Spring 2022

Evolutionary Dynamics of Rapid, Microgeographic Adaptation in an Amphibian Metapopulation

Adam Zachary Andis Arietta
Yale University Graduate School of Arts and Sciences, azandis@gmail.com

Abstract

Evolutionary Dynamics of Rapid, Microgeographic Adaptation in an Amphibian Metapopulation

Adam Zachary Andis Arietta

2022

Wild organisms can rapidly adapt to changing environments, even at fine spatial scales. This fact prompts hope that contemporary local adaptation may buffer some of the negative anthropogenic impacts to ecosystems. However, there are limits to the pace of adaptation. Understanding the adaptive potential—and limitations—of individual species at fine-resolution is an important task if we hope to accurately predict the repercussions of future climate and landscape change on biodiversity. My dissertation takes advantage of an uncommonly long-observed and closely-studied system to paint a comprehensive picture of evolution over time in association with shifts in ecological contexts.

In this dissertation, I show evidence of rapid, microgeographic evolution in response to climate within a metapopulation of wood frogs (*Rana sylvatica*). Critically, I show that populations separated by tens to hundreds of meters—well within the dispersal ability of the species—exhibited considerable shifts in development rates over a period of two decades, or roughly 6-9 generations. Using historical climate data and new methods of assessing landscape change, I show that these changes were mainly a response to warming climates.

The ecological contexts experienced by the metapopulation are associated with the evolution of physiological rates. Specifically, I show that climate change seems to have caused a counter-intuitive delay in spring breeding phenology while drought and warming

later in the larval development period correspond with a shift toward earlier metamorphosis. The picture that emerges is of a contracting developmental window, which is expected to select for faster intrinsic development rates.

Superimposed on the metapopulation-wide shift to faster development was a pattern of counter-gradient variation reflecting a similar pattern seen two decades prior. Furthermore, I empirically demonstrate a trade-off between faster development and a swimming performance trait that strongly contributes to fitness. This trade-off helps to explain why intrinsic development rates vary spatially with pond temperatures, but in the opposite direction of the relationship with temperature over time.

Though the evidence for rapid adaptation to climate change presented in this dissertation reveals that evolution can buffer populations from extinction, it also entreats caution. There is a clear trend of demographic decline among wood frog populations that experienced greater magnitudes of environmental change. In fact, the three populations that suffered local extinctions over the 20-year course of observations inhabited ponds characterized by the greatest change in temperature or canopy.

Evolutionary Dynamics of Rapid, Microgeographic Adaptation in an Amphibian Metapopulation

A Dissertation

Presented to the Faculty of the Graduate School

of

Yale University

In Candidacy for the Degree of

Doctor of Philosophy

By

Adam Zachary Andis Arietta

Dissertation Director: Dr. David K. Skelly

May 2022

In memory of my father, Arthur “Andy” F. Andis Jr. (1963-2004),
who taught me that anything in nature can inspire awe
if you are willing to look at it closely enough for long enough.

ACKNOWLEDGEMENTS

It is somewhat of a miracle that I ended up completing a doctoral degree. I have many people to thank for their roles, not only in achieving this milestone, but all of the preceding steps of my life and career that led me here.

First and foremost, I must recognize my dad, in whose memory this dissertation is dedicated. For a guy who never went to college, my dad always seemed to have the answer to any natural science question I could think to ask. I remember once sitting on his workbench while he drew out a Punnett square on a piece of scrap wood in an attempt to explain to me the principles of heredity and albinism. I was probably no older than eight at the time. He taught me that anything in nature will inspire ineffable awe if you are willing to look at it closely for long enough. That lesson rang true often during my research, especially during late nights watching frog embryos divide under the heat of the microscope lamps or while stooped in freezing water searching for the tiniest movement among the contents of my dipnet. The path of my dad's life was marked by wrong-turns and false-starts. While I wish he had navigated his life better, I have come to appreciate that his missteps enabled many of the opportunities that allowed me to map a clearer course for myself.

Secondly, I have to thank my partner, Bayla, for her enduring support and encouragement. We moved together from the fresh, steep peaks of Alaska to the continental mountains of Montana, then to the ancient rolling hills of the East Coast. Along the way, we picked up three degrees, two kids, a grumpy parrot, and a marriage certificate. In my wedding vows, I likened our love to the geology of the mountains we'd travelled through. I said that, even though the storms of life we would weather in the future might leave their marks, I could not wait to stand the test of time together—after all, a sunrise over the ancient and eroded Appalachians is no less beautiful than a sunset over the Pacific Coast

Range. This dissertation has certainly added a few wrinkles to the topology of our life together. Thank you for standing beside me through all of it.

Third, I have to thank my advisor David Skelly for his ever-equable mentorship. I could not have asked for a better advisor. He trusted me when I doubted myself, questioned me until my ideas were cogent, and encouraged me to always think bigger and broader. Over the last six years, I came to appreciate the advice he offered on life just as much as his thoughts on science. As with any great advisor, I think I have learned more from observing him as a mentor than his direct advice. I aspire to his level of patience, thoughtfulness, and affability.

None of my work could have been accomplished without Adriana Rubinstein who, in addition to her full-time task of keeping the lab running, put in many tireless hours toward my projects. Looking back on my PhD experience, some of my fondest memories are the times laughing with Adriana while driving back and forth from Yale Myers Forest, cleaning countless tadpole jars, and traipsing through frog ponds.

For me, the most fulfilling part of academia is the cast of amazing folks I have gotten to share wild (often foolish) ideas with—some of whom even took the bait and ended up as collaborators. I was fortunate to work with Kaija Gahm on a study that became an integral chapter of this dissertation. I hope to one day match her drive, enthusiasm, and organization. I think Yara Alshwairikh gets the prize for inspiring and enabling the greatest number of my fool-hardy ideas. Whether we were 10 days deep into the Alaska wilderness acting as mosquito food or months into a lab experiment pestering frogs, she emanated an unending positivity that I truly appreciate. Dr. Nate Edelman also deserves special thanks for the many hours he spent patiently coaching me on population genetics and Unix or

commiserating over the academic job market. Nate's optimism has been a perfect counter to my curmudgeonly pessimism.

Of course, the entirety of the Skelly Lab, past and present, have been a source of support and inspiration. It has been a pleasure to work closely with Logan Billet, Dr. Kealoha Freidenburg, Dahn-Young Dahn, Dr. Freya Rowland, Ryan Dougherty, Cesar Garcia Lopez, Brandon Sanchez, Kamau Walker, Paige Johnson, Dr. Jonathan Richarson, and Dr. Max Lambert. I also feel especially fortunate to have had a huge "extended family" here at Yale beyond the Skelly Lab, including my incoming cohort-mates, particularly Dan Kane, Paul Burow, Mary Burak, and Julia Monk. I thank my "Greeley Fam," including Dr. Annise Dobson, Dr. Rob Buchkowski, Kristy Ferraro, Nathalie Sommer, Dr. Noah Sokol, Chris Beltz, Dr. Adam Roddy, and all the folks in the Brodersen and Schmitz labs, for all of the many parties, conversations, and commiserations. In particular, I thank Sam Jordan, whose friendship is one of my most cherished outcomes from my time here at Yale Forestry and Environmental Studies.

The Ecology and Evolutionary Biology Department was often my home away from home. I thank Brooke Bodensteiner, The Eco-Evo Goonies (Dr. Franz Simon, Dr. Stephen Gaughran, Matthew Dougherty, Dr. Dan MacGuigan, plus Mary B.), and all of the folks in the Muñoz, Near, and Caccone labs for making it such a welcoming place.

It was a privilege to be able to work at the Peabody Museum of Natural History during my PhD. I learned a tremendous amount about museums and biology from the folks in the Vertebrate, Entomology, Invertebrate, and Botany divisions. I have to reserve special thanks for Greg Watkins-Colwell who became a surrogate mentor, ad hoc committee member, and dear friend over many hours sorting jars of mystery specimens, digging up sea turtle skeletons, and chatting over the prep table or back in the collections. I also thank Eric

Lazo-Wasem in Entomology for teaching me photomicrography and Dr. Larry Gall for helping me learn the museum cataloging software.

There are many people whose efforts behind-the-scenes made my work so much easier, including Elisabeth Barsa, Sara-Smiley Smith, Carol Mariani, Phillis Peterkins, Steve Prinn, Susan Bolden, Rachel Radin, and Joseph Orefice. I also thank Dr. Rob Denton, Dr. Lauren Oakes, Dr. Sam Snow, Dr. Marlyse Duguid, Ben Hamilton, Andrew Thoms, Dr. Dan Bolnick, Dr. Rowan Barrett, Marc-Olivier Beausoleil, Tim Thurman, David Hunt, Chandra Brown, Gabriel Garber, Jordan Wicker, Brett Simmons, and Megan McCormick for their support, academic or otherwise, along the way. I appreciate the advice of my committee members, Dr. Andrew Hendry, Dr. Tom Near, and Dr. Casey Dunn, and thank them whole-heartedly for guiding this dissertation.

Finally, I have to thank J and baby K. It was not until you two came along that I finally realized the true importance of conserving for future generations. Your addition to our family was also when I realized that dissertations don't need to be perfect, they just need to be done, because there are simply more important things in life waiting at home.

TABLE OF CONTENTS

ABSTRACT		*i*
TITLE PAGE		*iii*
DEDICATION		*v*
ACKNOWLEDGEMENTS		*vi*
TABLE OF CONTENTS		*x*
LIST OF TABLES AND FIGURES		*xi*
INTRODUCTION		1
CHAPTER 1:	Estimation of forest canopy structure and understory light using spherical panorama images from smartphone photography	14
CHAPTER 2:	Phenological delay despite warming in wood frog (*Rana sylvatica*) reproductive timing: a 20-year study	52
CHAPTER 3:	Climate change mediated effects on development and life history in an amphibian	84
CHAPTER 4:	Temperature-mediated tradeoff between development and performance in larval wood frogs (*Rana sylvatica*)	135
CHAPTER 5:	Rapid microgeographic evolution in response to climate change	160
CONCLUSION		210

LIST OF TABLES AND FIGURES

CHAPTER 1

Table 1 -	Comparisons of canopy metrics estimated from smartphone spherical panorama hemispherical photography to other methods	37
Table 2 -	Effects of downsampling smartphone spherical panoramas on canopy metrics	38
Figure 1 -	Representation of spherical panoramas projected into hemispherical photos	39
Figure 2 -	Comparisons of images generated with smartphone spherical panorama photography to other methods	40
Figure 3 -	Regression models comparing image methods	41
Figure 4 -	Differences in canopy metrics across image methods	42
Figure 5 -	Relative gaps size in relation to canopy openness	43
Appendix -	Supplemental materials	44

CHAPTER 2

Table 1 -	Seasonal trends in meteorological variables	74
Figure 1 -	Map of research sites	75
Figure 2 -	Trends in temperature, precipitation, and snow water equivalent 1980-2018	76
Figure 3 -	Seasonal trends and relative change in climate variables 1980-2018	77
Figure 4 -	Changes in oviposition date from 2000 to 2019	78
Figure 5 -	Hindcast model of oviposition dates since 1980	79
Appendix -	Supplemental material	80

CHAPTER 3

Table 1 -	Piecewise linear mixed effects models of physiological rates	114
Figure 1 -	Trends in ecological variables	115
Figure 2 -	Comparison of historical and contemporary measurements of embryo and larval size	116
Figure 3 -	Model estimates of allocation and developmental rate change 1999 to 2019	117
Appendix -	Supplemental material	118

CHAPTER 4

Table 1 -	Linear mixed effect model of swimming performance	156
Figure 1 -	Development rates and recorded temperatures of ponds and incubators	157
Figure 2 -	Geometric morphometric warps of larvae	158
Figure 3 -	Relationship between development rate and burst speed	159

CHAPTER 5

Table 1 -	Model used to predict embryonic development rates	187
Table 2 -	Model relating embryonic period to environmental gradients	188
Table 3 -	Results of structural equation model predicting the pond-wise change in embryonic period	189
Table 4 -	Model relating population growth to environmental change	190
Figure 1 -	Map of study sites, population growth rates, and environmental change	191
Figure 2 -	Embryonic period across environmental gradients	192
Figure 3 -	Change in embryonic period and volume	193
Figure 4 -	Change in population growth related to environmental change	194
Appendix -	Supporting material	195

INTRODUCTION

Organisms adapt over time as ecological change coerces them into novel environments and new selection regimes. Given enough time and space, populations adaptively react to such changes by shifting average trait values towards those that optimize fitness in the new environmental situation (Endler, 1986). It is this fine-tuning mechanism that has led to the diversity of species inhabiting the near and far reaches of the globe and permitted the continuation and proliferation of living communities.

The generative process of evolution is driven by the subtractive act of selection. Not all lineages persist. At times, selection can be so great as to lead to population collapse, as has occurred multiple times throughout biotic history in widespread response to rapid global change (Jablonski and Chaloner, 1994). Increasingly, human-derived impacts predominate the dynamics of environmental change, superseding natural patterns and placing new stressors on organisms and their communities (Vitousek et al., 1997; Sala et al., 2000). The effects of anthropogenic impacts are readily apparent (Walther et al., 2002; Parmesan and Yohe, 2003; Pereira et al., 2010; Cohen et al., 2018) and have lead some authors to ask if humans are engineering a sixth mass extinction (Barnosky et al., 2011). There is plenty of evidence to warrant gloom, and yet, in the augural words of Jurassic Park's Dr. Ian Malcom, "Life, uh, finds a way." Recent re-thinking about the pace of evolution offers a glimmer of hope—and potentially a new axis of impact—for some species (Reznick et al., 2019).

Traditionally, evolution was thought to operate on protracted timescales over hundreds or thousands of generations (Slobodkin, 1961). Darwin, himself, set the tone for conceptualizing the pace of evolution. He assumed evolution to be "extremely slow" overall, remarking on the unobservable (at the time) quality of evolutionary progression over the ages (Darwin, 1964, 302). Subsequent scholars refined the theme. Slobodkin (1961)

characterized rates of change on "evolutionary time" greater than 10 generations in contrast to shorter, "ecological time" in which evolution was unimportant. Similarly, Dobzhansky (1937) defined micro- and macroevolutionary time frames in contrast to shorter generational time frames.

No doubt, the species Darwin encountered on his daily walks were already in the process of evolving in response to human impacts. In the century after Darwin published *Origins*, iconic studies showed adaptive shifts in melanistic moth morphs in response to the Industrial Revolution (Haldane, 1956; Kettlewell, 1961), resistance to fertilizer treatment (Snaydon and Davies, 1972; Davies et al., 1976) and toxic mine tailing (Jain and Bradshaw, 1966; Antonovics, 1971, 2006) in plants, morphological divergence of mice introduced to islands (Berry, 1964) and of sparrows introduced to North America (Johnston and Selander, 1964), and apple maggot's preference for introduced host plants (Bush, 1969), all occurring within the previous 50 to 100 years.

More recent research has shown that populations can evolve rapidly, even within so-called ecological time frames (Thompson, 1998, 2013; Hendry and Kinnison, 1999; Hairston et al., 2005; Carroll et al., 2007; Fussmann et al., 2007; Hendry, 2016). Such examples have arisen across the animal kingdom in wild populations of invertebrates (Coldsnow et al.; Seeley, 1986; Carroll et al., 1997; Huey et al., 2000; Zuk et al., 2006), fish (Stearns, 1983; Reznick et al., 1997; Hendry et al., 2000; Kavanagh et al., 2010; Whitehead et al., 2017), amphibians (Phillips et al., 2006), reptiles (Campbell-Staton et al., 2017; Donihue et al., 2018), birds (Grant and Grant, 1995), and mammals (Epstein et al., 2016). The phenomenon of evolutionary change, once thought to occur over hundreds of human generations, has now been recapitulated in even short-term experiments (Barrett et al., 2011; Stuart et al., 2014).

With the proliferation of focused studies, we are now seeing more and more species

rapidly adapting to environmental conditions of which we are the creators (Palumbi, 2001; Darimont et al., 2009; Otto, 2018). The magnitude and rapidity of this change is unprecedented (Stuart et al., 2010; Barnosky et al., 2011), but the evidence of rapid adaptation offers some hope that species may be more resilient than our expectations. Evolution may be a means of circumventing extinction, a process—when successful—termed "evolutionary rescue" (Bell, 2013). There is further hope that evolution can occur at extremely small spatial scales (Richardson et al., 2014), such that even populations fragmented by human land use may harbor adaptive potential to avoid collapse (Cheptou et al., 2017).

Yet, we still see populations decline and disappear, suggesting that there must be limits to rapid evolutionary rescue. While the complexity and inherent unpredictability of the future of evolutionary processes under anthropogenic duress may be inscrutable, the task of modern conservation science should be to estimate and mitigate such impacts (Myers and Knoll, 2001; Barnosky et al., 2017). To do so, researchers must integrate contemporary rates of change into the frameworks of conservation biology and evolutionary theory to ask, what contexts allow for ecologically relevant rapid evolution? And alternatively, what are the limits to rapid adaptation with respect to evolutionary rescue?

To answer these questions in this dissertation, I use wood frogs (*Rana sylvatica*) as a model system. Wood frogs are an emerging model system for understanding evolution and ecology. The species occupies the largest native range of any amphibian, spanning extremely cold habitat in the Arctic to mild climates in the southeastern United States (Martof and Humphries, 1959). These cosmopolitan range limits are enabled by unique physiological adaptations to climate. Wood frogs are well-known for the ability to enter a state of suspended animation to survive through winter months (Storey, 1990). The most

northern populations can persist in this state for up to 6 months and tolerate temperatures as low as -18.1 °C (Larson et al., 2014). Wood frog embryos continue to develop at some of the coldest temperatures reported for Ranid species and can even survive super cooling (Frisbie et al., 2000). The larval stage, although not freeze-tolerant, exhibits a strongly cold-biased thermal performance curve (Cupp, 1980). These adaptations allowed the species to colonize the Arctic north when the Pleistocene glaciers receded as no other amphibians could (Lee-Yaw et al., 2008). Now that climate change is increasing temperatures on average, these remarkable physiological adaptations may be a liability for the species (Fitzpatrick et al., 2020). Hence, wood frogs provide an excellent natural system to consider the biotic effects of contemporary climate change.

Other life-history attributes of wood frogs make the species a particularly tractable model species. Wood frogs are explosive, synchronous breeders (Waldman, 1982; Petranka and Thomas, 1995) with propensity for philopatry (Berven and Grudzien, 1990) which results in relatively discrete populations. Larvae must remain isolated within natal ponds for much of the year. Most mortality, and perforce selection, occurs in this aquatic stage (Seigel, 1983; Berven, 1990). In fact, mass mortality events, due to desiccation, disease, or predation are common (Berven, 1995). This, in addition to their relatively short generation time of one to three years (Berven, 1990), makes it simple to identify and sample individual populations and natal habitat features driving natural selection. After larvae metamorphose, terrestrial dispersal allows for gene flow between populations (Berven and Grudzien, 1990; Newman and Squire, 2001). Given that all of the populations included in my research are located in very close proximity, this system also allows for tests of microgeographic evolution.

For this dissertation, I focus specifically on a metapopulation of wood frogs at Yale Myers Forest. About 60 of these populations have been consistently studied since the late

1990s providing a wealth of historical data with which to piece together the ecological and evolutionary history of this metapopulation. Most importantly, this metapopulation has been the focus of many studies demonstrating evidence of evolutionary divergence at small spatial scales with respect to development and growth rates (Skelly, 2004; Ligon and Skelly, 2009), thermal preference and performance (Skelly and Freidenburg, 2000; Freidenburg and Skelly, 2004), and morphology and predator avoidance (Urban et al., 2017).

Of the many prior studies demonstrating evolutionary divergence, canopy closure is one of the most commonly cited environmental correlates (e.g. Freidenburg and Skelly, 2004; Skelly, 2004). As both their common and Linnaean names suggest, wood frogs are strongly associated with forested environments. Given the ecological importance of forest habitat for wood frog evolution, it is important to be able to accurately estimate canopy structure. However, traditional methods of estimating canopies using hemispherical photos captured with DSLR cameras are expensive and tedious to deploy. In chapter 1, I develop a new method of estimating canopy closure using spherical panoramic images from smartphones. I found this method to out-perform traditional methods. For chapters 2 and 3, I use canopy estimates and historical data to examine the ecological effects of climate and habitat change on frog populations in the wild.

One of the most often reported biotic effects of climate change is advances in spring phenology associated with warmer weather (Parmesan and Yohe, 2003; While and Uller, 2014; Cohen et al., 2018). In chapter 2, I show that annual temperatures have increased considerably at Yale Myers Forest in recent decades. However, rather than an advance in spring breeding phenology, I found a counter-intuitive delay in the date at which wood frogs lay their eggs. Considering climate variables across the season rather than as annual averages illuminates a possible cause. Although temperatures increased overall, most of the warming effect is seen later in the year while the temperatures during oviposition warmed

relatively less. Meanwhile, average snowpack remained constant over the years, but shifted to persist later into spring, effectively blocking the frogs from breeding ponds.

In chapter 3, I consider the effects of phenological shifts for the larval stage of wood frogs. I compared museum specimens and historical records to contemporary field collected larvae to estimate changes in physiological rates (develop, growth, and allocation) and associate these with ecological correlates. I found that larval development rates have increased substantially. Allocation and growth rates have also shifted. The story that emerges from chapters 2 and 3 seems to be that climate change is compressing the seasonal window for aquatic stages and forcing the frogs to accelerate development. The physiological consequence of rapid development influences allocations rates, such that larvae tend to grow larger but ultimately metamorphose smaller which may lead to lower fitness in subsequent life stages (Berven, 1990).

In chapters 2 and 3, I consider phenotypes measured from wild specimens which are manifestations of both genetic and plastic responses to the environment and cannot be used to infer evolution. Although amphibians are remarkable in the breadth of plastic trait variation (Denver, 1997), prior research has shown that developmental rates are heritable, and diverge along environmental gradients in these populations (e.g. Skelly, 2004 and others). Most commonly reported is a pattern of countergradient variation with respect to canopy wherein individuals from closed-canopy ponds develop faster than conspecifics in open canopy ponds when raised in the same conditions. Countergradient variation is a well-known form of local adaptation seen in ectotherm populations along latitudinal gradients wherein elevated intrinsic development or growth rates in colder latitudes counteracts the lack of extrinsic thermal energy compared to warmer latitudes (Berven et al., 1979; Conover and Present, 1990; Laugen et al., 2003). However, countergradient variation offers somewhat of a paradox at microgeographic scales in which individuals can easily migrate

between all available habitats. For example, if faster development is advantageous, what keeps alleles conferring intrinsically accelerated development in dark ponds from conferring compound advantages when aligned with extrinsic factors in open ponds? One possible explanation is the existence of trade-offs wherein selection favors a trait at one end of the gradient but disfavors the trait at the opposite end. In chapter 4, I worked with an undergraduate student, Kaija Gahm, to test for a trade-off with rapid development. We induced either accelerated or depressed development rates in wood frog larvae and found that fast development was associated with poor burst-swimming performance—an important trait for avoiding predators.

Finally, in chapter 5, I explore the ultimate consequences of environmental change over time for populations of wood frogs at Yale Myers Forest. I repeated a common garden study conducted in 2001 seventeen years later to test for the persistence of microgeographic counter-gradient variation. Given the pattern of slower development in warmer ponds found among populations in 2001, one would expect warming temperatures to force more population to the performance-development rate trade-off demonstrated in chapter 4. If so, one would expect both an erosion of countergradient variation and decline in fitness. Instead, I found a similar pattern of microgeographic countergradient variation persisted. I also found that, over time, average intrinsic development rates had increased among all populations. Causal models indicated that shifts in development over time were driven primarily by changes in pond temperature. Thus, it appears that wood frogs have rapidly evolved in response to changing climates. However, this adaptive potential is not a silver bullet for circumventing the negative impacts of ecological change. Those populations that experienced the greatest environmental change exhibited the steepest population declines over time. In fact, the three populations that declined to local extinction during the experiments were those that saw the most environmental change.

In conclusion, I found that wood frogs exhibit profound plastic and rapid evolutionary responses to environmental change. And yet, the species cannot buffer the most extreme changes in habitats that broach physiological limits. We can draw lessons from this system to temper our prospectus for biotic responses to the current climate crisis. On one hand, some species may show remarkable ability to accommodate and even adapt to environmental change. On the other hand, there are limits to evolutionary rescue. Without more research we cannot know which species might adapt and which are perilously close to tipping past the threshold toward extinction.

REFERENCES

Antonovics, J. (1971). The effects of a heterogeneous environment on the genetics of natural populations. *Am. Sci.* 59, 593–599.

Antonovics, J. (2006). Evolution in closely adjacent plant populations X: long-term persistence of prereproductive isolation at a mine boundary. *Heredity* 97, 33–37.

Barnosky, A. D., Hadly, E. A., Gonzalez, P., Head, J., Polly, P. D., Lawing, A. M., et al. (2017). Merging paleobiology with conservation biology to guide the future of terrestrial ecosystems. *Science* 355. doi:10.1126/science.aah4787.

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., et al. (2011). Has the Earth's sixth mass extinction already arrived? *Nature* 471, 51–57.

Barrett, R. D. H., Paccard, A., Healy, T. M., Bergek, S., Schulte, P. M., Schluter, D., et al. (2011). Rapid evolution of cold tolerance in stickleback. *Proc. Biol. Sci.* 278, 233–238.

Bell, G. (2013). Evolutionary rescue and the limits of adaptation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368, 20120080.

Berry, R. J. (1964). The evolution of an island population of the house mouse. *Evolution* 18, 468–483.

Berven, K. A. (1990). Factors affecting population fluctuations in larval and adult stages of the wood frog (*Rana sylvatica*). *Ecology* 71, 1599–1608.

Berven, K. A. (1995). Population regulation in the wood frog, *Rana sylvatica*, in three diverse geographic localities. *Aust. J. Ecol.* 20, 385–392.

Berven, K. A., Gill, D. E., and Smith-Gill, S. J. (1979). Countergradient Selection in the Green Frog, Rana clamitans. *Evolution* 33, 609–623.

Berven, K. A., and Grudzien, T. A. (1990). Dispersal in the wood frog (Rana sylvatica):

implications for genetic populations structure. *Evolution* 44, 2047–2056.

Bush, G. L. (1969). Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (*diptera, tephritidae*). *Evolution* 23, 237–251.

Campbell-Staton, S. C., Cheviron, Z. A., Rochette, N., Catchen, J., Losos, J. B., and Edwards, S. V. (2017). Winter storms drive rapid phenotypic, regulatory, and genomic shifts in the green anole lizard. *Science* 357, 495–498.

Carroll, S. P., Dingle, H., and Klassen, S. P. (1997). Genetic differentiation of fitness-associated traits among rapidly evolving populations of the soapberry bug. *Evolution* 51, 1182–1188.

Carroll, S. P., Hendry, A. P., Reznick, D. N., and Fox, C. W. (2007). Evolution on ecological time-scales. *Funct. Ecol.* 21, 387–393.

Cheptou, P.-O., Hargreaves, A. L., Bonte, D., and Jacquemyn, H. (2017). Adaptation to fragmentation: evolutionary dynamics driven by human influences. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372. doi:10.1098/rstb.2016.0037.

Cohen, J. M., Lajeunesse, M. J., and Rohr, J. R. (2018). A global synthesis of animal phenological responses to climate change. *Nat. Clim. Chang.* 8, 224–228.

Coldsnow, K. D., Mattes, B. M., Hintz, W. D., and Relyea, R. A. Rapid evolution of tolerance to road salt in zooplankton. *Environ. Pollut.* doi:10.1016/j.envpol.2016.12.024.

Conover, D. O., and Present, T. M. C. (1990). Countergradient variation in growth rate: compensation for length of the growing season among Atlantic silversides from different latitudes. *Oecologia* 83, 316–324.

Cupp, P. V. (1980). Thermal Tolerance of Five Salientian Amphibians during Development and Metamorphosis. *Herpetologica* 36, 234–244.

Darimont, C. T., Carlson, S. M., Kinnison, M. T., Paquet, P. C., Reimchen, T. E., and Wilmers, C. C. (2009). Human predators outpace other agents of trait change in the wild. *Proc. Natl. Acad. Sci. U. S. A.* 106, 952–954.

Darwin, C. (1964). *On the Origin of Species: A Facsimile of the First Edition*. Cambridge, Massachusetts: Harvard University Press.

Davies, M. S., Snaydon, R., and Others (1976). Rapid population differentiation in a mosaic environment. *Heredity* 36, 59–66.

Denver, R. J. (1997). Proximate Mechanisms of Phenotypic Plasticity in Amphibian Metamorphosis. *Integr. Comp. Biol.* 37, 172–184.

Dobzhansky, T. (1937). *Genetics and the Origin of Species*. New York, NY: Columbia University Press.

Donihue, C. M., Herrel, A., Fabre, A.-C., Kamath, A., Geneva, A. J., Schoener, T. W., et al. (2018). Hurricane-induced selection on the morphology of an island lizard. *Nature* 560, 88–91.

Endler, J. A. (1986). *Natural Selection in the Wild*. Princeton University Press.

Epstein, B., Jones, M., Hamede, R., Hendricks, S., McCallum, H., Murchison, E. P., et al. (2016). Rapid evolutionary response to a transmissible cancer in Tasmanian devils. *Nat. Commun.* 7, 12684.

Fitzpatrick, M. J., Porter, W. P., Pauli, J. N., Kearney, M. R., Notaro, M., and Zuckerberg, B. (2020). Future winters present a complex energetic landscape of decreased costs and reduced risk for a freeze-tolerant amphibian, the Wood Frog (Lithobates sylvaticus). *Glob. Chang. Biol.* doi:10.1111/gcb.15321.

Freidenburg, L. K., and Skelly, D. K. (2004). Microgeographical variation in thermal preference by an amphibian. *Ecol. Lett.* 7, 369–373.

Frisbie, M. P., Costanzo, J. P., Lee, Jr., and Richard, E. (2000). Physiological and ecological aspects of low-temperature tolerance in embryos of the wood frog, Rana sylvatica. *Can. J. Zool.* 78, 1032–1041.

Fussmann, G. F., Loreau, M., and Abrams, P. A. (2007). Eco-evolutionary dynamics of communities and ecosystems. *Funct. Ecol.* 21, 465–477.

Grant, P. R., and Grant, B. R. (1995). Predicting microevolutionary responses to directional selection on heritable variation. *Evolution* 49, 241–251.

Hairston, N. G., Ellner, S. P., Geber, M. A., Yoshida, T., and Fox, J. A. (2005). Rapid evolution and the convergence of ecological and evolutionary time. *Ecol. Lett.* 8, 1114–1127.

Haldane, J. B. S. (1956). The theory of selection for melanism in *Lepidoptera*. *Proc. R. Soc. Lond. B Biol. Sci.* 145, 303–306.

Hendry, A. P. (2016). *Eco-Evolutionary Dynamics*. Princeton University Press.

Hendry, A. P., and Kinnison, M. T. (1999). Perspective: The pace of modern life: Measuring rates of contemporary microevolution. *Evolution* 53, 1637–1653.

Hendry, A. P., Wenburg, J. K., Bentzen, P., Volk, E. C., and Quinn, T. P. (2000). Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science* 290, 516–519.

Huey, R. B., Gilchrist, G. W., Carlson, M. L., Berrigan, D., and Serra, L. (2000). Rapid evolution of a geographic cline in size in an introduced fly. *Science* 287, 308–309.

Jablonski, D., and Chaloner, W. G. (1994). Extinctions in the Fossil Record [and Discussion]. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 344, 11–17.

Jain, S. K., and Bradshaw, A. D. (1966). Evolutionary divergence among adjacent plant populations I. The evidence and its theoretical analysis. *Heredity* 21, 407.

Johnston, R. F., and Selander, R. K. (1964). House sparrows: rapid evolution of races in North America. *Science* 144, 548–550.

Kavanagh, K. D., Haugen, T. O., Gregersen, F., Jernvall, J., and Vøllestad, L. A. (2010).

Contemporary temperature-driven divergence in a Nordic freshwater fish under conditions commonly thought to hinder adaptation. *BMC Evol. Biol.* 10, 350.

Kettlewell, H. B. D. (1961). The phenomenon of industrial melanism in *Lepidoptera*. *Annu. Rev. Entomol.* 6, 245–262.

Larson, D. J., Middle, L., Vu, H., Zhang, W., Serianni, A. S., Duman, J., et al. (2014). Wood frog adaptations to overwintering in Alaska: new limits to freezing tolerance. *J. Exp. Biol.* 217, 2193–2200.

Laugen, A. T., Laurila, A., Rasanen, K., and Merila, J. (2003). Latitudinal countergradient variation in the common frog (*Rana temporaria*) development rates - evidence for local adaptation. *J. Evol. Biol.* 16, 996–1005.

Lee-Yaw, J. A., Irwin, J. T., and Green, D. M. (2008). Postglacial range expansion from northern refugia by the wood frog, Rana sylvatica. *Mol. Ecol.* 17, 867–884.

Ligon, N. F., and Skelly, D. K. (2009). Cryptic divergence: countergradient variation in the wood frog. *Evol. Ecol. Res.* 11, 1099–1109.

Martof, B. S., and Humphries, R. L. (1959). Geographic Variation in the Wood Frog Rana sylvatica. *Am. Midl. Nat.* 61, 350–389.

Myers, N., and Knoll, A. H. (2001). The biotic crisis and the future of evolution. *Proc. Natl. Acad. Sci. U. S. A.* 98, 5389–5392.

Newman, R. A., and Squire, T. (2001). Microsatellite variation and fine-scale population structure in the wood frog (*Rana sylvatica*). *Mol. Ecol.* 10, 1087–1100.

Otto, S. P. (2018). Adaptation, speciation and extinction in the Anthropocene. *Proc. Biol. Sci.* 285. doi:10.1098/rspb.2018.2047.

Palumbi, S. R. (2001). Humans as the world's greatest evolutionary force. *Science* 293, 1786–1790.

Parmesan, C., and Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421, 37–42.

Pereira, H. M., Leadley, P. W., Proença, V., Alkemade, R., Scharlemann, J. P. W., Fernandez-Manjarrés, J. F., et al. (2010). Scenarios for global biodiversity in the 21st century. *Science* 330, 1496–1501.

Petranka, J. W., and Thomas, D. A. G. (1995). Explosive breeding reduces egg and tadpole cannibalism in the wood frog, Rana sylvatica. *Anim. Behav.* 50, 731–739.

Phillips, B. L., Brown, G. P., Webb, J. K., and Shine, R. (2006). Invasion and the evolution of speed in toads. *Nature* 439, 803.

Reznick, D. N., Losos, J., and Travis, J. (2019). From low to high gear: there has been a paradigm shift in our understanding of evolution. *Ecol. Lett.* 22, 233–244.

Reznick, D. N., Shaw, F. H., Rodd, F. H., and Shaw, R. G. (1997). Evaluation of the Rate of

Evolution in Natural Populations of Guppies (Poecilia reticulata). *Science* 275, 1934–1937.

Richardson, J. L., Urban, M. C., Bolnick, D. I., and Skelly, D. K. (2014). Microgeographic adaptation and the spatial scale of evolution. *Trends Ecol. Evol.* 29, 165–176.

Sala, O. E., Chapin, F. S., 3rd, Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., et al. (2000). Global biodiversity scenarios for the year 2100. *Science* 287, 1770–1774.

Seeley, R. H. (1986). Intense natural selection caused a rapid morphological transition in a living marine snail. *Proc. Natl. Acad. Sci. U. S. A.* 83, 6897–6901.

Seigel, R. A. (1983). Natural Survival of Eggs and Tadpoles of the Wood Frog, Rana sylvatica. *Copeia* 1983, 1096–1098.

Skelly, D. K. (2004). Microgeographic countergradient variation in the wood frog, *Rana sylvatica*. *Evolution* 58, 160–165.

Skelly, D. K., and Freidenburg, L. K. (2000). Effects of beaver on the thermal biology of an amphibian. *Ecol. Lett.* 3, 483–486.

Slobodkin, L. B. L. B. (1961). Growth and regulation of animal populations. Holt, Rinehart and Winston,.

Snaydon, R. W., and Davies, M. S. (1972). Rapid population differentiation in a mosaic environment. II. Morphological variation in *Anthoxanthum odoratum*. *Evolution* 26, 390–405.

Stearns, S. C. (1983). The evolution of life-history traits in mosquitofish since their introduction to Hawaii in 1905: rates of evolution, heritabilities, and developmental plasticity. *Am. Zool.* 23, 65–75.

Storey, K. B. (1990). Life in a frozen state: adaptive strategies for natural freeze tolerance in amphibians and reptiles. *Am. J. Physiol.* 258, R559–68.

Stuart, S. N., Wilson, E. O., McNeely, J. A., Mittermeier, R. A., and Rodríguez, J. P. (2010). The Barometer of Life. *Science* 328, 177–177.

Stuart, Y. E., Campbell, T. S., Hohenlohe, P. A., Reynolds, R. G., Revell, L. J., and Losos, J. B. (2014). Rapid evolution of a native species following invasion by a congener. *Science* 346, 463–466.

Thompson, J. N. (1998). Rapid evolution as an ecological process. *Trends Ecol. Evol.* 13, 329–332.

Thompson, J. N. (2013). *Relentless Evolution*. Chicago: University of Chicago Press.

Urban, M. C., Richardson, J. L., Freidenfelds, N. A., Drake, D. L., Fischer, J. F., and Saunders, P. P. (2017). Microgeographic Adaptation of Wood Frog Tadpoles to an Apex Predator. *Copeia*, 451–461.

Vitousek, P. M., Mooney, H. A., Lubchenco, J., and Melillo, J. M. (1997). Human domination of

Earth's ecosystems. *Science* 277, 494–499.

Waldman, B. (1982). Adaptive significance of communal oviposition in wood frogs (Rana sylvatica). *Behav. Ecol. Sociobiol.* 10, 169–174.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., et al. (2002). Ecological responses to recent climate change. *Nature* 416, 389–395.

While, G. M., and Uller, T. (2014). Quo vadis amphibia? Global warming and breeding phenology in frogs, toads and salamanders. *Ecography* 37, 921–929.

Whitehead, A., Clark, B. W., Reid, N. M., Hahn, M. E., and Nacci, D. (2017). When evolution is the solution to pollution: Key principles, and lessons from rapid repeated adaptation of killifish (Fundulus heteroclitus) populations. *Evol. Appl.* 10, 762–783.

Zuk, M., Rotenberry, J. T., and Tinghitella, R. M. (2006). Silent night: adaptive disappearance of a sexual signal in a parasitized population of field crickets. *Biol. Lett.* 2, 521–524.

CHAPTER 1

Estimation of forest canopy structure and understory light using spherical panorama images from smartphone photography

Publication Status: Arietta, A. Z. (2021). Estimation of forest canopy structure and understory light using spherical panorama images from smartphone photography. *Forestry: An International Journal of Forest Research* cpb034. DOI: 10.1093/forestry/cpab034. (Chapter formatted to journal standards).

ABSTRACT

Accurate estimates of forest canopy structure are central for a wide range of ecological studies. Hemispherical photography (HP) is a popular tool to estimate canopy attributes. However, traditional HP methods require expensive equipment, are sensitive to exposure settings, and produce limited resolution which dramatically affects the accuracy of gap fraction estimates. As an alternative, hemispherical images can be extracted from spherical panoramas produced by many smartphone camera applications. I compared hemispherical photos captured with a digital single lens reflex camera and 180° lens to those extracted from smartphone spherical panoramas (SSP). The SSP HP method leverages built-in features of current generation smartphones to produce sharper images of higher resolution, resulting in more definition of fine canopy structure. Canopy openness and global site factor from SSP HP are highly correlated with traditional methods ($R^2 > 0.9$), while leaf area index estimates are lower, especially in more closed canopies where traditional methods fail to capture fine gaps.

INTRODUCTION

Ecological patterns and processes in forests are mediated by the canopy in critical ways (Parker et al. 1995). The structure of the canopy directly alters the below-canopy light regime and microclimate which indirectly impacts microhabitat features such as soil moisture, snowpack and understory plant community (Jennings et al. 1999). Because of its fundamental importance, forest managers and ecologists alike require methods to accurately quantify canopy structure and light environments.

In recent decades, hemispherical photography (HP), and especially digital HP, has arisen as the most popular optical method for indirectly estimating canopy structure, primed by advances in digital photography equipment and software for image analysis and its relatively lower cost compared to other methods (e.g. laser scanning) (Promis 2013; Yan et al. 2019; Chianucci 2020). This technique uses extreme wide-angle lenses with a field-of-view of 180° that projects an entire hemisphere of view onto the camera sensor, resulting in a circular hemispherical image (Rich 1990). Pixels within the image are then classified into binary sky (white) or canopy elements (black) manually or algorithmically using global or local thresholds (Glatthorn and Beckschäfer 2014). From the binarized images, canopy structure measures such as gap fraction, canopy openness (CO), leaf area index (LAI), etc. can be estimated (Frazer et al. 1997; Gonsamo et al. 2013; Chianucci 2020). By plotting a sun path onto the binarized image, the canopy structure can be used to infer understory light regimes, or Site Factors, given estimates of prevailing above-canopy direct and diffuse radiation (Anderson 1964). Light values can be integrated over time to yield seasonal estimates of light environments from a single sampling event (Frazer et al. 1997) (or at least two samples in deciduous canopies (e.g. Halverson et al. 2003)). Thus, HP offers an efficient, non-destructive method of estimating forest microhabitat features.

However, HP suffers from methodological drawbacks relating to the difficulty of capturing images in field settings and the sensitivity of estimates to variation in image acquisition (Beckschäfer et al. 2013; Bianchi et al. 2017). Traditionally, HP requires a single lens reflex camera (or more commonly nowadays, digital single lens reflex (DSLR)) equipped with a specialized hemispherical lens and self-levelling tripod. For the most accurate estimates, images must be acquired against uniformly overcast skies or the fleeting light at dusk or dawn with the camera level to the horizon and with closely calibrated exposure. The reliance on DSLRs stems from the need to manually fine-tune exposure and the advantage of large sensors. Large sensors record more pixels per area of view and more information per pixel, which translates to more accurate classification of canopy elements. Careful tuning of exposure settings is critical to avoid major inaccuracies in final estimates (Zhang et al. 2005; Beckschäfer et al. 2013).

Typical camera and hemispherical lens systems are expensive and challenging to deploy in the field. In response, researchers have attempted to develop new methods including using smartphones with clip-on lenses (Tichý 2016; Bianchi et al. 2017) and eschewing specialized levelling devices (Origo et al. 2017), to varying success. These smartphone-based methods are limited by small sensors and the low quality of aftermarket fisheye lenses. Other researchers have developed methods that use standard cameras with reduced field-of-view lenses and account for non-hemispherical images (i.e. restricted view photography) to estimate canopy structure with less sensitivity to camera exposure while maximizing the full frame of the sensor (Chianucci 2020). Yet, without a full hemisphere of view, this method cannot be used for estimating light environments. Thus, despite the fact that almost all ecologists and foresters carry a high-powered, image processing device in our pockets, we have yet to fully employ it for the purpose of data collection.

Here, I test a novel method of acquiring HP images from smartphone spherical panoramas (SSP HP) for estimating canopy measures. This method leverages the advantages of restricted view photography and the utility of smartphones to produce true circular HP images at higher resolution than traditional DSLR HP without the need for additional equipment (i.e. levelling device or lens). I compare estimates from HP images extracted from SSP HP to traditional DSLR HP images and an alternative smartphone method with a fisheye lens that was proposed by Bianchi et al. (2017).

Spherical panoramas can be generated on any modern smart-phone with preinstalled software like Google Camera (Google LLC), or free applications like Google Street View (Google LLC) (available for Android OS or iOS). With Google Camera, spherical panoramas are composed from 36 individual images with limited field-of-view (57° field-of-view with Google Pixel 4a). Images are acquired by rotating the camera around a central point, guided by the camera's spatial mapping and aided by the device's internal gyroscope and compass. The smartphone camera software automatically merges the individual images into a spherical projection using interest point detection and scale invariant feature transformation to accommodate imperfect viewing distance and viewing angle between images (Szeliski and Shum 1997; Brown and Lowe 2007). Spherical panoramas are recorded in equirectangular projection from which the top half can be easily remapped into polar projection as a circular HP (Figure 1).

Modern smartphones overcome physical limitations of small optics and sensors by employing computational photography techniques that merge multiple images to create a single, high-resolution image (Barbero-García et al. 2018). The result is a composite image that retains the sharpest elements and most even exposure of each individual photo that is sharper (Gunturk 2017) and with greater dynamic range than any individual photo (Lukac

2017). Modern computational photography with small sensors can rival images produced by much larger, DSLR-sized sensors (Ignatov et al. 2017).

METHODS

I estimated canopy structure and light values from two sources: hemispherical photos captured with a DSLR camera (Canon 60D; Canon Inc., Tokyo, Japan) equipped with a circular hemispherical lens (Sigma 4.5 mm f2.8 EX DC; Sigma Corp., Ronkonkoma, NY, USA) and spherical panoramic images captured with a smartphone (Pixel 4a; Google LLC, Menlo Park, CA, USA) and native spherical panorama software (Google Camera v.8.1.011.342784911). In addition, I simulated images to approximate the method proposed by Bianchi et al. (2017) using a smartphone with clip-on fisheye lens.

I acquired images from 35 sites at Yale Nature Preserve, New Haven, CT, USA and 37 sites at Rockstock property, Woodstock, NY, USA on 4 July 2020 and 27 September 2020, respectively. Yale Nature Preserve is comprised of low-elevation rocky ridges separating wet valleys and vernal wetlands. The canopy is composed of mixed hardwoods, predominantly *Acer rubrum*, *Quercus rubra*, *Fagus grandifolia*, *Carya tomentosa*, *Quercus bicolor* and *Liriodendron tulipifera*. Rockstock property is an historical Northeastern farmstead and wetland on which a mixed forest of northern hardwoods (*Acer saccharum*, *Carya ovata*, *Quercus rubra*, *Acer pensylvanicum* and *Betula lenta*) and conifers (*Pinus strobus* and *Tsuga candensis*) have recolonized since the early twentieth century.

Photo sites were selected to represent a distribution of canopy species (deciduous, coniferous and mixed), openness and gap size. All images were captured at breast height (approximately 1.3 m) on uniformly overcast days with smartphone images taken immediately following each DSLR image.

DSLR protocol

Prior to image acquisition, I established exposure settings two stops overexposed relative to open sky following Brown et al. (2000) and Beckschäfer et al. (2013) with maximum ISO values of 1000 and minimum shutter speed of 1/100 s (Chianucci and Cutini 2012). Images were recorded in Canon RAW format (.CR2) with the camera oriented perpendicular to gravity enabled by a dual-axis gimbal and with the top of the image oriented towards magnetic north.

Even on uniformly overcast days, the sky brightness changes over time. To account for this, I adjusted white point values in Adobe Lightroom 5.7.1 to ensure that the grey value of the brightest sky pixels aligned to full white (Beckschäfer et al. 2013) and exported the images as full-resolution (5184 × 3456 pixels) JPEG files. These images are considered the standard reference for comparison throughout further analysis.

To compare the discrepancy due to image acquisition methods to the discrepancy due to incorrect exposure, I additionally created output files with exposure values adjusted 1, 2, 3, 4, or 5 values above and below the original exposure in Adobe Lightroom. I included a circular mask along the perimeter of the circular image to prevent glare at the margins from impacting downstream estimates.

SSP protocol

I created spherical panoramas using a Pixel 4a smartphone— Google's mid-range consumer-grade smartphone model—with the Google Camera application. Spherical panoramas are composed from 36 overlapping images of the entire 360° field-of-view. I began each spherical image sequence facing towards magnetic north (azimuth 0°; this becomes the top of the circular hemispherical image after processing). For this study, I ascertained a northern heading with an external compass first for DSLR HP and used this to

orient the SSP HP. However, the yaw angle from the metadata of the resulting SSP HP image can be used to rotate the image to the proper orientation in post-processing regardless of the direction of capture (Li and Ratti 2019). The first image of the panorama must be taken with the phone levelled to the horizon; the camera software facilitates this by placing a dot on the screen and disallowing images with too great of pitch or roll to the camera. Subsequent images are similarly guided by on-screen targets. Twelve images centered along the horizon comprise zenith 72°–108°. Nine images in the upper and lower hemisphere cover zenith 36°–72° or 108°–144°, respectively. Similarly, three images each comprise the remaining area at the poles. Although no order is specified by the application, I followed the same capture order for every spherical panorama, first rotating to capture the 12 images sequentially around the horizon. Next, I sequentially captured the nine images for zenith 36°–72°, followed by three images for zenith 0°–36°. I then followed the same order in the lower sphere. The phone's internal gyroscope is used to automatically level the horizon of the sphere. Care must be taken to rotate and pan the smartphone camera treating the camera as the centerpoint, rather than rotating the camera around one's body.

Spherical panoramas are created by tiling multiple planar images into a geodesic polyhedron and then stitching the images into a spherical, omnidirectional image which can be viewed in 3D (Fangi and Nardinocchi 2013). The spherical images are mapped into two dimensions following an equirectangular projection in which the zenith angle corresponds to the rectangular y-axis and azimuth angle corresponds to the rectangular x-axis (Fangi and Nardinocchi 2013) (Figure 1). Conveniently, when the sphere is levelled to the horizon, the top half of the rectangular panoramic image depicts the upper hemisphere of view and can be cropped and remapped into a polar projection (e.g. Li and Ratti 2019).

I extracted the top half of the equirectangular panorama JPEG file (i.e. the top hemisphere) and converted it to a circular hemispherical image via polar projection in GIMP

(Gnu Image Manipulation Program v.2.10.20) with batch processing implemented in BIMP (Batch Image Manipulation Plugin v.2.4) (setting files are included in the data archive along with scripts for an alternative processing from the command line with ImageMagick v.7.0.10).

Transformation from equirectangular to polar projection with square pixels requires either downsampling pixels closer to the pole or interpolating pixels near the horizon, or both. I retained the width of the equirectangular image in the transformation to polar projection, resulting in a SSP HP image with diameter equal to the width of the equirectangular projection. Thus, the circumference of the HP at zenith 57° (~1 radian), is equal to the width of the equirectangular image. Pixels circumscribing zenith angles greater or less than 57° are upscaled or downsampled, respectively. The result is a true HP with over 900 per cent of the resolution of traditional DSLR HP images. Because the area of the SSP HP image is larger than the upper half of the equirectangular image (i.e. more upscaling than downscaling), I test the impact of resolution below.

Unlike DSLR images, the white point does not need to be adjusted for smartphone images as this is automatically controlled by the device's high dynamic range (HDR) routine. However, the HDR routine yields a more homogenous histogram and consolidates pixel values at the mid-tones (Figure 2), which can make it difficult for binarization algorithms to differentiate between sky and canopy pixels. Contrast-stretching can facilitate pixel classification (Macfarlane et al. 2014). To test the effect of contrast-stretching, I output two sets of SSP HP images with and without expanding the tonal range by 5 (2 per cent) in GIMP prior to polar projection conversion.

Circular hemispherical images from DSLR photos (diameter: 2885 p; area: 6.5 MP) are considerably smaller than those produced from photo spheres (diameter: 8704 p; area:

59.5 MP). Resolution can impact canopy estimate because larger portions of the hemisphere are averaged into each pixel, leading to higher proportion of mixed pixels and underestimates of small gaps (Macfarlane 2011). To test the impacts of the resolution gain of the SSP HP, I exported an additional set of images downsampled to match the diameter of the DSLR photos (2885 p) in GIMP.

Fisheye lens simulation

Bianchi et al. (2017) proposed a method of approximating hemispherical photos from two perpendicular smartphone images using a fisheye lens adapter with 150° diagonal field-of-view. In order to compare this technique to the SSP HP method proposed in this study, I used the SSP HP images to simulate images from two perpendicular Pixel 4a photos captured with such a lens. I downsampled the SSP HP images to 6049 p diameter and applied a black mask that simulates Pixel 4a image dimensions (5802 × 4352 p) with 150° field-of-view.

ANALYSIS

Binarization and canopy estimates

The processing steps above yielded 16 sets of HP images in JPEG format for each of the 72 sites (Figure S1): standard DSLR HP images (no exposure adjustment), 10 sets of exposure-adjusted DSLR HP images, four sets of SSP HP images at full or low resolution with or without contrast adjustment, and one set of fisheye HP images with appropriate contrast adjustments (Figure S2). From this point, all images received the same processing steps.

I binarized images with the Hemispherical 2.0 plugin (Beckschäfer 2015) for ImageJ v.1.51 k which uses the 'Minimum' algorithm (Prewitt and Mendelsohn 1966), which was found to produce the most accurate threshold for HP images (Glatthorn and Beckschäfer

2014). This algorithm is applied to the blue color channel of the image to automatically classify pixels and output binary images in TIFF format (Beckschäfer 2015). During binarization, the program estimates total gap area and number of gaps, which I recorded for further analysis. I converted the binary TIFF files to BMP format in batch with ImageJ.

I used Gap Light Analyzer v.2.0 (Frazer et al. 1999) to estimate additional canopy structure measures and light transmittance, including CO, LAI and global site factor (GSF) measures, for further analysis. I used identical configuration parameters for all image sets with two exceptions. One, I adjusted the lens projection parameters per camera device. Two, I adjusted coordinates, elevation and declination per location. HP images created from spherical panoramas conform to true polar projection whereas the Sigma hemispherical lens used for DSLR HP images conforms to an equisolid projection (see parameter and lens configuration files in data archive). Gap Light Analyzer is implemented in a graphical interface without an option for command line input. So, I wrote a custom macro-script in AutoHotKey v.1.1.33.02 to batch process the images (script available in the data archive).

Statistical analysis

All statistical analyses were performed in R v.3.6.2 (R Core Team 2019). I used ordinary least squares and mixed effect regression models to compare differences between multiple image sets. I focused on three HP image characteristics—number of gaps, total gap area and relative gap size—to evaluate the difference in image quality between methods. I focused on three canopy measures—CO, effective LAI and GSF—to evaluate the similarity of canopy estimates between methods. Gap fraction, the ratio of white to total pixels, is the most foundational measure of HP. CO is similar to gap fraction, but weights pixels by zenith angle and is a more appropriate measure when comparing HP images with different lens distortion that biases gap size at different zenith angles (Frazer et al. 1997; Gonsamo et al.

2013). I calculated relative gap size as the area of the mean canopy gap, standardized to the total image size which is equal to CO divided by the number of gaps. LAI is a comparison of leaf area relative to horizontal ground area, formally defined as the ratio of single-side leaf area to a given area of horizontal ground footprint (Chen and Black 1992). In Gap Light Analyzer, LAI is integrated over zenith angles 0°–60° following Stenberg et al. (1994). Because no distinction is made between foliage and stem canopy elements nor between broad-leaves and needles in the calculation, LAI estimated in Gap Light Analyzer is effectively a plant area index. GSF is a measure of through-canopy radiation. It is the weighted average of the proportion of direct and indirect radiation transmitted through the canopy to that above the canopy (Anderson 1964). The three canopy measures are a function of the image quality measures and represent the range of the inference for which researchers use HP.

Indirect estimates of LAI from HP rely on the commonly violated assumption that leaves are randomly distributed in the canopy (Yan et al. 2019). To overcome this limitation, analytical methods have been developed to estimate and correct for the clumping effect in canopies by systematically excluding large between-crown gaps from small within-crown gaps to calculate a clumping index (e.g. Chen and Cihlar 1995). However, the poor resolution of traditional DSLR HP makes gap classification ineffective at low zenith angles and in dense canopies (Yan et al. 2019; Chianucci 2020). The high resolution of SSP HP is likely to overcome such limitations and produce more accurate calculations of clumping. To illustrate the differences, I additionally compare clumping index estimates from full-resolution SSP HP and DSLR HP, using the gap thresholding method established by Alvernini et al. (2018) (see Supplemental Materials for details).

I tested for differences in canopy structure and light environment measures across all image processing methods by fitting linear mixed effect models with the unique image

site as a random intercept (i.e. a repeated measures analysis). I regressed each canopy and image quality measure against the image acquisition and processing protocol as the independent variable with the standard DSLR HP image as the reference. I fit mixed effect models in 'lme4' (Bates et al. 2015) and estimated parameters and 95 per cent confidence intervals with 1000 non-parametric bootstrap iterations. Additionally, I fit bivariate ordinary least squares models regressing canopy measures estimated from each image processing protocol on the standard DSLR HP estimates. For these models, the effect size of the dissimilarity between image sets is measured by the magnitude of the deviation of the slope from 1 or the intercept from 0 and the coefficient of determination indicates the appropriateness of the image processing protocol as an alternative to traditional DSLR HP. In all model sets, the variability among estimates from under- or overexposed DSLR HP can be used as a qualitative gauge of the effect size between SSP HP methods.

RESULTS

The canopies surveyed in this study ranged from densely closed (CO_{min} = 1 per cent, GSF_{min} = 1) to moderately open (CO_{max} = 40 per cent, GSF_{max} = 63). Most sites skew towards denser canopies (CO_{IQR} = 2 per cent–6 per cent, GSF_{IQR} = 3–9). This is advantageous for the purpose of this study as estimates from dense canopies are substantially more sensitive to HP settings (Beckschäfer et al. 2013). There was no significant difference between locations (all $P > 0.08$; Table S1).

HP images generated from spherical panoramas were noticeably sharper than those captured with a DSLR and hemispherical lens (Figure 2C and E). The greater sharpness resulted in more definition of fine canopy structure even when scaled to the same resolution (Figure 2E). Close inspection of the smartphone HP reveals occasional infidelities

arising from the process of stitching the panoramas (Figure 2D). These artefacts appear as discontinuities or overlapping of canopy elements.

