

USING CO-LOCATED LAKE AND BOG RECORDS TO  
IMPROVE INFERENCES ON LATE QUATERNARY CLIMATE  
AND ECOLOGY

by  
Connor J. Nolan

---

Copyright © Connor J. Nolan 2019

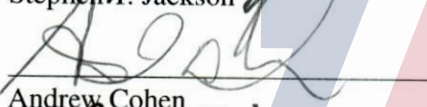
A Dissertation Submitted to the Faculty of the  
DEPARTMENT OF GEOSCIENCES  
In Partial Fulfillment of the Requirements  
For the Degree of  
DOCTOR OF PHILOSOPHY  
In the Graduate College  
THE UNIVERSITY OF ARIZONA

2019

THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Connor J. Nolan, titled Using co-located lake and bog records to improve inferences on late Quaternary climate and ecology and recommend that it be accepted as fulfilling the dissertation requirement for the degree of Doctor of Philosophy.

  
\_\_\_\_\_  
Stephen T. Jackson Date: April, 15, 2019


  
\_\_\_\_\_  
Andrew Cohen Date: April, 15, 2019

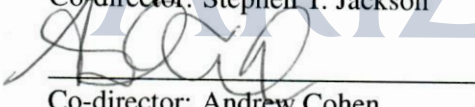
  
\_\_\_\_\_  
Jessica Tierney Date: April, 15, 2019

  
\_\_\_\_\_  
Valerie Trouet Date: April, 15, 2019

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

  
\_\_\_\_\_  
Co-director: Stephen T. Jackson Date: April 15, 2019

  
\_\_\_\_\_  
Co-director: Andrew Cohen Date: April 15, 2019

## ACKNOWLEDGMENTS

This dissertation would not have been possible without a huge group of helpful and encouraging people. First and foremost, thank you to my advisor, Steve Jackson, for great guidance and opportunities over these years. Thanks to Bryan Shuman and Bob Booth for hosting me in their labs and teaching me about lakes and bogs. Thanks to John Tipton for statistical wizardry and for always taking the time to Skype with me. Thank you to my committee members Andy Cohen, Jess Tierney, and Valerie Trouet. Thank you to all of the professors I have worked with or taken classes with at the University of Arizona including Jonathan Overpeck, Kevin Anchukaitis, David Breshears, Connie Woodhouse, Julie Cole, Katie Hirschboek, Don Falk, Joellen Russell, and Karl Flessa.

Thank you to my lab mates: Yao Liu, Teresa Krause, and Matt King.

Thank you to Nick McKay, Neil Pederson, Jenn Marlon, Britta Jensen, and Brian Chase for being great collaborators.

Thanks to PAEON leaders Jason McLachlan, Jack Williams, Dave Moore, Mike Dietze, Amy Hessel, and Jody Peters and fellow PAEON students and postdocs especially Ann Raiho, Dan Bishop, Malcolm Itter, Alex Dye, Simon Goring, and Andria Dawson.

Thank you to my undergraduate research advisors Larry Halverson and Carol and David Vleck for nurturing my passion for science. Thanks to Alan Wanamaker for introducing me to the field of paleoclimate.

Thank you fellow UA graduate students and postdocs: Zack Williams, Erin Harris-Parks, Jenna Shelton, Ted Cross, Jessie Pearl, Dan Griffin, Shelby Rader, Becki Beadling, Kristin Neff, Paul Goddard, Luke Parsons, Garrison Loope, Rachel Cajigas, Matt Meko, Talia Anderson, Haiyang Kehoe, Dervla Meegan Kumar, Audrey Dunham, Elizabeth Patterson, Grace Windler, Jonathan King, Alice Chapman, Sean O'Neal, Carson Richardson, Michael Kassela, Amy Hudson, Mallory Barnes, Brewster Malevich, Meg Mills-Novoa, Arica Crootof, Ethan Smith, and Matt Roby.

Thanks to my favorite people to hang out with at conferences: Jeremy Hoffman, Gaylen Sinclair, Shane Loeffler, Sylvia Dee, Kaustubh Thirumalai, and Sloan Coats.

Thank you Jessica Ray for all the coffee walks and Marty Pepper for the random conversations in the halls of Gould-Simpson.

Thanks to my great friends Chris Brus, Roy Grieg, Riley Millington, Ben Brus, Emily Mason, and Naty Krantz. Thank you Shayna Rosenblum.

Thank you Brandon Swift and Kelly Cruz for giving me a place to live.

Thanks GYM 244, especially the Burpee Boys (Reid Strickler, Manny Villegas, and Mark Anderson), Chris Gartrell and Leslie Johnson.

Thank you to my family: my parents, Becky and Art Nolan, and my brother, Dr. Trevor Nolan, and his family Breanna, Ady, and Hudson.

## DEDICATION

This dissertation is dedicated to my mom and dad for their unyielding support and enthusiasm.

## TABLE OF CONTENTS

LIST OF FIGURES . . . . .	8
LIST OF TABLES . . . . .	12
ABSTRACT . . . . .	13
1. INTRODUCTION . . . . .	15
2. PRESENT STUDY . . . . .	19
2.1. Comparing and improving methods for reconstructing peatland water-table depth from testate amoebae . . . . .	19
2.2. Using co-located lake-level and bog water-table depth records to understand Holocene climate and vegetation changes in Maine . . . . .	20
2.3. Coherent mid-to-late Holocene hydrologic variability across temporal scales in mid-latitude Eastern North America . . . . .	22
REFERENCES . . . . .	23
A. APPENDIX A . . . . .	27
A.1. Abstract . . . . .	28
A.2. Introduction . . . . .	29
A.3. Data and methods . . . . .	33
A.3.1. Modern surface sample training dataset . . . . .	33
A.3.2. Site Description and Lab Methods: Reconstruction Dataset	34
A.3.3. Models . . . . .	37
A.3.4. Cross-validation and scoring rules . . . . .	48
A.3.5. Computing methods . . . . .	50
A.4. Results and discussion . . . . .	50
A.4.1. Fits to training set . . . . .	50
A.4.2. Cross-validation experiment . . . . .	52
A.4.3. Will skill in cross-validation generalize to novel data? . .	54
A.4.4. Reconstruction . . . . .	57
A.4.5. Inference on past hydrology . . . . .	60
A.5. Conclusions and Future Directions: Or, Why Bayesian? Why Bother?	62
A.6. Acknowledgements . . . . .	64
REFERENCES . . . . .	65
B. APPENDIX B . . . . .	72

TABLE OF CONTENTS—*Continued*

C. APPENDIX C . . . . .	77
C.1. Abstract . . . . .	78
C.2. Introduction . . . . .	79
C.3. Sites and methods . . . . .	83
C.3.1. Giles Pond . . . . .	83
C.3.2. Caribou Bog . . . . .	87
C.3.3. Combining lake-level and water-table depth reconstructions . . . . .	90
C.4. Results . . . . .	92
C.4.1. Lake-level reconstruction . . . . .	92
C.4.2. Caribou Bog hydrology and disturbance records . . . . .	94
C.4.3. Integrated lake and bog paleohydrology . . . . .	95
C.4.4. Caribou Bog pollen record . . . . .	96
C.5. Discussion . . . . .	99
C.5.1. Long-term hydrologic trends and multi-centennial variability at Giles Pond . . . . .	99
C.5.2. High-frequency hydrologic variability and impacts at Caribou Bog . . . . .	102
C.5.3. Integrating lake-level and water-table depth records to reconstruct hydrologic variability across temporal scales . . . . .	103
C.5.4. Climate and ecological drivers of Holocene vegetation change at Caribou Bog . . . . .	104
C.6. Conclusions: What’s gained by having all of these records together? . . . . .	108
C.7. Acknowledgments . . . . .	109
REFERENCES . . . . .	110
D. APPENDIX D . . . . .	117
E. APPENDIX E . . . . .	120
E.1. Abstract . . . . .	121
E.2. Introduction . . . . .	122
E.3. Sites and Methods . . . . .	126
E.3.1. Knuckey Lake . . . . .	126
E.3.2. Ely Lake Bog . . . . .	127
E.3.3. TraCE-21k climate model analyses . . . . .	129
E.4. Results . . . . .	129
E.4.1. New Records . . . . .	129
E.4.2. New records in Minnesota context . . . . .	132

TABLE OF CONTENTS—*Continued*

E.4.3.	Bog water-table depth records from the eastern United States . . . . .	134
E.4.4.	Continental-scale lake-level synthesis . . . . .	139
E.4.5.	Multi-centennial precipitation variability in the TraCE model simulation . . . . .	140
E.5.	Discussion . . . . .	141
E.5.1.	Multi-scale Holocene hydrologic variability in Minnesota . . . . .	141
E.5.2.	Bog water-table depth records resolve regional hydroclimatic variability . . . . .	143
E.5.3.	Coherent multi-centennial lake-level variability across mid-latitude eastern North America . . . . .	143
E.6.	Acknowledgments . . . . .	145
REFERENCES	. . . . .	146
F. APPENDIX F	. . . . .	152

## LIST OF FIGURES

- FIGURE A.1. Map showing the sites that make up the modern training dataset sites (blue circles) and the site of the new water-table depth reconstruction, Caribou Bog (orange diamond). Panel A shows the continental USA, Panel B shows the training dataset sites in Alaska, and Panel C zooms in on Maine and the surrounding area. . . . . 35
- FIGURE A.2. Species abundance versus water-table depth (in cm) from the samples in the training dataset (points). The smooth curves are from Multivariate Gaussian Process for each species. The solid vertical lines are the raw optima from weighted averaging (WA). The dashed vertical lines are the deshrunk WA optima. The dot-dash vertical lines are the coefficients from weighted averaging of partial least squares after five partial least squares components. . . . . 51
- FIGURE A.3. Observed versus predicted water-table depth (cm) for each model from the 10-fold cross-validation experiment on the training dataset of 951 surface testate amoebae assemblages. Red lines in each panel indicate the one-to-one line. Blue lines in each panel are a linear fit of the observed vs. predicted water-table depths. Error bars denote the 95% credible interval . . . 53
- FIGURE A.4. A. Squared chord distances of the accepted analogs from Modern Analog Technique (MAT) for the training dataset (red) and reconstruction dataset from Caribou Bog (blue). Smaller squared chord distance corresponds to more similar assemblages. The training dataset contains many good analogs for the assemblages within the training dataset (distances for the red density curve are small) whereas the distances of the analog assemblages for the reconstruction dataset are larger and thus not as similar. B. Time series of the median squared chord distance from the accepted analog assemblages in the training dataset to each assemblage in the Caribou Bog reconstruction dataset. This panel shows temporal evolution of the analogs summarized by the blue density curve in panel A. . . . . 56
- FIGURE A.5. Reconstructions of water-table depth based on testate amoebae assemblages from Caribou Bog. For each panel, the underlying testate amoebae assemblage data is identical the only thing that changes is the transfer function model. The dark shading denotes 50% credible interval, the lighter shading denotes 95% credible interval. . . . . 58
- FIGURE B.1. Age-depth model for Caribou Bog . . . . . 74



## LIST OF FIGURES—Continued

- FIGURE B.2. Full testate amoebae data for Caribou Bog. The y-axis is calibrated age. Each x-axis is percent abundance. The species names abbreviations are as follows: ampfla = *Amphitrema flagellum*, ampwri = *A. wrightianum*, arccat = *Arcella catinus*, arcdis = *A. discoides*, arcvul = *A. vulgaris*, assmus = *Assulina muscorum*, asssem = *A. seminulum*, bulind = *Bulinaria indica*, cenacu = *Centropyxis aculeata*, cencas = *C. cassis*, ceneco = *C. ecornis*, cortri = *Corythion trinema*, cryovi = *Cryptodiffugia oviformis*, cycarc = *Cyclopyxis arcelloides*, difacu = *Diffugia acuminata*, difglo = *D. globulosa*, difluc = *D. lucida*, difobl = *D. oblonga*, eugstr = *Euglypha strigosa*, eugtub = *E. tubercleata*, helpet = *Heleopera petricola*, helsph = *H. sphagni*, helsyl = *H. sylvatica*, hyaele = *Hyalosphenia elegans*, hyapap = *H. papillio*, hyasub = *H. subflava*, nebcar = *Nebela carinata*, nebcoll = *N. collaris-bohemica*-type, nebgri = *N. griseola*, nebmil = *N. militaris*, nebtin = *N. tinctoria*, nebvitr = *N. vitraea*, phracr = *Phryganella acropodia*, plaspis = *Placcista spinosa*, pseful = *Pseudodiffugia fulva*, sphlen = *Sphenoderia lenta*, triarc = *Trigonopyxis arculeata* . . . . . 75
- FIGURE B.3. Z-scores of the water-table depth reconstructions resulting from the seven transfer function models applied to the testate amoeba data from Caribou Bog . . . . . 76
- FIGURE C.1. A. Map of sites used in this paper. Pollen data were used from Mansell Pond (very near Caribou Bog), Moulton Pond (south of Caribou Bog), Gould Pond (west of Caribou Bog), and Loon Pond (east of Caribou Bog). GGC30 is a marine core off the Scotian Margin with a published alkenone SST record (Sachs, 2007). Lake-level (LL) sites are shown as blue dots. B. Google Earth view of Giles Pond with approximate coring transect shown as a white line. B. Google Earth view of Caribou Bog with approximate coring location shown as a white square. . . . . 84
- FIGURE C.2. A. Titanium count data for Giles Pond cores 31A (thin black line), 39A (black line), 50A (gray line). B. Lake-level reconstruction from Giles Pond cores. Dark blue line is the reconstructed lake-level, light blue shading is the uncertainty, gray lines are some sample lake-level trajectories from the algorithm, black lines are the cores sedimentation history, orange dots are intervals in that core classified as littoral, tan dots are intervals classified as sub-littoral. C. Z-scores of lake-level reconstructions (offset for clarity) from New Long Pond (NLP) (Newby et al., 2009), Deep Pond (Marsicek et al., 2013), Davis Pond (Newby et al., 2011), and Giles Pond (this study). . . . . 91
- FIGURE C.3. A. Testate amoeba-inferred water-table depth from Caribou Bog with the smooth spline showing the low-frequency variability. B. The detrended water-table depth from Caribou Bog. C. The micro-charcoal from Caribou Bog. 94

LIST OF FIGURES—*Continued*

- FIGURE C.4. Density plot of the detrended water-table depth (higher values = drier conditions) for samples with high charcoal influx and samples with low charcoal influx. The difference between the distributions is statistically significant ( $P < 0.001$ ; Mann-Whitney rank-sum test). . . . . 96
- FIGURE C.5. A: Lake-level record from Giles Pond converted to z-scores and plotted with the fitted insolation curve. B: Detrended water-table depth record from Caribou Bog converted to z-scores. C: Integrated lake-level and bog water-table depth paleohydrology record created by adding the z-scores from panel A and B. . . . . 97
- FIGURE C.6. A: Integrated record that combines the Giles Pond lake-level history and Caribou Bog detrended water-table depth record (same as figure C.5C). B: Microcharcoal record from Caribou Bog. C: *Pinus* pollen from Caribou Bog. D: *Tsuga* pollen from Caribou Bog. E: *Betula* pollen from Caribou Bog. F: *Picea* pollen from Caribou Bog. G: *Fagus* pollen from Caribou Bog. H: *Quercus* pollen from Caribou Bog. . . . . 100
- FIGURE D.1. Ti counts from XRF versus depth for the three Giles Pond cores used in lake-level reconstruction annotated with the median calibrated radiocarbon dates. . . . . 119
- FIGURE E.1. Mid-latitude North America sites developed and used in this study. KL = Knuckey Lake, HB = Hole Bog, SL = Steel Lake, WOL = West Olaf Lake, SRB = South Rhody Bog, ISB = Irwin Smith Bog, MB = Minden Bog, PB = Pinhook Bog, CB = Caribou Bog, GP = Giles Pond, SiB = Sidney Bog, SaB = Saco Bog, GHB = Great Heath Bog, NLP = New Long Pond, DaP = Davis Pond, DeP = Deep Pond, BL = Blanding Lake. . . . . 126
- FIGURE E.2. A. Loss-on-ignition data from Knuckey Lake core 16A (thinnest line), core 19A (gray line), and core 22A (black line). B. Reconstructed lake level from Knuckey Lake (black line) with uncertainty (gray shading). The core data and accumulation history are shown as lines. Black dots represent an interval classified as littoral, gray dots represent an interval classified as sublittoral. C. Ely Lake Bog testate amoeba-based water-table depth reconstruction. D. Microcharcoal record from Ely Lake Bog. . . . . 130
- FIGURE E.3. Records of hydroclimate from Minnesota over the Holocene. A: Testate amoebae-inferred water-table depth from Ely Lake Bog. B: Testate amoebae-inferred water-table depth Hole Bog (Booth et al., 2006). C. Lake-level record from Knuckey Lake. D. Quartz record, a proxy for dustiness, from Elk Lake (Dean, 1993). E: Aragonite:Calcite ratio from West Olaf Lake (Nelson and Hu, 2008). F: Pine (shaded) and Oak (black line) pollen from Steel Lake (Nelson et al., 2004). . . . . 131

LIST OF FIGURES—*Continued*

- FIGURE E.4. Last 2000 years of hydrologic variability from Knuckey Lake and a network of bog water-table depths from the eastern United States. A: Lake-level reconstruction from Knuckey Lake. B: Water-table depth reconstructions from Ely Lake Bog (solid line; this study) and Hole Bog (dashed line; Booth et al. (2006)). C: Water-table depth reconstructions from Minden Bog (solid line; Booth et al. (2006)) and Irwin Smith Bog (dashed line; Booth et al. (2012)). D: Water-table depth reconstructions from Pinhook Bog (solid line; Booth et al. (2012)) and South Rhody Bog (dashed line; Booth et al. (2012)). E. Water-table depth reconstructions from Sidney Bog (solid line; Clifford and Booth (2013)), Great Heath Bog (dashed line; Clifford and Booth (2013)), Saco Bog (blue line; Clifford and Booth (2013)), and Caribou Bog (red line; Nolan, et al. 2019). . . . . 133
- FIGURE E.5. Z-Scores of lake-level records from the eastern US over the past 6200 years. The vertical lines indicate the periods of coherent multi-centennial variability among Massachusetts lake-level records identified by Newby et al. (2014). 136
- FIGURE E.6. Top: Mid-latitude North American lake-level records over the last 6200 years converted to z-scores and detrended. Bottom: Principal components of the detrended z-scores of the eastern US lake-level records. PC1 explains 43.5% of the variance, PC2 explains 21.1%, and PC3 explains 14.8%. 137
- FIGURE E.7. A. Smoothed z-scores of total precipitation for the grid points in the TraCE-21ka simulation with lake-level records used in this study. B. First principal component (59% of the variance) of the TraCE-21ka precipitation data for the four grid points in A. Second principal component (22% of the variance) of the TraCE-21ka precipitation data for the four grid points in A. C. Third principal component (16% of the variance) of the TraCE-21ka precipitation data for the four grid points in A. . . . . 138
- FIGURE F.1. A. First principal component (53% variance) of Upper Midwest bog records: Ely Lake Bog, Hole Bog, and Minden Bog. B. First principal component of two PCAs of Maine bog records. The PCA for present to 600 yr BP uses all four Maine bogs and the first PC represents 56% of the variance. The PCA for 1550 to 2000 yr BP uses Saco Bog, Sidney Bog, and Caribou Bog. The first PC represents 55% of the variance. . . . . 153
- FIGURE F.2. EOF loadings for the six lake-level records from the eastern US. . . . 153
- FIGURE F.3. TraCE-21ka total precipitation for the grid points that contain Giles Pond (A) and Knuckey Lake (C) and smoothed with a smooth spline with 40 degrees of freedom (B and D). . . . . 154
- FIGURE F.4. TraCE-21ka total precipitation for the grid points that contain Deep Pond (A) and Blanding Pond (C) and smoothed with a smooth spline with 40 degrees of freedom (B and D). . . . . 155

## LIST OF TABLES

TABLE A.1. Scoring results from 10-fold cross-validation experiment performed on the modern training dataset. Lower scores for MSPE, MAE, and CRPS indicate better performance. For coverage, better performance is indicated by values closer to the nominal rate (in this case, 95%). . . . .	54
TABLE B.1. Radiocarbon dates for Caribou Bog . . . . .	73
TABLE D.1. Radiocarbon dates for Giles Pond. . . . .	118
TABLE F.1. Radiocarbon dates for Knuckey Lake. . . . .	156
TABLE F.2. Radiocarbon dates for Ely Lake Bog. . . . .	156

## ABSTRACT

Records of past climate and ecological dynamics from lakes and bogs provide a long-term perspective that extends the modern observational record. Many different paleoclimatic and paleoecological proxies are used to characterize patterns, processes, and impacts of past climate variability, each with unique strengths and weaknesses. Integrating data across a network of different proxies requires a detailed understanding of each individual proxy. In the eastern United States two common proxies for Holocene hydroclimate conditions are testate amoeba-based water-table depth reconstructions and sedimentary lake-level reconstructions. Testate amoebae tests are preserved in sediments and can be identified to the species level. The community composition of testate amoebae on the surface of an ombrotrophic bog is sensitive to the surface moisture of the bog. Past fluctuations in lake level can be tracked by the changing elevation of the sand-mud boundary along a transect of near-shore sediment cores. Lake-level records have been developed mainly in southern New England and in the Rocky Mountains. Bog water-table depth records have been developed mainly in Maine and the Upper Midwest. For my dissertation, I have developed two pairs of co-located lake-level and bog water-table depth records from Maine and Minnesota. Co-located records allow direct comparison of proxy records that experienced the same past climate conditions. In the first chapter, I present a new testate amoeba record from Caribou Bog and a new Bayesian transfer function model that estimates water-table depth from testate amoeba assemblages. I compare seven different methods for estimating

water-table depth from testate amoebae in cross-validation and in reconstruction. In the second chapter, I present a suite of records from Caribou Bog and Giles Pond, both from south-central Maine. I validate and compare the hydrologic reconstructions from Caribou Bog and Giles Pond. Then, I leverage the strengths of each proxy reconstruction to produce an integrated record of past hydrology that captures hydroclimate variability and change from decadal to multi-millennial timescales. Finally, I use the integrated record of past hydroclimate to understand vegetation changes in a new pollen record from Caribou Bog. Holocene vegetation changes represent a complex interaction between large-scale biogeographic patterns, regional and local climate variability and change, and local ecological variability. In the final chapter, I introduce new records from Ely Lake Bog and Knuckey Lake in northeastern Minnesota. I analyze the new records in the context of mid-to-late Holocene changes in the Upper Midwest and mid-latitude North America. I find consistent sub-regional drought events over the past 2000 years and coherent multi-centennial variability in the eastern United States over the past 6,200 years. This overall combination of formal statistical modeling, detailed site-level reconstruction, and single-proxy network analyses represents ways to maximize the utility of proxy records for inferring past environmental variability and change.

## 1. INTRODUCTION

Paleoenvironmental reconstruction documents the patterns, processes, and impacts of past climate variability and change and represents a long-term historical perspective that extends the short observational record and provides empirical targets for modeling (Masson-Delmotte et al., 2013). Many different approaches are used to improve paleoclimate inference, including continental-to-global syntheses (PAGES, 2013; Marcott et al., 2013; PAGES2k, 2017; Cook, 2004; Cook et al., 2010), proxy system modeling (Evans et al., 2013), data-model intercomparisons (DiNezio and Tierney, 2013; Abram et al., 2016; Coats et al., 2016), and data assimilation (Hakim et al., 2016).

Proxies that cannot be directly calibrated to modern instrumental climate records have generally not been included in these efforts. This dissertation focuses on two such proxies: lake-level records based on tracking the sand-mud boundary from sedimentary evidence in a transect of near-shore cores (Digerfeldt, 1986; Pribyl and Shuman, 2014) and bog water-table depth records inferred from testate amoebae buried in peat (Woodland et al., 1998; Booth, 2010).

Lakes are powerful archives that can be used to reconstruct many paleoenvironmental variables (Cohen, 2003). Changes in lake level can be reconstructed by a multitude approaches (e.g., identification of paleoshorelines (McGee et al., 2018), biotic indicators such as the ratio of planktic to benthic diatoms (Stone and Fritz, 2004) aquatic plant macrofossils

Birks (2002), and stable isotope evidence (Yu et al., 1997; Gao et al., 2017)). Here I focus on using sedimentary evidence from a transect of near-shore cores to reconstruct the past position of the littoral-profundal boundary (Digerfeldt, 1986; Shuman et al., 2001; Pribyl and Shuman, 2014).

This lake-level reconstruction approach is best applied to small lakes (<100 ha) with shallowly sloping bathymetry, minimal surface inflow and outflow, and a well-defined drainage basin (Pribyl and Shuman, 2014). Target sites for coring are identified by ground penetrating radar (GPR) surveys to identify a pattern of reflective sand layers within the basin consistent with lake-level change. Features identified by GPR are targeted by a transect of near-shore sediment cores. Sand layers are identified in the cores visually, quantified using diagnostic data such as loss-on-ignition, grain size, or x-ray fluorescence (XRF), and dated via radiocarbon (Shuman et al., 2009; Pribyl and Shuman, 2014). All of this data is used iteratively in a decision-tree algorithm to quantitatively reconstruct the past elevation of the littoral-profundal boundary (Marsicek et al., 2013; Pribyl and Shuman, 2014).

Peatland archives also yield many different paleoenvironmental proxies (e.g., sphagnum-to-vascular ratio, humification, stable isotopes, and macrofossils (Chambers et al., 2012)). Here I use community composition of testate amoebae to infer changes in water-table depth. Testate amoebae are a polyphyletic group of protists that live on the surface of bogs and other wetland environments. Testate amoebae produce diverse tests (shells) from proteinaeous, calcareous, or siliceous material that are generally well-preserved in peat sediments (Charman, 2001; Mitchell et al., 2008). Changes in testate amoebae community compo-



sition are related to changes in bog surface wetness (Warner and Charman, 1994; Booth, 2008, 2010).

Transfer functions, developed from modern training samples where both the community composition and the water-table depth are measured, are used to quantitatively infer changes in bog-surface wetness from subfossil testate-amoebae assemblages (Birks et al., 1990; Juggins and Birks, 2012). Reconstructed water-table depth is sensitive to changes in summer moisture deficit (Charman, 2007; Charman et al., 2009; Booth, 2010). Bog water-table depth records skillfully register extreme hydroclimate events at decadal to multi-decadal time scales (Morris et al., 2015; Clifford and Booth, 2015), but long-term trends and low-frequency variability can be influenced by autogenic bog processes (Charman et al., 2006; Swindles et al., 2012; Morris et al., 2015).

Lake-level and bog water-table depth records are common in North America, but they have never been developed at co-located sites for comparison. Lake-level reconstructions have been developed in New England and the Rocky Mountains (Newby et al., 2014; Shuman and Burrell, 2017; Shuman and Serravezza, 2017), whereas bog water-table depth records have been developed in Maine, Michigan, Minnesota, and Wisconsin (Booth et al., 2006, 2012; Clifford and Booth, 2013). The lack of co-located records has limited the direct comparison of lake-level and bog water-table depth reconstruction. In order to compare paleoenvironmental inferences from lake-level and bog water-table depth records subject to the same past climate variability and change, I developed two new pairs of co-located lake and bog records from south-central Maine and northern Minnesota, where both suitable

lakes and bogs occur.

