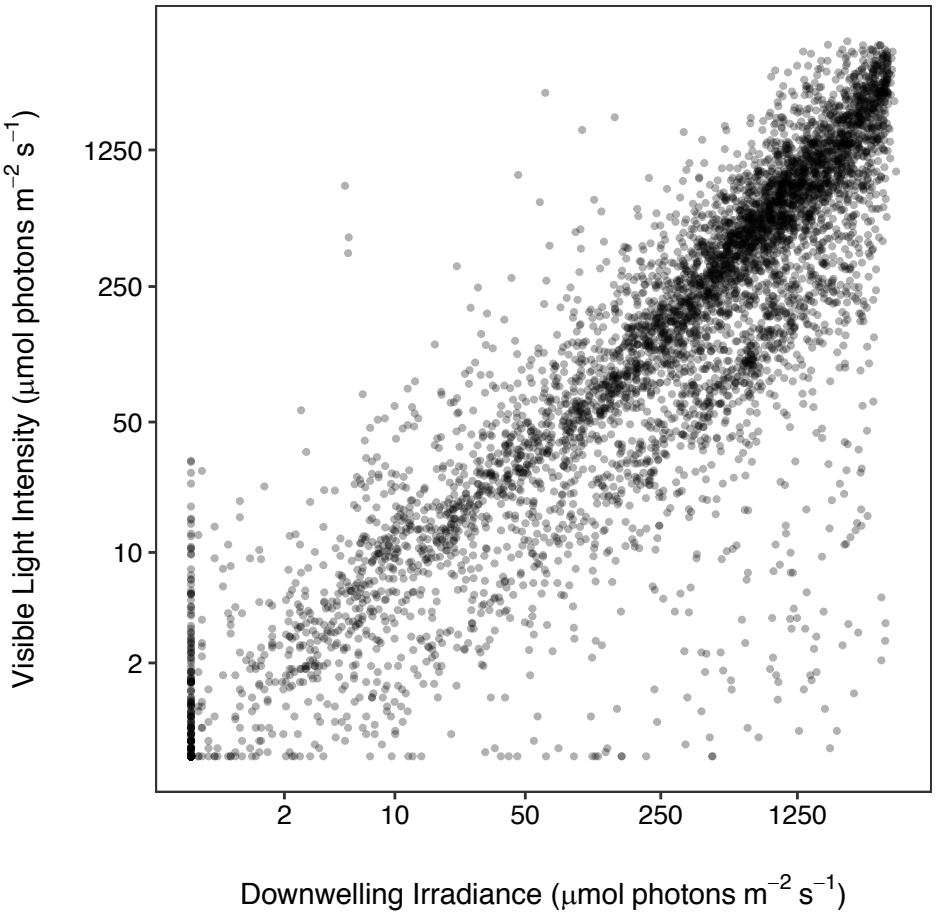

Figure B1. Visible light intensity measured by HOBO logger and downwelling irradiance from weather station, scalarly converted to approximate PAR equivalent.

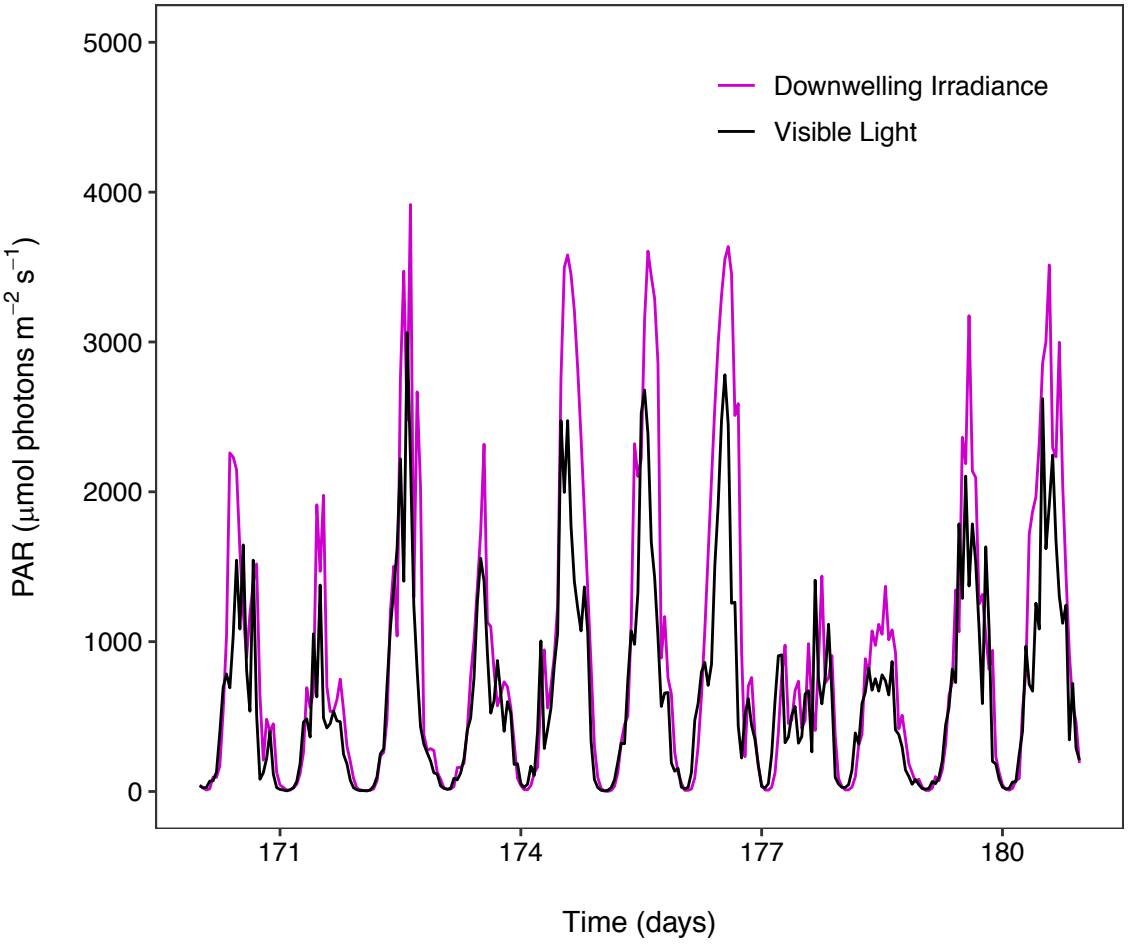

Figure B2. PAR equivalent for visible light intensity and downwelling irradiance plotted by hour for 10 days in 2013

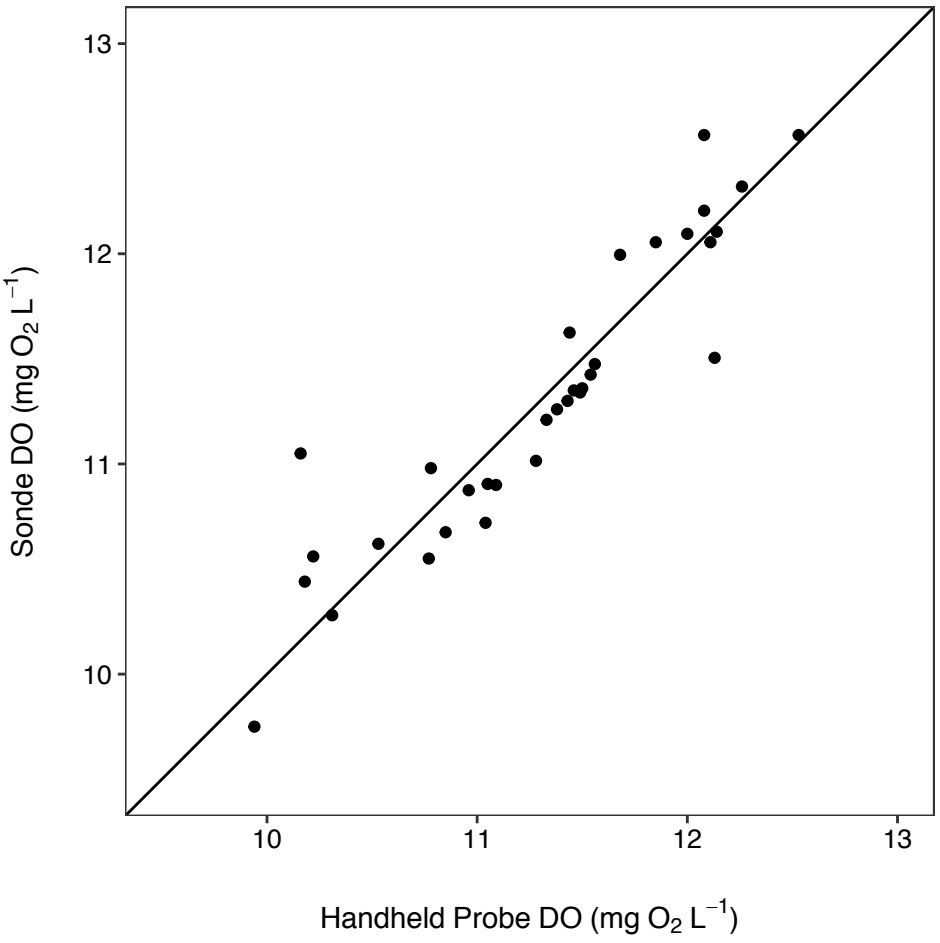

Figure B3. DO measured from the sonde and handheld probes, with the 1-to-1 line.

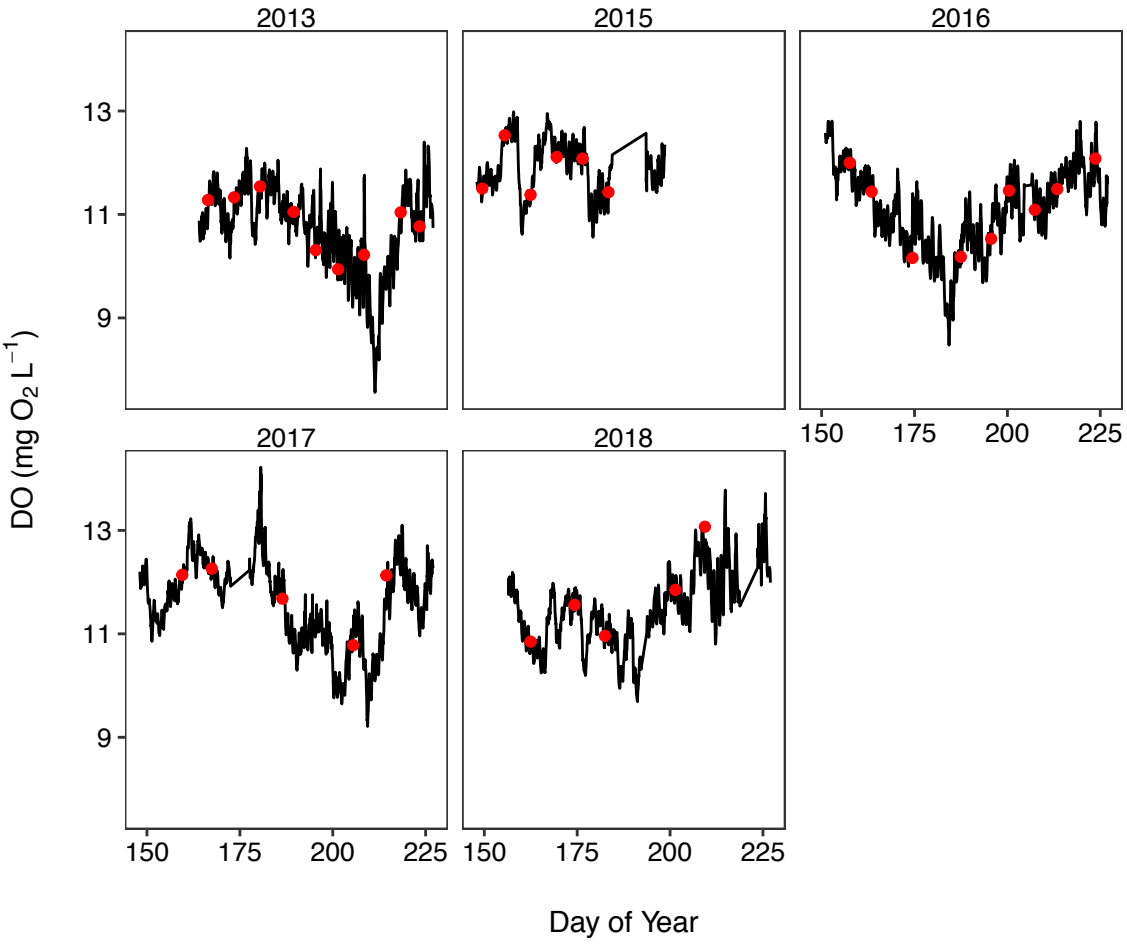

Figure B4. Hourly DO measurements plotted through time. The lines are the sonde measurements, while the points are the handheld probe measurements.

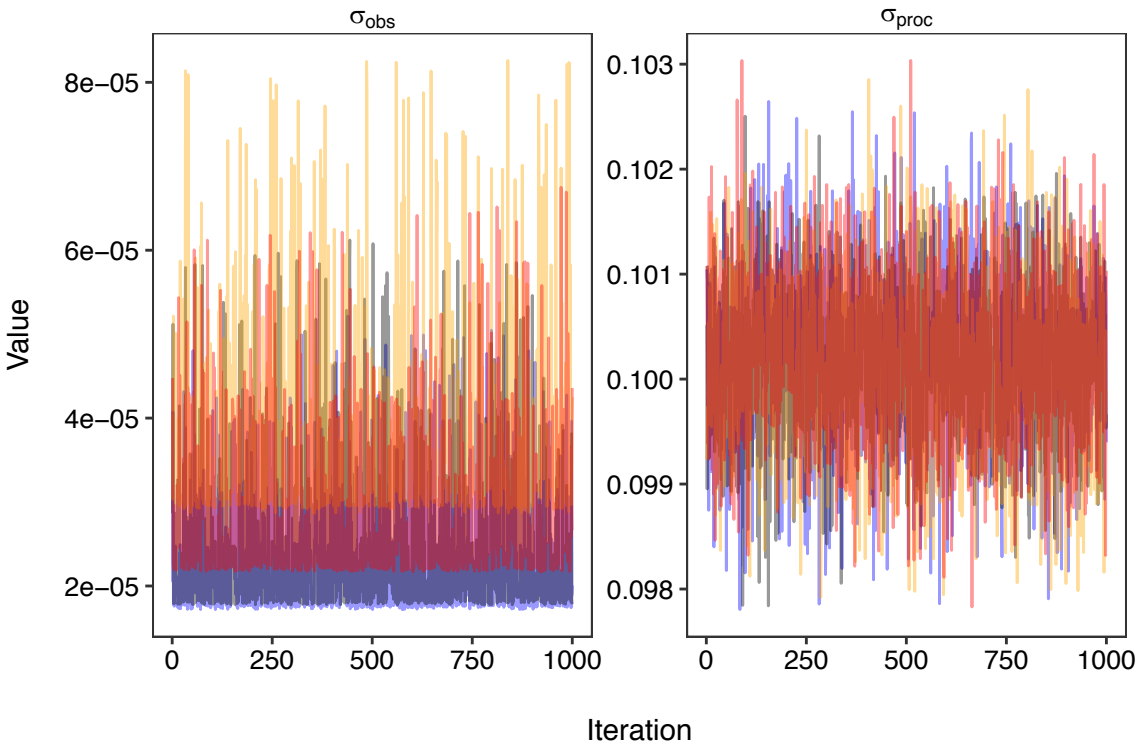

Figure B5. Trace plots for standard deviations of observation and process error from the MCMC. Colors indicate separately initialized chains (note the different y-axis scales).