Effect of contrast-stretching on SSP HP

Unlike most DSLR HP images, which exhibit bimodal distribution of pixel tones, the HDR pre-processing of smartphone cameras produces images with tonal values with a normal distribution centered around the midpoint tone (Figure 2A). This poses a problem for binarization algorithms that iteratively seek a global minimum along the histogram and can lead to extreme over-classification of sky pixels at lower resolutions or restricted field-of-view. One full-resolution SSP HP image (1 per cent) and three low resolution SSP HP images (4 per cent) were incorrectly classified and easily identified as outliers (Figure S3). Twelve fisheye HP images (17 per cent) were incorrectly classified. I applied a 2 per cent contrast-stretch to all SSP HP images and up to 8 per cent as needed for fisheye HP images to ensure no misclassifications. Without contrast adjustment, CO and GSF from full-resolution SSP HP images are overestimate by 18 per cent and 12 per cent whereas LAI is underestimated by less than −1 per cent relative to the corrected image set (Table 1). The effect of contrast-stretching is much stronger for estimates from low resolution images due to the greater number of incorrectly binarized outlier images. Relative to corrected images, CO and GSF are overestimated by 56 per cent and 37 per cent, respectively and LAI is underestimated by −2 per cent when low resolution SSP HP are not corrected with contrast-stretching (Table 1). Only the contrast adjusted images were retained for further analysis.

Effect of resolution on SSP HP

Downsampling SSP HP images from 59.5 MP to 6.5 MP (−815 per cent) to match the resolution of DSLR HP images resulted in nearly half the number of gaps (−48 per cent) of larger relative size (+142 per cent) compared to the full-resolution SSP HP image set (Table

1). Downsampling had minimal effect on canopy structure and light estimates, however (Table 1, Figure 3). Downsampling resulted in only a 1 per cent decrease in CO and 2 per cent decrease in GSF in relation to the standard reference compared to full-resolution SSP HP images (Table 1). Downsampling had a similarly minimal effect of increased LAI (1 per cent) (Table 1). Reduction in resolution increased variance in CO (1 per cent), GSF (1 per cent) and LAI (4 per cent) for the entire imageset, but decreased the variance in number of gaps (−61 per cent) and gap area (−89 per cent) considerably (Table 2, Figure 4). Downsampled images exhibited greater variance in relative gap size (+107 per cent) (Table 2).

Comparison to DSLR HP

Full-resolution SSP HP images retained more gaps (493 per cent) and smaller relative gap size (−79 per cent) compared to DSLR HP (Table 1, Figure 4). Although downsampling reduces the difference, SSP HP images with the same resolution as DSLR HP images exhibit more (209 per cent) gaps of smaller relative size (−49 percent) (Table 1). The difference in relative gap size increases with greater CO (Figure 5). Despite the larger difference in gap number, the total gap area in downsampled smartphone images was just 19 per cent larger than DSLR HP reference images on average (Table 1).

Estimates from SSP HP images were greater for CO (Full = +23 per cent, Low = +22 per cent), greater for GSF (Full = +18 per cent, Low = 16 per cent) and lower for LAI (Full = −19 per cent, Low = −18 per cent) compared to the reference DSLR HP images (Table 1, Figure 4). All canopy measures based on SSP HP images tend to overestimate the reference DSLR HP images at low values but underestimate at higher values (Figure 3). In general, these differences are comparable to the effects of overexposing DSLR images by 1–1.5 stops (Figures 4 and S4). SSP HP estimates of CO and GSF were highly correlated with DSLR HP

estimates ($R^2 > 0.9$) whereas LAI was moderately correlated ($R^2 = 0.71$) (Table S2, Figure S5).

Comparison to 150° field-of-view fisheye HP

HP images emulating perpendicular 150° field-of-view fisheye images were moderately to highly correlated with the reference DSLR HP images for CO ($R^2 = 0.87$), GSF ($R^2 = 0.89$) and LAI ($R^2 = 0.66$) but less so than true HP produced by spherical panoramas (Figure S5, Table S2). Fisheye HP image values tended to result in overestimates of CO (+6 per cent) and GSF (+7 per cent), relative to full-resolution SSP HP images, on average (Table 1). The restricted field-of-view of the fisheye images tend to underestimate LAI more so than full-resolution SSP HP (−7 per cent) (Table 1). Estimates from the fisheye HP method are comparable to DSLR HP images 1–2 stops overexposed (Figures 4 and S4).

Comparison of clumping index

SSP HP recovers more gaps in total (Table 1), but also tends to recover a higher proportion (+ 64 per cent) of between-canopy gaps (i.e. large gap fraction) relative to DSLR HP (Figure S6). Correspondingly, crown porosity tends to be greater (61 per cent) when estimated from SSP HP (Figure S6). As a result, DSLR HP tends to overestimate clumping (2 per cent), especially in denser canopies (3.4 per cent overestimation in canopies under 10 per cent openness) (Figure S6).

DISCUSSION

Smartphones have become nearly ubiquitous, yet researchers typically do not exploit even a fraction of their potential as a research tool. HP generated from SSP offers a highly accurate alternative to traditional DSLR HP with over 90 per cent correlation with traditional methods for common canopy and light measures. The difference between

estimates of canopy structure and light environment from spherical panoramas vary from the reference photos by about the same as over- or underexposing images by 1–1.5 stops.

The primary differences between SSP HP and DSLR HP are that the former produces larger images, sharper resolution and more even tonal range across the image. SSP HP images generated with the Google Pixel 4a smartphone are over nine times larger than those taken with a DSLR. However, the simple increase in resolution does not account for the difference in clarity, as downsampled SSP HP images still retain more fine structure than DSLR HP images of the same size. When comparing DSLR HP images to those from SSP HP at the same resolution, it is clear that DSLR HP tends to result in underestimation of the number of canopy gaps. This is due to low clarity causing adjacent pixels to bleed into each other, even when generated with industry standard lens and camera. Thus, the smallest gaps tend to be lost and small gaps tend to collapse into larger gaps more quickly, as evidenced by higher relative gap size across all canopy densities. This second point can be seen in the way relative gap size increases with CO much faster for DSLR HP. Thus, although SSP HP images contain many more total gaps, the total gap area is similar, albeit with slightly more relative gap area as a consequence of retaining small gaps. The retention of small gaps in SSP HP offers great potential in using crown clumping indices to correct LAI estimates for non-randomly distributed canopy elements.

The difference in clarity is most likely a product of the restricted field-of-view of individual photos included in the panorama, improved sharpening through computational photography, and homogenous tone across zenith regions. In contrast, DSLR HP images suffer from glare and hazing associated with extremely wide field-of-view lenses. In addition, DSLR cameras struggle to evenly expose the entire hemisphere in a single exposure. For these reasons, SSP HP is likely to be far more robust to non-optimal lighting conditions, but more studies in variable skies are needed. An additional advantage of SSP

HP is that, unlike DSLR methods, automatic exposure aided by HDR effectively obviates the need for tedious manual exposure settings. Although this can introduce errors in binarization, minimal contrast-stretching solves the issue. Improvements in pixel classification beyond simple thresholding (e.g. Díaz et al. 2021) are likely to make this a non-issue even in direct sunlight conditions.

Although many canopy and light estimation software programs require hemispherical images, panoramic images could be directly processed without re-projecting them into polar coordinates if estimates of light environment are not needed (e.g. Grotti et al. 2020). Directly processing panoramic projections has the advantage of retaining more pixels closer to the zenith which may improve accuracy of estimating gap fraction, especially in denser canopies (Chianucci et al. 2019). Retaining panoramic projections would additionally allow for direct comparison and validation with other methods, such as terrestrial scanning lidar (Grotti et al. 2020).

Artefacts generated by the panorama stitching process did not have a noticeable effect on canopy and light estimates but are a potential source of error. Care during image capture can reduce most cases of incongruity, however. Spherical panorama software assumes that all images are captured by rotating the camera around a single point in space (Fangi and Nardinocchi 2013). When taking images by hand, it is easy to shift the phone, and therefore the image plane, during rotation. Practice in steady positioning helps. Also, the ability to immediately review panoramas on one's screen phone or with stereoptic headsets lets researchers catch errors and retake panoramic images in the field. This problem will attenuate as smartphone stitching software continues to improve (Luhmann 2004).

SSP HP offers practical benefits over other methods in requiring no additional equipment other than a smartphone. Levelling is achieved automatically via the phone's gyroscope. Geographic coordinates, elevation and orientation are retained in the image metadata and could easily be integrated into an analysis pipeline. Furthermore, dedicated smartphone apps could be developed to make data acquisition accessible to non-experts. The fact that waterproof housings are cheap and readily available for smartphones is another benefit to field work not to be overlooked. Coupled with the fact that SSP HP requires no tedious exposure settings and can be easily georeferenced, this method is highly amenable to citizen science projects or widespread ground-truthing of remote sensing data.

Although there is likely to be variation between smartphone models which must be validated in future studies, the lack of lenses with idiosyncratic projections removes a major source of variability. All SSP HP images have polar projection by virtue of originating as spherical images.

The method presented here outperforms other methods of capturing HP with smartphones via clip-on fisheye lenses. Simulating Bianchi et al. (2017) method showed that the loss of information from restricting images to 150° field-of-view results in skewed but still relatively accurate estimates of canopy structure and light environment. However, the comparison is generous in that the fisheye HP images in this study were simulated from SSP HP images and therefore suffered no effects of poor optical quality associated with the small size of clip-on smartphone lenses.

Only the upper hemisphere was extracted from the spherical panoramas in this study, but other portions of the panorama could be extracted for other purposes. For instance, horizontal panorama could be used for estimating basal area (Fastie 2010) or mapping stands (Lu et al. 2019). The lower hemisphere could be useful in monitoring

understory plants or ground cover composition. Researchers can even enter the spherical image with a virtual reality headset to identify species after leaving the field.

CONCLUSION

High-quality hemispherical images can be acquired with modern smartphones via the following protocol:

• Capture photospheres following on-screen instructions of smartphone panorama application, being careful to rotate in place with the camera lens as the center of the sphere.

• Convert spherical panoramas to circular hemispherical images with provided scripts, or by manually extracting the upper half of the 2D Cartesian panorama, applying 2 per cent contrast-stretching, and projecting into polar coordinates.

SSP HP solves many of the problems associated with traditional HP while offering many practical benefits for field applications. The ubiquity of smartphones and their ever-improving quality of software and optical hardware will only widen the range of spherical panoramic imagery applications in silviculture and forest ecology into the future.

Supplementary Materials – The following supplementary material is available at Forestry online: additional data summary tables, additional figures, and additional analyses. R code used to conduct statistical analysis and generate figures, AutoHotKey macro scripts to automate Gap Light Analyzer software, and code to batch process spherical panoramas via BIMP recipes for GIMP or via ImageMagick are available in the data repository. Additional method tutorials are available at the author's website: azandisresearch.com.

Data Availability – Code for analysis and image processing is available at: github.com/andisa01/Arietta2021_Forestry.git. Due to the large file sizes, raw images are available upon request to the author.

Acknowledgements – I would like to thank Vince Mow for access to Rockstock property and Yale University for access to Yale Preserve. I thank Dr David Skelly, Logan Billet, Dahn-Young Dong, Dr Marlyse Duguid, Dr Jonathan Richardson and Dr Craig Brodersen for providing thoughts on the methods and manuscript.

Conflict of Interest Statement – The author declares no conflicts of interest.

Funding – Yale Institute for Biospheric Studies and the Kohlberg-Donohoe Research Fellowship.

REFERENCES

Alvernini, A., Fares, S., Ferrara, C. & Chianucci, F. (2018). 'An objective image analysis method for estimation of canopy attributes from digital cover photography', *Trees*, 32, pp. 713-723.

Anderson, M. C. (1964) 'Studies of the woodland light climate I. The photographic computation of light condition', *Journal of Ecology*, 52, pp. 27–41.

Barbero-García, I. *et al.* (2018) 'Smartphone-based close-range photogrammetric assessment of spherical objects', *Photogrammetric Record, The*, 33(162), pp. 283–299.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, Articles*, *67*(1), 1–48.

Beckschäfer, P. *et al.* (2013) 'On the exposure of hemispherical photographs in forests', *iForest*, 6(4), pp. 228–237.

Beckschäfer, P. (2015) *Hemispherical_2.0: Batch processing hemispherical and canopy photographs with ImageJ - User manual*. doi: 10.13140/RG.2.1.3059.4088.

Bianchi, S. *et al.* (2017) 'Rapid assessment of forest canopy and light regime using smartphone hemispherical photography', *Ecology and evolution*, 7(24), pp. 10556–10566.

Brown, M. and Lowe, D. G. (2007) 'Automatic panoramic image stitching using invariant features', *International journal of computer vision*, 74(1), pp. 59–73.

Brown, P. L., Doley, D. and Keenan, R. J. (2000) 'Estimating tree crown dimensions using digital analysis of vertical photographs', *Agricultural and Forest Meteorology*, 100(2), pp. 199–212.

Chen, J. M., and Black, T. A. (1992). Defining leaf area index for non-flat leaves. *Plant, Cell & Environment*, *15*(4), 421–429.

Chen, J. M., and Cihlar, J. (1995). Plant canopy gap-size analysis theory for improving optical measurements of leaf-are index. *Applied Optics*, 34(27), 6211–6222.

Chianucci, F. (2020) 'An overview of in situ digital canopy photography in forestry', *Canadian journal of forest research. Journal canadien de la recherche forestiere*, 50(3), pp. 227–242.

Chianucci, F. and Cutini, A. (2012) 'Digital hemispherical photography for estimating forest canopy properties: Current controversies and opportunities', *iForest - Biogeosciences and Forestry*, 5(6). doi: 10.3832/ifor0775-005.

Chianucci, F., Zou, J., Leng, P., Zhuang, Y., & Ferrara, C. (2019). A new method to estimate clumping index integrating gap fraction averaging with the analysis of gap size distribution. *Canadian Journal of Forest Research. Journal Canadien de La Recherche Forestiere*, *49*(5), 471–479.

Díaz, G. M., Negri, P. A. and Lencinas, J. D. (2021) 'Toward making canopy hemispherical photography independent of illumination conditions: A deep-learning-based approach', *Agricultural and Forest Meteorology*, 296, p. 108234.

Fangi, G. and Nardinocchi, C. (2013) 'Photogrammetric processing of spherical panoramas', *Photogrammetric Record, The*, 28(143), pp. 293–311.

Fastie, C. L. (2010) 'Estimating stand basal area from forest panoramas', in *Proceedings of the Fine International Conference on Gigapixel Imaging for Science*. *Fine International Conference on Gigapixel Imagingfor Science*, Carnegie Mellon University, pp. 1–7. Available at: https://doi.org/10.1184/ R1/6709391.v1.

Frazer, G. W., Canham, C. D. and Lertzman, K. P. (1999) *Gap Light Analyzer (GLA): Imaging software to extract canopy structure and gap light transmission indices from true-color fisheye photographs, users manual and documentation*. Burnaby, British Columbia: Simon Fraser University. Available at: http://rem-main.rem.sfu.ca/downloads/Forestry/GLAV2UsersManual.pdf.

Frazer, G. W., Trofymow, J. A. and Lertzman, K. P. (1997) *A method for estimating canopy*

openness, effective leaf area index, and photosynthetically active photon flux density using hemispherical photography and computerized image analysis techniques. BC-X-373. Pacific Forestry Centre. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.477.483&rep=rep1&type=pdf.

Glatthorn, J. and Beckschäfer, P. (2014) 'Standardizing the protocol for hemispherical photographs: accuracy assessment of binarization algorithms', *PloS one*, 9(11), p. e111924.

Gonsamo, A., D'odorico, P. and Pellikka, P. (2013) 'Measuring fractional forest canopy element cover and openness - definitions and methodologies revisited', *Oikos* , 122(9), pp. 1283–1291.

Grotti, M.; et al. (2020). An intensity, image-based method to estimate gap fraction, canopy openness and effective leaf area index from phase-shift terrestrial laser scanning. *Agricultural and Forest Meteorology*, *280*, 107766.

Gunturk, B. K. (2017) 'Super-resolution imaging', in Lukac, R. (ed.) *Computational Photography: Methods and Applications*. CRC Press, pp. 175–208.

Halverson, M. A. *et al.* (2003) 'Forest mediated light regime linked to amphibian distribution and performance', *Oecologia*, 134(3), pp. 360–364.

Ignatov, A. *et al.* (2017) 'DSLR-quality photos on mobile devices with deep convolutional networks', In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017. pp. 3277–3285. *arXiv [cs.CV]*. Available at: http://arxiv.org/abs/1704.02470.

Jennings, S. B., Brown, N. D. and Sheil, D. (1999) 'Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures', *Forestry* , 72(1), pp. 59–74.

Li, X. and Ratti, C. (2019) 'Mapping the spatio-temporal distribution of solar radiation within street canyons of Boston using Google Street View panoramas and building height model', *Landscape and urban planning*, 191, p. 103387.

Luhmann, T. (2004) 'A historical review on panorama photogrammetry', *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 34(5/W16), p. 8.

Lukac, R. (2017) *Computational Photography: Methods and Applications*. CRC Press.

Lu, M.-K. *et al.* (2019) 'Close-range photogrammetry with spherical panoramas for mapping spatial location and measuring diameters of trees under forest canopies', *Canadian journal of forest research. Journal canadien de la recherche forestiere*, 49(8), pp. 865–874.

Macfarlane, C. (2011) 'Classification method of mixed pixels does not affect canopy metrics from digital images of forest overstorey', *Agricultural and Forest Meteorology*, 151(7), pp. 833–840.

Macfarlane, C. *et al.* (2014) 'Digital canopy photography: Exposed and in the raw', *Agricultural and Forest Meteorology*, 197, pp. 244–253.

Origo, N. *et al.* (2017) 'Influence of levelling technique on the retrieval of canopy structural parameters from digital hemispherical photography', *Agricultural and Forest Meteorology*, 237-238, pp. 143–149.

Parker, G. G., Lowman, M. D., & Nadkarni, N. M. (eds) (1995). *Forest canopies*. Academic Press, San Diego. ISBN 0-12-457650-8.

Prewitt, J. M. S. and Mendelsohn, M. L. (1966) 'The analysis of cell images', *Annals of the New York Academy of Sciences*, 128(3), pp. 1035–1053.

Promis, A. (2013) 'Measuring and estimating the below-canopy light environment in a forest. a review', 19(1)), pp. 139–146.

R Core Team (2019) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/.

Rich, P. M. (1990) 'Characterizing plant canopies with hemispherical photographs', *Remote Sensing Reviews*, 5(1), pp. 13–29.

Stenberg, P., Linder, S., Smolander, H., & Flower-Ellis, J. (1994). Performance of the LAI-2000 plant canopy analyzer in estimating leaf area index of some Scots pine stands. *Tree Physiology*, *14*(7_9), 981–995.

Szeliski, R. and Shum, H.-Y. (1997) 'Creating full view panoramic image mosaics and environment maps', in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. USA: ACM Press/Addison-Wesley Publishing Co. (SIGGRAPH '97), pp. 251–258.

Tichý, L. (2016) 'Field test of canopy cover estimation by hemispherical photographs taken with a smartphone', *Journal of vegetation science: official organ of the International Association for Vegetation Science*. Edited by B. Collins, 27(2), pp. 427–435.

Yan, G., et al. (2019). Review of indirect optical measurements of leaf area index: Recent advances, challenges, and perspectives. *Agricultural and Forest Meteorology*, *265*, 390–411.

Zhang, Y., Chen, J. M. and Miller, J. R. (2005) 'Determining digital hemispherical photograph exposure for leaf area index estimation', *Agricultural and Forest Meteorology*, 133(1), pp. 166–181.

TABLES

Table 1. Linear mixed effect model coefficients and 95% confidence intervals for three canopy measures and three image quality metrics estimated from images acquired with traditional DSLR camera and hemispherical lens or extracted from smartphone spherical panoramas (SSP) and processed with difference protocols. For all models, the unique site of image acquisition was estimated as a random intercept (N_{obs} = 1152, N_{sites} = 72). Coefficients that significantly differ from the $DSLR_{standard}$ reference images (i.e. 95% confidence interval does not include 0) are indicated in bold. Coefficients and confidence intervals were estimated with 1000 non-parametric bootstraps.

Image set	Canopy openness (%)	Leaf area index	Global site factor	Gap Area (MP)	Number of gaps (x 10^3)	Relative gap size (x10^{-3})
$DSLR_{Standard}$	**5.49**	**4.19**	**8.23**	0.37	**4.14**	1.23
(Intercept)	**(3.45, 7.65)**	**(3.66, 4.69)**	**(5.68, 11.05)**	(-0.09, 0.88)	**(3.50, 4.80)**	(-7.19, 9.04)
SSP_{Full}	1.29	**-0.78**	1.47	**3.66**	**20.41**	-0.97
	(-1.15, 3.64)	**(-1.39, -0.16)**	(-0.88, 3.69)	**(2.98, 4.27)**	**(19.62, 21.22)**	(-12.61, 10.39)
SSP_{Low}	1.20	**-0.74**	1.32	0.07	**8.67**	-0.60
	(-1.15, 3.32)	**(-1.36, -0.15)**	(-0.97, 3.58)	(-0.56, 0.67)	**(7.92, 9.45)**	(-11.11, 11.13)
$SSP_{Full\ (uncorrected)}$	2.48	**-0.80**	**2.59**	**4.36**	**20.08**	0.12
	(0.12, 4.89)	**(-1.37, -0.21)**	**(0.32, 4.83)**	**(3.75, 4.97)**	**(19.21, 20.96)**	(-10.34, 11.24)
$SSP_{Low\ (uncorrected)}$	**4.95**	**-0.83**	**4.88**	0.30	**7.81**	**23.49**
	(2.66, 7.30)	**(-1.44, -0.20)**	**(2.62, 7.25)**	(-0.35, 0.92)	**(7.02, 8.62)**	**(12.94, 33.95)**
$SSP_{Fisheye}$	1.72	**-1.01**	2.15	**1.58**	**14.93**	-0.78
	(-0.40, 4.10)	**(-1.61, -0.45)**	(-0.11, 4.44)	**(0.94, 2.22)**	**(14.18, 15.73)**	(-12.55, 10.50)
$DSLR_{Exp\ -5}$	**-3.93**	**7.09**	**-5.92**	-0.26	**-3.42**	-0.31
	(-6.24, -1.72)	**(6.51, 7.72)**	**(-8.30, -3.60)**	(-0.92, 0.35)	**(-4.26, -2.61)**	(-11.73, 10.86)
$DSLR_{Exp\ -4}$	**-3.20**	**4.40**	**-4.90**	-0.23	**-2.81**	0.15
	(-5.25, -0.91)	**(3.76, 5.02)**	**(-7.26, -2.51)**	(-0.80, 0.44)	**(-3.61, -2.04)**	(-9.77, 11.60)
$DSLR_{Exp\ -3}$	-1.67	**2.60**	**-2.50**	-0.10	**-2.03**	0.47
	(-3.96, 0.54)	**(2.01, 3.22)**	**(-4.75, -0.25)**	(-0.72, 0.46)	**(-2.91, -1.25)**	(-10.81, 11.43)
$DSLR_{Exp\ -2}$	-1.39	**1.26**	-1.93	-0.10	**-1.73**	0.33
	(-3.67, 0.92)	**(0.66, 1.87)**	(-4.15, 0.19)	(-0.74, 0.54)	**(-2.51, -0.92)**	(-10.56, 11.69)
$DSLR_{Exp\ -1}$	-0.78 (-3.13, 1.63)	0.56 (-0.05, 1.15)	-1.14 (-3.42, 1.18)	-0.04 (-0.69, 0.61)	**-0.95 (-1.81, -0.17)**	0.44 (-10.90, 11.40)
$DSLR_{Exp\ +1}$	1.36 (-0.86, 3.51)	-0.58 (-1.18, 0.04)	1.92 (-0.37, 4.11)	0.09 (-0.54, 0.74)	**1.48 (0.71, 2.27)**	-0.02 (-11.40, 12.75)
$DSLR_{Exp\ +2}$	**3.22 (1.13, 5.63)**	**-1.07 (-1.71, -0.49)**	**4.59 (2.17, 6.98)**	0.22 (-0.38, 0.87)	**3.07 (2.30, 3.88)**	0.00 (-11.07, 11.65)
$DSLR_{Exp\ +3}$	**6.04 (3.65, 8.32)**	**-1.59 (-2.20, -0.96)**	**8.38 (6.11, 10.73)**	0.44 (-0.22, 1.08)	**4.58 (3.81, 5.41)**	0.31 (-10.31, 11.84)
$DSLR_{Exp\ +4}$	**9.94 (7.66, 12.10)**	**-2.00 (-2.58, -1.38)**	**13.54 (11.32, 15.86)**	**0.68 (0.06, 1.29)**	**5.80 (5.00, 6.57)**	0.63 (-11.06, 11.77)
$DSLR_{Exp\ +5}$	**15.87 (13.54, 18.08)**	**-2.38 (-2.99, -1.79)**	**20.53 (18.23, 22.90)**	**1.10 (0.49, 1.73)**	**7.41 (6.59, 8.21)**	1.37 (-9.42, 12.69)

Table 2. Comparison of canopy measures and image quality metrics estimated from full resolution smartphone spherical panorama hemispherical photography imagesets and the same imageset downsampled to the same area as DSLR hemispherical photography images.

Measure	Full Resolution Mean (S.D.)	Low Resolution Mean (S.D.)	% Difference of Mean	% Difference of S.D.
Canopy openness (%)	6.78 (5.35)	6.7 (5.41)	-1.15	0.96
Global site factor	9.69 (8.95)	9.56 (8.99)	-1.31	0.48
Leaf area index	3.19 (0.59)	3.23 (0.62)	1.01	4.25
Gap area (MP)	4.02 (3.18)	0.44 (0.36)	-89.16	-88.95
No. of gaps ($x10^3$)	24.53 (5.08)	12.82 (19.82)	-4.78	-6.10
Relative gap size ($x10^{-3}$)	0.29 (0.27)	0.55 (0.55)	91.01	106.89

FIGURES

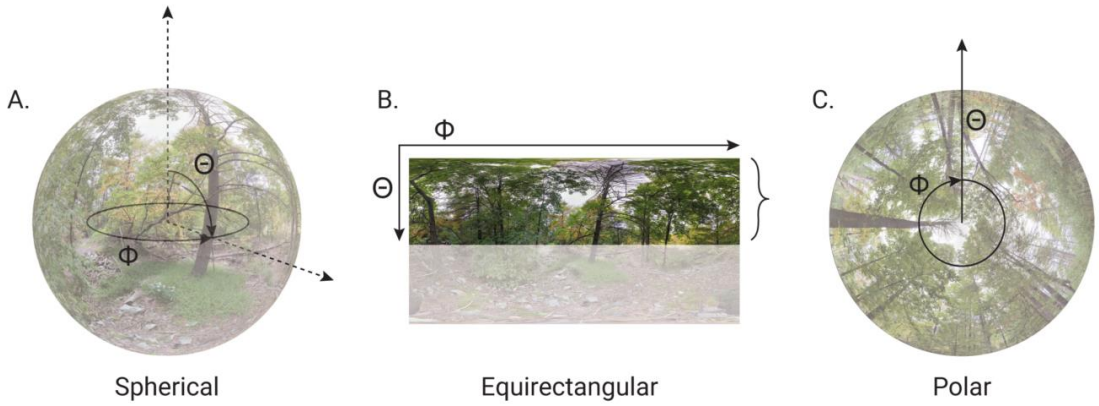

Figure 1. Spherical panoramas (A) are stored and output from smartphones as 2D images with equirectangular projection (B). Because spherical panoramas are automatically leveled using the phone gyroscope, the top half of the equirectangular image corresponds to the upper hemisphere of the spherical panorama. The top portion of the equirectangular image (B) can then be remapped onto the polar coordinate plane to create a circular hemispherical photo (C). In all images, zenith and azimuth are indicated by Θ and Φ, respectively.

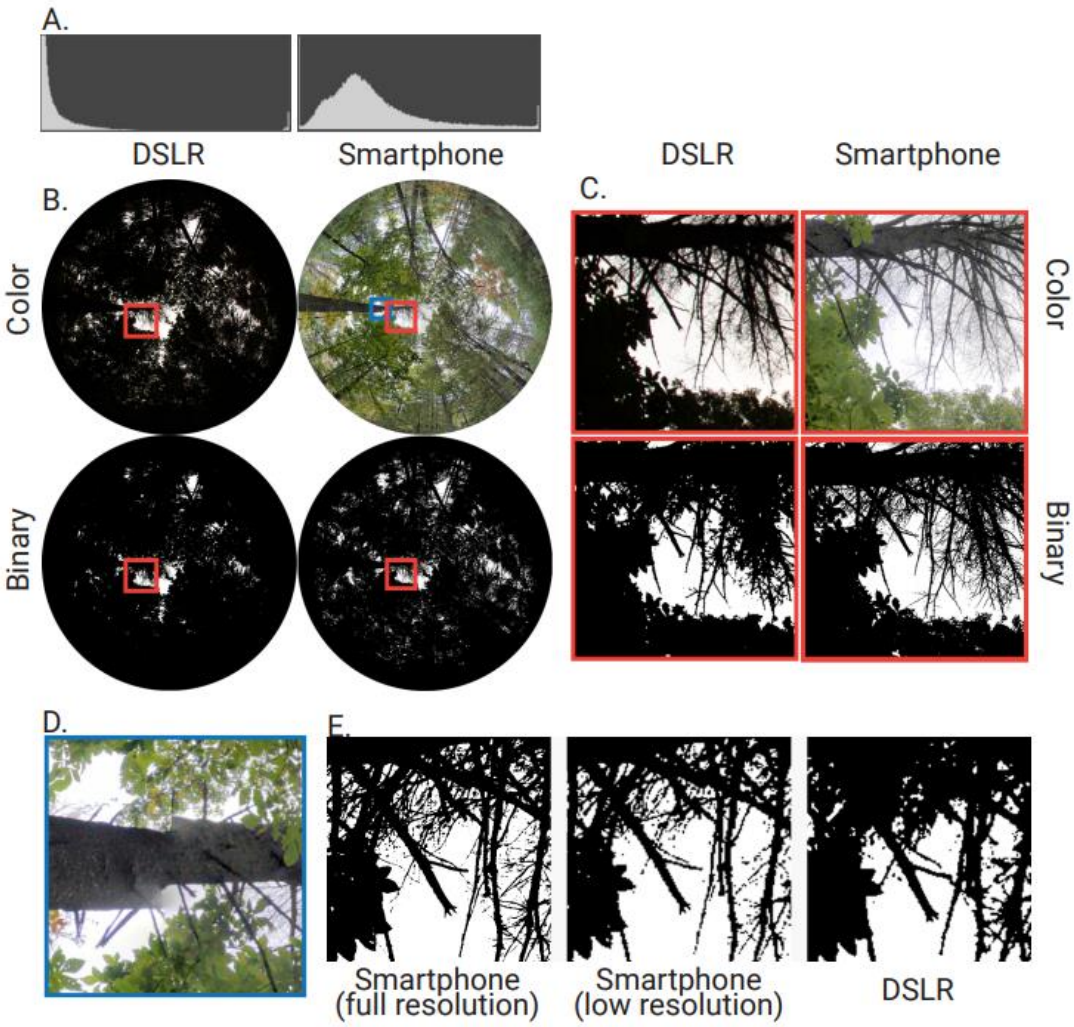

Figure 2. Comparisons of images generated with smartphone spherical panorama hemispherical photography (SSP HP) (right B and C) to traditional DSLR hemispherical photography (DSLR HP) (left B and C) captured at the same site. Details of a subsection of the canopy, indicated by orange boxes, are expanded in C. Binarized images are shown below color images in B and C. Image histograms differ in the distribution of luminance values in the blue color plane (A). In panel E, a section of the canopy from a binary full resolution SSP HP image (left), downsampled SSP HP image (middle), and DSLR HP image (right) is further expanded to demonstrate the effect of image clarity on pixel classification. An example of an incongruous artifact resulting from misalignment in the spherical panorama is outlined in blue in B and expanded in D. SSP HP images were generated from spherical panoramas taken with Google Pixel 4a and Google Camera. DSLR HP images were captured with Canon 60D and Sigma 4.5mm f2.8 EX DC hemispherical lens.

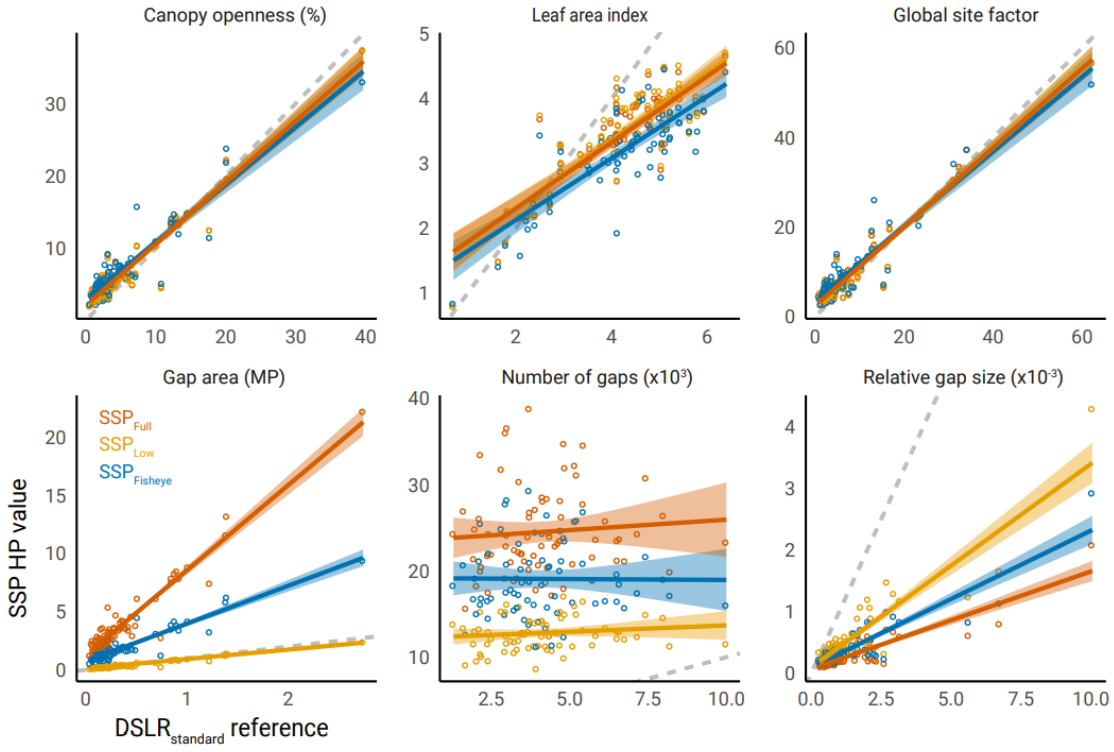

Figure 3. Linear regression estimates comparing three canopy structure measures (canopy openness, leaf-area index (LAI), and global site factor (GSF)) and three image quality metrics (total gap area, number of gaps, and relative gap size) from images hemispherical images extracted from smartphone spherical panoramas (SSP) to those acquired with traditional DSLR camera and hemispherical lens ($DSLR_{Standard}$). Three processing methods for SSP hemispherical images were compared, including full resolution (SSP_{Full}, dark orange), images down-sampled to match the resolution of DSLR images (SSP_{Low}, light orange), and images simulated to approximate Bianchi et al.'s (2017) method using a smartphone and clip-on fisheye lens ($SSP_{Fisheye}$, blue) are shown. The 1:1 line is indicated by a dashed grey line. Figure S4 additionally shows the comparison of exposure setting for DSLR images.

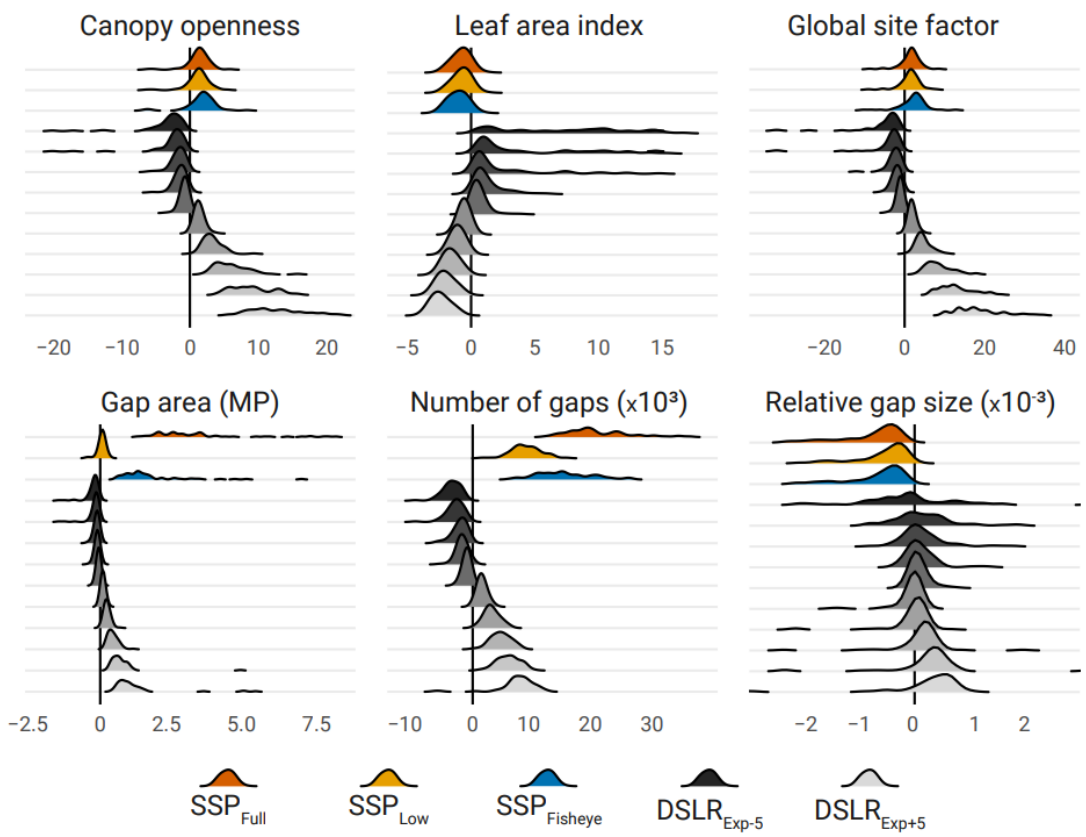

Figure 4. Difference in canopy structure and light environment estimates between reference (standard DSLR hemispherical photography (HP) images) and full resolution smartphone spherical panorama hemispherical photography (SSP HP) images (SSP_{Full}, dark orange), low resolution SSP HP images (SSP_{Low}, light orange) downsampled to match the standard DSLR resolution, fisheye HP ($SSP_{Fisheye}$, blue), and DSLR HP images with exposure adjusted from +5 to -5 ($DSLR_{Exp}$, light to dark).

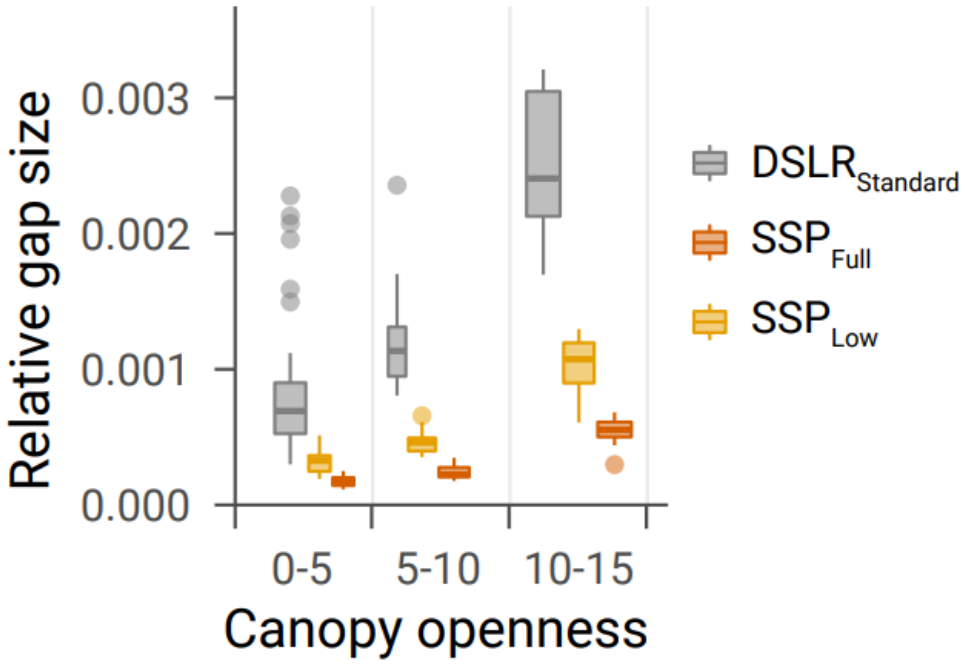

Figure 5. Relative gap size related to canopy openness for standard DSLR hemispherical photography (grey), downsampled smartphone spherical panorama hemispherical photography (SSP HP) images (SSP_{Low}, light orange), and full resolution SSP HP images (SSP_{Full}, dark orange). Values were binned into three canopy openness thresholds (0-5, 5-10, 10-15). Canopy openness values greater than 15% were excluded due to lack of sites with high values.

APPENDIX

Supplementary materials for:

Estimation of forest canopy structure and understory light using spherical panorama images from smartphone photography

A. Z. Andis Arietta
Yale School of the Environment, Yale University, New Haven, CT, USA

Table S1. Effect of location on canopy measures estimated from standard DSLR HP images fit with OLS regression. N = 72 for all models.

Measure	Difference	95% C.I.	*p*
Canopy openness (%)	-2.48	-5.24, 0.28	0.08
Leaf area index	0.49	-0.05, 1.04	0.08
Global site factor	-3.06	-7.57, 1.45	0.19

Table S2. Regression coefficients and coefficients of determination from linear OLS regression models predicting canopy structure and light environment values for reference (standard DSLR HP) images from values for the same measure for comparison data sets. Intercepts that significantly deviate from 0 and slopes that significantly deviate from 1 at the alpha = 0.05 level are indicated in bold.

	Canopy openness (%)			LAI			GSF		
Image set	Intercept (95% C.I.)	Slope (95% C.I.)	R^2	Intercept (95% C.I.)	Slope (95% C.I.)	R^2	Intercept (95% C.I.)	Slope (95% C.I.)	R^2
$Smartphone_{Full}$	**-1.86 (-2.48, -1.24)**	**1.09 (1.02, 1.16)**	**0.93**	-0.6 (-1.32, 0.13)	**1.4 (1.19, 1.61)**	**0.71**	**-2.06 (-2.89, -1.22)**	1.06 (1, 1.13)	0.94
$Smartphone_{Low}$	**-1.69 (-2.32, -1.07)**	1.08 (1, 1.15)	0.92	-0.47 (-1.19, 0.25)	**1.35 (1.14, 1.55)**	**0.71**	**-1.86 (-2.7, -1.03)**	1.06 (0.99, 1.12)	0.94
$Smartphone_{Fisheye}$	**-2.39 (-3.28, -1.5)**	1.09 (0.99, 1.19)	0.87	-0.18 (-0.93, 0.58)	**1.38 (1.14, 1.61)**	**0.66**	**-2.86 (-4.05, -1.67)**	1.07 (0.98, 1.16)	0.89

$DSLR_{Exp\ -5}$	**4.21 (2.82, 5.61)**	0.84 (0.45, 1.23)	0.2	**3.2 (2.58, 3.83)**	**0.09 0.04, 0.14)**	**0.14**	**6.34 (4.09, 8.58)**	0.84 (0.44, 1.24)	0.19
$DSLR_{Exp\ -4}$	**3.58 (2.07, 5.1)**	0.84 (0.47, 1.22)	0.22	**3.37 (2.83, 3.9)**	**0.09 (0.04, 0.15)**	**0.14**	**5.35 (2.98, 7.72)**	0.86 (0.49, 1.23)	0.23
$DSLR_{Exp\ -3}$	**1.64 (1.34, 1.95)**	1.02 (0.97, 1.06)	0.97	**2.91 (2.49, 3.33)**	**0.19 (0.13, 0.24)**	**0.4**	**2.49 (2.02, 2.97)**	1 (0.96, 1.04)	0.97
$DSLR_{Exp\ -2}$	**1.34 (1.1, 1.59)**	1.02 (0.99, 1.06)	0.98	**1.45 (1.06, 1.84)**	**0.5 (0.43, 0.57)**	**0.76**	**1.95 (1.64, 2.27)**	1.01 (0.98, 1.04)	0.99
$DSLR_{Exp\ -1}$	**0.7 (0.56, 0.85)**	1.02 (1, 1.04)	0.99	0.8 (0.51, 1.08)	**0.71 (0.66, 0.77)**	**0.9**	1.05 (0.85, 1.24)	1.01 (1, 1.03)	1
$DSLR_{Exp\ +1}$	**-1.09 (-1.28, -0.89)**	**0.97 (0.94, 0.99)**	**0.99**	**-0.28 (-0.46, -0.11)**	**1.23 (1.19, 1.28)**	**0.97**	**-1.59 (-1.84, -1.35)**	**0.97 (0.96, 0.99)**	**0.99**
$DSLR_{Exp\ +2}$	**-2.2 (-2.65, -1.76)**	**0.88 (0.84, 0.92)**	**0.96**	-0.25 (-0.57, 0.06)	**1.42 (1.33, 1.52)**	**0.92**	**-3.42 (-4.01, -2.83)**	**0.91 (0.87, 0.95)**	**0.97**
$DSLR_{Exp\ +3}$	**-3.41 (-4.2, -2.63)**	**0.77 (0.71, 0.83)**	**0.91**	0.13 (-0.3, 0.55)	**1.55 (1.4, 1.71)**	**0.85**	**-5.4 (-6.52, -4.36)**	**0.82 (0.76, 0.87)**	**0.92**
$DSLR_{Exp\ +4}$	-0.95 (-2.47, 0.56)	**0.42 (0.34, 0.5)**	**0.61**	**0.59 (0.11, 1.06)**	**1.64 (1.43, 1.85)**	**0.78**	**-5.09 (-7.31, -2.86)**	**0.61 (0.52, 0.7)**	**0.73**
$DSLR_{Exp\ +5}$	-0.51 (-1.79, 0.76)	**0.28 (0.24, 0.33)**	**0.67**	**1.29 (0.84, 1.74)**	**1.61 (1.37, 1.84)**	**0.72**	**-4.37 (-6.68, -2.05)**	**0.44 (0.37, 0.51)**	**0.7**

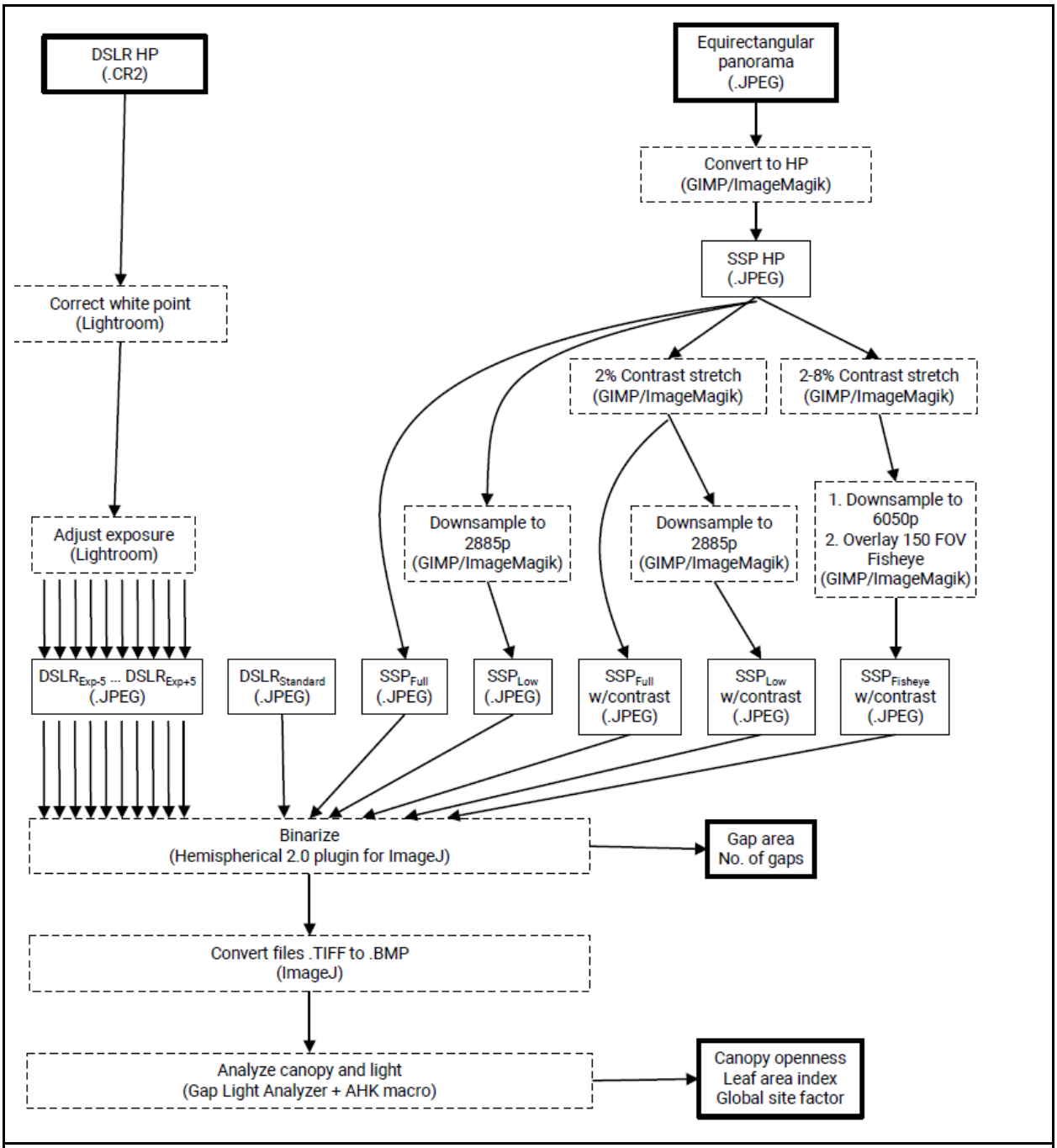

Figure S1. Workflow diagram for processing and analyzing hemispherical photos in this study. Intermediate files are indicated by solid boxes with file type in parentheses. Processing steps are indicated by dotted line boxes with the software required for the step in parentheses. Inputs and outputs are indicated with bold black boxes.

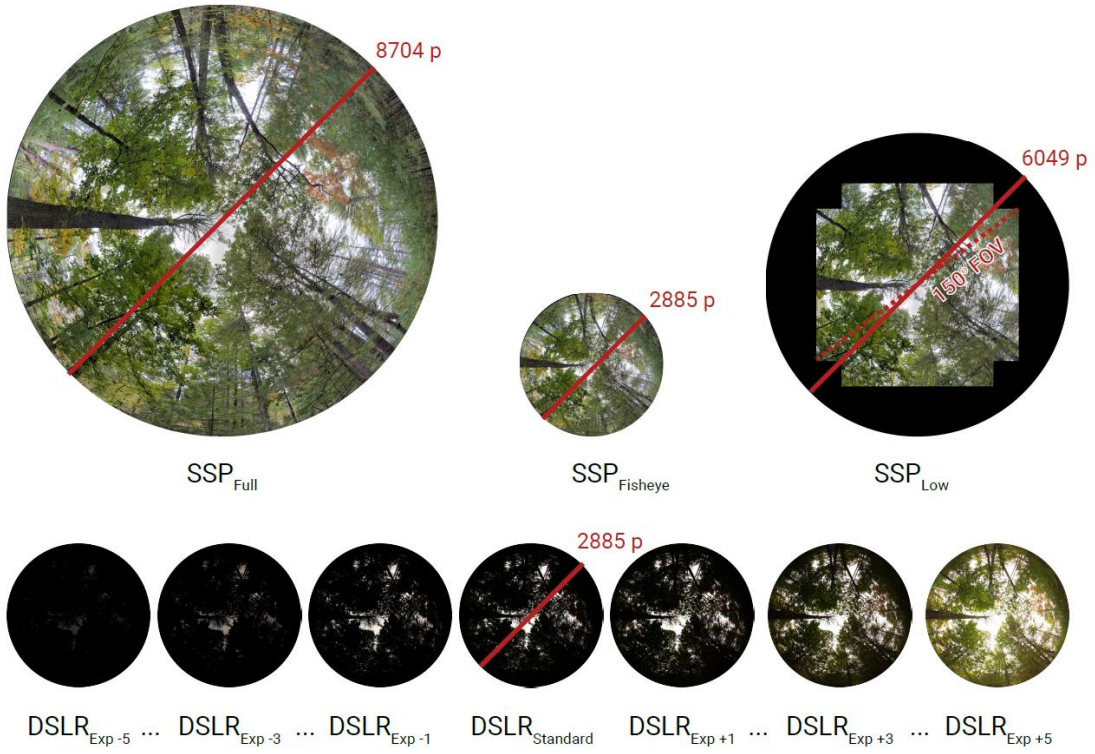

Figure S2. Comparison of the outputs following image processing for the same site. All images are scaled to relative size and pixel width of the diameter of the circular image is shown in red. Images in the top row were developed from smartphone spherical panoramas in full resolution (SSP_{Full}) or downsampled (SSP_{Low}) to the same resolution of DSLR images. Smartphone spherical panoramas were used to simulate the method proposed in Bianchi et al. (2017) ($SSP_{Fisheye}$) wherein a clip-on fisheye lens with 150 degree field-of-view attached to a smartphone was used to take perpendicular canopy images and merged into a partial hemispherical image. Image in the bottom row were acquired with a DSLR camera with a hemispherical lens and developed with standard protocols ($DSLR_{Standard}$). In order to compare relative impacts of incorrect exposure, DSLR images were over- or under-exposed by 1, 2, 3, 4, or 5 stops ($DSLR_{Exp\ -5}$ to $DSLR_{Exp\ +5}$).

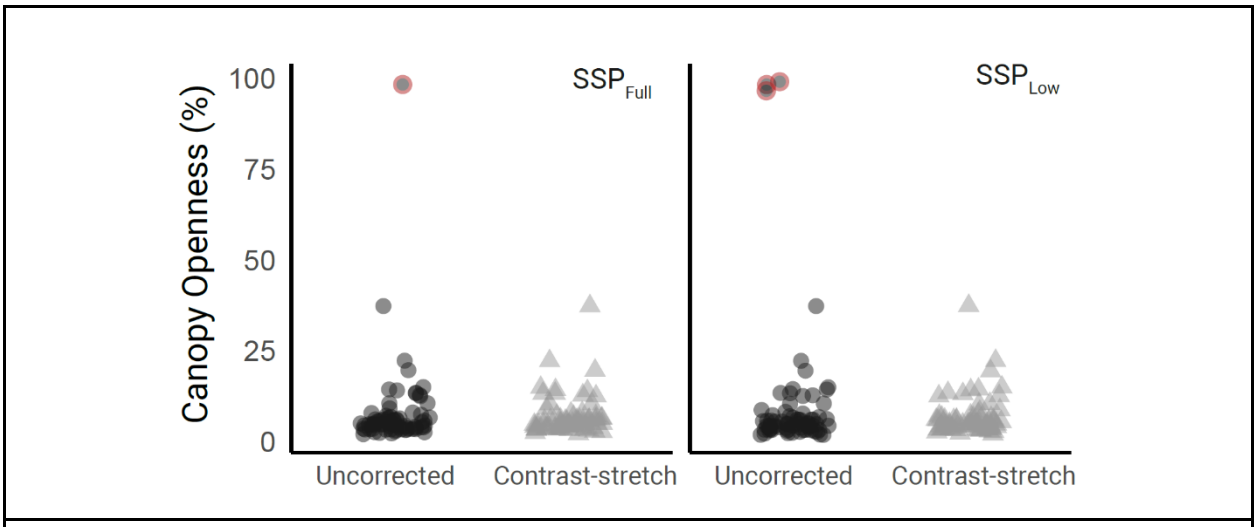

Figure S3. A small number of hemispherical images generated from smartphone spherical panoramas were incorrectly binarized by the thresholding algorithm (points indicated in red). This problem was more prevalent in the down-scaled images (SSP_{Low}) than the full resolution images (SSP_{Full}). Increasing the contrast of the images by 2% prior to binarization solved the problem.