I use several approaches to understand and improve inferences from lake-level and bog water-table depth records. First, I compare a new Bayesian model for reconstructing water-table depth from testate amoebae to 6 other transfer function methods. Bayesian statistical modeling provides an improved means of calibrating and quantitatively integrating bog records with other proxies (e.g., in a data assimilation framework), but there are inferential improvements that can be made without full Bayesian models. In the second chapter, I present a suite of records from Maine including water-table depth history from Caribou Bog, lake-level reconstruction from Giles Pond, and pollen and charcoal from Caribou Bog. These records overlap for the past 7,500 years and thus allow for detailed investigation of local to regional climate and ecological change. I apply expert knowledge of the proxy records to combine them in a way that emphasizes signals that are more confidently related to climate variability. I produce an integrated record of hydroclimate variability that I use to interpret the climate drivers of vegetation change. In the third chapter, I interpret new lake and bog records from Minnesota in the context of networks of existing lake and bog records. I find coherent multi-centennial variability across the eastern United States over the past 6,200 years. Using single-proxy networks allows for regional comparison and synthesis without explicit calibration. Altogether, this dissertation represents efforts to maximize the utility of proxy records to yield a more complete understanding of patterns, processes, and impacts of Holocene climate variability in eastern North America.

## 2. PRESENT STUDY

The dissertation is organized into three scientific papers included as appendices. The first paper, "Comparing and improving methods for reconstructing peatland water-table depth from testate amoebae," is accepted for publication in *The Holocene*. The second paper, "Using co-located lake-level and bog water-table depth records to understand Holocene climate and vegetation changes in Maine," is prepared for submission to *Quaternary Science Reviews*. The final paper, "Coherent mid-to-late Holocene hydrologic variability across temporal scales in mid-latitude Eastern North America," is prepared for submission to *Earth and Planetary Science Letters*. Each paper is co-authored, but I am the lead and corresponding author. I led the research design, performed the analyses, and wrote the manuscripts. This chapter provides a brief overview of each paper.

### 2.1. Comparing and improving methods for reconstructing peatland water-table depth from testate amoebae

Paper one presents a new Bayesian method for reconstructing peatland water-table depth from testate amoebae. I compare the performance of a new Bayesian method to six existing transfer function methods in cross-validation and in down-core reconstruction using a new testate amoebae record from Caribou Bog in Maine.

In cross-validation, the modern analog technique (MAT) performed best, followed by

weighted averaging (WA), weighted averaging of partial least squares (WAPLS), and the Bayesian method. MAT is not widely used for testate amoebae due to issues with down-core analog quality, related to differential preservation of some taxa. WA and WAPLS are commonly used and this analysis suggests they are reasonable models to use. The Bayesian method performs well and more fully accounts for uncertainties in reconstruction. Bayesian methods will be useful for future work integrating reconstructions across a network of testate amoeba-based water-table depth records, quantitative calibration of reconstructed water-table depths, and use in a data assimilation framework.

The reconstructions from Caribou Bog that result from the seven different transfer functions are broadly consistent, but they diverge in some details. Without an independent constraint, I cannot conclude that one model is "right" and another is wrong, but investigating the causes of differences between reconstructions in the underlying data yields testable ecological hypotheses and insights into transfer function sensitivities. Applying a range of different methods in reconstruction illustrates model choice as a source of uncertainty in paleoenvironmental proxies based on compositional data.

## 2.2. Using co-located lake-level and bog water-table depth records to understand Holocene climate and vegetation changes in Maine

The second paper presents and validates new co-located lake-level and bog water-table depth records from Maine. I compare the new reconstructions and combine them to produce

an integrated reconstruction of past hydroclimate variability and change. Finally, I use the integrated reconstruction to interpret a new pollen record and investigate the climate drivers of Holocene vegetation change in northern New England.

The lake-level record from Giles Pond shows an insolation-driven long-term trend toward wetter conditions with multi-centennial variability consistent with existing lake-level records from southern New England. The water-table depth record from Caribou Bog exhibits characteristic decadal-to-multidecadal variability superimposed on low-frequency trends that, in bog records, are often related to non-climatic processes. I removed the low-frequency variability to generate a detrended record that emphasizes decadal-to-multidecadal water-table depth variability. I smooth the detrended water-table depth record and compare it to the detrended lake-level record and find general agreement, suggesting these two diverse proxies sense similar latent climate variability. I combine the records by adding the z-scores of the lake-level and detrended water-table depth records together. The combined hydrologic record leverages the strengths of each record for a more complete understanding of past hydroclimate variability. The integrated record explains decadal to multicentennial variability in my new pollen record. Overall, this paper illustrates the utility of developing and combining multiple proxy records to better understand Holocene hydroclimate changes and their ecological implications

### 2.3. Coherent mid-to-late Holocene hydrologic variability across temporal scales in mid-latitude Eastern North America

The third paper introduces two new records from northeastern Minnesota, a 6,200-year lake level reconstruction from Knuckey Lake and a 1,000 year water-table depth record from 1500-500 cal yr BP from Ely Lake Bog. I introduce each record in the context of Minnesota paleoclimate and paleoecology, and then broaden the scope to compare the new records to the networks of existing bog water-table depth and lake-level records from mid-latitude North America.

Sub-regional coherency is dominant in the network of bog water-table depth records over the past 2000 years. Records from the Upper Midwest show major dry events at 1000 and 750 cal yr BP, whereas the records from from Maine agree on a dry event 500 years ago. In the lake-level network, there is coherent multi-centennial variability between lake-level reconstructions across the eastern United States over the past 6,200 years. This paper highlights the utility of consistent reconstruction methodologies and single-proxy networks to control for the complex responses of Holocene terrestrial hydroclimate proxies.

## REFERENCES

- Abram NJ, McGregor HV, Tierney JE, Evans MN, McKay NP, Kaufman DS, Thirumalai K, Martrat B, Goosse H, Phipps SJ et al. (2016) Early onset of industrial-era warming across the oceans and continents. *Nature* 536(7617): 411.
- Birks H, Ter Braak C, Line J, Juggins S and Stevenson A (1990) Diatoms and pH reconstruction. *Phil. Trans. R. Soc. Lond. B* 327(1240): 263–278.
- Birks HH (2002) Plant macrofossils. In: *Tracking environmental change using lake sediments*. Springer, pp. 49–74.
- Booth RK (2008) Testate amoebae as proxies for mean annual water-table depth in Sphagnum-dominated peatlands of North America. *Journal of Quaternary Science* 23(1): 43–57.
- Booth RK (2010) Testing the climate sensitivity of peat-based paleoclimate reconstructions in mid-continental North America. *Quaternary Science Reviews* 29(5-6): 720–731.
- Booth RK, Jackson ST, Sousa VA, Sullivan ME, Minckley TA and Clifford MJ (2012) Multi-decadal drought and amplified moisture variability drove rapid forest community change in a humid region. *Ecology* 93(2): 219–226.
- Booth RK, Notaro M, Jackson ST and Kutzbach JE (2006) Widespread drought episodes in the western Great Lakes region during the past 2000 years: geographic extent and potential mechanisms. *Earth and Planetary Science Letters* 242(3-4): 415–427.
- Chambers FM, Booth RK, De Vleeschouwer F, Lamentowicz M, Le Roux G, Mauquoy D, Nichols JE and Van Geel B (2012) Development and refinement of proxy-climate indicators from peats. *Quaternary International* 268: 21–33.
- Charman DJ (2001) Biostratigraphic and palaeoenvironmental applications of testate amoebae. *Quaternary Science Reviews* 20: 1753–1764.
- Charman DJ (2007) Summer water deficit variability controls on peatland water-table changes: implications for Holocene palaeoclimate reconstructions. *The Holocene* 17(2): 217–227.
- Charman DJ, Barber KE, Blaauw M, Langdon PG, Mauquoy D, Daley TJ, Hughes PD and Karofeld E (2009) Climate drivers for peatland palaeoclimate records. *Quaternary Science Reviews* 28(19-20): 1811–1819.
- Charman DJ, Blundell A, Chiverrell RC, Hendon D and Langdon PG (2006) Compilation of non-annually resolved Holocene proxy climate records: stacked holocene peatland

- palaeo-water table reconstructions from northern Britain. *Quaternary Science Reviews* 25(3-4): 336350.
- Clifford MJ and Booth RK (2013) Increased probability of fire during late Holocene droughts in northern New England. *Climatic change* 119(3-4): 693–704.
- Clifford MJ and Booth RK (2015) Late-Holocene drought and fire drove a widespread change in forest community composition in eastern North America. *The Holocene* 25(7): 1102–1110.
- Coats S, Smerdon JE, Cook BI, Seager R, Cook ER and Anchukaitis KJ (2016) Internal ocean-atmosphere variability drives megadroughts in western North America. *Geophysical Research Letters* 43(18): 98869894.
- Cohen AS (2003) *Paleolimnology: the history and evolution of lake systems*. Oxford University Press.
- Cook E, Anchukaitis K, Buckley B, D'Arrigo R, Jacoby G and Wright W (2010) Asian monsoon failure and megadrought during the last millennium. *Science* 328(5977): 486489.
- Cook ER (2004) Long-term aridity changes in the western United States. *Science* 306(5698): 10151018.
- Digerfeldt G (1986) Studies on past lake-level fluctuations. *Handbook of Holocene palaeoecology and palaeohydrology* 127: 143.
- DiNezio PN and Tierney JE (2013) The effect of sea level on glacial Indo-Pacific climate. *Nature Geoscience* 6(6): 485.
- Evans MN, Tolwinski-Ward SE, Thompson DM and Anchukaitis KJ (2013) Applications of proxy system modeling in high resolution paleoclimatology. *Quaternary science reviews* 76: 16–28.
- Gao L, Huang Y, Shuman B, Oswald WW and Foster D (2017) A high-resolution hydrogen isotope record of behenic acid for the past 16 kyr in the northeastern United States. *Quaternary International* 449: 1–11.
- Hakim GJ, Emile-Geay J, Steig EJ, Noone D, Anderson DM, Tardif R, Steiger N and Perkins WA (2016) The last millennium climate reanalysis project: Framework and first results. *Journal of Geophysical Research: Atmospheres* 121(12): 6745–6764.
- Juggins S and Birks HJB (2012) Quantitative environmental reconstructions from biological data. In: *Tracking environmental change using lake sediments*. Springer, pp. 431–494.



- Marcott SA, Shakun JD, Clark PU and Mix AC (2013) A reconstruction of regional and global temperature for the past 11,300 years. *science* 339(6124): 1198–1201.
- Marsicek JP, Shuman B, Brewer S, Foster DR and Oswald WW (2013) Moisture and temperature changes associated with the mid-Holocene Tsuga decline in the northeastern United States. *Quaternary Science Reviews* 80: 129–142.
- Masson-Delmotte V, Schulz M, Abe-Ouchi A, Beer J, Ganopolski A, Gonzalez Ruoco J, Jansen E, Lambeck K, Luterbacher J, Naish T, Osborn T, Otto-Bliesner B, Quinn T, Ramesh R, Rojas M, Shao X and Timmermann A (2013) Information from paleoclimate archives. In: Intergovernmental PoCC (ed.) *Climate Change 2013 - The Physical Science Basis*. Cambridge: Cambridge University Press, pp. 383–464.
- McGee D, Moreno-Chamarro E, Marshall J and Galbraith E (2018) Western us lake expansions during Heinrich stadials linked to pacific Hadley circulation. *Science advances* 4(11): eaav0118.
- Mitchell EA, Charman DJ and Warner BG (2008) Testate amoebae analysis in ecological and paleoecological studies of wetlands: past, present and future. *Biodiversity and Conservation* 17(9): 2115–2137.
- Morris PJ, Baird AJ, Young DM and Swindles GT (2015) Untangling climate signals from autogenic changes in long-term peatland development. *Geophysical Research Letters* 42(24): 10–788.
- Newby PE, Shuman BN, Donnelly JP, Karnauskas KB and Marsicek J (2014) Centennial-to-millennial hydrologic trends and variability along the North Atlantic Coast, USA, during the Holocene. *Geophysical Research Letters* 41(12): 4300–4307.
- PAGES kC (2013) Continental-scale temperature variability during the past two millennia. *Nature Geoscience* 6(5): 339346.
- PAGES2k C (2017) A global multiproxy database for temperature reconstructions of the common era. *Sci Data* 4: 170088.
- Pribyl P and Shuman BN (2014) A computational approach to Quaternary lake-level reconstruction applied in the central Rocky Mountains, Wyoming, USA. *Quaternary Research* 82(1): 249–259.
- Shuman B, Bravo J, Kaye J, Lynch JA, Newby P and Webb T (2001) Late Quaternary water-level variations and vegetation history at crooked pond, southeastern massachusetts. *Quaternary Research* 56(3): 401410.
- Shuman BN and Burrell SA (2017) Centennial to millennial hydroclimatic fluctuations in the humid northeast United States during the Holocene. *Quaternary Research* 88(3): 514–524.

- Shuman BN, Newby P and Donnelly JP (2009) Abrupt climate change as an important agent of ecological change in the Northeast US throughout the past 15,000 years. *Quaternary Science Reviews* 28(17-18): 1693–1709.
- Shuman BN and Serravezza M (2017) Patterns of hydroclimatic change in the Rocky Mountains and surrounding regions since the last glacial maximum. *Quaternary Science Reviews* 173: 58–77.
- Stone JR and Fritz SC (2004) Three-dimensional modeling of lacustrine diatom habitat areas: Improving paleolimnological interpretation of planktic: benthic ratios. *Limnology and Oceanography* 49(5): 1540–1548.
- Swindles GT, Morris PJ, Baird AJ, Blaauw M and Plunkett G (2012) Ecohydrological feedbacks confound peat-based climate reconstructions. *Geophysical Research Letters* 39(11).
- Warner BG and Charman DJ (1994) Holocene changes on a peatland in northwestern Ontario interpreted from testate amoebae (Protozoa) analysis. *Boreas* 23(3): 270–279.
- Woodland WA, Charman DJ and Sims PC (1998) Quantitative estimates of water tables and soil moisture in Holocene peatlands from testate amoebae. *The Holocene* 8(3): 261–273.
- Yu Z, McAndrews JH and Eicher U (1997) Middle Holocene dry climate caused by change in atmospheric circulation patterns: evidence from lake levels and stable isotopes. *Geology* 25(3): 251–254.

## A. APPENDIX A

COMPARING AND IMPROVING METHODS FOR RECONSTRUCTING PEATLAND  
WATER-TABLE DEPTH FROM TESTATE AMOEBAE

Manuscript published in *The Holocene*: Nolan, C., Tipton, J., Booth, R. K., Hooten, M. B., & Jackson, S. T. (2019), Comparing and improving methods for reconstructing peatland water-table depth from testate amoebae, *The Holocene*. DOI: 10.1177/0959683619846969.

Re-use by the Author in a dissertation is permitted via the SAGE Green Open Access Policy (<https://us.sagepub.com/en-us/nam/journal-author-archiving-policies-and-re-use>).

## **Comparing and improving methods for reconstructing peatland water-table depth from testate amoebae**

Connor Nolan<sup>1</sup>, John Tipton<sup>2</sup>, Robert K. Booth<sup>3</sup>, Mevin B. Hooten<sup>4,5,6</sup>, and Stephen T. Jackson<sup>7,1</sup>

<sup>1</sup> Department of Geosciences, University of Arizona, Tucson, AZ 85721, USA

<sup>2</sup> Department of Mathematical Sciences, University of Arkansas, Fayetteville, AR 72701, USA

<sup>3</sup> Department of Earth and Environmental Sciences, Lehigh University, Bethlehem, PA 18015, USA

<sup>4</sup> U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit, Fort Collins, CO 80523, USA

<sup>5</sup> Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO 80523, USA

<sup>6</sup> Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA

<sup>7</sup> U.S. Geological Survey, Southwest Climate Adaptation Science Center, Tucson, AZ 85721, USA

### A.1. Abstract

Proxies that use changes in the composition of ecological communities to reconstruct temporal changes in an environmental covariate are commonly used in paleoclimatology and paleolimnology. Existing methods, such as weighted averaging and modern analog technique, relate compositional data to the covariate in very simple ways, and different methods are seldom compared systematically. We present a new Bayesian model that better represents the underlying data and the complexity in the relationships between species' abundances and a paleoenvironmental covariate. Using testate amoeba-based reconstructions of water-table depth as a test case, we systematically compare new and existing models in a cross-validation experiment on a large training dataset from North America. We then apply the different models to a new 7500-year record of testate amoeba assemblages

from Caribou Bog in Maine and compare the resulting water-table depth reconstructions. We find that Bayesian models represent an improvement over existing methods in three key ways: more complete use of the underlying compositional data, full and meaningful treatment of uncertainty, and clear paths toward methodological improvements. Furthermore, we highlight how developing and systematically comparing methods leads to an improved understanding of the proxy system. This paper focuses on testate amoebae and water-table depth, but the framework and ideas are widely applicable to other proxies based on compositional data.

## A.2. Introduction

For more than a century, paleoecologists, paleolimnologists, and paleoceanographers have tried to infer environmental changes from temporal changes in the composition of ecological communities preserved in sediments (Edwards et al., 2017). Early work focused on qualitative assessments of the environmental affinities of different species or groups of species (von Post, 1918; Hustedt, 1937-39; Iversen, 1944; Phleger, 1953; Nygaard, 1956; Wright Jr. et al., 1963). By the 1970s, qualitative approaches began to give way to quantitative reconstructions. This shift was led by formative work of John Imbrie and colleagues. In a benchmark paper, Imbrie and Kipp (1971) introduced a framework for quantitative reconstruction of past environments from compositional data using transfer functions (Imbrie and Kipp, 1971; Imbrie et al., 1973). Transfer functions aim to formally and quantitatively

relate changes in the composition of an ecological community to changes in a paleoenvironmental covariate.

Imbrie's work centered on the marine realm, using foraminifera to reconstruct past sea surface temperatures. However, similar techniques were simultaneously being developed for paleoclimatological inference from terrestrial data (e.g. pollen) (Webb and Bryson, 1972; Imbrie and Webb, 1981) and were soon applied to other kinds of paleoecological data. Researchers have successfully used many different types of organisms to infer many different target variables. Some examples include: reconstructing temperature and precipitation changes via changes in the surrounding vegetation as recorded by pollen (Webb and Bryson, 1972), inferring bog water-table depth histories using changes in testate amoebae assemblages (Warner and Charman, 1994; Woodland et al., 1998), and generating records of lake pH and salinity based on the types of diatoms living in the lakes and preserved in sediment cores (Battarbee et al., 2002).

In particular, diatom-based pH reconstructions were a key testbed for development of quantitative methodologies. Researchers applied many possible methodologies to the problem of how to best reconstruct the pH of lakes based on diatom assemblages preserved in sediments (Battarbee, 1984; Charles, 1985; Davis et al., 1985; Renberg et al., 1985; Oehlert, 1988; Birks et al., 1990). This work led to application of diatom-based pH reconstructions in Europe and the United States as a key line of evidence in determining the effects of acid precipitation on lake systems. Ultimately, these paleolimnological studies were a major factor in environmental policy changes, including the 1990 amendments to the Clean Air

Act (Smol, 2009).

Applications of transfer functions have continued to the present, including important work to understand potential pitfalls and improve inferences. Examples include understanding the effects of spatial autocorrelation (Telford and Birks, 2009), differential preservation (Mitchell et al., 2008), and considering appropriate ecological models (Juggins, 2013). Furthermore, there has been important work on cross-validation and uncertainty estimation (Telford et al., 2004; Payne et al., 2012; Trachsel and Telford, 2016). Despite these advances and applications, there has been little fundamental rethinking of the underlying methodologies (Birks and Simpson, 2013): simple models like weighted averaging are still widely used in paleoenvironmental reconstructions from paleoecological data.

A logical step towards innovation for proxies based on compositional data is to take advantage of state-of-the-art Bayesian statistical techniques. Bayesian methods have started to be applied to compositional data (Toivonen et al., 2001; Salonen et al., 2012), but they have not seen wide use in the community. Bayesian models allow for a more realistic representation of underlying ecological responses to changes in the environment, inference on the underlying ecological processes, and a more formal and thorough treatment of all sources of uncertainty.

Uncertainties are a key part of any proxy reconstruction because they represent our confidence in the inferences being made. Paleoclimate records based on compositional data are important indicators of terrestrial climate over the Holocene (Marlon et al., 2017; Shuman et al., 2018). Thus, when these proxies are used in climatic synthesis and data

assimilation efforts such as PAGES2k (Hydro2k, 2017), the Last Millennium Reanalysis (Hakim et al., 2016), and the Paleoecological Observatory Network (PaleON) (Marlon et al., 2017), meaningful estimation of uncertainties is necessary to combine and weigh the data coming from the many different sources. It is therefore critical to question and improve inference and uncertainty estimation for proxies based on compositional data.

In this paper, we apply Bayesian statistical tools to the problem of reconstructing an environmental covariate from compositional data. We compare Bayesian methods to a range of existing methods in cross-validation and in reconstruction. The framework we use is general and potentially applicable to any proxy based on compositional data, but we focus on testate amoeba-based reconstructions of water-table depth in ombrotrophic peat bogs.

Testate amoebae are single-celled organisms that live on the surface of bogs and in lakes, wetlands, etc. The community composition of testate amoebae living on the surface of ombrotrophic bogs is sensitive to changes in the surface wetness of the bog. Our operational measure of surface wetness is water-table depth. Water-table depth is measured relative to the peat surface with high values corresponding to deep water-tables (dry surface conditions), low values indicating wetter surface conditions, and negative values indicating standing water (Booth, 2008).

Testate amoebae are a prime target for methodological comparison because we have a large training dataset of surface samples from bogs across North America and a new 7,500-year long testate amoeba record from Caribou Bog in Maine. First, we introduce



those datasets. Then, we review existing methods for developing transfer functions and introduce our new Bayesian model. Next, we apply these models to the training dataset and perform a cross-validation experiment to test their predictive skill. Finally, we use the models to reconstruct water-table depth and compare the resulting reconstructions. We discuss specific implications of our results for testate amoebae-based water-table depth reconstructions and general implications for proxies based on compositional data.

### A.3. Data and methods

#### A.3.1. Modern surface sample training dataset

Paleoenvironmental reconstruction from compositional data requires a modern training dataset, in which both the community assemblage and the environmental covariate are measured for a number of sites distributed in space. The training dataset should sample the full range of covariate values expected in reconstruction. Developing a training dataset entails measuring the depth to the water-table (the environmental covariate) at many sites on many peatlands, collecting surface samples from the locations of the water-table depth measurements, and, back in the lab, processing the surface sample and counting the testate amoeba assemblage under a microscope.

We used a large modern training dataset of 978 samples from 68 sites in North America (Figure A.1). This dataset contains samples from Booth (2002), Booth (2008), Booth and Zygmont (2005), and Markel et al. (2010). For the 378 samples from the Booth (2008)

dataset the water-table depths were measured using the method of PVC-tape discoloration (Booth et al., 2005). This technique produces a measurement of average water-table depth over a season, which can lead to a water-table depth estimate more directly analogous to the temporal integration of the testate amoeba surface samples, which are collected instantaneously but represent time-averaged death assemblages (Booth, 2008). The water-table depths for samples from datasets other than Booth (2008) were measured instantaneously at the time of collection. For 27 samples the true water-table depth was not measured because it was greater than 50cm below the surface of the bog. These samples were not used because the transfer function methods are not statistically robust to this type of censoring, and in particular, the censored observations affect the convergence of the Bayesian statistical models as formulated in this paper. We note that the elimination of these dry samples may have resulted in a loss of ecological information for some taxa, like *Hyalosphenia subflava*.

Our training dataset, then, consists of 951 samples. In practice, this looks like a vector,  $Y$ , of 951 measured water-table depths and a matrix,  $X$ , where each of the 951 lines corresponds to the observed testate amoeba assemblage from each site. Each column of the matrix  $X$  corresponds to a different testate amoebae species.

### A.3.2. Site Description and Lab Methods: Reconstruction Dataset

Caribou Bog (44.985 N, 68.814 W; elevation: 39 m) is a raised, ombrotrophic peat bog system located in south-central Maine, USA. The bog system (Caribou Bog, Mud Pond,

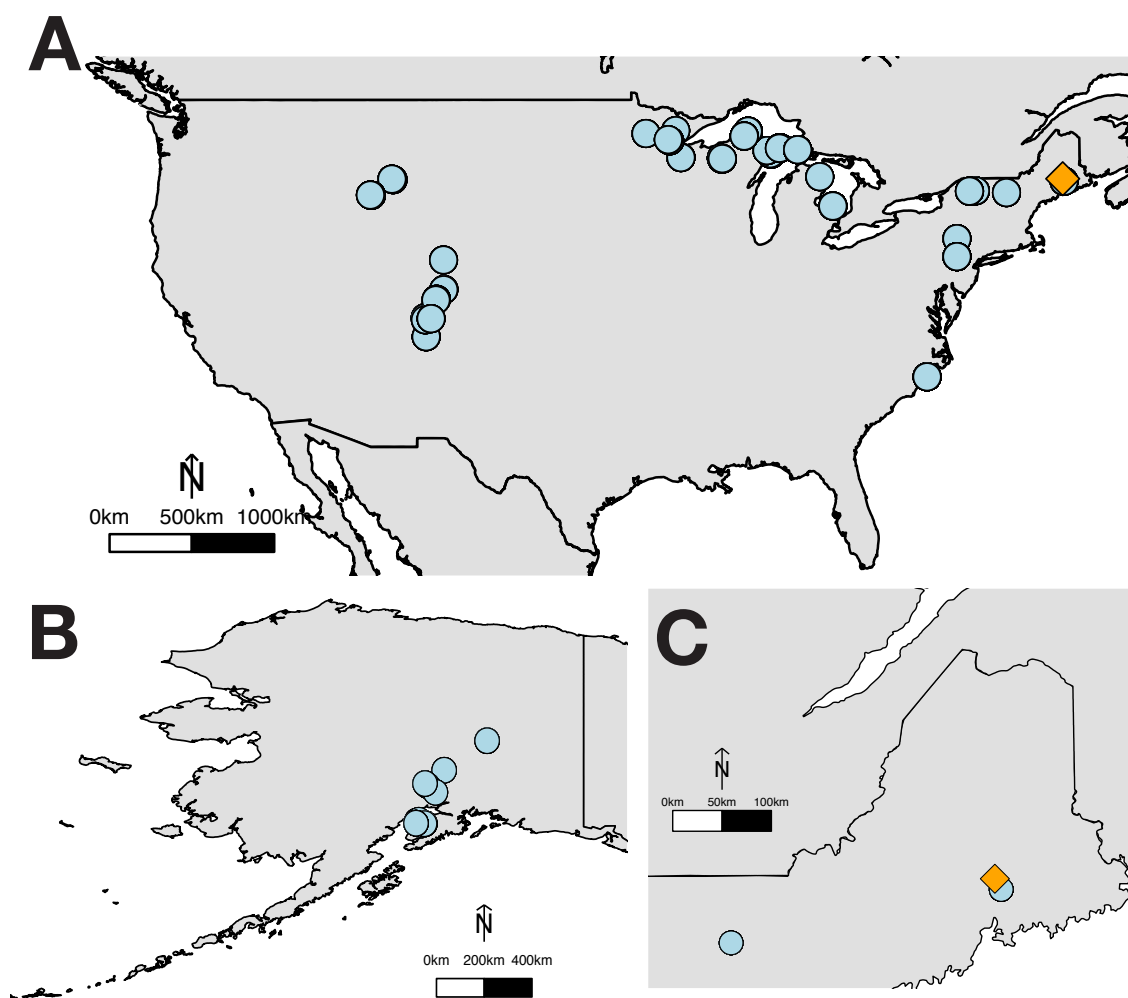


FIGURE A.1. Map showing the sites that make up the modern training dataset sites (blue circles) and the site of the new water-table depth reconstruction, Caribou Bog (orange diamond). Panel A shows the continental USA, Panel B shows the training dataset sites in Alaska, and Panel C zooms in on Maine and the surrounding area.

Pushaw Stream wetlands) lies in a series of north-northwest to south-southeast trending features and covers approximately 2400 hectares.