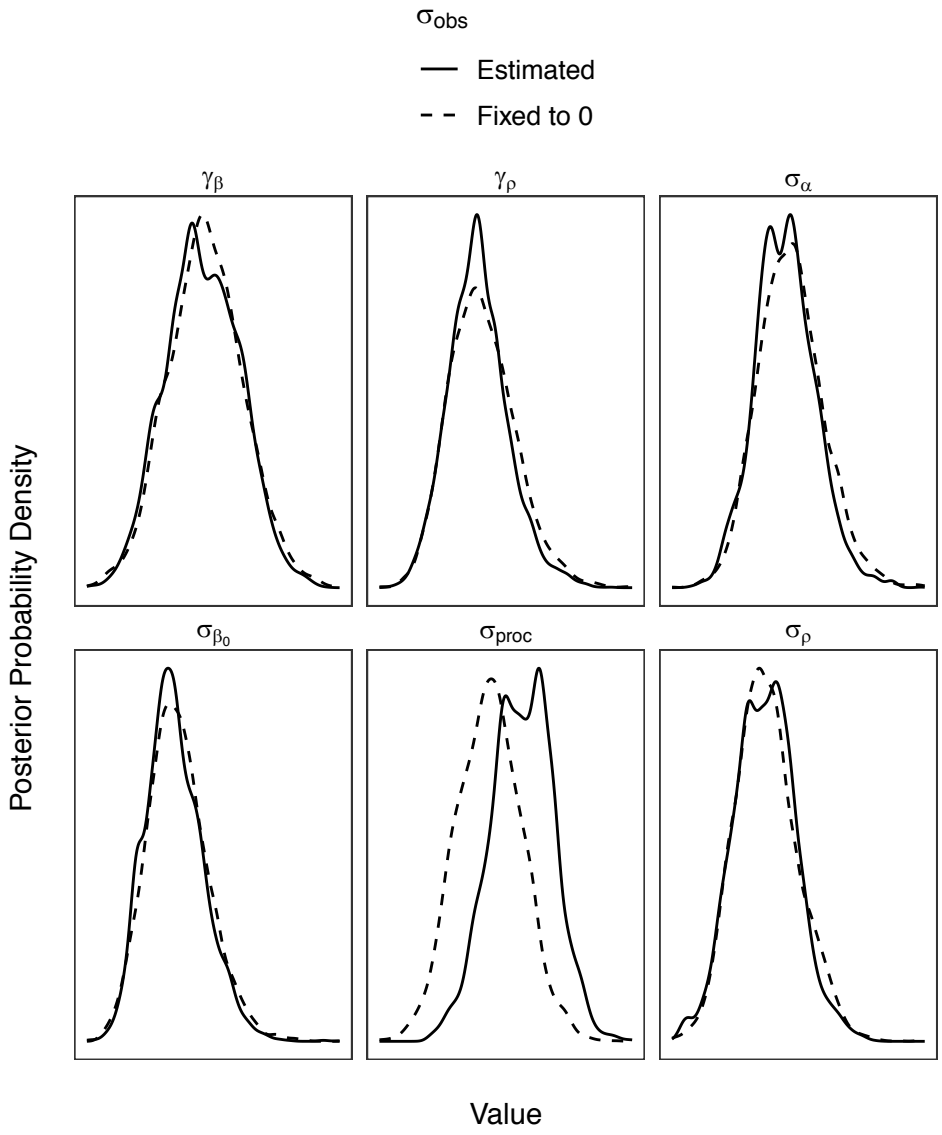

Figure B6. Posterior distributions for parameter estimates, with the standard deviation of observation error either estimated from data or fixed to 0.

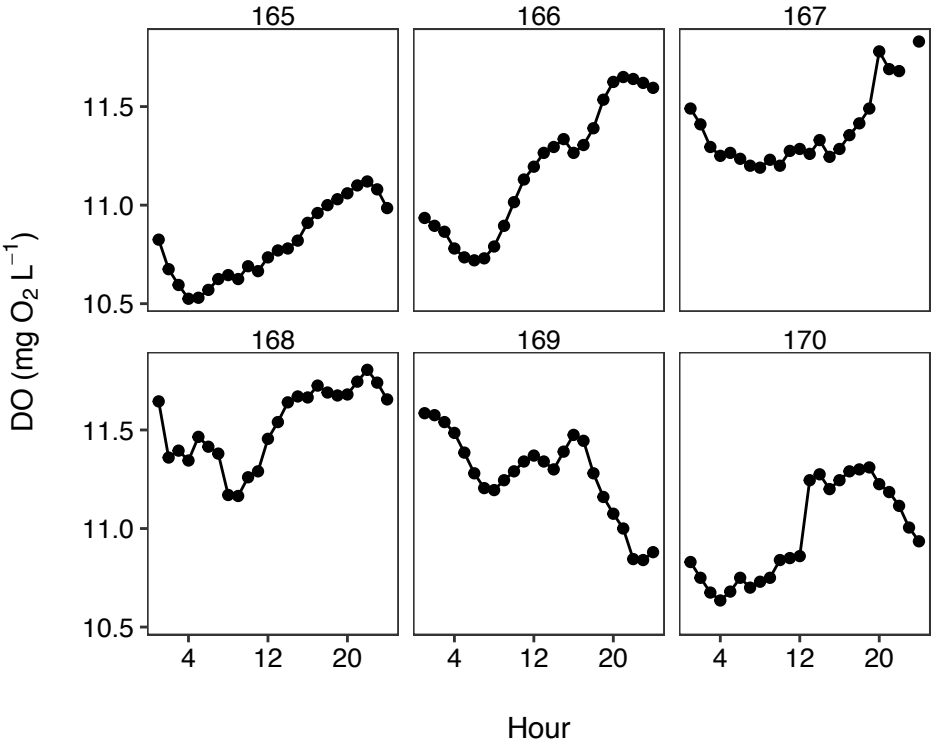

Figure B7. Observed DO time-series for first six days in 2012.

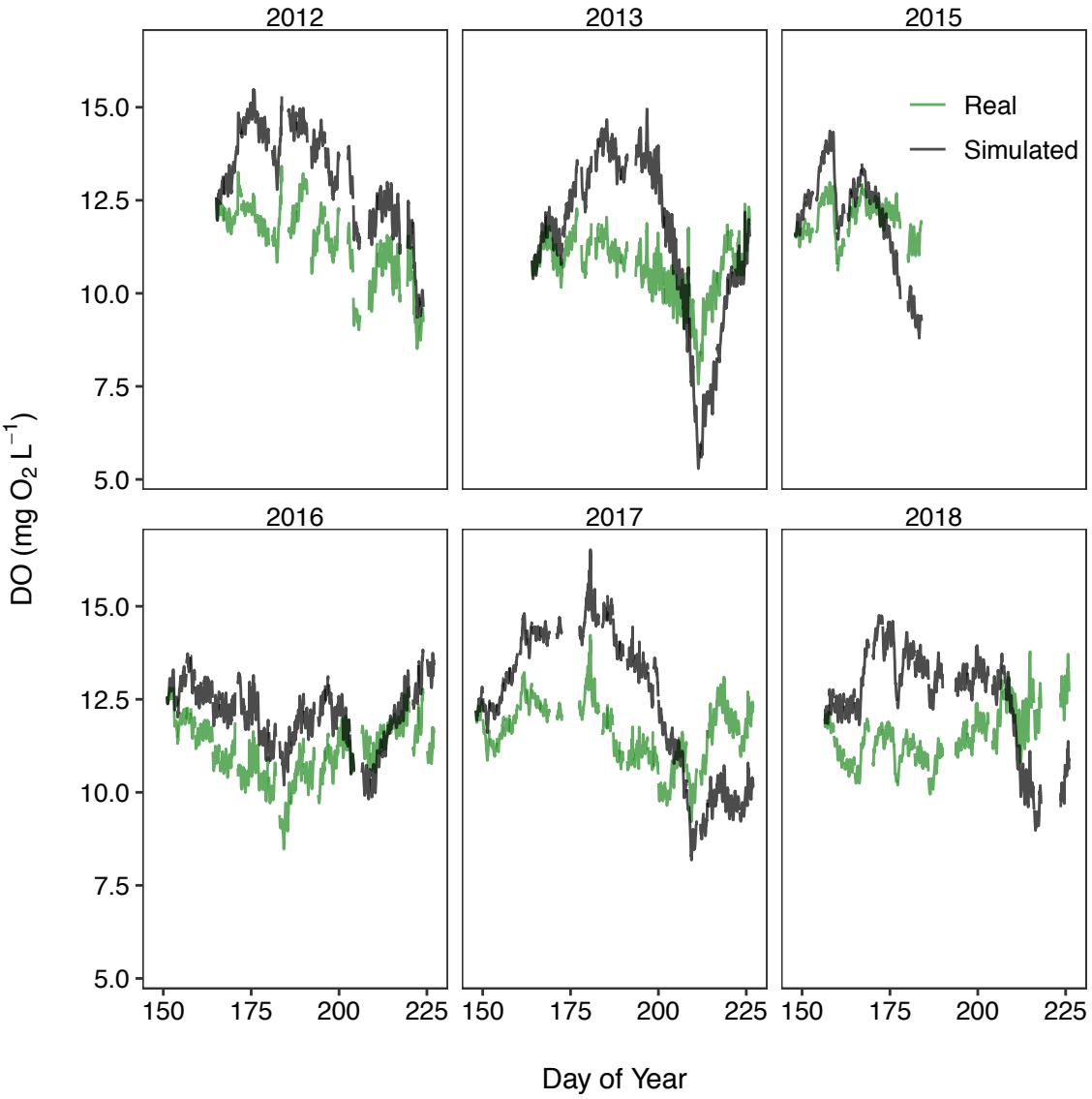

Figure B8. Time-series for real and simualted DO.

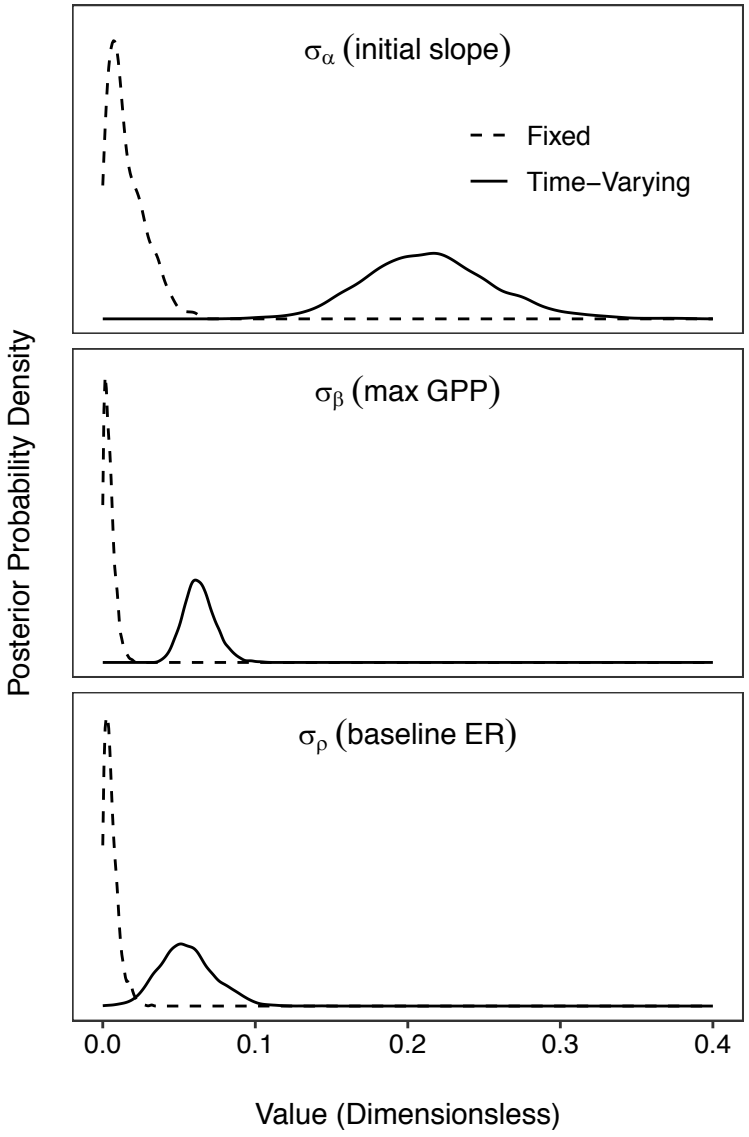

Figure B9. Posterior distributions for standard deviations of variation in ecosystem metabolism rates for real and simulated data.

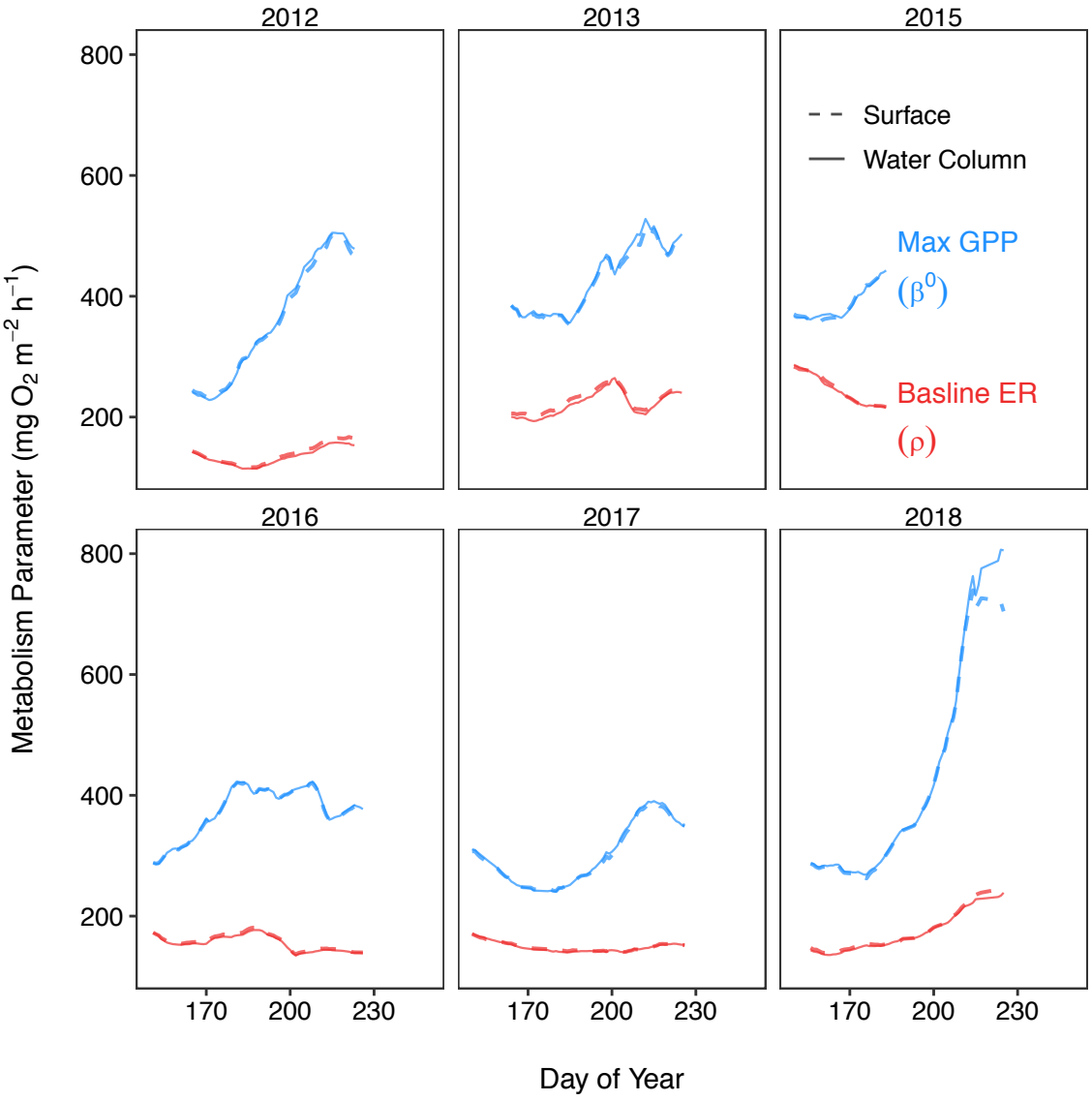

Figure B10. Ecosystem metabolism rates estimated using either surface or average water column light.