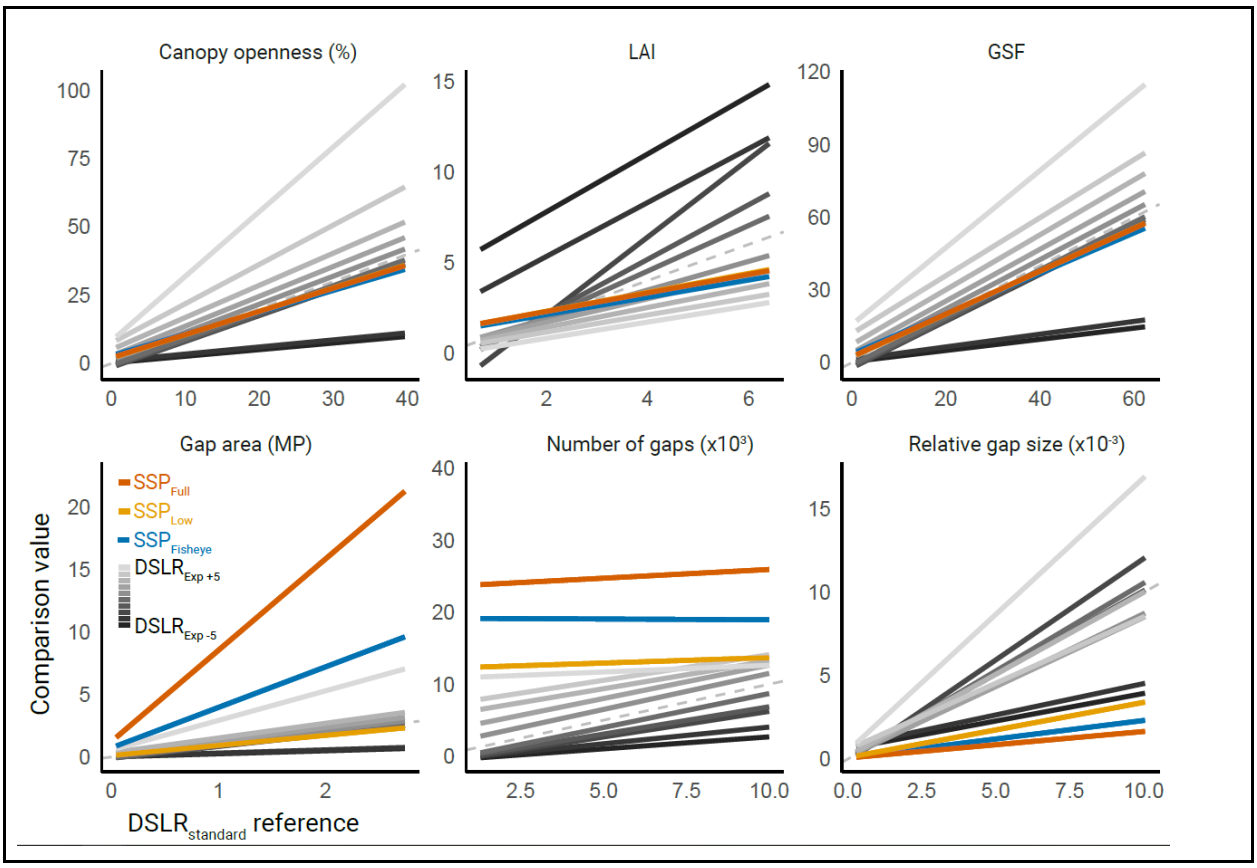

Figure S4. Linear regression estimates comparing three canopy structure metrics (canopy openness, leaf-area index (LAI), and global site factor (GSF)) and three image quality metrics (total gap area, number of gaps, and relative gap size) from images hemispherical images extracted from smartphone spherical panoramas (SSP) to those acquired with traditional DSLR camera and hemispherical lens ($DSLR_{Standard}$). Three processing methods for SSP hemispherical images were compared, including full resolution (SSP_{Full}, dark orange), images down-sampled to match the resolution of DSLR images (SSP_{Low}, light orange), and images simulated to approximate Bianchi et al.'s (2017) method using a smartphone and clip-on fisheye lens ($SSP_{Fisheye}$, blue) are shown. Estimates from DSLR images over- or under-exposed by -5 to 5 stops are shown for comparison. The 1:1 line is indicated by a dashed grey line.

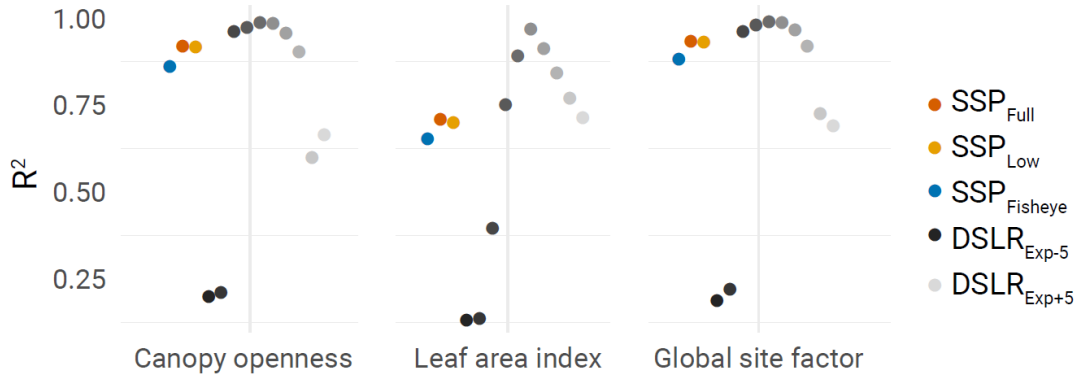

Figure S5. Coefficients of determination from independent OLS regression models predicting canopy structure and light environment values for reference (standard DSLR HP) from full resolution SSP HP (dark orange), low resolution SSP HP downsampled to match the standard DSLR resolution (light orange), fisheye HP (blue), and DSLR HP with exposure adjusted from +5 to -5 (light to dark).

Estimation of canopy clumping index

Methods

Canopy properties and clumping index were estimated by partitioning canopy gaps into within-canopy and between-canopy gaps (Chen and Cihlar, 1995) using a thresholding method derived from Alivernini *et al.* (2018).

Gap size was estimated using the Analyze Particles tool in ImageJ on inverted binary images output from Hemispherical 2.0 plugin. Gaps that greater than one standard error in excess of the mean of the distribution of gap sizes for an image were considered as large, between-canopy gaps. The proportion of total gap area (g_T) or large gap area (g_L)

to circular image area (p_r) was used to calculate total gap fraction (GF_T) and large gap fraction (GF_L), respectively. Crown cover (CC), foliage cover (FC), crown porosity (CP), and clumping index (CI) were estimated (following Alivernini *et al.* (2018)) as:

$$CC = 1 - GF_L$$

$$FC = 1 - GF_T$$

$$CP = 1 - FCCC$$

$$CI = (1 - CP) * \ln(1 - FC)\ln(CP) * FC$$

Results

Table S3. Estimates of canopy properties from hemispherical photos captured at 72 sites using smartphone spherical panoramas (SSP) HP or DSLR HP, and the difference in estimates (SSP HP - DSLR HP). The mean value is shown followed by standard deviation in parentheses.

Method	Number of large gaps	GF_T	GF_L	CC	FC	CP	CI
SSP HP	2067 (897)	6.76 (5.34)	5.85 (5.06)	0.94 (0.05)	0.93 (0.05)	0.010 (0.004)	0.64 (0.05)
DSLR HP	292 (174)	5.56 (6.12)	4.89 (5.84)	0.95 (0.06)	0.94 (0.06)	0.007 (0.005)	0.66 (0.08)
Difference	1774 (806)	1.20 (1.75)	1.00 (1.56)	-0.01 (0.02)	-0.01 (0.02)	0.003 (0.003)	-0.02 (0.02)

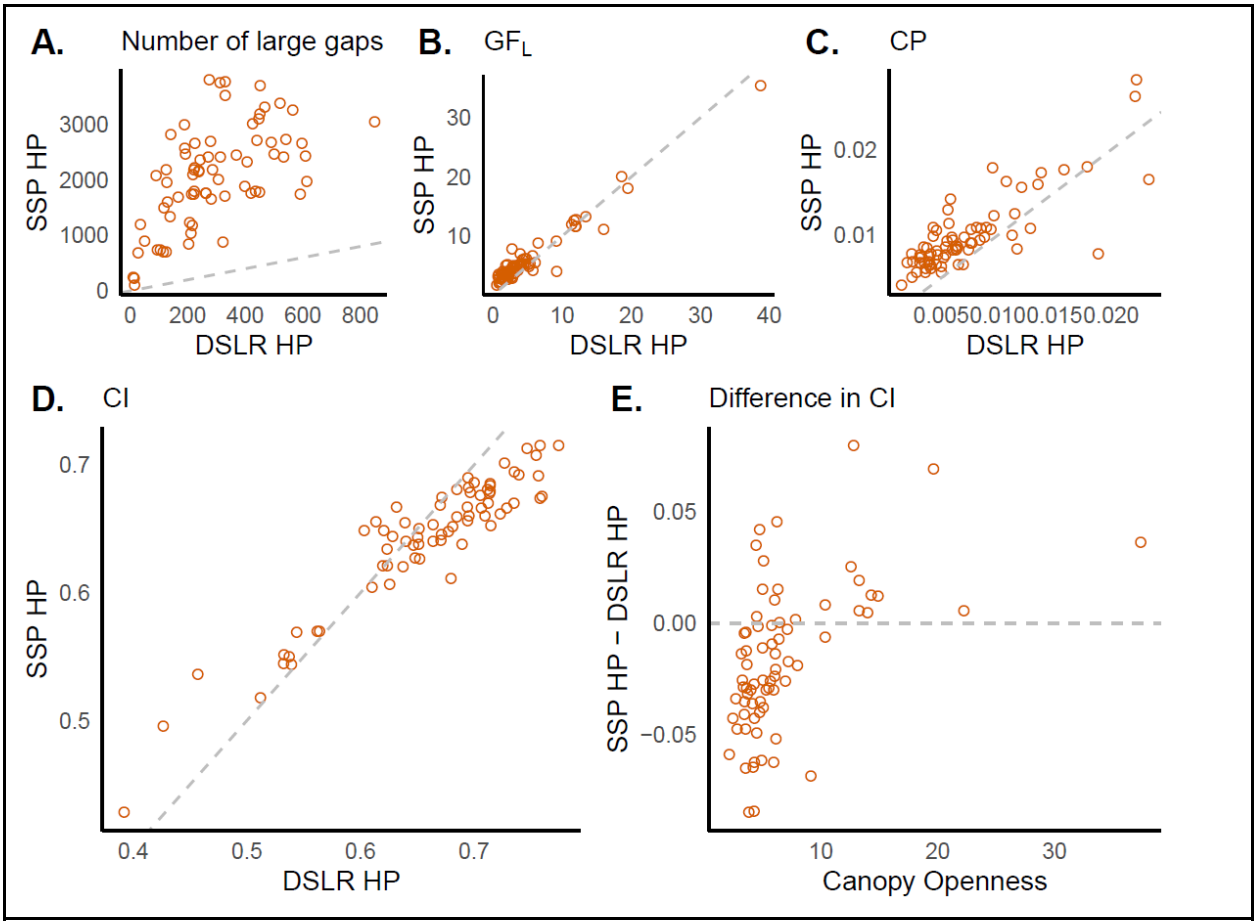

Figure S6. Comparison of canopy properties including the number of large (i.e. between-canopy) gaps (A), large gap fraction (B), canopy porosity (C), and clumping index (D) from hemispherical photos captured at 72 sites using DSLR HP or smartphone spherical panoramas (SSP) HP. The difference in clumping index estimated from images produced with SSP HP or DSLR HP relative to canopy openness is shown in E. The 1 to 1 line (dashed) is shown in all panels.

References:

Alvernini, A., Fares, S., Ferrara, C. & Chianucci, F. (2018). 'An objective image analysis method for estimation of canopy attributes from digital cover photography', *Trees*, 32, pp. 713-723.

Chen, J. M., and Cihlar, J. (1995). Plant canopy gap-size analysis theory for improving optical measurements of leaf-are index. *Applied Optics*, 34(27), 6211–6222.

CHAPTER 2

Phenological delay despite warming in wood frog (*Rana sylvatica*) reproductive timing: a 20-year study

Publication Status: Arietta, A. Z. A., Freidenburg, L. K., Urban, M. C., Rodrigues, S., Rubinstein, A., and Skelly, D. K. (2020). Phenological delay despite warming in wood frog Rana sylvatica reproductive timing: a 20-year study. *Ecography.* DOI: 10.1111/ecog.05297.

ABSTRACT

Across all taxa, amphibians exhibit some of the strongest phenological shifts in response to climate change. As climates warm, amphibians and other animals are expected to breed earlier in response to temperature cues. However, if species use fixed cues such as daylight, their breeding timing might remain fixed, potentially creating disconnects between their life history and environmental conditions. Wood frogs (*Rana* sylvatica) are a cold-adapted species that reproduce in early spring, immediately after breeding ponds are free of ice. We used long-term surveys of wood frog oviposition timing in 64 breeding ponds over 20 years to show that, despite experiencing a warming of 0.29 °C per decade in annual temperature, wood frog breeding phenology has shifted later by 2.8 days since 2000 (1.4 days per decade; 4.8 days per °C). This counterintuitive pattern is likely the result of changes in the timing of snowpack accumulation and melting. Finally, we used relationships between climate and oviposition between 2000 and 2018 to hindcast oviposition dates from climate records to model longer-term trends since 1980. Our study indicates that species can respond to fine-grained seasonal climate heterogeneity within years that is not apparent or counterintuitive when related to annual trends across years.

INTRODUCTION

As global temperatures and seasonal climatic variability continue to increase, so does interest in understanding how climate change will impact biodiversity. One major concern is that the extent of environmental change may outpace the limits of biological response with negative effects leading to population declines, changes in species distributions, and potentially extinction for species with limited response capacity (Pimm et al. 1995, Thuiller et al. 2004, Urban 2015). These biological changes could eventually affect not just global diversity patterns, but also entire ecosystems and the services that they provide to humans (Bongaarts 2019).

Many organisms exhibit an annual cycle of life history events, or phenology, that progress in concert with environmental conditions such as seasonal variation in temperature, precipitation, or photoperiod (Scranton and Amarasekare 2017). As global climate continues to shift, phenological events cued by temperature are expected to advance as long-term average temperatures increase and warming occurs earlier in the year (Root et al. 2003, Thackeray et al. 2016). In fact, many species have shown considerable phenological shifts, usually to earlier in the season (Walther et al. 2002, Parmesan and Yohe 2003, Cohen et al. 2018). These shifts are most pronounced for populations in the temperate zone and for ectotherms (Cohen et al. 2018).

If important components of a species' community and ecosystem all advance at similar rates, then these phenological shifts may benefit the species by allowing it to track shifting climates. However, different species respond to different phenological cues with different sensitivities, which can result in asynchronous shifts among interacting species (e.g. Ovaskainen et al. 2013). For example, the average flowering date of northeastern North American plants is advancing under climate change and wild bees that rely on pollen show

similarly advancing phenology, maintaining synchrony in the mutualistic relationship (Bartomeus et al. 2011). As a counter example, great tits (*Parus major*) rely on seasonal availability of caterpillar food resources during nesting; however, climate change is causing an advance in the caterpillar phenology. While great tits exhibit advances in laying dates over the same period, the change is not enough to take advantage of peak caterpillar food resources (Visser et al. 1998).

Amphibians exhibit the largest phenological shifts of any vertebrate taxa, which may make this group particularly vulnerable to phenological asynchronies (Parmesan 2006, While and Uller 2014, Thackeray et al. 2016). Disconcertingly, climate change and phenological mismatch have been suggested to contribute to the trend in global amphibians declines (Stuart et al. 2004). Life history and reproductive behavior are regulated primarily by temperature for most temperate amphibians, including emergence from hibernation, initiation of reproductive migration, oviposition timing, etc. (Reading 1998). Amphibian life cycle timing is expected to advance with warming climates as organisms emerge, breed, and initiate development earlier in the spring (Hughes 2000). However, at least some amphibians use photoperiod cues in combination with other environmental cues to time their breeding (Canavero and Arim 2009).

We analyzed climate and oviposition timing over 20 years in a metapopulation of wood frogs (*Rana sylvatica*) in northeastern Connecticut, USA to evaluate long-term patterns in reproductive phenology relative to changes in key climate variables. Climate data suggest that this region has rapidly warmed in concert with the rest of the northern hemisphere. Given the sensitivity of wood frog breeding phenology to temperature (Benard 2015), we hypothesize that wood frog breeding will have advanced to occur earlier in response to increasing temperatures. Alternatively, wood frogs might also use fixed cues such as photoperiod, which would mitigate phenological changes in response to climate.

METHODS

Natural History

In this study, we consider reproductive phenology of the wood frog, a wide-spread North American anuran (Layne and Lee 1995, Lee-Yaw et al. 2008, Larson et al. 2014). The phenology of the wood frog is closely tied to seasonal climate and selection favors the earliest possible reproduction and oviposition, balanced by the threat of egg mortality (if ponds freeze after oviposition) and pond desiccation at the end of the larval period. Most mortality (92-99%) occurs in pre-metamorphic life stages (Dodd 2013) and is often associated with predation or drying of breeding ponds (Camp et al. 1990, Berven 1995, DiMauro and Hunter 2002), which is, itself a function of climate (Brooks 2004). Early reproduction extends the larval period and increases the time available for development before ponds dry. In other amphibians, early oviposition also allows larvae to reach a critical body size threshold that provides refuge from gape-limited predators (Urban 2007). Intraspecific competition for mates also favors males that arrive earliest to breeding ponds (Waldman 1982).

Wood frogs possess a number of traits that facilitate early spring breeding and make the species particularly sensitive to climate. Adult wood frog males overwinter immediately adjacent to breeding ponds where they select hibernacula beneath forest litter (Regosin et al. 2003, Groff et al. 2017). Proximity to ponds enables breeding to commence as soon as the climate is favorable. Similarly, proximity to the ground surface enables individuals to emerge from hibernation as soon as possible, but simultaneously increases exposure to freezing temperatures (O'Connor and Rittenhouse 2016). Physiological adaptations for freeze-tolerance (Storey and Storey 1984, Storey 1990, Layne and Lee 1995, Costanzo et al. 2013), rapid reanimation post-freezing (Layne and First 1991), and fine-tuned microhabitat

selection (O'Connor and Rittenhouse 2016) ameliorate this problem. As a result, wood frogs are generally the earliest breeding amphibians, emerging long before ice and snow have diminished (Dodd 2013).

Seasonal timing of oviposition across the wood frogs' range varies positively with latitude and elevation (Berven 1982, Guttman et al. 1991). Thus, one would expect wood frog breeding and oviposition timing to similarly advance with increasing average temperatures. This is the case for the common frog (*Rana temporaria*), a related species found in Europe. Carroll et al. (2009) analyzed more than 70,000 records of oviposition dates for common frog populations across the UK between 1998 and 2007 and found an average advance of 10 days, although this did not correspond to earlier hatching (Scott et al. 2008).

The wood frog populations included in this study breed in vernal pools ranging from 190 m to 296 m in elevation across 0.06 degrees of latitude (8 km). Canopy cover, which also impacts local climate, ranges from nearly completely open (6.5%) to completely closed (98.3%) (see methods for estimating canopy cover below).

Climate

We extracted daily meteorological data, including maximum and minimum temperature, precipitation, and snowpack (quantified as snow water equivalent) from the DayMet database (v.3; Thornton et al. 2016). The Daymet algorithm uses daily meteorological observations from ground stations to interpolate estimates at 1km square cells across North America between 1980 and 2018. We extracted estimates for the cell centered nearest the geographical center of our site (tile: 11754, N 41.9198 W -72.1604). These data were indistinguishable from estimates extracted for tiles encompassing the

northernmost and southernmost extent of our site. We computed average daily temperatures as the midpoint between daily minimum and maximum temperatures.

In all cases, we transformed calendar dates into two variables, Year and day-of-year values, to enable modeling long-term trends versus annual variability and to account for leap years. We tested for long-term trends in meteorological variables over time by regressing the annual averages against Year. We confirmed that linear models were appropriate, first, by visually comparing linear and loess regression lines. Annual trends may mask important shifts in seasonal climate at shorter temporal scales. So, secondarily, we assessed non-linear, intra-annual trends by regressing the linear function of Year and penalized spline functions of day-of-year against each meteorological variable using generalized additive models implemented with 'mgcv' (v. 1.8.26; Wood 2011). For each model we include the interaction between the smoothing terms and Year. The interaction term allows the seasonal shape of the curve to vary across year while no interaction would indicate a shift in the mean but no seasonal change. All statistical analyses were performed in *R* version 3.5.2 (R Core Team 2017).

Oviposition phenology

We used records of annual egg-mass surveys collected between 2000 and 2019 for 64 wood frog breeding ponds at Yale Myers Forest, a 3,213 ha research forest in northeastern Connecticut, USA (Figure 1). During surveys, the entire pond was closely searched by observers. If eggs were found, the observers independently counted and then averaged the estimates (or, in cases of surveys by a single observer, egg masses were counted twice and averaged).

Wood frogs are explosive, aggregate breeders (reviewed in Dodd 2013). Breeding commences shortly after females emerge from hibernation and oviposition occurs in

synchrony over a short timespan (reviewed in Dodd 2013). Within a pond, oviposition lasts just a few days (~ 2-8, as reviewed in Dodd 2013) and, among proximal ponds, oviposition is generally concurrent (Petranka et al. 2004). At our site, eggs hatch between 9 to 18 days after oviposition, depending on water temperature (Skelly 2004). Surveys were conducted approximately 7 days after oviposition each year to ensure oviposition had concluded and before egg masses swelled too much to be individually distinguished. Thus, the survey dates closely correspond to oviposition timing and we use the survey dates as a proxy for mean annual oviposition within a pond population.

While survey dates are an imperfect proxy, other measures of oviposition timing, such as records of first or last appearance or calling, are similarly subject to both methodological and biological error (Tryjanowski et al. 2003). In ten years with protracted breeding, multiple surveys were conducted at each pond and cumulative egg mass counts were recorded. On average, counts differed for 3-4 ponds. In these cases, we estimated the oviposition date for each pond as the average survey day weighted by the number of new egg masses recorded in subsequent resurveys. We excluded observations in which no eggs were recorded in ponds.

Our oviposition data include 1103 observations across 64 ponds between 2000 and 2019. On average, we have records for 17.2 years (sd = 5.0, median = 19) for each pond, a total of 86% data coverage across all pond-year combinations. We excluded years for which we surveyed ponds but found no eggs (n = 278), leaving 825 oviposition date observations for further analysis. We tested the long-term change in wood frog phenology over time by fitting a linear mixed model in 'lme4' (v. 1.1.21**;** Bates et al. 2015) predicting oviposition date by Year, including the site ID from the pond of observation as a random intercept both with and without allowing the slopes of the relationship with Year to vary among ponds as a

random effect. We calculated 95% confidence intervals for this estimate with 1000 bootstrap iterations. This approach provided an estimated change in site-wide oviposition.

Hindcast modeling long-term phenology

Although the 20-year duration of our dataset is one of the longer continuous oviposition datasets available, it is a short window compared to the local ecological and evolutionary history of wood frogs. We tested the similarity between the short-term trend in our observed dataset and putative long-term change by comparing our results to a hindcast (or "retrodiction" *sensu* Green (2017)) dataset wherein we used our extant oviposition dates to train a random forest model to predict past oviposition dates.

Random forest is a powerful and flexible model-averaging technique that leverages many individual regression trees built with random subsets of the data and random subsets of predictor variables to estimates the outcome resulting from complex relationships among independent variables. Because these models can accommodate many correlated independent variables, we built our random forest from a robust set of pond-wise and year-wise predictors in order to capture the natural variability of pond micro-habitats at our site.

As pond-wise variables, we included latitude, elevation, aspect, and a suite of canopy metrics estimated from repeated hemispherical photos. Briefly, five hemispherical photographs were take along the shore at each cardinal point and at the center of each pond during leaf-off and leaf-on seasons. We used Gap Light Analyzer (Frazer 1999) to estimate average leaf-on and leaf-off global site factor (GSF; the ratio of above-canopy radiation to under-canopy radiation (Anderson 1964)) and a weighted GSF value integrated over the duration of wood frog larval and embryonic life cycle (Halverson et al. 2003). Additionally, we included the within-pond variance in GSF values, which captures important components of the canopy shape and structure above and surrounding the ponds. Point values for aspect

and elevation were estimated from the USGS National Elevation Dataset (U.S Geological Survey 2002). In addition, we included site-wide daily average temperature, precipitation, and SWE between day-of-year 0 and 120 (Jan 1 ~ April 29) from the DayMet dataset. This range spans the earliest and latest observed oviposition dates (day-of-year 74 and day-of-year 113) with a margin. Finally, we included a random noise variable in order to assess rank importance of all our included variables.

We grew our random forest from 1000 regression trees with package 'randomForest' (v. 4.10; Liaw and Wiener 2002) and assessed the predictive accuracy of our model through 5-fold cross validation. We then used the model fit to all our observations to hindcast oviposition dates between 1980 and 1999. We fit an ordinary least squares linear model to the combined datasets of hindcasted and observed site-wide average oviposition dates to test for long-term phenological change and compared these to the short-term trend displayed by our egg mass surveys. Our hindcast predictions assume that every pond hosted an oviposition event in every year which is almost certainly untrue. Thus, we consider the average oviposition date each year as an approximation of the metapopulation-wide oviposition date and use bootstrap resampling to estimate 95% confidence intervals.

Although random forests are useful for prediction, they are "black-box" models that are not easily interpretable. So, in order to assess the relative importance and direction of the relationship of climatological variables on oviposition dates, we computed 10-, 20-, 30-, and 40-day moving averages for daily temperature, precipitation, and snow water equivalent between 2000 and 2019 between day-of-year 0 and 160 (Jan 1 ~ June 9). For each variable, we calculated the Pearson's correlation between each window-average and observed oviposition timing averaged across all ponds for the 20-year overlap. We plotted the correlation coefficients over time to determine at what point in the season each variable is most likely to influence oviposition timing, determined by the span for non-overlapping

windows with the greatest absolute correlation to oviposition timing greater than a 95% confidence threshold (r = 0.45).

RESULTS

Climate

There has been a substantial increase in annual temperature over time (Figure 2a). Annual temperature increased at a rate of 0.29 °C per decade, or an estimated total increase of 1.1 °C since 1980 ($F_{1,36}$ = 10.19, slope = 0.029, p = 0.003). This rate is consistent with estimated increases in the northeast region of the US (Hartfield et al. 2018). There is no evidence of net change in annual precipitation ($F_{1,36}$ = 0.000, slope = -0.000, p = 1.0) or in annual snow water equivalent ($F_{1,36}$ = 0.002, slope = -0.009, p = 0.97) from 1980 to 2018 (Figure 2b, 2c). Although loess curves suggest a quadratic fit for precipitation and snowpack change, second order polynomial fits were also non-significant.

We do see seasonal shifts in daily temperatures and snowpack, but not in precipitation across years (Table 1). The greatest seasonal shift in daily temperatures is seen in the latter portion of the year, after the annual peak temperature around day-of-year 200 (July 19) (Figure 3). Daily temperatures shifted the least in the colder portion of the year from day-of-year 1 (Jan 1) to approximately day-of-year 100 (~April 10) (Figure 3). Although there is no evidence of an overall increase in the snowpack, the timing of accumulation and melt is delayed. The models estimate that snowpack accumulates approximately 28 days later, and persists approximately 27 days longer into spring, while the total snow water equivalent of the snowpack in mid-winter and early spring, between day-of-year 13 (Jan 13) and day-of-year 105 (~April 15) remains stable (Figure 3). The seasonal oviposition window coincides with a period of relatively low shift toward warmer temperatures, minimally decreased precipitation, and increased snowpack (Figure 3).

However, the magnitude of the change in air temperature (0.62 - 2.67 SD) is much larger than either precipitation (0 - 0.19 SD) or snowpack (0 - 0.50 SD) (Figure 3).

Oviposition phenology

The earliest and latest oviposition dates on record are day-of-year 74 (~ Mar 15) and day-of-year 113 (~ Apr 23), respectively, with a median of day-of-year 96 (~ Apr 6). In some years, ponds at the southern portion of the site or ponds with more open canopies breed earlier. On average, oviposition across ponds within a given year took place within a short period of 6 days (sd = 6.1 days). Fitting a mixed model allowing slopes to vary by pond did not increase predictive power (with (AIC = 5939.6) and without (AIC = 5937.6) random slopes). The multi-level mixed model fit to survey dates between 2000 and 2019 with random intercept by pond shows a delaying trend of 1.4 days per decade (4.8 days per °C) in site-wide oviposition timing (n = 825, slope = 0.139, se = 0.05, 95% CI = 0.03, 2.44) (Figure 4). This pattern contradicts our hypothesis and the general trend seen in most amphibians of phenological advancement in spring reproductive timing in response to warming temperatures.

Hindcast model

Our random forest model fit well (mean out-of-bag R^2 = 0.94) and yielded high predictive accuracy (RMSE = 2.34). On average our predicted oviposition dates were with 1.6 days of the true oviposition date of our hold-out cross-validation observations. The slope of the regression fit to observed oviposition dates (2000-2019) and hindcast predicted dates (1980-1999) provides no support for long-term directional change in phenology (n = 40, slope = -0.069, SE = 0.093, p = 0.45, 95% C.I. = -0.26, 0.17) with a 95% confidence interval from 1000 iterations that includes delay up to 1.7 days/decades and advance up to 2.4

days/decades (Fig. 5). Temperature was the most important site-wide variable in the random forest model, followed by precipitation and snowpack (Supplemental Figure S1). The maximum temperature on day-of-year 69, just prior to oviposition in most years, was the most important variable accounting for model accuracy, followed by precipitation on day-of-year 56. The importance of daily snowpack values was low but consistent. Only a single pond-wise variable, latitude, was ranked as high importance in the model.

When considered individually, all meteorological variables showed significant bivariate correlation with oviposition dates for at least one seasonal window (Supplemental Figure S2). The periods of strongest correlation between temperature and oviposition timing included a negative relationship at the outset of oviposition from day-of-year 66 to 106 and a weaker but significantly negative correlation earlier in the season from day-of-year 25 to 45. A negative relationship between temperature and oviposition means that warmer temperatures are correlated with earlier oviposition (i.e. lower day-of-year). There is a positive relationship between oviposition and precipitation from day-of-year 93 to113 which means that greater precipitation in this window is correlated with later oviposition dates. There is a positive relationship between oviposition and snow water equivalent from day-of-year 77 to 87, just prior to a period of significant correlation with radiation from day-of-year 80 to 90. Some climate windows with significant correlation to oviposition occur after the frogs breed, indicating that these are temporally autocorrelated, not causal.

DISCUSSION

Most studies across taxonomic groups, including all prior studies of amphibians, have documented advancement of breeding phenology in response to changing climate (e.g. Carroll et al. 2009, Green 2017). Although phenological delay has been documented in some amphibian species (Todd et al. 2011), these reports involve fall-breeding species in which

delay is expected with warmer summers. The present study is one of the first to find a delay in reproductive timing (2.8 d over 20 years) in a spring-breeding amphibian. As in other studies that have documented advances in timing, we are working in a system in which temperatures have increased in recent decades (0.6 C over 20 years).

We included a hindcast modelling procedure to evaluate whether estimated long-term patterns of oviposition negate the delaying trend we see in our shorter-term observations. Green (2017) used a similar approach to show that prior estimates of phenological delay in Fowler's Toads (*Anaxyrus fowleri*) were incompatible with long-term estimates from hindcasted breeding dates that suggested advancing phenology. In contrast to short-term estimates of phenological delay in oviposition, our hindcasted dataset of annual metapopulation-wide oviposition show no evidence of directional change. However, confidence intervals for the long-term hindcast model include our slope estimate of 1.4 days/decade. The fact that our model fit to hindcasted data showed no evidence of change in oviposition may be due to lack of power as this model cannot account for intra-annual variation in oviposition dates among ponds nor years when ponds host no breeding events.

Among animal taxa, amphibians exhibit some of the strongest relationships between changing climates and phenological shifts. A meta-analysis across all animals found that amphibians demonstrated an average spring advancement of 7.6 days per decade, almost double the average of all taxa combined (Parmesan 2007). An analysis of United Kingdom taxa found an average advancement of 2.6 days per °C, with amphibians exhibiting an average advancement of 3.4 days per °C, similar to other highly sensitive groups--plants, freshwater phytoplankton, and insects (Thackeray et al. 2016). A more recent meta-analysis of global responses estimated an average advancement of 3.23 days per decade for amphibians (Cohen et al. 2018). Similarly, a meta-analysis of amphibian breeding phenology found an average advancement of 6.09 days per decade +/- 1.65 days (While and Uller

2014). In the Northeast, Gibbs and Breisch (2001) found that four species of amphibians now initiate breeding calls 10-13 days earlier than a century earlier.

Although there has been a marked increase in annual average temperatures at our site over the past few decades, the magnitude of the shift varies across the season. Relative warming is greater later in the season while the least warming occurs in the late winter and early spring, coinciding with oviposition. Similarly, we see no evidence of annual change in snowpack, but we do see evidence of seasonal shifts, such that snow accumulates later and persists longer into spring by up to 27 days. The positive correlation between snow, precipitation, and low temperatures may explain why snowpack was not ranked as highly important in our hindcast model. Later persisting snow may be driving the counterintuitive phenological delay if snow cover prevents adults wood frogs from emerging and migrating to ponds. Alternatively, snow cover over ponds may insulate the ice, delaying ice-out and preventing breeding, even if frogs emerged earlier. Temperature and precipitation can interact to increase the snowpack when precipitation falls during periods below freezing. However, temperature and precipitation can rapidly reduce snowpack when rain falls and melts snow. While snowpack is certainly an important factor in wood frog reproductive timing, the interaction between temperature and precipitation exhibits a stronger effect.

Our results indicate that phenology is closely tied to complex, intersecting changes in climate across the season. Furthermore, our results highlight the danger in using annual averages or coarse-grained, *a priori* selection of seasonal windows as predictors of phenology. We show that climatic changes can be heterogeneous within years with or without annual climatic change across years. Indeed, predictive frameworks that rely on mean annual rates of climatic change have come under criticism for masking potentially relevant asynchronous seasonal variation (Straile et al. 2015, Senner et al. 2018). The relative impact of seasonal variations in climate depends on the ecology of the species in

question and how climatic shifts across the season interact with seasonally correlated life-history traits, like phenology. Thus, relating fine-grained, seasonal changes in interacting climatic variables, to trait change, can help resolve idiosyncratic effects of climate change and better relate these unexpected departures to general trends (Muths et al. 2017, Kirk et al. 2019).

Although our study site likely represents a single meta-population, variation in oviposition timing across ponds illustrates fine-grained spatial heterogeneity in responses to climate change (Senner et al. 2018). In our study, spatial difference in latitude between our ponds was an important variable in predicting oviposition timing. This variation is minimal at the scale of our study, but reflects similar variability at larger spatial scales (Sheridan et al. 2018). For example, Sheridan et al. (2018) tested for relationships between climatic variables and breeding phenology across the wood frog range from museum specimens dating back to 1901. They found that the day of first collection at a site (assumed to indicate breeding timing) was delayed in coastal Alaska and the coastal Northeast in association with either very large or very small proportional increases in frost free days, while average increases were associated with phenological advancement in breeding, indicating high spatial heterogeneity in wood frog responses to climate change. Examining phenological responses to climate change at small scales may help illuminate the processes driving range-wide patterns.

If organismal responses are cued by seasonal climates, variation in climate across space and time may result in lifehistory mismatch. For instance, asymmetric shifts in climate during early lifestages may delay or accelerate development and result in suboptimal timing for later life history events, like metamorphosis, reproduction, or hibernation (Van Dyck et al. 2015; Hale et al. 2016). Similarly, trophic mismatches can result if predators and prey respond to cues at different times in the season (Straile et al.

2015). Spatial variability in climate cues may affect some populations more or less and could lead to changes in migration rates or source-sink dynamics. Local extinction may increase in ponds that change unfavorably while populations in ponds with favorable change could increase due to immigration or intrinsic population growth (Miller-Rushing et al. 2010). If phenological shifts driven by climate are maladaptive, selection may cause species to compensate with other mechanisms (Root et al. 2003, Hoffmann and Sgrò 2011, Merilä and Hendry 2014). For example, some species exhibit adaptive differences in physiological rates which are both plastic and genetic (e.g. Berven 1987, Conover & Present 1990). Thus, the impacts of climate change will vary considerably at temporal and spatial scales not typically considered in global predictions.

In conclusion, we demonstrate that, despite seasonal shifts in temperature that would generally predict advances in breeding phenology, we find the opposite trend of phenological delay, likely due to the interacting effect of more persistent snowpack. Our results belie the strong presumption that global warming will shift phenology in a particular direction and highlight the need to closely investigate the drivers of biotic change, even when gross patterns match expectations. Otherwise, we may miss importance but unexpected mechanisms. Coupling long-term biological and climatological datasets at fine-spatial scales and temporal resolution aids our understanding of counterintuitive phenological patterns and can help generate predictions for future responses. The next step is to evaluate how shifts in phenology might translate into demographic differences. Such data are necessary to move beyond coarse estimates of species responses and start to generate more mechanistic models that make predictions based on biology rather than correlations (Urban et al. 2016). These insights will be critical for deciding which species to conserve and how to design conservation programs to mitigate the effects of climate change.

Acknowledgements – We thank Nicole Freidenfelds, Meredith Holgerson, Max Lambert, Jonathan Richardson, Meredith Smylie, Greg Watkins Colwell and many others for help with field sampling.

Funding – This study was supported by grants from the National Science Foundation, the Yale Institute for Biospheric Studies, Yale Natural Lands Grant, the Schiff Fund for Wildlife, and by support from Yale University. This project benefited from support by the Kohlberg-Donohoe Fellowship.

Author contributions – A. A., L. F., S. B., A. R., M. U., and D. S. contributed to data collection and field study. A. A. performed the analysis. D. S. conceived of the study. A. A. and D. S. wrote the manuscript with input from L. F., A. R., and M. U.

Permits – Animal handling was approved by Yale University IACUC (2004-10361, 2007-10361, 2010-10361, 2013-10361, 2016-10361, 2019-10361, 2006-11024, 2009-11024, 2009-11040) and field surveys were conducted with permission from Yale Myers Research Committee (SKE01, AND17).

REFERENCES

Anderson, M. C. 1964. Studies of the woodland light climate I. The photographic computation of light condition. - Ecology **52**: 27–41.

Bartomeus, I. et al. 2011. Climate-associated phenological advances in bee pollinators and bee-pollinated plants. - Proc. Natl. Acad. Sci. U. S. A. **108**: 20645–20649.

Bates, D. et al. 2015. Fitting linear mixed-effects models using lme4. – J. Stat. Softw. **67**: 1–48.

Benard, M. F. 2015. Warmer winters reduce frog fecundity and shift breeding phenology, which consequently alters larval development and metamorphic timing. Glob. Chang. Biol. **21**: 1058–1065.

Berven, K. A. 1982. The genetic basis of altitudinal variation in the wood frog *Rana sylvatica*. I. An experimental analysis of life history traits. - Evolution **36**: 962–983.

Berven, K. A. 1987. The heritable basis of variation in larval developmental patterns within populations of the wood frog (*Rana sylvatica*). - Evolution **41**: 1088–1097.

Berven, K. A. 1995. Population regulation in the wood frog, *Rana sylvatica*, in three diverse geographic localities. - Aust. J. Ecol. **20**: 385–392.

Bongaarts, J. 2019. IPBES, 2019. Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. - Population and Development Review **45**: 680–681.

Brooks, R. T. 2004. Weather-related effects on woodland vernal pool hydrology and hydroperiod. - Wetlands **24**: 104–114.

Camp, C. D. et al. 1990. Oviposition, larval development, and metamorphosis in the Wood Frog, *Rana sylvatica* (Anura: Ranidae) in Georgia. - Brimleyana **16**: 17–21.

Canavero, A. and Arim, M. 2009. Clues supporting photoperiod as the main determinant of seasonal variation in amphibian activity. - Journal of Natural History **43**: 2975–2984.

Carroll, E. A. et al. 2009. Influence of temperature on the spatial distribution of first spawning dates of the common frog (*Rana temporaria*) in the UK. - Glob. Chang. Biol. **15**: 467–473.

Cohen, J. M. et al. 2018. A global synthesis of animal phenological responses to climate change. - Nat. Clim. Chang. **8**: 224–228.

Conover, D. O. and Present, T. M. C. 1990. Countergradient variation in growth rate: compensation for length of the growing season among Atlantic silversides from different latitudes. - Oecologia **83**: 316–324.

Costanzo, J. P. et al. 2013. Hibernation physiology, freezing adaptation and extreme freeze tolerance in a northern population of the wood frog. - J. Exp. Biol. **216**: 3461–3473.

DiMauro, D. and Hunter, M. L. 2002. Reproduction of amphibians in natural and anthropogenic temporary pools in managed forests. - For. Sci. **48**: 397–406.

Dodd, C. K., Jr 2013. Frogs of the United States and Canada. - The Johns Hopkins Univ. Press.

Frazer, G. W. et al. 1999. Gap Light Analyzer (GLA): Imaging software to extract canopy structure and gap light transmission indices from true-color fisheye photographs, users manual and documentation. - Simon Fraser Univ.

Gibbs, J. P. and Breisch, A. R. 2001. Climate warming and calling phenology of frogs near

Ithaca, New York, 1900-1999. - Conserv. Biol. **15**: 1175–1178.

Green, D. M. 2017. Amphibian breeding phenology trends under climate change: predicting the past to forecast the future. - Glob. Chang. Biol. **23**: 646–656.

Groff, L. A. et al. 2017. Amphibian terrestrial habitat selection and movement patterns vary with annual life-history period. - Can. J. Zool. **95**: 433–442.

Guttman, D. et al. 1991. Observations on the breeding immigration of wood frogs *Rana sylvatica* reintroduced in east-central Missouri. - Am. Midl. Nat. **125**: 269–274.

Hale, R. et al. 2016. Evolutionary traps and range shifts in a rapidly changing world. - Biol. Lett. **12**: 20160003

Halverson, M. A. et al. 2003. Forest mediated light regime linked to amphibian distribution and performance. - Oecologia **134**: 360–364.

Hartfield, G. et al. 2018. State of the Climate in 2017. - Bull. Am. Meteorol. Soc. **99**: Si–S310.

Hoffmann, A. A. and Sgrò, C. M. 2011. Climate change and evolutionary adaptation. - Nature **470**: 479–485.

Hughes, I., I. 2000. Biological consequences of global warming: is the signal already apparent? - Trends Ecol. Evol. **15**: 56–61.

Kirk, M. A. et al. 2019. Seasonal differences in climate change explain a lack of multi-decadal shifts in population characteristics of a pond breeding salamander. - PLoS One **14**: e0222097.

Larson, D. J. et al. 2014. Wood frog adaptations to overwintering in Alaska: new limits to freezing tolerance. - J. Exp. Biol. **217**: 2193–2200.

Layne, J. R., Jr and First, M. C. 1991. Resumption of physiological functions in the wood frog (*Rana sylvatica*) after freezing. - Am. J. Physiol. **261**: R134–7.

Layne, J. R. and Lee, R. E. 1995. Adaptations of frogs to survive freezing. - Clim. Res. **5**: 53–59.

Lee-Yaw, J. A. et al. 2008. Postglacial range expansion from northern refugia by the wood frog, *Rana sylvatica*. - Mol. Ecol. **17**: 867–884.

Liaw, A. and Wiener, M. 2002. Classification and regression by randomForest. R News. **2**: 18-22.

Merilä, J. and Hendry, A. P. 2014. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. - Evol. Appl. **7**: 1–14.

Miller-Rushing, A. J. et al. 2010. The effects of phenological mismatches on demography. - Philos. Trans. R. Soc. Lond. B Biol. Sci. **365**: 3177–3186.

Muths, E. et al. 2017. Heterogeneous responses of temperate-zone amphibian populations

to climate change complicates conservation planning. - Sci. Rep. **7**: 17102.

O'Connor, J. H. and Rittenhouse, T. A. G. 2016. Snow cover and late fall movement influence wood frog survival during an unusually cold winter. - Oecologia **181**: 635–644.

Ovaskainen, O. et al. 2013. Community-level phenological response to climate change. - Proc. Natl. Acad. Sci. U. S. A. **110**: 13434–13439.

Parmesan, C. 2006. Ecological and Evolutionary Responses to Recent Climate Change. - Annu. Rev. Ecol. Evol. Syst. **37**: 637–669.

Parmesan, C. 2007. Influences of species, latitudes and methodologies on estimates of phenological response to global warming. - Glob. Chang. Biol. **13**: 1860–1872.

Parmesan, C. and Yohe, G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. - Nature **421**: 37–42.

Petranka, J. W. et al. 2004. Identifying the minimal demographic unit for monitoring pond-breeding amphibians. - Ecol. Appl. **14**: 1065–1078.

Pimm, S. L. et al. 1995. The future of biodiversity. - Science **269**: 347–350.

R Core Team 2017. R: A Language and Environment for Statistical Computing.

Reading, C. J. 1998. The effect of winter temperatures on the timing of breeding activity in the common toad Bufo bufo. - Oecologia **117**: 469–475.

Regosin, J. V. et al. 2003. Terrestrial habitat use and winter densities of the Wood frog (*Rana sylvatica*). - J. Herpetol. **37**: 390–394.

Root, T. L. et al. 2003. Fingerprints of global warming on wild animals and plants. - Nature **421**: 57–60.

Scott, W. A. et al. 2008. Long-term United Kingdom trends in the breeding phenology of the common frog, *Rana temporaria*. - J. Herpetol. **42**: 89–96.

Scranton, K. and Amarasekare, P. 2017. Predicting phenological shifts in a changing climate. - Proc. Natl. Acad. Sci. U. S. A. **114**: 13212–13217.

Senner, N. R. et al. 2018. Spatial and temporal heterogeneity in climate change limits species' dispersal capabilities and adaptive potential. - Ecography **41**: 1428–1440.

Sheridan, J. A. et al. 2018. Shifts in frog size and phenology: Testing predictions of climate change on a widespread anuran using data from prior to rapid climate warming. - Ecol. Evol. **8**: 1316–1327.

Skelly, D. K. 2004. Microgeographic countergradient variation in the wood frog, *Rana sylvatica*. - Evolution **58**: 160–165.

Storey, K. B. 1990. Life in a frozen state: adaptive strategies for natural freeze tolerance in amphibians and reptiles. - Am. J. Physiol. **258**: R559–68.

Storey, K. B. and Storey, J. M. 1984. Biochemical adaption for freezing tolerance in the wood frog, *Rana sylvatica*. - J. Comp. Physiol. B **155**: 29–36.

Straile, D. et al. 2015. Trophic mismatch requires seasonal heterogeneity of warming. – Ecology **96**: 2794-805.

Stuart, S. N. et al. 2004. Status and trends of amphibian declines and extinctions worldwide. - Science **306**: 1783–1786.

Thackeray, S. J. et al. 2016. Phenological sensitivity to climate across taxa and trophic levels. - Nature **535**: 241–245.

Thornton, P. E. et al. 2016. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3. in press. DOI: 10.3334/ORNLDAAC/1328

Thuiller, W. et al. 2004. Biodiversity conservation: uncertainty in predictions of extinction risk. - Nature **430**: 34.

Todd, B. D. et al. 2011. Climate change correlates with rapid delays and advancements in reproductive timing in an amphibian community. - Proc. Biol. Sci. **278**: 2191–2197.

Tryjanowski, P. et al. 2003. Changes in the first spawning dates of common frogs and common toads in western Poland in 1978—2002. - Ann. Zool. Fennici **40**: 459-464.

Urban, M. C. 2007. Predator size and phenology shape prey survival in temporary ponds. - Oecologia **154**: 571–580.

Urban, M. C. 2015. Climate change. Accelerating extinction risk from climate change. - Science **348**: 571–573.

Urban, M. C. et al. 2016. Improving the forecast for biodiversity under climate change. - Science **353**: 1113.

U. S. Geological Survey. 2002. National Elevation Database. Accessible at: https://www.sciencebase.gov/catalog/item/505a61cde4b0c8380cd71b8d

Van Dyck, H. et al. 2015. The lost generation hypothesis: could climate change drive ectotherms into a developmental trap? - Oikos **124**: 54–61.

Visser, M. E. et al. 1998. Warmer springs lead to mistimed reproduction in great tits (Parus major). - Proc. R. Soc. B **265**: 1867–1870.

Waldman, B. 1982. Adaptive significance of communal oviposition in wood frogs (*Rana sylvatica).* - Behav. Ecol. Sociobiol. **10**: 169–174.

Walther, G.-R. et al. 2002. Ecological responses to recent climate change. - Nature **416**: 389–395.

While, G. M. and Uller, T. 2014. Quo vadis amphibia? Global warming and breeding phenology in frogs, toads and salamanders. - Ecography **37**: 921–929.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. - J. R. Stat. Soc. Series B Stat. Methodol. **73**: 3–36.

TABLES

Table 1. Seasonal trends in meteorological variables (temperature, precipitation, and snow water equivalent) 1980 - 2018 fit by penalized general additive models to Year and day-of-year (DOY) in the form: $Y = \beta_0 + \text{Year} + f(\text{DOY}) + \text{Year} * f(\text{DOY}) + \varepsilon$, where Y is the meteorological variable of interest.

		Parametric coefficients				Smoothed functions		
Dependent Variable	n	Parameter	Estimate	Std. Error	p-value	Parameter	Est. DF	p-value
Temp (°C)	13,870	Intercept	-48.75	6.24	<< 0.001	$f(\text{DOY})$	1.635	< 0.001
		Year	0.01	0.003	0.005	$f(\text{DOY}) * \text{Year}$	9.181	<< 0.001
Precip (mm)	13,870	Intercept	3.71	13.28	0.780	$f(\text{DOY})$	1.402	0.804
		Year	-0.000	0.003	0.999	$f(\text{DOY}) * \text{Year}$	1.529	0.940
SWE (kg m^{-2})	7,752	Intercept	65.47	78.29	0.403	$f(\text{DOY})$	7.993	<< 0.001
		Year	-0.008	0.020	0.690	$f(\text{DOY}) * \text{Year}$	1.500	0.018

FIGURES

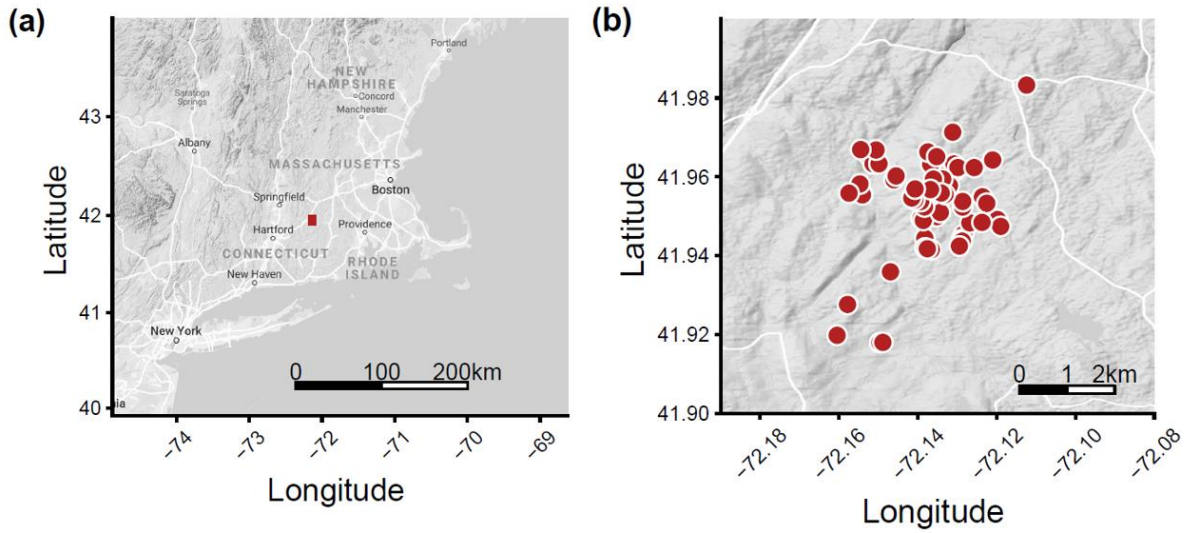

Figure 1. Location of Yale Myers Forest research site in northeastern Connecticut, USA (red square) (a). Locations of 64 wood frog breeding ponds at the site (red circles) (b).

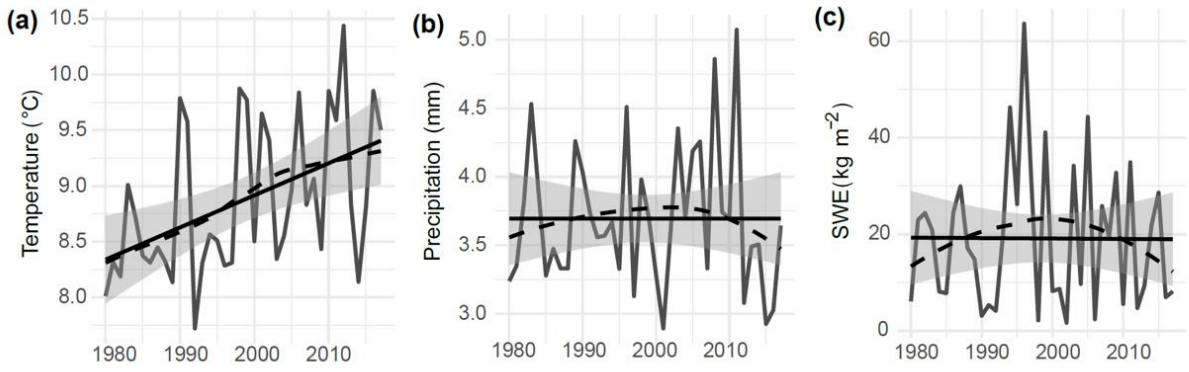

Figure 2. Trends in mean daily temperature (a) precipitation (b), and snow water equivalent (c) between 1980-2018 from DayMet meteorological database. Black solid lines and grey bands indicate linear regression line and 95% confidence intervals, respectively. Dashed lines indicates LOESS regression with 90% smoothing span for comparison. N = 39 for all models.

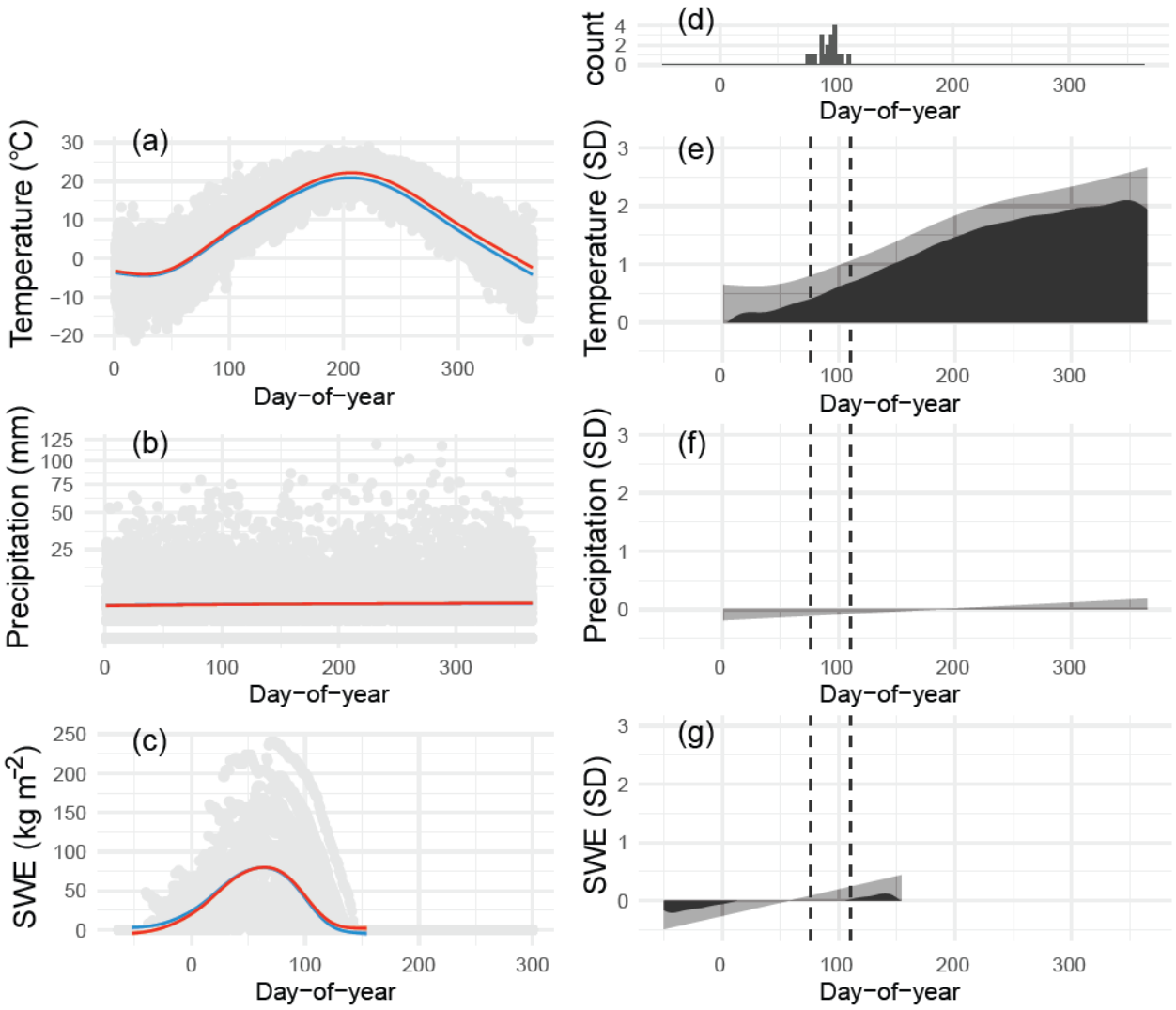

Figure 3. Seasonal trends in daily temperature (a), precipitation (square root scale) (b), and snow water equivalent (c) from 1980 (blue) to 2018 (red) as predicted by generalized additive model with interaction between Year and penalized spline smooth on day-of-year with 95% confidence intervals. Points represent daily values (N = 13,869 for all models). Annual mean oviposition dates (2000-2019) (d) in comparison to relative, seasonal change in temperature (e), precipitation (f), and snow water equivalent (SWE) (g) between 1980 and 2018. Seasonal change is the difference in daily values fit by generalized additive models for between 1980 and 2018. All differences are scaled to the standard deviation between annual averages for each variable in order to compare relative magnitude of change that coincides with the oviposition window (dotted lines). Dark bands indicate significant difference between 95% confidence intervals. Light bands indicate total difference. All meteorological observations from Daymet data between 1980 and 2018.