The post-glacial development of the peatland has been chronicled by Gajewski (1987) and Hu and Davis (1995). Briefly, after the Laurentide ice sheet receded, the isostatically depressed site was flooded by ocean water. Approximately 14,800 years before present, as the land surface emerged from the sea, an unproductive glacial lake formed. Eventually, organic sediment deposition began, with the oldest organic sediments found by Hu and Davis (1995) dating to approximately 12,300 cal yr BP. A peatland began to form via terrestrialization around 11,000 yr BP. This began first as a limnic sedge fen before transitioning to an ombrotrophic wooded fen (8300-9000 cal yr BP), and finally a *Sphagnum*-dominated peatland (~6400 cal yr BP) (Hu and Davis, 1995). The site has remained a *Sphagnum*-dominated ombrotrophic peatland to the present. The dates of the transitions varied in different parts of the bog (Hu and Davis, 1995) and our core suggests that the site was *Sphagnum*-dominated as far back as 7500 cal yr BP, but the above provides a general trajectory.

The modern vegetation at the site is an open *Sphagnum-Ericaceae* peatland with scattered, isolated *Picea*. There is a narrow *Alnus* swamp around the edge of the bog (Gajewski, 1987). The surrounding forests contain secondary *Betula papyrifera*, *Abies*, and *Populus*. Prior to European settlement *Tsuga canadensis* and *Pinus* were common in the area (Gajewski, 1987).

Sediment cores were collected in the summer of 2014. We used a 4-inch diameter

modified piston corer (Wright et al., 1984) to collect the top 251cm in three drives. We then used a Russian (Jowsey) peat corer to collect 96cm further in two drives for a total of 347cm. In the lab, the cores were cut into contiguous 1 cm sections and subsampled for testate amoebae and radiocarbon analyses.

We obtained a 1 cm<sup>3</sup> subsample from each 1 cm section of the core for analysis of testate amoebae and pollen and prepared them using standard techniques based on Booth (2010). Samples were boiled in water, sieved to retain the fraction between 15 and 250 micrometers, dyed with Safranin O, cleaned via two rounds of centrifugation, and stored in glycerol. Testate amoebae were identified and counted at 400x optical magnification to a count of at least 100 individuals per sample.

We obtained 24 radiocarbon dates on *Sphagnum* stems to develop a chronology for the core (Table B.1). Using these dates and the year of collection, we developed an age-depth model with Bchron (Parnell, 2014) (Figure B.1). The basal age is approximately 7500 calibrated years before present. The age model and core lithology show no signs of depositional hiatuses. Sedimentation times range from 5 yr/cm to 30 yr/cm with an median of 18 yr/cm. These are consistent with the regional priors established by Goring et al. (2012).

### A.3.3. Models

*Basic problem* There are a variety of methods to estimate environmental covariates from compositional data. Some of these tools have ties to formal statistics; others have developed

through domain experts' intuition (Birks, 1998; Juggins and Birks, 2012). Regardless of the complexity or connection to formal statistics, the most widely used methods are underlain by similar general assumptions (Imbrie and Webb, 1981; Juggins, 2013).

The basic idea, following Imbrie and Webb (1981), is that the ecological community assemblage,  $Y$ , is related to the value of some environmental covariate  $X$  via an ecological response function  $R_e$ , such that

$$Y = R_e(X).$$

In general, we are interested in the inverse prediction problem where we know the ecological assemblage and want to predict the covariate,

$$X = \phi(Y),$$

where  $\phi(Y)$  is a so-called transfer function. Having set up this basic problem, we introduce the models that have been used to predict  $Y$  based on  $X$ . First we discuss weighted averaging (WA), weighted averaging partial least squares (WAPLS), and maximum likelihood response curves (MLRC). These models treat the response of each species to the covariate independently. Then, we introduce two methods that approach the problem differently: modern analog technique (MAT) and random forest (RF). These models use the full assemblage simultaneously to estimate the covariate. Next, we discuss uncertainty estimation for WA, WAPLS, MLRC, MAT, and RF. Finally, we introduce the Bayesian models (BUMMER and MVGP) which are logical extensions and formalizations of the first set of

models. The Bayesian approaches model the ecological response of individual species, but in the Bayesian models the responses are modeled jointly to respect the sum-to-N nature of the data.

*Weighted Averaging, Weighted Averaging of Partial Least Squares, and Maximum Likelihood Response Curves* These three transfer function methods are based on the idea that the ecological response of each species to a covariate can be summarized by an optimum value at which the species is most abundant (Gauch and Whittaker, 1972). Weighted averaging (WA) is a very simple procedure that can be written in two equations. First, from the modern training dataset, the estimated optimum,  $\hat{\mu}_k$  for each of the  $k$  observed species is calculated by:

$$\hat{\mu}_k = \frac{\sum_i y_{ik} x_i}{\sum_i y_{ik}}$$

Then, these estimated optima,  $\hat{\mu}_k$  are then used to estimate the unobserved covariate,  $x$  based on the observed assemblage  $y$  as follows:

$$\hat{x} = \frac{\sum_k y_{ik} \hat{\mu}_k}{\sum_k y_{ik}}$$

The range of WA-inferred covariate at this point is greatly reduced ("shrunk") compared to the observed range in the training dataset. To correct for this, a linear regression is used to "deshrink" the predictions (ter Braak and van Dame, 1989; Birks et al., 1990). This can be done in two ways: so-called "classical" deshrinking regresses the inferred covariate values onto the observed covariate values from the training set, whereas "inverse" deshrinking

regresses the observed values onto the inferred values. Classical deshrinking produces a larger range of "deshrunk" values. Recent work has suggested using a smooth, monotonic spline regression instead of the standard linear regression. (Birks and Simpson, 2013). We suggest that deshrinking to improve the raw WA predictions using a simple regression can be thought of as "boosting" in the language of machine learning. Boosting is defined as using a set of "weak learners" to produce a "strong learner" (Schapire, 1990). The raw WA predictions and the linear regression on their own are weak models; combined together they produce a stronger model.

Despite the simplicity of WA, it has continued to be successful, with wide applications in paleoclimate inference (e.g., Birks and Simpson (2013), Clifford and Booth (2013), and Amesbury et al. (2016)). Some modifications to WA have been proposed including emphasizing species with narrow tolerances by weighting each species by the inverse of their tolerances (ter Braak and van Dame, 1989), but in practice this generally has little effect (Juggins and Birks, 2012). Another recent suggested modification was to perform WA on traits instead of species (van Bellen et al., 2017). The trait-based transfer function performed similarly to standard WA on species.

A modification of WA that can significantly improve predictions in some cases is weighted averaging of partial least squares (WAPLS) (ter Braak and Juggins, 1993). As the name suggests, WAPLS combines the weighted averaging method described above with partial least squares, a dimension-reduction algorithm comparable to principal components analysis. WAPLS with a single PLS component is equivalent to standard WA. Each ad-



ditional WAPLS component can be thought of as performing weighted averaging on the residuals of WA to improve the estimation of the optima. For details of the algorithm the reader is referred to ter Braak and Juggins (1993) and Juggins and Birks (2012). We suggest that, similar to the deshrinking, WAPLS is another example of boosting (Schapire, 1990). In this case, the weak-learner WA is being applied multiple times to produce a stronger learner.

Weighted averaging can be thought of as a simplification of maximum likelihood response curves (MLRC) (ter Braak and van Dame, 1989; Oksanen et al., 1990; Birks et al., 1990). Instead of summarizing a species' response to the covariate with a simple optima (as in WA), MLRC fits a Gaussian response curve for each species, in turn. Then, in prediction, the curves are used in a maximum likelihood framework to estimate the environmental covariate for an observed assemblage. MLRC is not widely used because the curves are difficult to fit within the Maximum Likelihood framework and the model often does not perform as well as WA or WAPLS (Juggins and Birks, 2012).

*Modern Analog Technique and Random Forest* Whereas WA, WAPLS, and MLRC use the response of each individual species to predict the covariate, a different approach, Modern Analog Technique (MAT), uses the full assemblage at once to estimate the covariate (Jackson and Williams, 2004). MAT uses the same training data set, but instead of estimating a species by species response to the covariate, MAT assumes that similar assemblages should be associated with similar covariate values. To estimate the covariate for a given

assemblage, MAT finds the  $k$ -closest analog assemblages in the training dataset based on a multidimensional distance metric. There are many possible choices for distance metrics, each with unique strengths and weaknesses (Prentice, 1980), but chord distance has become widely used for compositional data (Overpeck et al., 1985). The analog assemblages identified in the training set are each associated with a value of the covariate. For the assemblage of interest, the covariate is estimated by taking the mean (could use median or some weighted mean as well) of the covariates of the analog assemblages. In statistics, MAT is known as  $k$ -nearest neighbors. MAT has not been widely applied to testate amoebae, but it is commonly used in reconstruction of paleoclimate from pollen and other compositional data.

Another analytical approach we apply to the data is Random Forest (RF) (Breiman, 2001). RF is a common machine learning technique. Like MAT, RF generates predictions by identifying analogs, but RF uses a very different algorithm. Instead of using a simple multivariate distance metric, RF creates a large number of dichotomous decision trees to identify analogs. RF begins by creating bootstrapped versions of the modern training dataset: the modern training dataset is randomly sampled with replacement to create a bootstrapped training set of the same size as the real training dataset. Then for each bootstrapped training dataset, RF begins a dichotomous decision tree by choosing a random subset of species and identifying which of those species best splits the observed covariate. This process is repeated at each node until a full decision tree is grown. RF stops when each terminal node satisfies a "stop" condition by containing only a small number of observa-

tions or not being able to make a meaningful further split. Many different trees are created to make a "random forest". Then, to make a prediction an assemblage is run through all of the decision trees. The predicted value of the covariate is the average of the covariate value for each of the terminal nodes identified for the sample over all the trees. For details the reader is referred to Breiman (2001) and Liaw and Wiener (2002).

*Uncertainty Estimation* WA, MLRC, MAT, and RF do not produce full predictive distributions, so uncertainty must be estimated. This estimation is done by bootstrap re-sampling of the data (Birks et al., 1990; Liaw and Wiener, 2002; Juggins, 2017). This re-sampling of the data approximates the re-sampling of the parameters that is done in Bayesian statistics to produce a full predictive distribution (Hobbs and Hooten, 2015).

We illustrate the bootstrap process by describing its implementation for WA. The process is similar for the other models. Each bootstrap cycle begins by randomly drawing from the training set (with replacement) to create a new modern training dataset the same size as the full modern training dataset. The samples in the full training data that are not a part of the bootstrap sample are used as an out-of-sample validation test set (approximately one-third of total number of samples is out-of-sample per bootstrap sample). For each bootstrapped training set, the model is fit to the bootstrap sample and predictions of the covariate are generated on the out-of-sample bootstrap validation data. The bootstrapping is run for a large number of iterations, say, 1000, and the out-of sample predicted covariate values are compared to their respective observed values. Using the predictive distribution

and the observed distribution, the root mean-square prediction error is calculated. The first component of the mean square prediction error, called  $v_1$ , is intuitive:  $v_1$  is the variability in the estimation of the taxon optimum values  $\hat{\mu}_k$  across all of the bootstrap samples. This error component varies for each of the reconstruction samples. The  $v_1$  error approaches zero as the size of the calibration dataset increases. The second component,  $v_2$ , is the difference between the measured covariate values and the predicted covariate values when a sample is in the bootstrap test set. This error component arises from variation in taxon abundance for a given covariate value and does not vary between reconstruction samples. Using the language of statistical learning, we can call  $v_1$  the reducible error (the error that decreases as sample size increases) and  $v_2$  irreducible error. Irreducible error consists of variation that cannot be modeled by the method in question except as a random process.

WA, MLRC, MAT, and RF fail to account for the uncertainty in translating multinomial counts to proportions. For example, we are much more certain that an underlying composition is  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  if our count vector is  $(500, 500, 500)$  than if the observed count vector is  $(5, 5, 5)$ . WA, MLRC, MAT, and RF do not consider the underlying counts, instead they begin with observed proportions and thus lose some information about the uncertainty in those proportions. Previous work on the influence of count number on transfer function performance found that count totals from 50-150 performed similarly (Payne and Mitchell, 2009), but this work did not consider the uncertainties in going from counts to proportions.

*Bayesian approaches* Bayesian statistics allows us to efficiently fit more complex and realistic mechanistic models. Bayesian approaches begin by mathematically specifying all parts of the process and their interrelations. These processes generally form a natural hierarchy that in Bayesian models manifests as a series of conditional probabilities. This is called a Bayesian hierarchical model (Hobbs and Hooten, 2015). Bayesian hierarchical models can flexibly account for all sources of uncertainty within a model (or even between models if Bayesian Model Averaging is used (Hoeting et al., 1999)). Bayesian hierarchical models can capture more complexity in ecological responses to an environmental covariate, jointly estimate taxon responses to the environmental covariate, and model the data in the form they were generated. Bayesian models used in this paper are also generative models (Gelman et al., 2017), meaning they can be used to simulate data. Simulation studies can be useful to better understand the data and system (Tipton et al., 2019). Here we consider two Bayesian models: BUMMER (Toivonen et al., 2001) and a new model we call multivariate Gaussian process (MVGP; (Tipton et al., 2019)).

BUMMER (Toivonen et al., 2001) is a Bayesian unimodal response model that can be thought of as a Bayesian version of MLRC in that it models the response of each taxon to the changes in the covariate as a Gaussian kernel. BUMMER assumes the functional relationship between taxon abundance and environmental covariate can be reasonably approximated using a bell-curve shape. BUMMER has been applied to paleoenvironmental reconstructions based on chironomids (Toivonen et al., 2001) and pollen (Salonen et al., 2012) with some success, but has not been used on testate amoebae previously.

Multivariate Gaussian process (MVGP) is a Bayesian hierarchical model that relaxes the assumption of a unimodal response. MVGP requires only a smooth response and uses a flexible B-spline or Gaussian process (here we use the B-spline version) to fit each taxon's response to the covariate. This allows for more complex and realistic functional responses of species' abundance to the covariate (Tipton et al., 2019). Other choices of functional response, not explored in this paper, include low-rank correlated Gaussian processes, penalized B-splines, and shape-restricted splines constrained to be unimodal but not necessarily bell-shaped.

The process for fitting and predicting with these Bayesian models is similar for both MVGP and BUMMER. As with all of the other methods, these models begin with the modern calibration dataset, and like WA, WAPLS, and MLRC these Bayesian approaches seek to model the relationship between the abundance of each taxon and the covariate (Figure A.2). These Bayesian models fit the relationship jointly meaning that if one species has a high abundance at a particular value of the covariate at least one or more other species must be lower. The Bayesian models start with the counts collected by the analyst then the model accounts for large variation in composition at a given covariate value by using a Dirichlet-Multinomial distribution. Both BUMMER and MVGP model the latent functional response between abundance and the covariate. Both Bayesian models are fit to the training data using Markov Chain Monte Carlo to generate a posterior distribution. The posterior estimates of parameters are then used to generate posterior predictions of the covariate given an assemblage either from the test data (in cross-validation) or from sub-fossil

samples downcore (in reconstruction). The posterior predictions are generated by inverting the functional relationship between abundance and **the environmental covariate**, producing a proper probability distribution for each covariate in the dataset.

MVGP and BUMMER have a few notable features. First, instead of converting counts to proportions, both models begin with compositional count data. It is better to model the data in the manner the data were collected than to transform the data to fit a specific model because data transformations done before the modeling can introduce uncertainty that can be difficult to account for. For example, we know that the compositional count data have variability that is unaccounted for in the traditional models. By modeling the compositional counts with a Dirichlet-Multinomial likelihood, the Bayesian hierarchical models are better able to account for the overdispersion observed in the data while automatically satisfying the sum-to-one constraint. MVGP jointly fits a functional relationship between abundance and the covariate using B-splines. The B-spline functional relationship is capable of capturing potentially asymmetric and/or multi-modal responses in a species' relative abundance along an ecological gradient (Hefley et al., 2017). These potentially more complex functional responses to the covariate are not accommodated by models based on unimodal distributions or a simple mean (i.e. BUMMER, WA, MLRC). There are many potential avenues for modifying these Bayesian models to improve the predictions. Some ideas include accounting for interactions among species, adding spatially/temporally correlated random effects to model overdispersion in the data, and clustering species with similar responses to the environmental covariate to reduce the number of parameters that need to be estimated.

#### A.3.4. Cross-validation and scoring rules

We use cross-validation to evaluate the performance of all of these models introduced above. Cross-validation consists of five steps: i) setting aside a subset of the data (the test set), ii) fitting a model using the remaining data (the training set), iii) generating predictions about the test set, iv) evaluating the predictions about the test set, and v) repeating the process with different test and training sets. In statistics and machine learning, cross-validation performance of a model gives a good indication of the performance of a predictive model on novel data (Hooten and Hobbs, 2015). If a model's predictions are accurate and precise under cross-validation then it is generally reasonable to assume that the quality of the predictions will generalize to new data. This generalizability is a critical property in paleoclimate reconstruction because there are often no direct measurements of the covariate of interest for the reconstruction samples. Recent work on transfer function validation has highlighted the usefulness of independent test sets, such as modern samples not included in the training set Payne et al. (2012); Van Bellen et al. (2014); Tsyganov et al. (2016) or instrumental data (Swindles et al., 2015). In the case where these are unavailable, cross-validation remains the state-of-the-art and is commonly used to evaluate competing models in statistics and machine learning. For this study, no independent test set or instrumental data are available, so we perform a 10-fold cross-validation and use the following scoring rules to evaluate the strengths and weaknesses of the predictions coming from the seven models under comparison



Evaluation of model performance using cross-validation requires a choice of scoring rule. Commonly used scoring rules include mean square prediction error (MSPE) and mean absolute error (MAE). MSPE measures the model's skill in predicting the mean while MAE measures skill in predicting the median. Smaller MSPE and MAE indicate better model performance.

A desirable property of a scoring rule is propriety. A scoring rule is proper if the scoring rule chooses the best predictive model on average. In practice, proper scoring rules have a similar interpretation to many other statistical quantities: for any given dataset a proper scoring rule might not pick the best model due to sampling variability, but a proper scoring rule will pick the best predictive model over repeated re-sampling. MSPE and MAE are generally not strictly proper scoring rules because if there are two models with the same predictive mean but different predictive standard deviations, the models will have the same MSPE but the model with 95% predictive coverage intervals closest to a 95% frequency is the better predictive model. Therefore, MSPE is unable to determine the better predictive model and is not strictly proper. The Continuous Ranked Probability Score (CRPS) was developed to be a proper scoring rule that is both accurate (predictive mean is centered on the latent quantity) and precise (the predictive distribution is calibrated so the the  $\alpha\%$  predictive distribution has empirical coverage near  $\alpha\%$  (Gneiting, 2011). For the non-Bayesian methods the CRPS score is equivalent to MAE because these methods do not produce full predictive distributions.

### A.3.5. Computing methods

Analyses were done in R (R Core Team, 2017). WA, MLRC, and MAT were fit with the rioja package (Juggins, 2017). RF was fit using the package randomForest (Liaw and Wiener, 2002). BUMMER was fit using Stan (Carpenter et al., 2016). MVGP was fit using the BayesComposition package, available on GitHub at [github.com/jtipton25/BayesComposition](https://github.com/jtipton25/BayesComposition). The BayesComposition repository also includes code and data to replicate the figures and analyses in this paper.

## A.4. Results and discussion

Because WA and WAPLS are the most commonly used transfer function methods in the analysis of testate amoebae, we focus our analysis on the performance of WA and WAPLS compared to the Bayesian models with some brief discussion of the broader suite of models.

### A.4.1. Fits to training set

Figure A.2 illustrates the model fits to the training dataset by comparing MVGP, WA, and WAPLS. Each taxon's distribution of abundances is plotted as a function of water-table depth (WTD). The response curves generated by MVGP and the raw WA optima, deshrunk WA optima, and WAPLS optima are plotted for each taxon. The plots of abundance versus water-table depth illustrate the strongly over-dispersed nature of the data and the difficulty of summarizing these distributions with a simple mean. The distributions of some taxa may

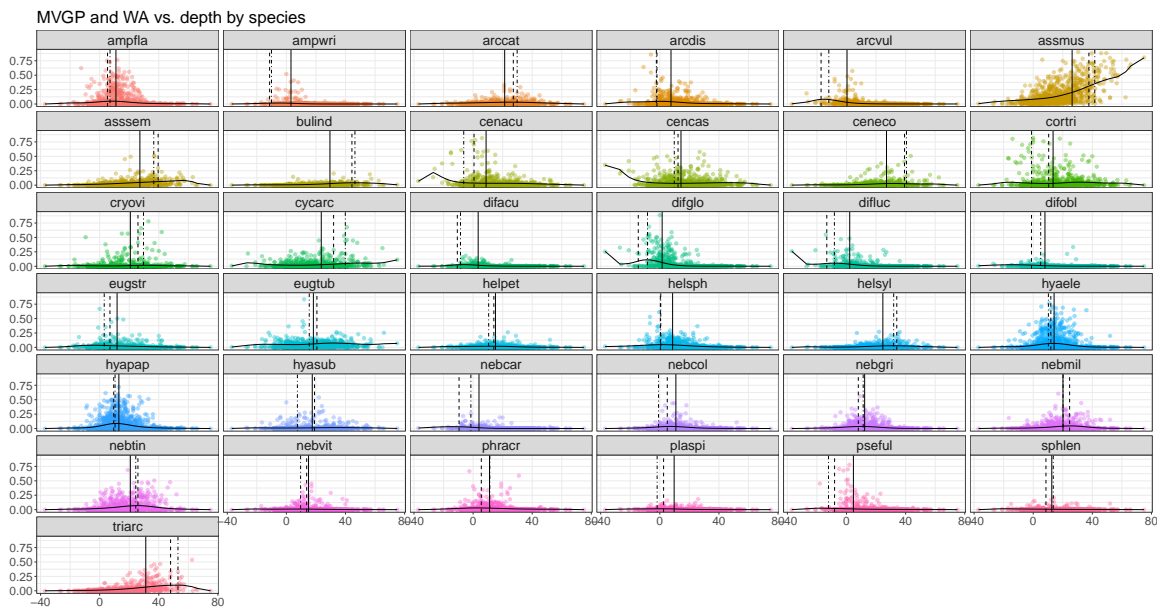


FIGURE A.2. Species abundance versus water-table depth (in cm) from the samples in the training dataset (points). The smooth curves are from Multivariate Gaussian Process for each species. The solid vertical lines are the raw optima from weighted averaging (WA). The dashed vertical lines are the deshrunk WA optima. The dot-dash vertical lines are the coefficients from weighted averaging of partial least squares after five partial least squares components.

be reasonably summarized by an optimal water-table depth (e.g., *Amphitrema wrightianum* and *Pseudodiffugia fulva*-type). The distributions of other species, however, are clearly not well-represented by the WA optima (e.g., *Assulina muscorum* and *Trigonopyxis arcuata*). For many species, the deshrunk WA optima and the WAPLS-derived coefficient are comparable, but there are many notable differences. For some species (e.g., *Hyalosphenia subflava*, *Nebela collaris-bohemica*-type, and others) the WAPLS-derived coefficients are significantly different from what a visual inspection of the distribution might suggest. The species coefficients derived from WAPLS with multiple components (here we are using two PLS components) become no longer interpretable as optima of the ecological distributions. Instead, they are regression coefficients meant to give a better overall prediction of the covariate (Juggins and Birks, 2012).

#### A.4.2. Cross-validation experiment

We performed a 10-fold cross-validation experiment. Each of the 951 samples in the training set were in the test set for exactly one of the cross-validation folds. Figure A.3 shows the observed versus predicted water-table depths with a one-to-one line and a linear fit. We also calculated four scoring metrics to numerically evaluate the cross-validation results (Table A.1).

The results of the cross-validation show a clear bias-variance trade-off in both the plots of observed vs. fitted WTD (Figure A.3) and in the scores (Table A.1). WA, WAPLS, MAT, and RF all show significant bias in the plots of observed vs. predicted WTD as evidenced

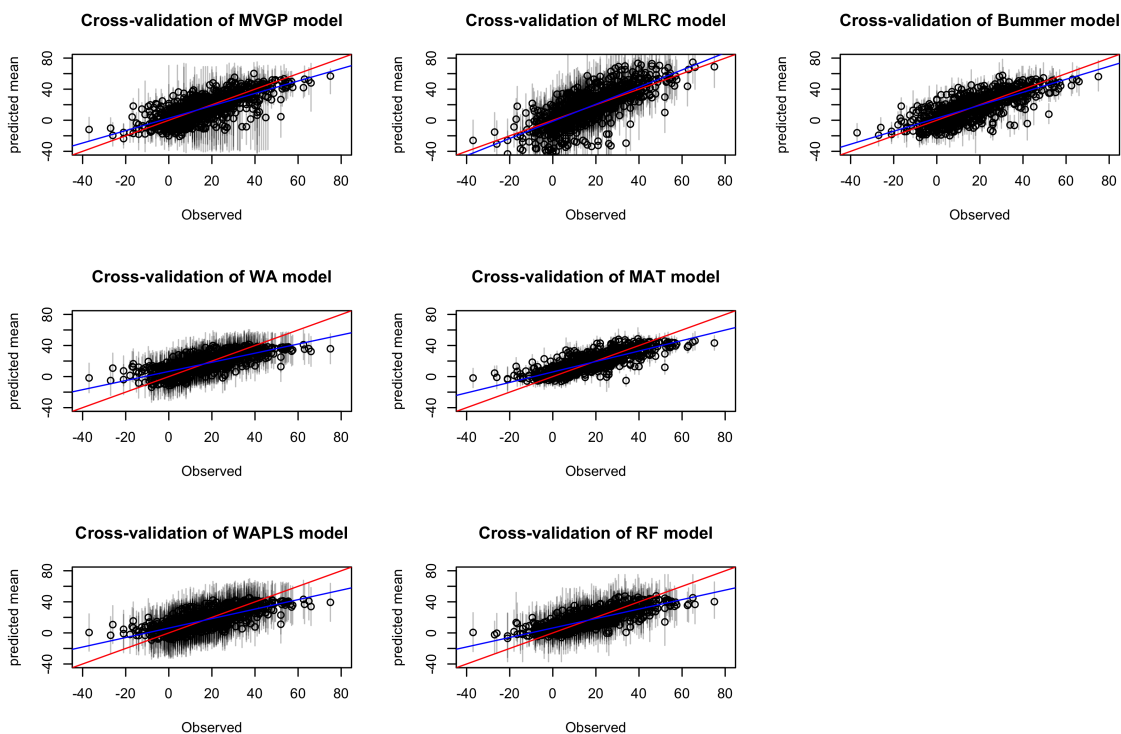


FIGURE A.3. Observed versus predicted water-table depth (cm) for each model from the 10-fold cross-validation experiment on the training dataset of 951 surface testate amoebae assemblages. Red lines in each panel indicate the one-to-one line. Blue lines in each panel are a linear fit of the observed vs. predicted water-table depths. Error bars denote the 95% credible interval

by an offset of the linear fit (blue line) from the one-to-one line (red line). This induced bias, a shrinkage to the mean, allows for generally smaller standard deviations. In contrast, MVGP, Bummer, and MLRC have smaller bias but larger variance.

This bias-variance tradeoff is also apparent in the cross-validation scores (Table A.1). MAT, RF, WA, and WAPLS have the smallest MSPE and MAE (metrics of prediction accuracy for the mean) compared to MLRC, which has the largest MSPE and MAE, and MVGP and BUMMER which are intermediate.

TABLE A.1. Scoring results from 10-fold cross-validation experiment performed on the modern training dataset. Lower scores for MSPE, MAE, and CRPS indicate better performance. For coverage, better performance is indicated by values closer to the nominal rate (in this case, 95%).

	<b>MVGP</b>	<b>BUMMER</b>	<b>WA</b>	<b>WAPLS</b>	<b>MLRC</b>	<b>MAT</b>	<b>RF</b>
MSPE	121.05	111.65	89.99	86.38	193.14	71.99	74.30
MAE	8.32	8.08	7.20	7.02	10.04	6.24	6.22
95% CI coverage	75%	75%	95%	98%	94%	73%	95%
CRPS	6.27	6.06	7.20	7.02	10.04	6.24	6.22

WA, WAPLS, MLRC, and RF all have 95% CI coverage close to the nominal rate. The Bayesian methods (MVGP and BUMMER) had the lower coverages, approximately 75%, for their 95% credible intervals. This is likely due to over-dispersion in the count data beyond what can be accounted for by the Dirichlet-Multinomial likelihood.