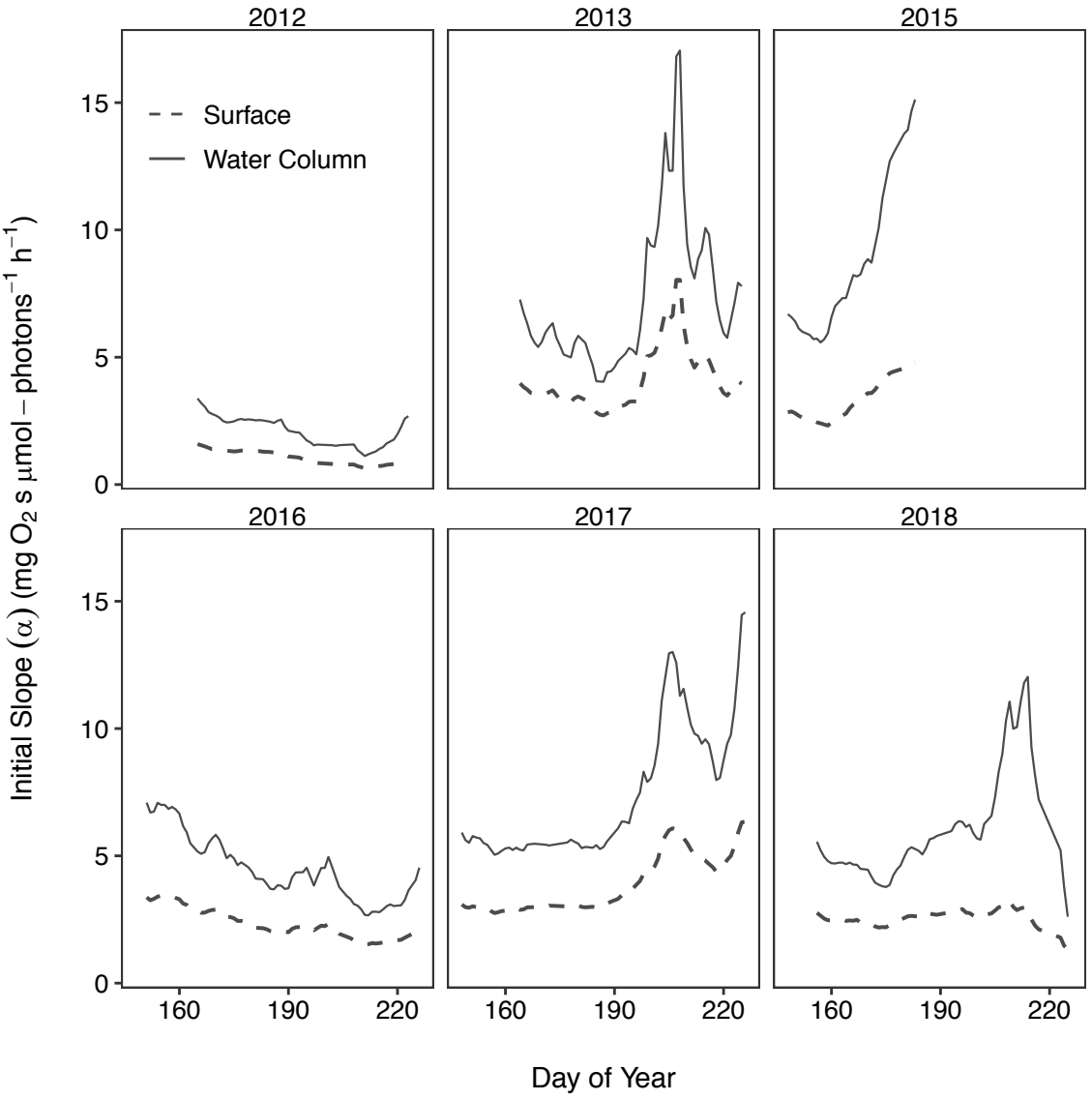

Figure B11. Initial slope of the P-I curve estimated using either surface or average water column light.

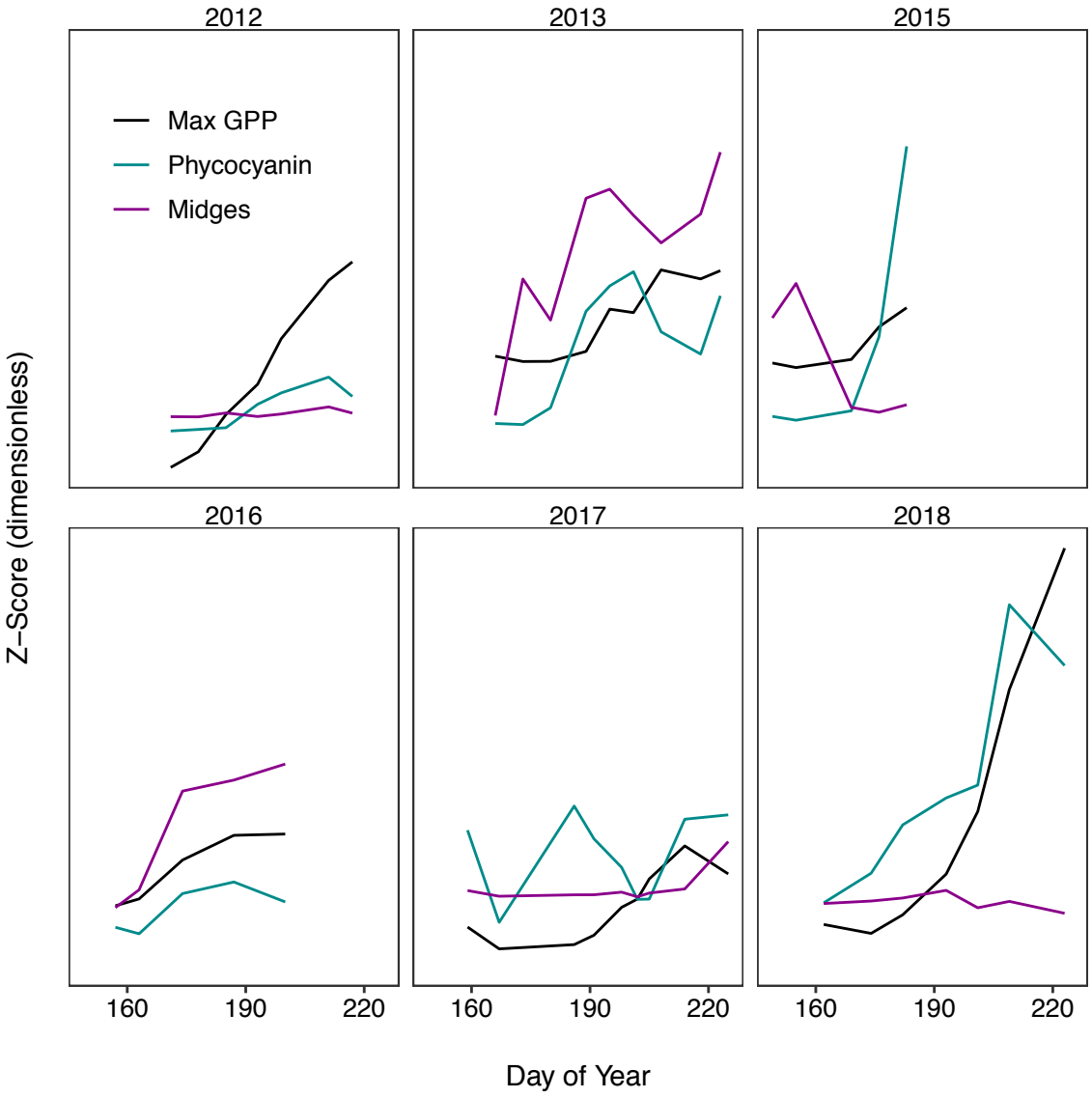

Figure B12. Time-series of maximum GPP (as inferred from the model), phycocyanin, and midge larva.

Chapter 3

Estimating time-varying demographic rates from age-structured abundance: application to a historically variable fishery

Joseph S. Phillips[1], Anthony R. Ives[1], and Guðni Guðbergsson[2]

[1]Department of Integrative Biology, University of Wisconsin-Madison, Madison, WI, 53706

[2]Marine and Freshwater Research Institute, Reykjavik, Iceland

Status: In preparation (near submission) for *Methods in Ecology & Evolution*

Abstract

A variety of methods have been developed for estimating recruitment and survival in populations from field data. One approach is to fit a model to age- or stage-structured time-series data; recruitment is inferred from the appearance of the earliest stages, while survival is inferred from the loss of individuals from cohorts. However, most statistical methods for time series assume that demographic rates are fixed through time, while those that attempt to incorporate temporal variation have proven to implement. This is a limitation because demographic rates are almost certainly variable for any natural population, and this variability can have important implications for population dynamics. Here, we present a flexible method for incorporating time-varying recruitment and survival into structured models that can be fit to time series of observed population abundance. The demographic rates are modeled as random walks, allowing them to vary smoothly through time while taking advantage of the full set of data for estimating all parameters simultaneously. We illustrate the utility of this method through application to age-structured observations of an arctic char fishery that was historically heavily exploited and has experienced substantial fluctuations over the 32-year time series. When fit to the arctic char time series, the model inferred substantial variation in recruitment and survival. This was corroborated by comparing the full model to reduced models with different combinations of demographic rates fixed through time, which all produced poorer fits to the data than the full model even when accounting for model complexity. Temporal variation in per capita recruitment was the dominant contributor to variation in the per capita population growth rate, while temporal variation in survival was relatively less important. These results illustrate the ability of our approach to estimate time-varying recruitment and survival in age- or stage- structured models. The method is flexible in its description of temporal variation in demographic rates and is readily

generalizable to a variety of population structures, which should make it a useful tool for inferring temporal variation in a variety of natural populations.

Keywords: Arctic char, matrix population model, recruitment, survival, state-space model

Introduction

Quantifying temporal variation in the age- and stage-specific demographic rates underpins efforts to characterize both exogenous and endogenous drivers of population dynamics (Twombly 1994, Zeng et al. 1998, Koons et al. 2016). For some taxa it is possible to repeatedly detect and identify individual organisms (i.e. mark-recapture and related methods), which enables relatively direct estimation of detection probabilities and transition rates between life-stages (Lebreton et al. 1992, Nichols et al. 1992, Fujiwara and Caswell 2002). While mark-recapture methods often assume that demographic rates are fixed through time, when applied over sufficiently long time periods they can be used to infer temporal variation in demographic rates (Hadley et al. 2007, Regehr et al. 2010). Recent work has used this approach to link temporal variation in age- and stage-specific demographic rates to environmental drivers such as climate (Forcada et al. 2008, Hunter et al. 2010) and harvesting (Michielsens et al. 2006, Whitlock and McAllister 2009, Whitlock et al. 2012).

Unfortunately, repeated observation of individuals is often labor-intensive and only applicable to a limited set of taxa typically comprising large or sessile organisms. Therefore, in many contexts estimation of demographic rates must rely on time series of observed population abundance (Zipkin et al. 2014). It is possible to estimate detection probabilities and transition rates between life-stages in a manner analogous to mark-recapture from repeated or spatially replicated observations of abundance using N-mixture models (Royle 2004, Dail and Madsen 2011, Zipkin et al. 2014). However, such methods are only applicable to populations (or subpopulations within a metapopulation) that are closed to immigration and reproduction during the period of repeated observations. Furthermore, these methods require multiple observations within each time period of interest and therefore are relatively labor-intensive. Some studies have

fit age- or stage-structure models to time-series of population abundance without the structural restrictions of N-mixture models (sometimes known as “inverse methods”) (Gross et al. 2002, Udevitz and Gogan 2012, Gudmundsson and Gunnlaugsson 2012, Smart et al. 2018). However, this approach generally assumes that demographic rates are fixed through time (e.g.Udevitz & Gogan, 2012) or change in to response measured covariates (e.g. Hansen, Ives, Vander Zanden, & Carpenter, 2013).

There have been a few previous attempts to incorporate temporal variation in demographic parameters into age- and stage-structured models fit to observed population abundance. One approach is to fit matrix population models to moving windows containing subsets of the full data, providing a sequence of time-dependent demographic rates (Twombly and Caswell 1991, Twombly 1994). However, this method fails to utilize all of the data to inform each time-specific demographic rate, resulting in potentially biased estimates and difficulties in quantifying parameter uncertainty (Wood 1994, Twombly 1994). Wood (1994) presents an alternative to matrix projection methods that entails fitting a smooth surface to observed abundances across organism age/stage and sample time, from which time-dependent recruitment and survival can be inferred from the shape of the surface. However, this method is computationally difficult to implement and is quite different from more conventional approaches to population modeling, which could explain why it has not been widely utilized (Caswell, 2001; but see Ohman & Wood, 1996).

In this study, we present a flexible method for incorporating time-varying recruitment and survival into structured population models that can be estimated using age- or stage-structured time series of observed abundance. The demographic rates are modeled as random walks, an approach that has previously been applied to infer population growth rates from non-

structured abundance (Zeng et al. 1998) and for age-specific mortality rates inferred from fisheries stock assessments (Nielsen and Berg 2014). This method takes advantage of the entire time series for estimating the parameters while allowing them to vary smoothly through time, potentially reducing bias and facilitating quantification of parameter uncertainty (Ives and Dakos 2012). Furthermore, this approach is a direct extension of widely used matrix population methods and therefore should be relatively intuitive to implement and interpret. The model parameters are estimated in a state-space framework, which in principle allows real changes in abundance to be distinguished from apparent changes due to observation error (Harvey 1989)

We illustrate the utility of this method through application to an age-structured time series of arctic char (*Salvelinus alpinus*) abundance from Lake Mývatn, a shallow eutrophic lake in Iceland. This population of char was an important fishery prior to a decline that followed very high exploitation rates and other changes in the lake ecosystem (Gudbergsson 2004). Fishing regulations have limited harvesting in recent years, and a key question is whether the population has started to recover. Furthermore, the dominant food sources for the char (midge larva and large-bodied zooplankton) fluctuate dramatically in abundance between years (Ornolfsdottir and Einarsson 2004, Ives et al. 2008), which may play an important role in char recruitment and survival. An important step to understanding the decline and possible recovery of the char population is quantifying the temporal variation in demographic rates and their resulting effects on variation in the population growth rate, making the Mývatn char fishery a good test case for our method.