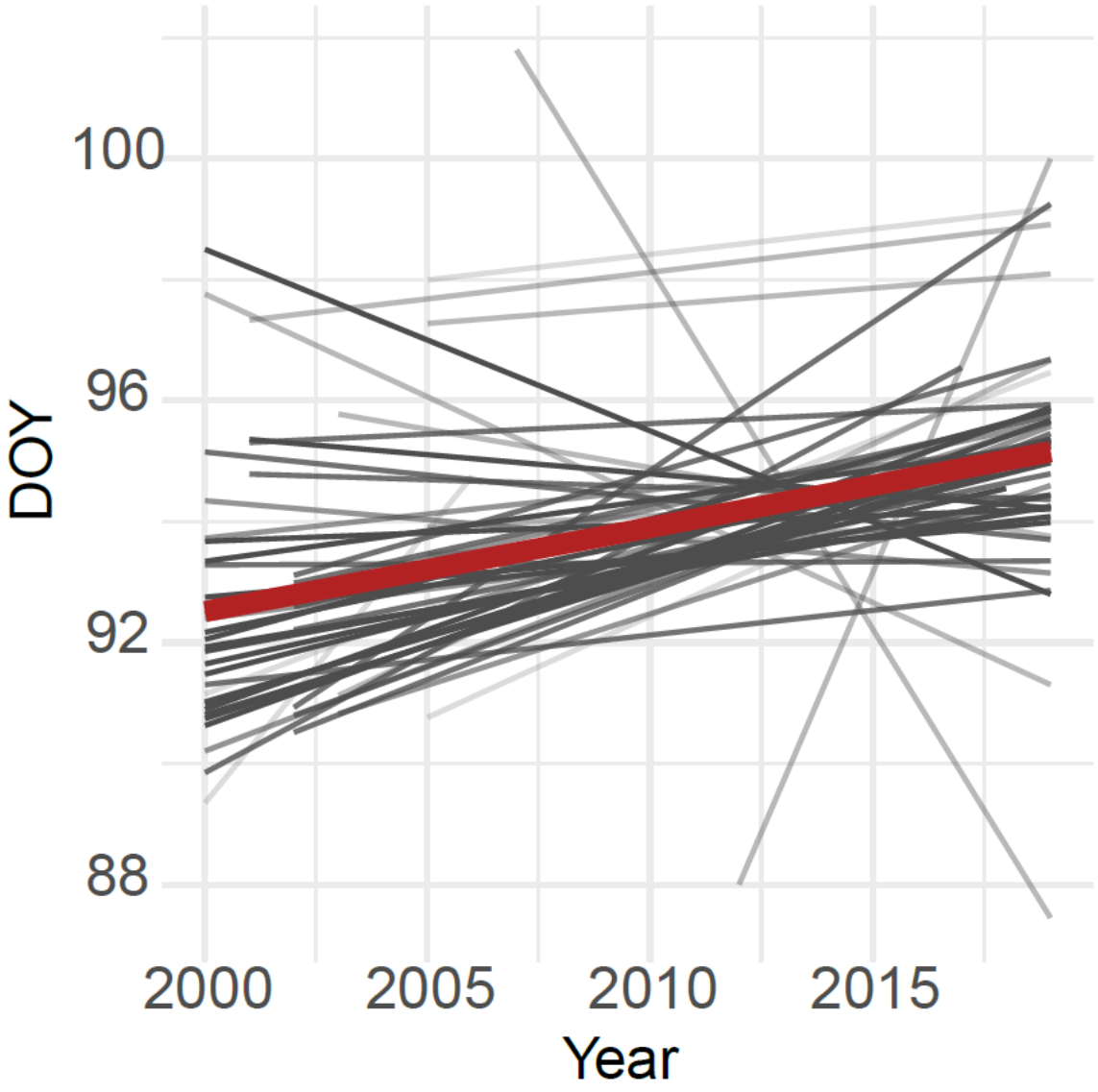

Figure 4. Trends in observed oviposition dates (DOY = day-of-year) across years between 2000 and 2019. Grey lines represent trends for individual ponds with transparency indicating number of observations (darker lines = more observations) (excludes ponds with fewer than 4 years of observation for simplicity). The red line represents the trend across all ponds, estimated as the fixed effect of a multilevel mixed effect model (n = 825, slope = 0.139, se = 0.05, 95% CI = 0.03, 2.44).

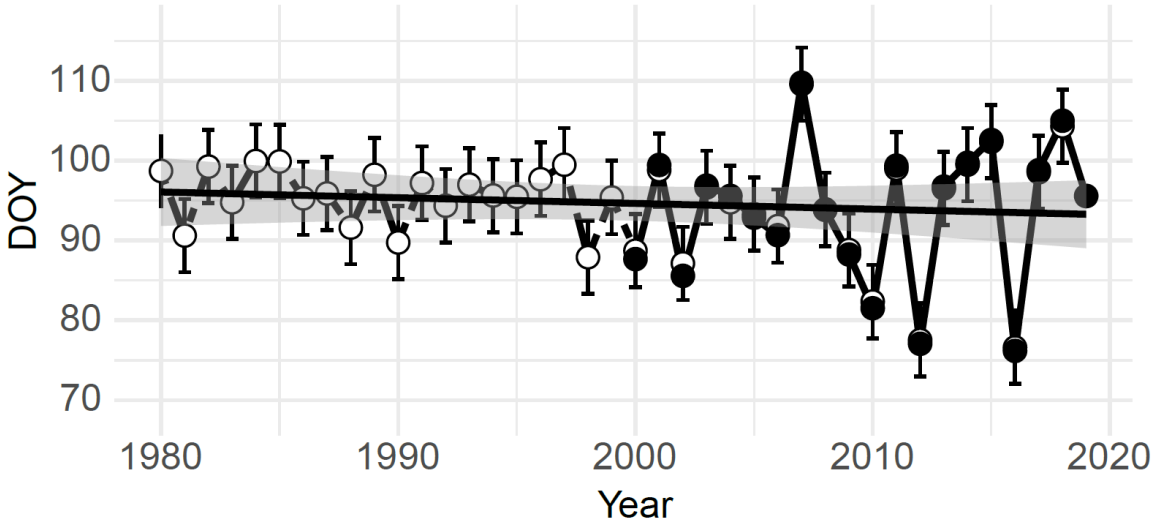

Figure 5. Predicted (open circles (1980-2018)) and observed (filled circles (2000-2019)) oviposition dates. Error bars indicate 95% prediction intervals (estimated as +/- 1.96 * cross-validated RMSE). Simple linear regression and 95% confidence intervals fit to predicted (1980-1999) and observed (2000-2019) dates (n = 40) are shown. Note that meteorological variables for 2019 were unavailable at the time of publishing, so no were predictions made for that year.

APPENDIX

Supplementary material

Ecography
ECOG-05297
Arietta, A. Z. A., Freidenburg, L. K., Urban, M. C., Rodrigues, S. B., Rubinstein, A. and Skelly, D. K. 2020. Phenological delay despite warming in wood frog *Rana sylvatica* reproductive timing: a 20-year study. – Ecography doi: 10.1111/ecog.05297

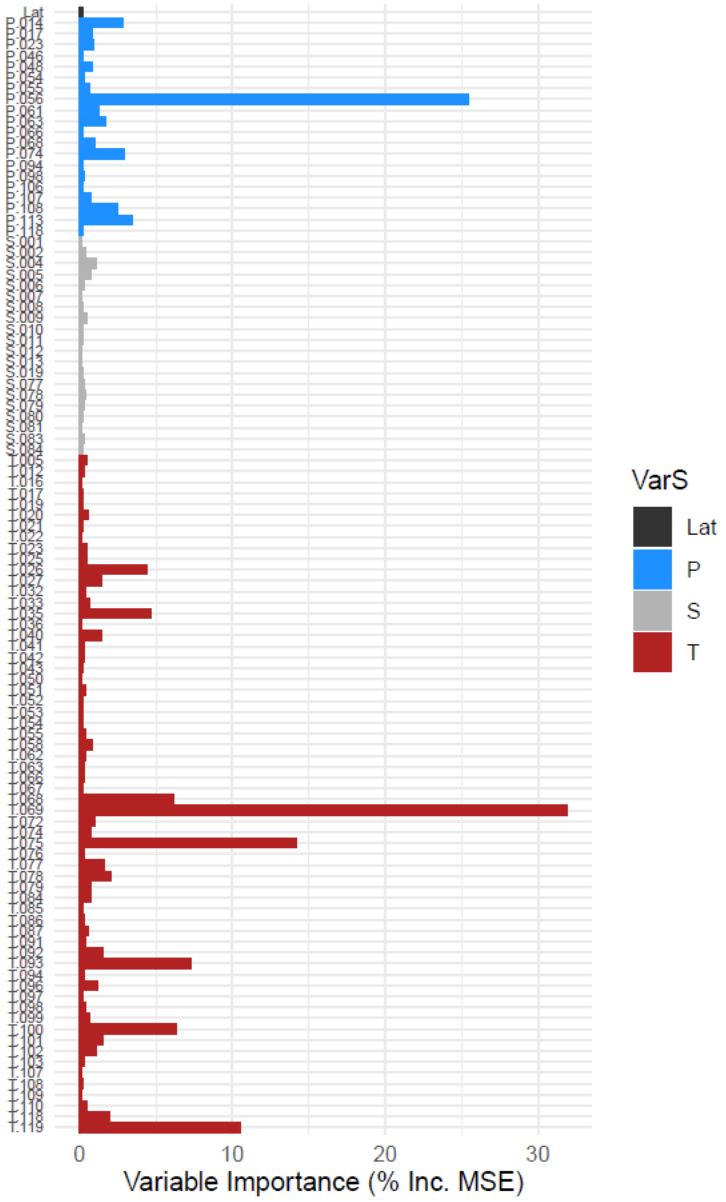

Figure A1. Random forest regression variable importance reported as the percent increase in predictive accuracy (% Inc MSE). The variable important metric estimates the proportional increase in predictive accuracy as the decrease in mean squared error from including a variable compared to a random permutation of that variable and is calculated from crossvalidation on the out-of-bag samples for each tree in the forest. Variable codes are in the format of variable letter code followed by day-of-year (example: T.093 = daily mean temperature at day-of-year 93; P.017 = daily mean precipitation at day-of- year 17). Variables in the models include average daily temperature (T), daily precipitation (P), daily snow water equivalent (S), latitude (Lat), aspect, elevation, leaf-on global site factor (GSF), leaf-off GSF, weighted GSF, GSF variance, and an a site identifier (Pond). A randomly generated variable (Noise1) was included to assess rank importance. Only variables with importance greater than random noise are shown.

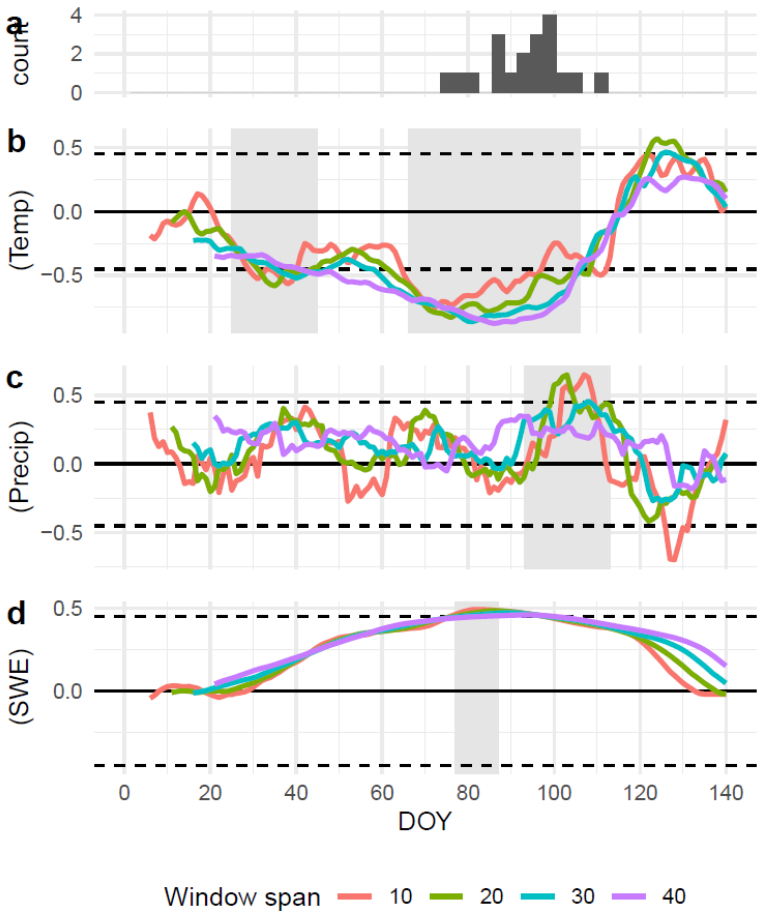

Figure A2. The correlation between 10-, 20-, 30-, and 40-day rolling averages of daily mean temperature (b), precipitation (c), and snow water equivalent (d) between 2000 and 2018 with oviposition timing (annual averages 2000-2019, 3-day bin width)(a). Dotted lines indicate 95% confidence interval (+/- 0.45) for Pearson's correlation for n = 20 pairs and 18 degrees of freedom. Light grey bands indicate non-overlapping windows of greatest correlation.

Table A1. Correlation between oviposition timing and seasonal windows for daily average temperature, precipitation, snow water equivalent (SWE).

Climate variable	Window	Pearson's correlation
Temperature	66, 106	-0.88
Temperature	66, 96	-0.86
Temperature	66, 86	-0.83
Temperature	67, 77	-0.76
Precipitation	102, 112	0.65
Precipitation	93, 113	0.65
Temperature	39, 79	-0.60
Temperature	25, 45	-0.58
Temperature	33, 43	-0.57
Precipitation	94, 114	0.53
Temperature	25, 55	-0.52
SWE	77, 87	0.49
SWE	76, 96	0.48
SWE	75, 105	0.47
SWE	74, 114	0.46
Precipitation	93, 123	0.46

CHAPTER 3

Climate change mediated effects on development and life history in an amphibian

Publication Status: A. Z. Andis Arietta; L. Kealoha Freidenburg; Susan B. Rodrigues; Adiana Rubinstein; Ella T. Schimdt; David K. Skelly. (*In prep.*). Climate change mediated effects on development and life history in an amphibian. Target Journal: *Ecology*. (Chapter formatted to journal standards)

ABSTRACT

Climate change is expected to have a range of impacts on organisms yet most prior research has focused on just a few responses (e.g. distribution, breeding phenology, body size). A deeper understanding of climate change effects depends on closer study of attributes such as development and life history from long term studies. During a 20-year study of the wood frog (*Rana sylvatica*) across 37 populations, we found that climate changed rapidly (0.49 C increase in water temperature), including a severe climate event (second worst drought in the northeastern United States in the last 120 years). Over the same time period, wood frog population densities declined by 40%. Wood frog physiological rates demonstrated profound changes in response. Larval period declined by 18%; in general, wood frogs in 2019 reached metamorphosis 12 days sooner than in 1999. Following the drought, embryo size doubled and subsequently attenuated to pre-drought levels within three years. Larval growth patterns also changed: peak size was elevated but metamorphic size was slightly smaller. Partitioning effects allows further interpretation of some responses. Drought appears to have the strongest effect on both rates of development and allocation (i.e. growth per developmental stage), with prolonged larval periods and larger larvae throughout ontogeny in extremely dry or extremely wet years. Although less pronounced of an effect, larvae tend to grow larger over a longer period at moderate pond temperatures.

Moderate levels of conspecific densities are associated with slower development. Nevertheless, there is residual variation in development rate not explained by these factors. Despite these remarkable effects on our study species, a superficial assessment of wood frogs might conclude that little had changed. The species has remained widely distributed in our study area throughout two decades. While some of these responses likely represent phenotypic plasticity, others may constitute evolved responses. Further study will be needed to understand how performance responses contribute to persistence of climate-affected species and may have future negative impacts (e.g., through lost genetic variation).

INTRODUCTION

In recent decades, the impacts of climate change on wild organisms have become widely apparent. The most frequently studied have been shifts in species' distributions (Thomas 2010, Chen et al. 2011) as well as phenology (Parmesan and Yohe 2003, Cohen et al. 2018). Changes in body size have also been reported (Sheridan and Bickford 2011, Ryding et al. 2021). While these responses represent important attributes of species, one of the primary reasons they have been the focus of study is their relatively widespread availability in long term data sources. Meanwhile, there is broad acknowledgement that to better understand and predict impacts from climate change, we will need to evaluate how changes in the environment wrought by changing climate interact with species' biology (Chevin et al. 2010, Huey et al. 2012). Many studies have examined the effect of, for example, elevated temperature or CO^2 on the physiology of a species (Harkey and Semlitsch 1988, Mitchell and Seymour 2000). However, there is a gap in our understanding due to the relative scarcity of long-term observational studies of traits and performance that connect climate to species biology more closely (see Crozier and Hutchings 2014, Seebacher et al. 2015, Lefevre 2016 for exceptions).

Much of the evidence for biological responses to climate change relies on associating changes in organismal phenotypes or life-history traits across temporal or spatial climatic gradients (e.g. Walther et al. 2002, Ryding et al. 2021). However, direct associations are difficult as changing climates impact many facets of ecological systems and can elicit multiple organismal responses. For instance, the "temperature-size rule" pattern for ectotherms is commonly explained by different temperature dependencies of growth and development (Forster and Hirst 2012, Zuo et al. 2012). However, in the wild, development and growth rates can be strongly impacted by hydroperiod and density-mediated competition for resources (Teplitsky et al. 2008, Rogers et al. 2011, DeLong et al. 2014). These effects are often interactive. For instance, change in growth rates may be directly affected by the environment, or the indirect effect of environmentally-induced changes in development (e.g. Mueller et al. 2012) or carry-over effects from prior life-stages (e.g. Benard 2015, Freidenburg 2017). In addition to plastic changes, there is evidence that variation in intrinsic rates of growth and development in pond-breeding amphibians has a genetic component and diverges at microgeographic scales (Skelly 2004, Orizaola and Laurila 2009, Richter-Boix et al. 2015, Urban et al. 2017).

Researchers most often consider only the change in average temperature or precipitation that is associated with trait variation, possibly overlooking profound impacts of climate variability (Thornton et al. 2014). However, extreme climatic events such as heatwaves, droughts, and flooding are expected to increase as a consequence and extension of increased climate variability (Seneviratne et al. 2021). It has been suggested that these events may foster larger consequences and more pronounced biotic responses than even long-term shifts in averages (Bailey and van de Pol 2016).

Wood frog growth and development rates vary considerably across the species' broad geographic range (reviewed in Dodd 2013). This variation is partly a plastic response

to temperature and conditions such as competition, predation, and hydroperiod in vernal pools where larvae develop and partly a result of local adaptation to these same factors (Berven 1982a, 1982b, 1987, Berven and Chadra 1988, Skelly 2004). Growth rates, size, and ontogenetic timing of the aquatic larval stage are strongly related to fitness in wood frogs, as in most ectotherms (Werner 1986, Kingsolver and Huey 2008). Many taxa, especially pool-breeding amphibians, have evolved mechanisms to regulate ontogenetic timing and optimal size through compensatory growth or development (Murillo-Rincón et al. 2017, Burraco et al. 2017b, 2017a). However, compensation often leads to fitness consequences in the larval stage (Gahm et al. 2021) or later life stages (Metcalfe and Monaghan 2001, Hector et al. 2012).

In this study we take advantage of a two-decade long study of a species, the wood frog (*Rana sylvatica*) for which we have both climate information as well as life history records for many populations living in the northeastern United States. Over the past two decades, this region has experienced a long-term trend of exceptional warming (Karmalkar and Horton 2021). Amidst this background of warming are occurrences of extreme weather. For example, in 2016 the northeastern United States experienced one of the most severe droughts in the past 150 years, exceeded only by the drought of the 1960s which caused severe municipal water shortages and agricultural loss for multiple years (Janes and Brumbach, 1965).

Droughts pose unique ecological threats to amphibians due to their sensitivity to water balance as terrestrial and semi-terrestrial adults (Feder and Burggen 1992). In addition, droughts may harbor negative consequences for aquatic larvae, especially in species that breed in temporary—or vernal—wetlands. Vernal pool obligates like the wood frog are restricted in their life history timing by the hydroperiod of natal pools, racing to metamorphose into terrestrial stages before ponds dry. Because the hydroperiod of these

pools are largely driven by short-term weather patterns (Brooks 2004), droughts may cause pools to dry up before all larvae metamorphose. In the worst instances, droughts can cause catastrophic population failure and local extinctions (Donald et al. 2011, Scheele et al. 2012, Cayuela et al. 2016).

We focus here on embryos and larvae collected from a wild metapopulation of 37 wood frog breeding populations in the Northeastern USA from the last two decades between 1999 and 2005 and recent surveys in 2017 and 2019. We pair these data with long-term climate, water temperature, and population data to hindcast pond microhabitat, conspecific density, drought conditions, and life history timing. We use an algorithmic modelling approach to partition the relative effects of putative ecological drivers of shifts in physiological rates of growth allocation (i.e. size per developmental stage) and development (i.e. stage per unit time).

MATERIALS AND METHODS

Field site and study system

We focus on 37 wood frog breeding populations at Yale Myers Forest, a 3,213-ha research forest in northeastern Connecticut, USA (N 41.9198 W −72.1604). Wood frogs are the most cosmopolitan amphibian in North America, ranging from Georgia to the Arctic Coast—the most northern range limit of any anuran (Lee-Yaw et al. 2008, Dodd 2013). The species is well-known for the ability to overwinter by freezing before reanimating in the spring (Storey and Storey 1984, Layne et al. 1998, Costanzo et al. 2013).

Wood frogs are vernal pool obligates that breed in explosive aggregations and deposit eggs in synchrony (Dodd 2013). Within their range, they are typically the earliest spring-breeding amphibian. The embryos can develop remarkably quickly at cold temperatures in procession to free-swimming larvae (Frisbie et al. 2000). As with other

poikilothermic amphibian larvae, development rates are highly plastic and positively associated with water temperature and negatively associated with hydroperiod (Amburgey et al. 2012, Lambert et al. 2018).

Wood frogs undergo indirect development from embryos, to aquatic larvae, to terrestrial juveniles, and finally, mature adults. Within the larval period, there are characteristic ontogenetic phases, although there is some ambiguity in demarcation and nomenclature (McDiarmid and Altig 1999). Here, we follow the conventions of (Feder and Burggen 1992, McDiarmid and Altig 1999) in defining four phases within the larval stage. These phases, identified by Gosner (1960) developmental stages are: hatchling (approx. 21-24), premetamorphic (approx. 25-35), prometamorphic (approx. 35-42), and metamorphic (42-46). Herein, we define the "metamorphic transition" as the point of development between pre- and prometamorphoic stages when larvae reach maximum size. This transition initiates shifts in physiological rates that continue through the metamorphosis as all body systems orient to terrestrial life.

Ecological drivers

We assessed broad trends among potential ecological drivers of development rates by estimating annual site-wide values of pond water temperature, drought, conspecific density, and oviposition timing and fitting mixed effect models against time. We further incorporate pond-wise annual estimates in these variables to model rates of development and growth per developmental stage (i.e. allocation). All statistical analysis was performed in R (v3.6.2, R Core Team 2019).

Pond temperature

Air temperature at our site has increased by 0.6 C (0.3 C/decade) over the 20 year duration of our study (Arietta et al. 2020). We expect the water temperature of wood frog breeding pools to have similarly increased. To evaluate this prediction, we analyzed long-term water temperatures recorded in the ponds. Starting in 2001, submersible temperature loggers (HOBO 8K Pendant; Onset Computer Corporation) were deployed in a subset of ponds each year. Loggers were suspended 10cm below the surface at the point of maximum depth and recorded water temperature every 1 or 0.5 hours. Loggers were deployed during annual egg mass counts; so, records begin prior to the larval period. Loggers were generally removed well after larvae had metamorphosed and left the ponds, coinciding in some years with ponds drying at the end of the season. Thus, the logger data span the larval period for a population, but the data occupy different temporal windows across ponds and years.

To maximize the greatest overlap in logger data among ponds and years, we restricted our analysis of annual temperature data to the day of year (DOY) of the earliest observation of free-swimming tadpoles (i.e. stage 25; DOY = 108) through the first record of forelimb emergence (i.e. stage 42; DOY 167) in our dataset. This filter constrains the time window of comparison to a 59 day period for which we know larvae were present in all ponds across all years. Water temperature was recorded for 242 out of a potential 798 pond/years (and 59 of the 147 pond cohorts in our study). For the remaining pond/years, we imputed data with a random forest model implemented in **randomForest** (v4.6.14, Liaw and Wiener 2002) (Appendix S1). Our model was grown from 500 regression trees. We used 10-fold cross validation (holding back 10% of data points as a testing dataset) to evaluate the predictive accuracy of our random forest model for daily observations. We used an annual temperature value from averaging temperatures over the larval window rather than daily observations in subsequent models, which dampens the error variance

associated with daily estimates. We tested the accuracy of annual average estimates with leave-one-out cross validation for 100 randomly sampled ponds with complete logger data.

Pond water temperatures are subject to many influences, including convection from the air, direct solar radiation, mixing with precipitation and groundwater, etc. (Brooks 2004). In the case of northeastern vernal pools, forest canopy over the ponds can exert a strong influence by shading radiation, blocking wind, and insulating against daily air temperature variability (Brooks 2004). To account for these factors, we included both climate and canopy variables in our random forest model along with pond specific latitude, elevation, and aspect. We incorporated daily estimates of mean, minimum, and maximum air temperature, precipitation, radiation, snowpack, and atmospheric pressure from the Daymet database (v.3 (Thornton et al. 2016); tile: 11754, N 41.9198 W -72.1604). We included 1- and 2-day lags for all climate variables. Canopy variables included mean and variance of leaf-on and leaf-off global site factor (GSF; the ratio of below-canopy radiation to above-canopy radiation (Anderson 1964)) and a weighted average GSF for the larval period (Halverson et al. 2003) for each pond. Briefly, five hemispherical photographs were taken along the shore at each cardinal point and at the center of each pond during leaf-off and leaf-on seasons. We used Gap Light Analyzer (Frazer 1999) to estimate GSF integrated over the larval period. Canopy estimates were unique to each pond but the same across years, while climatological variables were unique to each day and year but the same across all ponds.

We combined our observed and imputed missing data from our hindcast model to estimate an annual mean water temperature value for each pond in each year for the period between day of year 108 and 167. We tested for a trend of increasing water temperatures over the years by fitting a linear mixed model, allowing a random slope and intercept for each pond. We implemented our mixed effect model with **lme4** (v1.1.21, Bates et al. 2015)

and assessed the significance of our model by constructing 95% nonparametric confidence intervals from 1000 bootstrap iterations, as with all following mixed effect model analyses.

Drought

We considered the direct impact of drought at the pond level by estimating changes in pond water depth over time with a generalized additive mixed model accounting for uneven sampling. Pond depth was recorded in each breeding pond during most annual population census surveys (i.e., egg mass counts) at a fixed point in each pond. However, pond depths were not recorded consistently throughout our dataset. Thus, we developed a composite drought index for further modelling.

The Palmer Drought Severity Index was developed in response to the 1965 drought and has been retroactively estimated back to 1895. Because wood frogs reach sexual maturity in two to three years, we calculated the 3-year cumulative drought index for our field site in Windham and Tolland Counties, Connecticut, USA to determine the relative magnitude of the 2016 drought.

We incorporated annual estimates of PDSI averaged over the larval period (DOY 90-167) with mean precipitation and snow water equivalent (DOY 0 - 167) from the Daymet database into a single drought variable using principal component analysis (Appendix S2). These variables were highly correlated (all $r > 0.5$) and the first principal component explained 67% of the variance. We employed this composite drought variable in subsequent models for ease of interpretation.

Conspecific density

Wood frog populations at our site fluctuate over time and between neighboring pond populations (Rowland et al. *in press*). These dynamics result in variable conspecific

densities experienced by each larval cohort which can also influence growth and development (Berven 1990).

We estimated conspecific density for each larval cohort as the population size divided by pond surface area. Population sizes were estimated from annual egg mass counts. Each year, egg mass surveys take place approximately 7 days after first oviposition. This timing ensures that the laying event is complete and all egg masses are still distinguishable. We estimated pond size as the spring bank-full water surface area as an ellipse from linear measurements along the long axis of the pond and its perpendicular. As with pond temperatures, we test for long-term change in conspecific density over time by fitting a linear mixed model with random slopes and intercepts for each pond and constructed bootstrap confidence intervals (Appendix S3).

Oviposition timing

Larval rates of growth and development may be constrained by preceding life history phenology. For instance, a trend of delayed oviposition timing observed at our site (Arietta et al. 2020) constrained hatching phenology and limited the window for larval development. To test for this, we include oviposition timing, estimated from egg mass surveys (as in Arietta et al. 2020) for each pond in our dataset.

Embryo size

We compared embryo size between historical (2000 and 2001) and contemporary (2017, 2018, and 2019) embryos collected from the field shortly after oviposition. Up to twelve embryos from up to eight clutches in each pond were individually photographed under a dissecting microscope. We recorded developmental stage and size from the photos in ImageJ (Schneider et al. 2012). We estimated embryo size by measuring the vertical and

horizontal diameter of each embryo and then used the average diameter value to calculate a spherical volume. Measurements from the year 2000 were an exception. In 2000, diameter records were estimated via ocular micrometer and clutch of origin was not recorded. We tested for differences in historical and contemporary embryo size with a mixed effect model accounting for the nesting of observations within clutches and clutches within ponds. We treated each observation from 2000 as a unique clutch. Because there is uneven sampling across years, we used a weighted, nonparametric bootstrapping procedure by refitting the model to 1000 datasets resampled from our data to ensure equal representation between timepoints and years (Appendix S4).

Larval size and developmental stage

We estimated rates of development and growth using data from wild-caught wood frog larvae and metamorphs collected during field surveys between 1999 and 2019 (1999-2005, 2017, 2019). We included only datasets in which larvae were collected with unbiased survey methods including standardized dipnet surveys and pipe sampling (Skelly 1996, Gunzburger 2007, Werner et al. 2007). Field sampling ceased when larvae reached the metamorphic stage and began to leave the natal ponds between stage 42 and 46. Sampling of post-metamorphic individuals required targeted capture of a haphazard sample of terrestrial individuals along the pond margins. In all cases, larvae and metamorphs were euthanized in the field and returned to the lab where the samples were assigned a developmental stage and measured for snout-to-vent length following standard procedures. We checked for observer bias in larval measurements by reexamining a subset of larvae and metamorphs (n = 637) from historical samples (Appendix S5).

The data were recorded for each individual with the exception of one dataset representing 0.5% of total observations in which stage and size measurements were

recorded as mean values for a random sample of larvae collected in pipe samples (count of larvae included in pipe: average n = 8.7, median n = 4). In these cases, we treated the pipe averages (n = 74) as individual records. We excluded all embryo and hatchling observations (Gosner stage <25), which left 16,271 observations representing 147 pond cohorts for the analysis.

We tested for differences between historical (1999 to 2005) and contemporary (2017 and 2019) size by grouping observations by developmental stage. To reduce sampling disparity, we aggregated observations so that each pond cohort was represented by a single value and then applied Welch's t-tests.

Rates of development and growth

Divergence in physiological rates of growth and development have been implicated as causes of body size change in response to recent climate change. The trade-off between growth and development may explain differences between historical and contemporary wood frog populations at our site. To test this and estimate overall changes in size and larval period, we fit linear mixed effect models with a continuous piecewise function to estimate physiological rates of development (stage by day) and allocation (size by stage) (Appendix S6). For each rate, we fit a model allowing rates to vary linearly with year for *i* observations from *j* pond (Equation 1). This model enabled us to estimate total change from 1999 to 2019. We estimated the larval period from the development rate model as the predicted duration between stage 25 and 42 for each year.

Equation 1:

$$Y_{ij} = \beta_0 + b_{0_j} + (\beta_1 + b_{1_j}) \cdot X_{Pre_{ij}} + (\beta_2 + b_{2_j}) \cdot X_{Post_{ij}} + \beta_3 \cdot Year_{ij} + (\beta_4 \cdot X_{Pre_{ij}} + \beta_5 \cdot X_{post_{ij}}) \cdot Year_{ij} + \varepsilon_{ij}$$

Where *Y* is size measured as snout-to-vent length or developmental stage, and *X* is developmental stage or day-of-the-year with the piecewise function allowing slopes to differ before and after the metamorphic transition *T* with the transformation:

$$X_{Pre} = f(X) = \begin{Bmatrix} X - T, & if\ X < T \\ 0, & if\ X \geq T \end{Bmatrix}, X_{Post} = f(X) = \begin{Bmatrix} 0, & if\ X \geq T \\ X - T, & if\ X > T \end{Bmatrix}$$

and β represents fixed effect parameters while *b* represents parameters estimated as random effects with:

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} \sim N(0, \Omega_b), \Omega_b = \begin{bmatrix} \sigma^2_{b_0} & & \\ \rho\sigma_{b_0}\sigma_{b_1} & \sigma^2_{b_1} & \\ \rho\sigma_{b_0}\sigma_{b_2} & \rho\sigma_{b_1}\sigma_{b_2} & \sigma^2_{b_2} \end{bmatrix}$$

$$\varepsilon \sim N(0, \sigma^2_{\varepsilon_0})$$

Our models include random effects accounting for variation in the slopes of physiological rates and intercepts for each pond population. Rates of growth and development do not scale continuously as amphibian larvae approach the metamorphic climax from aquatic larvae to the terrestrial body plan (Werner 1986). During development, growth rates change and often inflect at the metamorphic transition. In order to account for different rates before and after the transition, we fit our multiple regressions as piecewise models that allow for different slopes at an established breakpoint. Rather than assume the transition stage *a priori*, we estimated it from the data. To determine the midpoint stage at which the metamorphic transition occurs, we initially fit our allocation rate model using stage 30 through 46 as potential transition points. We determined the best fit model with the lowest AIC and used that transition stage in following models (Appendix S5). In order to model rates of growth and development over the season, we must relate the average stage of metamorphic transition, determined previously, to the day at which the transition occurred in each pond cohort. To do so, we fit a linear mixed model predicting the day of

year that larvae reach the transition stage using only observations prior to the metamorphic transition, essentially fitting only the pre-transition slope. We used this model to predict the transition day for each pond cohort (Appendix S5). We assessed significance of model parameters by refitting our models to 1000 bootstrap iterations of the data and estimating 95% confidence intervals.

Drivers of physiological rate change

We compare the relative impacts of putative ecological drivers with partial predictions from random forest models learned from a dataset including observed larval size and stage, and ecological predictors: pond temperatures, composite drought score, conspecific density, and oviposition date. Models were grown from 1000 trees and assessed with out-of-bag cross validation. We then use the models to estimate either the stage from DOY (development) or size from stage over a range of an ecological predictor variable (min, max, and quartiles) while holding all other ecological predictors at the median value. This allows us to ask the question of how a given predictor impacts development or growth independent of the other predictors? We refit our random forest models to 1000 bootstraps of our dataset to estimate 90% confidence intervals.

RESULTS

Ecological drivers

Pond water temperatures have increased by an estimated 0.49°C (0.25°C/decade) over 20 years (95% CI: 0.27, 0.71; Fig. 1B and Table S1.1). There was no support for the inclusion of random slopes in the model (**Δ**AIC = 1.5, likelihood ratio test p = 0.29). Predicted annual estimates were within 0.31°C of the observed value on average (MSR = 0.60) based on hold-one-out cross validation.

Over the same 20-year period, wood frog population densities have decreased by 40%, from 0.10 to 0.06 clutch/m^2, at a rate of -2.8% per year (95% CI: -1.6, -4.0; Fig. 1C and Table S3.1). While the site-wide trend is significantly negative, there is high variability in the direction of change across ponds (Fig. S3.1) and there was no support for including random slope estimates in the model (**Δ**AIC = 2.0, likelihood ratio test p = 0.65).

The cumulative three years of drought conditions leading up to 2017 were the most severe seen at our field site since the major drought of 1965 (Figure 1A). Other than 1965 and 1966, 2016 is the only year since the turn of the century in which average annual PDSI values remained below -3. Across all ponds spring water levels dropped significantly in 2016 and remained well below the pre-drought average (Figure 1D).

Embryonic and larval size

Contemporary embryos tend to be larger than historical samples by approximately 31% (1.66ml; 95% C.I. = 1.51, 1.82) (Fig 2A and Table S4.1). However, most of this difference is driven by the 2017 cohort immediately following the drought for which average embryo size was double (8.01ml; 95% C.I. = 7.37 - 8.74) that of embryos prior to the drought (Fig. S4.1). In subsequent years, embryo sizes diminished across all observed populations back to pre-drought size. There was no significant difference in embryo size between historical and the 2019 cohorts (C.I. of difference = -0.2 - 0.09).

Remeasuring historical specimens showed no difference in size among larval stages ($p = 0.51$) but an average difference of -0.5 mm ($p < 0.001$) among metamorphic stages compared to original estimates (Table S5.1). We adjusted contemporary measurements of metamorphs by +0.5 mm to account for this disparity.

Contemporary larvae also tend to be larger, but the effect varies over the course of development (Fig. 2). Overall, larvae after the drought tend to be larger up to the point of

metamorphic transition (Fig. 2). Thereafter, differences diminish upon metamorphosis to the point that contemporary metamorphs are slightly smaller than during prior decades, noting however, that sample sizes are smaller for these stages (Table S5.2).

Larval physiological rates

The metamorphic transition occurs close to Gosner stage 35 as the piecewise model predicting size at a given stage with this breakpoint had the best fit (Table S6.1). In general, tadpoles allocate more growth per stage by 1.01 mm stage^{-1} up to the metamorphic transition, at which point average size is relatively constant (Table 1).

In the past 20 years, allocation rates have accelerated by 0.01 mm stage^{-1} year^{-1} leading up to the metamorphic transition, but have decelerated by -0.02 mm stage^{-1} year^{-1} thereafter (Table 1). Thus, we estimate that contemporary tadpoles in 2019 enter the larval period (stage 25) 1.7 mm (28%) longer and are 2.7 mm (17%) longer at maximum size (stage 35) (Table 1). However, because tadpoles are larger upon entering the larval period in recent years, they still tend to metamorph at about the same size (Fig. 3A). At metamorphic climax (stage 42), contemporary larvae are only 0.3 mm (0.2%) smaller. By the end of metamorphosis, contemporary larvae are 1.0 mm (6.2%) smaller (Fig. 3A).

Developmental rates over the season are also affected by ontogenetic shifts in physiology, although the effect of the metamorphic transition is not as pronounced (Fig. 3B). The baseline development rate is 0.11 stage day^{-1}, increasing to 0.16 stage day^{-1} after stage 35 (Table 1). The rate of development has accelerated over time, especially after the metamorphic transition, by 0.01 stages day^{-1} year^{-1}. We estimate that the larval period has been reduced from 65 days in 1999 to 53 days in 2019, a reduction of 12 days (18%), overall.

Our random forest models fit well and were able to predict larval allocation and development from ecological variables, accounting for 93% and 92% of the variation in out-of-bag samples, respectively (Fig. S7.1). When considering the partial effects of ecological drivers, drought appears to have the strongest effect on both rates of development and allocation (Fig. S7.2C and Fig. S7.3C). Drought years tend to prolong the larval period and result in larger larvae throughout ontogeny. Development is fastest in moderate drought years while larval size is smallest at median drought levels, especially later in ontogeny.

Pond water temperature, conspecific densities, and oviposition dates shifted the absolute seasonal timing of developmental stages, but did not have strong, consistent effects on larval period (Fig S7.3A). Similarly, conspecific densities did not have noticeable effects on allocation (Fig S7.2B). Allocation rates tended to increase at moderate levels of pond temperature, especially later in development (Fig S7.2A). Conversely, allocation rates tend to be lowest at moderate oviposition timing (Fig S7.2D). The strongest effect of oviposition timing can be seen in the size larvae hatch and enter the larval period. Hatchlings tend to be largest in years of most advanced breeding phenology, and to a lesser extent, are also larger in years of extreme phenological delay.

After accounting for the putative ecological drivers, there is considerable variance in both allocation and development between years that is unaccounted for (Fig S7.2D and Fig S7.3D).

DISCUSSION

The two-decade period captured by our study was one of substantial change in climate and included an extreme climate event. In a specific place and time, our study system represents what climate scientists project as commonplace in the future (van der Wiel and Bintanja 2021): continuing, accelerating change in climate variables like

temperature (increased 0.49 C) punctuated by an extreme event (second worst drought in more than a century). Our study species survived the challenges posed by recent climate but it has been affected. Wood frog populations at our site have undergone considerable changes in growth and development rates over 20 years punctuated by a major drought event. In fundamental ways, the wood frogs we study are different now than they were when we started our work. The populations, which have declined in density by 40% since 1999, now complete larval development 18% faster than they did early in the study. Immediately following an extreme drought, embryo sizes were twice what they had been earlier; within 3 years following the conclusion of the drought, embryo sizes were indistinguishable from those measured in pre-drought populations (Fig. S4.1). The pattern of growth is also altered. Larvae reach larger peak sizes but metamorphose slightly smaller.

Some of these responses are understandable in terms of the forcing variables we have measured. In other dimensions, it is not clear why the species has responded as it has. As an example, the large increase in embryo size following an extreme drought is a predictable response to selection on body size in adult females (Berven 1988). The phenotypic of shifts in physiological rates is only weakly associated with the effects of pond temperature and conspecific density, compared to other changes over time. We find that contemporary wood frog larvae tend to be larger upon entering the larval period, consistent with larger embryonic volume, and remain larger throughout most of larval ontogeny compared to larvae collected in years prior to the drought. Our models estimate a 17% increase in larval size at the metamorphic transition since 1999. However, size differences decrease approaching metamorphosis and we estimate no difference in size at metamorphic climax. Our data suggest that contemporary wood frogs are 10% smaller by the end of metamorphosis. We estimate that accelerated developmental rates have reduced the larval period by 18% (12 days) since 1999. The net effect is that contemporary tadpoles

tend to enter the larval period at a larger size later in the season, but exit the larval period at smaller sizes than pre-drought cohorts.

In addition to the drought, these shifts in physiological rates are concomitant with an estimated 0.49° C increase in water temperatures of breeding ponds. Warming water temperature trends concur with estimates of air temperature increases for the region (Hartfield et al. 2018). Over the same period, site-wide densities have declined by 40% as the metapopulation size declined and oviposition has been delayed by 3 days (Arietta et al. 2020).

The contributions of drought, temperature, and oviposition timing on changes in developmental timing and body size changes are evident. Extremely wet or dry years and moderate pond temperatures are associated with protracted larval development. Oviposition timing is negatively associated with larval period length. Drought years and years with earlier oviposition are expected to yield larger larvae, especially earlier in the premetamorphic phase, whereas larger prometamorphic larvae are expected at moderate pond temperatures. Conspecific density did not have strong effects on either growth or development. The lack of density-dependent effects is surprising given expectations from controlled lab and field experiments that show high conspecific density and low per capita food resources result in protracted development and smaller size at metamorphosis (Crump 1981, Murray 1990, Riha and Berven 1991, Audo et al. 1995, Loman 2002, Dahl et al. 2012).

Finally, persistent effects from all of these factors can drive selection for adaptive changes in intrinsic growth and development rates, even at extremely small spatial scales (Freidenburg and Skelly 2004, Orizaola and Laurila 2009, Richter-Boix et al. 2015, Urban et al. 2017). Heritable differences in physiological rates with respect to temperature gradients (i.e. elevation, latitude, and canopy) have been consistently demonstrated in wood frog

populations (Berven 1982b, Riha and Berven 1991, Ligon and Skelly 2009, Arietta and Skelly 2021). Although there is little direct evidence of genetic adaptation to recent climate change (see Gienapp et al. 2008, Urban et al. 2014), common garden and reciprocal transplant studies find that evolutionary divergence in these rates are negatively associated with the environmentally-induced effects of temperatures (Merilä et al. 2000, Skelly 2004, Orizaola and Laurila 2009).

When we interrogate shifts in rates over time while controlling for putative ecological drivers, we find that temporal trends can be either additive or counteracting. Historically, spatial (elevation or latitude) or temporal (year) changes in traits that associate with temperature were used for inferring the impact of warming, a practice that can mask important, counteracting effects (Yom-Tov and Geffen 2011). Our results highlight the necessity of explicitly testing for both temporal trends and spatio-temporal changes in temperature simultaneously.

Long-term datasets are uniquely useful in assessing biotic responses to climate change, but the nature of our data imposes limitations on our inferences. For example, we measure size and stage as populations rather than tracking individuals. Thus, size or stage dependent mortality during the larval period could bias our growth and development estimates, given that over 90% of larvae perish prior to metamorphosis (Berven 1990). There is evidence of size dependent mortality influencing body size in birds (Gardner et al. 2014).

Our sampling focuses on the larval period because this is a period of rapid growth and development when larvae are of a known cohort age and can be conveniently resampled. Sampling aquatic larvae also avoids the confounding interaction of temperature, precipitation, and water balance issues with which physiological studies of terrestrial amphibians must contend. However, as larvae metamorph and disperse from ponds, we

have imperfect data at later stages and no data on juveniles or adults. Therefore, we do not know how changes in larval development impact adult size or ultimate fitness. Although sample sizes for late-stage metamorphs in our data are small, we see some evidence that contemporary larvae may be smaller at the end of metamorphosis as an effect of commensurate size reduction from larger tadpoles. Prior studies of wood frogs indicate that juvenile and adult survival correlate with size at metamorphosis (Berven 1990). As such, we might expect that changes in physiological rates in the larval phase results in reduced fitness for contemporary adults. However, while larger tadpoles must allocate more absolute energy to development at the cost of growth during metamorphosis, they do so with more efficiency per unit mass which may yield more energy reserves at the completion of metamorphosis (Pandian and Marian 1985, Beck and Congdon 2003, Orlofske and Hopkins 2009). Individuals with greater lipid reserves upon entering the juvenile phase tend to have greater lifetime fitness (Scott et al. 2007), suggesting that size at metamorphosis may trade-off with an energetic advantage in the terrestrial phase. Similarly, there may be trade-offs between body size, development rate, and performance (Beck and Congdon 2000). For example, faster development is associated with reduced larval performance (Gahm et al. 2021), but it is unclear how larval performance relates to later life stages (Watkins 1997, Gomez-Mestre et al. 2010).

That contemporary embryos are also larger suggests that maternal effect may play a large role in differences in development and growth rates, as has been shown in other high-latitude anuran populations (Berven and Chadra 1988, Laugen et al. 2002). Larger eggs confer larger hatchlings and faster development rates (Bradford 1990).

Lastly, we do not know the sex of the larvae in our sample. Wood frogs exhibit dimorphic body size and sex ratios are influenced by temperature dependent sex

determination (Lambert et al. 2018). Inasmuch as climate may affect sex rations in our populations, the sex skew may correlate with body size changes we see.

The 0.5 °C change in water temperatures is only a small fraction of the 6 °C increase in climate predicted by the end of the century, yet we find marked shifts in developmental timing and body size. Continued climate warming will have multifarious effects on organisms. While our study considers a small geographical area, our results reflect those of body size change in reproductive age wood frogs across the species range over the past century (Sheridan et al. 2018). We show that predictions based on temperature-dependencies of physiological rates can only partly account for shifting body size and life history timing in natural populations. The effects of climate change may be confounded with important ecological effects, like density mediated competition, phenological shifts and a multitude of other indirect ecological factors that may also be moderated by climate. Most importantly, we find large temporal effects which enervate the assumption that the effects of climate change can be accomplished by simply scaling the effects of physiological processes alone.

Acknowledgements – We would like to thank Gregory Watkins-Colwell and the Yale Peabody Museum of Natural history for access to historical samples. We thank Eric Lee, Nicole Freidenfelds, Jonathan Richardson, Steven Brady, Kamau Walker, Danielle de Haerne, and Baasim Zafar for past field collections or assistance with contemporary field collections and analysis.

Permits – Animal handling and collections were approved by Yale University IACUC (2004-10361, 2007-10361, 2010-10361, 2013-10361, 2016-10361, 2019-10361, 2006-11024, 2009-11024, 2009-11040) and Connecticut DEEP (205001, 208001, 9902006, 0103019,

010619, 0119019, 121019). Field work was conducted with permission from Yale Myers Research Committee (SKE01, AND17).

REFERENCES

Amburgey, S., W. C. Funk, M. Murphy, and E. Muths. 2012. Effects of Hydroperiod Duration on Survival, Developmental Rate, and Size at Metamorphosis in Boreal Chorus Frog Tadpoles (Pseudacris maculata). Herpetologica 68:456–467.

Anderson, M. C. 1964. Studies of the woodland light climate I. The photographic computation of light condition. Journal of Ecology 52:27–41.

Arietta, A. Z. A., L. K. Freidenburg, M. C. Urban, S. B. Rodrigues, A. Rubinstein, and D. K. Skelly. 2020. Phenological delay despite warming in wood frog *Rana sylvatica* reproductive timing: a 20-year study. Ecography 52:27.

Arietta, A. Z. A., and D. K. Skelly. 2021. Rapid microgeographic evolution in response to climate change. Evolution; international journal of organic evolution 75:2930–2943.

Audo, M. C., T. M. Mann, T. L. Polk, C. M. Loudenslager, W. J. Diehl, and R. Altig. 1995. Food deprivation during different periods of tadpole (Hyla chrysoscelis) ontogeny affects metamorphic performance differently. Oecologia 103:518–522.

Bailey, L. D., and M. van de Pol. 2016. Tackling extremes: challenges for ecological and evolutionary research on extreme climatic events. The Journal of animal ecology 85:85–96.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, Articles 67:1–48.

Beck, C. W., and J. D. Congdon. 2000. Effects of age and size at metamorphosis on performance and metabolic rates of Southern Toad, Bufo terrestris, metamorphs. Functional ecology 14:32–38.

Beck, C. W., and J. D. Congdon. 2003. Energetics of metamorphic climax in the southern toad (Bufo terrestris). Oecologia 137:344–351.

Benard, M. F. 2015. Warmer winters reduce frog fecundity and shift breeding phenology, which consequently alters larval development and metamorphic timing. Global change biology 21:1058–1065.

Berven, K. A. 1982a. The genetic basis of altitudinal variation in the wood frog rana sylvatica. I. An experimental analysis of life history traits. Evolution; international journal of organic evolution 36:962–983.

Berven, K. A. 1982b. The genetic basis of altitudinal variation in the wood frog Rana sylvatica II. An experimental analysis of larval development. Oecologia 52:360–369.

Berven, K. A. 1987. The heritable basis of variation in larval developmental patterns within populations of the wood frog (*Rana sylvatica*). Evolution; international journal of organic evolution 41:1088–1097.

Berven, K. A. 1988. Factors Affecting Variation in Reproductive Traits within a Population of Wood Frogs (Rana sylvatica). Copeia 1988:605–615.

Berven, K. A. 1990. Factors affecting population fluctuations in larval and adult stages of the wood frog (*Rana sylvatica*). Ecology 71:1599–1608.

Berven, K. A., and B. G. Chadra. 1988. The relationship among egg size, density and food level on larval development in the wood frog (*Rana sylvatica*). Oecologia 75:67–72.

Bradford, D. F. 1990. Incubation Time and Rate of Embryonic Development in Amphibians: The Influence of Ovum Size, Temperature, and Reproductive Mode. Physiological zoology 63:1157–1180.

Brooks, R. T. 2004. Weather-related effects on woodland vernal pool hydrology and hydroperiod. Wetlands 24:104–114.

Burraco, P., C. Díaz-Paniagua, and I. Gomez-Mestre. 2017a. Different effects of accelerated development and enhanced growth on oxidative stress and telomere shortening in amphibian larvae. Scientific reports 7:7494.

Burraco, P., A. E. Valdés, F. Johansson, and I. Gomez-Mestre. 2017b. Physiological mechanisms of adaptive developmental plasticity in Rana temporaria island populations. BMC evolutionary biology 17:164.

Cayuela, H., D. Arsovski, E. Bonnaire, R. Duguet, P. Joly, and A. Besnard. 2016. The impact of severe drought on survival, fecundity, and population persistence in an endangered amphibian. Ecosphere 7.

Chen, I.-C., J. K. Hill, R. Ohlemüller, D. B. Roy, and C. D. Thomas. 2011. Rapid range shifts of species associated with high levels of climate warming. Science 333:1024–1026.

Chevin, L.-M., R. Lande, and G. M. Mace. 2010. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. PLoS biology 8:e1000357.

Cohen, J. M., M. J. Lajeunesse, and J. R. Rohr. 2018. A global synthesis of animal phenological responses to climate change. Nature climate change 8:224–228.

Costanzo, J. P., M. C. F. do Amaral, A. J. Rosendale, and R. E. Lee Jr. 2013. Hibernation physiology, freezing adaptation and extreme freeze tolerance in a northern population of the wood frog. The Journal of experimental biology 216:3461–3473.

Crozier, L. G., and J. A. Hutchings. 2014. Plastic and evolutionary responses to climate change in fish. Evolutionary applications 7:68–87.

Crump, M. L. 1981. Energy accumulation and amphibian metamorphosis. Oecologia 49:167–169.

Dahl, E., G. Orizaola, A. G. Nicieza, and A. Laurila. 2012. Time constraints and flexibility of growth strategies: geographic variation in catch-up growth responses in amphibian larvae. The Journal of animal ecology 81:1233–1243.

DeLong, J. P., T. C. Hanley, and D. A. Vasseur. 2014. Competition and the density dependence of metabolic rates. The Journal of animal ecology 83:51–58.

Dodd, C. K., Jr. 2013. Frogs of the United States and Canada, 2-vol. set. The Johns Hopkins University Press.

Donald, D. B., W. T. Aitken, C. Paquette, and S. S. Wulff. 2011. Winter snowfall determines the occupancy of northern prairie wetlands by tadpoles of the Wood Frog (Lithobates sylvaticus). Canadian journal of zoology 89:1063–1073.

Feder, M. E., and W. W. Burggen, editors. 1992. Environmental Physiology of the Amphibians. University of Chicago Press, Chicago.

Forster, J., and A. G. Hirst. 2012. The temperature-size rule emerges from ontogenetic differences between growth and development rates. Functional ecology 26:483–492.

Freidenburg, L. K. 2017. Environmental drivers of carry-over effects in a pond-breeding amphibian, the Wood Frog (Rana sylvatica). Canadian journal of zoology 95:255–262.

Freidenburg, L. K., and D. K. Skelly. 2004. Microgeographical variation in thermal preference by an amphibian. Ecology letters 7:369–373.

Frisbie, M. P., J. P. Costanzo, Lee, Jr., and E. Richard. 2000. Physiological and ecological aspects of low-temperature tolerance in embryos of the wood frog, Rana sylvatica. Canadian journal of zoology 78:1032–1041.

Gahm, K., A. Z. A. Arietta, and D. K. Skelly. 2021. Temperature-mediated trade-off between development and performance in larval wood frogs (Rana sylvatica). Journal of Experimental.

Gardner, J. L., T. Amano, B. G. Mackey, W. J. Sutherland, M. Clayton, and A. Peters. 2014. Dynamic size responses to climate change: prevailing effects of rising temperature drive long-term body size increases in a semi-arid passerine. Global change biology 20:2062–2075.

Gienapp, P., C. Teplitsky, J. S. Alho, J. A. Mills, and J. Merilä. 2008. Climate change and evolution: disentangling environmental and genetic responses. Molecular ecology 17:167–178.

Gomez-Mestre, I., V. L. Saccoccio, T. Iijima, E. M. Collins, G. G. Rosenthal, and K. M. Warkentin. 2010. The shape of things to come: linking developmental plasticity to post-metamorphic morphology in anurans. Journal of evolutionary biology 23:1364–1373.

Gosner, K. L. 1960. A simplified table for staging Anuran embryos and larvae with notes on

identification. Herpetologica 16:183–190.

Gunzburger, M. S. 2007. Evaluation of seven aquatic sampling methods for amphibians and other aquatic fauna. Applied Herpetology 4:47–63.

Halverson, M. A., D. K. Skelly, J. M. Kiesecker, and L. K. Freidenburg. 2003. Forest mediated light regime linked to amphibian distribution and performance. Oecologia 134:360–364.

Harkey, G. A., and R. D. Semlitsch. 1988. Effects of Temperature on Growth, Development, and Color Polymorphism in the Ornate Chorus Frog Pseudacris ornata. Copeia 1988:1001–1007.

Hartfield, G., J. Blunden, and D. S. Arndt. 2018. State of the Climate in 2017. Bulletin of the American Meteorological Society 99:Si–S310.

Hector, K. L., P. J. Bishop, and S. Nakagawa. 2012. Consequences of compensatory growth in an amphibian: Compensatory growth in tadpoles. Journal of zoology 286:93–101.

Huey, R. B., M. R. Kearney, A. Krockenberger, J. A. M. Holtum, M. Jess, and S. E. Williams. 2012. Predicting organismal vulnerability to climate warming: roles of behaviour, physiology and adaptation. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 367:1665–1679.

Janes, B. E., and J. J. Brumbach. 1965. The 1964 Agricultural Drought in Connecticut. Storrs Agricultural Experiment Station Bulletin 390:21.

Karmalkar, A. V., and R. M. Horton. 2021. Drivers of exceptional coastal warming in the northeastern United States. Nature climate change 11:854–860.

Kingsolver, J. G., and R. B. Huey. 2008. Size, temperature, and fitness: three rules. Evolutionary ecology research 10:251–268.

Lambert, M. R., M. S. Smylie, A. J. Roman, L. K. Freidenburg, and D. K. Skelly. 2018. Sexual and somatic development of wood frog tadpoles along a thermal gradient. Journal of experimental zoology. Part A, Ecological and integrative physiology 329:72–79.

Laugen, A. T., A. Laurila, and J. Merila. 2002. Maternal and genetic contributions to geographical variation in Rana temporaria larval life-history traits. Biological journal of the Linnean Society. Linnean Society of London 76:61–70.

Layne, J. R., Jr, J. P. Costanzo, and R. E. Lee Jr. 1998. Freeze duration influences postfreeze survival in the frog Rana sylvatica. The Journal of experimental zoology 280:197–201.

Lee-Yaw, J. A., J. T. Irwin, and D. M. Green. 2008. Postglacial range expansion from northern refugia by the wood frog, Rana sylvatica. Molecular ecology 17:867–884.