For CRPS, the integrated mean and variance score, BUMMER performs best, followed by MVGP, RF, and MAT, then WA and WAPLS. MLRC performs worst.

#### A.4.3. Will skill in cross-validation generalize to novel data?

The goals of model evaluation under cross-validation are to estimate how well the predictive model can generate predictions on novel data. In our experiment, we consider a variety of methods that have relative predictive strengths and weaknesses. WA, WAPLS, MAT, and RF are generally good predictive methods, but are not based on a statistical likelihood. Thus, their predictive ability does not come with any of the statistical guarantees like the "law of large numbers."

Furthermore because MAT and RF rely on finding analogs to generate predictions, their

predictive skill can only generalize to new data when there are reasonable analogs for the new data in the training set. In cross-validation this is unlikely to be important, but non-analog issues can become a source of bias when generating predictions down-core. MAT is not widely used for reconstructions with testate amoeba datasets due to these difficulties of finding good analogues downcore (Charman, et al. 2007), but because MAT is still widely used with other sources of compositional data (Simpson, et al. 2005; Marsicek, et al. 2018), we have retained it here for completeness. There are many possible sources of non-analog assemblages, but one potential source for testate amoebae-based reconstructions is differential preservation of certain types of tests (Mitchell et al., 2008). Thus, it is important to check if there are good analogs for the down-core samples in the training dataset. To do this we calculated the squared chord distance of the accepted analogs for each sample in the training set in cross-validation and for each sample in the Caribou Bog core. We found that the Caribou Bog samples had systematically worse analogs than the training dataset (Figure A.4A). We then looked at these analogs in the time domain and found there is no clear temporal trend in accepted analog distance (Figure A.4B), but there does appear to be some serial autocorrelation with some periods having better analogs and some having worse analogs in the training dataset.

This analysis provides evidence that the strong predictive skill of MAT, RF, WA, and WAPLS in cross-validation may not generalize to predictive skill in reconstruction, a feature that is common in Quaternary paleoclimate and paleoecology (Jackson, 2012). On the other hand, the likelihood-based approaches like MLRC, MVGP, and BUMMER, are ca-

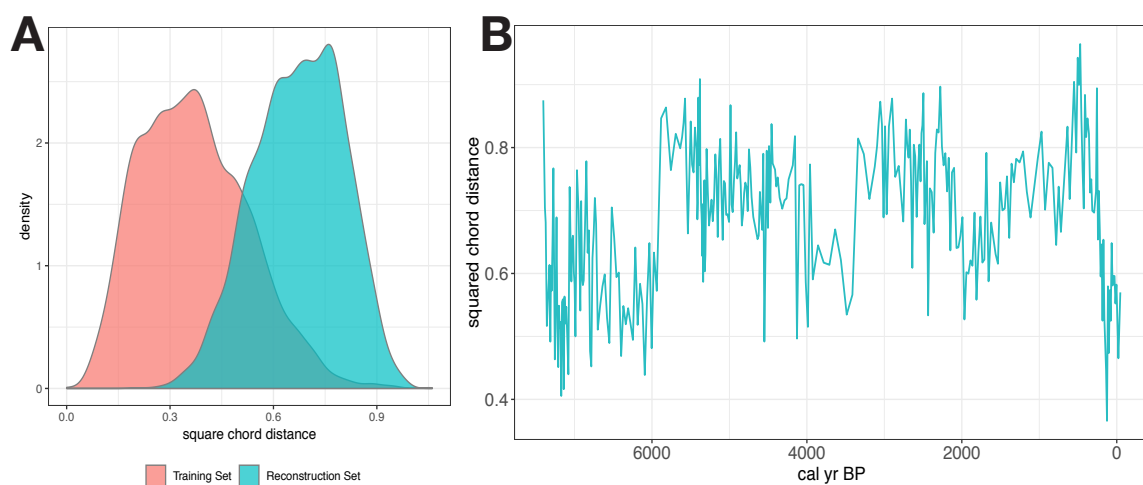


FIGURE A.4. A. Squared chord distances of the accepted analogs from Modern Analog Technique (MAT) for the training dataset (red) and reconstruction dataset from Caribou Bog (blue). Smaller squared chord distance corresponds to more similar assemblages. The training dataset contains many good analogs for the assemblages within the training dataset (distances for the red density curve are small) whereas the distances of the analog assemblages for the reconstruction dataset are larger and thus not as similar. B. Time series of the median squared chord distance from the accepted analog assemblages in the training dataset to each assemblage in the Caribou Bog reconstruction dataset. This panel shows temporal evolution of the analogs summarized by the blue density curve in panel A.



pable of generalizing to novel data where there are not appropriate analogs and generating predictions with uncertainties that account for the how well the covariate value is known for a given assemblage. From this perspective, likelihood-based methods are a relatively conservative approach to prediction because they are potentially less prone to issues of analog similarity and predictive bias (Tipton et al., 2019).

#### A.4.4. Reconstruction

We now move from cross-validation, where we have both the testate amoeba assemblage and the measured water-table depth (WTD), to reconstruction. In reconstruction, we only have a testate amoeba assemblage and we wish to estimate a water-table depth. The 7500 years of testate amoeba assemblages from Caribou Bog are shown in Figure B.2. We then apply the seven transfer function models to infer a WTD history from Caribou Bog. The resulting reconstructions are shown in Figure A.5 and the z-scores of the reconstructions are shown in Figure B.3. The methods break out into three subsets. The first subset is formed by the two Bayesian models – MVGP and BUMMER; the resulting reconstructions look similar to each other in terms of both means and uncertainties. The main difference between MVGP and Bummer is the magnitude of some of the smaller deviations. MVGP tends to have a somewhat damped response compared to BUMMER. This damped response could be due to the more flexible ecological responses allowed by MVGP. The uncertainties, as denoted by the 50% and 95% credible intervals are generally smaller and more variable from sample to sample compared to the other methods. It makes

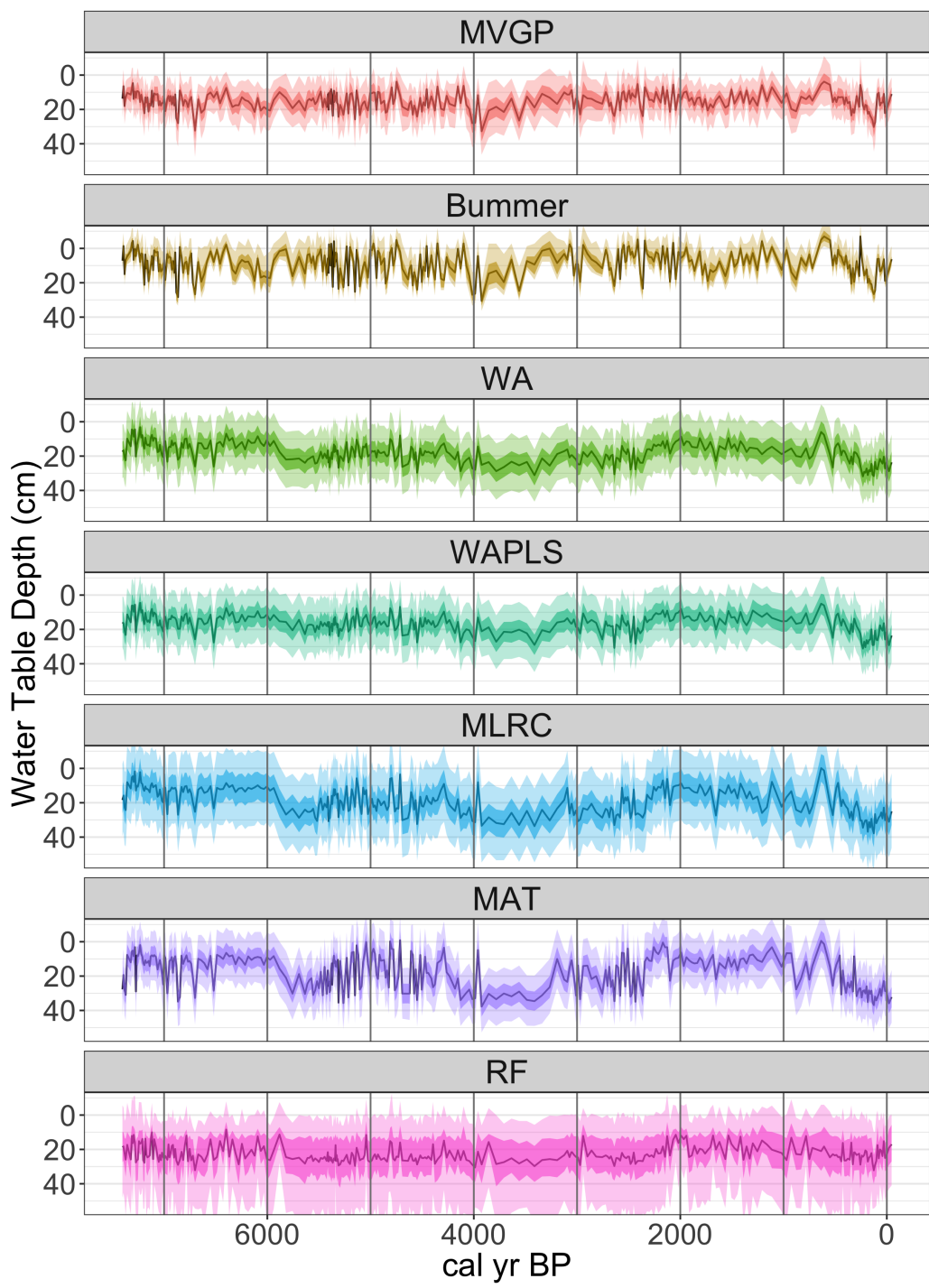


FIGURE A.5. Reconstructions of water-table depth based on testate amoebae assemblages from Caribou Bog. For each panel, the underlying testate amoebae assemblage data is identical the only thing that changes is the transfer function model. The dark shading denotes 50% credible interval, the lighter shading denotes 95% credible interval.

intuitive sense that some predictions should be more certain than others for many potential reasons, including, for example, the particular set of species present and their distributions, the similarity to samples in the training dataset, or the number of individuals counted. With their uncertainties that vary from sample to sample, Bayesian models capture these features of the data.

The next subset is WA, WAPLS, MLRC, and MAT. The reconstructions for Caribou Bog that result from these WA, MLRC and MAT all look similar to each other, again with some variability in the magnitude of the changes in WTD. The reconstructed pattern from WAPLS appears to be intermediate between the Bayesian models and WA. For this set of four models, the uncertainties from sample to sample are nearly constant. These unchanging uncertainties are due to the large  $v_2$  irreducible error component. The  $v_2$  component is responsible for 98% of the root mean squared error of prediction in this dataset. This large irreducible error component limits the usefulness of the uncertainties estimated from these methods.

Lastly, RF produces predictions with a mean near the mean of the training dataset, little variability around that mean, and wide uncertainties. This is likely due to the lack of good analogs in the training set for the reconstruction samples as discussed above. The samples with smaller uncertainties for RF (e.g., around 2000 and 6000 yr BP) are associated with samples that have better analogs in the training dataset (Figure A.4B).

#### A.4.5. Inference on past hydrology

To a first order, when the different methods are applied to the same subfossil dataset (Figure S2), the resulting estimates of past hydrological variability are similar (Figure A.5, Figure S3). This is reassuring because the underlying data are identical between the different reconstructions, but there are some junctures at which one might draw different inferences depending on the method used to reconstruct water-table depth. For example, one would infer a significant dry event at ca. 4000 yr BP using MVGP and BUMMER. But using WA, MAT, and other methods, this event is not nearly as pronounced. This divergence is related to differing model responses to an increase in *Assulina muscorum*. *A. muscorum* occurs across the moisture gradient, but obtains significantly higher abundance in dry conditions (Figure A.2). This affinity for dry conditions leads the Bayesian models to infer drier conditions in response to a small increase in *A. muscorum*. In contrast, WA does not respond strongly because the percentage increase of *A. muscorum* is relatively small and therefore its influence on the WA reconstruction is small. This event illustrates how MVGP and BUMMER can draw out potentially important single-species anomalies that may be missed by other models. This sensitivity leads to a more complete use of the full assemblage, but it comes with a caveat: we need to carefully consider whether species, like *A. muscorum*, that can leverage a WTD reconstruction are faithful indicators of the hydrologic status of the peatland. These are testable ecological hypotheses and avenues for future improvement of mechanistic models (e.g., accommodation of more complex relationships

between the abundance of certain species and the covariate).

In contrast to the event at 4000 yr BP, at ca. 6000 yr BP, WA, MAT, and MLRC show a large drop in WTD, but MVGP and BUMMER do not. The large change in WTD predicted by WA and others is driven by a change in dominance from *Archerella flavum* to *Hyalosphe-  
nia subflava*. These species occur in high percentages (~40%) and thus the exchange of the two drives large changes in WTD for WA. However, the change in the dominant species does not drive a large change in the MVGP and BUMMER reconstructions because the difference between the affinities of the two species, when considering their full distribution in this calibration dataset, is relatively small (Figure A.2). Numerous modern and paleoecological studies suggest that the shift from *A. flavum* to *H. subflava* is a meaningful indicator of a change from wet to dry conditions (e.g. Charman et al. (1999); Booth and Jackson (2003); Swindles et al. (2009); Talbot et al. (2010); Sullivan and Booth (2011); Clifford and Booth (2013)), suggesting that the Bayesian methods likely underestimate the covariate response to this change in the testate amoebae assemblage. Incorporating this expert knowledge is possible in the Bayesian hierarchical framework, and future model development should seek to understand and refine the response of these models to changes in dominant species. These two events represent examples of differential sensitivity of different types of transfer function methods. These kinds of differences in reconstructions based on choice of transfer function model are a source of uncertainty that deserves increased attention in paleoenvironmental reconstructions from proxies based on compositional data. Understanding where and why different models diverge can lead to improved inference on

both modern ecology and past hydrology.

#### A.5. Conclusions and Future Directions: Or, Why Bayesian? Why Bother?

We have presented a detailed analysis of methods for reconstructing water-table depth from testate amoebae community composition. We reviewed existing methods and introduced a new Bayesian method and a Bayesian method that had not yet been applied to testate amoebae data. We tested the models in a cross-validation framework and then applied them to reconstructions.

We found that the existing methods compare well in cross-validation with more complex Bayesian models, but for the simpler existing methods we found that their predictive skill in cross-validation may not transfer to skill in reconstruction (Tipton et al., 2019). When we applied the models in reconstruction, we found generally similar inferences of past hydrology with some key differences. We investigated the sources of some of the discrepancies between the reconstructions from different models. Ultimately, when comparing competing models at a single site without an independent target for validation it is unclear whether one reconstruction is correct and another is wrong. By presenting a systematic comparison of results from different models on the same datasets we illustrate the importance of model choice as a source of uncertainty in records based on compositional data (Jackson, 2012; Gelman and Loken, 2013).

Bayesian models are worth developing and considering for testate amoebae and other

paleoenvironmental proxies based on compositional data. Bayesian models like MVGP and BUMMER are mechanistic models driven primarily by assumptions about the unobserved ecological processes. Developing these types of models makes our assumptions clear and written in the formal language of statistics and can yield improved mechanistic understanding of the system being studied. Mechanistic models underpinned by formal statistics like these are worth the extra effort involved because they allow for a more complete use of the compositional data we spend so much time collecting and result in more robust paleoenvironmental inferences. Including meaningfully estimated uncertainties makes these inferences even more useful. Bayesian models formally account for uncertainty throughout the modeling process (Hobbs and Hooten, 2015). This more complete treatment of uncertainty quantifies how well we really know the past values of the environmental covariate of interest. In secondary analyses and syntheses, uncertainties are necessary to weigh and combine information from many different types of proxies (Hakim et al., 2016; Shuman et al., 2018). When Bayesian models are used on a network of sites, it is possible to add an explicit spatial component to borrow strength across sites and gain new inference on the covariate over space and time with complete uncertainties. With traditional methods, there is a less clear path to achieve a similar borrowing of space and time while properly propagating uncertainty.

We encourage users of all proxies based on compositional data to carefully examine their data and methods in a framework like the one presented here. Testing and comparing new and existing models in a systematic framework leads to a deeper understanding of the

strengths and weaknesses of available tools. Regardless of whether the process results in dramatic improvements of inferences, the process itself is clarifying and important: questioning and testing models and assumptions that underly our important scientific inferences is critical for minimizing "ignorance creep" (Jackson, 2012).

#### A.6. Acknowledgements

This work is part of the Paleoecological Observatory Network (PaleON) Project ([paleonproject.org](http://paleonproject.org)). It was funded by the National Science Foundation Macrosystems Biology program under grant nos. DEB-1241851 and DEB-1241856. We thank Melissa Berke for field assistance; Tom Webb for historical perspective on the development of transfer functions; and Don Charles, Dan Charman, and an anonymous reviewer for useful comments that improved the manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.



## REFERENCES

- Amesbury MJ, Swindles GT, Bobrov A, Charman DJ, Holden J, Lamentowicz M, Mallon G, Mazei Y, Mitchell EA, Payne RJ et al. (2016) Development of a new pan-European testate amoeba transfer function for reconstructing peatland palaeohydrology. *Quaternary Science Reviews* 152: 132–151.
- Battarbee RW (1984) Diatom analysis and the acidification of lakes. *Phil. Trans. R. Soc. Lond. B* 305(1124): 451–477.
- Battarbee RW, Jones VJ, Flower RJ, Cameron NG, Bennion H, Carvalho L and Juggins S (2002) Diatoms. In: *Tracking environmental change using lake sediments*. Springer, pp. 155–202.
- Birks H (1998) Numerical tools in palaeolimnology—Progress, potentialities, and problems. *Journal of paleolimnology* 20(4): 307–332.
- Birks H, Ter Braak C, Line J, Juggins S and Stevenson A (1990) Diatoms and pH reconstruction. *Phil. Trans. R. Soc. Lond. B* 327(1240): 263–278.
- Birks HJB and Simpson GL (2013) Diatoms and pH reconstruction(1990) revisited. *Journal of Paleolimnology* 49(3): 363–371.
- Booth RK (2002) Testate amoebae as paleoindicators of surface-moisture changes on Michigan peatlands: modern ecology and hydrological calibration. *Journal of Paleolimnology* 28(3): 329–348.
- Booth RK (2008) Testate amoebae as proxies for mean annual water-table depth in Sphagnum-dominated peatlands of North America. *Journal of Quaternary Science* 23(1): 43–57.
- Booth RK (2010) Testing the climate sensitivity of peat-based paleoclimate reconstructions in mid-continental North America. *Quaternary Science Reviews* 29(5-6): 720–731.
- Booth RK, Hotchkiss SC and Wilcox DA (2005) Discoloration of polyvinyl chloride (PVC) tape as a proxy for water-table depth in peatlands: validation and assessment of seasonal variability. *Functional Ecology* 19(6): 1040–1047.
- Booth RK and Jackson ST (2003) A high-resolution record of late-Holocene moisture variability from a Michigan raised bog, usa. *The Holocene* 13(6): 863–876.
- Booth RK and Zygmunt JR (2005) Biogeography and comparative ecology of testate amoebae inhabiting sphagnum-dominated peatlands in the Great Lakes and Rocky Mountain regions of North America. *Diversity and Distributions* 11(6): 577–590.

- Breiman L (2001) Random forests. *Machine learning* 45(1): 5–32.
- Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, Brubaker MA, Guo J, Li P and Riddell A (2016) Stan: A probabilistic programming language. *Journal of Statistical Software* 20(2): 1–37.
- Charles DF (1985) Relationships between surface sediment diatom assemblages and lake-water characteristics in Adirondack lakes. *Ecology* 66(3): 994–1011.
- Charman DJ, Hendon D and Packman S (1999) Multiproxy surface wetness records from replicate cores on an ombrotrophic mire: implications for Holocene palaeoclimate records. *Journal of Quaternary Science* 14(5): 451–463.
- Clifford MJ and Booth RK (2013) Increased probability of fire during late Holocene droughts in northern New England. *Climatic change* 119(3-4): 693–704.
- Davis RB, Anderson DS and Berge F (1985) Palaeolimnological evidence that lake acidification is accompanied by loss of organic matter. *Nature* 316(6027): 436.
- Edwards KJ, Fyfe RM and Jackson ST (2017) The first 100 years of pollen analysis. *Nature Plants* 3: 17001.
- Gajewski K (1987) Environmental history of Caribou Bog, Penobscot Co., Maine. *Naturaliste Canadien* 114: 133–140.
- Gauch H and Whittaker R (1972) Coenocline simulation. *Ecology* 53(3): 446–451.
- Gelman A and Loken E (2013) The garden of forking paths: Why multiple comparisons can be a problem, even when there is no fishing expedition or p-hacking and the research hypothesis was posited ahead of time. *Department of Statistics, Columbia University* .
- Gelman A, Simpson D and Betancourt M (2017) The prior can often only be understood in the context of the likelihood. *Entropy* 19(10): 555.
- Gneiting T (2011) Making and evaluating point forecasts. *Journal of the American Statistical Association* 106(494): 746–762.
- Goring S, Williams J, Blois J, Jackson S, Paciorek C, Booth R, Marlon J, Blaauw M and Christen J (2012) Deposition times in the northeastern United States during the Holocene: establishing valid priors for Bayesian age models. *Quaternary Science Reviews* 48: 54–60.
- Hakim GJ, Emile-Geay J, Steig EJ, Noone D, Anderson DM, Tardif R, Steiger N and Perkins WA (2016) The Last Millennium Climate Reanalysis Project: Framework and first results. *Journal of Geophysical Research: Atmospheres* 121(12): 6745–6764.

- Hefley TJ, Broms KM, Brost BM, Buderman FE, Kay SL, Scharf HR, Tipton JR, Williams PJ and Hooten MB (2017) The basis function approach for modeling autocorrelation in ecological data. *Ecology* 98(3): 632–646.
- Hobbs NT and Hooten MB (2015) *Bayesian models: a statistical primer for ecologists*. Princeton University Press.
- Hoeting JA, Madigan D, Raftery AE and Volinsky CT (1999) Bayesian model averaging: a tutorial. *Statistical Science* : 382–401.
- Hooten MB and Hobbs N (2015) A guide to Bayesian model selection for ecologists. *Ecological Monographs* 85(1): 3–28.
- Hu FS and Davis RB (1995) Postglacial development of a Maine bog and paleoenvironmental implications. *Canadian Journal of Botany* 73(4): 638–649.
- Hustedt F (1937-39) Systematische und kologische unter- suchungen ber die diatomeenflora von Java, Bali und Sumatra nach dem material der deutschen limnologischen Sunda-Expedition. i. *Arch Hydrobiol. (suppl.)* 15-16.
- Hydro2k P (2017) Comparing proxy and model estimates of hydroclimate variability and change over the Common Era. *Climate of the Past* 13(12): 1851–1900.
- Imbrie J and Kipp N (1971) A new micropaleontological method for quantitative paleoclimatology: Application to a late Pleistocene Caribbean core. *The Late Cenozoic Glacial Ages*. Yale Univ. Press, New Haven 3: 71–181.
- Imbrie J, van Donk J and Kipp NG (1973) Paleoclimatic investigation of a late pleistocene Caribbean Deep-Sea Core: Comparison of isotopic and faunal methods 1. *Quaternary Research* 3(1): 10–38.
- Imbrie J and Webb T (1981) Transfer functions: calibrating micropaleontological data in climatic terms. In: *Climatic Variations and Variability: Facts and Theories*. Springer, pp. 125–134.
- Iversen J (1944) Viscum, Hedera and Ilex as climate indicators: A contribution to the study of the post-glacial temperature climate. *Geologiska Föreningen i Stockholm Förhandlingar* 66(3): 463–483.
- Jackson ST (2012) Representation of flora and vegetation in Quaternary fossil assemblages: known and unknown knowns and unknowns. *Quaternary Science Reviews* 49: 1–15.
- Jackson ST and Williams JW (2004) Modern analogs in Quaternary paleoecology: here today, gone yesterday, gone tomorrow? *Annual Review of Earth and Planetary Sciences* 32.

- Juggins S (2013) Quantitative reconstructions in palaeolimnology: new paradigm or sick science? *Quaternary Science Reviews* 64: 20–32.
- Juggins S (2017) *rioja: Analysis of Quaternary Science Data*. URL <http://www.staff.ncl.ac.uk/stephen.juggins/>. R package version 0.9-15.1.
- Juggins S and Birks HJB (2012) Quantitative environmental reconstructions from biological data. In: *Tracking environmental change using lake sediments*. Springer, pp. 431–494.
- Liaw A and Wiener M (2002) Classification and Regression by randomForest. *R News* 2(3): 18–22. URL <http://CRAN.R-project.org/doc/Rnews/>.
- Markel ER, Booth RK and Qin Y (2010) Testate amoebae and  $\delta^{13}\text{C}$  of Sphagnum as surface-moisture proxies in Alaskan peatlands. *The Holocene* 20(3): 463–475.
- Marlon JR, Pederson N, Nolan C, Goring S, Shuman B, Robertson A, Booth R, Bartlein PJ, Berke MA, Clifford M et al. (2017) Climatic history of the northeastern United States during the past 3000 years. *Climate of the Past* 13(10): 1355.
- Mitchell EA, Payne RJ and Lamentowicz M (2008) Potential implications of differential preservation of testate amoeba shells for paleoenvironmental reconstruction in peatlands. *Journal of Paleolimnology* 40(2): 603–618.
- Nygaard G (1956) Ancient and recent flora of diatoms and chrysophyceae in Lake Gribso. Studies on the humic acid Lake Gribso. *Folia Limnologica Scandinavica* 8: 32–94.
- Oehlert GW (1988) Interval estimates for diatom-inferred lake pH histories. *Canadian Journal of Statistics* 16(1): 51–60.
- Oksanen J, Läärä E, Huttunen P and Meriläinen J (1990) Maximum likelihood prediction of lake acidity based on sedimented diatoms. *Journal of Vegetation Science* 1(1): 49–56.
- Overpeck J, Webb T and Prentice I (1985) Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs. *Quaternary Research* 23(1): 87–108.
- Parnell A (2014) Bchron: Radiocarbon dating, age-depth modelling, relative sea level rate estimation, and non-parametric phase modelling. *R package version* 4(1).
- Payne RJ and Mitchell EA (2009) How many is enough? Determining optimal count totals for ecological and palaeoecological studies of testate amoebae. *Journal of Paleolimnology* 42(4): 483–495.