Methods

We begin with a description of the study system and data to motivate the formulation of the model. However, the approach is a direct extension of matrix population models (Caswell 2001) that have been applied to a wide variety of taxa and is similarly generalizable. After explaining the procedure for estimating the parameters of the full model, we illustrate how comparison to reduced models can provide an assessment of the evidence for temporal variation in the demographic rates. Finally, we describe how the estimated sequence of time-varying recruitment and survival can used to calculate time-varying per capita population growth rates and how these can in turn be partitioned into contributions from each demographic parameter.

Study System

Mývatn is located in northeastern Iceland (65°40'N 17°00'W) and has a tundra-subarctic climate (Björnsson and Jónsson 2004). The lake is moderately large (37 km2) and is dived into north (8 km2) and south (29 km2) connected by a narrow channel (Einarsson et al. 2004). Mývatn is shallow (mean depth = 2m) and is fed by cold and warms springs rich in phosphorous (south basin N and P inputs of 1.4 and 1.5 g m-2 y-1, respectively) (Ólafsson 1979). Consequently, the lake has high levels of benthic primary production that supports large populations of midges comprising >30 species. Midges are an important food source for Mývatn's vertebrate populations, including arctic char, three-spined sticklebacks (*Gasterosteus aculeatus*), brown trout (*Salmo trutta*), and various waterfowl (Einarsson et al. 2004).

Arctic char (Salmonidae: Salmoninae) have a circumpolar distribution and are important consumers in many boreal and arctic lakes. While some char populations are anadromous, Mývatn's population appears to reside strictly within the lake, despite a major outflow that

connects the lake to the Greenland Sea (Gudbergsson 2004). Spawning occurs in the fall among individuals aged 4 years or older. There are two char morphs within Mývatn: “regular” (size at maturation 35-50cm) and “dwarf” (20-25cm). However, the dwarf morph is restricted to a small and relatively isolated part of the S basin; therefore, the “regular” morph is the focus of this study. Gut content data reveal a diverse diet, including midges, snails, clams, zooplankton, and sticklebacks. The large bodied zooplankton *Eurycercus* is thought to be a particularly important prey item. While many char populations are cannibalistic, especially in the absence of other large-bodied prey (e.g. fish), the char in Mývatn appear to lack cannibalism, possibly because of the abundance of other prey.

Data

Systematic surveys of Mývatn’s arctic char have been conducted with gill nets every year from 1986-2017. Most captured individuals were measured for length, mass, age (either based on otoliths or estimated from size), and a binary index of maturity. Surveys were conducted at 12 sites around the lake, most of which were located in the south basin. For our analysis, we pooled all observations across all sites and assumed that the population is well mixed across years. While this assumption is not completely realistic, there were no consistent differences across sites and variation through time was fairly synchronous (Supporting Information). Surveys were conducted in fall (late August through September) of every year (just before spawning) and in May a subset of years. For this analysis we used only the September data. Therefore, in our age classifications a fish of age x is a fish that survived to that age (e.g. an “age 1” fish was born in the previous fall). The oldest fish in the data set were 12 years old, although observations of fish older than 6 years are very sparse. Therefore, we grouped the oldest fish into a single

reproductive age class, called “adults”. In Mývatn, char reach maturity between 4 and 5 years of age, so all aged 4 and older were categorized as adults (denoted “age 4” for simplicity).

Model Structure

The model includes four age classes with three juvenile stages and one adult stage. Individuals aged 1 through 3 progress to the next age i with survival probability ϕ_t^{i}, while adults survive and return to the adult class with probability ϕ_t^4. Adults contribute to the age 1 fish of the next time step with per capita recruitment ρ_t. Assuming a population closed to immigration, this structure implies the following set of equations

$$\begin{aligned} n_{t+1}^1 &= \rho_t n_t^4 \\ n_{t+1}^2 &= \phi_t^1 n_t^1 \\ n_{t+1}^3 &= \phi_t^2 n_t^2 \\ n_{t+1}^4 &= \phi_t^3 n_t^3 + \phi_t^4 n_t^4 \end{aligned} \tag{1}$$

where n_t^i denotes the number of individuals of age i at time t. The per capita recruitment (ρ_t) varies through time as random walk on a log scale:

$$\begin{aligned} \log(\rho_{t+1}) &= \log(\rho_t) + \varepsilon_t^\rho \\ \varepsilon_t^\rho &\sim \mathcal{N}(0, \sigma_\rho) \end{aligned} \tag{2}$$

where σ_ρ is the standard deviation of Gaussian-distributed steps in the random walk. Modeling the random walk on a log scale ensures that ρ_t remains positive. Similarly, temporal variation in

age-specific survival probabilities (ϕ_t^i) is modeled as random walk on a logit scale (constraining ϕ_{t+1}^{i} between 0 and 1):

$$\mathrm{logit}(\phi_{t+1}^{\mathrm{i}}) = \mathrm{logit}\left(\phi_t^{\mathrm{i}}\right) + \varepsilon_t^{\phi_i}$$
$$\varepsilon_t^{\phi_i} \sim \mathcal{N}\left(0, \sigma_\phi\right) \quad (3)$$

where σ_ϕ is the standard deviation of the random walk. A single standard deviation is estimated for survival probabilities of all ages to reduce the number of parameters to estimate, although the actual degree of temporal variability is informed by the data and therefore can differ between ages. In principle different values of σ_ϕ could be used for each age class; however, in our application to the arctic char data this resulted in poor convergence during model fitting, so we used a single value of σ_ϕ for all life stages.

We fit the model in a state-space framework, using log-normally distributed observation error:

$$\log\left(n_{obs,t}^i\right) = \log\left(n_t^{\mathrm{i}}\right) + \varepsilon_t^{n_{obs}^i}$$
$$\varepsilon_t^{n_{obs}^i} \sim \mathcal{N}(0, \sigma_{obs}) \quad (4)$$

where $n_{obs,t}^i$ is the observed abundance of age i fish at time t and σ_{obs} is the observation error standard deviation (on a log-scale). The observation error is assumed to be symmetrically distributed about the log population size with the same standard deviation across life stages. This is an unrealistic assumption, as smaller fish are less likely to be captured in gill nets than large

fish. Indeed, for a given cohort the observed number of age 1 fish ($n^1_{obs,t}$) is generally much less than $n^2_{obs,t+1}$, implying an increase in age 2 fish despite the fact that the population is closed (Figure 1). Furthermore, $n^2_{obs,t}$ is generally only slightly higher than $n^3_{obs,t+1}$, implying an unrealistically high survivorship for age 2 fish. Therefore, it appears that both age 1 and 2 fish are systematically underrepresented in the survey, as is often the case for age- and stage-structured observations of abundance.

We addressed systematic detection bias across age in two ways. First, because the detection bias for age 1 fish was so great and the data so "noisy" in comparison to other age classes, we did not use those data to fit the model (i.e. n^1_t was treated as a latent variable) and set the survival probability of age 1 fish (ϕ^1_t) equal to 1. Therefore, ρ_t can be interpreted as the per capita recruitment and survival of fish to age 2 at time t + 1. Second, we evaluated the sensitivity of the parameter estimates to systematic under-detection of age 2 fish by imposing a bias of 0.5 relative to other ages (i.e. by doubling the observed number of age 2 fish prior to model fitting). This model resulted in elevated estimates of recruitment and reduced estimates of age 2 survival (Figure C3-4), which is a straightforward consequence of assuming relatively greater abundance of age 2 individuals. However, the pattern of temporal variation in recruitment was almost identical to the model without assuming systematic detection bias, as were the estimates of survival probabilities of age 3 and adults. The survival probability of age 2 fish was more strongly affected by imposing the systematic detection bias, primarily because the temporal variability of age 2 survival was constrained by its upper bound when estimated without detection bias. Because the true detection bias is unknown and could not be estimated from the data, we present the model fit to the observed data without imposing the detection bias, which

was reasonable given our focus on pattern of temporal variation that was largely insensitive to this bias (but see Appendix C for analysis of the model with systematic detection bias).

Model Fitting

We fit the model with a Bayesian approach using Stan 2.17.0 run from R 3.5.1 (R Core Team 2018) with the 'rstan' package (Stan Development Team 2018). Stan utilizes Hamiltonian Monte Carlo to generate Markov chains that converge to a stationary distribution approximating the joint posterior distribution for the unobserved variables (Carpenter et al. 2017).Weakly informative Gaussian priors were used for all parameters, truncated at 0 for parameters that must be nonnegative (σ_{obs}, σ_{ϕ}, and σ_{ρ}). We scaled the observation error standard deviation (σ_{obs}) by the standard deviation of the observed data (on a log scale) and scaled the per capita recruitment (ρ_t) by the mean abundance of age 2 individuals divided by the by the mean abundance of adults before fitting the model to facilitate selection of weakly informative priors on unit scale; these parameters were then transformed back to their natural scales following model fitting. We used four independent chains with diffuse initial values and 2000 iterations including a 1000-iteration "warm-up". Convergence was assessed by examining trace plots for parameter estimates, the number of divergent transitions of the Markov chain, and the potential scale reduction factor ($\hat{R}$), which quantifies the relative variance within and between chains. All parameters had $\hat{R} \leq 1.01$, which is regarded as a conservative threshold for assessing convergence as measured by $\hat{R}$. Detailed specifications for the priors, fitting routine, and additional diagnostics can be found at https://github.com/jsphillips2/arctic_char.

For most parameters we used posterior medians as point estimates and 16% and 84% quantiles for the bounds of 68% uncertainty intervals (UI68%), matching the coverage of

standard errors. However, for parameters with lower bounds of 0 for which quantiles provide poor summaries (σ_{obs}, σ_{ϕ}, and σ_{ρ}) we used posterior modes and 68% highest density intervals calculated (HDI68%) with the 'hdrcde' package in R 3.5.1 (R Core Team 2018).

Model Comparison

Modeling time-varying demographic rates as walks provides in a single parameter that characterizes the magnitude temporal variation in each demographic rate (σ_{ϕ} for recruitment and σ_{ρ} for survival), and fitting the model in a Bayesian framework yields in a posterior distribution characterizing the uncertainty in σ_{ϕ} and σ_{ρ}. However, the fact that σ_{ϕ} and σ_{ρ} are bounded by zero means that there is some ambiguity in assessing the strength of evidence for temporal variation in the demographic rates, which is important for the statistical objective of determining whether the data contain meaningful information about changes in demographic rates.

To provide a more formal assessment of the evidence for temporal variation in the demographic rates, we fit reduced versions of the model including only time-varying recruitment ("Variable Recruitment), only time-varying survival ("Variable Survival)", or with both fixed through time ("Both Fixed"). We then compared the fit of the models using the Leave-One-Out Information Criterion (LOOIC), which is related to WAIC but has been shown to be more robust (Vehtari et al. 2017). LOOIC is an approximation to leave-one-out cross validation and is calculated from the log pointwise predictive density penalized by an estimate of the effective number of parameters. In this way, LOOIC is conceptually similar to likelihood-based information criteria (e.g. AIC and BIC), and it is reported in units of deviance so that it can be interpreted in a similar manner. However, the calculation of LOOIC integrates across the full posterior distribution of

the model and therefore is fully Bayesian. We calculated LOOIC using the "loo" package in R 3.5.1 (R Core Team 2018).

Population Growth Rate

Quantifying time-varying demographic rates from Mývatn's arctic char fishery was motivated by a desire to characterize the dynamics of the population during its decline and possible recovery. To illustrate the utility of estimating time-varying demographic rates for this objective, we used the parameter estimates to partition the variance in the per capita population growth rate into contributions from different demographic parameters. Over a given time step, the realized per capita growth rate is calculated as

$$\lambda_t^{real} = \frac{N_t}{N_{t-1}} \tag{5}$$

where N_t is the summed abundance of all age classes. Changes in λ_t^{real} result from both changes in the demographic rates (ρ_t and ϕ_t^{i}) and shifting age structure when the population has yet to reach a stable age distribution, which is the case when demographic rates change though time (Caswell 2001, Koons et al. 2016). While it is possible to partition variance in λ_t^{real} to both changes in demographic rates and non-equilibrium stage structure (Koons et al. 2016), we take the alternative approach of calculating the asymptotic growth rate (λ_t^{asym}) determined by a given set of demographic rates under the assumption that the population has reached its stable age distribution. Temporal variation in λ_t^{asym} reflects temporal variation in the demographic parameters *per se* and so provides a synthetic measure of their collective influence on the

dynamics of the population (see Figure C1 for comparison of λ_t^{asym} and λ_t^{real} from the full model).

The asymptotic growth rate equals the leading eigenvalue of the demographic projection matrix $\mathbf{A_t}$ parameterized with recruitment and survival probabilities in a given time step (Caswell 2001):

$$\mathbf{A_t} = \begin{pmatrix} 0 & 0 & 0 & \rho_t \\ \phi_t^1 & 0 & 0 & 0 \\ 0 & \phi_t^2 & 0 & 0 \\ 0 & 0 & \phi_t^3 & \phi_t^4 \end{pmatrix} \quad (6)$$

We calculated λ_t^{asym} for the full model and the two reduced models that allowed at least one demographic rate to vary through time (i.e. Variable Recruitment and Variable Survival).
To partition variance in λ_t^{asym} for the full model resulting from temporal variation in the recruitment and survival, we define a vector $\boldsymbol{\lambda^{asym}}$ containing asymptotic growth rates and vectors $\boldsymbol{\theta^j}$ for each demographic parameter j (e.g. $\boldsymbol{\theta}^{\rho} = \{\rho_1, \rho_2, \ldots \rho_{\mathrm{tmax}}\}$). The contribution of $\boldsymbol{\theta^j}$ to the variance in $\boldsymbol{\lambda^{asym}}$ is then approximated as

$$\text{Contribution from } \boldsymbol{\theta^j} \approx \frac{1}{\mathrm{Var}[\boldsymbol{\lambda^{asym}}]} \sum_k \mathrm{Cov}\left[\boldsymbol{\theta^j}, \boldsymbol{\theta^k}\right] \frac{\partial \lambda_{\bar{\boldsymbol{\theta}}}^{asym}}{\partial \overline{\theta^j}} \frac{\partial \lambda_{\bar{\boldsymbol{\theta}}}^{asym}}{\partial \overline{\theta^k}} \quad (7)$$

where $\mathrm{Cov}\left[\boldsymbol{\theta^j}, \boldsymbol{\theta^k}\right]$ is the covariance between the parameter vectors $\boldsymbol{\theta^j}$ and $\boldsymbol{\theta^k}$ (the covariance when j = k is the variance of $\boldsymbol{\theta^j}$), $\overline{\theta^j}$ is the mean value for parameter vector $\boldsymbol{\theta^j}$, $\lambda_{\bar{\boldsymbol{\theta}}}^{asym}$ is the asymptotic growth rate for a projection matrix consisting of the mean parameter values, and

$\mathrm{Var}[\boldsymbol{\lambda^{asym}}]$ is the total variance in $\boldsymbol{\lambda^{asym}}$ calculated as the sum of the contributions for all $\boldsymbol{\theta^j}$ (Caswell 2001). The partial derivatives $\left(\frac{\partial \lambda_{\bar{\boldsymbol{\theta}}}^{asym}}{\partial \bar{\theta}^j}\right)$ are the "sensitives" of the asymptotic growth with respect to mean value of a given demographic parameter. Because the variance terms are generally larger than the covariances, the contribution of a given demographic parameter to variance in $\lambda_{\bar{\boldsymbol{\theta}}}^{asym}$ is largely determined by the variance of the parameter and the sensitivity of $\lambda_{\bar{\boldsymbol{\theta}}}^{asym}$ to that variance. We calculated the contributions from each parameter for 1000 draws from the posterior distribution of the full model, which we summarized using medians and 16% and 84% quantiles.