Lefevre, S. 2016. Are global warming and ocean acidification conspiring against marine ectotherms? A meta-analysis of the respiratory effects of elevated temperature, high CO2 and their interaction. Conservation physiology 4:cow009.

Liaw, A., and M. Wiener. 2002. Classification and Regression by randomForest.

Ligon, N. F., and D. K. Skelly. 2009. Cryptic divergence: countergradient variation in the wood frog. Evolutionary ecology research 11:1099–1109.

Loman, J. 2002. Temperature, genetic and hydroperiod effects on metamorphosis of brown frogs Rana arvalis and R. temporaria in the field. Journal of zoology 258:115–129.

McDiarmid, R. W., and R. Altig. 1999. Tadpoles: the biology of anuran larvae. Chicago ; University of Chicago Press, 1999.

Merilä, J., A. Laurila, M. Pahkala, K. Räsänen, and A. Timenes Laugen. 2000. Adaptive phenotypic plasticity in timing of metamorphosis in the common frog Rana temporaria. Écoscience 7:18–24.

Metcalfe, N. B., and P. Monaghan. 2001. Compensation for a bad start: grow now, pay later? Trends in ecology & evolution 16:254–260.

Mitchell, N. J., and R. S. Seymour. 2000. Effects of temperature on energy cost and timing of embryonic and larval development of the terrestrially breeding moss frog, Bryobatrachus nimbus. Physiological and biochemical zoology: PBZ 73:829–840.

Mueller, C. A., S. Augustine, S. A. L. M. Kooijman, M. R. Kearney, and R. S. Seymour. 2012. The trade-off between maturation and growth during accelerated development in frogs. Comparative biochemistry and physiology. Part A, Molecular & integrative physiology 163:95–102.

Murillo-Rincón, A. P., N. A. Kolter, A. Laurila, and G. Orizaola. 2017. Intraspecific priority effects modify compensatory responses to changes in hatching phenology in an amphibian. The Journal of animal ecology 86:128–135.

Murray, D. L. 1990. The effects of food and density on growth and metamorphosis in larval wood frogs (Rana sylvatica) from central Labrador. Canadian journal of zoology 68:1221–1226.

Orizaola, G., and A. Laurila. 2009. Microgeographic variation in temperature-induced plasticity in an isolated amphibian metapopulation. Evolutionary ecology 23:979.

Orlofske, S. A., and W. A. Hopkins. 2009. Energetics of metamorphic climax in the pickerel frog (Lithobates palustris). Comparative biochemistry and physiology. Part A, Molecular & integrative physiology 154:191–196.

Pandian, T. J., and M. P. Marian. 1985. Time and Energy Costs of Metamorphosis in the Indian Bullfrog Rana tigrina. Copeia 1985:653–662.

Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421:37–42.

R Core Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Richter-Boix, A., M. Katzenberger, H. Duarte, M. Quintela, M. Tejedo, and A. Laurila. 2015. Local divergence of thermal reaction norms among amphibian populations is affected by pond temperature variation. Evolution; international journal of organic evolution 69:2210–2226.

Riha, V. F., and K. A. Berven. 1991. An Analysis of Latitudinal Variation in the Larval Development of the Wood Frog (Rana sylvatica). Copeia 1991:209–221.

Rogers, L. A., L. C. Stige, E. M. Olsen, H. Knutsen, K.-S. Chan, and N. C. Stenseth. 2011. Climate and population density drive changes in cod body size throughout a century on the Norwegian coast. Proceedings of the National Academy of Sciences of the United States of America 108:1961–1966.

Ryding, S., M. Klaassen, G. J. Tattersall, J. L. Gardner, and M. R. E. Symonds. 2021. Shape-shifting: changing animal morphologies as a response to climatic warming. Trends in ecology & evolution 36:1036–1048.

Scheele, B. C., D. A. Driscoll, J. Fischer, and D. A. Hunter. 2012. Decline of an endangered amphibian during an extreme climatic event. Ecosphere 3:art101.

Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. Nature methods 9:671–675.

Scott, D. E., E. D. Casey, M. F. Donovan, and T. K. Lynch. 2007. Amphibian lipid levels at metamorphosis correlate to post-metamorphic terrestrial survival. Oecologia 153:521–532.

Seebacher, F., C. R. White, and C. E. Franklin. 2015. Physiological plasticity increases resilience of ectothermic animals to climate change. Nature climate change 5:61–66.

Seneviratne, S. I., X. Zhang, M. Adnan, W. Badi, C. Dereczynski, A. Di Luca, S. Ghosh, I. Iskandar, J. Kossin, S. Lewis, F. Otto, I. Pinto, M. Satoh, S. M. Vicente-Serrano, M. Wehner, and B. Zhou. 2021. Weather and Climate Extreme Events in a Changing Climate. *in* V. MassonDelmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, Caud, N, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou, editors. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Sheridan, J. A., and D. Bickford. 2011. Shrinking body size as an ecological response to climate change. Nature climate change 1:401–406.

Sheridan, J. A., N. M. Caruso, J. J. Apodaca, and L. J. Rissler. 2018. Shifts in frog size and phenology: Testing predictions of climate change on a widespread anuran using data from prior to rapid climate warming. Ecology and evolution 8:1316–1327.

Skelly, D. 1996. Pond Drying, Predators, and the Distribution of Pseudacris Tadpoles 1996:599–605.

Skelly, D. K. 2004. Microgeographic countergradient variation in the wood frog, *Rana*

sylvatica. Evolution; international journal of organic evolution 58:160–165.

Storey, K. B., and J. M. Storey. 1984. Biochemical adaption for freezing tolerance in the wood frog,Rana sylvatica. Journal of comparative physiology. B, Biochemical, systemic, and environmental physiology 155:29–36.

Teplitsky, C., J. A. Mills, J. S. Alho, J. W. Yarrall, and J. Merilä. 2008. Bergmann's rule and climate change revisited: Disentangling environmental and genetic responses in a wild bird population. Proceedings of the National Academy of Sciences 105:13492–13496.

The northeast water supply crisis of the 1960's. 1968. . US Department of the Interior Geologic Survey.

Thomas, C. D. 2010. Climate, climate change and range boundaries. Diversity & distributions 16:488–495.

Thornton, P. E., M. M. Thornton, B. W. Mayer, Y. Wei, R. Devarakonda, R. S. Vose, and R. B. Cook. 2016, July 15. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3. ORNL Distributed Active Archive Center.

Thornton, P. K., P. J. Ericksen, M. Herrero, and A. J. Challinor. 2014. Climate variability and vulnerability to climate change: a review. Global change biology 20:3313–3328.

Urban, M. C., J. L. Richardson, and N. A. Freidenfelds. 2014. Plasticity and genetic adaptation mediate amphibian and reptile responses to climate change. Evolutionary applications 7:88–103.

Urban, M. C., J. L. Richardson, N. A. Freidenfelds, D. L. Drake, J. F. Fischer, and P. P. Saunders. 2017. Microgeographic Adaptation of Wood Frog Tadpoles to an Apex Predator. Copeia:451–461.

Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent climate change. Nature 416:389–395.

Watkins, T. B. 1997. The effect of metamorphosis on the repeatability of maximal locomotor performance in the Pacific tree frog Hyla regilla. The Journal of experimental biology 200:2663–2668.

Werner, E. E. 1986. Amphibian Metamorphosis: Growth Rate, Predation Risk, and the Optimal Size at Transformation. The American naturalist 128:319–341.

Werner, E. E., D. K. Skelly, R. A. Relyea, and K. L. Yurewicz. 2007. Amphibian species richness across environmental gradients. Oikos 116:1697–1712.

van der Wiel, K., and R. Bintanja. 2021. Contribution of climatic changes in mean and variability to monthly temperature and precipitation extremes. Communications Earth & Environment 2:1–11.

Yom-Tov, Y., and E. Geffen. 2011. Recent spatial and temporal changes in body size of

terrestrial vertebrates: probable causes and pitfalls. Biological reviews of the Cambridge Philosophical Society 86:531–541.

Zuo, W., M. E. Moses, G. B. West, C. Hou, and J. H. Brown. 2012. A general model for effects of temperature on ectotherm ontogenetic growth and development. Proceedings. Biological sciences / The Royal Society 279:1840–1846.

TABLES

Table 1. Piecewise linear mixed model results allocation (stage-specific size) and development (stage over time) rates before (Pre) and after (Post) the metamorphic transition at stage 35 of larval wood frogs from 1999 to 2019. N_{obs} = 14844, N_{ponds} = 37. Confidence intervals were estimated from 1000 nonparametric bootstraps. Significant coefficients for which confidence intervals do not include zero are shown in bold.

Parameters	Allocation rate model (Y = SVL, X = Stage)	Development rate model (Y = Stage, X = DOY)
Fixed effects	Estimate [95% C.I.]	Estimate [95% C.I.]
β_0	**16.10 [15.87, 16.33]**	**33.97 [33.74, 34.20]**
X_{Pre}	**1.01 [0.97, 1.05]**	**0.21 [0.20, 0.22]**
X_{Post}	0.02 [-0.03, 0.07]	**0.34 [0.32, 0.37]**
Year	**0.14 [0.12, 0.15]**	**0.06 [0.04, 0.08]**
X_{Pre}* Year	**0.01 [0.00, 0.01]**	**0.00 [0.00, 0.00]**
X_{Post}* Year	**-0.02 [-0.02, -0.01]**	**0.01 [0.01, 0.01]**
Random effects	Variance	Variance
b_{pond}	2.83 [2.56, 3.60]	1.62 [1.43, 2.07]
X_{Pre}	0.07 [0.07, 0.09]	0.00 [0.00, 0.00]
X_{Post}	0.05 [0.04, 0.08]	0.06 [0.05, 0.07]
Residual	2.20 [2.09, 2.27]	3.41 [3.27, 3.51]
$R^2_{Conditional}$	0.87	0.89
$R^2_{Marginal}$	0.80	0.80

FIGURES

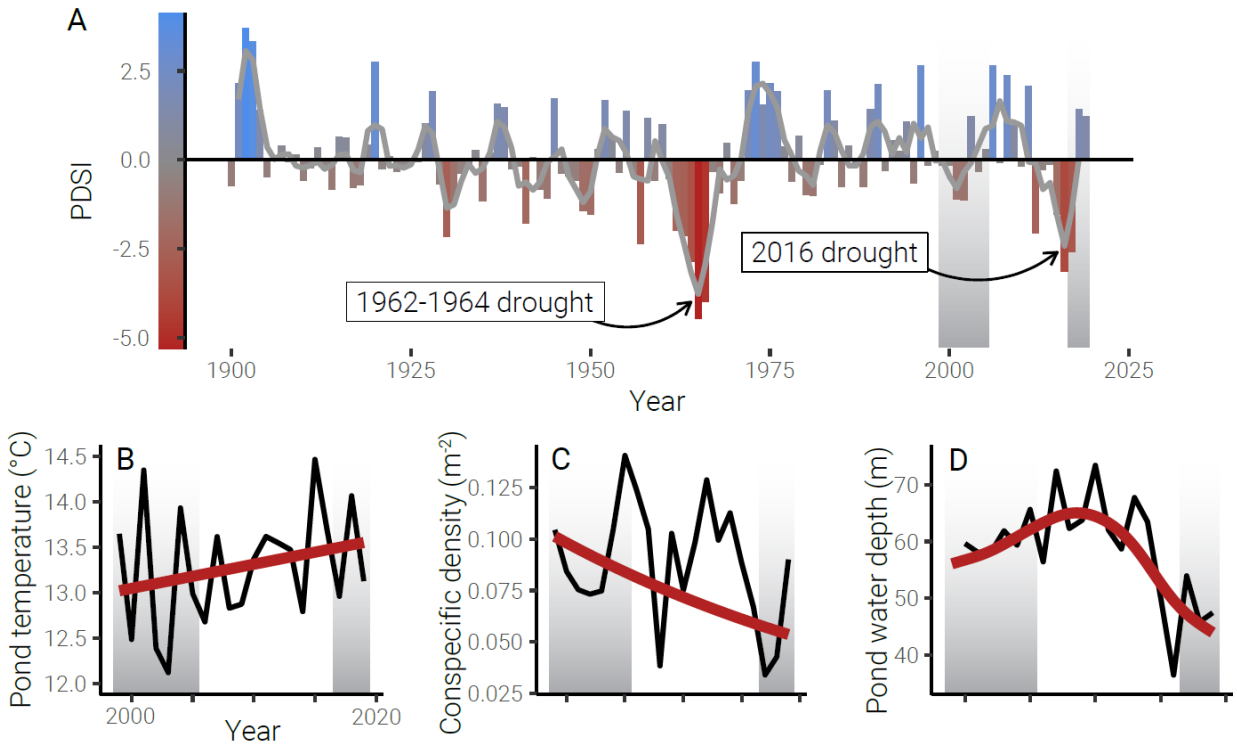

Figure 1. Palmer Drought Severity Index (PDSI) values for Yale Myers Forest, Connecticut from 1900 to 2019 (A) and estimates of pond-wise mean water temperature from day-of-the-year 108 to 167 (B), conspecific density (C), and spring water depth (D). In all panels, the time spans for which historical (1999-2005) and contemporary (2017 and 2019) samples were collected are indicated by grey boxes. The 3-year rolling average PDSI value is shown as a grey trend line in A. The major drought of 2016 and the last major historical drought in 1962-1964 are shown. The black lines in B, C, and D connect site-wide mean values, while the red lines represent the marginal estimates across all ponds from mixed effect models.

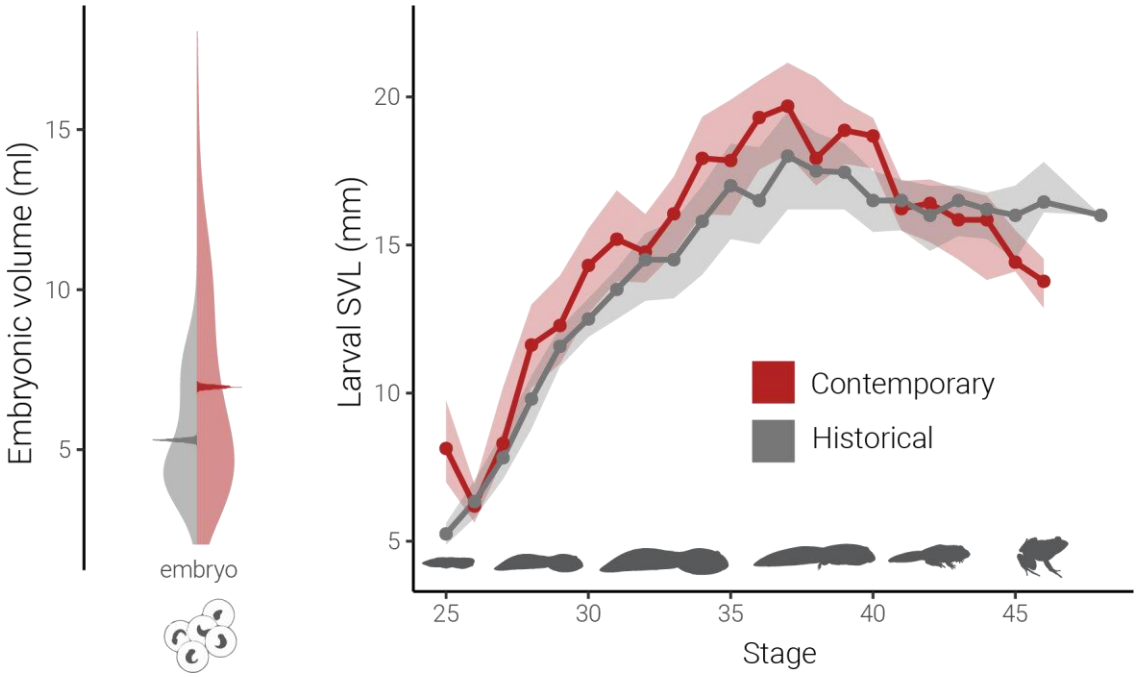

Figure 2. Differences in size between historical (grey) and contemporary (red) wood frog embryos (A) and larvae (B). Embryo size was estimated as volume (ml) and larval size was estimated as snout-to-vent length (mm). Density curves of embryo size (A) represent the raw data while the overlaid histograms represent the estimated means after accounting for nested data structure from 1000 bootstraps. In B, observations from the same pond and day were averaged prior to estimating site-wide average values per stage. Points represent mean values and bands represent the upper and lower quartile of observed cohort means. Asterisks in B indicate significant differences between larval size between timepoints based on Welch's t-test with alpha = 0.01.

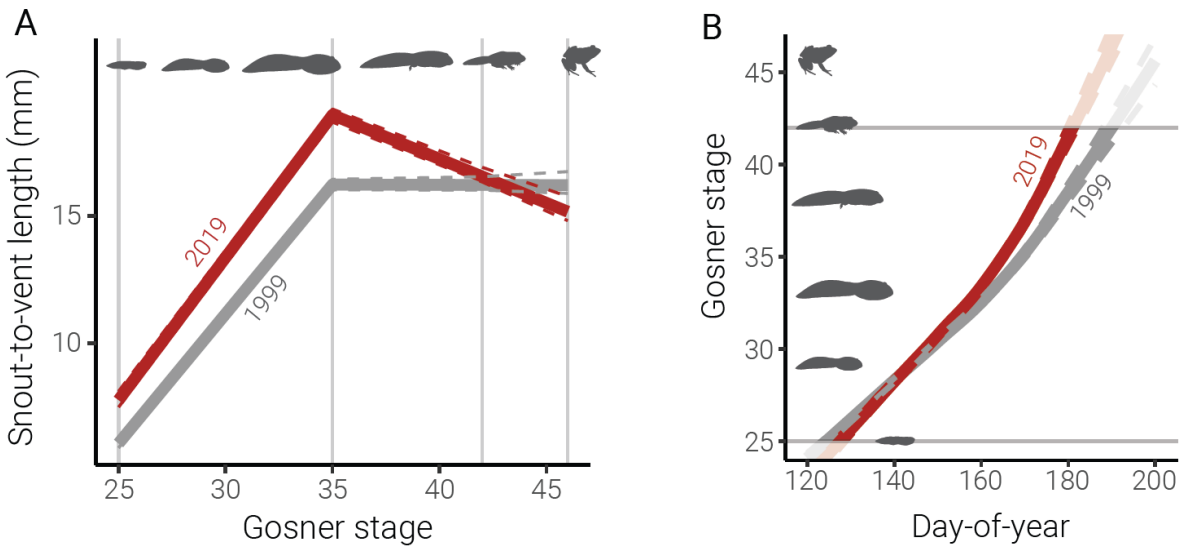

Figure 3. Estimated allocation rates (snout-to-vent length in mm per Gosner stage)(A) and development rates (Gosner stages per day)(B) from piecewise mixed effect models. Solid lines represent marginal trends for years 1999 (grey) or 2019 (red) while dotted lines represent 95% nonparametric bootstrap confidence intervals.

APPENDIX

Supplementary materials for
"Climate change mediated effects on development and life history in an amphibian"

Appendix S1. Pond temperature

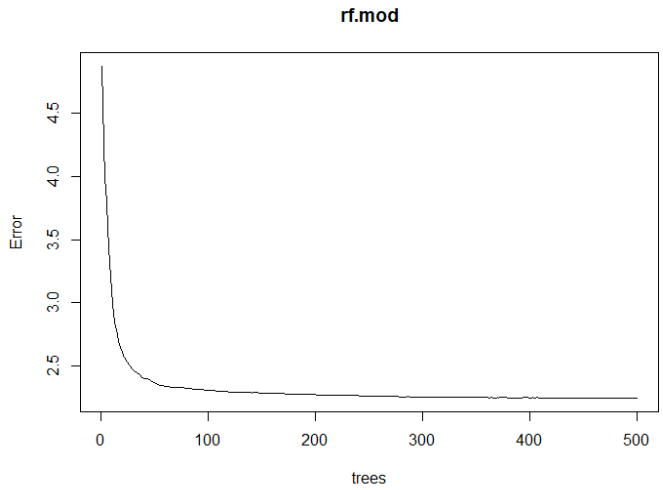

Figure S1.1. Error plot for random forest regression grown from 500 individual regression trees predicting daily-average water temperatures. Including more trees improves accuracy at the cost of computation time. Optimal forest sizes minimize the error without including too many trees, as indicated by the error rate levelling off after approximately 400 trees. The model explained 82% of the variance in out-of-bag observations across trees (RMSE = 2.31).

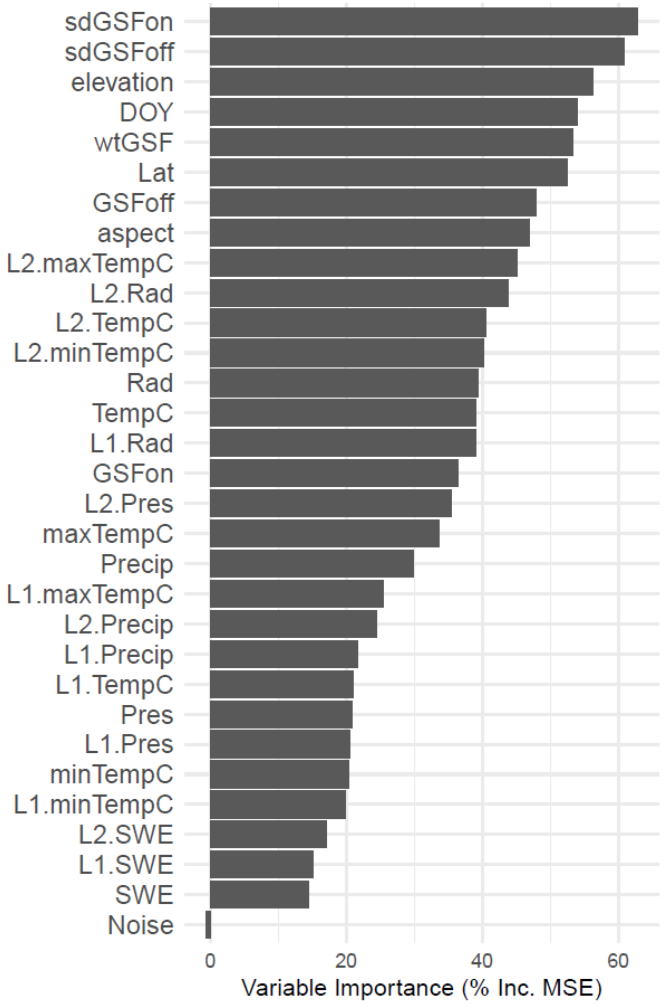

Figure S1.2. Included variable importance of predictors used in a random forest regression model to predict daily water temperature of wood frog breeding pools. Variable codes: TempC = mean air temperature, max.TempC/min.TempC = maximum/minimum air temperature, Precip = mean precipitation, Rad = mean radiation, SWE = mean snow water equivalent, L1/L2 = 1-/2-day lag of a given variable, GSFon/GSFoff = mean leaf-on/leaf-off global site factor, sdGS = standard deviation of the variance in global site factor estimates, wtGSF = weighted mean of leaf-on and leaf-off global site factor for the larval period. A random noise variable (Noise) was included as a measure of null importance.

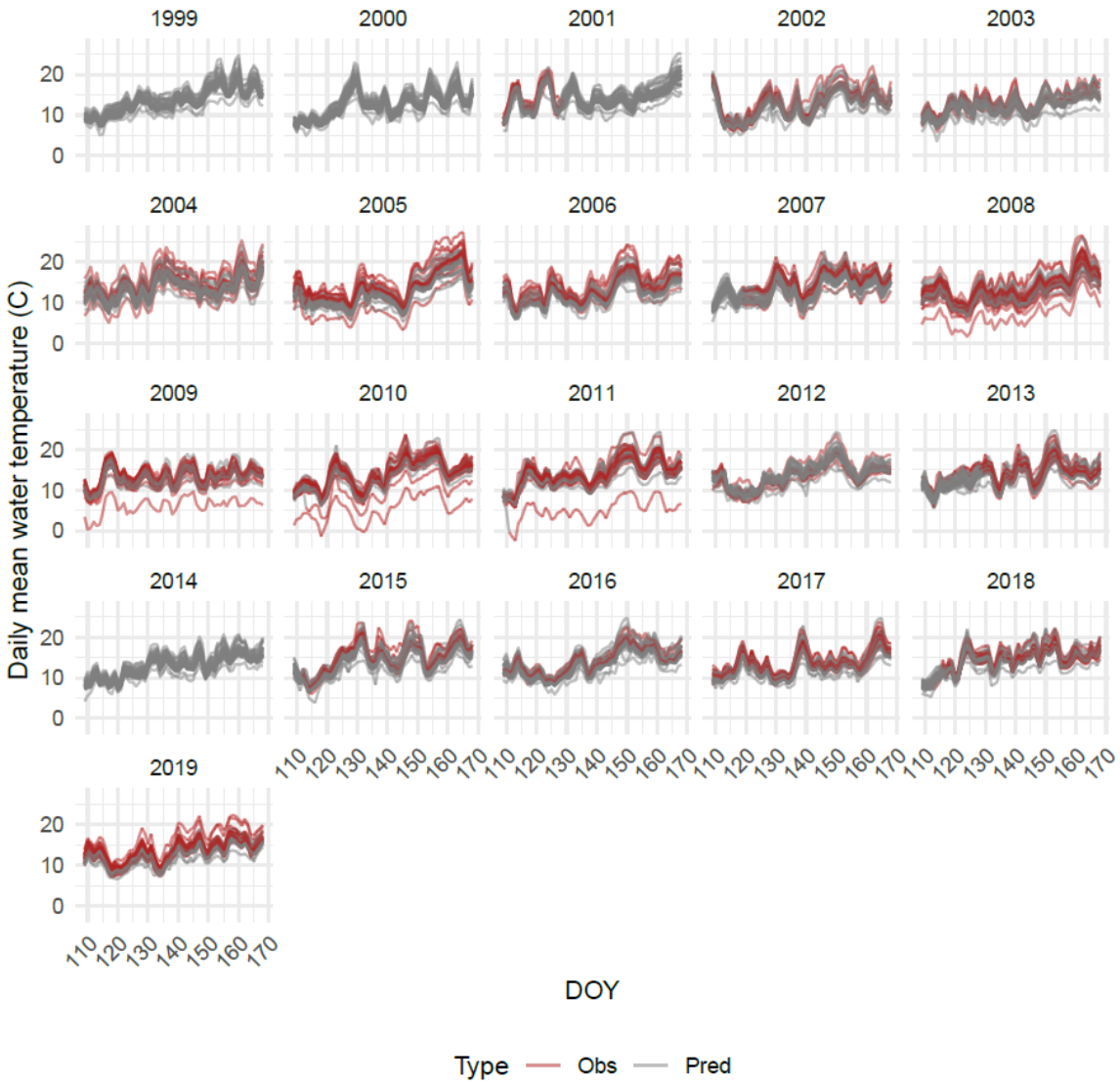

Figure S1.3. Daily mean temperature of pond water as estimated by random forest regression (grey) and observed (red) from submersible loggers. Daily estimates are averaged into annual estimates for each pond/year for inclusion in subsequent models of larval rates.

Table S1.1. Linear mixed model results of the change in annual pond water temperature by year (1999 to 2019) with random intercepts for each pond. The model was fit to the estimated annual average pond water temperature for the larval period from DOY 108 to 167.

Fixed effects	Estimate	S.E.	DF	95% C.I.
Intercept	13.03	0.196	66.39	12.6, 13.3
Year	0.025	0.005	4.44	0.014, 0.035
Random effects	Variance	S.D.		
$Intercept_{pond}$	1.572	1.254	0.66, 1.74	0.81, 1.32

Residual	0.821	0.906	0.74, 0.90	0.86, 0.95
$R^2_{Conditional}$	0.58			
$R^2_{Marginal}$	0.01			

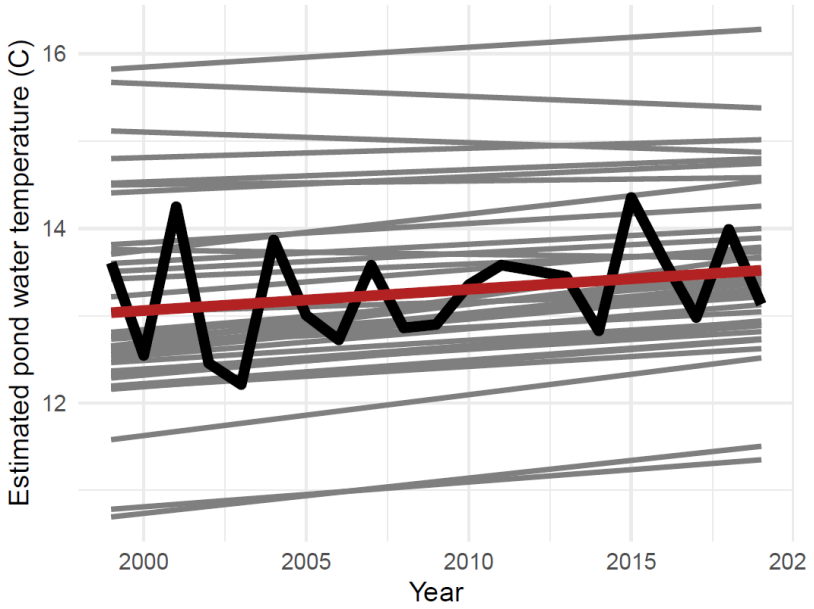

Figure S1.4. Change in average annual pond water temperature (day of year 108-167) between 1999 and 2019. The marginal estimate across all ponds from the mixed effects model (red) and individual linear model estimates for each pond (grey) are shown. The black line connects site-wide average annual pond water temperatures.

Appendix S2. Drought

Table S2.1. Person's correlations between climate variables and drought. All p << 0.001		
	SWE	PDSI
Precipitation	0.56	0.55
SWE		0.90

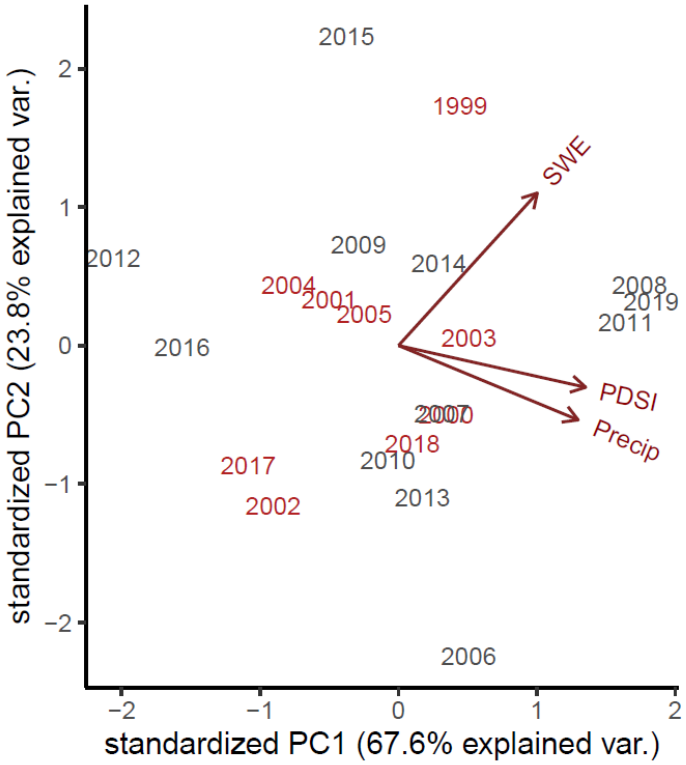

Figure. Biplot of the first two principal components of the relationship between climate and drought values. Years in which data are included in the larval dataset are denoted in red.

Appendix S3. Conspecific density

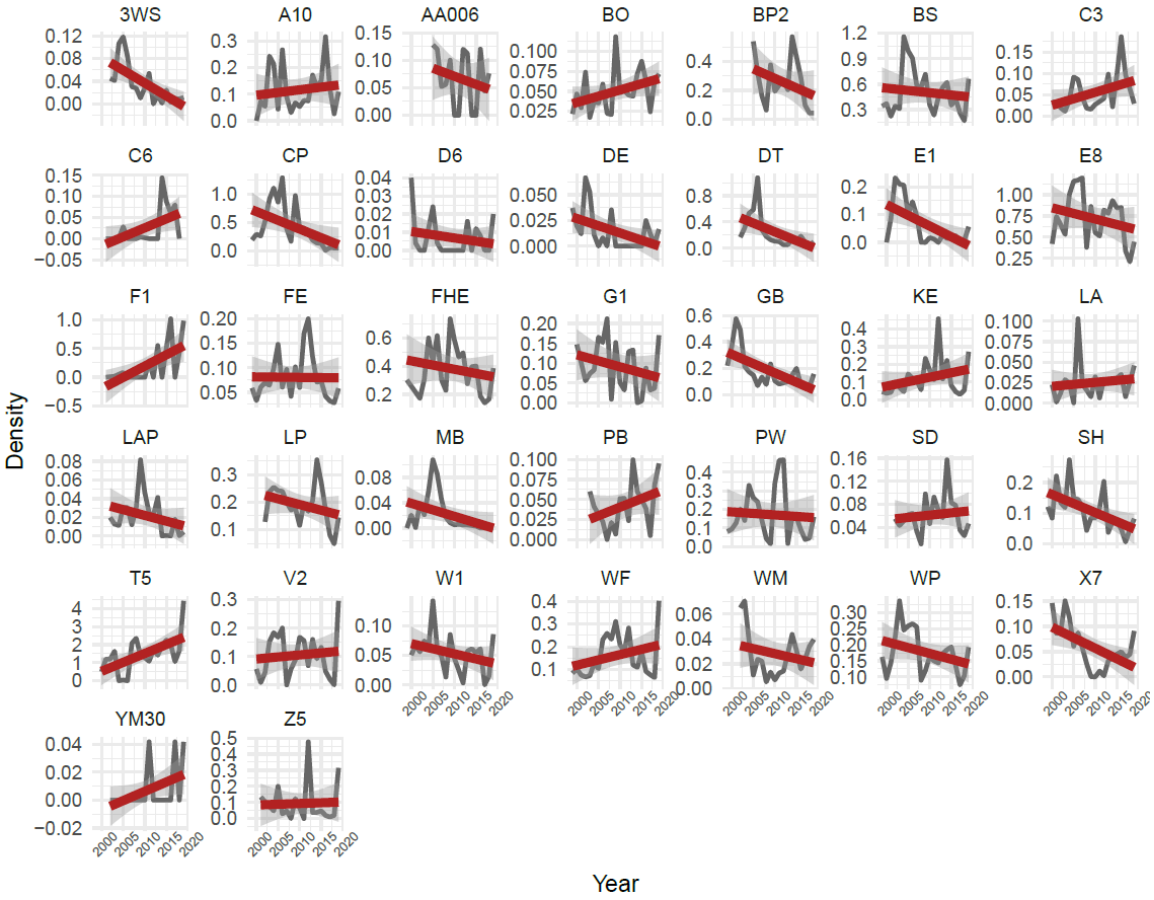

Figure S3.1. Annual wood frog density estimates (egg masses per m^2 of pond area) for 32 populations between 1999 and 2019. Linear regression models (red line) are shown. Note that y-scales differ between plots.

Table S3.1. Linear mixed model results of the change in population density (log-transformed) by year (1999 to 2019) with random intercepts for each pond. (N_{obs} = 634, N_{ponds} = 34).

Fixed effects	Estimate	S.E.	DF	95% C.I.
Intercept	54.906	13.321	598.9	30.157, 80.326
Year	-0.029	0.007	598.6	-0.041, -0.016
Random effects	Variance	S.D.		
$Intercept_{pond}$	1.309	1.144		0.858, 1.406
Residual	0.907	0.952		0.899, 1.004
$R^2_{Conditional}$	0.60			
$R^2_{Marginal}$	0.01			

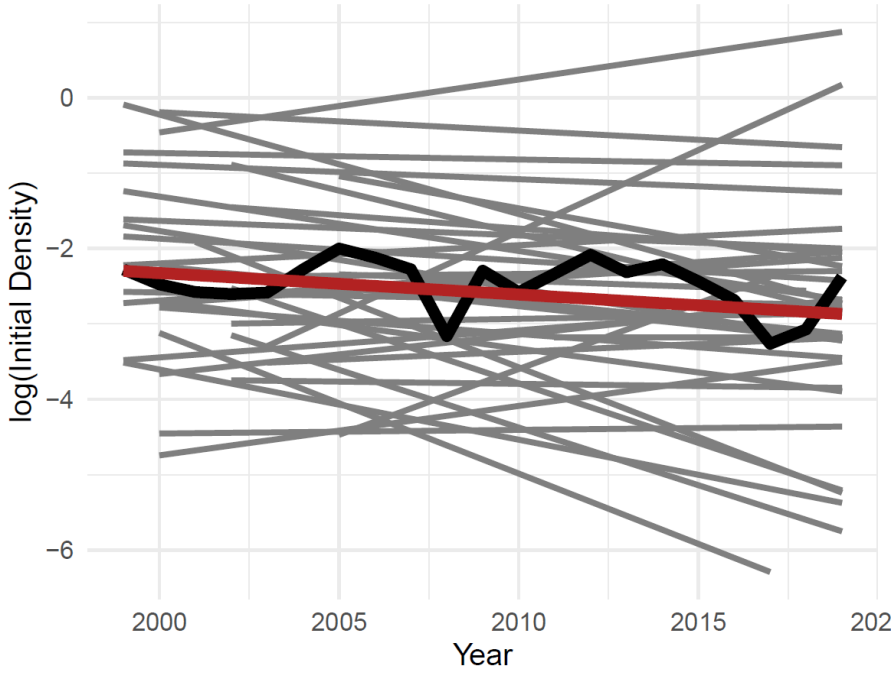

Figure S3.2. Change in wood frog density (egg masses per m^2) between 1999 and 2019 (note that density values are log-transformed). The marginal estimate across all ponds from the mixed effects model (red) and individual linear model estimates for each pond (grey) are shown. The black line connects site-wide median annual densities.

Appendix S4. Embryo size

Table S4.1. Linear mixed model results of the difference in embryo size between historical (2000 and 2000) and contemporary (2017 to 2019) samples with random intercepts accounting for observations nested within clutches and clutches nested within ponds. Parameter estimates and confidence intervals were estimated by fitting the model to 1000 bootstrap samples of the data, weighted to ensure even representation across timepoints and years.

Fixed effects	Estimate	95% C.I.
Intercept (Contemporary)	6.96	6.85, 7.06
Difference	-1.66	-1.82, -1.51
Random effects	Variance	
$Intercept_{pond}$	2.50	2.24, 2.80
$Intercept_{clutch:ond}$	2.63	2.41, 2.86
Residual	2.97	2.68, 3.27

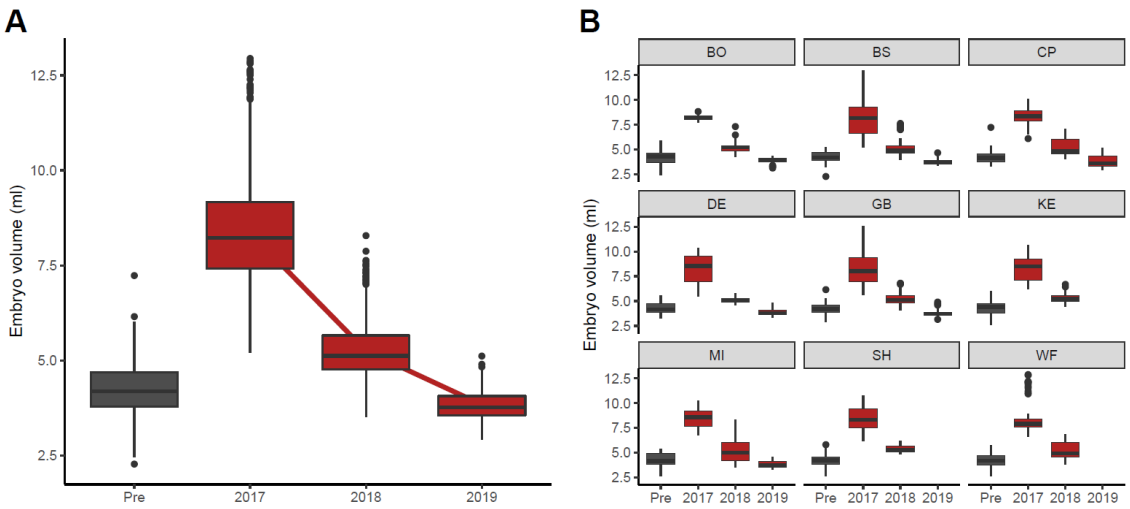

Figure S4.1. Embryo volumes are over twice as large as pre-drought in 2017, but decrease in size each year back to pre-drought levels. The figure on the left (A) shows the marginal trend across all ponds (from the linear mixed model). The figure on the right (B) shows that this trend is the same for all ponds with available data.

Appendix S5. Larval size

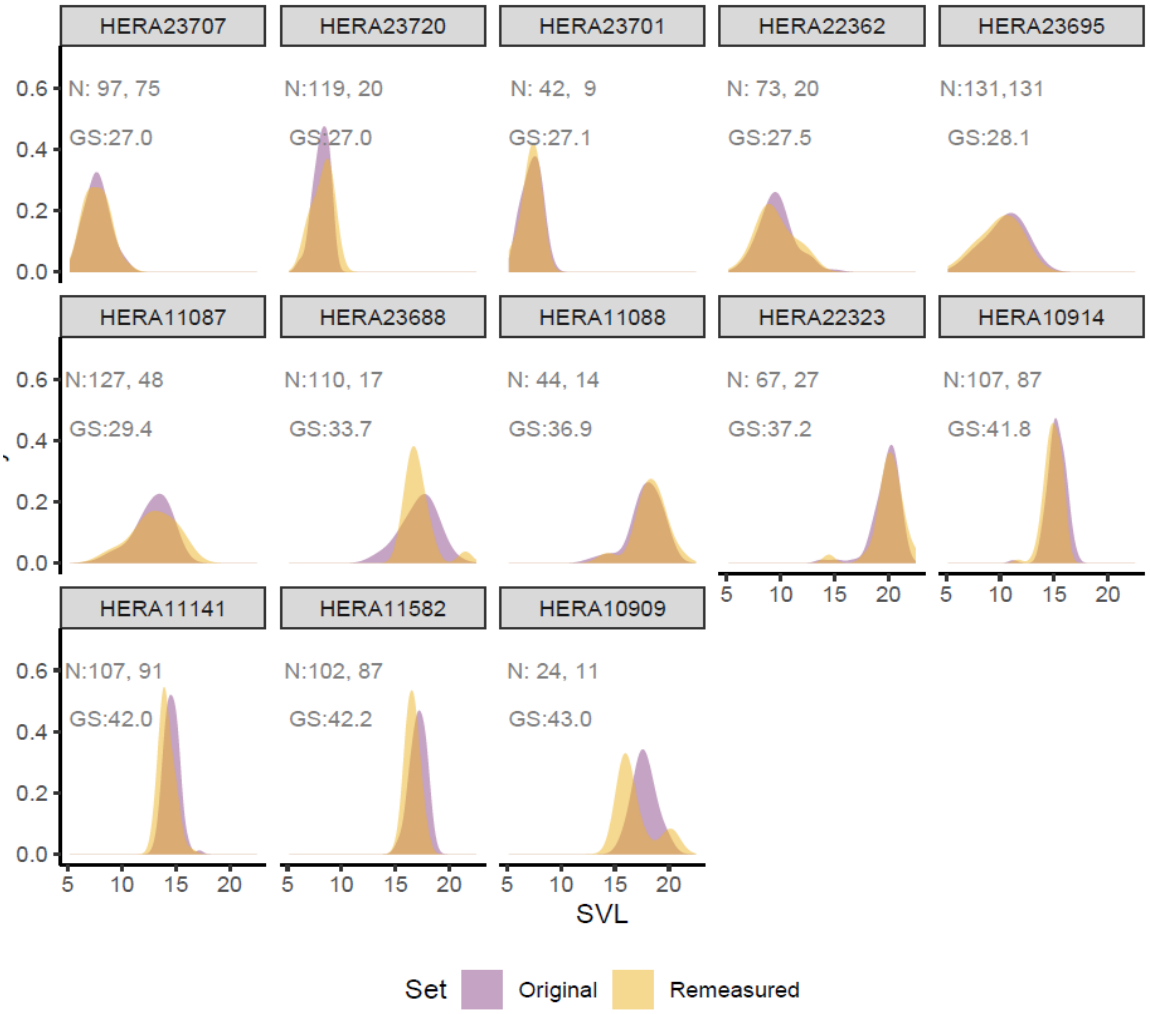

Figure S5.1. Difference in original measurements of larval and metamorphic snout-to-vent length (blue) compared to a subset of the samples that were remeasured in 2020 (yellow) as a validation between observers. Each panel represents a multi-specimen lot that was collected from a single pond on a single date, indicated by a Peabody Museum of Natural History catalog number (i.e. HERA10909). The number of samples associated with the lot in the original dataset and the number that were resamples and the average Gosner stage recorded from the original dataset are included on each panel. Note, all remaining specimens in a lot were re-measured, but only a subset of those larvae that were originally measured were placed in the collection, and could therefore be re-measured.

Table S5.1. Linear mixed model results of snout-to-vent length measurements in the historical datasets as originally recorded compared to remeasurements of a subsample of museum specimens. Measurements were validated separately for larval (Gosner stage < 42) and metamorphic (Gosner stage 42+) samples.

	Larval		Metamorphic	
Fixed effects	Estimate	p-value	Estimate	p-value
Intercept (Original)	13.37	<<0.01	16.25	<<0.01
Difference	-0.07	0.51	-0.49	<<0.01
Random effects	Variance		Variance	
$Intercept_{Lot}$	21.94		1.45	
Residual	2.32		0.51	

Table S5.2. Size differences between contemporary (2017 and 2019) and historical (1999-2005) wood frog larvae at each Gosner developmental stage of the larval period. Observations were aggregated at the level of pond cohort (each N in the dataset represents the aggregate of 11 larval observations on average). P-values were estimated with Welch's t-test.

Stage	$N_{Hist.}$	$N_{Cont.}$	$SVL_{Hist.}$	$SVL_{Cont.}$	**Δ**	p
25	**27**	**15**	**5.53**	**8.27**	**2.74**	**0.001**
26	70	8	6.75	8.00	1.25	0.093
27	82	11	8.14	10.14	2.00	0.014
28	**91**	**37**	**9.54**	**11.51**	**1.96**	**<0.001**
29	**76**	**32**	**10.95**	**13.15**	**2.20**	**<0.001**
30	**68**	**35**	**11.96**	**14.43**	**2.46**	**<0.001**
31	**69**	**39**	**13.02**	**15.40**	**2.38**	**<0.001**
32	**59**	**21**	**13.70**	**15.91**	**2.21**	**<0.001**
33	**63**	**32**	**14.32**	**16.63**	**2.30**	**<0.001**
34	**58**	**30**	**15.01**	**17.93**	**2.92**	**<0.001**
35	**67**	**34**	**15.96**	**18.53**	**2.57**	**<0.001**
36	**57**	**32**	**16.63**	**19.67**	**3.03**	**<0.001**
37	**47**	**17**	**17.07**	**19.19**	**2.12**	**0.003**
38	**52**	**16**	**17.20**	**19.23**	**2.03**	**0.008**
39	41	10	17.53	18.81	1.28	0.059
40	40	9	17.10	18.01	0.91	0.111
41	38	11	16.36	15.91	-0.45	0.401
42	33	13	15.58	16.44	0.86	0.098
43	24	10	15.72	16.39	0.67	0.276
44	20	8	15.31	15.71	0.40	0.691

45	6	3	15.12	14.56	-0.56	0.645
46	**4**	**7**	**16.88**	**14.49**	**-2.39**	**0.010**

Appendix S6. Metamorphic transition, rates of development and allocation

Larval physiology shifts dramatically as the tadpole enters the metamorphic transition. This transition is marked by a reversal in growth per stage and reduction in absolute size. We incorporate the form of this relationship into our analysis by modelling piecewise regressions that allow for independent slopes before and after a breakpoint point. In order to determine the optimum stage at which this breakpoint should occur, we repeatedly fit our maximal model predicting size from stage with the breakpoint iterating through stage 30 to 46 and consider the stage used in the best fit model as the point of metamorphic transition.

Table S6.1. Information criteria for fits of piecewise regression models from Equation 2 with different slopes before after a transition stage.

Stage	AIC	ΔAIC
30	56932.32	NA
31	55936.75	-995.573
32	55063.93	-872.824
33	54627.01	-436.92
34	54307.05	-319.954
35	54277.86	-29.1937
36	54311.66	33.8054
37	54599.07	287.4053
38	55415.75	816.6809
39	56519.03	1103.278
40	57974.25	1455.227
41	59775.46	1801.207
42	61441.74	1666.282
43	62532.53	1090.782
44	63137.95	605.42
45	63421.6	283.6512
46	64076.98	655.3785

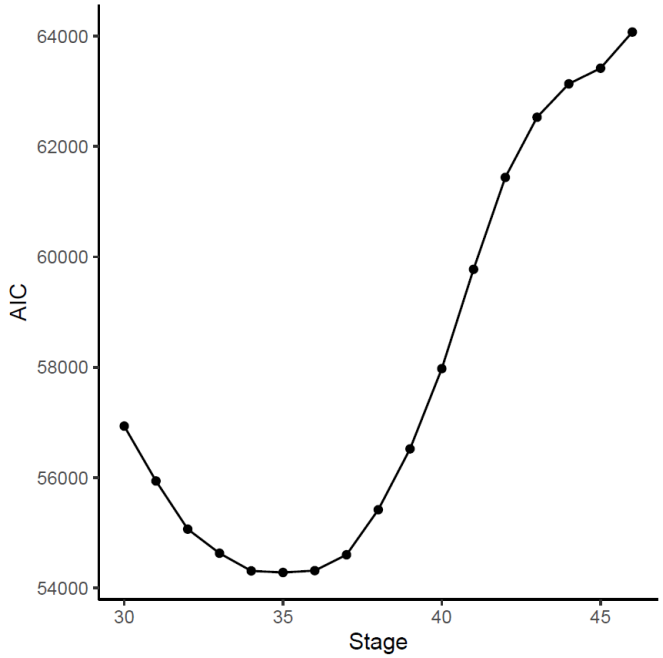

Figure S6.1. AIC values for piecewise regression models models predicting larvae size at a given developmental stage, of the form Size = PondTemp + Year + Stage + PondTemp * Stage + Year * Stage + (Stage|Pond) with different slopes before and after the transition stage. The transition stage for the best model fit (lowest AIC) was considered the best estimate of the stage at which the metamorphic transition climaxes and was used in subsequent models.

In order to model rates of growth and development over the season, we must relate the average stage of metamorphic transition, determined previously, to the day at which the transition occurred in each population and generation. However, we did not collect stage 35 larvae from every population in every year, so we use the developmental rate of observations leading up to the metamorphic transition to infer the day of year that the population attained stage 35 (i.e. the apogee of the metamorphic transition).
First, we fit a model to only the observations prior to the metamorphic transition (i.e. Stage less than 36) with the form:

$$DOY = Stage + (1|Pond) + (1|Year)$$

This specification is essentially a repeated measures model predicting the day of year from developmental stage with a random intercept for each pond and year.

Table S6.2. Linear mixed effects model predicting the seasonal timing of metamorphic transition for larval wood frogs. N_{obs} = 11883, N_{ponds} = 37, N_{years} = 9.

Fixed effects	Estimate	S.E.	D.F.	95% C.I.
Intercept	34.93	2.51	9.7	29.98, 39.87
Stage	3.80	0.03	11863.4	3.75, 3.85
Random effects	Variance	S.D.		

$Intercept_{pond}$	32.61	5.71		4.38, 7.01
$Intercept_{year}$	39.08	6.25		3.00, 9.62
Residual	54.34	7.37		7.28, 7.46
$R^2_{Conditional}$	0.77			
$R^2_{Marginal}$	0.47			

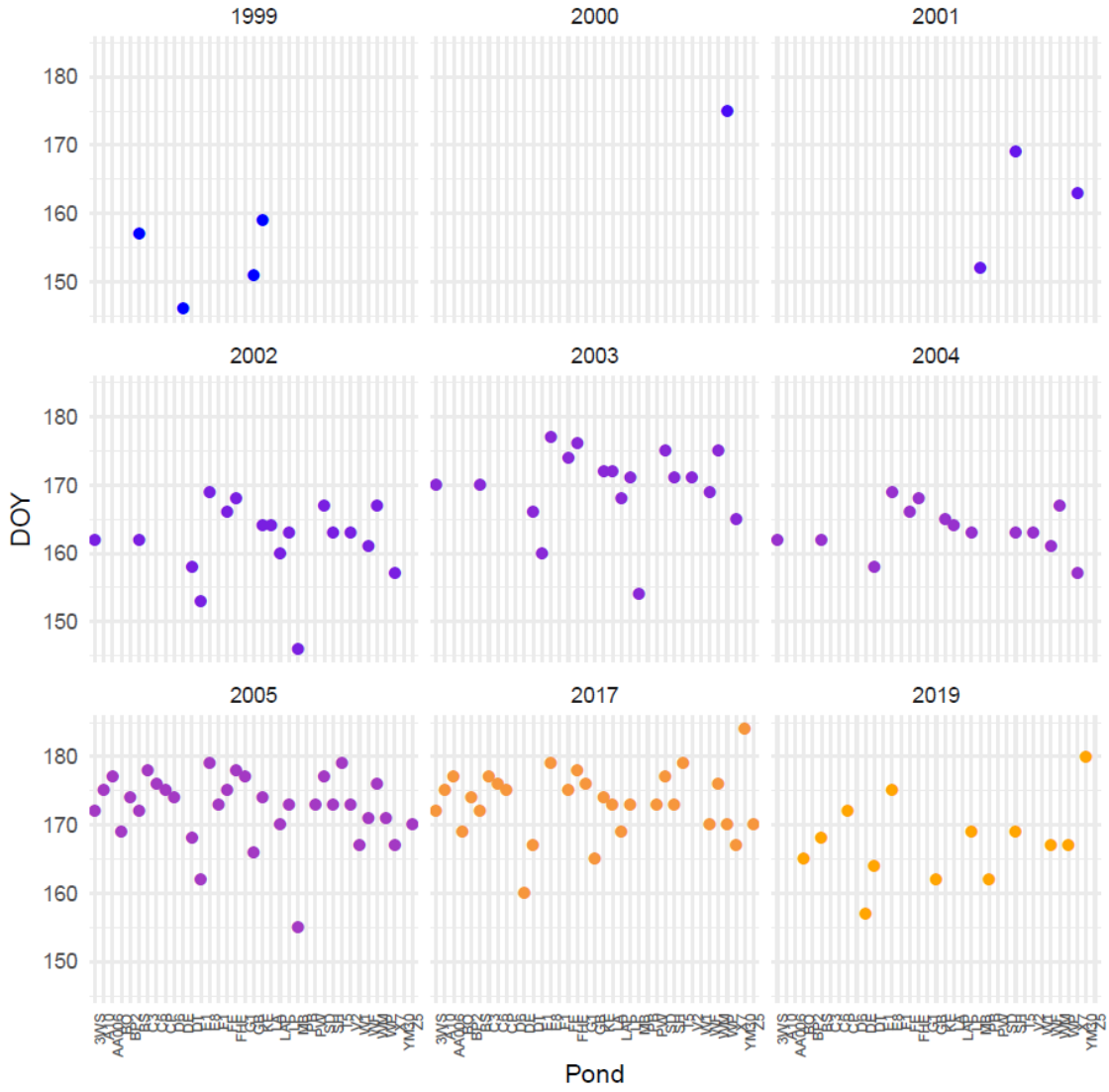

Figure S6.2. Predicted day of year at which the average wood frog larvae from a given pond cohort attained peak metamorphic transition at stage 35.

Table S6.3. Estimated change in larval size between 1999 and 2019 for four developmental stages. Values were estimated from Equation 1.

Stage	SVL_{1999}	SVL_{2019}	Difference

25	6.07	7.8	1.70
35	16.2	19.0	2.73
42	16.2	16.6	0.33
46	16.2	15.2	-1.04

Appendix S7. Ecological drivers

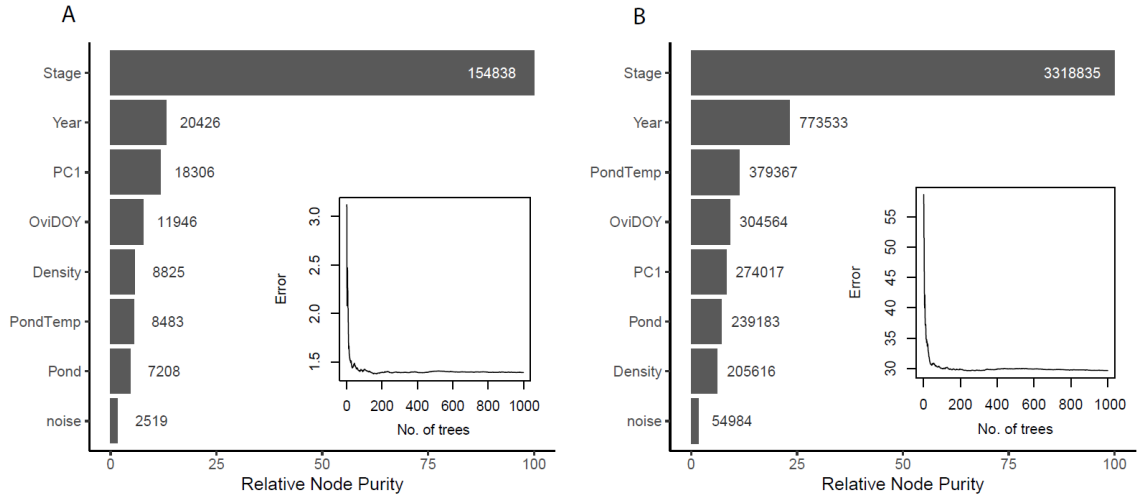

Figure S7.1. Importance of variables in predicting SVL for each stage in the allocation model (A) or day-of-year larvae attain a given stage in the development rate model (B). Out-of-bag model error for increasing numbers of trees included in the ensemble is displayed as insets, respectively. Raw importance scores are shown and variables are ranked by node purity, relative to the maximum variable. A random noise variable ("noise") was included in each model as a baseline reference of importance.