- Payne RJ, Telford RJ, Blackford JJ, Blundell A, Booth RK, Charman DJ, Lamentowicz Ł, Lamentowicz M, Mitchell EA, Potts G et al. (2012) Testing peatland testate amoeba transfer functions: appropriate methods for clustered training-sets. *The Holocene* 22(7): 819–825.
- Phleger FB (1953) North Atlantic core foraminifera. *Reports on the Swedish Deep-Sea Expedition* : 1–122.
- Prentice IC (1980) Multidimensional scaling as a research tool in Quaternary palynology: a review of theory and methods. *Review of Palaeobotany and Palynology* 31: 71–104.
- R Core Team (2017) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Renberg I, Hellberg T and Nilsson M (1985) Effects of acidification on diatom communities as revealed by analyses of lake sediments. *Ecological Bulletins* : 219–223.
- Salonen JS, Ilvonen L, Seppä H, Holmström L, Telford RJ, Gaidamavičius A, Stančikaitė M and Subetto D (2012) Comparing different calibration methods (WA/WA-PLS regression and Bayesian modelling) and different-sized calibration sets in pollen-based quantitative climate reconstruction. *The Holocene* 22(4): 413–424.
- Schapire RE (1990) The strength of weak learnability. *Machine learning* 5(2): 197–227.
- Shuman BN, Routson C, McKay N, Fritz S, Kaufman D, Kirby ME, Nolan C, Pederson GT and St-Jacques JM (2018) Placing the Common Era in a Holocene context: Millennial-to-centennial patterns and trends in the hydroclimate of North America over the past 2000 years. *Clim. Past* .
- Smol JP (2009) *Pollution of lakes and rivers: a paleoenvironmental perspective*. John Wiley & Sons.
- Sullivan ME and Booth RK (2011) The potential influence of short-term environmental variability on the composition of testate amoeba communities in Sphagnum peatlands. *Microbial Ecology* 62(1): 80–93.
- Swindles G, Charman D, Roe H and Sansum P (2009) Environmental controls on peatland testate amoebae (Protozoa: Rhizopoda) in the North of Ireland: implications for Holocene palaeoclimate studies. *Journal of Paleolimnology* 42(1): 123–140.
- Swindles GT, Holden J, Raby CL, Turner TE, Blundell A, Charman DJ, Menberu MW and Kløve B (2015) Testing peatland water-table depth transfer functions using high-resolution hydrological monitoring data. *Quaternary Science Reviews* 120: 107–117.

- Talbot J, Richard P, Roulet N and Booth R (2010) Assessing long-term hydrological and ecological responses to drainage in a raised bog using paleoecology and a hydrosere. *Journal of Vegetation Science* 21(1): 143–156.
- Telford R, Andersson C, Birks H and Juggins S (2004) Biases in the estimation of transfer function prediction errors. *Paleoceanography* 19(4).
- Telford R and Birks H (2009) Evaluation of transfer functions in spatially structured environments. *Quaternary Science Reviews* 28(13-14): 1309–1316.
- ter Braak CJ and Juggins S (1993) Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. In: *Twelfth International Diatom Symposium*. Springer, pp. 485–502.
- ter Braak CJ and van Dam H (1989) Inferring pH from diatoms: a comparison of old and new calibration methods. *Hydrobiologia* 178(3): 209–223.
- Tipton JR, Hooten MB, Nolan C, Booth RK and McLachlan J (2019) Predicting paleoclimate from compositional data using multivariate Gaussian process inverse prediction. *arXiv e-prints* : arXiv:1903.05036.
- Toivonen HT, Mannila H, Korhola A and Olander H (2001) Applying Bayesian statistics to organism-based environmental reconstruction. *Ecological Applications* 11(2): 618–630.
- Trachsel M and Telford RJ (2016) Estimating unbiased transfer-function performances in spatially structured environments. *Climate of the Past* 12(5): 1215–1223.
- Tsyganov AN, Mityaeva OA, Mazei YA and Payne RJ (2016) Testate amoeba transfer function performance along localised hydrological gradients. *European Journal of Protistology* 55: 141–151.
- Van Bellen S, Mauquoy D, Payne RJ, Roland TP, Daley TJ, Hughes PD, Loader NJ, Street-Perrot FA, Rice EM and Pancotto VA (2014) Testate amoebae as a proxy for reconstructing Holocene water table dynamics in southern Patagonian peat bogs. *Journal of Quaternary Science* 29(5): 463–474.
- van Bellen S, Mauquoy D, Payne RJ, Roland TP, Hughes PD, Daley TJ, Loader NJ, Street-Perrott FA, Rice EM and Pancotto VA (2017) An alternative approach to transfer functions? Testing the performance of a functional trait-based model for testate amoebae. *Palaeogeography, Palaeoclimatology, Palaeoecology* 468: 173–183.
- von Post L (1918) Forest tree pollen in south Swedish peat bog deposits, translated by Margaret B. Davis and Knut Faegri, 1967. *Pollen et Spores* 9: 375–401.

- Warner BG and Charman DJ (1994) Holocene changes on a peatland in northwestern Ontario interpreted from testate amoebae (Protozoa) analysis. *Boreas* 23(3): 270–279.
- Webb T and Bryson RA (1972) Late- and postglacial climatic change in the northern Midwest, USA: quantitative estimates derived from fossil pollen spectra by multivariate statistical analysis. *Quaternary Research* 2(1): 70–115.
- Woodland WA, Charman DJ and Sims PC (1998) Quantitative estimates of water tables and soil moisture in Holocene peatlands from testate amoebae. *The Holocene* 8(3): 261–273.
- Wright HE, Mann D and Glaser P (1984) Piston corers for peat and lake sediments. *Ecology* 65(2): 657–659.
- Wright HE, Winter TC and Patten HL (1963) Two pollen diagrams from southeastern Minnesota: problems in the regional late-glacial and postglacial vegetational history. *Geological Society of America Bulletin* 74(11): 1371–1396.

B. APPENDIX B

SUPPLEMENTARY INFORMATION FOR COMPARING AND IMPROVING  
METHODS FOR RECONSTRUCTING PEATLAND WATER-TABLE DEPTH FROM  
TESTATE AMOEBAE



Radiocarbon Lab Sample ID	Depth (cm)	Material dated	14C Age (14C years)	14C Error
OS-122006	35.5	<i>Sphagnum</i> stems	315	15
UCIAMS-190619	44.5	<i>Sphagnum</i> stems	490	20
UCIAMS-190620	53.5	<i>Sphagnum</i> stems	880	20
OS-122007	64.5	<i>Sphagnum</i> stems	1390	15
UCIAMS-190621	76.5	<i>Sphagnum</i> stems	1740	20
UCIAMS-190622	93.5	<i>Sphagnum</i> stems	2120	15
OS-122008	107.5	<i>Sphagnum</i> stems	2430	20
UCIAMS-190623	120.5	<i>Sphagnum</i> stems	2530	15
UCIAMS-194158	129.5	<i>Sphagnum</i> stems	2675	20
OS-122009	133.5	<i>Sphagnum</i> stems	3800	20
UCIAMS-194159	141.5	<i>Sphagnum</i> stems	2935	20
UCIAMS-190624	152.5	<i>Sphagnum</i> stems	3620	15
OS-122010	167.5	<i>Sphagnum</i> stems	3910	20
UCIAMS-190625	188.5	<i>Sphagnum</i> stems	4150	20
OS-122011	208.5	<i>Sphagnum</i> stems	4430	20
UCIAMS-190626	227.5	<i>Sphagnum</i> stems	4595	20
OS-122012	247.5	<i>Sphagnum</i> stems	4660	25
UCIAMS-190627	258.5	<i>Sphagnum</i> stems	4915	20
UCIAMS-190628	263.5	<i>Sphagnum</i> stems	5190	15
UCIAMS-190629	288.5	<i>Sphagnum</i> stems	5895	20
OS-122013	298.5	<i>Sphagnum</i> stems	6020	25
UCIAMS-190630	317.5	<i>Sphagnum</i> stems	6175	15
OS-121990	342.5	<i>Sphagnum</i> stems	6410	25

TABLE B.1. Radiocarbon dates for Caribou Bog

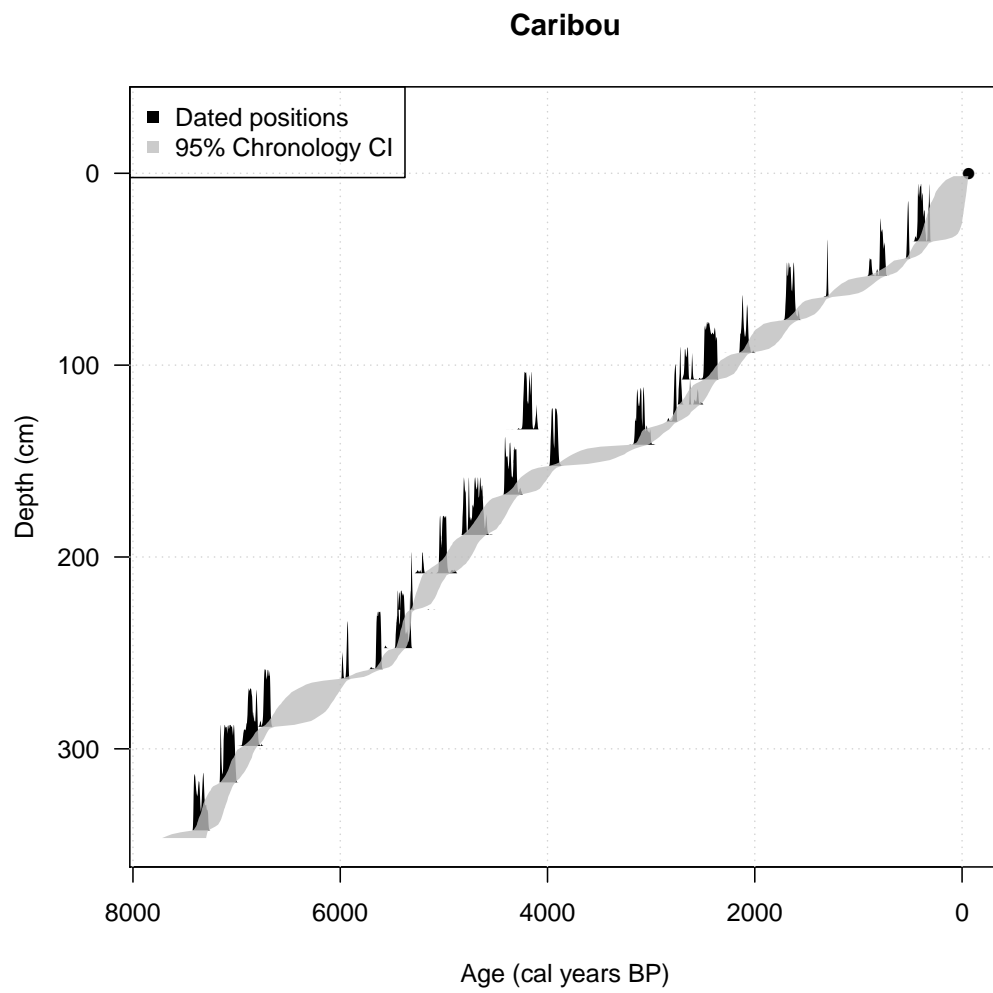


FIGURE B.1. Age-depth model for Caribou Bog

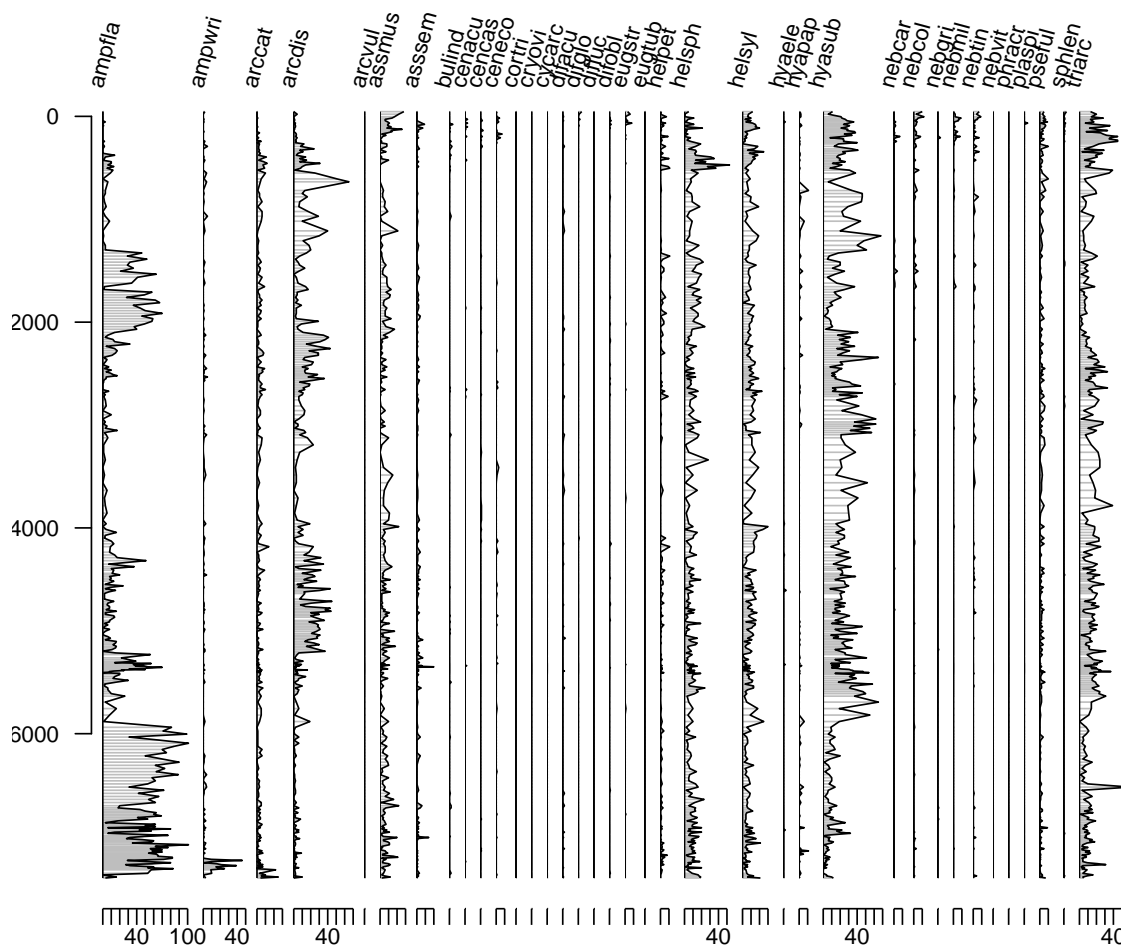


FIGURE B.2. Full testate amoebae data for Caribou Bog. The y-axis is calibrated age. Each x-axis is percent abundance. The species names abbreviations are as follows: ampfla = *Amphitrema flagellum*, ampwri = *A. wrightianum*, arccat = *Arcella catinus*, arcdis = *A. discoides*, arcvul = *A. vulgaris*, assmus = *Assulina muscorum*, assem = *A. seminulum*, bulind = *Bulinaria indica*, cenacu = *Centropyxis aculeata*, cencas = *C. cassis*, ceneco = *C. ecornis*, cortri = *Corythion trinema*, cryovi = *Cryptodiffugia oviformis*, cycarc = *Cyclopyxis arcelloides*, difacu = *Diffugia acuminata*, difglo = *D. globulosa*, difluc = *D. lucida*, difobl = *D. oblonga*, eugstr = *Euglypha strigosa*, eugtub = *E. tubercleata*, helpet = *Heleopera petricola*, helsph = *H. sphagni*, helsyl = *H. sylvatica*, hyaele = *Hyalosphenia elegans*, hyapap = *H. papillio*, hyasub = *H. subflava*, nebcar = *Nebela carinata*, nebcoll = *N. collaris-bohemica*-type, nebgri = *N. griseola*, nebmil = *N. militaris*, nebtin = *N. tincta*, nebvit = *N. vitraea*, phracr = *Phryganella acropodia*, plaspi = *Placcista spinosa*, pseful = *Pseudodiffugia fulva*, sphlen = *Sphenoderia lenta*, triarc = *Trigonopyxis arculeata*

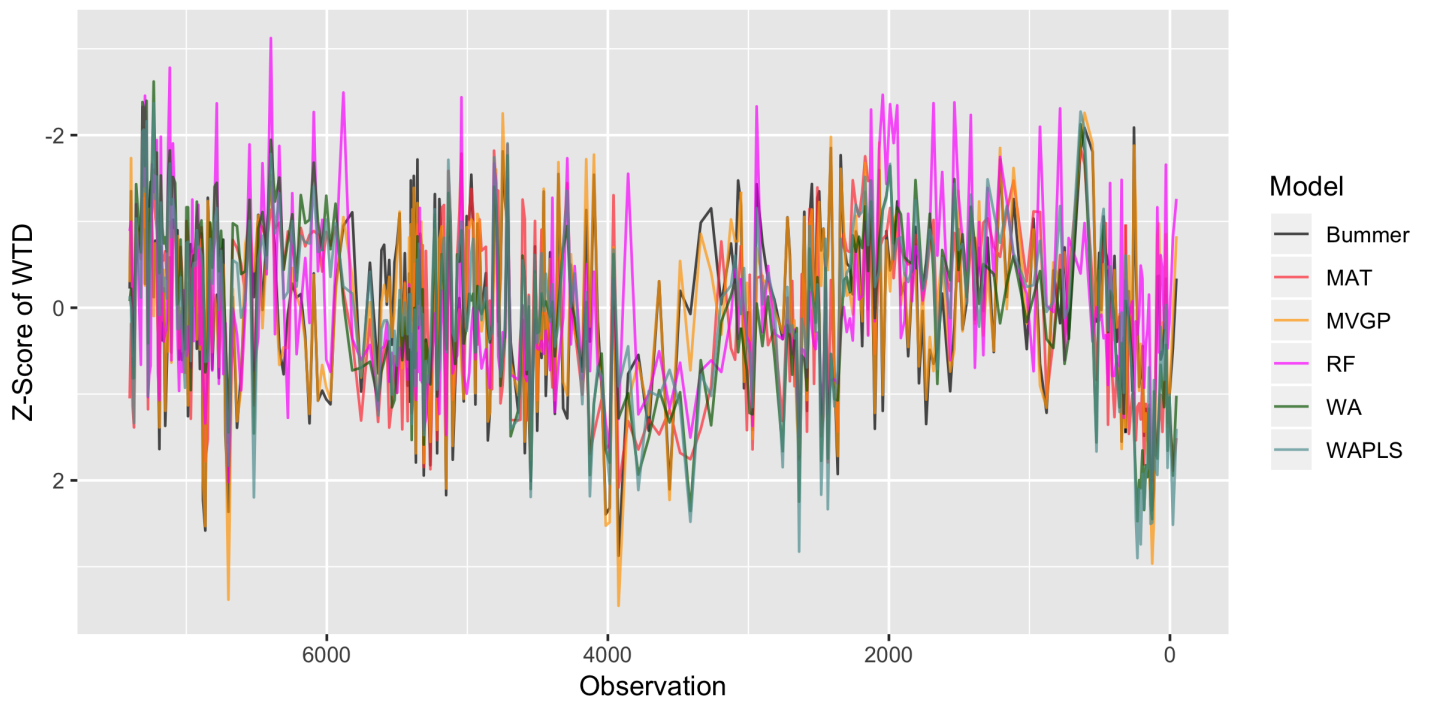


FIGURE B.3. Z-scores of the water-table depth reconstructions resulting from the seven transfer function models applied to the testate amoeba data from Caribou Bog

C. APPENDIX C

USING CO-LOCATED LAKE-LEVEL AND BOG WATER-TABLE DEPTH RECORDS  
TO UNDERSTAND HOLOCENE CLIMATE AND VEGETATION CHANGES IN  
MAINE

Manuscript prepared for submission to *Quaternary Science Reviews*

## **Using co-located lake-level and bog water-table depth records to understand**

### **Holocene climate and vegetation changes in Maine**

Connor Nolan<sup>1</sup>, Robert K. Booth<sup>2</sup>, Bryan N. Shuman<sup>3</sup>, and Stephen T. Jackson<sup>4,1</sup>

<sup>1</sup> Department of Geosciences, University of Arizona, Tucson, AZ 85721, USA

<sup>2</sup> Department of Earth and Environmental Sciences, Lehigh University, Bethlehem, PA 18015, USA

<sup>3</sup> Department of Geology and Geophysics, University of Wyoming, Laramie, WY 82071, USA

<sup>4</sup> U.S. Geological Survey, Southwest Climate Adaptation Science Center, Tucson, AZ 85721, USA

#### C.1. Abstract

Paleoecological and paleoclimate research produces long-term records that provide concrete examples of complex interactions between climate drivers and ecological change at multiple spatial and temporal scales. To better understand these interactions, networks of high-quality ecological records and independent paleoclimate records are required, together with detailed understanding of the records that make up those networks. To this end, we present a new set of co-located climatic and ecological records from a lake and bog in Maine. A lake-level record from Giles Pond and, from 40 km to the northeast, a testate amoeba-inferred water-table depth record from Caribou Bog support inference on

past hydroclimate conditions. These two proxies have never been systematically compared at nearby sites. A pollen record of past vegetation and a micro-charcoal record of past fire activity, both from Caribou Bog, provide vegetation and disturbance history. We find that the new lake-level and bog water-table depth records provide robust, though sometimes contrasting, information on past climate. For example, the lake-level record from Giles Pond is consistent with other lake-level records from New England, suggesting it recorded regional-scale hydrological variability. The bog water-table depth record is consistent with the bog micro-charcoal record, suggesting it recorded ecologically relevant hydrologic variability. The pollen record contains many familiar features of Holocene vegetation sequence of the region. When used in combination, this pair of records explains more about the overall vegetation and climate history than either individual record alone. Altogether, these new records illustrate the richness and complexity of the last 7,500 years of climate and vegetation in Maine.

## C.2. Introduction

Understanding interactions between past changes in climate and past changes in vegetation has been a key interest of palynology for more than a century (Edwards et al., 2017). Pollen preserved in sediment cores showed that vegetation of the past was often vastly different from that of the present. This was a powerful and tangible insight into the scale of past changes in vegetation and the changes in climate necessary to drive those changes (von

Post, 1918; Deevey, 1939; Cain, 1939). In the early 1970s, it became possible to quantitatively infer climate variables (e.g., annual and seasonal precipitation and temperature) from pollen assemblages (Imbrie and Kipp, 1971; Webb and Bryson, 1972). Paleoclimate inferred from pollen records around the globe have become a major source of data for understanding terrestrial climate over the last 20,000 years (COHMAP Members, 1988).

In the 1970s and 1980s, independent terrestrial paleoclimate proxies (e.g., lake levels, stable isotopes, diatoms, chironomids) were developed and refined (Digerfeldt, 1974; Winkler et al., 1986; Harrison, 1989; Guiot et al., 1993; Singer et al., 1996). Multiproxy studies often combined evidence from pollen with one or more independent paleoclimate proxies to understand local and regional climate sequences (Levesque et al., 1993; Fritz et al., 1991; Laird et al., 1998).

Ecologists were interested in using pollen records to investigate ecological questions (Davis, 1981; Davis et al., 1986; Jackson and Whitehead, 1991; Davis et al., 1992). To do this effectively, pollen could not be relied on as evidence for both past climate and vegetation change. Independent inference on past climate changes was necessary to understand details of the interactions between climate and vegetation change (Davis, 1978; Ritchie, 1986). Independent use of paleoclimate and paleoecological records began in earnest in the late 1990s and early 2000s including Weng and Jackson (1999); Shuman et al. (2001); Jackson and Booth (2002); Webb III et al. (2003); Booth and Jackson (2003); Booth et al. (2004).

In eastern North America two prominent independent terrestrial paleoclimate proxies



are lake-level records (Digerfeldt, 1986; Shuman et al., 2018) and peatland water-table depth records based on testate amoebae (Booth and Jackson, 2003; Clifford and Booth, 2013). Changes in lake level are related to annual changes in the relative balance of precipitation and evaporation (Digerfeldt, 1986). Lake-level records are particularly good at recording slow variations, such as long-term trends and multi-centennial-to-millennial variability. Lake-level records may not capture multi-decadal and shorter variability. In some cases, lake-level variations can be driven by non-climatic processes (Huybers et al., 2016). Variability in reconstructed lake level at an individual site can be limited by many factors including lake morphometry, coring locations, catchment size, and the presence of outlets.

Lake-level records were compared synoptically to vegetation as a part of the COHMAP synthesis (Harrison, 1989; Guiot et al., 1993), and were also compared in more detail at the site level (Weng and Jackson, 1999; Almquist et al., 2001; Shuman et al., 2004). Lake-level reconstruction from sites in southern New England generated using the same field, lab, and numerical techniques (Pribyl and Shuman, 2014) has led to records that are directly comparable and reproducible across the region (Newby et al., 2014; Shuman and Burrell, 2017). Lake-level records have also shown to be consistent with precipitation minus evaporation inferred from pollen records (Marsicek et al., 2013; Shuman et al., 2019b).

Peatland records of testate amoeba-inferred water-table depth have been used as proxies for past hydrologic conditions in North America and around the globe (Booth and Jackson, 2003). Testate amoebae-inferred water-table depth records from ombrotrophic peatlands are particularly sensitive to decadal-to-multidecadal wet and dry events (Charman et al.,

2006; Booth, 2010; Morris et al., 2015). Ombrotrophic, "cloud fed", bogs are raised relative to the surrounding topography and thus the sole moisture input is precipitation onto the surface of the bog. The major driver of testate-amoebea community composition is summer moisture deficit (Charman, 2007; Charman et al., 2009). Bogs are insensitive to winter precipitation as long as the bog is saturated during the fall (Charman, 2007). Centennial-to-millennial and longer trends in water-table depth reconstructions may be attributable to changes in peat accumulation rates or peatland size (Waddington et al., 2015; Morris et al., 2015; Swindles et al., 2012; Booth, 2002; Charman et al., 2006). When used in combination with pollen, testate amoebae records have proven useful in understanding the climatic context for observed changes in vegetation at decadal to centennial scales (Wilmshurst et al., 2002; Booth et al., 2012,a; Minckley et al., 2012).

Lake-level work in North America has focused on southern New England and the Rocky Mountains. Bog water-table depth reconstructions have focused on northern New England and the Upper Midwest. These two important terrestrial paleoclimate proxies have never been developed from adjacent sites subject to the same regional climate variability and change. Both lake-level and bog water-table depth records have 'worked' in the past to explain independently verifiable changes in climate and vegetation, but the relationship and potential complementarity of lake-level and bog water-table depth records has not been explored.

Here we present a pair of co-located lake-level and bog water-table depth records, developed from sites 40 km apart in Maine. We validate the inferences from each independent

paleoclimate record and seek to integrate these two unique proxies to better understand the climate and vegetation history of the area across temporal scales. The vegetation history of the northeast US has events that appear to be clearly related to climate at both long and short time scales. Holocene patterns, (e.g., in *Pinus*, *Picea*, *Tsuga*, *Fagus*) may be explained by first-order climate changes (Shuman et al., 2004). Short-term events have also been explained by decadal-scale climate changes (Clifford and Booth, 2015; Booth et al., 2012). Other changes are more complicated or dependent on ecological context and may not be easily explainable by broad-scale climate control (Jackson and Whitehead, 1991; Booth et al., 2012a). We show that using these records in concert leads to a more clear and textured understanding of past climate and vegetation change.

### C.3. Sites and methods

In this paper, we present new data from two sites in Maine: Giles Pond, a lake near Aurora, Maine and Caribou Bog, a raised ombrotrophic peatland near Old Town, Maine. Caribou Bog is 40 km ENE of Giles Pond. In this section, we introduce each site and the methods used to develop and compare and integrate the new records.

#### C.3.1. Giles Pond

Giles Pond (44.83536 N, -68.31715 W; elevation 115 m) is a small, shallow lake (23 hectares, max depth 2.4 m) formed in a glacial depression with a gently sloping bottom.