Data and code availability

Data and R code for all analyses can be found at https://github.com/jsphillips2/arctic_char.

Results

Model Fit

The full model provided a good visual fit to the observed data, with short-term fluctuations attributed to observation error and longer-term changes attributed to real changes abundance (Figure 1). However, there was some apparent temporal autocorrelation in the observation error, which could either be due genuine autocorrelation in the observation process or from the model failing to fully separate observation error from changes in the true population size. The standard deviation for per capita recruitment (σ_ρ) indicated substantial temporal variation in recruitment (Figure 2) with a peak in the early 1990's followed by two periods of steep decline in the late 1990's and late 2000's (Figure 3). This manifested in the observed abundance of age 2 and 3 fish, which declined in the late 1990's and remained fairly steady

thereafter despite increases in the abundance of adults. The decline in recruitment was also visible in the observed abundance of age 1 individuals, even though these data were not used to fit the model (Figure 1). Age-specific survival probabilities (ϕ_t^{i}) also appeared to vary through time (Figures 2 and 4), with an increase in survival for age 3 and adult individuals in the early 2000s, followed by modest declines. However, the posterior distribution for σ_ϕ had high variance when compared to σ_ρ, and the precision in the estimates of ϕ_t^{i} was low relative to the magnitude of temporal variability, which indicates greater uncertainty in the magnitude of variation in survival than for recruitment.

Model Comparison

To determine the strength of evidence for temporal variation in survival and recruitment, we fit reduced versions of the full model and compared their fits to the observed data using LOOIC. The full model had by far the lowest LOOIC, with ΔLOOIC > 30 for the other models, even though its effective number of parameters was much greater. The improved fit of the full model relative to the reduced models can be seen in Figure 5, especially for the peak in the abundance of age classes two and three in the 1990's and the increase in adults beginning in the mid 2000's. Altogether, this indicates evidence in favor of the model including temporal variation in both survival and recruitment.

The LOOIC for Variable Survival was two units lower than for Variable Recruitment, indicating modest support for the former over the latter (Table I). This was due in part to the higher number of effective parameters for Variable Recruitment, which itself was a result of the estimates for recruitment being much more temporally variable than the estimates for survival (Figures C7-8). The temporal patterns for the demographic rates estimated by Variable

Recruitment and Variable Survival were largely similar to the corresponding estimates in the full model, with the main exception being that Variable Recruitment did not detect the decline in recruitment over the last decade that was inferred by the full model (Figure C7). Furthermore, the effective number of parameters in the full model was nearly the same as the combined effective number of parameters from Variable Recruitment and Variable Survival (Table I). These results show that the estimates of time-varying recruitment and survival were largely independent of each other, which is consistent with both contributing substantially to the fit of the full model.

Population Growth Rate

The asymptotic growth rate for the full model varied substantially through time (Figure 6), indicating periods of both population growth ($\lambda_t^{asym} > 1$) and decline ($\lambda_t^{asym} < 1$). The population growth rate peaked in the mid 1990's, followed by a steep decline and a subsequent recovery towards positive growth in the mid 2000's. However, the population growth rate gradually declined over the last decade and was slightly below the replacement rate in the most recent year of observation (2017). This recent decline in the asymptotic growth rate coincided with relatively stable abundance of age classes two and three despite an increase in the adults (Figure 1), consistent with the decline in recruitment inferred by the full model (Figure 3). The relative importance of time-varying recruitment for the per capita growth rate can be seen by comparing λ_t^{asym} calculated for the reduced models Variable Recruitment and Variable Survival (Figure 6). Despite the fact that Variable Survival provided a slightly better fit to the observed data, Variable Recruitment provided estimates of λ_t^{asym} that were qualitatively more similar to the full model than those for Variable Survival. Only allowing temporal variation in survival

resulted in a substantial underestimate of the overall temporal variability in λ_t^{asym}, which failed to show the peak in the early 1990's and the decline over the last decade inferred by the two models allowing variable recruitment. The variance partitioning of λ_t^{asym} for the full model into contributions from time-varying recruitment and survival corroborates this point, with recruitment accounting for nearly two-thirds of the variance in the asymptotic growth rate (Figure 7). Survival probability of age three fish to adulthood made the next largest contribution to the variance in λ_t^{asym}, with the contribution of adult survival being small and the contribution of survival of age two fish being almost negligible.

The full model assuming systematic under-detection of age 2 fish produced estimates of $\boldsymbol{\lambda^{asym}}$ nearly identical to those assuming no such bias (Figure C5). The model with system under-detection also resulted in a very similar partitioning of variance, but with a slightly greater contribution of age 2 individuals (Figure C6).

Discussion

In this study, we present an approach for using observations of abundance to estimate temporal variation in recruitment and survival with age- or stage-structured state-space models. The approach characterizes time-varying demographic rates as random walks, which has the advantage of using all of the data to inform the estimates of all parameter values simultaneously while allowing the parameters to vary smoothly through time (Zeng et al. 1998, Nielsen and Berg 2014). The method is a direct extension of matrix population models (Caswell 2001) that have been applied to a wide variety of taxa and is similarly generalizable. Furthermore, the method is straightforward to implement in the Stan statistical software (Carpenter et al. 2017) and could be readily adapted for other modeling-fitting software environments that have become

increasingly used by ecologists, such as WinBUGS, JAGS, and AD Model Builder (Bolker et al. 2013). When applied to the Mývatn arctic char fishery, the method inferred substantial temporal variation in both recruitment and age-specific survival. Variation in these demographic rates was responsible for large fluctuations in the growth rate of the char population, including a steady decline over the last of the time series decade despite efforts to alleviate harvesting pressure.

We illustrated our method in the context of state-space models fit to time series of observed abundance, which are often the only data available in both basic and applied contexts. However, when appropriate data are available, using random walks to model temporal variation in recruitment and survival could also be incorporated into mark-recapture and N-mixture models. This is conceptually similar to modeling variation in demographic rates using random effects or splines, as has previously been done for mark-recapture type models either alone or in conjunction with explicit covariates (Gimenez et al. 2006, Schofield et al. 2009, Grosbois et al. 2009, Regehr et al. 2010). The advantage of modeling demographic rates as random walks is the explicit assumption that the demographic rates are autocorrelated through time, as is likely the case for most populations since the biological (e.g. organismal traits) and environmental (e.g. climate) drivers of recruitment and survival are themselves temporally autocorrelated.

Any method for estimating temporal variation in demographic rates should provide an approach for assessing the strength of evidence for such variation (Fieberg and Staples 2006, Grosbois et al. 2009, Regehr et al. 2010). This is true not only for the ecological objective of determining whether the demographic rates actually vary, which is likely the case to at least some extent for any wild population, but also for the statistical objective of determining whether the data contain meaningful information about changes in demographic rates. Modeling temporal variation in demographic rates as random walks provides a single parameter characterizing the

extent to which a given demographic rate varies through time: the standard deviation of step sizes in the random walk. When fit to Mývatn's arctic char data, the posterior distribution of the standard deviations for variation in recruitment (σ_ρ) provided clear evidence of variation, while the evidence for variation in survival (σ_ϕ) was more equivocal due to low precision and much of the posterior density being concentration near the lower bound of 0. To more formally assess the evidence for temporal variation in the demographic rates, we fit reduced versions of the model (Variable Recruitment, Variable Survival, and Both Fixed) and then compared their fits to the observed data using LOOIC. The full model fit the data substantially better than any of the reduced models, which provides evidence for temporal variation in both recruitment and survival. Comparison of the full model to Variable Recruitment and Variable Survival indicated that the estimates of time-varying recruitment and survival were largely independent of each other. This is a desirable statistical property, as extreme interdependence of the estimated demographic rates could lead to poor identifiability and make it difficult to evaluate their relative importance to the model fit, which has limited previous attempts to estimate time-varying demographic rates in age structured models (as discussed in Wood 1994).

We applied our method to Mývatn's arctic char fishery to characterize its dynamics during a decline and possible recovery following relaxed harvesting pressure. The per capita growth rate of the char population varied substantially through time, with large oscillations between population growth ($\lambda_t^{asym} > 1$) and decline ($\lambda_t^{asym} < 1$) preceding a period of modest population growth over the last decade. This period of growth was accompanied by a steady increase in the abundance of adult fish, which has been taken as a sign of possible recovery by some of the local fisherman. However, this increase in the abundance of adult fish coincided with relatively little change in the abundance of younger fish, which is the consequence of a

large decline in recruitment. This shift towards a relatively older age structure has manifested in the gradual but steady decline in the asymptotic growth rate from values unambiguously greater than one in the late 2000's to very slightly below one in 2017. While this decline was mostly driven by recruitment, the survival probabilities of age 3 fish and adults plateaued around 2010 and have slightly declined in the last few years. A continuation of these trends would lead to a decline of the population in the near future, despite the hope for the population to return to its historically high abundances that could support increased harvest (Gudbergsson 2004). A crucial step for projecting the future dynamics of the population is identifying the underlying causes for changes in demographic rates beyond harvesting pressure, which are currently unknown but are speculated to be connected to large fluctuations in zooplankton and aquatic insects (midges) that are major components of the char diet (Gudbergsson 2004, Einarsson and Örnólfsdóttir 2004, Gardarsson et al. 2004).

In this study, we presented a method for inferring time-varying demographic rates in age- and stage-structured models from time series of observed abundance. Our method performs well given the limited nature of the data, providing reasonable estimates for temporal variation in recruitment and survival. Application to Mývatn's arctic char population illustrates the utility of this approach for investigating the possible recovery of a historically heavily exploited fishery. This approach is readily generalizable to other populations and may help managers to make important policy decisions that often rely on limited and noisy data (Doak et al. 2005).

References

Björnsson, H., and T. Jónsson. 2004. Climate and climatic variability at Lake Myvatn. Aquatic Ecology 38:129–144.

Bolker, B. M., B. Gardner, M. Maunder, C. W. Berg, M. Brooks, L. Comita, E. Crone, S. Cubaynes, T. Davies, P. de Valpine, J. Ford, O. Gimenez, M. Kéry, E. J. Kim, C. Lennert-Cody, A. Magnusson, S. Martell, J. Nash, A. Nielsen, J. Regetz, H. Skaug, and E. Zipkin. 2013. Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. Methods in Ecology and Evolution 4:501–512.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell. 2017. Stan: A Probabilistic Programming Language. Journal of Statistical Software 76.

Caswell, H. 2001. Matrix Population Models. 2nd edition. Sinaur Associates, Sunderland, MA, USA.

Dail, D., and L. Madsen. 2011. Models for Estimating Abundance from Repeated Counts of an Open Metapopulation. Biometrics 67:577–587.

Doak, D. F., K. Gross, and W. F. Morris. 2005. Understanding and Predicting the Effects of Sparse Data on Demographic Analyses. Ecology 86:1154–1163.

Einarsson, A., and E. B. Örnólfsdóttir. 2004. Long-term changes in benthic Cladocera populations in Lake Myvatn, Iceland. Aquatic Ecology 38:253–262.

Einarsson, Á., G. Stefánsdóttir, H. Jóhannesson, J. S. Ólafsson, G. M. Gíslason, I. Wakana, G. Gudbergsson, and A. Gardarsson. 2004. The ecology of Lake Myvatn and the River Laxá: Variation in space and time. Aquatic Ecology 38:317–348.

Fieberg, J., and D. F. Staples. 2006. The role of variability and uncertainty in testing hypotheses involving parameters in stochastic demographic models. Canadian Journal of Zoology 84:1698–1701.

Forcada, J., P. N. Trathan, and E. J. Murphy. 2008. Life history buffering in Antarctic mammals and birds against changing patterns of climate and environmental variation. Global Change Biology 14:2473–2488.

Fujiwara, M., and H. Caswell. 2002. Estimating Population Projection Matrices from Multi-Stage Mark–Recapture Data. Ecology 83:3257–3265.

Gardarsson, A., Á. Einarsson, G. M. Gíslason, T. Hrafnsdóttir, H. R. Ingvason, E. Jónsson, and J. S. Ólafsson. 2004. Population fluctuations of chironomid and simuliid Diptera at Myvatn in 1977–1996. Aquatic Ecology 38:209–217.

Gimenez, O., C. Crainiceanu, C. Barbraud, S. Jenouvrier, and B. J. T. Morgan. 2006. Semiparametric Regression in Capture-Recapture Modeling. Biometrics 62:691–698.

Grosbois, V., M. P. Harris, T. Anker-Nilssen, R. H. McCleery, D. N. Shaw, B. J. T. Morgan, and O. Gimenez. 2009. Modeling survival at multi-population scales using mark–recapture data. Ecology 90:2922–2932.

Gross, K., B. A. Craig, and W. D. Hutchison. 2002. Bayesian Estimation of a Demographic Matrix Model from Stage-Frequency Data. Ecology 83:3285–3298.

Gudbergsson, G. 2004. Arctic charr in Lake Myvatn: the centennial catch record in the light of recent stock estimates. Aquatic Ecology 38:271–285.

Gudmundsson, G., and T. Gunnlaugsson. 2012. Selection and estimation of sequential catch-at-age models. Canadian Journal of Fisheries and Aquatic Sciences 69:1760–1772.

Hadley, G. L., J. J. Rotella, and R. A. Garrott. 2007. Influence of maternal characteristics and oceanographic conditions on survival and recruitment probabilities of Weddell seals. Oikos 116:601–613.

Hansen, G. J. A., A. R. Ives, M. J. Vander Zanden, and S. R. Carpenter. 2013. Are rapid transitions between invasive and native species caused by alternative stable states, and does it matter? Ecology 94:2207–2219.

Harvey, A. C. 1989. Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge, UK.

Hunter, C. M., H. Caswell, M. C. Runge, E. V. Regehr, S. C. Amstrup, and I. Stirling. 2010. Climate change threatens polar bear populations: a stochastic demographic analysis. Ecology 91:2883–2897.

Ives, A. R., and V. Dakos. 2012. Detecting dynamical changes in nonlinear time series using locally linear state-space models. Ecosphere 3:art58.

Ives, A. R., Á. Einarsson, V. A. A. Jansen, and A. Gardarsson. 2008. High-amplitude fluctuations and alternative dynamical states of midges in Lake Myvatn. Nature 452:84–87.

Koons, D. N., D. T. Iles, M. Schaub, and H. Caswell. 2016. A life-history perspective on the demographic drivers of structured population dynamics in changing environments. Ecology Letters 19:1023–1031.

Lebreton, J.-D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling Survival and Testing Biological Hypotheses Using Marked Animals: A Unified Approach with Case Studies. Ecological Monographs 62:67–118.