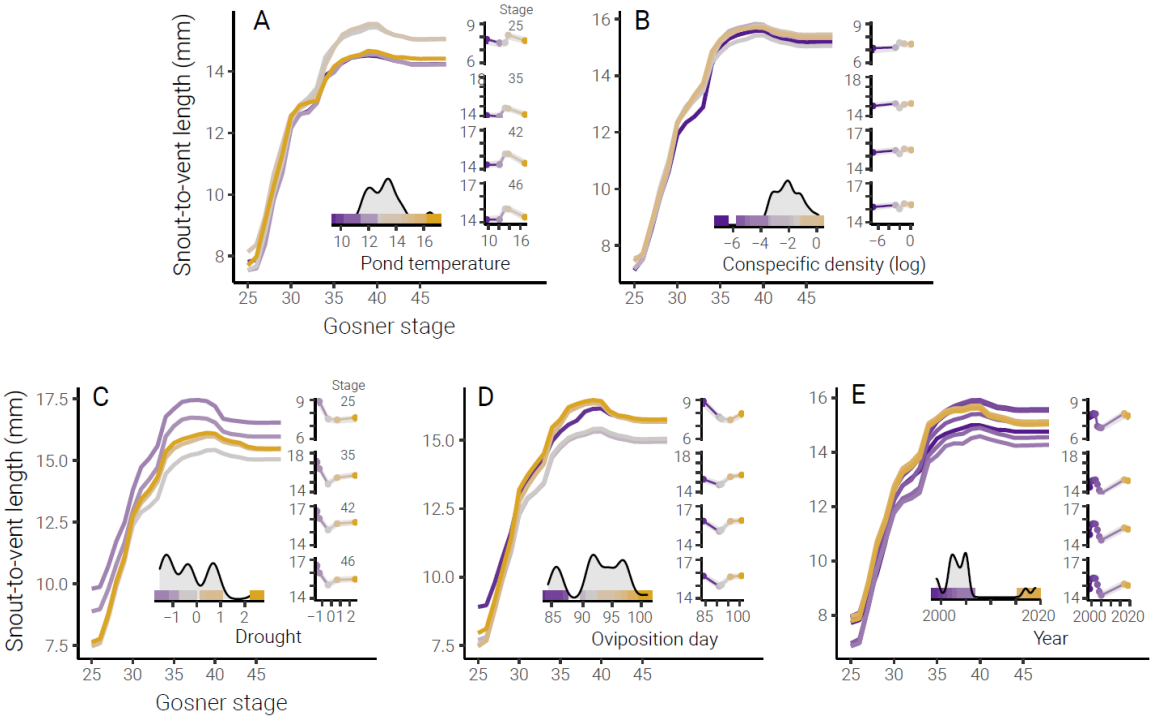

Figure S7.2. Partial estimates of allocation (snout-to-vent length in mm per Gosner stage) from a random forest model of the effect of pond temperature (in °C) (A), conspecific density (log-transformed) (B), composite drought index (C), oviposition timing (D), or year (E) while holding all other variables at the median. Lines in the main

figure of each panel represent estimates for the minimum (purple), maximum (yellow), and quartiles for the focal variable (estimates are shown for all years 1999-2005, 2017, and 2019 included in the dataset). The insets display the estimated snout-to-vent length over the range of the focal variable at stage 25 (hatching), stage 35 (metamorphic climax), stage 42 (metamorphosis), and stage 46 (juvenile). Density plots display the distribution of the focal variable in the dataset and associated color legend.

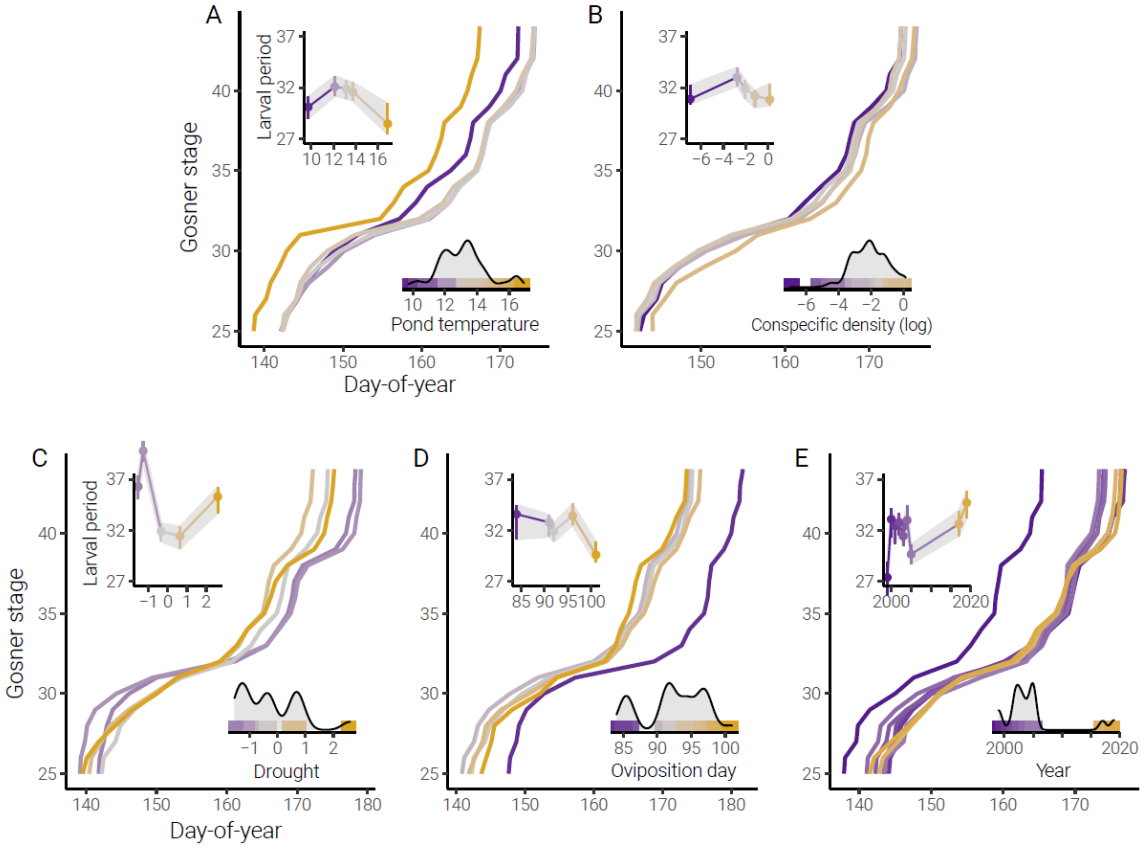

Figure S7.3. Partial estimates of development (Gosner stages per day) from a random forest model of the effect of pond temperature (in °C) (A), conspecific density (log-transformed) (B), composite drought index (C), oviposition timing (D), or year (E) while holding all other variables at the median. Lines in the main figure of each panel represent estimates for the minimum (purple), maximum (yellow), and quartiles for the focal variable (estimates are shown for all years 1999-2005, 2017, and 2019 included in the dataset). The insets display the estimated larval period as the duration between Gosner stage 25 and 42 for each focal variable value. Density plots display the distribution of the focal variable in the dataset and associated color legend.

CHAPTER 4

Temperature-mediated tradeoff between development and performance in larval wood frogs (*Rana sylvatica*)

Publication Status: Gahm, K., Arietta, A. Z. A., and Skelly, D. K. (2021). Temperature-mediated trade-off between development and performance in larval wood frogs (*Rana sylvatica*). *Journal of Experimental Zoology A*. 335:146-157. DOI: 10.1002/jez.2434. (Chapter formatted to journal standards).

ABSTRACT

Countergradient variation has been detected in diverse taxa. In a common manifestation, individuals from colder environments develop faster than conspecifics from warmer environments when placed in a common garden. Where such a pattern exists, it implies a tradeoff: individuals from warmer environments have intrinsic rates of development lower than those demonstrated by other individuals of the same species. We explored a tradeoff between development rate and locomotor performance in the wood frog (*Rana sylvatica*), an amphibian for which countergradient variation has been well documented. We reared wood frogs from 10 populations under two temperature regimes, bracketing the temperatures observed in local natural ponds. Individuals reared under warmer conditions developed more rapidly but exhibited burst speeds 20% lower than individuals reared under colder conditions. The shape of the relationship was consistent across the 10 populations and thus, we found no evidence of countergradient variation in performance. Burst speed assays of wild-caught tadpoles from the same populations showed a similar but nonsignificant trend, with greater variability among ponds. Overall, our findings support the

existence of a tradeoff that may be of broad importance and which may help explain widespread occurrence of countergradient variation.

1. INTRODUCTION

Countergradient variation emerges when phenotypic and genotypic variation are in opposition (Conover & Schultz, 1995). In the first documented example, an altitudinal cline in fruit fly size was reversed when the populations were reared under identical conditions in the laboratory (Levins, 1969). It is not easy to observe countergradient variation because, in the wild, it tends to diminish phenotypic difference and thus obscure its existence. A number of instances have been discovered as an accident while carrying out experiments for other purposes (e.g. Conover & Schultz, 1995; Levins, 1969). In the last decade or so, the number of examples has increased more rapidly as research on countergradient patterns has gained wider notice (Richardson et al., 2014).

In spite of the attention, critical aspects of the phenomenon remain poorly understood. Countergradient variation implies a tradeoff. In a typical example, higher temperatures are associated with increased development rate but populations from colder environments develop faster when reared in a common garden. While implying local adaptation to increase development in a challenging environment, the explanation for lower intrinsic rates of development in warmer conditions is seldom investigated or understood. Because of seasonal and other time constraints on organisms, rapid development should confer advantages regardless of thermal conditions. Yet populations with decreased intrinsic rates are commonly seen in common garden experiments. This widely observed pattern potentially implies an equally widespread mechanism driving lower intrinsic development rates. The goal of this study is to evaluate the potential for a tradeoff between development rate and performance. Specifically, we hypothesize that higher development

rates will be associated with decreased locomotor performance (e.g. Arendt, 2003; Arnott, Chiba, & Conover, 2006; Watkins & Vraspir, 2006).

In this study, we focus on the wood frog. Wood frogs breed in a wide variety of wetlands that vary in thermal conditions (Halverson, Skelly, Kiesecker, & Friedenburg, 2003; Skelly, 2004), variation that is at least partially driven by variation in canopy cover (Skelly, Friedenburg, & Bolden, 2014). On a microgeographic scale, development rates of both embryos and larvae (i.e. tadpoles) vary in countergradient fashion among populations—individuals from colder natal environments develop faster when reared under common garden conditions (Ligon & Skelly, 2009; Skelly, 2004). Here, we study the same populations in which countergradient variation in development rate has been observed. We ask: (1) whether burst speed performance is decreased when tadpoles are lab-reared at higher temperatures, (2) whether we see variation between pond populations in burst speed among lab-reared tadpoles, and if so, whether it is countergradient with natal pond temperatures, and (3) whether wild-caught tadpoles from the same ponds vary in burst speed performance in a countergradient fashion.

2. METHODS

2.1. Study site

Wood frog embryos were collected within 24 hours of oviposition from ten wood frog breeding ponds at Yale Myers Forest (YMF; 3213 ha, northeastern Connecticut, USA). Ponds were selected to represent a gradient of local water temperatures. We recorded water temperature in natal ponds by placing submersible loggers (HOBO 8K Pendant; Onset Computer Corporation) 10 cm below the water surface at the point of maximum depth. Loggers were placed in the ponds before breeding began.

2.2. Experiment

Lab-reared tadpoles. From each pond, we collected a portion of each of five egg masses (i.e. clutches) and stocked these across two temperature treatments in a split-clutch design in order to induce a contrast in accelerated and depressed development rates. Incubators (I-36VL; Percival Scientific, Inc.) were used to create two temperature treatments under a 12:12 h light:dark cycle. We increased incubator temperatures gradually over the course of development, approximating the temperature trajectories that wood frog tadpoles experience in the warmest and coolest natal ponds at YMF (Fig. 1). Every 14 days, we ramped the temperature, ensuring no more than 1°C increase each day. We ceased ramping incubators at 17°C.

Upon hatching, tadpoles were stocked individually into glass jars (four per treatment per clutch) containing 500 ml of reconstituted distilled (RDi) water (0.06 g aquarium salt (API Aquarium Pharmaceuticals) per 1 L of distilled water (final conductivity ~120 μS). Each incubator was stocked with 200 tadpoles (ten natal ponds, five clutches per pond, four individuals per clutch). We blocked tadpoles by clutch across five shelves in each incubator with individuals randomized on each shelf. We cleaned jars and replaced water every 3 or 4 days. At the same time, we fed each tadpole approximately 10% of body mass (3:1 ratio of powdered Kaytee Rabbit Chow (Kaytee Products, Inc.) and TetraPro Goldfish Food (Tetra Gmbh). We estimated body mass from a subset of tadpoles not included in the experiment but housed in the incubators alongside the experimental tadpoles. Jars were rotated within blocks after each cleaning. The day before conducting burst speed assays, we removed tadpoles from the incubators and fasted them for 24 hours while allowing them to acclimate to the conditions in a temperature controlled animal room maintained at the same light and temperature as the final incubator temperatures (17 °C with 12:12 h light:dark cycle).

Wild-caught tadpoles. We revisited the natal ponds at YMF every two weeks. Once wild tadpoles reached Gosner (1960) stage 30 or greater, we collected 15-20 individuals and stocked them into the animal room within individual 500ml glass jars, feeding them approximately 10% of body mass every 3 days. We maintained wild tadpoles in the lab for 1 to 5 days prior to performance assays to acclimate them to experimental temperatures (Brattstrom, 1968). Wild tadpoles were also fasted for 24 hours immediately prior to their burst speed assays.

In two natal ponds, E1 and MI, we were not able to recover any wild-caught tadpoles, despite several attempts. For the ponds from which we were able to collect wild tadpoles, the median number of tadpoles analyzed was 18.5 (minimum = 9; maximum = 20).

2.3. Developmental rates

We estimated development rates as the difference in developmental stage from oviposition (stage 1) to the time of the performance assay, divided by the number of days since oviposition. We tested for differences in development rates between tadpoles from laboratory temperature treatments and the wild with an ANOVA. We further tested the proximate effect of average incubator temperature or average pond temperature on development rates with ordinary least squares regression.

Ectotherm growth and development proceed interactively and are affected by many environmental stimuli (Kingsolver & Huey, 2008). As a result, allocation of resources to growth or development is variable across tadpole ontogeny in different environments (Feder & Burggen, 1992). We estimated allocation as deviance from an average mass at a given stage by extracting the residuals of linear models fit within and between lab-reared

and wild-caught tadpoles. We used these size-at-stage residuals fit within groups as covariates in subsequent models of performance.

2.3 Performance

Burst speed assays. We assessed performance as the burst speed of a tadpole's startle response to a simulated predator attack. We conducted assays of burst on both lab-reared and wild-caught tadpoles in a custom-built arena setup. The arenas consisted of two 45 x 65 cm white, shallow trays filled with RDi water to 1.5 cm depth in order to confine tadpoles to a horizontal swimming plane, following Arendt (2003). Arenas were illuminated using LED light strips and overhead lamps, and a camera (Hero 4 (GoPro, Inc.), 1080p, 60 f/s, narrow FOV) was fixed above each arena.

To estimate each tadpole's burst swimming speed, we placed a translucent acrylic cylinder over the tadpole to confine it to the center of the arena for a 45 s acclimation period, plus an additional 15 s motionless period. If the tadpole moved, the 15 s count was restarted. Following the 15 s motionless period, we lifted the cylinder and tapped the tadpole's tail from behind with a blunt probe to induce a startle response (Arendt, 2010). Once the tadpole reached the edge of the arena or stopped swimming, we recaptured it in the cylinder and returned it to the center of the arena. If the tadpole bolted after the cylinder had been lifted but before its tail had been tapped, it was returned to the center and allowed an additional 30 s motionless rest period before initiating another startle response.

We conducted at least three trials for each tadpole, allowing a 15 s motionless rest period between trials. If the tadpole reacted minimally to the probe, we returned it to the center and conducted an additional trial after an additional rest period.

We measured water temperature in the arena using an infrared probe at the beginning of the sequence of trials for each tadpole. We maintained water temperature in the arenas at a mean of 17.0 °C (range = (15.5 – 18.1 °C), std = 0.54 °C).

Morphometrics. We euthanized tadpoles within 24 h after their burst speed trials and recorded tadpole wet-mass (to 0.001 g). We determined developmental stages with a dissecting microscope and captured lateral photographs for morphometric analysis. We used 26 morphometric landmarks (Fig. 2) to estimate each tadpole's body shape, following Arendt (2010), Johansson, Lederer, & Lind, (2010), and Urban et al. (2017). Landmarks were digitized from the photographs, using tpsDig (v2.31). We used the **geomorph** package (Adams & Otárola-Castillo, 2013) to perform a generalized procrustes analysis on the landmarks and to generate principal components of lateral body and tail shape variables for further analysis (Adams, Collyer, & Kaliontzopoulou, 2020). We considered the first two principal components generated from the morphometric landmarks, which together explained 66.4% of the variation in tadpole shape, for inclusion in the model. PC1 roughly corresponded to the ratio of body length to total length, and PC2 was an approximate measure of the ratio of total length to body depth. Of these two, PC2 was a stronger predictor of burst speed in a univariate regression, so we selected it for inclusion in the model (Fig. 2).

Burst speed analysis. We used Adobe Premiere Pro to edit the videos of performance assays into separate trials. Video sequences were corrected to account for lens distortion from the cameras that might otherwise have affected measurements of distance and speed, especially near the edges of the tray. We used a custom script in MATLAB to identify the position of the tadpole's centroid in each video frame and track its trajectory through each video clip.

We extracted burst speeds as the straight-line distance covered by the tadpole in the initial half-second of startle response measured from the first frame in which the tadpole moved from its starting position. Consistent with (Arendt, 2003), paired t-tests between the first and second half second showed that tadpoles began to slow after the first half second.

2.4. Performance analysis

We estimated performance and developmental trade-offs by fitting linear mixed effect models predicting performance from development rates while accounting for potential confounding variables and pseudoreplication. Mixed models were fit in **lme4** (v1.1.21; Bates, Mächler, Bolker, & Walker, 2015)**.** Based on prior work, we hypothesized that developmental stage, shape, size-at-stage (allocation), and the water temperature of the experimental arena, would affect tadpole burst speeds (Goldstein, Hoff, & Hillyard, 2017; Van Buskirk & McCollum, 2000). Accordingly, we included developmental stage, size-at-stage, arena temperature, and a principal component of shape from our morphometric analysis as additive covariates. Because allocation rates may impact the trade-off between performance and development rates, we also included the interaction between size-at-stage and development rates. We treated pond, clutch, and individual identity as nested random intercepts to account for hierarchical sample structure and repeated measures. We fit identical models for lab-reared and wild-caught tadpoles with the exception that clutch affiliation was unknown for wild tadpoles, so the random effect included only the individual identifier nested within pond.

From these models, we used partial regression to estimate burst speeds while holding confounding variables (all independent variables except development rate) constant. These corrected burst speed values were then used to test for countergradient variation in performance and to estimate performance plasticity of lab-reared tadpoles.

To test for countergradient variation in burst speed with respect to pond temperature, we examined variance estimates for the nested random effects in our lab and wild model outputs. If the pond-level random effect coefficient was estimated to be greater than zero, indicating differences among ponds, we tested for a relationship between pond temperature and corrected burst speed.

For all mixed effect models, we calculated 95% confidence intervals and p values for main effect coefficients from 1000 bootstrap iterations with the **parameters** package (v0.5.0; Makowski, Ben-Shachar, & Lüdecke, 2019), and estimated conditional and marginal R^2 values with Nakagawa & Schielzeth's (2013) method implemented in the **MuMIn** package (v1.43.15; Bartón, 2019). We calculated 95% confidence intervals for random effect coefficients from 1000 bootstrap iterations in **lme4** (v1.1.21; Bates et al., 2015). All modelling and statistical analyses were performed in R (v3.6.2; R Core Team 2019).

We estimated repeatability and overall variance of our performance metric from the random effect estimates of our models. For both lab-reared and wild-caught tadpoles, we estimated repeatability as the proportion of total variation due to individual differences (Bell, Hankison, & Laskowski, 2009). Repeatability quantifies the consistency of a trait within an individual over time relative to total measurable differences, and may serve two purposes. It can be interpreted as an upper bound for the heritability of a trait and an index of the accuracy of the trait value estimate (Bell et al., 2009). Because our measurements were taken rapidly over the course of a single trial rather than over an acclimation period (Biro, 2012), our estimate primarily serves the latter purpose.

To estimate plasticity of burst speed with respect to development rates, and to compare the plastic response across ponds and clutches, we evaluated burst speeds for clutch-mates reared at high versus low temperatures. First, we estimated mean burst speeds for each tadpole and for each clutch-temperature treatment combination. The scope

of the plastic response, *D*, is defined as the mean at high resource availability minus the mean at low resource availability (Stearns, 1992; Valladares, Sanchez-Gomez, & Zavala, 2006). Because we hypothesized a negative relationship between developmental temperature and burst speed, we considered the low temperature treatment to correspond to "high resource availability" for this calculation. In order to preserve our sample size, we subtracted the high-temperature clutch mean from each of the low-temperature individual means, and each of the high-temperature individual means from the low-temperature clutch mean. No tadpoles from one of the clutches from CPS pond remained in the low temperature treatment, so we excluded this clutch from our analyses of plasticity for lack of a basis of comparison for the high-temperature tadpoles.

3. RESULTS

3.1. Performance trade-offs

Among lab-reared tadpoles, performance was negatively associated with development rate (Table 1, Fig. 3). After accounting for confounding variables, tadpoles reared at the low temperature exhibited 20.6% faster burst speeds (mean = 70.89 mm/s, SD = 8.75 mm/s) compared to tadpoles in the high temperature treatment (mean = 58.76 mm/s, SD = 11.19 mm/s). Development rates in the low temperature treatment (mean = 0.40 stage/day) were 39.9% slower than those reared at warm temperatures (0.67 stage/day). Thus, our experiment revealed a tradeoff: we measured a reduction of 4.6 mm/s in burst speed for each 0.1 stage/day increase in development rates (Table 1, Fig. 3A).

In the lab, development rates increased by 0.09 stage/day for 1°C ($F_{1, 305}$ = 3501, R^2 = 0.92, $p < 0.001$). In the wild, development rates (0.50 stage/day) fell between those of lab-reared treatments and were positively associated with average 2019 pond temperatures, increasing by 0.02 stage/day for each 1 C ($F_{1, 126}$ = 32.2, R^2 = 0.20, p < 0.001).

The model suggests a similar performance trade-off in wild-caught tadpoles (Fig. 3B). Burst speed decreased by 6.31 mm/s for every 0.1 stage/day increase in development rate, however confidence intervals for the slope coefficient included 0 (Table 1).

3.2. Countergradient variation

We found no evidence of variation in performance, and therefore no evidence of countergradient variation, among pond populations for lab-reared tadpoles (σ = 0.00, C.I = 0.00 - 2.64)(Table 1). Average deviation in performance between populations for wild-caught tadpoles was 5.00, although the confidence interval includes zero (C.I. 0.00 - 9.53)(Table 1). Although non-significant, the association between performance of wild-caught tadpoles and natal pond temperatures is negative, as would be expected in the case of countergradient variation (-2.58 mm/s/°C, $F_{1, 126}$ = 2.715, R2 = 0.01, p = 0.10).

3.3 Covariates of performance

For both wild-caught and lab-reared tadpoles, slope coefficients for performance were positively associated with arena temperature, size-at-stage, and developmental stage, although all coefficients with the exception of developmental stage in the model of lab-reared tadpole performance had confidence intervals including 0 (Table 1). Tadpoles were of similar developmental stage across both lab treatment groups (High median = 35, range = 31-37; Low median = 34, range = 29-36) and wild-caught tadpoles (median = 34, range = 28-38). Likewise, all groups had similar ranges of mass (High: mean = 0.39, SD = 0.10; Low: mean = 0.48, SD = 0.13; Wild: mean = 0.57, SD = 0.19).

Morphology was also weakly associated with performance (Table 1). Body shape varied considerably between wild and lab-reared tadpoles (Fig. 2B). PC2 primarily differentiated wild-caught tadpoles, which had a high ratio of tail surface area to body size

and a low ratio of tail surface area of tail muscle depth, from lab-reared tadpoles, which tended to have smaller tails but deeper tail muscles (Figure 2B). More 'lab-like' morphology tended to confer faster burst speeds (Table 1).

Lab-reared tadpoles exhibited nearly identical allocation rates across temperature treatments, with tight correlation (High rate = 0.067 g/stage, Low rate = 0.053 g/stage, $F_{2, 304} = 457.5$, $R^2 = 0.75$, $p < 0.001$). Although the rate of allocation among wild tadpoles was similar to those in the lab (0.055 g/stage, $F_{1, 127} = 147.1$, $R^2 = 0.53$, $p < 0.001$), tadpoles collected from the wild tended to be significantly larger at most stages and the variation much greater (Supplemental Figure 2). The difference in size at a given stage between wild-caught and lab-reared diminished for tadpoles approaching the metamorphic transition, however (Supplemental Figure 2).

3.4 Repeatability and variation

Our dataset included an average of 2.88 burst speed trials per tadpole for 307 lab-reared tadpoles ($N_{High} = 164$ and $N_{Low} = 143$) and 128 wild-caught tadpoles, totalling 1251 observations. Repeatability of burst speeds was similarly high for measurements of both wild (0.81) and lab-reared tadpole (0.73) burst speeds.

After accounting for variation in performance among measurements for an individual tadpole, within-pond variation of performance of wild-caught tadpoles was much higher ($\sigma = 17.14$, 95% C.I. = 14.69 - 19.69) than between-pond variation ($\sigma = 5.00$, 95% C.I. = 0.00 - 9.53) (Table 1). For lab-reared tadpoles, there was no discernable variation between ponds ($\sigma = 0.00$, 95% C.I. = 0.00 - 2.64), nor between clutches within-ponds ($\sigma^2 = 0.00$, 95% C.I. = 0.00 - 3.57), indicating low heritability of performance traits in these populations. Likewise, we did not find a strong indication of divergence in performance among populations (Table 1). The total variance in performance in lab-reared tadpoles was

driven by variation among siblings (σ = 11.22, 95% C.I. = 9.86 - 12.40) (Table 1). In general, total variation in performance was 2.3 times greater in the wild than in the lab, indicating strong environmentally-induced plasticity of performance among wild ponds.

3.5. Plasticity

Plasticity in performance was apparent in most clutches and roughly equal among ponds (Supplemental figure 3). Most of the variance in plasticity is attributable to differences between individuals within clutches (variance = 8.6) and between clutches within ponds (variance = 5.0).

4. DISCUSSION

Temperature-mediated development rate trades off with swimming performance in wood frog tadpoles. Our analysis suggests that tadpoles reared at higher temperatures develop more rapidly and swim 20% more slowly than their low-temperature counterparts, a difference of 12.1 mm/s [4.03 mm/s/C]. In magnitude, our findings are largely congruent with the few prior studies of amphibians that have detected decreased swimming performance with more rapid development [fire-bellied toads (*Bombina orientalis* [approx. 1.34 mm/s/C (Parichy & Kaplan, 1995)], Pacific tree frogs (*Hyla regilla*) [6.46 mm/s/C (Watkins, 2000)], and wood frogs (*Rana sylvatica*) [1.2 mm/s/C (Watkins & Vraspir, 2006), see Seebacher & Grigaltchik (2015) for an exception]. The consistency of effect among these studies is matched, within our study, by the strong consistency across 10 natural populations of the relationship between development rate and locomotor performance. While we observed a comparable trend among wild-caught individuals from the same populations, measurements for wild individuals were more variable, a situation we discuss below.

What could be the proximate mechanism for this tradeoff? Burst swimming integrates complex interactions of biophysical processes, including neuromuscular recruitment of muscle fibers and the timing of contraction and release of bilateral muscles, all of which are developmentally dependent (Feder & Burggen 1992). The aerobic properties of the muscles may also differ between treatment groups. Arnott et al. (2006) proposed that disparity in swimming speeds may result from competing metabolic demands of growth and performance. Watkins (2000) showed that ATPase activity, which is related to muscle force and contraction velocity, was elevated in *Hyla regilla* tadpoles reared at a cooler temperature. The same relationship has also been demonstrated in teleost fish (e.g. Johnson, Bennett, & McLister, 1996) and crocodilians (Seebacher & James, 2008). While the mechanism underlying our findings remains unexplored, the potential role of temperature in the development of muscle morphology and physiology deserves further study and could prove to be more widely relevant among ectotherms.

Rapid development rates are generally advantageous for pond breeding amphibians, whose larvae must navigate waters filled with gape-limited predators and reach metamorphosis before their pond dries. Despite these strong incentives for maximizing development rates, larval amphibians from warmer habitats tend to develop more slowly than their cold-habitat counterparts when reared at a common temperature. Several independent studies of wood frogs have found this countergradient effect in embryonic (Skelly, 2004) and larval (Ligon & Skelly, 2009) development rates. Similar findings are known in other species (Berven, Gill, & Smith-Gill, 1979; Laugen, Laurila, Räsänen, & Merilä, 2003; Rödin-Mörch et al., 2019). The fact that populations in warmer ponds counteract extrinsic factors through depressed intrinsic development rates indicates that there is a cost to rapid development.

Our results suggest that in wood frogs, this cost comes, at least partly, in the form of a detriment to locomotor performance. Indeed, environmentally-mediated growth increases have been shown to compromise performance in a range of ectothermic vertebrates (Álvarez & Metcalfe, 2007; Arendt, 2003; Li et al., 2007; Parichy & Kaplan, 1995; Watkins, 2000; Watkins & Vraspir, 2006). Additionally, a growing body of literature has examined how performance trades off with genetically-mediated growth (Arendt, 1997; Billerbeck, Lankford, & Conover, 2001; Gregory, Gregory, & Wood, 1998; Kolok & Oris, 1995). Among potential implications, rapid development could compromise the ability of larvae to evade capture by predators––a potential direction for future study. Other aspects of performance, not included in our study, could also be compromised, including terrestrial locomotion: several studies in anurans have indicated that costs of rapid development persist after the metamorphic transition (Álvarez & Nicieza, 2003; Ficetola & De Bernardi, 2006). These costs may exert a selection pressure toward slower intrinsic growth rates to mitigate the performance disadvantage.

Notably, we find no evidence of divergence in burst speed among pond populations. Neither mean performance nor plasticity in performance varies among lab-reared tadpoles originating from different ponds (Table 1, Supplemental Figure 3). Studies of this metapopulation have consistently found microgeographic divergence in other traits, including thermal preference (Friedenburg & Skelly, 2004) and critical thermal maximum (Skelly & Friedenburg, 2000), in addition to embryonic and larval development rates (Skelly, 2004; Ligon & Skelly, 2009). That burst speed alone does not differ among populations suggests that it may drive the formation of countergradient variation in development rates.

The relationship between development rate and performance in wild-caught tadpoles was weaker than in lab-reared individuals, and variability was higher (Figure 3).

Environmental factors in the wild ponds likely account for this variation. Compared to their lab-reared cohort-mates, wild tadpoles experienced more variability in temperature during development. Pond temperatures represent means across several months, obscuring both diurnal temperature variability and large swings, such as those that occurred in late April and mid-May 2019 (Figure 1A). The temperature environments in our study ponds are spatially heterogeneous (Skelly et al., 2014) and tadpoles exhibit thermal preference through their behavior (Freidenburg & Skelly, 2004; Herreid & Kinney, 1967). Pond temperatures estimated from a single logger in a fixed location likely do not reflect tadpoles' experiences of the pond thermal environment and may contribute to the greater variation in burst speed observed relative to lab-reared individuals. (Supplemental Figure 6).

Still, wild tadpoles' development rates show only a weak negative relationship with performance, indicating that other environmental factors may be in play (Figure 3B). In the lab, survival was high, conditions controlled, and predators absent. By contrast, the tadpoles that we captured from the wild were those that had survived whatever episodes of selection had whittled down the initial cohort by the time they were collected. These selective pressures almost certainly differed among ponds. Selection on burst speed likely affected the distribution of speeds in our sample of wild tadpoles, altering the development-performance relationship compared with that observed in their lab-reared cohort-mates.

Temperature is an external force that shapes nearly every feature of ectotherms. If thermal biology could be understood as a simple function of distinct biological reactions, we could scale up the temperature dependencies of individual processes to understand how temperature impacts the function and fitness of organisms, populations, or even species. However, physiological functions depend on temperature in different ways that are sometimes in conflict, as our study demonstrates. Understanding tradeoffs like the one considered here helps us to better grasp the complex interrelated processes involved in

ectotherm physiology. The consistency of the development-performance tradeoff among populations in our study, together with similar findings across taxa, suggests that performance costs of rapid growth and development are a fundamental feature of ectotherm life history.

In a warming world, it is critical to understand how climate change will affect organisms. Amphibians are especially vulnerable to climate change (Li, Cohen, & Rohr, 2013), and ectotherms in general are likely to be particularly affected by warming temperatures (Parmesan, 2006). In the case of pond-breeding amphibians like the wood frog, the dual pressures of warming temperatures and shortening hydroperiods should favor more rapid development. But as our study shows, developing more rapidly can come with performance costs. Conceivably, organisms could counteract the costs of temperature-mediated rapid development by evolving adaptively slower intrinsic development rates. The degree to which rapid evolution will support the persistence of species facing changing climate is a challenging but critical topic for further study.

Authorship Statement – This study was conceived by KG, AA, and DS. KG and AA conducted the experiment and analyses. KG, AA, and DS wrote the manuscript.

Acknowledgements – We thank Adriana Rubinstein for assistance in the field and laboratory, Simon Stump and Kirby Broderick for assistance with MATLAB code, Baasim Zafar for assistance in data analysis, and Bayla Arietta for the illustration in Figure 2. Funding support was provided by Yale Institute for Biospheric Studies, the Kohlberg-Donohoe Research Fellowship, the Pierson Richter Fellowship, the Yale College Dean's Research Fellowship in the Sciences, and the Yale Peabody Museum. All methods described conform

to Yale IACUC protocol 2019-10361, and specimens were collected under CT DEEP Permit 0116019b.

REFERENCES

Adams, D. C., Collyer, M. L., & Kaliontzopoulou, A. (2020). Geomorph: Software for geometric morphometric analyses. R package version 3.2.1. https://cran.r-project.org/package=geomorph

Adams, D. C., & Otárola-Castillo, E. (2013). Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*, 4, 393–399. doi: 10.1111/2041-210X.12035

Álvarez, D., & Metcalfe, N. B. (2007). The tradeoff between catch-up growth and escape speed: variation between habitats in the cost of compensation. *Oikos,* 7, 1144–1151. doi: 10.1111/j.0030-1299.2007.15861.x

Álvarez, D., & Nicieza, A. G. (2003). Predator avoidance behaviour in wild and hatchery-reared brown trout: the role of experience and domestication. *Journal of Fish Biology*, *63*, 1565–1577. doi: 10.1111/j.1095-8649.2003.00267.x

Arendt, J. D. (1997). Adaptive intrinsic growth rates: an integration across taxa. *The Quarterly Review of Biology*, *72*, 149–177. doi: 10.1086/419764

Arendt, J. D. (2010). Morphological correlates of sprint swimming speed in five species of spadefoot toad tadpoles: comparison of morphometric methods. *Journal of Morphology*, *271*, 1044–1052. doi: 10.1002/jmor.10851

Arendt, J. D. (2003). Reduced burst speed is a cost of rapid growth in anuran tadpoles: problems of autocorrelation and inferences about growth rates. *Functional Ecology*, *17*, 328–334. doi: 10.1046/j.1365-2435.2003.00737.x

Arnott, S. A., Chiba, S., & Conover, D. O. (2006). Evolution of intrinsic growth rate: metabolic costs drive trade-offs between growth and swimming performance in *Menidia menidia*. *Evolution*, *60*, 1269–1278. doi: 10.1111/j.0014-3820.2006.tb01204.x

Bartón, K. (2019). MuMIn: Multi-Model Inference. https://CRAN.R-project.org/package=MuMIn

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. doi: 10.18637/jss.v067.i01

Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, *77*, 771–783. doi: 10.1016/j.anbehav.2008.12.022

Berven, K. A., Gill, D. E., & Smith-Gill, S. J. (1979). Countergradient selection in the green frog, *Rana clamitans*. *Evolution*, *33*, 609–623. doi: 10.2307/2407784

Billerbeck, J. M., Lankford, T. E., Jr, & Conover, D. O. (2001). Evolution of intrinsic growth and energy acquisition rates. I. Trade-offs with swimming performance in *Menidia menidia*. *Evolution*, *55*, 1863–1872. doi: 10.1554/0014-3820(2001)055[1863:EOIGAE]2.0.CO;2

Biro, P. A. (2012). Do rapid assays predict repeatability in labile (behavioural) traits? *Animal Behaviour*, *83*, 1295–1300. doi: 10.1016/j.anbehav.2012.12.021

Brattstrom, B. H. (1968). Thermal acclimation in anuran amphibians as a function of latitude and altitude. *Comparative Biochemistry and Physiology*, *24*(1), 93–111. doi: 10.1016/0010-406X(68)90961-4

Conover, D. O., & Schultz, E. T. (1995). Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology & Evolution*, *10*(6), 248–252. doi: 10.1016/S0169-5347(00)89081-3

Feder, M. E., & Burggen, W. W. (Eds.). (1992). *Environmental Physiology of the Amphibians*. Chicago, IL: University of Chicago Press.

Ficetola, G. F., & De Bernardi, F. (2006). Trade-off between larval development rate and post-metamorphic traits in the frog *Rana latastei*. *Evolutionary Ecology, 20*(2), 143–158. doi: 10.1007/s10682-005-5508-6

Freidenburg, L. K., & Skelly, D. K. (2004). Microgeographical variation in thermal preference by an amphibian. *Ecology Letters*, *7*(5), 369–373. doi: 10.1111/j.1461-0248.2004.00587.x

Goldstein, J. A., Hoff, K. von S., & Hillyard, S. D. (2017). The effect of temperature on development and behaviour of relict leopard frog tadpoles. *Conservation Physiology*, *5*(1), cow075. doi: 10.1093/conphys/cow075

Gosner, K. L. (1960). A simplified table for staging Anuran embryos and larvae with notes on identification. *Herpetologica*, 16, 183–190. doi: 10.2307/3890061

Gregory, T. R., Ryan Gregory, T., & Wood, C. M. (1998). Individual variation and interrelationships between swimming performance, growth rate, and feeding in juvenile rainbow trout (*Oncorhynchus mykiss*). *Canadian Journal of Fisheries and Aquatic Sciences*, 55(7), 1583–1590. doi: 10.1139/f98-044

Halverson, M. A., Skelly, D. K., Kiesecker, J. M., & Freidenburg, L. K. (2003). Forest mediated light regime linked to amphibian distribution and performance. *Oecologia*, *134*(3), 360–364.

Herreid, C. F., II, & Kinney, S. (1967). Temperature and Development of the Wood Frog, Rana Sylvatica, in Alaska. *Ecology*, *48*(4), 579–590. doi: 10.1007/s00442-002-1136-9

Johansson, F., Lederer, B., & Lind, M. I. (2010). Trait performance correlations across life stages under environmental stress conditions in the common frog, *Rana temporaria*. *PloS One*, *5*(7), e11680. doi: 10.1371/journal.pone.0011680

Johnson, T. P., Bennett, A. F., & McLister, J. D. (1996). Thermal dependence and acclimation of fast start locomotion and its physiological basis in rainbow trout (*Oncorhynchus*

mykiss). *Physiological Zoology*, *69*(2), 276–292. doi: 10.1086/physzool.69.2.30164184

Kingsolver, J. G., & Huey, R. B. (2008). Size, temperature, and fitness: three rules. *Evolutionary Ecology Research*, *10*(2), 251–268.

Kolok, A. S., & Oris, J. T. (1995). The relationship between specific growth rate and swimming performance in male fathead minnows (*Pimephales promelas*). *Canadian Journal of Zoology*, *73*(11), 2165–2167. doi: 10.1139/z95-254

Laugen, A. T., Laurila, A., Räsänen, K., & Merilä, J. (2003). Latitudinal countergradient variation in the common frog (*Rana temporaria*) development rates--evidence for local adaptation. *Journal of Evolutionary Biology*, *16*(5), 996–1005. doi: 10.1046/j.1420-9101.2003.00560.x

Levins, R. (1969). Thermal acclimation and heat resistance in *Drosophila* species. *The American Naturalist*, *103*(933), 483–499. doi: 10.1086/282616

Li, D., Fu, C., Hu, W., Zhong, S., Wang, Y., & Zhu, Z. (2007). Rapid growth cost in "all-fish" growth hormone gene transgenic carp: reduced critical swimming speed. *Chinese Science Bulletin*, *52*,(11), 1501–1506. doi: 10.1007/s11434-007-0217-x

Ligon, N. F., & Skelly, D. K. (2009). Cryptic divergence: countergradient variation in the wood frog. *Evolutionary Ecology Research*, *11*(7), 1099–1109.

Li, Y., Cohen, J. M., & Rohr, J. R. (2013). Review and synthesis of the effects of climate change on amphibians. *Integrative Zoology*. doi: 10.1111/1749-4877.12001

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). Describe and understand your model's parameters. R package. https://github.com/easystats/parameters

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. doi: 10.1111/j.2041-210x.2012.00261.x

Parichy, D. M., & Kaplan, R. H. (1995). Maternal investment and developmental plasticity: functional consequences for locomotor performance of hatchling frog larvae. *Functional Ecology*, *9*(4), 606–617. doi: 10.2307/2390151

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, *37*(1), 637–669. doi: 10.1146/annurev.ecolsys.37.091305.110100

R Core Team (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, J. L., Urban, M. C., Bolnick, D. I., & Skelly, D. K. (2014). Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology & Evolution*, *29*(3), 165–176. doi: 10.1016/j.tree.2014.01.002

Rödin-Mörch, P., Luquet, E., Meyer-Lucht, Y., Richter-Boix, A., Höglund, J., & Laurila, A. (2019). Latitudinal divergence in a wide-spread amphibian: contrasting patterns of

neutral and adaptive genomic variation. *Molecular Ecology*. doi: 10.1111/mec.15132

Seebacher, F., and Grigaltchik, V. S. (2015). Developmental thermal plasticity of prey modifies the impact of predation. *Journal of Experimental Biology, 218,* 1402–1409.

Seebacher, F., & James, R. S. (2008). Plasticity of muscle function in a thermoregulating ectotherm (*Crocodylus porosus*): biomechanics and metabolism. *American Journal of Physiology*, *294*(3), 1024–1032. doi: 10.1152/ajpregu.00755.2007

Skelly, D. K. (2004). Microgeographic countergradient variation in the wood frog, *Rana sylvatica*. *Evolution; International Journal of Organic Evolution*, *58*(1), 160–165. doi: 10.1554/03-425

Skelly, D. K., Bolden, S. R., & Freidenburg, L. K. (2014). Experimental canopy removal enhances diversity of vernal pond amphibians. *Ecological Applications, 24*(2), 340–345. doi: 10.1890/13-1042.1

Stearns, S. C. (1992). The Evolution of Life Histories (2nd ed.). Oxford, OX2: Oxford University Press.

Urban, M. C., Richardson, J. L., Freidenfelds, N. A., Drake, D. L., Fischer, J. F., & Saunders, P. P. (2017). Microgeographic adaptation of wood frog tadpoles to an apex predator. *Copeia, 105*(3), 451–461. doi: 10.1643/CG-16-534

Valladares, F., Sanchez-Gomez, D., & Zavala, M. A. (2006). Quantitative estimation of phenotypic plasticity: bridging the gap between the evolutionary concept and its ecological applications. *Journal of Ecology, 94*(6), 1103-1116). doi: 10.1111/j.1365-2745.2006.01176.x

Van Buskirk, J., & McCollum, S. A. (2000). Influence of tail shape on tadpole swimming performance. *The Journal of Experimental Biology*, *203*(14), 2149–2158.

Watkins, T. B. (2000). The effects of acute and developmental temperature on burst swimming speed and myofibrillar ATPase activity in tadpoles of the Pacific tree frog, *Hyla regilla*. *Physiological and Biochemical Zoology*, *73*(3), 356–364. doi: 10.1086/316744

Watkins, T. B., & Vraspir, J. (2006). Both incubation temperature and posthatching temperature affect swimming performance and morphology of wood frog tadpoles (*Rana sylvatica*). *Physiological and Biochemical Zoology*, *79*(1), 140–149. doi: 10.1086/498182

TABLES

Table 1. Linear mixed effect models of swimming performance (burst speed) for lab-reared and wild-caught tadpoles by development rate, Gosner stage, tadpole shape, size-at-stage, and arena temperature as fixed effects. The model of lab-reared tadpoles, includes nested random intercepts for each pond, clutch, and individual tadpole (n = 887). Burst speed is measured in mm/second. The model of wild-caught tadpoles included nested random intercepts for each pond and individual tadpole (n = 367). Burst speed is measured in mm/second.

	Lab-reared			Wild-caught		
Fixed effects	Estimate	95% CI†	*p*†	Estimate	95% CI†	*p*†
Intercept	-96.659	-151.739, -46.651	< 0.001	48.806	-67.157, 171.715	0.480
Size at stage	83.038	-5.826, 182.536	0.082	422.223	-17.812, 858.335	0.064
Development rate	-45.704	-57.821, -33.573	< 0.001	-63.064	-183.852, 47.149	0.296
Gosner stage	5.167	4.133, 6.283	< 0.001	1.408	-0.503, 3.224	0.130
Shape PC2	72.834	-6.087, 152.037	0.003	36.109	-122.168, 196.574	0.653
Arena temperature	0.235	-2.481, 3.046	0.855	-0.182	-6.533, 6.173	0.995
Size at stage : Development rate	-131.747	-308.008, 27.322	0.128	-786.560	-1678.603, 70.338	0.076
Random effects	SD	95% CI†		SD	95% CI†	
Individual:(Clutch:Pond)	11.22	9.862, 12.399				
Individual: Pond				17.140	14.680, 19.810	
Clutch:Pond	0.00	0.000, 3.573				
Pond	0.00	0.000, 2.637		4.998	0.000, 9.930	
Residual	10.93	10.287, 11.517		10.325	9.348, 11.252	
Marginal R^2	0.237‡			0.080‡		
Conditional R^2	0.629‡			0.769‡		

†*The 95% confidence intervals and* p-*values for main effects were estimated from 1000 bootstrap iterations.*

‡*Marginal and conditional R2 values were estimated with Nakagawa & Schielzeth's (2013) method.*

FIGURES

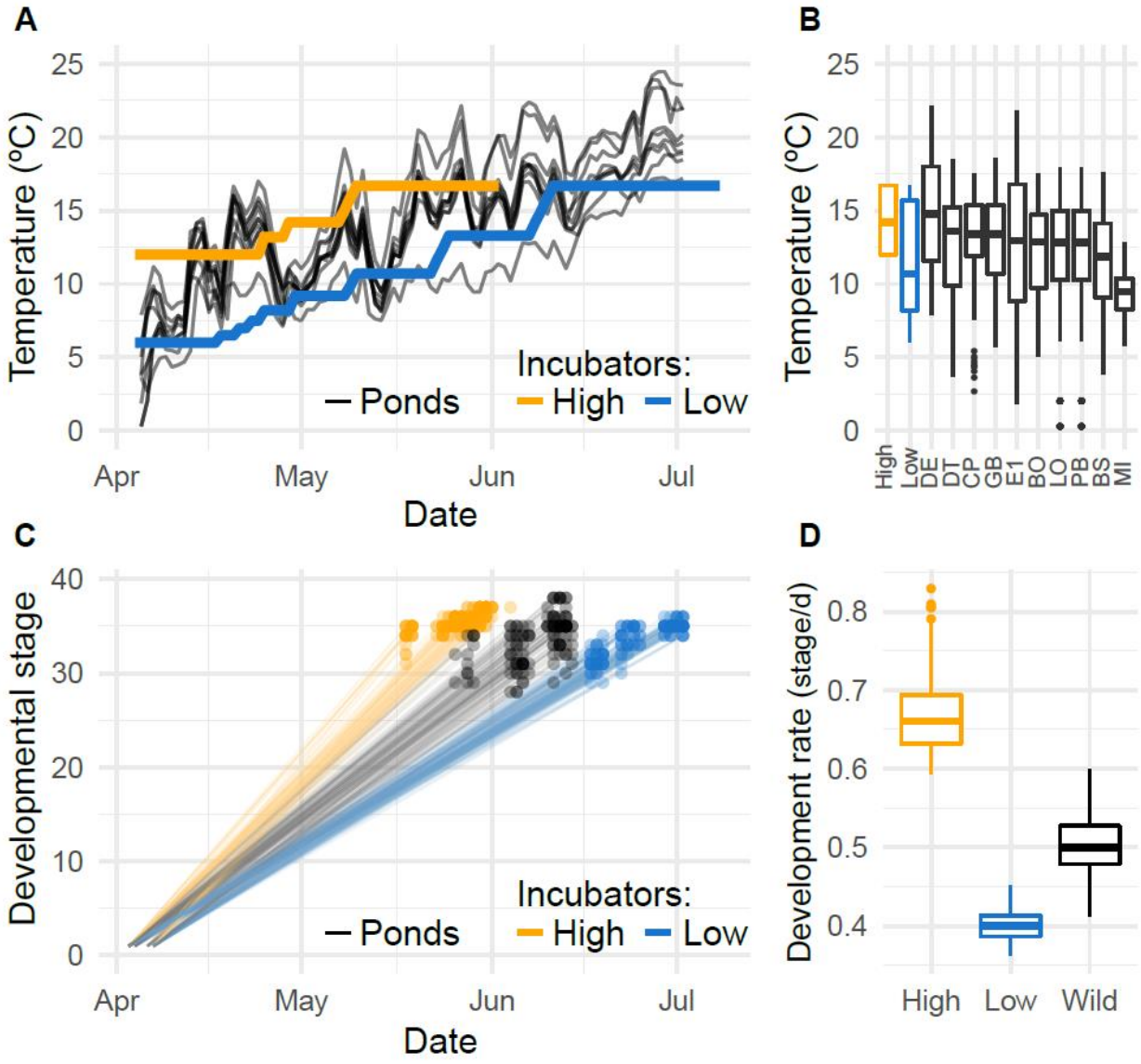

Figure 1. (A, B) Temperatures in incubators and natal ponds during the 2019 season. (C, D) Development rates of warm-treatment, cold-treatment, and wild tadpoles.

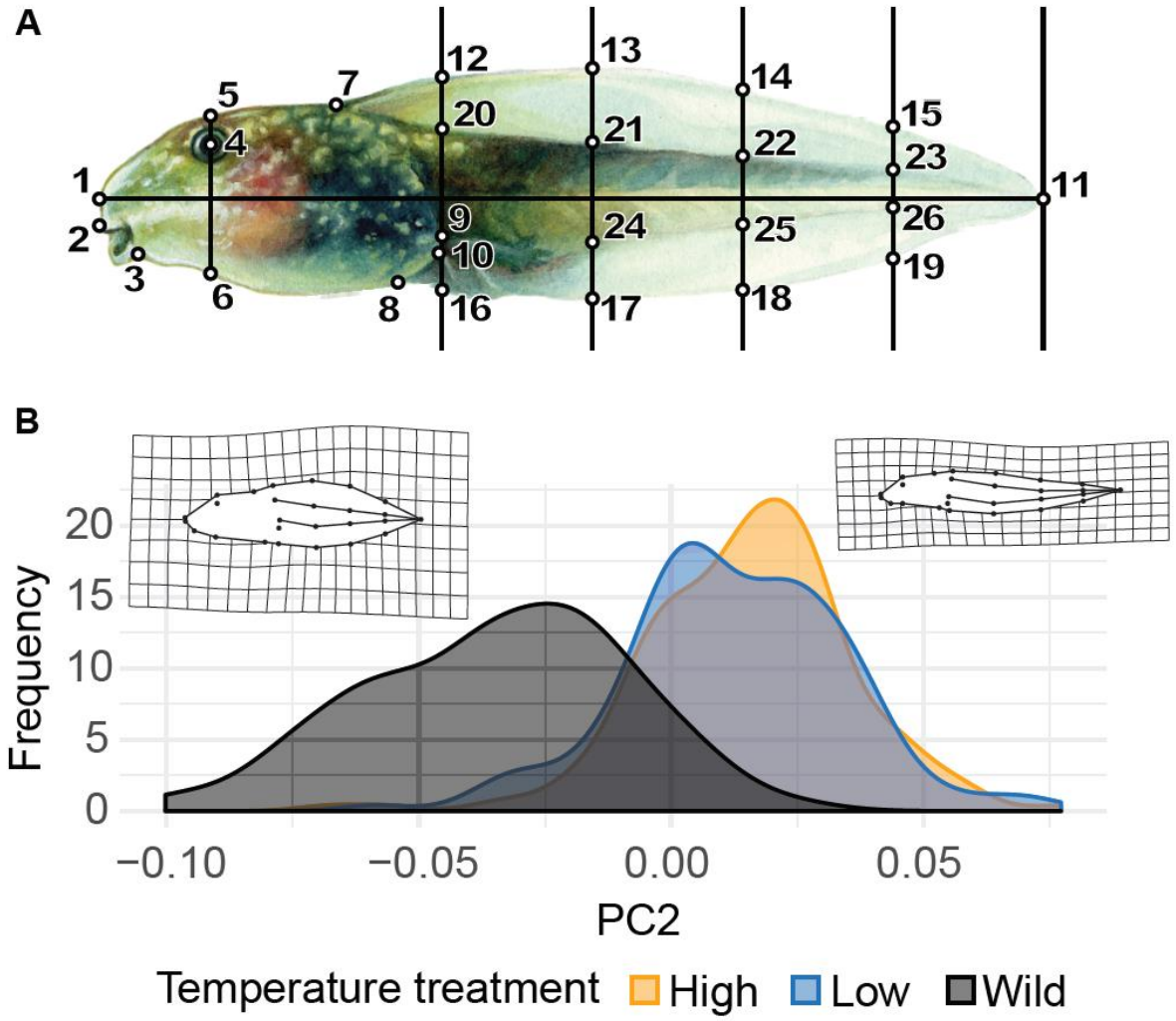

Figure 2. (A) Geometric morphometric landmarks. Landmarks 1-11 were treated as true landmarks because they represented homologous body parts between specimens. Landmarks 12-15 (upper edge of the tail fin), 16-19 (lower edge of the tail fin), 20-23 (upper edge of the tail muscle), and 24-26 (lower edge of the tail muscle), were digitized based on their placement relative to other landmarks (e.g. landmarks 14, 18, 22, and 25 are located along a vertical line at the halfway point between the base and tip of the tail). Therefore, these landmarks were treated as semilandmarks and allowed to "slide" along their respective curves (Adams 2020). (B) Density plot of PC scores for wild and lab-reared tadpoles along the second size-independent principal component constructed from the landmarks shown in (A). Warps represent minimum and maximum scores along the PC2 axis, respectively.

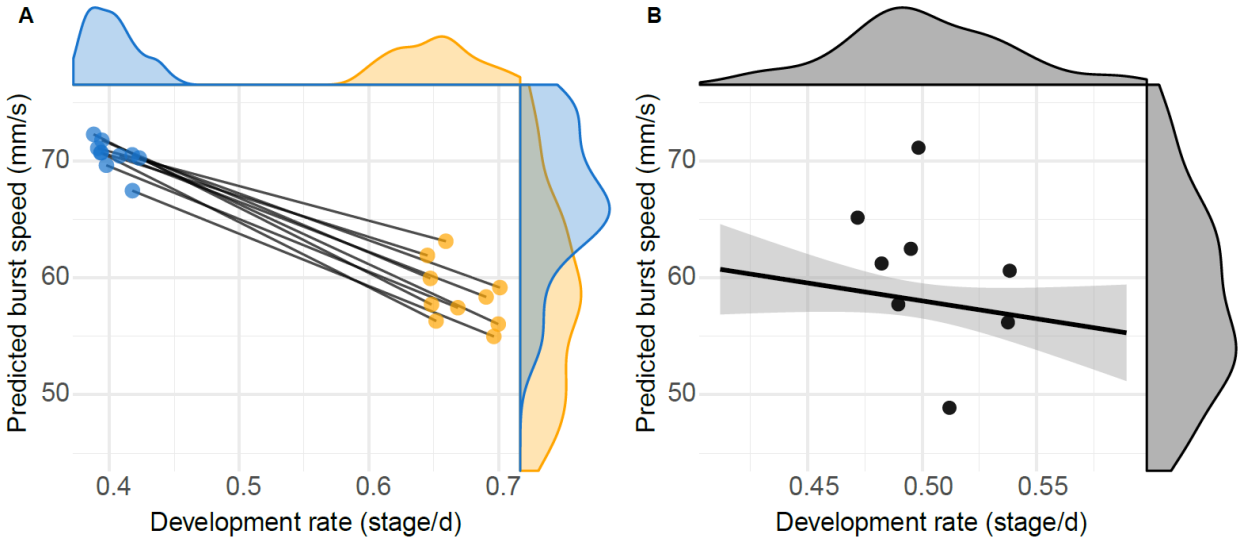

Figure 3. Relationship between development rate and burst speed for (A) lab tadpoles and (B) wild tadpoles. Dots represent pond-wise means, and in (A), lines connect means from the same pond. Marginal density plots are based on individual tadpoles rather than pond-wise means. Orange and blue represent tadpoles reared in the high- and low-temperature incubators, respectively.