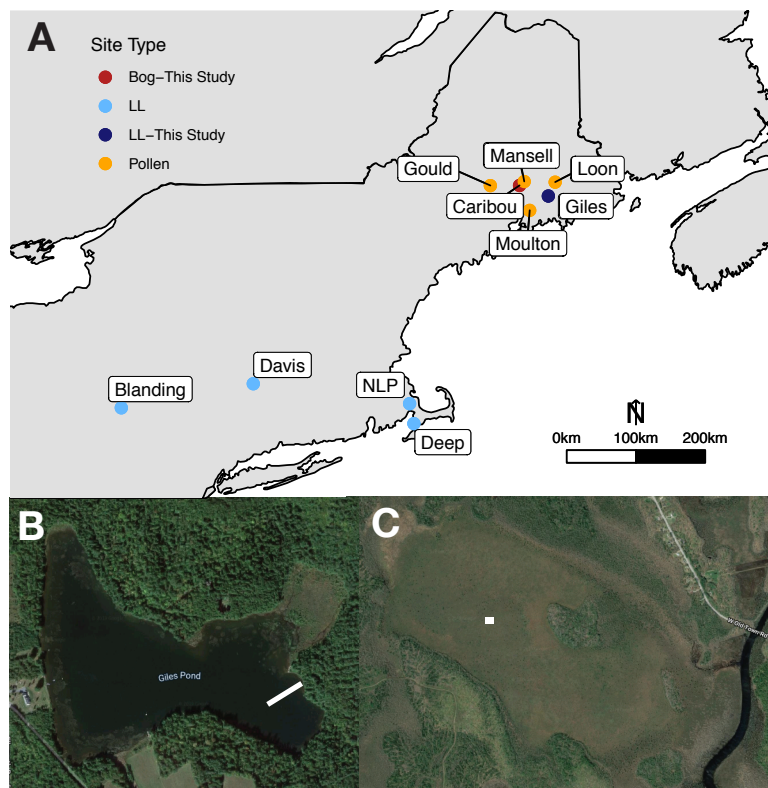


FIGURE C.1. A. Map of sites used in this paper. Pollen data were used from Mansell Pond (very near Caribou Bog), Moulton Pond (south of Caribou Bog), Gould Pond (west of Caribou Bog), and Loon Pond (east of Caribou Bog). GGC30 is a marine core off the Scotian Margin with a published alkenone SST record (Sachs, 2007). Lake-level (LL) sites are shown as blue dots. B. Google Earth view of Giles Pond with approximate coring transect shown as a white line. C. Google Earth view of Caribou Bog with approximate coring location shown as a white square.

Today, there is a small natural stream outlet with a shallow channel. Once the outlet was cut, it became a spillover point that limits the water level from rising much above the modern level. During dry periods, the lake can drop below the level of the outlet, shutting it off seasonally or annually.

The major idea of lake-level reconstruction is to track the elevation of the sand-mud boundary near the edge of the lake through time. Changes in the elevation of this boundary between the littoral (sandy sediment accumulated in shallow water) and profundal (gyttja and other organic-rich sediment deposited in deep water) track changes in the overall hydrologic status of the lake. As lake-level rises, the boundary moves laterally (toward the shoreline) and up in elevation. When lake-level decreases, the boundary retreats toward the interior of the basin and down in elevation.

In the field, we surveyed the lake with ground penetrating radar (GPR). The GPR survey showed many reflective layers that appeared in a consistent pattern around the basin. This indicates the presence of sand layers caused by past lake-level declines. To capture the salient GPR features in sediment cores we chose a representative transect in the southwest corner of the lake. Using a piston corer, we collected four cores from along the transect, one each at 20 m, 31 m, 39 m, and 50 m from shore. We collected a total of 600cm of sediment.

Cores were described, photographed, and run through the scanning micro-XRF at Woods Hole Oceanographic Institution before shipped to University of Arizona where they were sampled for radiocarbon dating. A total of 25 radiocarbon dates were obtained on charcoal

and bulk sediment samples (Table D.1). The levels for dating were chosen to bracket sand layers of interest. Age-depth models for each core were built using Bchron (Parnell, 2014). Cores 31A and 39A span from present to 15,000 cal yr BP. Core 50A spans from present to 9000 cal yr BP. Core 20 contained a hiatus from 1000 cal yr BP to 11,000 cal yr BP, and thus, was not used in the lake level reconstruction procedure.

Titanium counts from the XRF data were used as a proxy for sand content. Titanium is a conservative tracer for terrestrially-derived material (Shuman et al., 2009). Higher titanium counts are related to increased sand content. We compared the XRF-derived titanium counts for the Giles Pond to the lab photographs and descriptions to confirm that high titanium counts are associated with sand layers. In the deglaciation, much higher Ti content was related to glacial gray clays. Over our target period, the past 10,000 years, Ti is related to sand content in the cores.

To use these data to reconstruct lake level, we use an algorithm from Pribyl and Shuman (2014). The algorithm takes an input file consisting of the depth of each sample in each core, the age of that sample (based on the Bchron-derived age-depth model), and the value of the proxy for sand content (in this case, Ti counts from XRF) (Figure C.2A, D.1). The user sets a range of candidate thresholds for littoral and sub-littoral sediment. For the Giles Pond reconstruction we set candidate thresholds for littoral ranging from 250 to 350 Ti counts and candidate thresholds for sublittoral sediment ranging from 120 to 250 Ti counts. For each iteration, one littoral and one sublittoral threshold is used to classify each sample from each core as either littoral ( $Ti > \text{littoral threshold}$ ), sublittoral ( $\text{littoral threshold} > Ti$

>sublittoral threshold), or profundal ( $T_i < \text{sublittoral threshold}$ ). The algorithm interpolates the core data to 50-year intervals and iteratively classifies the data at each time point based on the thresholds. These classifications and the accumulation history of the cores are used to determine possible lake-level histories consistent with all of the data. For example, if, for a given 50-year interval, the deepest core is classified as profundal and the two cores nearer to shore are classified as littoral then the sand-mud boundary at that time must have been somewhere between the elevation of the deepest core and the second-deepest core. These classifications of each core are shown in figure C.2B. After many iterations of the entire process, a consensus lake-level history and uncertainty are produced by a weighted mean of trajectories consistent with the core data. For further details and implementation the reader is referred to Pribyl and Shuman (2014)

We compare the new lake-level record from Giles Pond to existing lake-level records around the region (Figure C.2C). Figure 1 shows a map of the sites used. Long-term trends in lake-level change are driven by ice-sheet retreat and insolation changes (Shuman and Plank, 2011). The insolation-driven trends are estimated by regressing Northern Hemisphere June insolation (Berger and Loutre, 1991) on each lake-level record, following Shuman and Burrell (2017).

### C.3.2. Caribou Bog

Caribou Bog (44.985 N, 68.814 W; elevation 35 m) is an ombrotrophic raised bog near Old Town, Maine. Previous researchers have documented the Holocene developmental

history of the site (Gajewski, 1987; Hu and Davis, 1995), the hydrologic structure (Comas et al., 2005), and the mercury fluxes over the last 10,000 years (Roos-Barraclough et al., 2006).

Sediment cores for this study were collected in summer 2014. We recovered the top 250 cm using a 10 cm-diameter modified piston corer (Wright et al., 1984). We recovered another 97 cm with a Russian (Jowsey) corer for a total of 347 cm of sediment. Cores were extruded and described in the field.

In the lab, the cores were cut into 1 cm contiguous slices, which were each subsampled for testate amoebae and pollen (1 cm<sup>3</sup>), loss-on-ignition (LOI) (1cm<sup>3</sup>), and humification (3 cm<sup>3</sup>). The testate amoebae/pollen subsamples were processed via standard methods (Booth, 2010).

We obtained 24 radiocarbon dates to develop a chronology for the site. The age model was developed using Bchron Parnell (2014). The basal date of the record is 7,500 cal yr BP. The age model and the core lithology show no signs of depositional hiatuses. The radiocarbon dates and age-depth model are in Nolan et al. (2019).

Testate amoebae were identified and counted under a microscope at 400x magnification based on a standard key (Charman et al., 2000) to a minimum of 100 individuals. A total of 327 samples (nearly every centimeter) were counted for testate amoebae, for a median temporal resolution of 17 years per sample (range: 4-60 years).

The testate amoebae-based water-table depth (WTD) record from Caribou Bog was first published in Nolan et al. (2019), which showed a comparison of seven different transfer



functions. For the purposes of this paper we present only the water-table depth reconstruction based on the simplest model, weighted averaging (ter Braak and van Dame, 1989). We use weighted averaging here because it makes the analyses comparing the water table depth record to pollen and charcoal records more directly analogous to other such examples in the literature (e.g. Clifford and Booth (2013); Booth et al. (2012a)). Furthermore, using WA allows use the full set of samples and species in the North American training dataset, some of which must be lumped together or eliminated in order to fit Bayesian transfer function models.

Because low-frequency variability in the reconstructed WTD record is potentially associated with non-climatic processes (e.g., lateral bog growth and changes in peat accumulation rate), we removed it by fitting a flexible smooth spline to the raw WTD record. Subtracting the spline leaves only the high-frequency variability that is more confidently attributable to climate.

On the same slides used for testate amoebae analysis, micro-charcoal (15-250 micrometers) was also identified and counted. This size class of charcoal is generally related to regional fire activity (Clifford and Booth, 2013; Tinner and Hu, 2003; Innes et al., 2004). Charcoal counts were converted to charcoal accumulation rate (CHAR; pieces  $\text{cm}^{-2} \text{year}^{-1}$ ) by indexing them to the counts of *Lycopodium* spores added at a known concentration and data about the sediment accumulation rate derived from the age model. There were 16 samples in Core 1 Drive 1 to which no *Lycopodium* spores were added and thus charcoal influx cannot be calculated.

To understand the relationship between water-table depth conditions and charcoal influx we divided the data into high charcoal (charcoal influx greater than median charcoal influx) and low charcoal (charcoal influx less than median charcoal influx). We tested the null hypothesis that the water-table depth conditions for the high charcoal and low charcoal groups was indistinguishable from zero using the Mann-Whitney rank-sum test.

From a subset of the samples prepared for testate amoeba analysis, pollen was identified and counted to a minimum count of 250 arboreal grains per sample. A total of 124 samples were counted for pollen for a median resolution of 56 years per sample (range: 10-120 years).

### C.3.3. Combining lake-level and water-table depth reconstructions

We combine the long-term trend and low-frequency variability of the lake-level record with the high-frequency variability of the detrended bog water-table depth record to produce an integrated record of past hydrologic variability. We interpolate each record to common 20-year time steps, convert each record to z-scores where positive z-scores represent wetter conditions and negative z-scores represent drier conditions. Then, we combine the lake-level and detrended water-table depth records via simple addition of the z-scores.

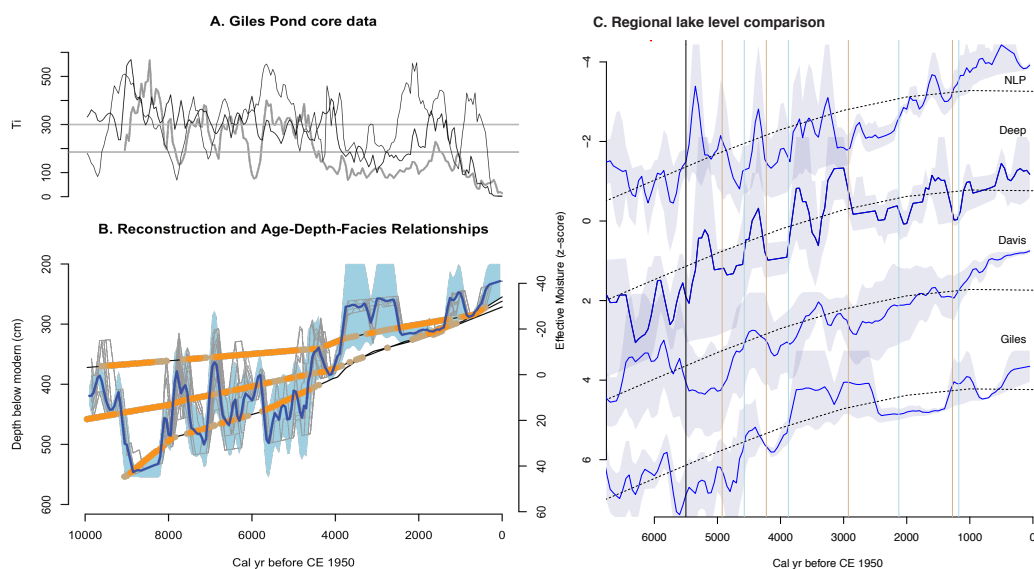


FIGURE C.2. A. Titanium count data for Giles Pond cores 31A (thin black line), 39A (black line), 50A (gray line). B. Lake-level reconstruction from Giles Pond cores. Dark blue line is the reconstructed lake-level, light blue shading is the uncertainty, gray lines are some sample lake-level trajectories from the algorithm, black lines are the cores sedimentation history, orange dots are intervals in that core classified as littoral, tan dots are intervals classified as sub-littoral. C. Z-scores of lake-level reconstructions (offset for clarity) from New Long Pond (NLP) (Newby et al., 2009), Deep Pond (Marsicek et al., 2013), Davis Pond (Newby et al., 2011), and Giles Pond (this study).

## C.4. Results

### C.4.1. Lake-level reconstruction

The cores from Giles Pond contain numerous transitions between sand and mud with coherent patterns between the cores. The Titanium counts between cores (Figure C.2A, D.1) show consistent patterns. The deepest core, core 50A, shows less sand than core 39A, which shows less sand than 31A, the shallowest core. Ti counts were compared in detail to photos and descriptions of the cores to confirm that Ti counts were closely related to visually-identified sand layers.

Figure C.2B presents the 10,000-year lake-level history developed using the reconstruction algorithm. Lake-levels were low in the early Holocene and increased toward the present. This trend is consistent with other reconstructions from the region and is driven by the retreat of the ice sheets and insolation changes over the Holocene (COHMAP Members, 1988; Shuman and Plank, 2011). Superimposed on the long-term trend are several multicentennial-to-millennial modulations toward wetter or drier conditions. Many variations are shared between other lake-level reconstructions from New England (Figure C.2C).

Early Holocene sediments of all three cores from Giles Pond contain sand (Figure C.2B) suggesting low lake levels. After 8000 yr BP, the lake level began rising with variable conditions from 7800-5800 yr BP. Lake level decreased at 5700 yr BP, suggesting drying conditions consistent with the inland southern New England sites (Davis and Blanding). However, this event is absent at the coastal sites (Deep Pond and New Long Pond). Lower

lake-level conditions persisted until 4700 yr BP when lake levels began to increase at Giles Pond and at all of the other sites in New England, reaching high points just after 4500 yr BP. These lake-level increases were followed by a decrease at Giles Pond and all of the other sites to a lowered lake level at 4100 yr BP. Lake levels at Giles Pond increased to near modern levels at 3800 yr BP. The core data clearly suggest an increase in lake level and relatively higher lake levels from 3800 yr BP until 2600 yr BP, but the magnitude and dynamics of this interval are not well constrained by the available core data (Figure C.2B). The timing of the beginning of this wet interval is consistent with the other lake-level records from the region, but Giles Pond stayed high longer (until around 2600 yr BP) than the southern New England sites (which decreased around 2900 yr BP). From 2600 to 1400 yr BP lake levels at Giles appear to have been relatively steady and lower than modern. The sediment limit was between core 31A and core 39A or possibly between core 39A and core 50A from 1800-1600 yr BP. The accumulation history of the cores brings their relative elevations close together during this interval, making the amplitude of any lake-level changes ambiguous. Lake level increased at 1400 yr BP and remained high for about 500 years. This change is in the opposite direction of the southern New England lake-level records, which suggest drying during this interval. Lake levels at Giles were lower than present from about 800-500 yr BP before they rose to modern levels about 350 yr BP.

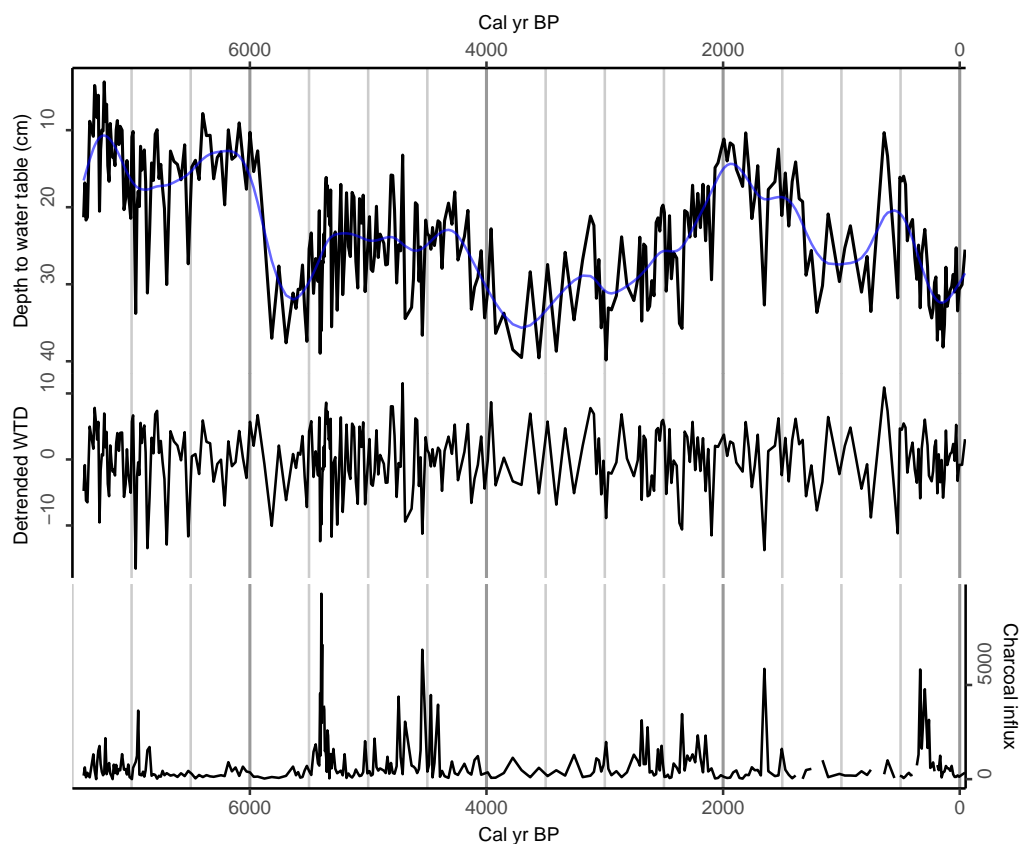


FIGURE C.3. A. Testate amoeba-inferred water-table depth from Caribou Bog with the smooth spline showing the low-frequency variability. B. The detrended water-table depth from Caribou Bog. C. The micro-charcoal from Caribou Bog.

#### C.4.2. Caribou Bog hydrology and disturbance records

*Water table depth reconstruction* Figure C.3A shows the testate amoebae-based water-table depth reconstruction from Caribou Bog. Deeper water-table depth is related to a drier bog surface, and shallower water-table depth is related to a wetter bog surface. Strong decadal-to-multidecadal variability is superimposed on lower frequency trends. Because the low-frequency variability in bog records is potentially related to non-climatic factors, we fit

a flexible, smooth spline with 30 degrees of freedom to the water-table depth record to remove low-frequency variability and target decadal-scale events that are more confidently attributable to climate variability (Figure C.3B).

*Charcoal record* The microcharcoal record shows large swings between periods of minimal charcoal influx and periods of extremely high charcoal influx (Figure C.3C). Moderate charcoal influx from 7500-6800 yr BP was followed by low charcoal from 6800-5500. Maximum charcoal influx in the record occurred around 5400 yr BP, followed by a high but variable period of charcoal until 4300 yr BP. Charcoal influx was generally low in the late Holocene except for moderate influx from 2700-2100 yr BP and large peaks at 1700 and 350 yr BP. High charcoal counts are associated with drier water-table depth conditions ( $P < 0.001$ ; Mann-Whitney rank-sum test; Figure C.4).

#### C.4.3. Integrated lake and bog paleohydrology

The lake-level record and water-table depth records are both related to local and regional hydrologic variability, but at face value they look different from each other (Figure C.2B vs. Figure C.3A). After the removal of the non-climatic low-frequency variability, only decadal-to-multidecadal variability remains in the water-table depth record. This variability complements the long-term trend and the multicentennial-to-millennial scale variability of the lake-level record. We combine these two reconstructions into an integrated record of past hydrologic variability across scales (Figure C.5). The combined record doc-

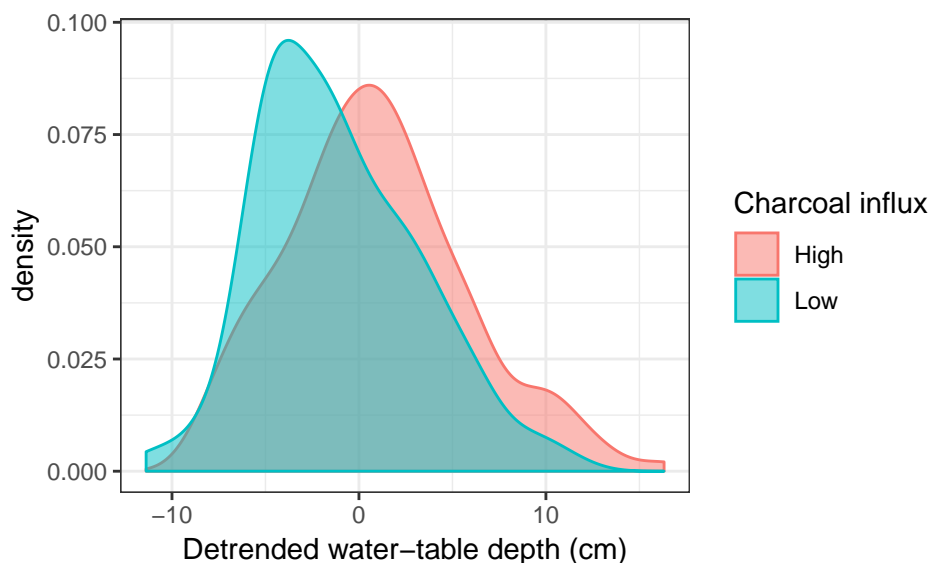


FIGURE C.4. Density plot of the detrended water-table depth (higher values = drier conditions) for samples with high charcoal influx and samples with low charcoal influx. The difference between the distributions is statistically significant ( $P < 0.001$ ; Mann-Whitney rank-sum test).

uments hydrologic variability on timescales of decades to millennia.

#### C.4.4. Caribou Bog pollen record

Here we present the Holocene record for the six primary arboreal pollen taxa from Caribou Bog: *Pinus*, *Tsuga*, *Betula*, *Picea*, *Fagus*, and *Quercus*.

*Pinus*, *Pinus strobus* in particular, is the dominant taxon at the start of the record, representing 50-60% of arboreal pollen from 7.5 to 6.4 ka and generally 40% or greater until 3.2 ka. At 3.2 ka, coincident with wet conditions in the combined lake and bog record, *Pinus* shifted to a new lower percentage, around 30% and *Tsuga* became dominant. Low-frequency variability modulates the relationship between *Pinus* and *Tsuga* in this interval,



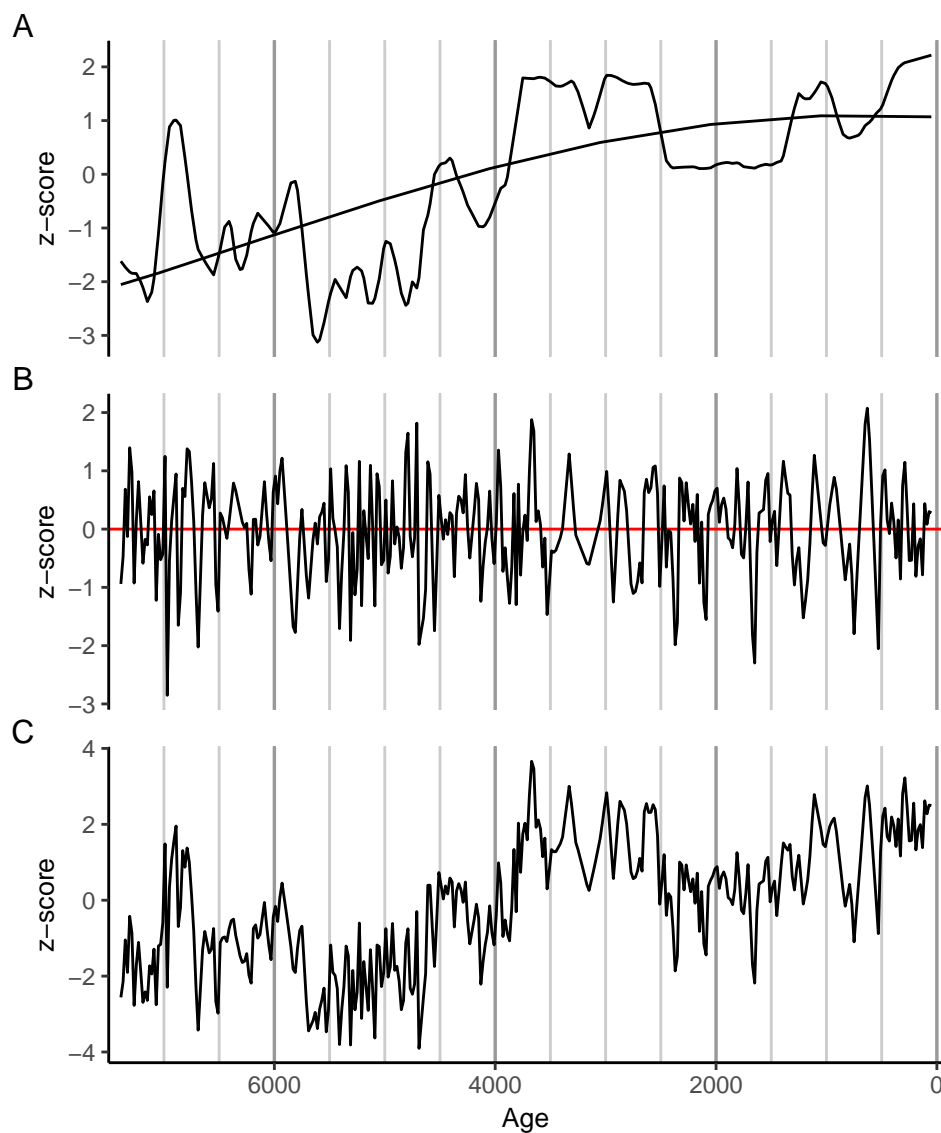


FIGURE C.5. A: Lake-level record from Giles Pond converted to z-scores and plotted with the fitted insolation curve. B: Detrended water-table depth record from Caribou Bog converted to z-scores. C: Integrated lake-level and bog water-table depth paleohydrology record created by adding the z-scores from panel A and B.

including increases in *Pinus* 2000-1800 yr BP and 500 yr BP to present.

*Tsuga* pollen varies from less than 10% to greater than 50% arboreal pollen, sometimes within just a few hundred years, and is strongly anti-correlated with *Pinus*. *Tsuga* began the record low, around 20%, and at 6750 cal yr BP began increasing to its mid-Holocene maximum. From 6.4 to 5.2 ka *Tsuga* was abundant but highly variable. At 5.9 ka, *Tsuga* declined from 44% to 7% and rose back to 45% in less than 250 years.

At 5.2 ka, *Tsuga* sharply declined again and did not immediately recover. This is the local expression of the range-wide mid-Holocene *Tsuga* decline (Bennett and Fuller, 2002; Foster et al., 2006; Booth et al., 2012a). At 4.6 ka, *Tsuga* briefly increases to 40% for 150 years, sharply declined again, and then began a variable, upward trajectory to a maximum at 2.7 ka. *Tsuga* remains dominant, generally above 40%, until 500 yr BP. *Tsuga* drops precipitously from 50% at 650 yr BP to 2% at 250 yr BP and remains low to the present.

*Betula* begins the record around 5% and increases to a maximum of greater than 30% at 4200 yr BP. After 4200 yr BP, *Betula* decreases slightly and remains consistently variable around 20% through the late Holocene. *Betula* attains a short, high peak at 250 yr BP. Macrofossil evidence from Gould Pond suggests *Betula alleghaniensis* generally replaced *Betula papyrifera* after 6000 yr BP (Anderson et al., 1992). Thus, most of the *Betula* pollen in the Caribou Bog record is likely *B. alleghaniensis*.

From 7.5-1.5 ka, *Picea* is generally less than 5% with occasional brief increases up to 15%. From 2000 yr BP to present, *Picea* generally increases with a number of sharp ups and downs. *Picea* pollen in this record derives from both trees growing on the bog surface

(*P. mariana*) and *Picea* in local-to-regional upland forests (mostly *P. rubens* today).