Michielsens, C. G. ., M. K. McAllister, S. Kuikka, T. Pakarinen, L. Karlsson, A. Romakkaniemi, I. Perä, and S. Mäntyniemi. 2006. A Bayesian state-space mark-recapture model to estimate exploitation rates in mixed-stock fisheries. Canadian Journal of Fisheries and Aquatic Sciences 63:321–334.

Nichols, J. D., J. R. Sauer, K. H. Pollock, and J. B. Hestbeck. 1992. Estimating Transition Probabilities for Stage-Based Population Projection Matrices Using Capture-Recapture Data. Ecology 73:306–312.

Nielsen, A., and C. W. Berg. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research 158:96–101.

Ohman, M. D., and S. N. Wood. 1996. Mortality estimation for planktonic copepods: Pseudocalanus newmani in a temperate fjord. Limnology and Oceanography 41:126–135.

Ólafsson, J. 1979. The Chemistry of Lake Mývatn and River Laxá. Oikos 32:82–112.

Ornolfsdottir, E. B., and A. Einarsson. 2004. Spatial and temporal variation of benthic Cladocera (Crustacea) studied with activity traps in Lake Myvatn, Iceland. Aquatic Ecology 38:239–251.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Regehr, E. V., C. M. Hunter, H. Caswell, S. C. Amstrup, and I. Stirling. 2010. Survival and breeding of polar bears in the southern Beaufort Sea in relation to sea ice. Journal of Animal Ecology 79:117–127.

Royle, J. A. 2004. N-Mixture Models for Estimating Population Size from Spatially Replicated Counts. Biometrics 60:108–115.

Schofield, M. R., R. J. Barker, and D. I. MacKenzie. 2009. Flexible hierarchical mark-recapture modeling for open populations using WinBUGS. Environmental and Ecological Statistics 16:369–387.

Smart, J. J., A. E. Punt, M. Espinoza, W. T. White, and C. A. Simpfendorfer. 2018. Refining mortality estimates in shark demographic analyses: a Bayesian inverse matrix approach. Ecological Applications 28:1520–1533.

Twombly, S. 1994. Comparative demography and population dynamics of two coexisting copepods in a Venezuelan floodplain lake. Limnology and Oceanography 39:234–247.

Twombly, S., and H. Caswell. 1991. Demographic analysis of tropical zooplankton populations. SIL Proceedings, 1922-2010 24:1183–1187.

Udevitz, M. S., and P. J. P. Gogan. 2012. Estimating survival rates with time series of standing age-structure data. Ecology 93:726–732.

Vehtari, A., A. Gelman, and J. Gabry. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing 27:1413–1432.

Whitlock, R., and M. McAllister. 2009. A Bayesian mark–recapture model for multiple-recapture data in a catch-and-release fishery. Canadian Journal of Fisheries and Aquatic Sciences 66:1554–1568.

Whitlock, R., M. K. McAllister, and B. A. Block. 2012. Estimating fishing and natural mortality rates for Pacific bluefin tuna (*Thunnus orientalis*) using electronic tagging data. Fisheries Research 119–120:115–127.

Wood, S. N. 1994. Obtaining Birth and Mortality Patterns From Structured Population Trajectories. Ecological Monographs 64:23–44.

Zeng, Z., R. M. Nowierski, M. L. Taper, B. Dennis, and W. P. Kemp. 1998. Complex population dynamics in the real world: modeling the influence of time-varying parameters and time lags. Ecology 79:2193–2209.

Zipkin, E. F., J. T. Thorson, K. See, H. J. Lynch, E. H. C. Grant, Y. Kanno, R. B. Chandler, B. H. Letcher, and J. A. Royle. 2014. Modeling structured population dynamics using data from unmarked individuals. Ecology 95:22–29.

Author Contributions

GG conducted the surveys of Mývatn's arctic char population. JSP conducted the analysis, with input from ARI. JSP wrote the first draft of the paper, with substantial contributions from all authors to subsequent drafts.

Acknowledgements

The data used in this manuscript were collected during routine sampling of Mývatn's char population supported X and the Mývatn Research Station directed by Árni Einarsson. Further support came from NSF LTREB DEB-1556208 to ARI and NSF Graduate Research Fellowship (DGE-1256259) supporting JSP

Table I: Comparisons of the full and reduced versions of the model fit to the arctic char data. LOOIC is in units of deviance (similar to AIC) and ΔLOOIC shows the difference in LOOIC between a given model and the full model.

Model	Effective parameters	LOOIC	ΔLOOIC
Full	26.1	223	0
Variable Survival	12.1	256	33
Variable Recruitment	15.0	258	35
Both Fixed	6.01	272	49

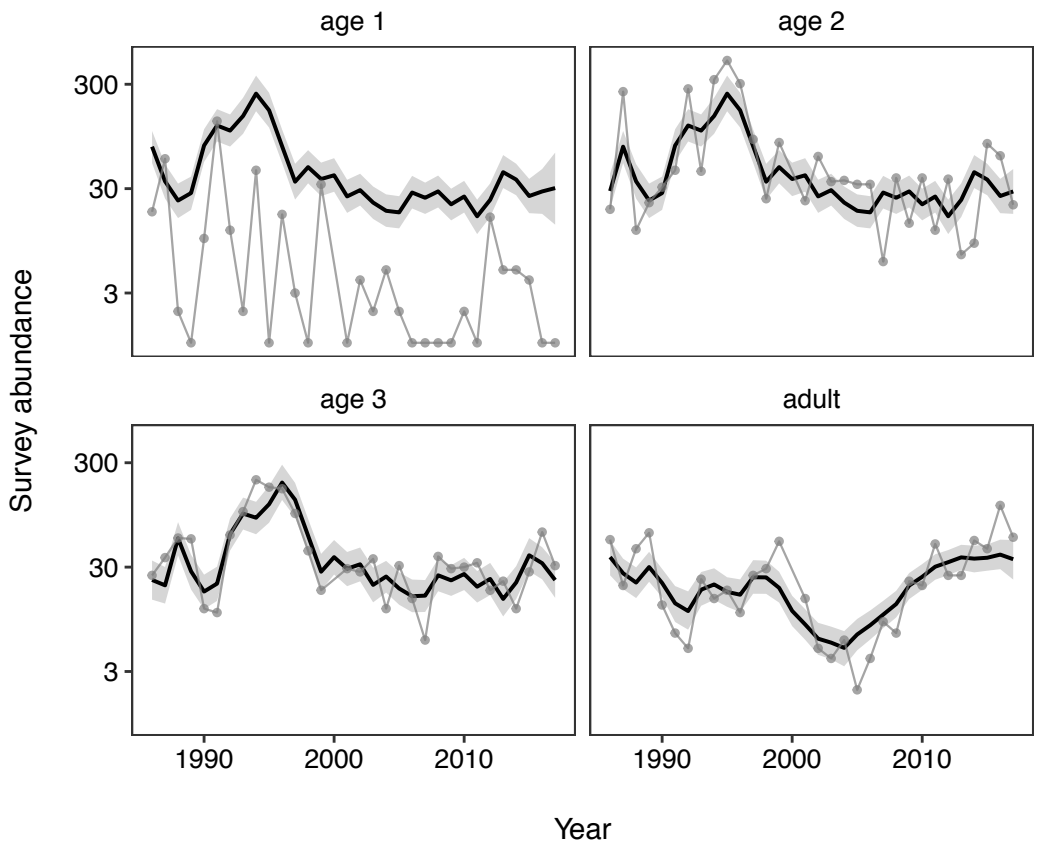

Figure 1. Model fits (thick lines) to the surveyed abundance of Mývatn's arctic char population (gray lines and points), on a log scale. The shaded regions correspond to $UI_{68\%}$, matching the coverage of standard errors. Data for age 1 individuals were not used to fit the model due to large detection bias and "noisy" fluctuations relative to other age classes.

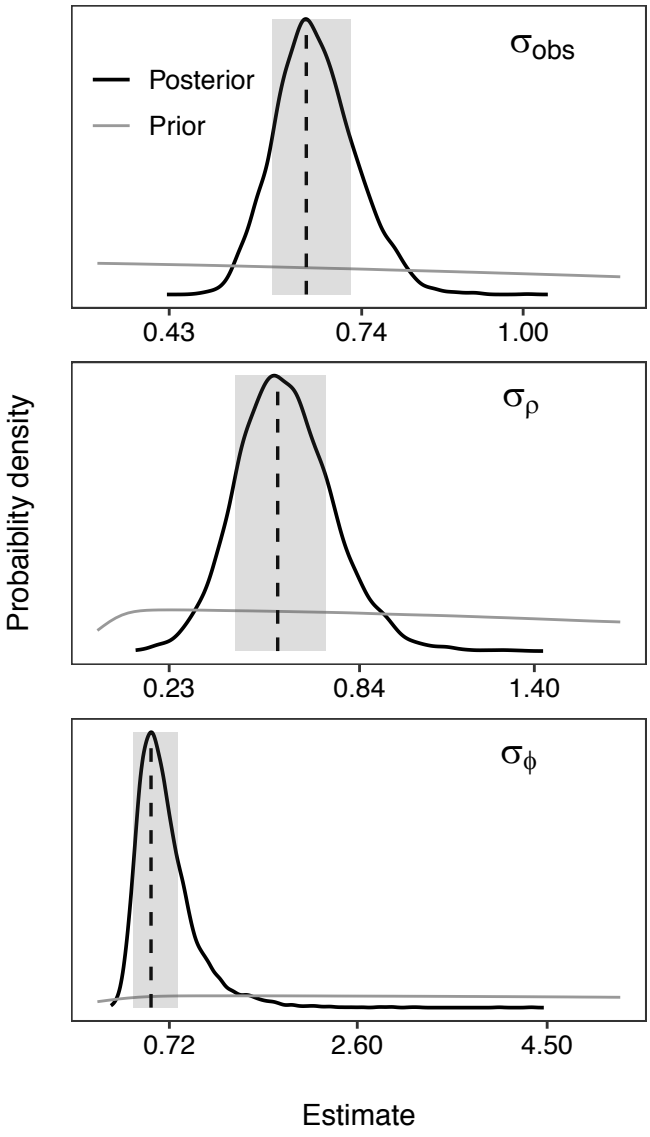

Figure 2. Prior and posterior densities for standard deviation of observation error (σ_{obs}), recruitment (σ_{ρ}), and survival probability (σ_{ϕ}) from the model fit the real data. The dashed vertical lines correspond to the posterior modes and the shaded regions correspond to the $HDI_{68\%}$, matching the coverage of standard errors. Probability densities are scaled to the same maximum value across panels for ease of visualization.

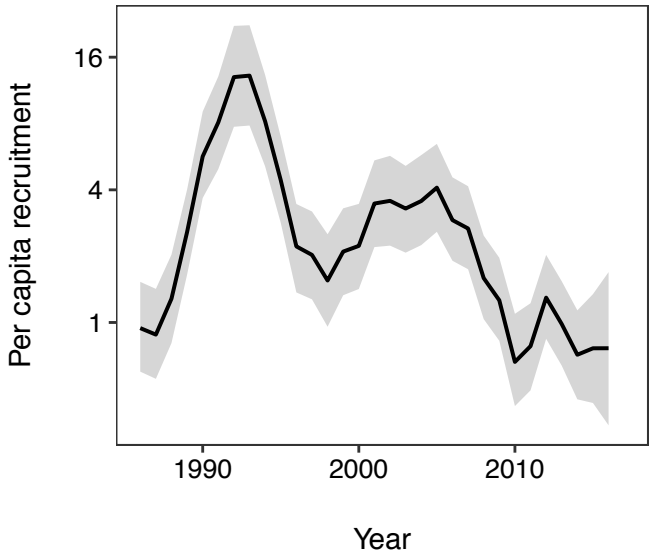

Figure 3. Estimated per capita recruitment (ρ_t) of Mývatn's arctic char population, on a log scale. The shaded region corresponds to $UI_{68\%}$.

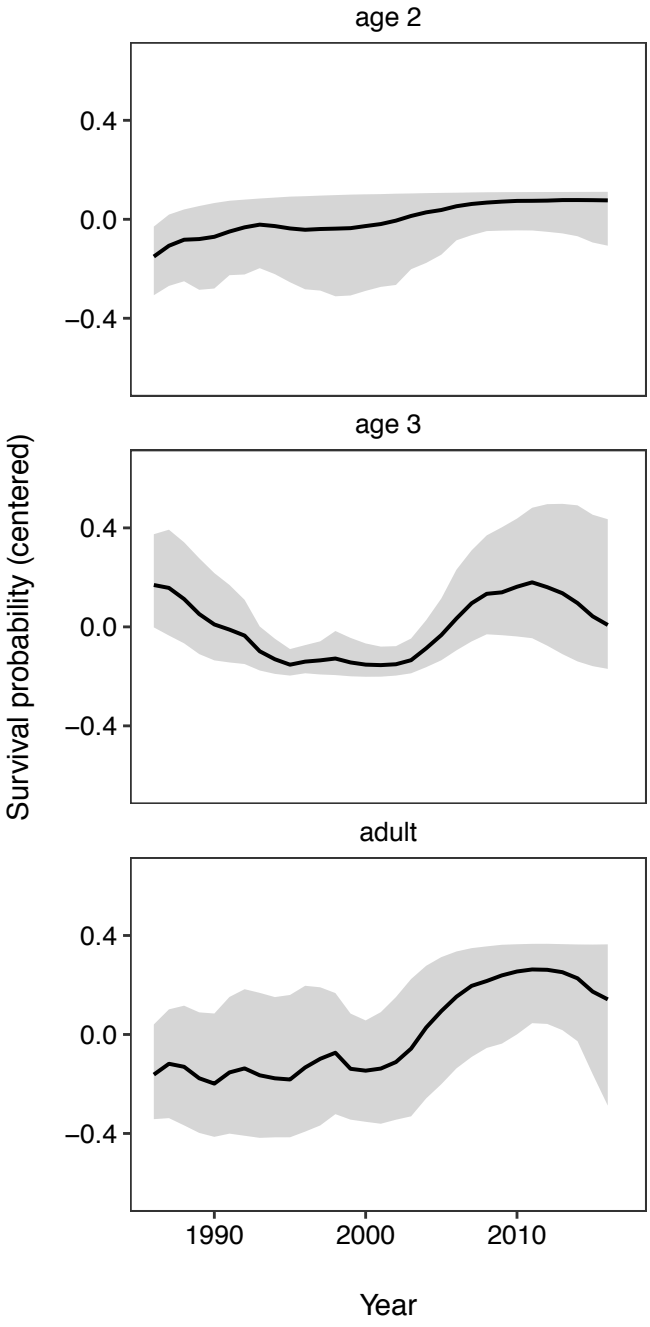

Figure 4. Estimated age-specific survival probabilities (ϕ_t^i) of Mývatn's arctic char population. The shaded regions correspond to $UI_{68\%}$. The estimates are centered on 0, as the overall magnitudes of the estimates are biased by differences in detection probability across different ages.