CHAPTER 5

Rapid microgeographic evolution in response to climate change

Publication Status: Arietta, A. Z. A., and Skelly, D. K. (2021). Rapid microgeographic evolution in response to climate change. *Evolution* 75, 2930–2943. DOI: 10.1111/evo.14350. (Chapter formatted to journal standards).

ABSTRACT

Environmental change is predicted to accelerate into the future and will exert strong selection pressure on biota. While many species may be fated to extinction, others may survive through their capacity to evolve rapidly at highly localized (i.e. microgeographic) scales. Yet, even as new examples have been discovered, the limits to such evolutionary responses have not often been evaluated. One of the first examples of microgeographic variation involved pond populations of wood frogs (*Rana sylvatica*). Although separated by just tens to hundreds of meters, these populations exhibited countergradient variation in intrinsic embryonic development rates when reared in a common garden. We repeated this experiment 17 years (approx. 6-9 generations) later and found that microgeographic variation persists in contemporary populations. Furthermore, we found that contemporary embryos have evolved to develop 14% to 19% faster than those in 2001. Structural equation models indicate that the predominant cause for this response is likely due to changes in climate over the intervening 17 years. Despite potential for rapid and fine-scale evolution, demographic declines in populations experiencing the greatest changes in climate and habitat imply a limit to the species' ability to mitigate extreme environmental change.

1. INTRODUCTION

The pace of environmental change experienced by contemporary wildlife populations is unprecedented and expected to accelerate (Sala et al., 2000; Urban, 2015). Although climate and habitat change will exert strong selection pressure, we lack an understanding of the rate at which plastic and genetic adaptation can keep pace with environmental change (Gienapp et al., 2008; Hoffmann & Sgrò, 2011; Meester et al., 2018). Currently, most models of extinction risk fail to account for adaptation and fine-grained variation which could drastically alter predicted risks (Urban, 2015) and inform management decisions (Gaitán-Espitia & Hobday, 2020).

Countergradient variation, wherein organism's phenotypes counter environmentally induced effects along a gradient such as temperature, indicates a remarkable ability for traits to be fine-tuned to local conditions even when phenotypes vary relatively little *in situ* (Conover & Schultz, 1995). Common garden experiments that compare populations from across environmental gradients in a common setting have shown that genetic (as opposed to plastic) variation can account for a substantial portion of a population's ability to buffer environmental effects (e.g. Berven et al., 1979; Muir et al., 2014; Orizaola et al., 2010), even at surprisingly small scales (Ficetola & Bernardi, 2005; Richardson et al., 2014; Skelly, 2004). In addition to uncovering evolutionary divergence at small spatial scales, recent research indicates that ecologically relevant evolution can occur over short time scales, as well (Hendry & Kinnison, 1999; Reznick et al., 2019). Although the intersection of fine-grained and rapid adaptation remains largely unexplored, this phenomenon prompts hope in the capacity of natural populations to react to contemporary, often anthropogenic, environmental change (Hoffmann & Sgrò, 2011; Razgour et al., 2019). However, it is unclear if organisms can keep pace with such manifold changes (Carlson et al., 2014; Catullo et al., 2019).

Declining populations may avoid extinction through rapid adaptation—a process termed evolutionary rescue (Gomulkiewicz & Holt, 1995). However, there are limits to a population's ability to adapt (Klausmeier et al., 2020; Meester et al., 2018). Contemporary climates are changing at an unprecedented pace, while at the same time, land-use change alters habitat composition. For evolutionary rescue to occur, adaptive alleles must increase in a population fast enough to avoid falling below a demographic threshold where stochastic processes make extinction likely (Carlson et al., 2014). Populations that are slow to adapt remain below the threshold longer and are more likely fated to extinction (Gomulkiewicz & Holt, 1995). In heterogenous landscapes, influxes of migrants from nearby dissimilar habitats can introduce maladapted alleles and protract the time a population remains below the demographic threshold, as well (Bolnick & Nosil, 2007; Schiffers et al., 2013). Thus, the coincidence of rapid adaptation at microgeographic scales is expected to facilitate evolutionary rescue.

One of the first examples of microgeographic countergradient variation was demonstrated in development rates of wood frogs (*Rana sylvatica*) with respect to a gradient of canopy openness (Skelly, 2004). As in most ectotherms, developmental rates of aquatic wood frog larvae are strongly influenced by water temperatures of their resident ponds. Canopy structure has large impacts on the ecology of temporary wetlands (i.e. vernal pools) in which wood frogs breed (Werner et al., 2007). Canopy closure partially mediates pond water temperature through shading, and can alter hydroperiod via evapotranspiration, limiting the duration of the developmental window of aquatic larvae (Brooks & Hayashi, 2002).

In Skelly's (2004) common garden experiment, wood frogs from dark, ostensibly colder, ponds exhibited faster embryonic development than conspecifics from warm, open ponds separated by tens to hundreds of meters. The interpretation of these results was that

faster intrinsic development confers a fitness advantage in shaded ponds, allowing larvae to advance into the terrestrial life-stage prior to ponds drying up and in relative synchrony with competing warm-pond conspecifics after metamorphosis.

While common garden experiments and other synchronic comparisons (i.e. comparing populations across space at the same time) are powerful tools to infer evolutionary divergence, they cannot provide direct evidence of evolutionary response to climate change, which is inherently a change over time (Kawecki & Ebert, 2004; Merilä & Hendry, 2014). Repeating common-garden studies across time allows for allochronic comparisons (i.e. comparing the same populations at different times) to directly associate changes in the environment with the genetic component of phenotypic change (e.g. Bradshaw & Holzapfel, 2001; Nevo et al., 2012).

In the intervening decades since Skelly's (2004) experiment, ambient annual temperatures experienced by wood frog populations at the site have increased by 0.6 °C (Arietta et al., 2020) which is likely to have similarly warmed the water temperature of natal ponds. The climate has also become more variable and extreme. For example, drought conditions between 2015 and 2017 were some of the most extreme of the last century, second only to the catastrophic drought of the 1960s that caused crop failure and water shortages across the nation (Rowland et al. *in press*). The forest canopies above wood frogs' natal ponds are likely to have changed, too. Disturbances such as blowdowns and forest clearing by beaver or humans can immediately and drastically increase light penetration to vernal pools. Conversely, as the composition of shoreline vegetation changes and individual plants mature, canopies can quickly shade ponds within two decades. Thus, this system offers a unique opportunity to test for rapid, microgeographic evolution to contemporary climate and habitat changes.

In this experiment, we repeat a common garden experiment from 2001 (Skelly, 2004) with as much fidelity as possible to the original experiment (e.g. following the same protocol, using the original lab equipment, testing the same populations) and reanalyze the original data in order to compare to our contemporary results. We first evaluate whether wood frogs maintain microgeographic variation in development rates with respect to canopy or pond temperature after 17 years. Countergradient variation offers somewhat of a paradox for predicting responses to climate change in contexts in which adaptive alleles could be easily shared among all populations in a handful of generations. For microgeographic countergradient variation to persist, there must be not only directional selection for a trait at one end of the gradient, but also selection against that trait at the opposite end. Otherwise, the pattern would erode as populations at one extreme benefited from intrinsic and extrinsic effects in the same direction. In the case of development rates, tradeoffs with performance have been invoked to explain microgeographic countergradient variation (Gahm et al., 2021). Over time, environmental change shifts the range of the gradient experienced by populations, setting more populations up against tradeoff thresholds. Unless further adaptation shifts the threshold for tradeoffs, we would expect microgeographic countergradient patterns to further erode.

Second, we test for environmental change in pond water temperature and canopy closure between experimental timepoints. Third, we test for demographic and evolutionary response of wood frogs to recent environmental change by allochronic comparison of intrinsic embryonic development between experimental cohorts. We use structural equation models to decouple the direct impacts of environmental change and parental effects on changes in developmental timing. We hypothesize that if rapid local adaptation has occurred, the magnitude of trait change in a population should correspond to the magnitude of local environmental change, which varied among ponds.

Finally, we relate population growth or decline to the degree of environmental change. If there are limits to evolutionary rescue, we hypothesize that populations experiencing greater environmental change will exhibit negative population growth.

2. METHODS

2.1 Embryo collection and care

In the spring of 2018, we conducted a common garden experiment following the methods in Skelly (2004). We collected embryos from 13 wood frog breeding ponds at Yale Myers Forest, Connecticut, USA (Fig. 1A) within 24 hours of oviposition between 31 March and 14 April 2018. Nine of these ponds overlap with those studied in the 2001 experiment. Three of the ponds (MB, CC, LT) included in the previous study did not host breeding aggregations in 2018. We added four ponds (DT, E1, LA, WP) not included in 2001 to represent the range of canopy and temperature gradients present across our field site (Fig. 1B). As in 2001, we collected 12 embryos from each of up to 6 clutches, with the exception of "LA" pond, from which we collected embryos from two clutches due to stocking space limitations.

We excised embryos from the egg mass, being careful not to puncture the vitelline membrane, and placed them individually into wells of 6-well culture plates with 15 ml reconstituted distilled water so that each plate contained 6 full-sib embryos. We split clutches across two temperature treatments (high: 13.7°C; low: 11.7°C). Incubator temperatures were set to replicate the original study and represent a large portion of the range in temperatures measured at wood frog oviposition sites. We stocked one plate from each clutch into each light-controlled incubator treatment with a 12:12 H photoperiod centered at 1200 EST. The incubators (I-36VL; Percival Scientific, Inc.) were the same units

used in the original study. Plates were placed randomly in the incubators and rotated daily. In total, we included 888 embryos representing 74 clutches in the experiment.

We monitored embryos daily and estimated hatching as the moment when a larva breaks through the vitelline membrane. Embryos that died or exhibited abnormal development were noted and excluded from further analysis.

2.2 Embryonic period and embryo volume

We photographed embryos upon stocking, twice during larval development, upon hatching, and 3 days post-hatch in order to record developmental stages (Fig. 2B). From these photos, we estimated developmental stage (Gosner, 1960) and size. We estimated initial embryo size as the spherical volume of the mean diameter measured on the x and y axis in ImageJ (v 1.51k (Schneider et al., 2012)). This protocol follows Skelly (2004) with one modification—whereas Skelly (2004) did not include photos taken after hatching, we included photos taken 3 days post-hatch to make up for 123 individuals that were not photographed immediately upon hatching.

Because embryos varied in initial stage and hatching stage, we standardized embryonic period estimates by fitting a logarithmic mixed effect regression model of developmental rate to predict the elapsed time between Gosner stage 1 (fertilization) and 20 (hatching) while correcting for the slight difference in realized treatment temperatures between experimental years (11.3 °C and 13.4 °C in 2001 versus 11.7 °C and 13.7 °C in 2018) (Equation 1). To do so, we estimated a parameter for incubator temperature and held this parameter value constant at the mid-point between years (Low = 11.50 °C, High = 13.55 °C) when predicting embryonic periods. This correction assumes that the reaction norm of temperature on development is linear over the 0.3-0.4 °C interval of difference between experiments (Fig. S8).

We fit our models by regressing the embryonic stage on the interaction between the natural log of days since oviposition (*Day*) and incubator temperature (*IncTemp*) (Equation 1). We considered the collection date as day 1. We included random intercepts for each clutch (*j*) nested within pond (*k*) and random slopes for individuals (*i*) nested within clutch (*j*), respective of treatment group (*l*) that were uncorrelated with the intercept. This specification assumes that all embryos from the same clutch were fertilized at the same time, accounts for repeated measures of individuals, and accounts for nested data structure. Here, our analysis departs from Skelly (2004) who fit logarithmic models of development rate to each embryo, independently. In order to compare datasets, we re-estimated embryonic periods for the 2001 experiment using our model. Pond "MB" was excluded from the 2001 dataset as an outlier following Skelly (2004), because it was assumed that these individuals represented first generation immigrants from nearby ponds.

Equation 1.

$$Stage = \beta_0 + \beta_1 \ln Day + \beta_2 IncTemp + \beta_3 \ln Day \cdot IncTemp + b_{0_k} + b_{1_{jk}} + b_{2_{jkl}} \ln Day + b_{3_{ijkl}} \ln Day + \varepsilon$$

All regression models in this study were fit with lme4 (v 1.1.23 (Bates et al., 2015)) with 95% confidence intervals for model parameters estimated from 1000 nonparametric bootstrap iterations. Marginal and conditional pseudo-R^2 values were estimated with MuMIn (v 1.43.17 (Barton, 2019)).

2.3 Canopy

We estimated canopy closure as the ratio of above-canopy radiation to below-canopy radiation expressed as a percent. This ratio, termed global site factor (GSF) ranges from 0% to 100% (i.e. complete closure to completely open) (Anderson, 1964). We

estimated GSF values from hemispherical photos of the leaf-off canopy (i.e. prior to deciduous leaf emergence) with Gap Light Analyzer (Frazer et al., 1999), by plotting a sun path over the hemispherical images and integrating incident radiation from day-of-year 91 to 135 (approx. April 1 to May 15), as in Skelly (2004). In addition, we estimated GSF over the duration of the larval period for the duration of April 1 through August 30 as the weighted average of the leaf-on and leaf-off estimates with respect to the proportion of days before and after leaf emergence (May 15).

Skelly (2004) estimated canopy closure in 2001 from multiple photos taken at 5 m intervals along a cartesian grid. In 2018, we captured five photos for each pond—four photos 1 m in from the springtime shoreline at each cardinal point and one photo in the center. In order to compare canopy estimates across timepoints, we subsampled the five grid points corresponding to the cardinal point and pond center and recalculated mean GSF. These subsampled estimates tightly correlate with those including all photos from the grid (R^2 = 0.93) (see Supplement 1). We averaged these five estimates for pond-wise values and compared closure estimates for 14 ponds with data from 2001 and 2018, including all nine ponds included in the embryonic development experiments. We used a paired t-test to test for a site-wide change in canopy closure over time.

2.4 Water temperature

Temperature is directly related to developmental rates in amphibians. However, at the time of the 2001 experiment, no long-term temperature data were available to parse the relative impacts of temperature or canopy gradients. Starting in 2001, we recorded water temperatures in a subset of the 60 wood frog breeding ponds at our field site. We recorded water temperature at the deepest point in the pond every 0.5 or 1 h with submersible temperature loggers (HOBO 8K Pendant; Onset Computer Corporation) suspended 10 cm

below the surface. Loggers were deployed within days of oviposition and removed after larvae had metamorphosed or the pond dried. We use these data to estimate long-term water temperature gradients among our ponds.

Because we did not record water temperature in all ponds included in this study in all years, we imputed missing temperature values with a random forest model implemented in randomForest (v4.6.14 (Liaw & Wiener, 2002)) trained on all ponds at our site. Regression trees included both climate and site predictor variables. Climate variables included daily maximum, minimum and average temperature, radiation, precipitation, snowpack, and atmospheric pressure from the Daymet database (v.3 (Thornton et al., 2016)). We included a 1-day lag for all climate variables. Site variables included elevation, aspect, latitude, mean and variance of leaf-on and leaf-off GSF, and a factor for individual ponds (see Supplement 2).

We grew our random forest from 300 regression trees and used 10-fold cross validation to evaluate the predictive accuracy for imputed points. Combining observed and imputed daily temperature values allowed us to calculate annual mean temperatures for the spring, leaf-off period (day-of-year 91 - 135) and approximate larval period (day-of-year 91 - 168). We combined these values for the three years preceding each experiment (i.e. 1998-2000 for 2001 and 2015-2017 for 2018) into a single average temperature value. This timespan captures the average temperatures experienced by the parental cohort and reflects long-term heterogeneity in temperature among ponds without year-to-year variation.

2.5 Countergradient variation

We tested for countergradient variation and estimated the effects of canopy and pond water temperatures on the duration of embryonic period by fitting linear mixed

models with embryonic period as the dependent variable. For each experimental cohort, we fit multiple models including combinations of canopy, pond water temperature, and the interaction. All models included an interaction term to allow the effect along the environmental gradient to vary between incubator temperature treatments and random intercepts for pond of origin. We fit the same models using environmental variables estimated for seasonal window encompassing the embryonic and larval period but results did not differ substantially (Table S5, Fig. S9). Clutch membership was not included as a random effect in these models as the variance among siblings was accounted for in the best linear unbiased predictors of the development rate model used to estimate embryonic periods. We consider a significant and positive relationship between embryonic period and pond temperature or canopy openness as evidence of countergradient variation.

2.6 Temporal comparisons

We combined the embryonic period and embryo volume estimates with the environmental datasets generated in the synchronic analyses into an allochronic dataset to test for evolution in embryonic development across time. Change in canopy and temperature were estimated as the pond-wise difference between 2018 and 2001 values as described above.

We estimated population-wise evolutionary rates in *haldanes* following equations in Kinnison & Hendry (2001). We estimated change in developmental rates by subtracting the individual values of embryonic period in 2001 from pond-wise means in 2018, and the reverse (i.e. individual 2018 traits minus pond-wise 2001 means). We then estimated clutch-wise values as the mean among siblings within each temperature treatment for further analysis.

We built structural equation models (SEM) to test how the magnitude of changes in canopy, temperature, and embryo volume impact the shifts in embryonic period, given an *a priori* causal model (Table S6). Our causal path model was predicated on the hypothesis that changes in canopy, temperature, and/or embryonic volume may have directly induced a response in intrinsic embryonic development. We further hypothesized that the change in embryo volume may be a response to change in canopy and/or temperature. Finally, we hypothesized that change in pond temperature may be a result of changes in canopy. We fit our SEMs as a piecewise model with paths estimated as locally-independent relationships in piecewiseSEM (v. 2.1.0 (Lefcheck, 2016)). We fit an identical SEM for each temperature treatment.

2.7 Limits of evolutionary response

There may be a point at which the pace of environmental change outpaces wood frogs' ability to respond, leading to demographic effects, and ultimately, extirpation of breeding populations. In fact, three ponds ceased to host breeding populations between experimental cohorts. We tested for a correlation between environmental change and demographic decline using annual egg clutch counts. Because each female produces only a single clutch each year, these surveys serve as an excellent index of the size of the female breeding population (Berven, 2009). Briefly, egg mass counts were conducted in all ponds within one week after oviposition. Since wood frogs are explosive breeders, this timing ensures that all oviposition has ceased but egg masses are not too swollen to identify distinct clutches. Ponds were completely searched by two observers whose independent counts were then averaged (see detailed methods in Arietta et al., 2020).

We assessed the relationship between environmental change and population growth in a mixed model framework. We regressed logarithmic population growth curves—

annual egg mass counts (scaled to population size) against the log-transformed year—allowing the slopes to vary with total environmental change via an interaction parameter. We computed a single metric of environmental change by dividing pond water temperature and canopy measures by their respective standard deviations and summing their absolute values. We modeled the variance in growth curves among pond as random effects and assessed significance of parameter estimates with 1000 non-parametric bootstraps. Thus, a significant parameter estimate for the interaction between environmental change and the log-transformed year would indicate a relationship between the magnitude of environmental change and the rate of population growth or decline.

3. RESULTS

3.1 Embryonic development and size

In total, we stocked 888 embryos into the 2018 experiment. Mortality was low (n = 31) with 96% survival in the high temperature treatment and 97% survival in the low treatment. We removed 46 (5%) observations from the dataset due to deformities or irregularities during development and six due to missing data (1%). In total, we retained 805 embryos (91%) in the analysis (high: n = 405, low: n = 400).

Of the 780 embryos included in the 2001 experiment, 59 (8%) were excluded because the embryo did not survive to hatch (n = 58), or due to missing data (n = 1). Embryos from "MB" pond (n = 47) were determined to be outliers by Skelly (2004), and therefore, were not included in the countergradient analysis. "MB" pond did not host a breeding population in 2018; and so, does not pertain to the allochronic analysis. In total, 674 embryos (86%) were included in the analysis (high: n = 339, low: n = 335).

Upon stocking the 2018 experiment, embryos ranged in stage from GS 1 to GS 10 (mean = GS 5.5), similar to the initial stages of embryos in the 2001 experiment (mean = 7.8,

range = 3 to 11). Embryos were larger on average in 2018 (mean = 5.27 ml, sd = 1.0 ml) than in 2001 (mean = 4.2 ml, sd = 0.7 ml) ($p < 0.001$, Table S2). The models predicting embryonic period from repeated measures of developmental stage fit very well for both experimental years (Year [mR^2, cR^2]: 2001 [0.82, 0.98], 2018 [0.87, 0.98]) (Table 1, Fig. S7).

Our models of development estimate that an average embryo from the 2018 experiment attained hatching stage in 10.70 days in the low temperature treatment and 7.47 days in the high temperature treatment (Table S3). In comparison, the average embryonic period in the 2001 cohort was 12.46 days and 9.07 days, respectively (Table S3). Embryonic development was faster, and perforce, embryonic periods shorter in 2018 than 2001 by 1.7 days (-14.2%) for the low temperature treatment and 1.6 days (-17.6%) for the high temperature treatment (low: $p < 0.001$, high: $p < 0.001$; Table S3).

3.2 Pond temperature and canopy

Our random forest model accounted for 87% of variance in daily pond temperatures with predictive accuracy within +/- 0.54 °C of daily temperature. Across all ponds at our field site, pond water temperatures averaged an increase of 0.22 C (95% C.I. = 0.05 to 0.40) since 1999, in concordance with a 0.46 °C rise in air temperature for the same seasonal window. However, temperature change during the embryonic period exhibited increases and decreases among populations, with extremes of similar magnitude (range = -0.33 °C to 0.38 °C) (Fig. 1C, Table S1). On average, canopies have become more closed, and consequently, ponds received 9.2% less light on average during the embryonic period in 2018 than in 2001 (range = -36.2 to 0.59) (Fig. 1C, Table S1). While all ponds experienced positive or no change in canopy, there was no correlation between canopy change and temperature change during the embryonic period.

3.3 Countergradient variation

The best fit model predicting embryonic period for the 2018 dataset included the interactive effects of pond temperature and canopy and fit much better than the next best model (ΔAIC = 39.6) (Table S4). The best fit model for the 2001 dataset included canopy but not temperature (Table S4); however, this model had only slightly better fit than the model with both environmental variables and the interaction (ΔAIC = 0.5). Thus, we report the results of the full model to be consistent with the 2018 results.

For experiments in both years, the effects of pond temperature and canopy were significant for embryos reared in the low temperature treatment but not the high temperature treatment (Table 2). This may be due to the increased range of embryonic periods resulting from protracted development in the low temperature treatment yielding higher power to discern differences. Here, we interpret the response among the low temperature treatment.

Among the 2018 cohort, holding canopy constant, we estimated embryonic periods increased by 0.32 day for ponds that are 1 °C warmer, in a countergradient fashion (Fig. 2C). Functionally, this difference in intrinsic development rate predicts about a 1 day difference in embryonic period between the warmest and coldest ponds in our experiment. The effect of temperature is stronger in darker ponds. Among 2001 populations, we estimated embryonic periods to be longer by 0.19 day for 1 °C warmer ponds (Fig. 2E).

Embryonic period, holding pond temperature constant, showed countergradient variation with respect to canopy in 2001 (Fig. 2F), but this relationship reversed in the 2018 experiment (Fig. 2D). For pond populations in 2018, we estimated embryonic periods to be 0.10 day shorter in ponds with 10% more canopy openness, but 0.30 day longer for the 2001 cohort (Table 2).

3.4 Temporal comparisons

Considering only the nine populations represented in both the 2001 and 2018 experiments, all populations exhibited faster development and reduced embryonic periods in both temperature treatments during the recent experiment (Fig. 3B). Compared to 2001, population-wise average embryonic periods evolved at a rate of -0.27 haldanes (range = -0.44 — -0.15) in the high temperature treatment and -0.15 haldanes (range = -0.38 — -0.03) in the low treatment, assuming a generation time of two years. All but one pond (BS) exhibited an increase in embryonic volume over time (Fig 3A) with an average increase of 1.05 ml (25%).

Our SEMs fit reasonably well, accounting for 54-57% of the variance in the difference in embryonic period across years (Table 3, Fig. 3C). For both temperature treatments, the greater the increase in pond temperature, the greater the increase in embryonic period across years. Thus, all ponds exhibited faster development over time, but those populations that experienced the least amount of warming exhibited the most negative change in embryonic periods (i.e. greatest acceleration in developmental rates), relatively. Increases in embryonic volume elicited decreases in embryonic period, although this relationship was not significant for the high temperature treatment. There was no evidence that changes in canopy caused changes in pond temperature. Changes in canopy did not have a measurable effect on embryonic period directly nor mediated through the effect on embryonic volume.

3.5 Limits of evolutionary response

There was a negative relationship between population growth and the pace of environmental change of natal ponds (Table 4, Fig. 4B). We estimate that, in general, populations that experience less than 0.87 standard deviations of change in temperature or

canopy (or a combination thereof) exhibit stable or growing populations while environmental change greater than that threshold leads to population declines (Fig. 4A). The three populations that went extinct prior to 2018 (MB, LT, and CC) experienced the greatest magnitude of environmental change (Fig. 1C) and some of the steepest population declines (Fig. 1D).

4. DISCUSSION

A pattern of microgeographic variation first detected in 2001 persists in the same populations nearly two decades later. The results of our experiment affirm that evolution takes place over spatial scales easily traversed by an individual of our study species and that populations are evolving over time. While Skelly (2004) was one of the first to document such microgeographic patterns, the phenomenon has since been observed in other systems (Richardson et al., 2014) and, owing to the cryptic nature of countergradient variation, is likely to be uncovered in still more as further field research is carried out. Within the context of climate change, our fundamental finding that a species is capable of rapid, fine-scale evolution in response to thermal variation should provide a reason to expect that adaptation to changing climate is possible and may insulate species from some of its consequences. While that is a reasonable interpretation, we also have evidence to conclude that there are limits to such adaptation and that the failure of evolutionary rescue may be linked to the extinction of some populations (Fig. 4) (Klausmeier et al., 2020; Meester et al., 2018).

Despite the complexity of environmental change taking place between 2001 and 2018 (canopy closure, increased water temperature, period of intense drought), the relationship between embryonic period and temperature in our experiment conformed to expectations based on the pattern of countergradient variation seen in 2001. Most critically,

embryos from warmer ponds took relatively longer to develop, while embryos from colder ponds developed relatively more rapidly. However, in absolute terms, embryonic periods decreased overall from 2001 to 2018 while temperatures across the site increased. This result violates both expectations from space-for-time inference and expectations that selection would erode spatial gradients over time as selection drives populations at the maladapted end of the gradient to extinction.

One potential explanation for this counterintuitive difference between spatial and temporal gradients is that while accelerated development may be advantageous in general, there may be trade-offs when organisms develop too quickly (Gahm et al., 2021). The microgeographic pattern uncovered in 2001 showed that populations in the warmest ponds exhibited depressed intrinsic development rates relative to those in colder ponds. Given warming and canopy change over time, selection may have favored faster development across the metapopulation as a whole, leading to shorter embryonic periods overall. However, those population that faced large increases in temperature more quickly may have approached the threshold at which accelerated development results in performance costs, limiting their ability to adapt and attenuating the resultant shift in embryonic development across timepoints.

The potential for populations to circumvent extinction through evolutionary rescue has become ever more salient in the face of anthropogenic climate change. Despite the general focus on climate, our results suggest that landscape change can interact with climate change in important ways. Since 2001, ponds tended to become darker as the surrounding canopy matured. The effect of shading did not have a straightforward effect on spring pond temperature, perhaps because other factors influence water temperature, such as ground water and snowmelt runoff. In fact, the direction of the relationship between canopy closure and embryonic development reversed between experimental years. While

embryos developed more quickly in darker ponds in 2001, embryos from more open ponds tended to develop most quickly in 2018. Notably, our SEMs did not indicate that changes in canopy resulted in changes in embryonic period directly nor mediated through change in egg volume. Three ponds populations that failed prior to 2018 exhibited some of the largest shifts in canopy, but consequently were not included in the SEM dataset. The compound effects of canopy closure and increasing air temperatures may impact wood frog development indirectly, however, by altering the hydroperiod of breeding ponds.

Both higher air temperature and denser canopies lead to increased evaporation and evapotranspiration of vernal pools (Brooks & Hayashi, 2002), especially after deciduous leaves emerge. The shift to warmer air and pond water temperature and darker canopies is much greater for the larval period than the spring embryonic period (Arietta et al. 2020 and supplemental figures). The net result is likely a decrease in hydroperiod.

The aquatic developmental period of wood frogs is limited to the ice-free hydroperiod of vernal pools. Oviposition occurs shortly after the surface ice melts and larvae must metamorphose prior to the pond drying out. While climate warming allows some species to maintain the developmental window by commencing breeding earlier in the spring, this is not possible at our site during at least some years because more persistent snowpack counterintuitively delays oviposition (Arietta et al., 2020). Thus, constricting hydroperiods would exert extreme selective pressure for faster development to avoid mass mortality events when ponds desiccate prior to metamorphosis. There is considerable experimental evidence of plastic and genetic effects on developmental timing in response to hydroperiod (Lent & Babbitt, 2020; Richter-Boix et al., 2011).

In addition to shifts in embryonic period, embryonic volume has also shifted over time. Across amphibians as a group, larger embryos are associated with faster development (Bradford, 1990). We see this same relationship—within a given cohort, larger embryos

tend to develop more quickly (i.e. shorter embryonic periods), and between cohorts, the increase in embryo size is associated with a decrease in embryonic periods.

The shift to larger embryos partially counteracts the direct effect of increases in temperature on embryonic periods in our populations. Interestingly, the SEMs did not indicate that changes in embryo size were a result of shifts in pond temperature nor canopy. Nevertheless, climate may have played a role in shifting embryonic size. The interim period between our experiments, saw some of the most extreme drought conditions in the past century. Droughts may have selected for larger, older females which are better able to cope with water balance due to smaller surface area-to-volume ratio and lower area-specific evaporative water loss rate (Claussen, 1969). In ranid frogs, there is a strong, positive correspondence between the size and age of females and larger embryos (Berven, 1988).

Although we see evidence that populations exhibit putatively adaptive shifts in intrinsic embryonic development rate, it is notable that three populations that did not persist between experimental years (MB, CC, and LT) also experienced the greatest changes in canopy and temperature (Fig. 4). For both the embryonic and larval period, "LT" saw the greatest increase in canopy closure, but moderate change in temperature. "MB" and "CC" saw the greatest increase and decrease in temperature, respectively, but moderate canopy closure. "DE", a pond with comparable change to "CC", has seen declining populations since 2001. Yet, these ponds are not outliers in temperature or canopy. It will be worth exploring further whether it is the pace of changes in conditions, and not the conditions *per se*, that determine population persistence and the extent to which microevolutionary adaptation can rescue populations. Current methods of predicting species range shifts in the future focus almost entirely on changes in habitat suitability but not the pace at which it occurs (Wiens et al., 2009). Thus, this is a critical question for conservation management and reveals the importance of tracking evolutionary responses over time.

Despite the power of allochronic common garden studies to demonstrate responses to contemporary climate change (Hendry & Kinnison, 1999), very few researchers have undertaken such experiments to date (Merilä & Hendry, 2014). One reason is the difficulty of exactly replicating experimental conditions between timepoints. Despite using the same incubators in both experimental years, the daily variation in rearing temperatures was lower in 2018 than in 2001 and the mean temperature was slightly higher. Embryonic development in ectotherms is influenced by both mean temperature and fluctuation in temperature in non-additive ways (Massey & Hutchings, 2021). We corrected for the shift in mean temperature analytically, but could not correct for variance. At moderate temperatures like those used in our experiment both theory (Georges et al., 2005; Liu & Meng, 2000; Massey & Hutchings, 2021) and empirical findings (e.g. Arrighi et al., 2013; Hall & Warner, 2020; Niehaus et al., 2006) suggest that temperature fluctuations have either no effect or cause accelerated developmental rates. In our experiments we find the opposite, embryos in 2001 experience more thermal variability yet slower development that embryos in 2018, making it unlikely that the differences we see are a result of experimental conditions.

A second reason replication is rarely attempted is the inherent difficulty in defining any field ecology study conducted over time as truly 'direct' replication, given that all ecology takes place on a non-replicated planet (Filazzola & Cahill, 2021). In our case, this same problem means that it is impossible to conclusively state that allochronic difference are the result of evolution and not artifacts. However, our study approached direct replication as closely as possible and multiple lines of evidence point to interesting biological change.

The extent to which the pace and scale of microevolution will play in conservation and management of biodiversity in the future is an open question (Stockwell et al., 2003).

Answering it will require ongoing theory and empirical studies of evolutionary responses to climate change. Vernal pond species are particularly vulnerable to climate change because their developmental window is time-constrained by developmental rates which are closely tied to water temperatures. Unlike other organisms, vernal pool species cannot simply migrate upstream or down the water column to cooler waters and are relatively isolated over very small spatial scales. Thus, adaptation is the only recourse, making vernal pool species excellent models to understand the limits of adaptation to climate change. This study joins a growing body of research showing that contemporary evolution is common and often rapid relative to ecological processes. While there are limits on the extent to which evolution and plasticity can buffer climate change (Radchuk et al., 2019), this study suggests a potential role for evolutionary rescue in circumventing biodiversity loss in some cases.

Author Contributions – AA conducted the experiments and analysis. AA and DS conceived of the study and wrote the manuscript.

Acknowledgements – We would like to thank Adriana Rubinstein, Dr. L. Kealoha Freidenburg, and Greg Watkins-Colwell for their generous help with lab and field work. Thanks to Dr. Freya Rowland, Dr. Nate Edleman, Yara Alshwairikh, Dahn-Young Dong, Logan Billet, and Dr. Annise Dobson for early reviews of this manuscript.

Funding for this study was provided by Yale Institute for Biospheric Studies and Yale Forests Kohlberg-Donohoe Research Fellowship.

Animal handling and collections were approved by Yale University IACUC (2004-10361, 2016-10361, 2019-10361, 2006-11024, 2009-11024, 2009-11040) and Connecticut

DEEP (205001, 211001b, 0112019d). Field work was conducted with permission from Yale Myers Research Committee (SKE01, AND17).

Data Accessibility Statement – The data that support the findings of this study are openly available in Dryad at https://doi.org/10.5061/dryad.08kprr53b.

REFERENCES

Anderson, M. C. (1964). Studies of the woodland light climate I. The photographic computation of light condition. *Journal of Ecology*, *52*, 27–41.

Arietta, A. Z. A. & Skelly, D. K. (2021). Data from: Rapid microgeographic evolution in response to climate change. – Dryad Digital Repository, <https://doi.org/10.5061/dryad.08kprr53b>.

Arietta, A. Z. A., Freidenburg, L. K., Urban, M. C., Rodrigues, S. B., Rubinstein, A., & Skelly, D. K. (2020). Phenological delay despite warming in wood frog *Rana sylvatica* reproductive timing: a 20-year study. *Ecography*, *52*, 27.

Arrighi, J. M., Lencer, E. S., Jukar, A., Park, D., Phillips, P. C., & Kaplan, R. H. (2013). Daily temperature fluctuations unpredictably influence developmental rate and morphology at a critical early larval stage in a frog. *BMC Ecology*, *13*, 18.

Barton, K. (2019). *MuMIn: Multi-Model Inference*. https://CRAN.R-project.org/package=MuMIn

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, Articles*, *67*(1), 1–48.

Berven, K. A. (1988). Factors Affecting Variation in Reproductive Traits within a Population of Wood Frogs (Rana sylvatica). *Copeia*, *1988*(3), 605–615.

Berven, K. A. (2009). Density Dependence in the Terrestrial Stage of Wood Frogs: Evidence from a 21-Year Population Study. In *Copeia* (Vol. 2009, Issue 2, pp. 328–338). https://doi.org/10.1643/ch-08-052

Berven, K. A., Gill, D. E., & Smith-Gill, S. J. (1979). Countergradient Selection in the Green Frog, Rana clamitans. *Evolution; International Journal of Organic Evolution*, *33*(2), 609–623.

Bolnick, D. I., & Nosil, P. (2007). Natural selection in populations subject to a migration load. *Evolution; International Journal of Organic Evolution*, *61*(9), 2229–2243.

Bradford, D. F. (1990). Incubation Time and Rate of Embryonic Development in

Amphibians: The Influence of Ovum Size, Temperature, and Reproductive Mode. *Physiological Zoology*, *63*(6), 1157–1180.

Bradshaw, W. E., & Holzapfel, C. M. (2001). Genetic shift in photoperiodic response correlated with global warming. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(25), 14509–14511.

Brooks, R. T., & Hayashi, M. (2002). Depth-area-volume and hydroperiod relationships of ephemeral (vernal) forest pools in southern New England. *Wetlands. 22(2): 247-255.*, *22*(2). https://link.springer.com/article/10.1672/0277-5212(2002)022[0247:DAVAHR]2.0.CO;2

Carlson, S. M., Cunningham, C. J., & Westley, P. A. H. (2014). Evolutionary rescue in a changing world. *Trends in Ecology & Evolution*, *29*(9), 521–530.

Catullo, R. A., Llewelyn, J., Phillips, B. L., & Moritz, C. C. (2019). The Potential for Rapid Evolution under Anthropogenic Climate Change. *Current Biology: CB*, *29*(19), R996–R1007.

Claussen, D. L. (1969). Studies on water loss and rehydration in anurans. *Physiological Zoology*, *42*(1), 1–14.

Conover, D. O., & Schultz, E. T. (1995). Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology & Evolution*, *10*(6), 248–252.

Ficetola, G. F., & Bernardi, F. (2005). Supplementation or in situ conservation? Evidence of local adaptation in the Italian agile frog Rana latastei and consequences for the management of populations. *Animal Conservation*, *8*(1), 33–40.

Filazzola, A., & Cahill Jr., J. F. (2021). Replication in field biology: Identifying challenges and proposing solutions. *Methods in Ecology and Evolution, 00*, 1–13.

Frazer, G. W., Canham, C. D., & Lertzman, K. P. (1999). *Gap Light Analyzer (GLA): Imaging software to extract canopy structure and gap light transmission indices from true-color fisheye photographs, users manual and documentation* (Version 2.0). Simon Fraser University. http://rem-main.rem.sfu.ca/downloads/Forestry/GLAV2UsersManual.pdf

Gahm, K., Arietta, A. Z. A., & Skelly, D. K. (2021). Temperature-mediated trade-off between development and performance in larval wood frogs (Rana sylvatica). *Journal of Experimental Zoology. Part A, Ecological and Integrative Physiology*, *335*(1), 146–157.

Gaitán-Espitia, J. D., & Hobday, A. J. (2020). Evolutionary principles and genetic considerations for guiding conservation interventions under climate change. *Global Change Biology*. https://doi.org/10.1111/gcb.15359

Georges, A., Beggs, K., Young, J. E., & Doody, J. S. (2005). Modelling development of reptile embryos under fluctuating temperature regimes. *Physiological and Biochemical Zoology: PBZ*, *78*(1), 18–30.

Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A., & Merilä, J. (2008). Climate change and

evolution: disentangling environmental and genetic responses. *Molecular Ecology*, *17*(1), 167–178.

Gomulkiewicz, R., & Holt, R. D. (1995). When does evolution by natural selection prevent extinction? *Evolution; International Journal of Organic Evolution*, *49*(1), 201–207.

Gosner, K. L. (1960). A simplified table for staging Anuran embryos and larvae with notes on identification. *Herpetologica*, *16*(3), 183–190.

Hall, J. M., & Warner, D. A. (2020). Ecologically relevant thermal fluctuations enhance offspring fitness: biological and methodological implications for studies of thermal developmental plasticity. *The Journal of Experimental Biology*, *223*(Pt 19). https://doi.org/10.1242/jeb.231902

Hendry, A. P., & Kinnison, M. T. (1999). Perspective: The pace of modern fife: measuring rates of contemporary microevolution. *Evolution; International Journal of Organic Evolution*, *53*(6), 1637–1653.

Hoffmann, A. A., & Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, *470*(7335), 479–485.

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*(12), 1225–1241.

Kinnison, M. T., & Hendry, A. P. (2001). The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica*, *112-113*, 145–164.

Klausmeier, C. A., Osmond, M. M., Kremer, C. T., & Litchman, E. (2020). Ecological limits to evolutionary rescue. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *375*(1814), 20190453.

Lefcheck, J. S. (2016). piecewiseSEM : Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Methods in Ecology and Evolution / British Ecological Society*, *7*(5), 573–579.

Lent, E., & Babbitt, K. J. (2020). The effects of hydroperiod and predator density on growth, development, and morphology of wood frogs (Rana sylvatica). *Aquatic Ecology*, *54*, 369–386.

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. In *R News* (Vol. 2, Issue 3, pp. 18–22). https://CRAN.R-project.org/doc/Rnews/

Liu, S. S., & Meng, X. D. (2000). Modelling development time of Lipaphis erysimi (Hemiptera: Aphididae) at constant and variable temperatures. *Bulletin of Entomological Research*, *90*(4), 337–347.

Massey, M. D., & Hutchings, J. A. (2021). Thermal variability during ectotherm egg incubation: A synthesis and framework. *Journal of Experimental Zoology. Part A, Ecological and Integrative Physiology*, *335*(1), 59–71.

Meester, L. D., Stoks, R., & Brans, K. I. (2018). Genetic adaptation as a biological buffer

against climate change: Potential and limitations. *Integrative Zoology*, *13*(4), 372–391.

Merilä, J., & Hendry, A. P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications*, *7*(1), 1–14.

Muir, A. P., Biek, R., Thomas, R., & Mable, B. K. (2014). Local adaptation with high gene flow: temperature parameters drive adaptation to altitude in the common frog (Rana temporaria). *Molecular Ecology*, *23*(3), 561–574.

Nevo, E., Fu, Y.-B., Pavlicek, T., Khalifa, S., Tavasi, M., & Beiles, A. (2012). Evolution of wild cereals during 28 years of global warming in Israel. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(9), 3412–3415.

Niehaus, A. C., Wilson, R. S., & Franklin, C. E. (2006). Short- and long-term consequences of thermal variation in the larval environment of anurans. *The Journal of Animal Ecology*, *75*(3), 686–692.

Orizaola, G., Quintela, M., & Laurila, A. (2010). Climatic adaptation in an isolated and genetically impoverished amphibian population. *Ecography*, *33*(4), 730–737.

Radchuk, V., Reed, T., Teplitsky, C., van de Pol, M., Charmantier, A., Hassall, C., Adamík, P., Adriaensen, F., Ahola, M. P., Arcese, P., Miguel Avilés, J., Balbontin, J., Berg, K. S., Borras, A., Burthe, S., Clobert, J., Dehnhard, N., de Lope, F., Dhondt, A. A., … Kramer-Schadt, S. (2019). Adaptive responses of animals to climate change are most likely insufficient. *Nature Communications*, *10*(1), 3109.

Razgour, O., Forester, B., Taggart, J. B., Bekaert, M., Juste, J., Ibáñez, C., Puechmaille, S. J., Novella-Fernandez, R., Alberdi, A., & Manel, S. (2019). Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(21), 10418–10423.

Reznick, D. N., Losos, J., & Travis, J. (2019). From low to high gear: there has been a paradigm shift in our understanding of evolution. *Ecology Letters*, *22*(2), 233–244.

Richardson, J. L., Urban, M. C., Bolnick, D. I., & Skelly, D. K. (2014). Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology & Evolution*, *29*(3), 165–176.

Richter-Boix, A., Tejedo, M., & Rezende, E. L. (2011). Evolution and plasticity of anuran larval development in response to desiccation. A comparative analysis. *Ecology and Evolution*, *1*(1), 15–25.

Sala, O. E., Chapin, F. S., 3rd, Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L. F., Jackson, R. B., Kinzig, A., Leemans, R., Lodge, D. M., Mooney, H. A., Oesterheld, M., Poff, N. L., Sykes, M. T., Walker, B. H., Walker, M., & Wall, D. H. (2000). Global biodiversity scenarios for the year 2100. *Science*, *287*(5459), 1770–1774.

Schiffers, K., Bourne, E. C., Lavergne, S., Thuiller, W., & Travis, J. M. J. (2013). Limited evolutionary rescue of locally adapted populations facing climate change. *Philosophical*

Transactions of the Royal Society of London. Series B, Biological Sciences, *368*(1610), 20120083.

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675.

Skelly, D. K. (2004). Microgeographic countergradient variation in the wood frog, *Rana sylvatica*. *Evolution; International Journal of Organic Evolution*, *58*(1), 160–165.

Stockwell, C. A., Hendry, A. P., & Kinnison, M. T. (2003). Contemporary evolution meets conservation biology. *Trends in Ecology & Evolution*, *18*(2), 94–101.

Thornton, P. E., Thornton, M. M., Mayer, B. W., Wei, Y., Devarakonda, R., Vose, R. S., & Cook, R. B. (2016). *Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3*. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/1328

Urban, M. C. (2015). Climate change. Accelerating extinction risk from climate change. *Science*, *348*(6234), 571–573.

Werner, E. E., Skelly, D. K., Relyea, R. A., & Yurewicz, K. L. (2007). Amphibian species richness across environmental gradients. *Oikos* , *116*(10), 1697–1712.

Wiens, J. A., Stralberg, D., Jongsomjit, D., Howell, C. A., & Snyder, M. A. (2009). Niches, models, and climate change: assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences of the United States of America*, *106 Suppl 2*, 19729–19736.

TABLES

Table 1. Results of a mixed-effect regression of embryonic development with random intercepts for clutch-mates and random slopes for individuals nested within clutch fit to data from the 2018 experiment ($N_{observation}$ = 3186, $N_{individual}$ = 805, N_{clutch} = 146, N_{pond} = 13) and 2001 experiment ($N_{observation}$ = 3971, $N_{individual}$ = 721, N_{clutch} = 128, N_{pond} = 12). Fixed effect predictor variables include the natural log transformed days since stocking (ln(Day)), the realized incubator temperature in °C (IncTemp), and the interaction. These models were used to predict the duration of embryonic periods from stage 1 to 20. Confidence intervals represent the 95 percentile values from 1000 non-parametric bootstraps.

	Stage (2018)			**Stage (2001)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
Intercept	4.04	2.79 – 5.31	**<0.001**	3.59	2.71 – 4.56	**<0.001**
ln(Day)	1.32	-0.35 – 2.91	0.13	2.84	1.63 – 4.01	**<0.001**
IncTemp	0.15	0.08 – 0.21	**<0.001**	0.32	0.26 – 0.37	**<0.001**
ln(Day) * IncTemp	0.40	0.28 – 0.53	**<0.001**	0.19	0.10 – 0.29	**<0.001**
Random Effects				**Random Effects**		
σ^2	1.13			0.60		
$b_{0\ Pond}$	2.27			1.32		
$b_{1\ Clutch:Pond}$	1.90			1.06		
$b_{3\ Inc:Pond:Clutch*lnDay}$	0.58			0.35		
$b_{4\ Inc:Pond:Clutch:ID*lnDay}$	0.00			0.00		
Marginal R^2 / Conditional R^2	0.87 / 0.98			0.825 / 0.976		

Table 2. Model results predicting embryonic period from spring pond temperatures (PondTemp), leaf-off canopy (Canopy), and treatment group ($Treatment_{Low}$, the High treatment is the baseline category) for the 2018 ($N_{observation}$ = 805, N_{pond} = 13) and 2001 ($N_{observation}$ = 674, N_{pond} = 11) experimental cohorts. Confidence intervals represent the 95 percentile values from 1000 non-parametric bootstraps.

	EP (2018)			**EP (2001)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	6.64	-3.21 – 16.66	0.20	-2.13	-42.27 – 39.58	0.96
$Treatment_{Low}$	**-11.37**	**-15.62 – -7.21**	**<<0.01**	**-9.72**	**-20.75 – -0.11**	**0.04**
PondTemp	0.04	-0.87 – 0.95	0.92	0.97	-3.00 – 4.57	0.62
Canopy	-0.02	-0.19 – 0.15	0.84	0.14	-0.43 – 0.69	0.64
$Treatment_{Low}$ * PondTemp	**1.46**	**1.08 – 1.86**	**<<0.01**	**1.09**	**0.18 – 2.12**	**0.01**
$Treatment_{Low}$ * Canopy	**0.22**	**0.15 – 0.29**	**<<0.01**	**0.19**	**0.06 – 0.33**	**<0.01**
PondTemp * Canopy	0.00	-0.01 – 0.02	0.79	-0.01	-0.06 – 0.04	0.66
$Treatment_{Low}$ * PondTemp * Canopy	**-0.02**	**-0.03 – -0.02**	**<<0.01**	**-0.02**	**-0.03 – 0.00**	**<0.01**
Random Effects				**Random Effects**		
σ^2	0.27			0.66		
$b_{0\ Pond}$	0.08			0.64		
Marginal R^2 / Conditional R^2	0.89 / 0.91			0.70 / 0.85		

Table 3. Results of structural equation model predicting the pond-wise change in embryonic period from 2001 to 2018.

		Low			**High**		
Response	Predictor	Std. Est.	Est.	P value	Std. Est.	Est.	P value
Δ Volume	Δ Canopy	-0.26	-0.04	0.09	-0.17	-0.03	0.27
Δ Volume	Δ Temp	0.28	1.44	0.07	0.27	1.4	0.08
Δ Temp	Δ Canopy	-0.04	0.00	0.92	-0.04	0.00	0.92
Δ EP	Δ Temp	**0.73**	**5.81**	**<<0.01**	**0.47**	**1.71**	**<0.01**
Δ EP	Δ Canopy	0.14	0.03	0.12	0.03	0.00	0.84
Δ EP	Δ Volume	**-0.22**	**-0.34**	**<0.01**	-0.17	-0.12	0.05
		Marginal R2/ Conditional R2			Marginal R2/ Conditional R2		
Δ Volume		0.14 / 0.38			0.10 / 0.36		
Δ Temp		0.00 / NA			0.00 / NA		
Δ EP		0.50 / 0.54			0.20 / 0.57		

Table 4. Model results estimating the relationship between environmental change (sum of absolute standardized change in pond temperature and canopy) and female breeding population size changes over time for wood frog populations at Yale Myers Forest (N_{pond} = 14, N_{obs} = 266). Confidence intervals represent the 95 percentile values from 1000 non-parametric bootstraps.

Predictors	*Estimates*	*CI*	*p*
Intercept	0.019	-0.003 – 0.050	0.5
Environmental Change	0.039	0.020 – 0.062	<0.01
log(Year)	0.016	0.003 – 0.026	0.16
Environmental Change * log(Year)	-0.019	-0.028 – -0.011	<0.01
Random Effects			
σ^2	0.002		
$b_{0\ Pond}$	0.002		
$B_{1\ log(Year)}$	0.001		
$b_{0\ Pond:log(Year)}$	-0.001		
Marginal R^2 / Conditional R^2	0.15 / 0.25		

FIGURES

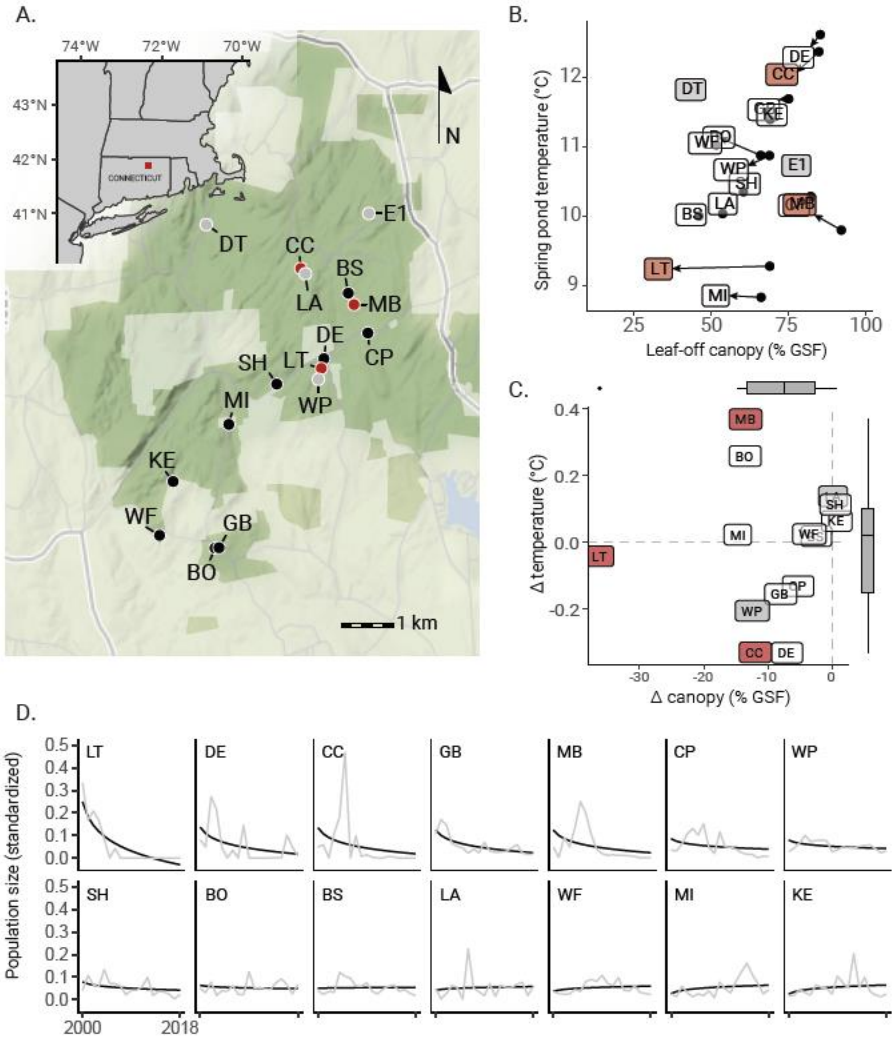

Figure 1. Embryonic development rates and natal pond environment were measured for 16 wood frog breeding populations (A). Cohorts sampled in both 2001 and 2018 for nine ponds were included in the experiment. Three ponds that were included in the 2001 experiment ceased to host breeding populations in 2018 (red) and four additional populations were included in the 2018 experiment (grey). Absolute (B) and relative (C) change in leaf-off canopy closure (measured as percent global site factor) and water temperature during the embryonic period (day-of-year 91 to 135) are shown for those ponds with records. Marginal distribution of environmental change are indicated by the box plots in C. Population trends (D) are reported as relative proportion of total population counts egg mass counts from 2000 to 2018 (grey lines), with logarithmic growth curves for each population (black lines).

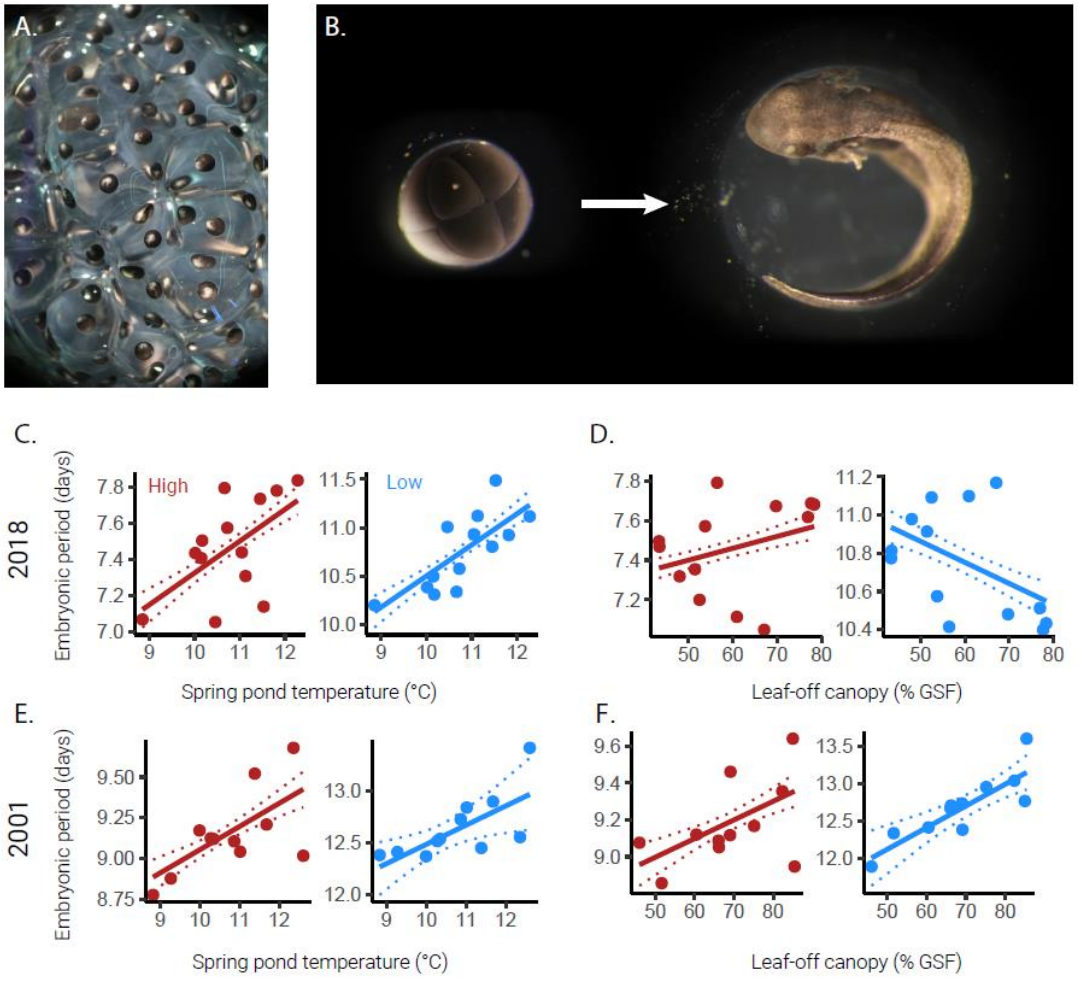

Figure 2. Embryonic periods were estimated by rearing individual embryos excised from wood frog egg masses (A) and fitting development rate models to Gosner (1960) stages assigned from photographs taken of embryos between oviposition and hatching (B). Countergradient variation was assessed by regressing estimated embryonic periods against pond water temperature and canopy closure values for each cohort with mixed-effect multiple regression. The partial effects of pond water temperature (C, E) while holding canopy constant and canopy (D, F) while holding temperature constant are shown for embryos in the high (red) and low (blue) temperature treatments. Points indicate mean population-wise estimates (model fit + conditional residuals). 95% confidence intervals (dashed lines) for the regression line were estimated by fitting the models to 1000 non-parametric bootstrap replicates.