*Fagus* increases from less than 5% at 7.5 ka to 10% at 5.7 ka, then drops suddenly and remains less than 5% until 4.9 ka. From 4.9 ka to 4.0 ka, *Fagus* is more prominent, around 10% before a another sudden drop to less than 5%. *Fagus* increases from 4.0 ka to 3.3 ka and then generally decreases from 3.3 ka to 500 yr BP. At 500 yr BP, *Fagus* increases suddenly to greater than 10% and then, at 350 yr BP, dramatically drops to 1%.

*Quercus* generally ranges from 2-7% throughout the record with peaks up to around 13% at 4450 yr BP, during the early return of *Tsuga*, and at 2800 yr BP, coincident with an increase in *Betula*. The general clustering of higher *Quercus* pollen between 5000 and 4000 yr BP is coincident with higher *Fagus* pollen percentages.

*Acer*, *Alnus*, and *Abies* are the next most abundant taxa, but are all well below 5%.

## C.5. Discussion

### C.5.1. Long-term hydrologic trends and multi-centennial variability at Giles Pond

The lake-level history of Giles Pond exhibits a known insolation-driven trend towards wetter conditions over the past 10,000 years (Shuman and Plank, 2011). The long-term trend is modulated by centennial-to-millennial scale variability (Newby et al., 2014; Shuman and Burrell, 2017), in particular at Giles Pond: (1) dry conditions from 5.7 to 5.5 ka, (2) the sequence of wet conditions at 4.5 ka and dry conditions from 4.2 to 3.9 ka, (3) wet conditions from 3.7 to 2.5 ka with a short dry interruption in the middle, and (4) dry

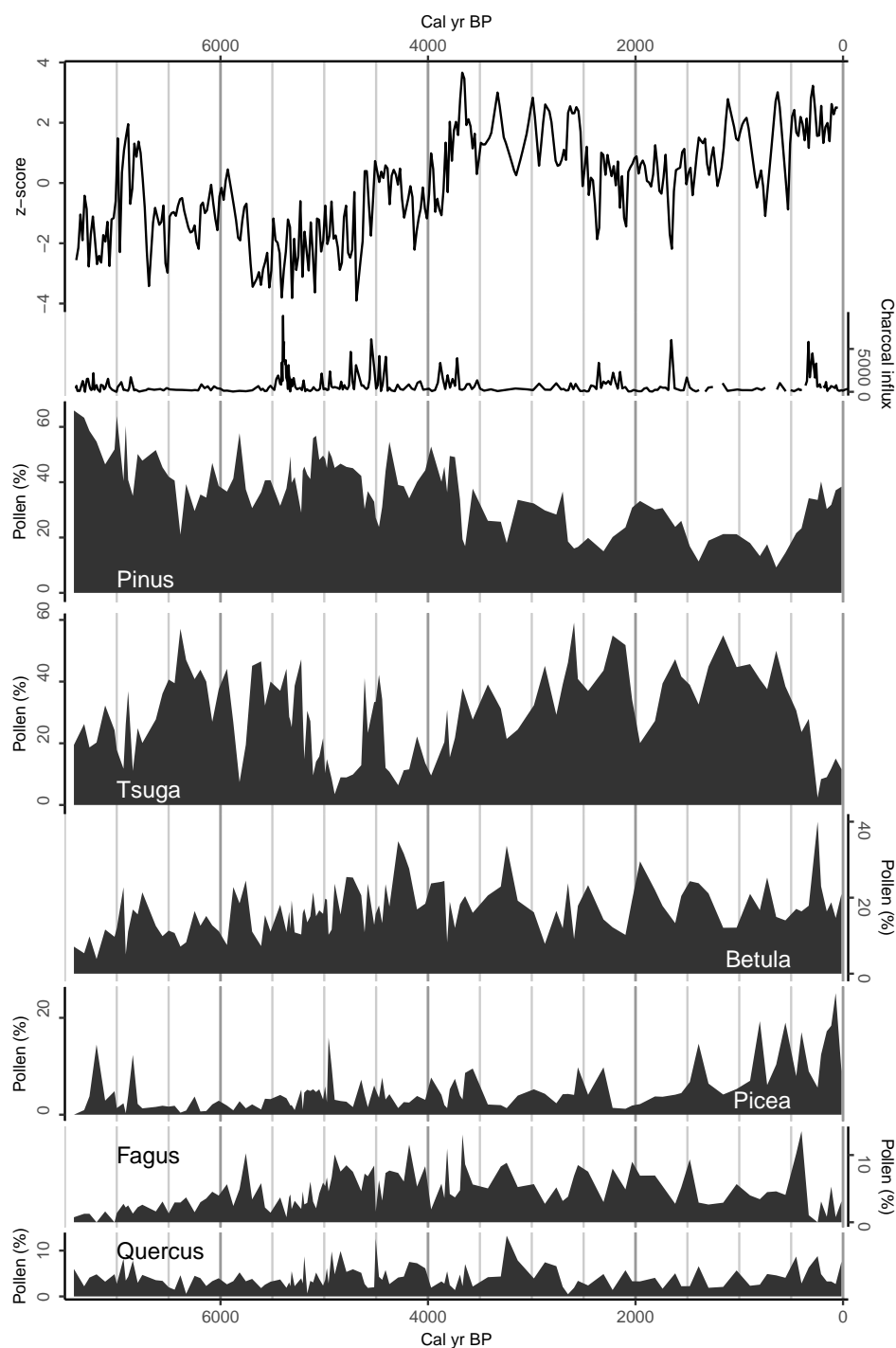


FIGURE C.6. A: Integrated record that combines the Giles Pond lake-level history and Caribou Bog detrended water-table depth record (same as figure C.5C. B: Microcharcoal record from Caribou Bog. C: *Pinus* pollen from Caribou Bog. D: *Tsuga* pollen from Caribou Bog. E: *Betula* pollen from Caribou Bog. F: *Picea* pollen from Caribou Bog. G: *Fagus* pollen from Caribou Bog. H: *Quercus* pollen from Caribou Bog.

conditions at 1.8 ka.

Qualitative lake-level reconstructions from nearby Mansell Pond suggested similar patterns to Giles Pond: specifically, dry conditions prior to 5 ka followed by a variable increase to near modern levels by 3 ka (Almquist et al., 2001). Giles Pond extends the existing network of quantitative lake-level records in southern New England (Newby et al., 2014; Shuman and Burrell, 2017) northward by 500 km. The expanded network allows for assessment of patterns of lake-level change between southern and northern New England. Some multi-century changes appear consistent between records across the northeast US, lending confidence to the interpretation of lake-level records as sensitive to regional climatic variations. Not all events appear in all records because droughts and pluvials have a range of spatial patterns. Sometimes a drought may affect the entire northeastern US, but other times the spatial extent may be more limited. For example, the 1960s drought affected the entire region from eastern Pennsylvania to northern Maine (Seager et al., 2012), but the 1947 drought primarily affected Maine and New Hampshire.

Due to the small perennial stream outlet, Giles Pond lacks sensitivity to record conditions wetter than modern day. The reconstructed lake level rises to near modern levels (e.g. from 3700-2500 BP, 1200 BP, and 500 BP to present), but it is possible that some of these wet periods were actually wetter or more variable than our reconstruction suggests. During drier times, the lake level drops below the level of the outlet stream. Blanding Pond in Pennsylvania also has an outlet that similarly limits its sensitivity in the late Holocene (Shuman and Burrell, 2017).

### C.5.2. High-frequency hydrologic variability and impacts at Caribou Bog

The testate amoebae-based water-table depth reconstruction from Caribou Bog extends the local network of similar records in Maine from Sidney Bog, Saco Bog, and Great Heath Bog (Clifford and Booth, 2013, 2015). Existing records focused only on the most recent 2000 years, particularly the major changes at 500 yr BP. Caribou Bog provides a 7,500-year record of hydrologic variability. The water-table depth reconstruction exhibits both fast and slow changes. Low-frequency variability in bog records is susceptible to non-climatic drivers such as changes in bog growth rate (Morris et al., 2015). Thus, in order to emphasize the variability that is more confidently related to hydroclimate variability, we detrended the water-table depth reconstruction using a smooth spline. Future work could develop more sophisticated approaches to separating climatic and non-climatic drivers of centennial-to-millennial trends in bog water-table depth records.

The simple detrending applied here yields a water-table depth record with many decadal to multi-decadal dry and wet events and more closely matches qualitative interpretation applied to bog water-table depth records. Dry conditions in the detrended water-table depth are related to increased microcharcoal, suggesting that the testate amoebae-based water-table depth reconstruction is sensitive to ecologically meaningful hydrologic variations over the past 7,500 years.

### C.5.3. Integrating lake-level and water-table depth records to reconstruct hydrologic variability across temporal scales

Lake-level and bog water-table depth records in general, and our new records from Giles Pond and Caribou Bog in particular, represent validated proxies for past hydrologic conditions. Although their respective hydrologic histories appear to differ, we can apply expert knowledge about each proxy to synthesize and integrate these records (Jackson, 2012). Lake-level records are smoothed representations of hydrologic variability because it takes large hydroclimatic changes to affect lake levels. Thus, interpretation of lake-level records should focus on long-term trends and multi-centennial to millennial scale variability. In contrast, low-frequency variability in bog records derives primarily from autogenic bog processes and their interaction with climate variability. Thus, for bogs, interpretation should focus on wet versus dry conditions on decadal to multi-decadal time scales (Charman et al., 2009; Swindles et al., 2012).

Given their differential timescales, combining the lake-level and bog water-table depth records results in an integrated history that reasonably describes decadal to multi-millennial hydrologic variability and change over the last 7,500 years (Figure C.5). This integrated record is more like the latent, true hydroclimate than either record alone (e.g., the combined record explains some of the clustering of peaks in the charcoal record) (Figure C.6).

Incorporating expert knowledge in proxy interpretation makes possible equitable comparisons and combinations of records. The simple addition of z-scores ignores some poten-

tial mismatches between the bog and lake proxies, especially in three potentially important areas: seasonality, hydrologic inputs, and influence of temperature. Lake-level changes are sensitive to annual water balance whereas the bog water-table depth reconstructions are most sensitive to the summer season (Charman, 2007). Lakes receive hydrologic input from both runoff and precipitation while precipitation is the only input for ombrotrophic bogs. Changes in bog water-table depth and lake level are sensitive to a balance between precipitation and evaporation. This balance has a temperature component, but the details require additional investigation (Pribyl and Shuman, 2014; Charman et al., 2009). Future work should interrogate the assumptions about proxy sensitivities and mechanisms that underlie the simple addition of z-scores via modern observational studies of lakes and bogs, development of more pairs of co-located records, and development of terrestrial temperature records.

#### C.5.4. Climate and ecological drivers of Holocene vegetation change at Caribou Bog

Our new pollen record from Caribou Bog is similar to a lower-resolution sequence from the site (Gajewski, 1987), but the increased resolution (one sample every 60 years compared to one sample per 700 years) allows for investigation of both slow and fast changes in vegetation over the past 7,500 years. Interpretation of vegetation changes in the light of our independent paleoclimate records from Caribou Bog and Giles Pond provides new insights into climate-vegetation interactions.

Some of the observed changes are consistent with a relatively simple climate-forcing



explanation. For example, the decrease in *Pinus* and replacement by *Tsuga* after 3.5 ka is likely related to a long-term wetting trend (Figure C.2 and C.6a) modulated by disturbance and local ecological interaction. The decline of *Pinus* in the mid-to-late Holocene is common in records around the region, but the timing and magnitude of the decline varies.

*Fagus*'s decline from 5.8 to 5.0 ka may be related to its intolerance to drought and fire. This may be analogous to dynamics in the Great Lakes region related to droughts and fires in the Medieval Climate Anomaly (Booth et al., 2012) and throughout the Holocene (Wang et al., 2016). These events have not been clearly documented in existing pollen records from lakes in Maine, possibly due to the smoothing of lake records (discussion below).

The Holocene history of *Tsuga*, particularly mid-Holocene dynamics, is complicated. A recent study suggests a niche bimodality, where *Tsuga* prefers warm and dry conditions or cool and wet conditions (Shuman et al., 2019a). The mid-Holocene *Tsuga* maximum from 7.0 to 5.2 ka may represent optimal warmer and drier conditions and the later Holocene *Tsuga* maximum from 3.5 to 0.5 ka may represent optimal cooler and wetter conditions. Shuman et al. (2019a) also use this niche bimodality to predict the Holocene dynamics of *Tsuga* pollen in southern New England. Attempts to apply this index to the Caribou Bog record resulted in relatively poor agreement between the timing of climatic changes and the mid-Holocene dynamics of *Tsuga* at Caribou Bog. Major discrepancies between the (Shuman et al., 2019a) approach and *Tsuga* pollen changes at Caribou Bog suggests simple, climate-only explanations may be insufficient to explain mid-Holocene dynamics of *Tsuga* in northern New England.

Booth et al. (2012a) have suggested treating the *Tsuga* decline, low period, and recovery separately as epiphenomena and evaluating their potential drivers independently (Booth et al., 2012a). Theoretical modeling suggests intrinsic ecological processes can also generate abrupt quasi-synchronous tree declines (Ramiadantsoa et al., 2019). The true causes may be some combination of climatic and ecological processes interacting across scales in potentially unpredictable ways. Expanding the network of highly-resolved, precisely-dated pollen and paleoclimate records will further elucidate the drivers of mid-Holocene *Tsuga* dynamics (Booth et al., 2012a).

*Betula*'s low abundance around 7.5 ka and rise to a steady, higher level of abundance after 5.0 ka is consistent with regional patterns. In particular, immigration and expansion of *Betula alleghaniensis* in the Adirondacks, the White Mountains, and Upper Michigan coincided with sustained *Betula* pollen increases (Davis et al., 1980; Jackson and Whitehead, 1991; Jackson et al., 2014). However, in our new record from Caribou Bog *Betula* only reaches 20-30%. This is lower than nearby sites, particularly Mansell Pond and Gould Pond where *Betula* is consistently greater than 40% after 5500 yr BP and much less variable than Caribou Bog (Almquist-Jacobson and Sanger, 1995; Anderson et al., 1992). Differences in *Betula* pollen abundance between these sites could be due to climatic or edaphic differences between sites, but fundamental differences between pollen records from bogs versus lakes may also contribute to the discrepancies. In lakes, pollen grains are deposited on the water surface, become inundated, and then sink in the water column to eventually be deposited in the sediment (Davis et al., 1971). In contrast, the pollen record from an

ombrotrophic bog consists solely of pollen that is deposited onto the surface of the bog at the coring site. Differential imbibition and sinking rates between pollen taxa can then affect the ratio of different types of pollen grains that ultimately are preserved in lake sediment relative to the ratio that would be preserved in surface samples or bog sediments (Davis and Brubaker, 1973). Furthermore, pollen records from lakes are smoothed because they represent a broader sample of airborne pollen and because sediment concentration and mixing smooths the record that is ultimately preserved in lake sediment. Future work could systematically compare sets of nearby pollen records and surface pollen samples from different depositional environments (e.g. lake vs. bog vs. moss pollster) to better understand the differences in the pollen records preserved in these two archives.

*Picea* in the bog pollen records represents both local (bog surface) and regional upland populations. In local lake pollen records *Picea* is present in only trace amounts (<2%) between 10 ka and 2ka (Almquist-Jacobson and Sanger, 1995; Anderson et al., 1992), although there is some evidence for mid-Holocene coastal refugia (Schauffler and Jacobson, 2002). *Picea* in the Caribou record prior to 2 ka likely derives mainly from local bog populations, while the general rise after 2 ka represents regional re-expansion, although contribution from expansion on the bog surface in the last 2 ka cannot be ruled out. *Picea*'s late Holocene return has been attributed to temperature, particularly cooling summer temperatures (Webb, 1986). The trend towards wetter conditions may have also helped facilitate the return of *Picea*, but the lack of directional response to the major dry event at 500 yr BP suggests that temperature may have been more important to *Picea* than hydroclimate.

Many species exhibit synchronous changes at 500 yr BP, including an increase in *Pinus* and major decreases in *Tsuga* and *Fagus*. Similar dry events have been documented at three other bogs in Maine, and caused consistent ecological response (increased *Pinus* and *Quercus*, decreased *Tsuga* and *Fagus*) across New England (Clifford and Booth, 2015).

#### C.6. Conclusions: What's gained by having all of these records together?

These new records provide a detailed climate and vegetation history for south-central Maine, and add detail and texture to our understanding of the last 7,500 years of environmental change in northern New England. Lake-level and bog water-table depth records have been applied independently to understand climate drivers of the vegetation dynamics that underlie pollen data, but this study represents the first dual application of these two powerful and complementary paleoclimate proxies. Taken at face value, the lake-level and the bog water-table depth records may seem to be saying very different things about the paleoenvironments, but when applied in the light of knowledge of proxy processes they can be combined to yield an integrated record of hydrologic variability and change. The combined record explains dynamics of the pollen and charcoal records more comprehensively than either paleoclimate record alone.

Developing and combining unique sets of proxy records illustrates the rich possibilities and many challenges of paleoenvironmental reconstruction. Integration of diverse sets of archives and proxies yields both improved paleoenvironmental inference and clearer under-

standing of proxy sensitivity and bias (Jackson, 2012). Detailed site level work provides insights into climate and contingencies that led to today's environmental state (Jackson et al., 2009) and is critical to underpin regional to global syntheses. This process does not necessarily result in simple unifying conclusions, but rather, generates paleoenvironmental histories that more closely match the complex, multivariate, and non-linear nature of ecological responses to environmental change.

#### C.7. Acknowledgments

This research was supported by the National Science Foundation Macrosystems Biology Program (DEB-1241851). Thanks to Stephanie Madsen and Jeff Donnelly for lab and field support. John Tipton, Neil Pederson, and Matthew Roby provided useful comments that improved the manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

## REFERENCES

- Almquist H, Dieffenbacher-Krall AC, Flanagan-Brown R and Sanger D (2001) The holocene record of lake levels of Mansell Pond, central Maine, USA. *The Holocene* 11(2): 189–201.
- Almquist-Jacobson H and Sanger D (1995) Holocene climate and vegetation in the Milford drainage basin, Maine, USA, and their implications for human history. *Vegetation History and Archaeobotany* 4(4): 211–222.
- Anderson RS, Jacobson Jr GL, Davis RB and Stuckenrath R (1992) Gould Pond, Maine: late-glacial transitions from marine to upland environments. *Boreas* 21(4): 359–371.
- Bennett K and Fuller J (2002) Determining the age of the mid-holocene *Tsuga canadensis* (hemlock) decline, eastern North America. *The Holocene* 12(4): 421–429.
- Berger A and Loutre MF (1991) Insolation values for the climate of the last 10 million years. *Quaternary Science Reviews* 10(4): 297–317.
- Booth RK (2002) Testate amoebae as paleoindicators of surface-moisture changes on Michigan peatlands: modern ecology and hydrological calibration. *Journal of Paleolimnology* 28(3): 329–348.
- Booth RK (2010) Testing the climate sensitivity of peat-based paleoclimate reconstructions in mid-continental North America. *Quaternary Science Reviews* 29(5-6): 720–731.
- Booth RK, Brewer S, Blaauw M, Minckley TA and Jackson ST (2012a) Decomposing the mid-holocene *Tsuga* decline in eastern North America. *Ecology* 93(8): 1841–1852.
- Booth RK and Jackson ST (2003) A high-resolution record of late-Holocene moisture variability from a Michigan raised bog, usa. *The Holocene* 13(6): 863–876.
- Booth RK, Jackson ST and Gray CE (2004) Paleoecology and high-resolution paleohydrology of a kettle peatland in upper michigan. *Quaternary Research* 61(01): 113. DOI: 10.1016/j.yqres.2003.07.013.
- Booth RK, Jackson ST, Sousa VA, Sullivan ME, Minckley TA and Clifford MJ (2012b) Multi-decadal drought and amplified moisture variability drove rapid forest community change in a humid region. *Ecology* 93(2): 219–226.
- Cain SA (1939) Pollen analysis as a paleo-ecological research method. *The Botanical Review* 5(12): 627–654.

- Charman DJ (2007) Summer water deficit variability controls on peatland water-table changes: implications for Holocene palaeoclimate reconstructions. *The Holocene* 17(2): 217–227.
- Charman DJ, Barber KE, Blaauw M, Langdon PG, Mauquoy D, Daley TJ, Hughes PD and Karofeld E (2009) Climate drivers for peatland palaeoclimate records. *Quaternary Science Reviews* 28(19-20): 1811–1819.
- Charman DJ, Blundell A, Chiverrell RC, Hendon D and Langdon PG (2006) Compilation of non-annually resolved holocene proxy climate records: stacked Holocene peatland palaeo-water table reconstructions from northern Britain. *Quaternary Science Reviews* 25(3-4): 336–350.
- Charman DJ, Hendon D, Woodland WA, Association QR et al. (2000) *The identification of testate amoebae (Protozoa: Rhizopoda) in peats*. Quaternary Research Association.
- Clifford MJ and Booth RK (2013) Increased probability of fire during late Holocene droughts in northern New England. *Climatic change* 119(3-4): 693–704.
- Clifford MJ and Booth RK (2015) Late-Holocene drought and fire drove a widespread change in forest community composition in eastern North America. *The Holocene* 25(7): 1102–1110.
- COHMAP Members (1988) Climatic changes of the last 18,000 years: observations and model simulations. *Science* : 1043–1052.
- Comas X, Slater L and Reeve A (2005) Stratigraphic controls on pool formation in a domed bog inferred from ground penetrating radar (GPR). *Journal of Hydrology* 315(1-4): 40–51.
- Davis M, Woods K, Webb S and Futyma R (1986) Dispersal versus climate: Expansion of *Fagus* and *Tsuga* into the Upper Great Lakes region. *Vegetatio* 67(2): 93–103.
- Davis MB (1978) Climatic interpretation of pollen in Quaternary sediments. In: *Biology and Quaternary Environments*. Australian Academy of Science, pp. 35–51.
- Davis MB (1981) Outbreaks of forest pathogens in Quaternary history. *Proceedings of the Fourth International Palynological Conference* .
- Davis MB and Brubaker LB (1973) Differential sedimentation of pollen grains in lakes. *Limnology and Oceanography* 18(4): 635–646.
- Davis MB, Brubaker LB and Beiswenger JM (1971) Pollen grains in lake sediments: Pollen percentages in surface sediments from southern Michigan. *Quaternary Research* 1(4): 450–467.

- Davis MB, Spear RW and Shane LC (1980) Holocene climate of new england. *Quaternary Research* 14(2): 240–250.
- Davis MB, Sugita S, Calcote RR and Frelich L (1992) Effects of invasion by *Tsuga canadensis* on a North American forest ecosystem. In: *Responses of forest ecosystems to environmental changes*. Springer, pp. 34–44.
- Deevey ES (1939) Studies on connecticut lake sediments; part 1, a postglacial climatic chronology for southern New England. *American Journal of Science* 237(10): 691–724.
- Digerfeldt G (1974) The Post-Glacial development of the Ranviken bay in Lake Immeln; I. The history of the regional vegetation, and II. the water-level changes. *Geologiska Föreningen i Stockholm Förhandlingar* 96(1): 3–32.
- Digerfeldt G (1986) Studies on past lake-level fluctuations. *Handbook of Holocene palaeoecology and palaeohydrology* 127: 143.
- Edwards KJ, Fyfe RM and Jackson ST (2017) The first 100 years of pollen analysis. *Nature Plants* 3: 17001.
- Foster DR, Oswald WW, Faison EK, Doughty ED and Hansen BC (2006) A climatic driver for abrupt mid-Holocene vegetation dynamics and the hemlock decline in New England. *Ecology* 87(12): 2959–2966.
- Fritz S, Juggins S, Battarbee R and Engstrom D (1991) Reconstruction of past changes in salinity and climate using a diatom-based transfer function. *Nature* 352(6337): 706.
- Gajewski K (1987) Environmental history of Caribou Bog, Penobscot Co., Maine. *Naturaliste Canadien* 114: 133–140.
- Guiot J, Harrison SP and Prentice IC (1993) Reconstruction of Holocene precipitation patterns in Europe using pollen and lake-level data. *Quaternary Research* 40(2): 139–149.
- Harrison SP (1989) Lake levels and climatic change in eastern North America. *Climate Dynamics* 3(3): 157–167.
- Hu FS and Davis RB (1995) Postglacial development of a Maine bog and paleoenvironmental implications. *Canadian Journal of Botany* 73(4): 638–649.
- Huybers K, Rupper S and Roe GH (2016) Response of closed basin lakes to interannual climate variability. *Climate dynamics* 46(11-12): 3709–3723.
- Imbrie J and Kipp N (1971) A new micropaleontological method for quantitative paleoclimatology: Application to a late Pleistocene Caribbean core. *The Late Cenozoic Glacial Ages*. Yale Univ. Press, New Haven 3: 71–181.



- Innes J, Blackford J and Simmons I (2004) Testing the integrity of fine spatial resolution palaeoecological records: microcharcoal data from near-duplicate peat profiles from the North York Moors, UK. *Palaeogeography, Palaeoclimatology, Palaeoecology* 214(4): 295–307.
- Jackson ST (2012) Representation of flora and vegetation in Quaternary fossil assemblages: known and unknown knowns and unknowns. *Quaternary Science Reviews* 49: 1–15.
- Jackson ST, Betancourt JL, Booth RK and Gray ST (2009) Ecology and the ratchet of events: climate variability, niche dimensions, and species distributions. *Proceedings of the National Academy of Sciences* : pnas-0901644106.
- Jackson ST and Booth RK (2002) The role of Late Holocene climate variability in the expansion of yellow birch in the western Great Lakes region. *Diversity and Distributions* 8(5): 275–284.
- Jackson ST, Booth RK, Reeves K, Andersen JJ, Minckley TA and Jones RA (2014) Inferring local to regional changes in forest composition from holocene macrofossils and pollen of a small lake in central Upper Michigan. *Quaternary Science Reviews* 98: 60–73.
- Jackson ST and Whitehead DR (1991) Holocene vegetation patterns in the Adirondack Mountains. *Ecology* 72(2): 641–653.
- Laird KR, Fritz SC and Cumming BF (1998) A diatom-based reconstruction of drought intensity, duration, and frequency from Moon Lake, North Dakota: a sub-decadal record of the last 2300 years. *Journal of Paleolimnology* 19(2): 161–179.
- Levesque AJ, Mayle FE, Walker IR and Cwynar LC (1993) A previously unrecognized late-glacial cold event in eastern North America. *Nature* 361(6413): 623.
- Marsicek JP, Shuman B, Brewer S, Foster DR and Oswald WW (2013) Moisture and temperature changes associated with the mid-Holocene Tsuga decline in the northeastern United States. *Quaternary Science Reviews* 80: 129–142.
- Minckley TA, Booth RK and Jackson ST (2012) Response of arboreal pollen abundance to late-holocene drought events in the Upper Midwest, USA. *The Holocene* 22(5): 531–539.
- Morris PJ, Baird AJ, Young DM and Swindles GT (2015) Untangling climate signals from autogenic changes in long-term peatland development. *Geophysical Research Letters* 42(24): 10–788.
- Newby PE, Donnelly JP, Shuman BN and MacDonald D (2009) Evidence of centennial-scale drought from southeastern Massachusetts during the Pleistocene/Holocene transition. *Quaternary Science Reviews* 28(17-18): 1675–1692.