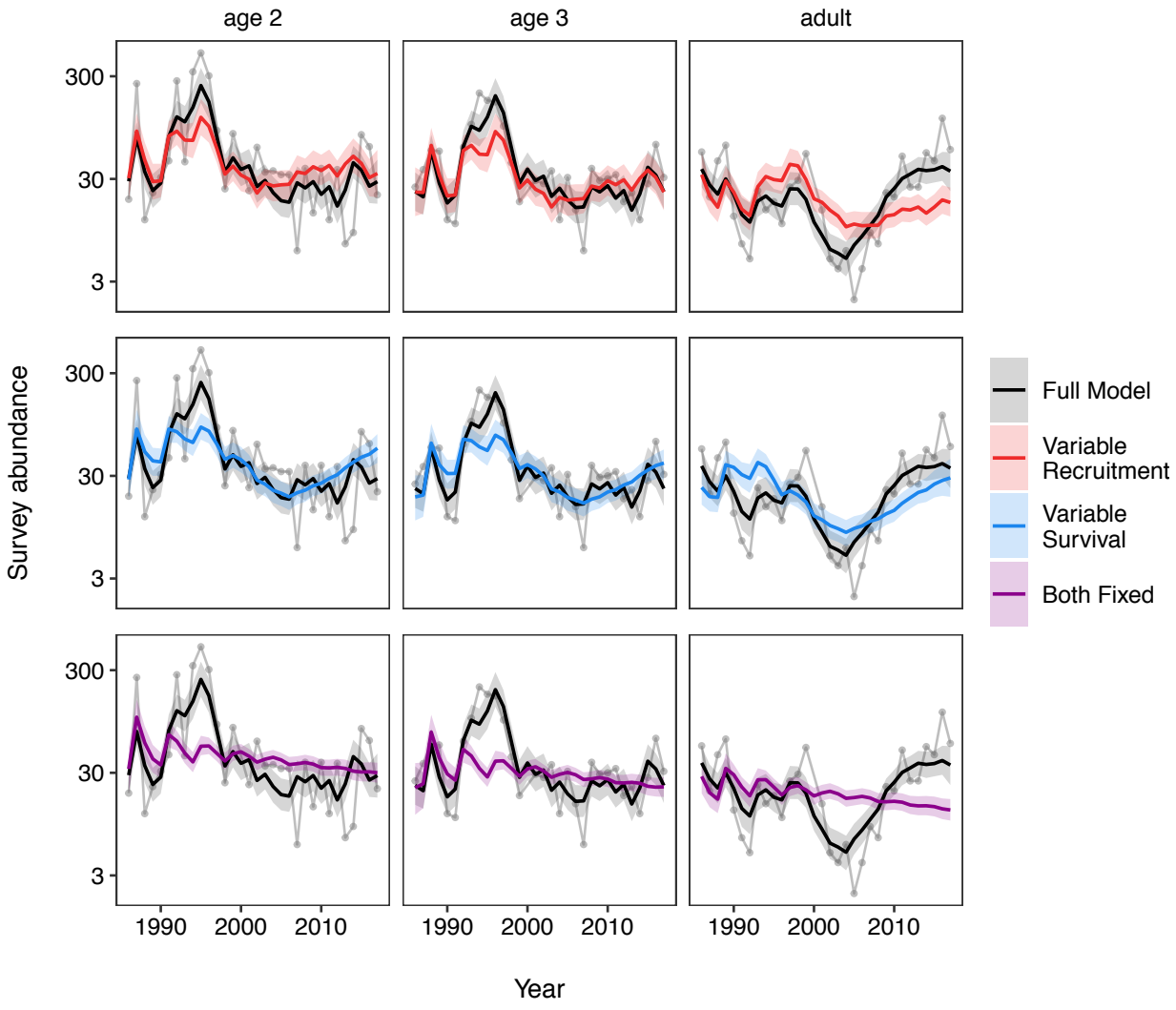

Figure 5. Comparison of the full and reduced models to the observed data. The observed data are shown as points and light gray lines. The full model is shown in each panel to facilitate comparison with each reduced model. The shaded regions correspond to $UI_{68\%}$.

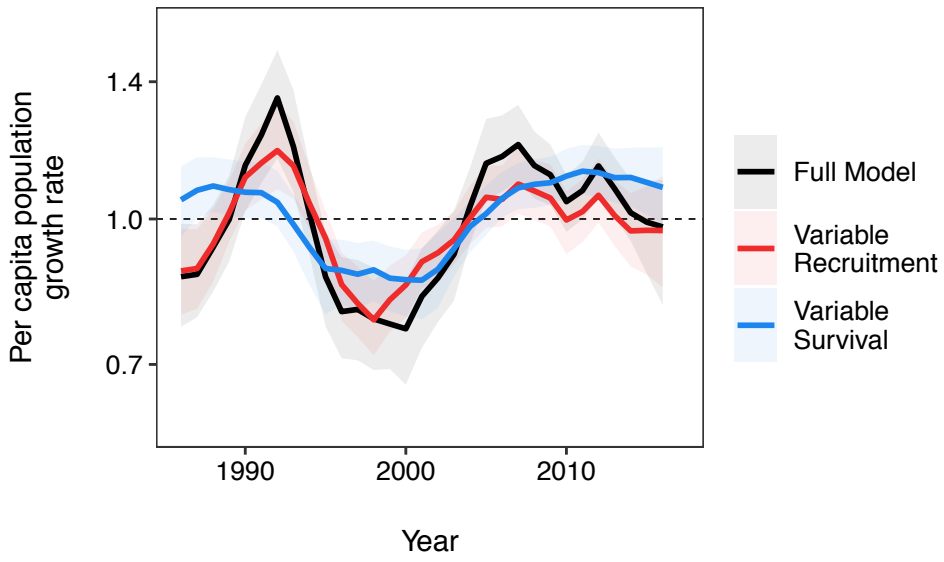

Figure 6. Asymptotic per capita growth rates (λ_t^{asym}) for the full model and the two reduced models that included at least one time-varying demographic parameter, on a log scale. The shaded regions correspond to $UI_{68\%}$. Per capita growth rates of 1 (horizontal dashed line) indicate no change in population size (asymptotically), with values above or below 1 indicating population increase or decrease, respectively.

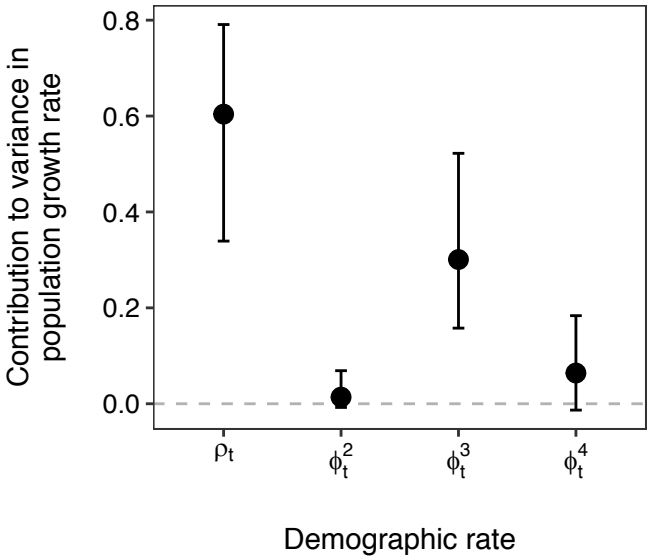

Figure 7. Partitioning of variance in the asymptotic per capita growth rate (λ_t^{asym}) into contributions from time-varying recruitment (ρ_t) and survival (ϕ_t^i). The error bars correspond to $UI_{68\%}$, matching the coverage of standard errors. The survival probability of age 1 fish does not contribute to variance in the per capita growth rate because it is fixed through time.

Appendices

Appendix C: Supplementary methods and results for Chapter 3

Appendix C: Supplementary methods and results for Chapter 3

Realized versus asymptotic population growth rate

Figure C1 compares the realized (λ_t^{real}) and asymptotic (λ_t^{asym}) per capita population growth rates for the full model, as defined in the main text of Chapter 2. The realized growth rate includes variation due to both changes in demographic rates and non-equilibrium age structure, while the asymptotic growth rate assumes equilibrium age structure and therefore only reflects changes in the demographic rates. Therefore, the realized growth rate provides information about the actual history of the population, while the asymptotic growth rate isolates the contribution of the demographic rates *per se* (which was of primary interest in this study). This difference is reflected in the fact that the realized growth rate is substantially more variable than the asymptotic growth rate. Nonetheless, the overall temporal pattern for the two is similar, which is a reflection of the relative importance of the time-varying demographic rates.

Systematic detection bias for age class 2

In the arctic char survey data, the observed number of age 1 fish in a given cohort ($n_{obs,t}^1$) is generally much less than $n_{obs,t+1}^2$, implying an increase in age 2 fish despite the fact that the population is closed (Chapter 2 figure 1). Furthermore, $n_{obs,t}^2$ is generally only slightly higher than $n_{obs,t+1}^3$, implying an unrealistically high survivorship for age 2 fish. Therefore, it appears that both age 1 and 2 fish are systematically underrepresented in the survey, as is often the case for age- and stage-structured observations of abundance. We address this in two ways: (1) by not using the data on age 1 fish to fit the model and instead treating it as a latent variable and (2) by exploring the consequences of imposing a systematic detection bias on age 2 fish relative to the

other stages. The true bias in unknown, so for the sake of this exercise we imposed a bias of 0.5 relative to the other life stages, which was equivalent to doubling the observed number of age 2 fish prior to fitting the model.

With the exception of age 2 survival, the temporal patterns of all parameters from the model fit with the systematic detection bias were nearly identical to those for the model without the bias (Figures C2-4). The abundance of age 2 fish and the per capita recruitment estimated from the model with bias were larger than those for the model without, which is a straightforward consequence of the imposed bias. Survival of the age 2 fish was substantially different between the two models, which is a consequence of the survival probability being close to its upper threshold of 1 when no bias was imposed, because the abundance of age 2 fish was so close to the abundance of age three fish.

The asymptotic per capita growth rate was nearly identical between the models with and without systematic detection bias (C5). The partitioning of variance in the per capita growth rate was also very similar between the two (C6), although with a greater contribution from age 2 survival for the model with the bias. Nonetheless, for both models the survival probability of age 3 fish was more important than the survival probability of the other two age classes, and the contribution of recruitment was greatest overall.

Time-varying recruitment and survival for the reduced models

Figures C7-8 compare the estimated demographic rates from the full model to each of the reduced models allowing either recruitment (Figure C7) or survival (Figure C8) to vary through time. The pattern of temporal variation in recruitment was generally similar between the full model and Variable Recruitment, although the latter failed to detect the large decline in

recruitment over the last decade of the time series. The estimates of survival from the full model and Variable Survival were quite similar, although the latter had much greater uncertainty in the estimated survival probability of adults.

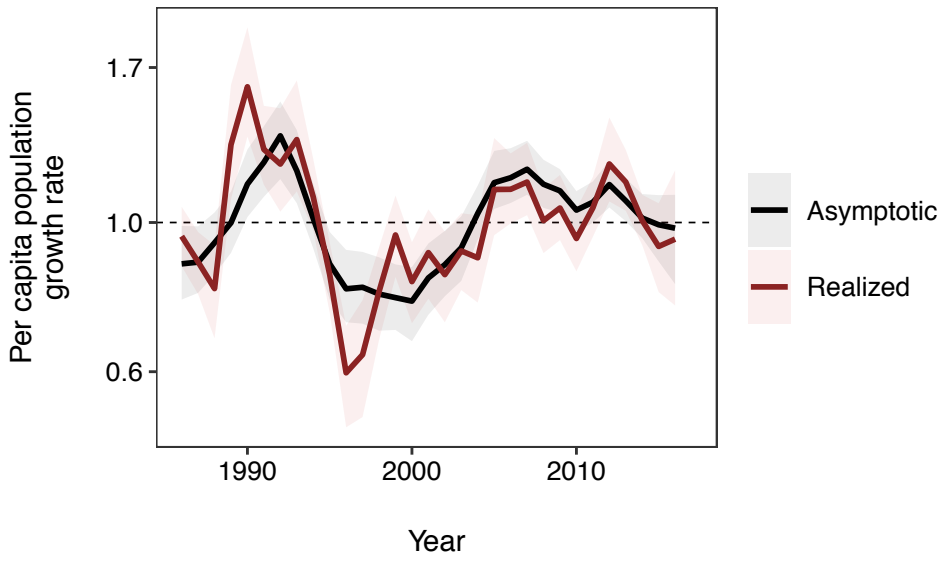

Figure C1. Realized (λ_t^{real}) and asymptotic (λ_t^{asym}) per capita growth rates for the full model. The shaded regions correspond to $UI_{68\%}$.

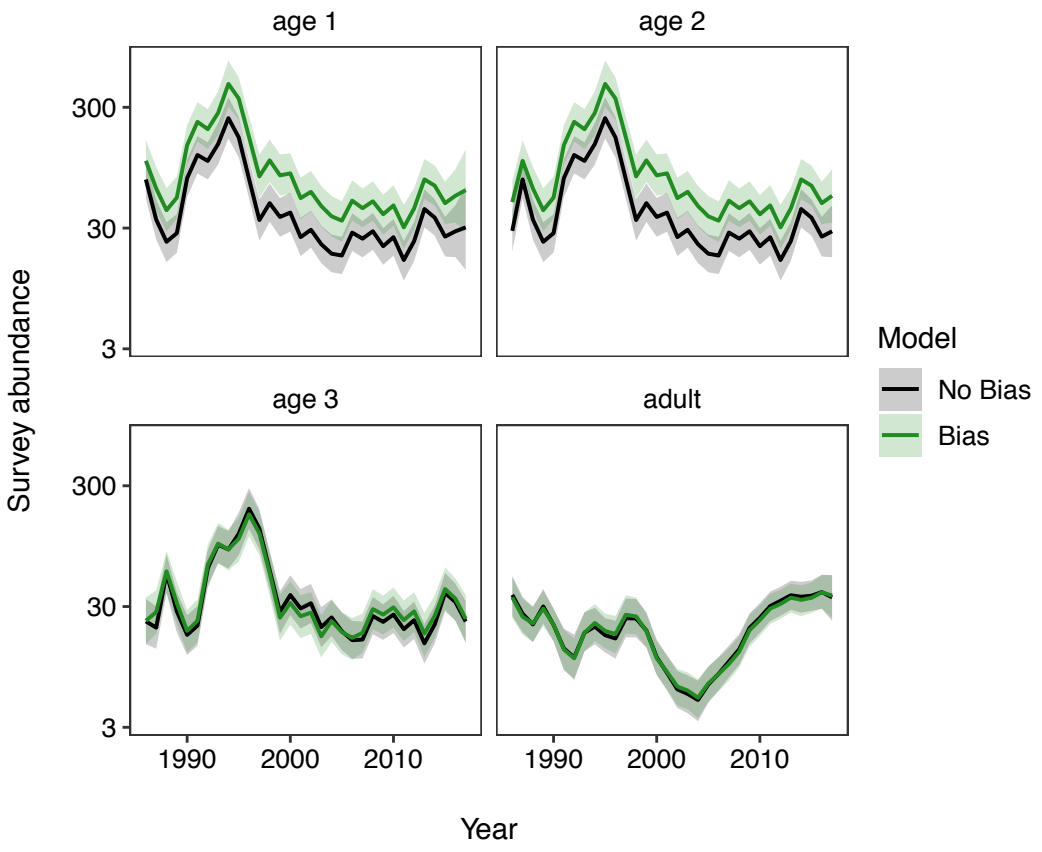

Figure C2. Estimated abundance from the models with and without the systematic detection bias. The shaded regions correspond to $UI_{68\%}$.