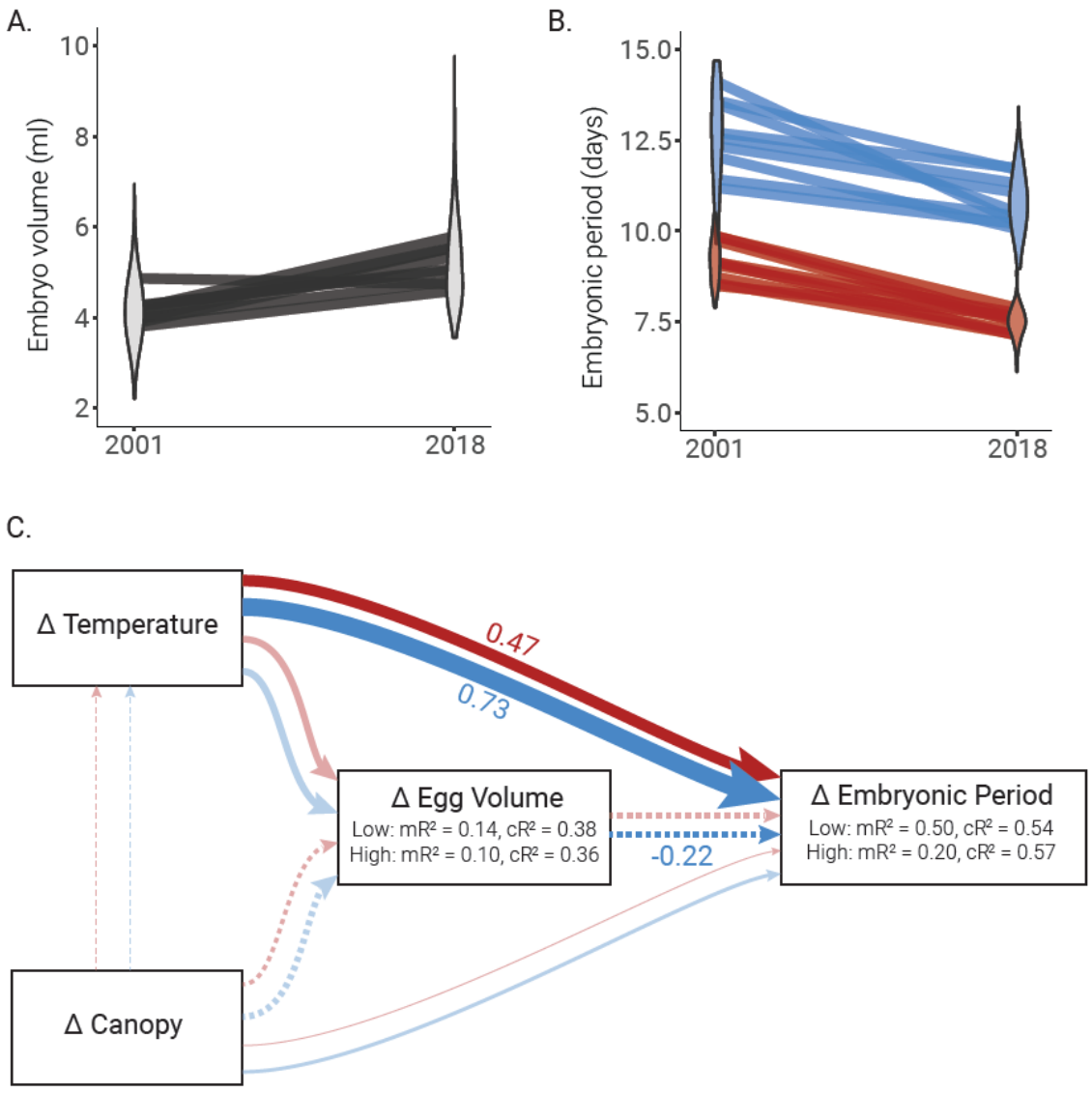

Figure 3. In general, embryonic volume increased (A) and embryonic periods decreased (B) between cohorts. Violin plots indicate overall change between cohorts while lines between years show the change in means for each of the nine populations included in both the 2001 and 2018 common garden experiments. Structural equation models were used to estimate the strength and direction of variables hypothesized to cause a change in embryonic period (C). Paths in C represent directional causal paths. The widths of the paths are scaled to the standardized coefficient estimates. Solid paths indicate positive relationships while dashed lines are negative. Significant paths are opaque and include standardized coefficient values while nonsignificant paths are shown as transparent. Marginal and conditional coefficients of determination for piecewise models are shown. In all figures, colors represent high (red) and low (blue) temperature treatments.

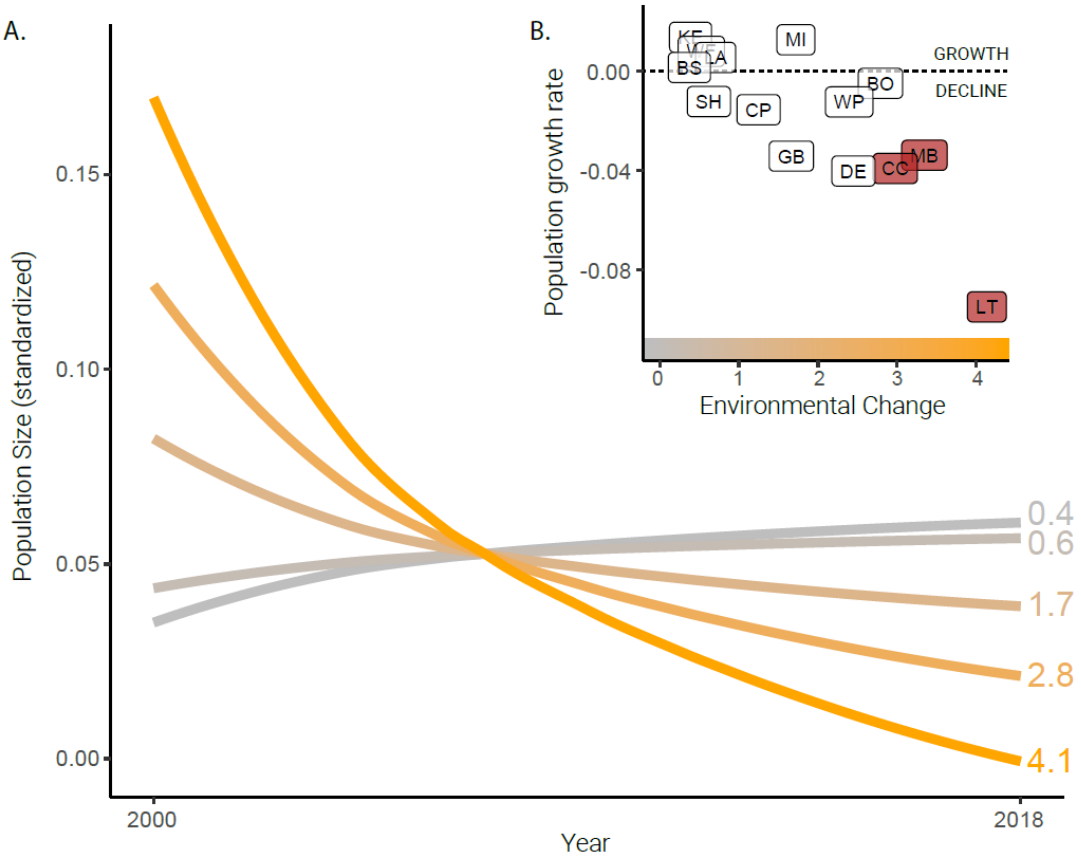

Figure 4. Population growth rates tend to decrease (B) and populations decline (A) with greater environmental change (sum of the scaled change in pond temperature and canopy). Growth curves (A) were predicted for the quartile values of environmental change experience by populations in the study. Conditional estimates of the log transformed slope of population growth (i.e., population growth rates) for each pond population are shown along the environmental change gradient in B. Three pond populations that declined to extinction during the course of this study are indicated in red (B).

APPENDIX

Supporting Material

Rapid microgeographic evolution in response to climate change

Authors

A. Z. Andis Arietta, Yale University, School of the Environment (corresponding author)
David K. Skelly, Yale University, School of the Environment

Correspondence: a.andis@yale.edu; 370 Prospect St., New Haven, CT 06511

1. Canopy cover

Methods

Canopy changes over time through successional changes in species composition, maturation, or short-term changes due to disturbance like beaver alterations, fire, wind-fall, or timber harvest. Above-canopy radiation is affected by the time of day, sky clarity, and season (Promis, 2013). The understory light environment is a function of the interaction between the light blocked by the canopy and the magnitude and angle of solar radiation above the canopy.

Hemispherical photos can be used to characterize canopy closure at a given point in the understory (Jennings et al., 1999) and can be used in conjunction with solar radiation models to integrate total or average radiation incident at that point over any time interval (Rich, 1989; Promis et al., 2012). Individual points estimates can be treated as spatial samples and averaged to estimate light environment over space.

We estimated contemporary light environment for Wood frog breeding ponds as the proportion of direct and indirect radiation incident upon a point below the canopy to the same point absent the canopy (Anderson, 1964) using hemispherical photos captured in 2017 and 2018.

For each pond, we took photos at each of the cardinal points (N, S, E, W) 1m in from the bank along the spring high-water perimeter and in the center (C) of the pond at the intersection of the N-S and E-W lines. We captured one set of 5 photos during the summer while the canopy was in full leaf and 5 photos at the same points during the winter after the leaves had fallen. All photos were taken approximately 1.5 vertical meters above the spring high water height with a handheld leveling system.

Hemispherical photos require processing into binarized images prior to analysis (Rich, 1990; Glatthorn and Beckschäfer, 2014). Many errors can arise when camera settings, corrective editing, or binarization techniques are inappropriate. In the field, we captured photos on uniformly overcast days. We standardized exposure settings to the sky by setting ISO, aperture, and shutter speed so that the open sky was at exposure value 0, then

down-stepped shutter speed by two stops. This ensures that the grey value in the photos is standardized to the sky (Wagner, 1998). Even on uniformly overcast days, the sky brightness changes over time. To account for this, we captured photos in RAW format and adjusted exposure values in Adobe Lightroom 5.7.1 to ensure that the brightest pixel grey value aligned to full white (Beckschäfer et al., 2013).

We exported images as full sized JPEG files from Lightroom and used Hemispherical_2.0 (Beckschäfer, 2015) plugin for ImageJ. Hemispherical_2.0 uses a "minimum" thresholding algorithm on the blue color plane of the image to classify each pixel as black or white. Binarized images were then converted to BMP format and analyzed in Gap Light Analyzer (Frazer et al., 1999). We parameterized the GLA models with estimates from the 1991-2010 Update to the National Solar Radiation Database (2010) for Hartford, Connecticut (approximately 40 km from the study site) [Citation: National Solar Radiation Database (2010) https://rredc.nrel.gov/solar/old_data/nsrdb/1991-2010/].

The typical pre-metamorphic period for wood frogs at our site begins with oviposition in early April and ends with metamorphosis by the end of August. Leaves typically emerge in Mid-may. Therefore, we estimated leaf-on and leaf-off GSF estimates from the set of summer photos over the period April 1 to May 15 and winter photos over the period May 16 to August 30, respectively. We calculated a spring + summer GSF for the period April 1 through August 30 as the weighted average of the leaf-on and leaf-off estimates with respect to the proportion of days before and after leaf emergence. Analyses for both seasonal windows and subsequent models incorporating spring + summer canopy are included in this appendix, but were not sufficiently different from the spring estimates to include in the manuscript.

These methods closely replicate those used by Halverson (Halverson et al., 2003) for the canopy closure estimates used by Skelly in the 2001 embryonic development rate experiment (Skelly, 2004), with two exceptions that may bias comparisons across time. First, the historical estimates were calculated with proprietary HemiView software rather than Gap Light Analyzer. These programs are analogous and we set identical parameters to mitigate bias. Second, the sampling layout for canopy photos differed between timespoints. Canopy values used in Skelly (2004) were estimated from hemispherical photos taken along the intersections of 5 m cartesian grids (2 to 172 photos per pond) and analyzed in HemiView software (Halverson et al., 2003) while contemporary photos were sampled at cardinal points and the center of the pond.

In order to compare these historical data to current estimates, we subsampled 5 points from the cartesian grid estimates, corresponding to the 4 cardinal point and 1 central point and tested for a difference in GSF estimates. We compared the pairwise differences between spring GSF estimated from the two methods using a paired t-test and estimated pearson's correlation between estimates.

Results

Our dataset includes historical canopy estimates for 14 of the 16 ponds included in either the 2001 or 2018 experiment (excludes DT and E1). On average, leaf-off estimates of canopy closure from photos at the cardinal points and center of the ponds are 1.8% darker than estimates using photos taken along a cartesian grid, but the difference is not significant

(p = 0.08, 95% C.I. = -3.9 - 0.3%). The estimates of leaf-off GSF from subsampled photos are tightly correlated with the original estimates (R^2 = 0.92) (Fig S1, A).

Larval period canopy closure is estimated to be significantly darker by 2.2% when comparing photos from cardinal points to those captured along a cartesian grid (p = 0.04, 95% C.I. = -4.3 - 0.1%) with similarly strong correlation between methods (R^2 = 0.97) (Fig S1, B).

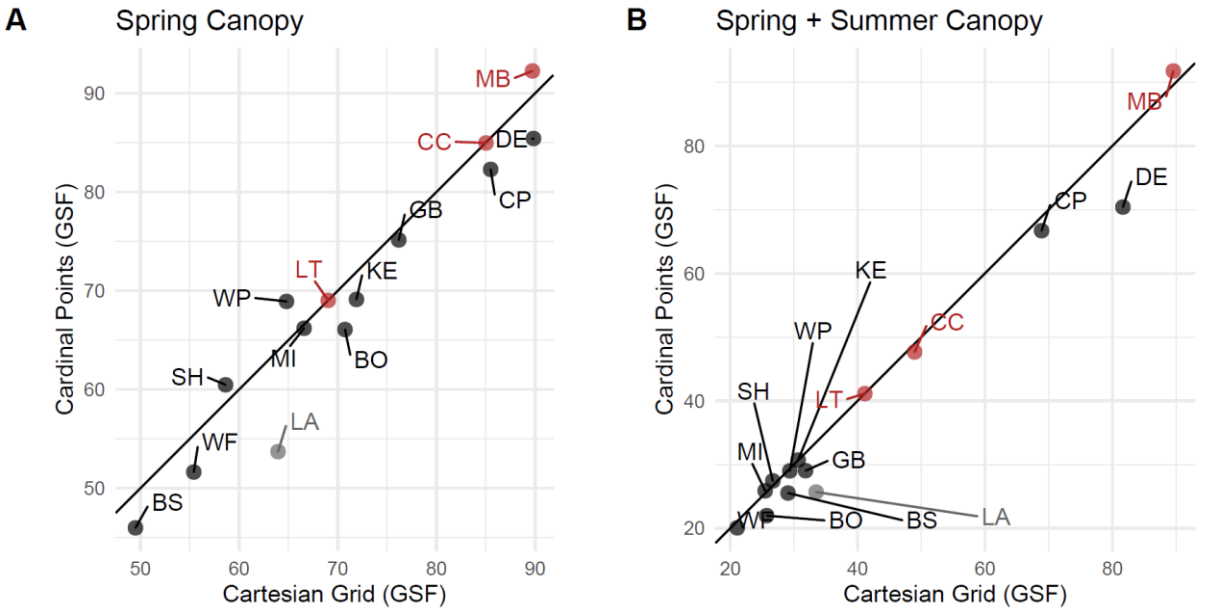

Figure S1. Comparison of spring (A, leaf-off) and spring + summer (B, weighted mean leaf-on and leaf-off over wood frog aquatic development period) canopy closure estimates (GSF = Global Site Factor) for fourteen wood frog breeding ponds averaged over 2 to 172 photos captured at the intersection of 5m cartesian grids (Cartesian Grid) or a subsample of 5 photos captured at the 4 cardinal points and center (Cardinal Points) of each pond during 1999 and 2000 seasons. Colors indicate ponds that are represented in both experimental timepoints (black), only the 2001 experiment (red), and only the 2018 experiment (grey).

2. Pond temperature

Methods

Over the past few decades, air temperatures in our region have been increasing, but the magnitude varies across the season. Summer temperatures during the wood frog larval period have increased more than two-fold compared to temperatures during the embryonic period (Arietta et al., 2020). Temperature-mediated selection during either period may drive evolutionary changes in embryonic development. Therefore, we estimated long-term average temperatures for two seasonal windows from day-of-year 91 to 135 and day-of-year 91 to

168. These windows correspond to the leaf-off period of embryonic and early larval development and the aquatic development period, respectively.

Starting in 2001, we recorded water temperatures in a subset of the 60 wood frog breeding ponds at our field site. We recorded water temperature in breeding ponds every 0.5 or 1 hour with submersible temperature loggers (HOBO 8K Pendant (Onset Computer Corporation)) suspended 10cm below the surface at the deepest point in the pond. Loggers were deployed within days of oviposition (approximately DOY 91) and removed after larvae had metamorphosed or the pond dried (approximately DOY 168) (Fig S2). We computed daily means from the hourly data and restricted our data to the window of greatest data saturation (DOY 91 to 168), resulting in 18,524 daily observations from 34 ponds over 18 years between 2001 and 2019 for our training dataset.

We imputed missing temperature values with a random forest model implemented in randomForest (v4.6.14 (Liaw and Wiener, 2002)) trained on all ponds at our site. Regression trees included both climate and site predictor variables. Climate variables included daily maximum, minimum and average temperature, radiation, precipitation, snowpack, and atmospheric pressure from the Daymet database (v.3 (Thornton et al., 2016)). We included a 1-day lag for all climate variables. Site variables included elevation, aspect, latitude, and a factor for individual ponds. A random noise variable was included in the set of predictor variables to assess relative importance. We grew our random forest from 300 regression trees (Fig S3) and used 10-fold cross validation to evaluate the predictive accuracy for imputed points.

Results

We grew our random forest from 300 trees with 8 predictor variables randomly tried at each split. Forest explained 87.5% of the variance in the data (RMSR = 1.32). Air temperature was the most important variable in both accuracy (i.e. reducing out-of-bag error) and node purity (i.e. correctly splitting trees) (Fig S4). On average, our predicted daily pond temperatures are off by +/- 0.54 C as assessed by 10-fold cross validation. In total, we imputed 29,212 observations, 61% of the total dataset.

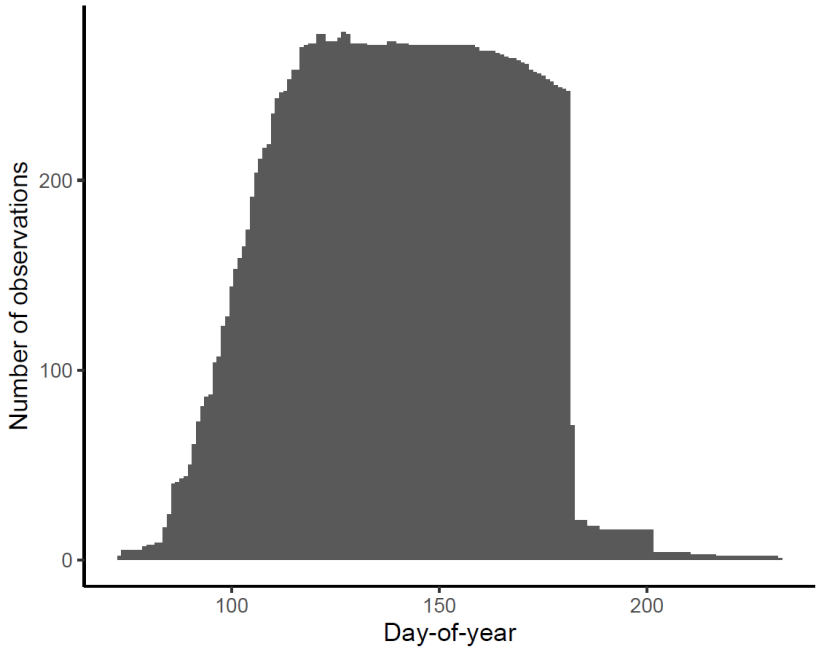

Figure S2. The number of daily observations across all 34 ponds and 18 years (2001 to 2019) included in the random forest training dataset. Loggers were deployed coincident with oviposition (+/- 2-3 days) and removed after larvae metamorphosed and left the pond or the pond dried.

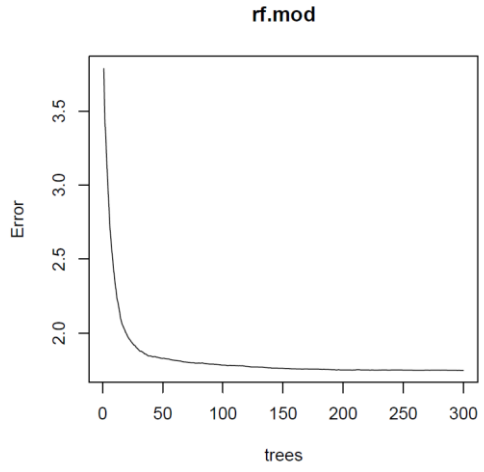

Figure S3. Decrease in error, estimated by out-of-bag cross-validation, for increasing numbers of trees included in the random forest model predicting daily pond water temperature.

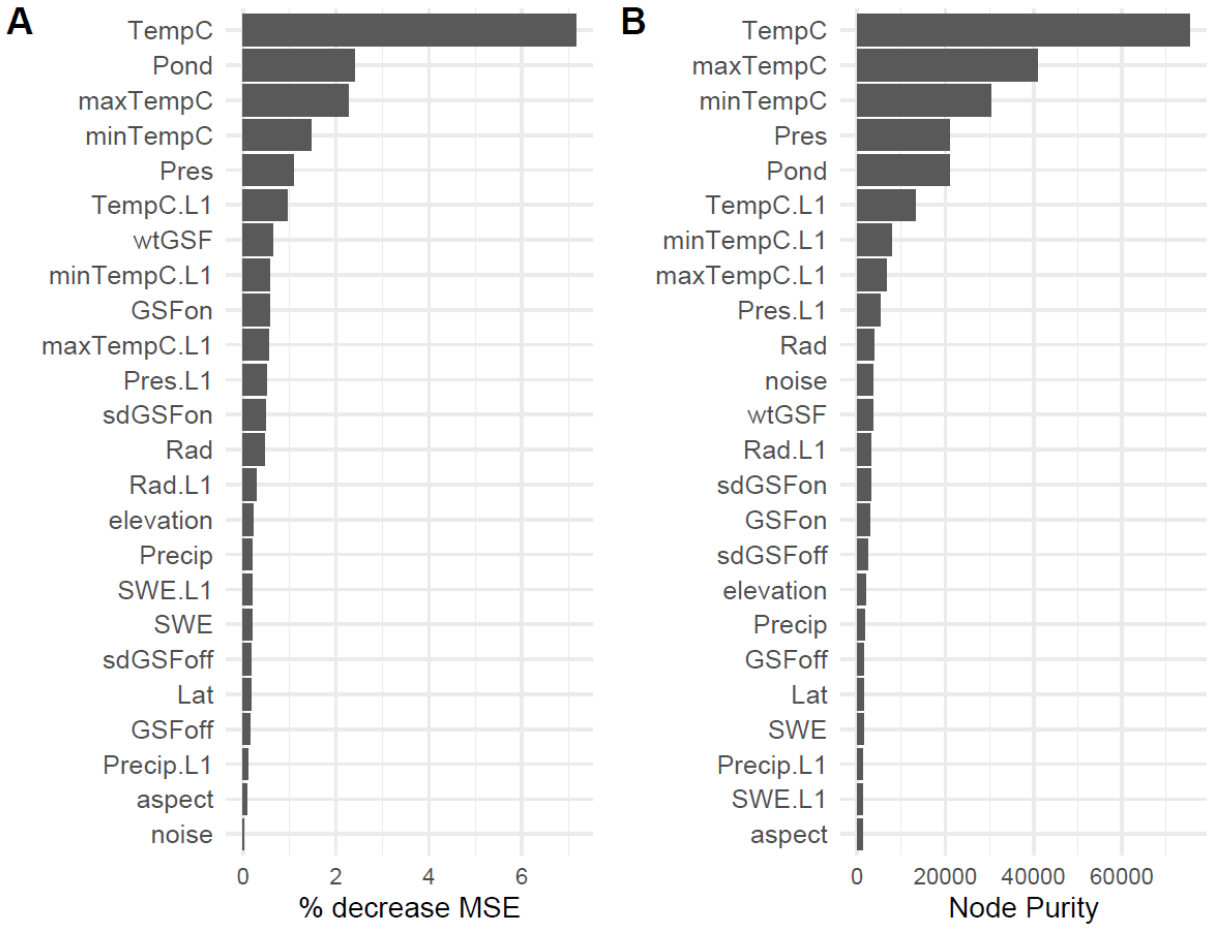

Figure S4. Importance of variable in predictive accuracy of the random forest model prediction daily pond water temperatures measured by the percentage decrease in mean square error of the out-of-bag cross-validation estimates when each variable is included (A) and the total decrease in node impurity (i.e. number of correctly estimated leaves) by splitting on each variable (B). Both importance estimates were averages across all trees. Variables included average, minimum, and maximum daily pond water temperature (TempC, minTempC, maxTempC), daily atmospheric pressure (Press), daily solar radiation incident upon Earth's service (Rad), daily precipitation (Precip), and daily snow water equivalent (SWE). For all daily climate variables, a 1-day lag was also included (L1). Canopy variables included Global Site Factor values for the leaf-on (GSFon), leaf-off (GSFoff), and weigted average over the growing season (wtGSF) and the standard deviations in these measures across the five estimates for each pond (sd). Variable for Pond identity, elevation, aspect, and latitude (Lat) were included. A random noise variable (noise) was included to assess relative importance.

3. Environmental change

Results

Table S1. Change in environmental variables between the 2001 and 2018 experiments.

Pond	Δ Leaf-off GSF	Δ Weighted GSF	Δ Spring water temperature	Δ Aquatic period water temperature	Experiment
BO	-13.6	-5.36	0.26	0.34	2001, 2018
BS	-2.57	-1.96	0.02	0.38	2001, 2018
CC	-12.05	-5.86	-0.33	-0.11	2001
CP	-5.38	-16.66	-0.13	0.14	2001, 2018
DE	-7.1	-2.09	-0.33	-0.09	2001, 2018
GB	-8.1	-4.88	-0.15	0.03	2001, 2018
KE	0.59	-3.26	0.06	0.31	2001, 2018
LA	0.02	-8.48	0.14	0.18	2018
LT	-36.19	-26	-0.04	0.22	2001
MB	-13.56	-14.98	0.37	1.27	2001
MI	-14.77	-8.57	0.02	0.16	2001, 2018
SH	0.41	-8.21	0.11	0.31	2001, 2018
WF	-3.65	0.19	0.03	0.19	2001, 2018
WP	-12.49	-6.64	-0.2	-0.02	2018
Mean	-9.17	-8.05	-0.01	0.24	NA

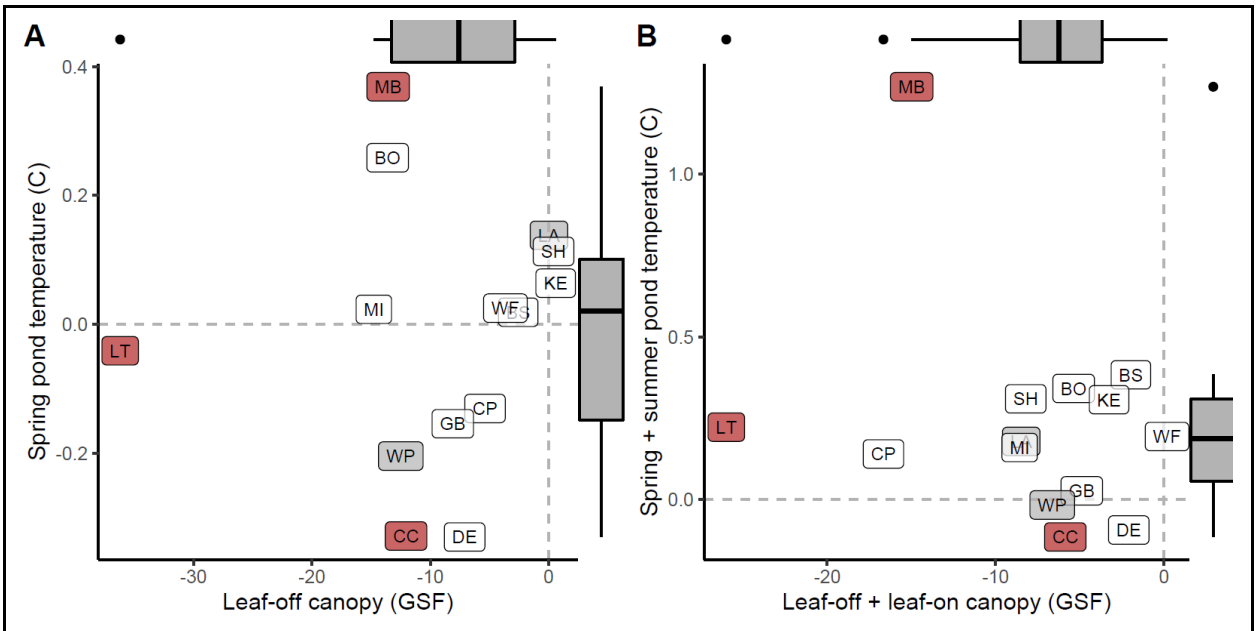

Figure S5. Change in pond temperature in relation to canopy for 14 wood frogs ponds between 2001 and 2018 for spring (A) and spring and summer (B) seasonal windows. Red points indicate populations that ceased to host breeding populations between 2001 and 2018. Circles represent ponds represented in both 2001 and 2018 datasets. Spring pond temperatures are averaged over the period of DOY 91 to 135 and spring canopy is the average leaf-off canopy closure values for

the same period. This corresponds to the embryonic development period of wood frogs. Spring and summer temperatures are averaged over the period DOY 91 to 168. Spring and summer canopy values are the weighted average of leaf-off and leaf-on canopy closure values from DOY 91 to 243. Marginal boxplots show the mean and interquartile range of pond-wise values.

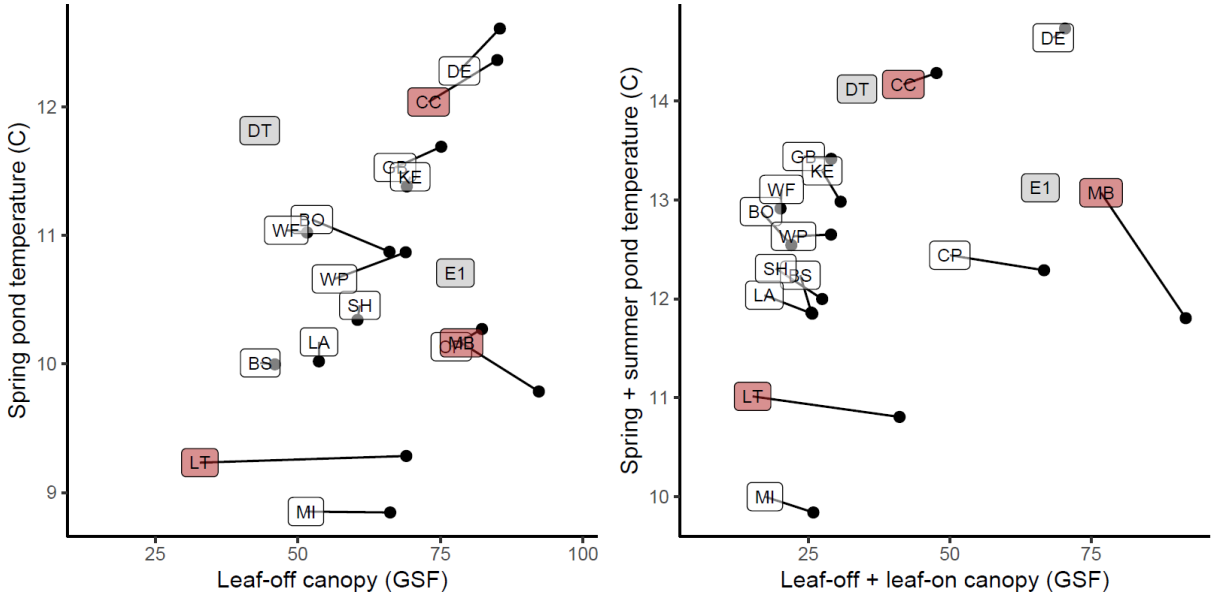

Figure S6. Change in pond temperature in relation to canopy for 16 wood frogs ponds between 2001 (point) and 2018 (label) for spring (A) and spring and summer (B) seasonal windows. Red labels indicate populations that ceased to host breeding populations between 2001 and 2018. Grey labels indicate ponds for which conditions in 2001 are unknown. Spring pond temperatures are averaged over the period of DOY 91 to 135 and spring canopy is the average leaf-off canopy closure values for the same period. This corresponds to the embryonic development period of wood frogs. Spring and summer temperatures are averaged over the period DOY 91 to 168. Spring and summer canopy values are the weighted average of leaf-off and leaf-on canopy closure values from DOY 91 to 243.

4. Embryonic development and size

Results

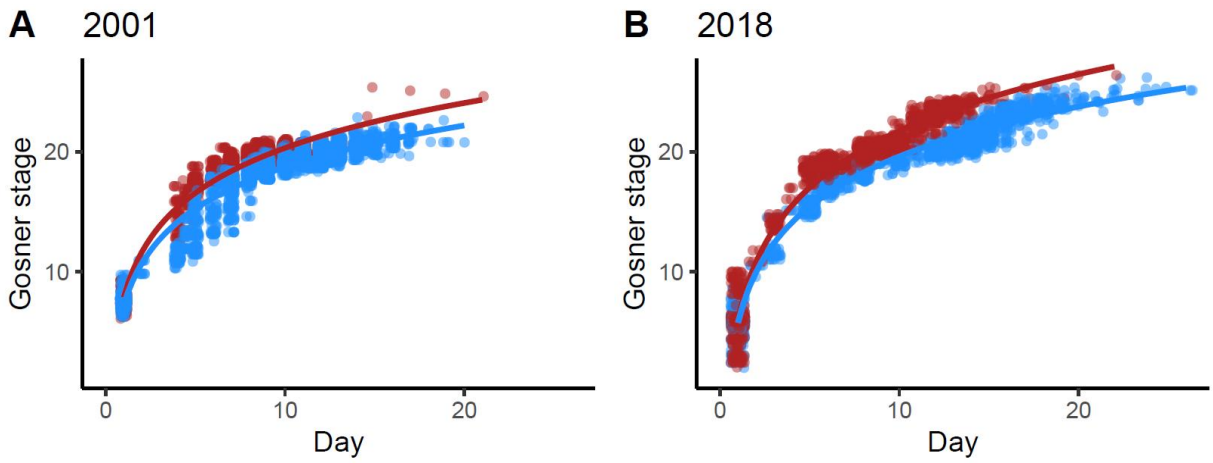

Figure S7. Embryonic development rates of wood frog embryos collected within 24 hours of oviposition in 2001 (A) and 2018 (B) and reared in incubators representing high (red) or low (blue) temperatures experienced across natal ponds until hatching (approximately Gosner state 20). Curves represent marginal fits of mixed effect models. Embryos were photographed throughout development in order to estimate developmental stage.

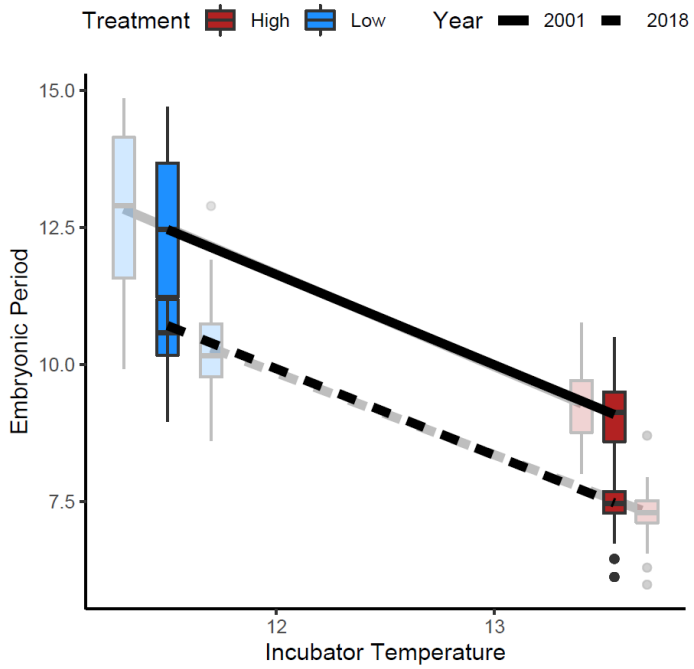

Figure S8. Embryonic period duration for each embryo was estimated from developmental growth rates corrected for differences in realized incubator temperatures. This was accomplished by including incubator temperature as a continuous variable in the development rate models. The effect of incubator temperature was then held constant at the midpoint between years (a difference of +/-0.2 C for the low temperature treatment or +/-0.15 C for the high temperature treatment) when predicting embryonic periods. Uncorrected estimates are shown (transparent) behind the corrected estimates (solid). Lines connect treatment groups from the same year.

Table S2. Mixed effect model results testing for difference in initial embryo volume between 2001 and 2018 experiment with random intercepts for clutchmates nested within pond. Year 2001 is the baseline value for the categorical Year variable. N_{year} = 2, N_{pond} = 24, N_{clutch} = 134, N_{obs} = 1526.

	Vol.mm3		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	4.18	3.85 – 4.53	**<0.001**
Year [2018]	1.15	0.67 – 1.63	**<0.001**
Random Effects			
σ^2	0.16		
$b_{1\ Year:Pond:Clutch}$	0.44		
$b_{0\ Year:Pond}$	0.24		
Marginal R^2 / Conditional R^2	0.28 / 0.86		

Table S3. Mixed effect model results testing for difference in initial embryonic period between 2001 and 2018 experiment with random intercepts for ponds. Year 2001 is the baseline value for the categorical Year variable. N_{pond} = 15, $N_{obs(Low)}$ = 735, $N_{obs(High)}$ = 744.

	EP (Low temperature treatment)			**EP (High temperature treatment)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	12.46	11.99 – 12.95	**<0.001**	9.07	8.82 – 9.32	**<0.001**
Year [2018]	-1.77	-2.43 – -1.09	**<0.001**	-1.60	-1.93 – -1.28	**<0.001**
Random Effects				**Random Effects**		
σ^2	0.69			0.12		
$b_{Year:Pond}$	0.64			0.17		
Marginal R^2 / Conditional R^2	0.37 / 0.67			0.69 / 0.87		

5. Countergradient variation

Results

Model selection tables

Table S4. Model selection table predicting embryonic period by leaf-off canopy (Can) and spring water temperatures (Temp) for the 2001 and 2018 experimental cohorts and vary with temperature treatment (Treat). Pond-wise random intercepts were included in all models. The difference in AIC compared to the best model in each year is represented as Δ AIC.

		2018		2001	
Model	Parameters	AIC	Δ AIC	AIC	Δ AIC
Treat * Can	6	1355.1	80.6	1689.3	0
Treat * Temp	6	1335.2	60.7	1702.7	13.4
Treat * Can + Treat * Temp	8	1314.1	39.6	1692.8	3.5
Treat * Can * Temp	10	1274.5	0	1689.8	0.5

Larval season countergradient regressions

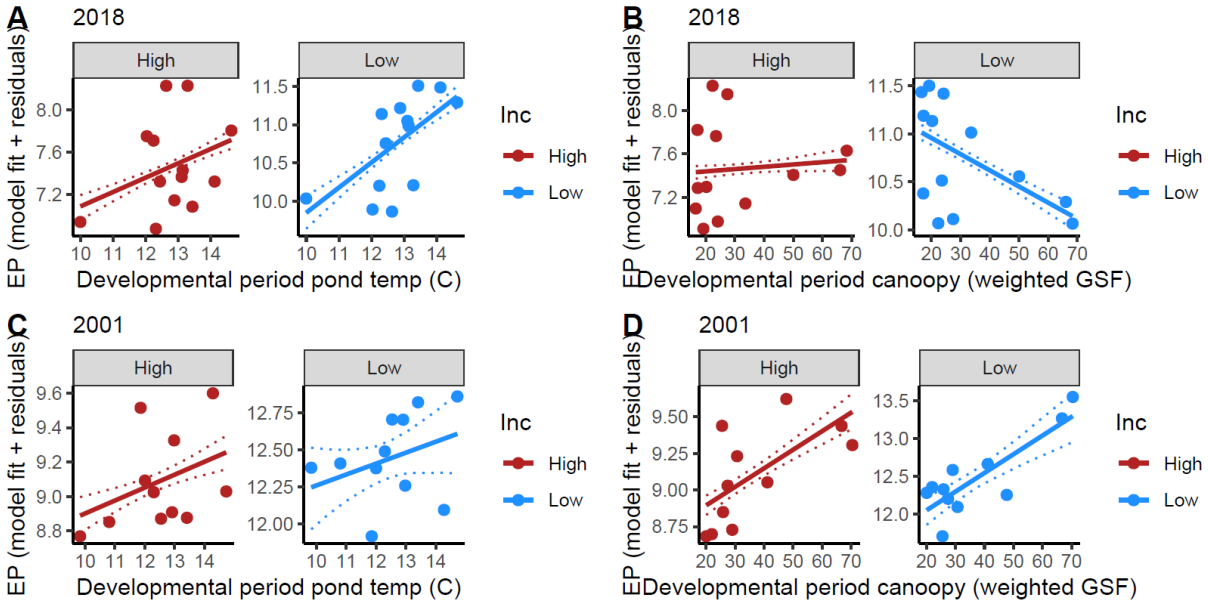

Figure S9. Partial effect plots for embryonic period duration estimated with linear mixed effect models for 2018 (top row, A & B) and 2001 (bottom row, C & D) cohorts when reared in a common garden at high (red) and low (blue) temperature treatments. The effect of larval period pond

temperature gradient while holding canopy at the mean (A & C) and the effect of spring and summer canopy gradient while holding temperature at the mean (B & D) are shown. Points represent pond-wise average fit (i.e. model fit plus conditional residuals). 95% confidence intervals for the regression line were estimated by fitting the models to 1000 non-parametric bootstrap replicates.

Table S5. Mixed effect regression results for estimating the effect of larval period pond temperatures (PondTemp) and spring and summer canopy cover (Canopy) on embryonic period for the 2018 and 2001 experimental cohorts.

	EP (2018)			**EP (2001)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
Intercept	**5.78**	**1.92 – 9.50**	**<0.01**	6.71	-6.93 – 19.34	0.32
$Treatment_{Low}$	-1.51	-3.24 – 0.09	0.07	**3.78**	**0.18 – 7.76**	**0.04**
PondTemp	0.13	-0.17 – 0.43	0.40	0.15	-0.81 – 1.25	0.77
Canopy	0.00	-0.12 – 0.12	0.99	0.04	-0.28 – 0.37	0.77
$Treatment_{Low}$ * PondTemp	**0.42**	**0.30 – 0.55**	**<<0.01**	-0.07	-0.38 – 0.21	0.65
$Treatment_{Low}$ * Canopy	**0.07**	**0.03 – 0.13**	**<0.01**	-0.01	-0.10 – 0.07	0.82
PondTemp * Canopy	0.00	-0.01 – 0.01	0.96	0.00	-0.03 – 0.02	0.82
$Treatment_{Low}$ * PondTemp * Canopy	**-0.01**	**-0.01 – 0.00**	**<0.01**	0.00	0.00 – 0.01	0.62
Random Effects				**Random Effects**		
$b_{0\ Pond}$	0.08			0.63		
σ^2	0.26			0.67		
$N_{\ Pond}$	13			11		
$N_{\ Observations}$	805			674		
Marginal R^2 / Conditional R^2	0.89 / 0.91			0.69 / 0.84		

6. Temporal comparisons

Table S6. Exposition of causal assumption tested by the structural equation models.

Path	Causal assumption	Reference
Δ Canopy → Δ Temperature	Canopy influences pond site microclimate by shading radiation and altering evaporative cooling	(De Frenne et al., 2021)
Δ Canopy → Δ EP	Canopy drives microgeographic variation in embryonic development rates	(Skelly, 2004)
Δ Temperature → Δ EP	Temperature drives microgeographic variation in embryonic development rates	*This study*
Δ Egg Vol. → Δ EP	Larger embryos represent greater provisioning which influences embryonic development rates	(Berven, 1988; Berven and Chadra, 1988; Bradford, 1990)
Δ Canopy → Δ Egg Vol.	Microgeographic variation in canopy influences maternal condition and provisioning	(Werner and Glennemeier, 1999)
Δ Temperature → Δ Egg Vol.	Microgeographic variation in temperature influences maternal condition and provisioning	(Berven, 1988; Rowiński et al., 2020)

References

Anderson, M. C. (1964). Studies of the woodland light climate I. The photographic computation of light condition. *Journal of Ecology* 52, 27–41.

Arietta, A. Z. A., Freidenburg, L. K., Urban, M. C., Rodrigues, S. B., Rubinstein, A., and Skelly, D. K. (2020). Phenological delay despite warming in wood frog *Rana sylvatica* reproductive timing: a 20-year study. *Ecography* 52, 27.

Beckschäfer, P. (2015). Hemispherical_2.0: Batch processing hemispherical and canopy photographs with ImageJ - User manual. doi:10.13140/RG.2.1.3059.4088.

Beckschäfer, P., Seidel, D., Kleinn, C., and Xu, J. (2013). On the exposure of hemispherical photographs in forests. *iForest* 6, 228–237.

Berven, K. A. (1988). Factors Affecting Variation in Reproductive Traits within a Population of Wood Frogs (Rana sylvatica). *Copeia* 1988, 605–615.

Berven, K. A., and Chadra, B. G. (1988). The relationship among egg size, density and food level on larval development in the wood frog (*Rana sylvatica*). *Oecologia* 75, 67–72.

Bradford, D. F. (1990). Incubation Time and Rate of Embryonic Development in Amphibians: The Influence of Ovum Size, Temperature, and Reproductive Mode. *Physiol. Zool.* 63, 1157–1180.

De Frenne, P., Lenoir, J., Luoto, M., Scheffers, B. R., Zellweger, F., Aalto, J., et al. (2021). Forest microclimates and climate change: Importance, drivers and future research agenda. *Glob. Chang. Biol.* doi:10.1111/gcb.15569.

Frazer, G. W., Canham, C. D., and Lertzman, K. P. (1999). Gap Light Analyzer (GLA): Imaging software to extract canopy structure and gap light transmission indices from true-color fisheye photographs, users manual and documentation. Burnaby, British Columbia: Simon Fraser University Available at: http://rem-main.rem.sfu.ca/downloads/Forestry/GLAV2UsersManual.pdf.

Glatthorn, J., and Beckschäfer, P. (2014). Standardizing the protocol for hemispherical photographs: accuracy assessment of binarization algorithms. *PLoS One* 9, e111924.

Halverson, M. A., Skelly, D. K., Kiesecker, J. M., and Freidenburg, L. K. (2003). Forest mediated light regime linked to amphibian distribution and performance. *Oecologia* 134, 360–364.

Jennings, S. B., Brown, N. D., and Sheil, D. (1999). Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures. *Forestry* 72, 59–74.

Liaw, A., and Wiener, M. (2002). Classification and Regression by randomForest. *R News* 2, 18–22. Available at: https://CRAN.R-project.org/doc/Rnews/.

Promis, A. (2013). Measuring and estimating the below-canopy light environment in a forest. a review. 19, 139–146.

Promis, A., Caldentey, J., and Cruz, G. (2012). Evaluating the usefulness of hemispherical photographs as a means to estimate photosynthetic photon flux density during a growing season in the understorey of Nothofagus pumilio forests. *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology* 146, 237–243.

Rich, P. M. (1989). A Manual for Analysis of Hemispherical Canopy Photography. Los Alamos National Laboratory.

Rich, P. M. (1990). Characterizing plant canopies with hemispherical photographs. *Remote Sens. Rev.* 5, 13–29.

Rowiński, P. K., Laurila, A., Gotthard, K., Sowersby, W., Lind, M. I., Richter-Boix, A., et al. (2020). Parental effects influence life history traits and covary with an environmental

cline in common frog populations. *Oecologia* 192, 1013–1022.

Skelly, D. K. (2004). Microgeographic countergradient variation in the wood frog, *Rana sylvatica*. *Evolution* 58, 160–165.

Thornton, P. E., Thornton, M. M., Mayer, B. W., Wei, Y., Devarakonda, R., Vose, R. S., et al. (2016). Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3. doi:10.3334/ORNLDAAC/1328.

Wagner, S. (1998). Calibration of grey values of hemispherical photographs for image analysis. *Agric. For. Meteorol.* 90, 103–117.

Werner, E. E., and Glennemeier, K. S. (1999). Influence of Forest Canopy Cover on the Breeding Pond Distributions of Several Amphibian Species. *Copeia* 1999, 1–12.

CONCLUSION

Wild organisms can rapidly adapt to changing environments, even at fine spatial scales (Richardson et al., 2014). However, there is a limit to the capacity for adaptation to buffer extreme environmental change (Bell, 2013; De Meester et al., 2018). Understanding the adaptive potential of individual species at fine-resolution is an important task if we hope to accurately predict the impacts of future climate and landscape changes and gauge the risk to biodiversity (Urban et al., 2016). My dissertation takes advantage of an uncommonly long-observed and closely-studied metapopulation to paint a comprehensive picture of evolution over time in association with shifts in ecological contexts.

Throughout North America, the clucking chorus of breeding wood frogs and appearance of gelatinous egg masses are one of the earliest harbingers of spring. Many springtime biotic events have advanced earlier into spring as the climate warms, especially amphibian species and species in temperate regions (Cohen et al., 2018). However, in chapter 2, I examined the date at which wood frogs deposited eggs in communal aggregations. I found that, contrary to expectations, oviposition was delayed by almost 1.4 days per decade despite a 0.3 °C per decade increase in temperatures. I used a combination of sliding-window correlation analysis and ensemble regression trees to consider seasonal effects that are ordinarily subsumed in analyses that rely on annual averages. This analysis indicated that wood frog oviposition timing corresponded to a timepoint in the season characterized by minimal increase in temperatures but longer persisting snowpack.

In chapter 3, I investigated the climate and drought patterns over the latter part of the wood frog's aquatic lifecycle. When considered alongside the data in chapter 2, which shows that the bulk of warming occurs later in the season, the picture that emerges is of vernal ponds that are increasingly inaccessible in the spring due to late snowpack and tend

to dry earlier due to warmer climates and drought. Thus, hydroperiods may be contracted and the timespan available for larval development is shortened.

Amphibians are marvels of phenotypic plasticity in response to drying ponds, as well as other environmental cues (Denver, 1997). In chapter 3, I explore how shifts in physiological rates track changes in ecological drivers. Rates of development and growth have changed dramatically over time. Most notably, larvae develop much faster, conforming to expectations of environmentally-contracted hydroperiods. In other words, frogs fit their aquatic ontogeny into the time frame available—which is decreasing over time—to avoid desiccation and mass mortality events as ponds dry.

Most mortality in wood frogs occurs in pre-metamorphic life stages (92-99%) and desiccation is a significant culprit (Dodd, 2013). Thus, there is strong selection pressure for wood frogs to adapt to changing climate and habitat. In chapter 5, I show that a pattern of microgeographic variation persists after nearly two decades and is maintained by local adaptation tracking shifts in realized pond temperatures. This study overcomes a common challenge in inferring evolution in that I was able to track phenotypes over time while also controlling for plasticity with a common garden design. In doing so, these results add another example to the growing list of cases evincing the phenomenon of rapid evolution.

Relative to prior research, this wood frog system is unique in that evolution is demonstrable not only at short temporal scale, but also at small spatial scale. In general, evolutionary theory would suggest that population-level difference in traits should attenuate with gene flow and that gene flow is usually a function of dispersal density (Lenormand, 2002). This study is a clear counter-example, with countergradient local adaptation present among populations with overlapping dispersal kernels.

For countergradient local adaptation to be maintained at microgeographic scales, there must be opposing directional selection on a trait at both ends of the environmental

gradient. Otherwise, advantageous alleles for the maximum trait values would be shared across all populations within a few generations, eroding local differences. Chapter 4 demonstrates one potential trade-off mechanism to explain why faster intrinsic development is disfavored in warmer ponds. Specifically, developing too fast results in slower burst swimming speeds used to avoid predators, so slower intrinsic development counteracts faster exogenous development in warm ponds.

The results in this dissertation have important applications in conservation. First, rapid adaptation to climate change is possible, even in structured populations. In addition to providing a hopeful outlook, this conclusion highlights the need to account for rapid evolution in predicting and managing biodiversity.

Second, evolutionary rescue via rapid adaptation has limits. The pace of contemporary environmental change will outpace the ability for some populations to adapt.

Third, the entirety of this dissertation was conducted in populations spanning just a few kilometers. All of this interesting biology would be entirely lost into one or a few grid cells of most spatial ecological models. Thus, we must consider that our large-scale predictions are averaging over fine-scale eco-evolutionary processes which could produce spurious results.

Fourth, long-term datasets are crucial for understanding ecological and evolutionary consequences of contemporary environmental change. Most of the results in this dissertation would have been indecipherable without detailed historical data.

Fifth, modern technology can empower anyone to collect high-quality, scientifically salient ecological data, as demonstrated in chapter 1. Greater adoption of citizen science could provide rich datasets for tracking eco-evolutionary processes in real time for conservation management.

Several questions remain. For instance, the trend of delayed oviposition evident in chapter 2 was overwhelming, but a few populations deviated from the general trend. Further work to understand why some populations deviate from the regional trend and role of spatial and temporal variability in breeding commencement is needed. Understanding phenology in the context of metapopulations could be an important advance, especially for organisms in which phenology may be influenced by behavioral decisions in addition to physiological cues.

Many amphibians are sensitive to hydroperiod with respect to developmental timing. The results in chapter 3 indicate that drought and its influence on hydroperiod are likely a strong driver of physiological rates in our study populations. In particular, it seems that a major "Black Swan" drought in recent years may have had lasting impacts. The importance of such rare but extreme events have largely been neglected in ecology and eco-evolutionary studies (Anderson et al., 2017). As longer datasets and remotely sensed data become available, mining these datasets with efficient algorithmic tools could allow us to move beyond modelling just the mean and typical variance of environmental variables in eco-evolutionary research.

Chapter 4 highlighted one physiological trade-off with performance related to temperature, but it stands to reason that most trade-offs experienced by organisms in the wild are not always simple one-to-one relationships. Furthermore, it is important to consider the way trade-offs develop over the entirety of an organism's thermal performance curve (Taylor et al., 2020). Future work following chapter 4 could address both of these issues with mutlifactoral experimental designs across temperature treatments. One area worthy of future research is in understanding how trade-offs themselves can evolve. For instance, repeating similar methods from chapter 4 with wood frog populations from across

the range could uncover the extent to which thermal performance is more or less conserved by physiological trade-offs or phylogeny (Bodensteiner et al., 2021).

Although the common garden design used in chapter 5 mostly controls for plastic or parental effects that could facilitate such a pattern of intraspecific variation, they cannot be entirely precluded. Future work will involve experimentally manipulating genotypes and environments in the wild to test for evolutionary responses allowing for strict causal inference. An additional next step is to interrogate the genomes of these frogs to uncover nonrandom patterns of gene flow, and perhaps, the genetic basis of intrinsic developmental rates.

In conclusion, the work included in this dissertation illustrates that we need to understand fine-scale spatial eco-evo dynamics in order to understand the future of biodiversity.

REFERENCES

Anderson, S. C., Branch, T. A., Cooper, A. B., and Dulvy, N. K. (2017). Black-swan events in animal populations. *Proc. Natl. Acad. Sci. U. S. A.* 114, 3252–3257.

Bell, G. (2013). Evolutionary rescue and the limits of adaptation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368, 20120080.

Bodensteiner, B. L., Agudelo-Cantero, G. A., Arietta, A. Z. A., Gunderson, A. R., Muñoz, M. M., Refsnider, J. M., et al. (2021). Thermal adaptation revisited: How conserved are thermal traits of reptiles and amphibians? *J Exp Zool A Ecol Integr Physiol* 335, 173–194.

Cohen, J. M., Lajeunesse, M. J., and Rohr, J. R. (2018). A global synthesis of animal phenological responses to climate change. *Nat. Clim. Chang.* 8, 224–228.

De Meester, L., Stoks, R., and Brans, K. I. (2018). Genetic adaptation as a biological buffer against climate change: Potential and limitations. *Integr. Zool.* 13, 372–391.

Denver, R. J. (1997). Proximate Mechanisms of Phenotypic Plasticity in Amphibian Metamorphosis. *Integr. Comp. Biol.* 37, 172–184.

Dodd, C. K., Jr (2013). *Frogs of the United States and Canada, 2-vol. set*. The Johns Hopkins University Press.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends Ecol. Evol.* 17, 183–189.

Richardson, J. L., Urban, M. C., Bolnick, D. I., and Skelly, D. K. (2014). Microgeographic adaptation and the spatial scale of evolution. *Trends Ecol. Evol.* 29, 165–176.

Taylor, E. N., Diele-Viegas, L. M., Gangloff, E. J., Hall, J. M., Halpern, B., Massey, M. D., et al. (2020). The thermal ecology and physiology of reptiles and amphibians: A user's guide. *J Exp Zool A Ecol Integr Physiol*. doi:10.1002/jez.2396.

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J.-B., Pe'er, G., Singer, A., et al. (2016). Improving the forecast for biodiversity under climate change. *Science* 353. doi:10.1126/science.aad8466.