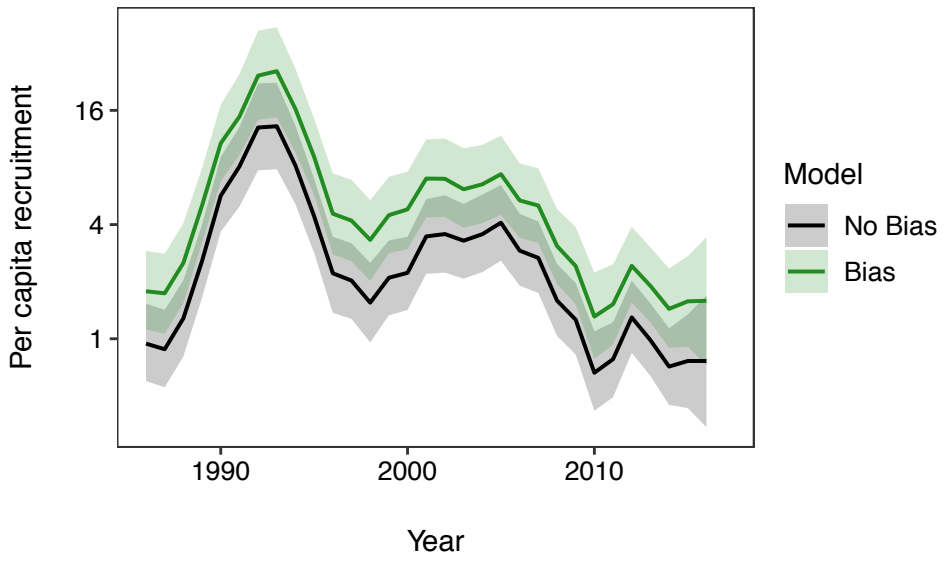

Figure C3. Estimated per capita recruitment from the models with and without the systematic detection bias. The shaded regions correspond to $UI_{68\%}$.

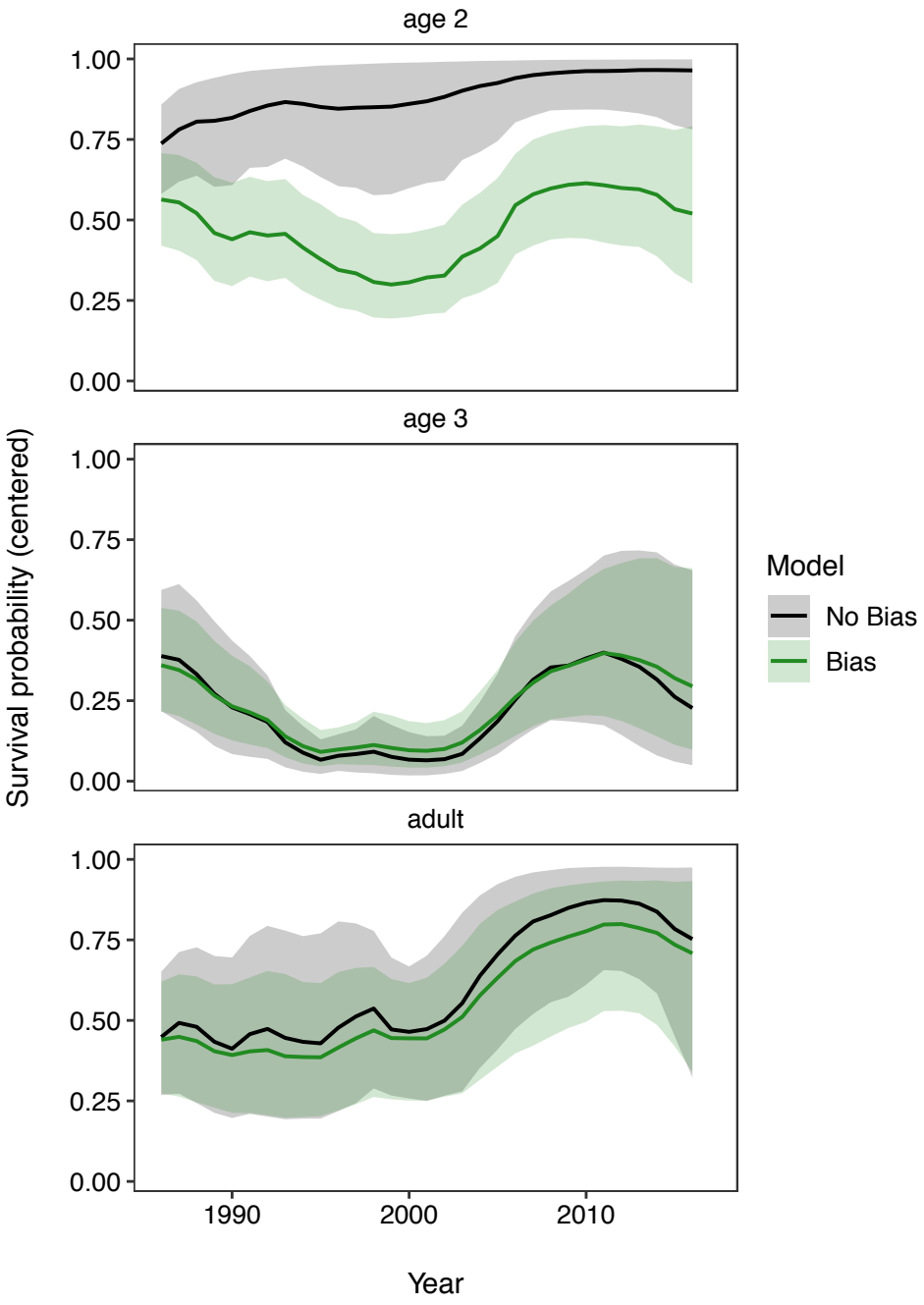

Figure C4. Estimated survival probabilities from the models with and without the systematic detection bias. Unlike in the Chapter 2 figure 4, the survival probabilities are not centered on their means, to highlight the difference between the overall magnitudes of the estimates between the two models. The shaded regions correspond to $UI_{68\%}$.

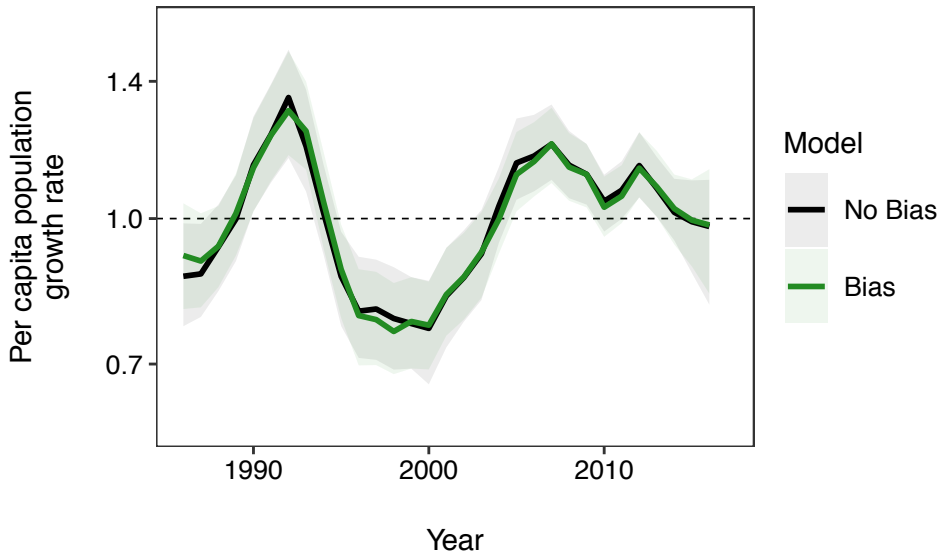

Figure C5. Asymptotic per capita growth rates from the models with and without the systematic detection bias. The shaded regions correspond to $UI_{68\%}$.

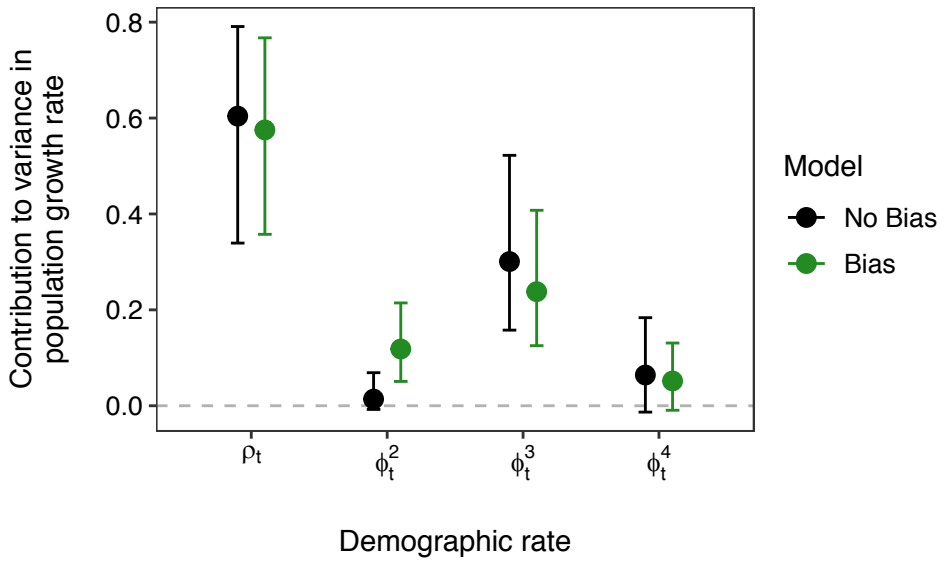

Figure C6. Partitioning of variance in the asymptotic per capita growth rate (λ_t^{asym}) into contributions from time-varying recruitment (ρ_t) and survival (ϕ_t^i), from the models with and without the systematic detection bias. The error bars correspond to $UI_{68\%}$.

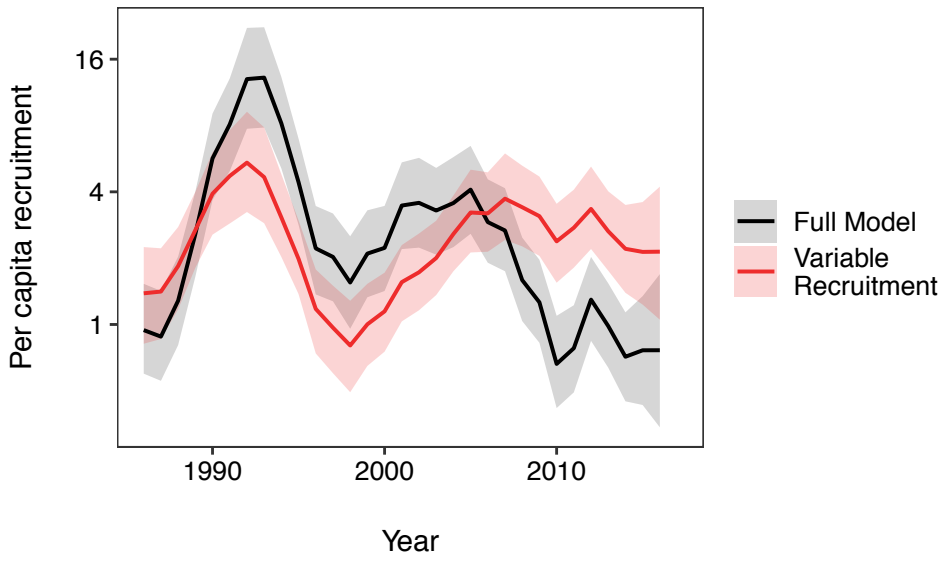

Figure C7. Estimated per capita recruitment from the full model and reduced model with variable recruitment. The shaded regions correspond to $UI_{68\%}$.

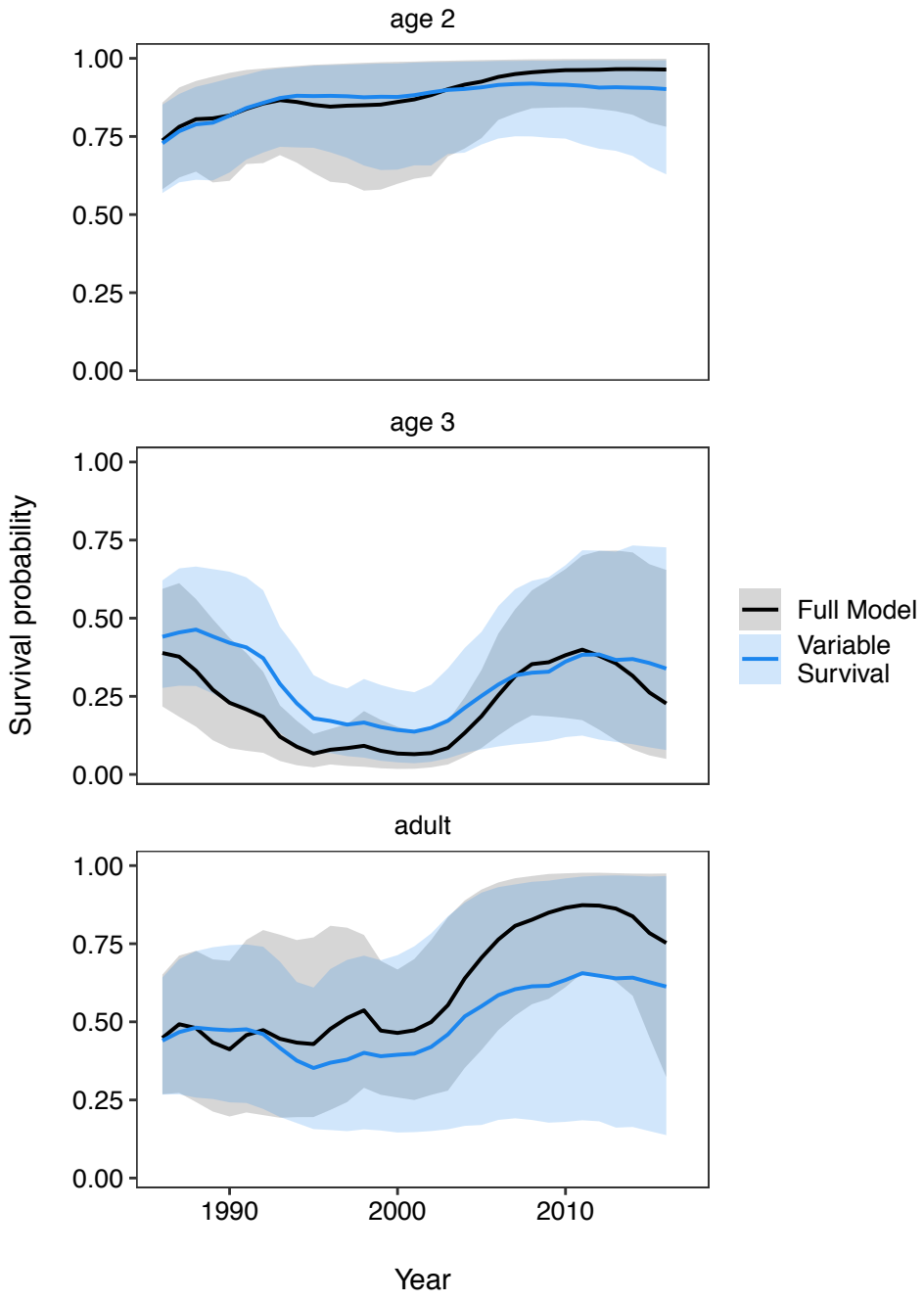

Figure C8. Estimated survival probabilities from the full model and reduced model with variable survival. The shaded regions correspond to $UI_{68\%}$.