- Newby PE, Shuman BN, Donnelly JP, Karnauskas KB and Marsicek J (2014) Centennial-to-millennial hydrologic trends and variability along the North Atlantic Coast, USA, during the Holocene. *Geophysical Research Letters* 41(12): 4300–4307.
- Newby PE, Shuman BN, Donnelly JP and MacDonald D (2011) Repeated century-scale droughts over the past 13,000 yr near the Hudson River watershed, USA. *Quaternary Research* 75(3): 523–530.
- Nolan C, Tipton J, Booth RK, Hooten MB and Jackson ST (2019) Comparing and improving methods for reconstructing peatland water-table depth from testate amoebae. *The Holocene* : accepted.
- Parnell A (2014) Bchron: Radiocarbon dating, age-depth modelling, relative sea level rate estimation, and non-parametric phase modelling. *R package version* 4(1).
- Pribyl P and Shuman BN (2014) A computational approach to Quaternary lake-level reconstruction applied in the central Rocky Mountains, Wyoming, USA. *Quaternary Research* 82(1): 249–259.
- Ramiadantsoa T, Stegner MA, Williams JW and Ives AR (2019) The potential role of intrinsic processes in generating abrupt and quasi-synchronous tree declines during the Holocene. *Ecology* 100(2): e02579.
- Ritchie J (1986) Climate change and vegetation response. *Vegetatio* 67(2): 65–74.
- Roos-Barraclough F, Givélet N, Cheburkin AK, Shotyk W and Norton S (2006) Use of Br and Se in peat to reconstruct the natural and anthropogenic fluxes of atmospheric hg: A 10000-year record from Caribou Bog, Maine. *Environmental science & technology* 40(10): 3188–3194.
- Sachs JP (2007) Cooling of Northwest Atlantic slope waters during the holocene. *Geophysical Research Letters* 34(3).
- Schauffler M and Jacobson GL (2002) Persistence of coastal spruce refugia during the Holocene in northern New England, USA, detected by stand-scale pollen stratigraphies. *Journal of Ecology* 90(2): 235–250.
- Seager R, Pederson N, Kushnir Y, Nakamura J and Jurburg S (2012) The 1960s drought and the subsequent shift to a wetter climate in the Catskill Mountains region of the New York City watershed. *Journal of Climate* 25(19): 6721–6742.
- Shuman B, Bravo J, Kaye J, Lynch JA, Newby P and Webb T (2001) Late Quaternary water-level variations and vegetation history at Crooked Pond, southeastern Massachusetts. *Quaternary Research* 56(3): 401–410.

- Shuman B, Newby P, Huang Y and Webb III T (2004) Evidence for the close climatic control of New England vegetation history. *Ecology* 85(5): 1297-1310.
- Shuman B, Oswald WW and Foster DR (2019a) Multivariate climate change, the climate niche, and the Holocene history of eastern hemlock (*Tsuga canadensis*). *Ecological Monographs* .
- Shuman B and Plank C (2011) Orbital, ice sheet, and possible solar controls on Holocene moisture trends in the North Atlantic drainage basin. *Geology* 39(2): 151–154.
- Shuman BN and Burrell SA (2017) Centennial to millennial hydroclimatic fluctuations in the humid northeast United States during the Holocene. *Quaternary Research* 88(3): 514–524.
- Shuman BN, Marsicek J, Oswald WW and Foster DR (2019b) Predictable hydrological and ecological responses to holocene north atlantic variability. *Proceedings of the National Academy of Sciences* 116(13): 5985–5990.
- Shuman BN, Newby P and Donnelly JP (2009) Abrupt climate change as an important agent of ecological change in the Northeast US throughout the past 15,000 years. *Quaternary Science Reviews* 28(17-18): 1693–1709.
- Shuman BN, Routson C, McKay N, Fritz S, Kaufman D, Kirby ME, Nolan C, Pederson GT and St-Jacques JM (2018) Placing the Common Era in a Holocene context: Millennial-to-centennial patterns and trends in the hydroclimate of North America over the past 2000 years. *Clim. Past* .
- Singer DK, Jackson ST, Madsen BJ and Wilcox DA (1996) Differentiating climatic and successional influences on long-term development of a marsh. *Ecology* 77(6): 1765–1778.
- Swindles GT, Morris PJ, Baird AJ, Blaauw M and Plunkett G (2012) Ecohydrological feedbacks confound peat-based climate reconstructions. *Geophysical Research Letters* 39(11).
- ter Braak CJ and van Dame H (1989) Inferring pH from diatoms: a comparison of old and new calibration methods. *Hydrobiologia* 178(3): 209–223.
- Tinner W and Hu FS (2003) Size parameters, size-class distribution and area-number relationship of microscopic charcoal: relevance for fire reconstruction. *The Holocene* 13(4): 499–505.
- von Post L (1918) Forest tree pollen in south Swedish peat bog deposits, translated by Margaret B. Davis and Knut Faegri, 1967. *Pollen et Spores* 9: 375–401.

- Waddington J, Morris P, Kettridge N, Granath G, Thompson D and Moore P (2015) Hydrological feedbacks in northern peatlands. *Ecohydrology* 8(1): 113–127.
- Wang Y, Gill JL, Marsicek J, Dierking A, Shuman B and Williams JW (2016) Pronounced variations in *Fagus grandifolia* abundances in the Great Lakes region during the Holocene. *The Holocene* 26(4): 578–591.
- Webb T (1986) Is vegetation in equilibrium with climate? how to interpret late-Quaternary pollen data. *Vegetatio* 67(2): 75–91.
- Webb T and Bryson RA (1972) Late-and postglacial climatic change in the northern Midwest, USA: quantitative estimates derived from fossil pollen spectra by multivariate statistical analysis. *Quaternary Research* 2(1): 70–115.
- Webb III T, Shuman B and Williams JW (2003) Climatically forced vegetation dynamics in eastern North America during the late Quaternary period. *Developments in Quaternary Sciences* 1: 459–478.
- Weng C and Jackson ST (1999) Late Glacial and Holocene vegetation history and paleoclimate of the Kaibab Plateau, Arizona. *Palaeogeography, Palaeoclimatology, Palaeoecology* 153(1-4): 179–201.
- Wilmshurst JM, McGlone MS and Charman DJ (2002) Holocene vegetation and climate change in southern New Zealand: linkages between forest composition and quantitative surface moisture reconstructions from an ombrogenous bog. *Journal of Quaternary Science: Published for the Quaternary Research Association* 17(7): 653–666.
- Winkler MG, Swain AM and Kutzbach JE (1986) Middle Holocene dry period in the northern midwestern United States: lake levels and pollen stratigraphy. *Quaternary Research* 25(2): 235–250.
- Wright HE, Mann D and Glaser P (1984) Piston corers for peat and lake sediments. *Ecology* 65(2): 657–659.

D. APPENDIX D

SUPPLEMENTAL INFORMATION FOR USING CO-LOCATED LAKE-LEVEL AND  
BOG WATER-TABLE DEPTH RECORDS TO UNDERSTAND HOLOCENE CLIMATE  
AND VEGETATION CHANGES IN MAINE

Sample ID	Core	Depth (cm)	Material dated	14C Age	14C Error	notes
122625	31A	32.5	Bulk sediment	755	20	
122626	31A	57.5	Bulk sediment	2660	20	
122627	31A	71.5	Bulk sediment	3480	20	
122628	31A	88.5	Bulk sediment	3940	30	
122629	31A	122.5	Bulk sediment	9230	20	
122630	31A	149.5	Bulk sediment	9830	20	
122631	31A	160.5	Bulk sediment	11000	20	
122632	31A	179.5	Bulk sediment	12350	30	
137468	39A	36.5	Bulk sediment	1090	15	
137469	39A	68.5	Bulk sediment	2270	20	
137470	39A	119.5	Bulk sediment	3860	20	
137472	39A	165.5	Bulk sediment	6710	25	
137471	39A	174.5	Bulk sediment	7180	25	
137475	39A	199	Bulk sediment	8950	35	
137473	39A	244	Bulk sediment	10800	45	
137474	39A	253	Bulk sediment	10950	40	
165955	50A	54.5	Charcoal pieces	2005	35	
137476	50A	74.5	Bulk sediment	2920	20	
137477	50A	89.5	Charcoal pieces	3250	25	
165956	50A	100.5	Bulk sediment	1740	25	
165957	50A	118.5	Charcoal pieces	3640	25	
165958	50A	178.5	Bulk sediment	5135	25	
137478	50A	218.5	Charcoal pieces	7060	30	
137479	50A	241	Bulk sediment	7440	25	
137480	50A	283	Bulk sediment	8090	35	

TABLE D.1. Radiocarbon dates for Giles Pond.

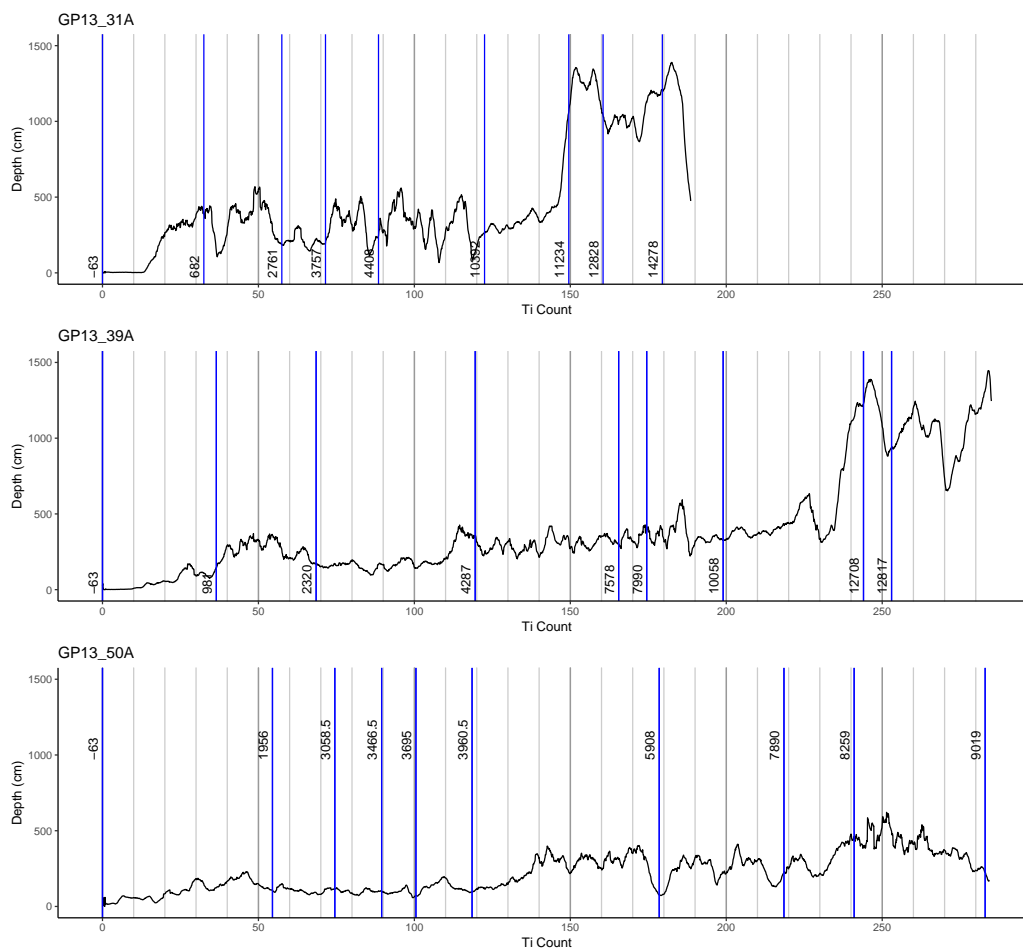


FIGURE D.1. Ti counts from XRF versus depth for the three Giles Pond cores used in lake-level reconstruction annotated with the median calibrated radiocarbon dates.

E. APPENDIX E

COHERENT MID-TO-LATE HOLOCENE HYDROLOGIC VARIABILITY ACROSS  
TEMPORAL SCALES IN MID-LATITUDE EASTERN NORTH AMERICA

Manuscript prepared for submission to *Earth and Planetary Science Letters*



**Coherent mid-to-late Holocene hydrologic variability across temporal scales in  
mid-latitude eastern North America**

Connor Nolan<sup>1</sup>, Bryan N. Shuman<sup>2</sup>, Robert K. Booth<sup>3</sup>, and Stephen T. Jackson<sup>4,1</sup>

<sup>1</sup> Department of Geosciences, University of Arizona, Tucson, AZ 85721, USA

<sup>2</sup> Department of Geology and Geophysics, University of Wyoming, Laramie, WY 82071,  
USA

<sup>3</sup> Department of Earth and Environmental Sciences, Lehigh University, Bethlehem, PA  
18015, USA

<sup>4</sup> U.S. Geological Survey, Southwest Climate Adaptation Science Center, Tucson, AZ  
85721, USA

E.1. Abstract

Paleoclimate records provide long-term historical context for climate variability from the Holocene and beyond. Lake-level and bog water-table depth records are two key proxies for terrestrial hydroclimate in North America. These proxies have generally not been developed in the same places, but for both proxies there is a developing network of sites in mid-latitude North America. Here we present a new pair of co-located records from northeastern Minnesota: a 6,200-year long lake-level reconstruction from Knuckey Lake and a 1000-year long testate amoeba-based water-table depth reconstruction spanning 1500 to

500 cal yr BP from Ely Lake Bog. These two proxies have different hydroclimatic sensitivities and time scales of response, but together they provide a new record of local Holocene hydroclimate and represent new datapoints in their respective single-proxy networks. The new records provide details about known Minnesota climatic and ecological changes over the Holocene, such as the warm-dry mid-Holocene. When compared to bog water-table depth records from Minnesota and Michigan, the Ely Lake Bog record supports inference of synchronous Upper Midwest dry events in the last millennium. The network of eastern US bog records is generally consistent at the sub-regional scale, but lacks clear coherence between records from Maine and records from the Upper Midwest. The new lake-level record from Knuckey Lake connects a network of lake-level reconstructions from Wyoming to Maine generated using consistent methodologies. We analyze this network and find coherent multi-centennial hydroclimate variability, especially among sites across eastern North America. Altogether, we illustrate the utility of single-proxy networks in explaining the complex spatial and temporal variability of hydroclimate over the Holocene in mid-latitude North America.

## E.2. Introduction

Records of past hydroclimate extend the short observational record and place modern observations in a long-term context (Masson-Delmotte et al., 2013). Tree ring records in particular have provided annually-resolved centennial to millennial perspective on past hy-

droclimate (Woodhouse and Overpeck, 1998; Meko et al., 2007). Tree ring-based drought atlases have provided continental-scale records of spatial patterning of droughts over the last millennium (Cook, 2004; Cook et al., 2010, 2015). Analyses of these networks in conjunction with climate modeling has led to insights about the dynamical forcing of past droughts (Coats et al., 2013, 2016; Stevenson et al., 2015).

However, tree-ring drought atlases extend no more than 2000 years, and in some regions where old-growth forests have disappeared (e.g. the northeast US), the drought atlases are only trustworthy for the past 200-300 years. Attempts to understand regional- to continental-scale hydroclimate variability on longer time scales have lagged behind tree-ring efforts and temperature syntheses (PAGES, 2013; Marcott et al., 2013; Shakun et al., 2012). Initial syntheses of non-tree ring hydroclimate proxies have shown limited coherence (Shuman et al., 2018; Rodysill et al., 2018; Marlon et al., 2017). In a recent synthesis from North America, nearby proxy records did not agree on trends or variability over the last 2000 years and loaded differently onto principal component axes (Shuman et al., 2018).

Hydroclimate patterns are complex in space and time, and even modern precipitation and drought observations and projections have significant disagreements and uncertainties (Hartmann et al., 2013; Stocker et al., 2013). Thus, it is not surprising that using proxies to understand hydroclimate hundreds to thousands of years in the past presents significant challenges (PAGES, 2017). Hydroclimate has a shorter decorrelation distance and more complex spatial pattern than temperature; and such, continental and global averages are less meaningful for hydroclimate than for temperature (Büntgen et al., 2010; Gomez-Navarro

et al., 2015). Furthermore, hydroclimate proxies have a wide range of sensitivities and record different aspects of hydroclimate at different temporal scales.

Some useful approaches for integrating hydroclimatic records include: treating proxies qualitatively (DiNezio and Tierney, 2013; Thirumalai et al., 2018), using co-located proxy records (Nolan et al., 2019), and focusing on proxy-specific networks (Booth et al., 2006; Newby et al., 2014). Specifically, analysis of groups of lake-level sites in the northeastern US and the central Rocky Mountains have found coherent multicentury variability (Shuman and Burrell, 2017; Shuman and Serravezza, 2017). Multicentury variability in the Holocene is well-documented from many sites, but the patterns, extent, and mechanisms are not well-constrained (Mayewski et al., 2004; Wanner et al., 2011; Fawcett et al., 2011; Shuman, 2012; Ault et al., 2013).

In this paper, we present new co-located lake-level and bog water-table depth records from north-eastern Minnesota. The new records build on a long history of paleoclimate and paleoecological research in the region. Pollen records from Minnesota show a general sequence of spruce-dominated forests in the late Pleistocene transitioning to pine-dominated forests in the early Holocene (Wright Jr. et al., 1963; Webb et al., 1983). Pine dominance was interrupted by warm and dry conditions in the mid-Holocene that shifted the prairie-forest boundary more than 100 km to the east (Webb et al., 1983). The warm and dry mid-Holocene has been identified in many proxies including pollen (Webb et al., 1983), lake levels (Digerfeldt et al., 1992), stable isotopes (Henderson et al., 2010), ostracodes (Smith et al., 2002), diatoms (Saros et al., 2000), geochemistry (Nelson et al., 2004), and

dust flux (Keen and Shane, 1990; Dean, 1993). Aridity generally set in around 8.0 ka, reached its maximum around 7.0 ka, and ended around 6.0 ka (Nelson and Hu, 2008). Pollen-based estimates suggest a 10-25 % decrease in precipitation and a mean July temperature increase of 0.5-2.0 degrees Celsius relative to pre-Industrial conditions (Bartlein et al., 1984). Lake-level hydrologic modeling suggests lake-level decreases of 3-6 meters below present and decreases in precipitation of nearly 30% (Almendinger, 1993).

After 6.0 ka, precipitation began to increase and mean July temperature decreased in northern Minnesota and increased in southern Minnesota. *Pinus* dominated the forests again and the cooling temperatures allowed the return of *Picea* to parts of the northern Upper Midwest by 3.0 ka (Webb et al., 1983; Bartlein et al., 1984). High-resolution records of the past 1000-2000 years such as (Brugam et al., 1988; Fritz et al., 2000; Jacques et al., 2008; Booth et al., 2006) document recent variability and identify coherent dry events around 1000 cal yr BP and 750 cal yr BP (Booth et al., 2006).

We analyze our new lake-level and bog water-table records in the context of regional paleoclimate and paleoecology. In particular, our new lake-level record is the first of its kind in this region. It bridges the gap between records of the last 2000 years and records of mid-Holocene aridity and provides insights into the multi-centennial to millennial scale variability over the past 6000 years in the Upper Midwest. Then, we connect the new lake-level and bog water-table depth records to existing networks of lake-level and bog water-table depth records from across the eastern United States to analyze spatial patterns of decadal to multi-centennial hydroclimate variability at regional to continental scales over

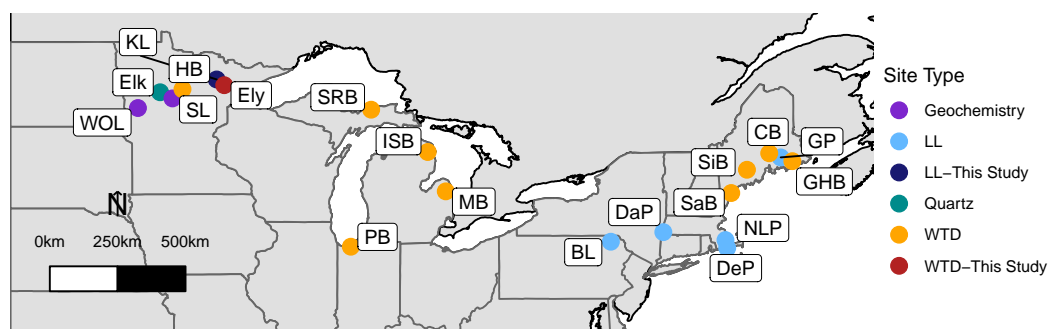


FIGURE E.1. Mid-latitude North America sites developed and used in this study. KL = Knuckey Lake, HB = Hole Bog, SL = Steel Lake, WOL = West Olaf Lake, SRB = South Rhody Bog, ISB = Irwin Smith Bog, MB = Minden Bog, PB = Pinhook Bog, CB = Caribou Bog, GP = Giles Pond, SiB = Sidney Bog, SaB = Saco Bog, GHB = Great Heath Bog, NLP = New Long Pond, DaP = Davis Pond, DeP = Deep Pond, BL = Blanding Lake.

the past 6000 years. Lastly, we compare the patterns and variability from the lake-level network to a long climate model simulation to understand how the proxy reconstructions compare to simulated mid-to-late Holocene hydrologic variability.

### E.3. Sites and Methods

#### E.3.1. Knuckey Lake

Knuckey Lake (47.651625, -92.766152; 438 m elevation) is a small kettle pond (28 ha, max depth 4 meters) with two perennial, small stream outlets. The lake is located northwest of Virginia, MN on the outwash of the St. Croix moraine association. We performed ground penetrating radar surveys and found reflective sand layers consistent with past lake-level change. Using a piston corer, we collected three near-shore cores from the north-northwest

side of the lake: core KL15-16A is 129 cm long and from a water depth of 68cm, core KL15-19A is 179 cm long and from a water depth of 76 cm, and KL15-22A is 235 cm long and from a water depth of 91 cm.

The cores were returned to the lab where they were described and scanned on a GeoTek core scanner to obtain photographs and gamma density. Loss-on-ignition (LOI) analysis was performed on every centimeter of sediment. Low LOI values are associated with increased sand content; high LOI is associated with increased organic content. The LOI data corresponded well with sand and organic-rich gyttja layers identified in visual descriptions of the cores. We obtained 15 radiocarbon dates (Table F.1), most positioned immediately above and below sand layers to date drawdown periods. Age-depth models were fit using Bchron (Parnell, 2014).

The core accumulation histories and LOI data were used in a decision-tree lake-level reconstruction algorithm (Pribyl and Shuman, 2014) to generate lake-level trajectories consistent with the core data. We used cut-offs for littoral ranging from 5-20% LOI and cut-offs for sublittoral ranging from 20-35% LOI (Shuman, 2003).

### E.3.2. Ely Lake Bog

Ely Lake Bog (47.445, -92.439; 420 m elevation; approx. 500 ha) is an ombrotrophic peatland 30 km southeast of Knuckey Lake that formed on glacial lake sediment. Sediment cores were collected in the summer of 2015. We collected the top 52 cm as a monolith and the next 152cm in two drives using a modified piston corer (Wright et al., 1984). The cores

were extruded and described in the field.

In the lab, the cores were sliced into one-centimeter thick sections. From each section, a subsample ( $1\text{cm}^3$ ) was prepared for testate amoebae analysis following standard methods (Booth, 2010). Testate amoebae were identified based on a key from Charman et al. (2000) and counted to a minimum of 100 individuals per sample. To translate testate amoeba community composition from the Ely Lake Bog record to an estimate of water-table depth, we used weighted averaging trained on a large modern surface-sample dataset from North America. We used this simple transfer function because it performs well for single-site reconstructions and makes our WTD record directly comparable to existing records from the region (Booth et al., 2006).

Microcharcoal particles (15-250 micrometers) were counted on the same slides as the testate amoeba. Microcharcoal is a proxy for regional fire activity (Innes et al., 2004; Tinner and Hu, 2003; Clifford and Booth, 2013). The microcharcoal counts were converted to charcoal influx based on the accumulation rates derived from the age-depth model and *Lycopodium* spores added to the testate amoebae preparation at a known concentration.

We obtained 16 radiocarbon dates (Table F.2) and fit an age-depth model using Bchron (Parnell, 2014). There is a hiatus from 140 and 450 yr BP (note: all dates are expressed as cal yr BP, where present is defined as 1950 CE), the timing of which is similar to a hiatus at Hole Bog (Booth et al., 2006; Booth, 2010), another raised bog in Minnesota about 200 km west of Ely Lake Bog. We focus on the interval from 1500-500 yr BP to capture regional drought events that occurred around 1000 and 750 yr BP (Booth et al., 2006; Shuman et al.,



2009).

### E.3.3. TraCE-21k climate model analyses

TraCE-21k (Simulation of the Transient Climate of the Last 21,000 Years) is a 22,000 year transient simulation using the Community Climate System Model version 3 (CCSM3) (He, 2011; Liu et al., 2009). This is the only publicly-available, fully-coupled transient climate model simulation of the last deglaciation and Holocene. We used the decadal-mean annual total precipitation (PRECT) from the National Center for Atmospheric Research Climate Data Gateway. We extracted time-series for the four gridpoints that contain the six lake-level records from the eastern US. We smoothed the decadal-averaged time series with a smooth spline to emphasize variability on 200+ year timescales. We performed a principal component analysis on the decadal averages from 10,000 yr BP to present and smoothed the resulting PCs with a smooth spline.

## E.4. Results

### E.4.1. New Records

*Knuckey Lake lake-level record.* The Knuckey Lake cores show a coherent pattern of loss-on-ignition (Figure E.2A) that is consistent with sand layers identified visually in the cores. Basal dates for core 22A, 19A, and 16A are 6200 BP, 5300 BP, and 4400 BP, respectively.

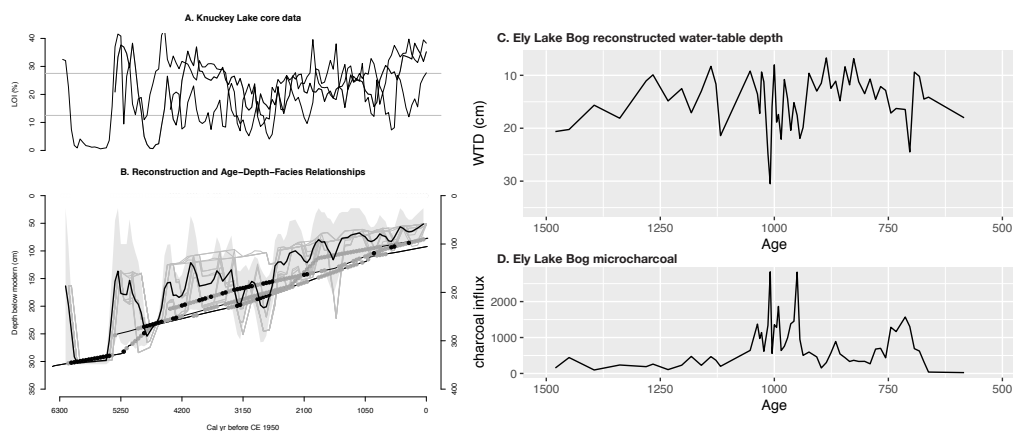


FIGURE E.2. A. Loss-on-ignition data from Knuckey Lake core 16A (thinnest line), core 19A (gray line), and core 22A (black line). B. Reconstructed lake level from Knuckey Lake (black line) with uncertainty (gray shading). The core data and accumulation history are shown as lines. Black dots represent an interval classified as littoral, gray dots represent an interval classified as sublittoral. C. Ely Lake Bog testate amoeba-based water-table depth reconstruction. D. Microcharcoal record from Ely Lake Bog.

Using the core accumulation history and LOI data, the decision-tree lake-level reconstruction algorithm (Pribyl and Shuman, 2014) produces a 6,200-year lake-level history. An overall trend towards wetter conditions over the past 6200 years was modulated by multi-centennial scale variability. At 6.2 ka, the lake first reached our cores, as shown by gyttja at the base of core 22A. A transition to sand then indicates a dry conditions from 6.0 to 5.5 ka. Gyttja at the base of core 19A indicates lake levels rose from 5.4 to 4.9 ka. 4.9 to 3.9 ka is characterized by a sequence of alternating dry and wet intervals. Clear sand layers return at 3.3 ka indicating drier conditions, followed by dry and variable conditions through 2.7 ka. Lake levels rose through 2.2 ka. After 2.2 ka, most of the cores contain mainly gyttja, except for a few short intervals of decreased LOI (more sand) including: 2100-2000, 1650-1550, 1050-950, 700-600, and 450-300 yr BP.

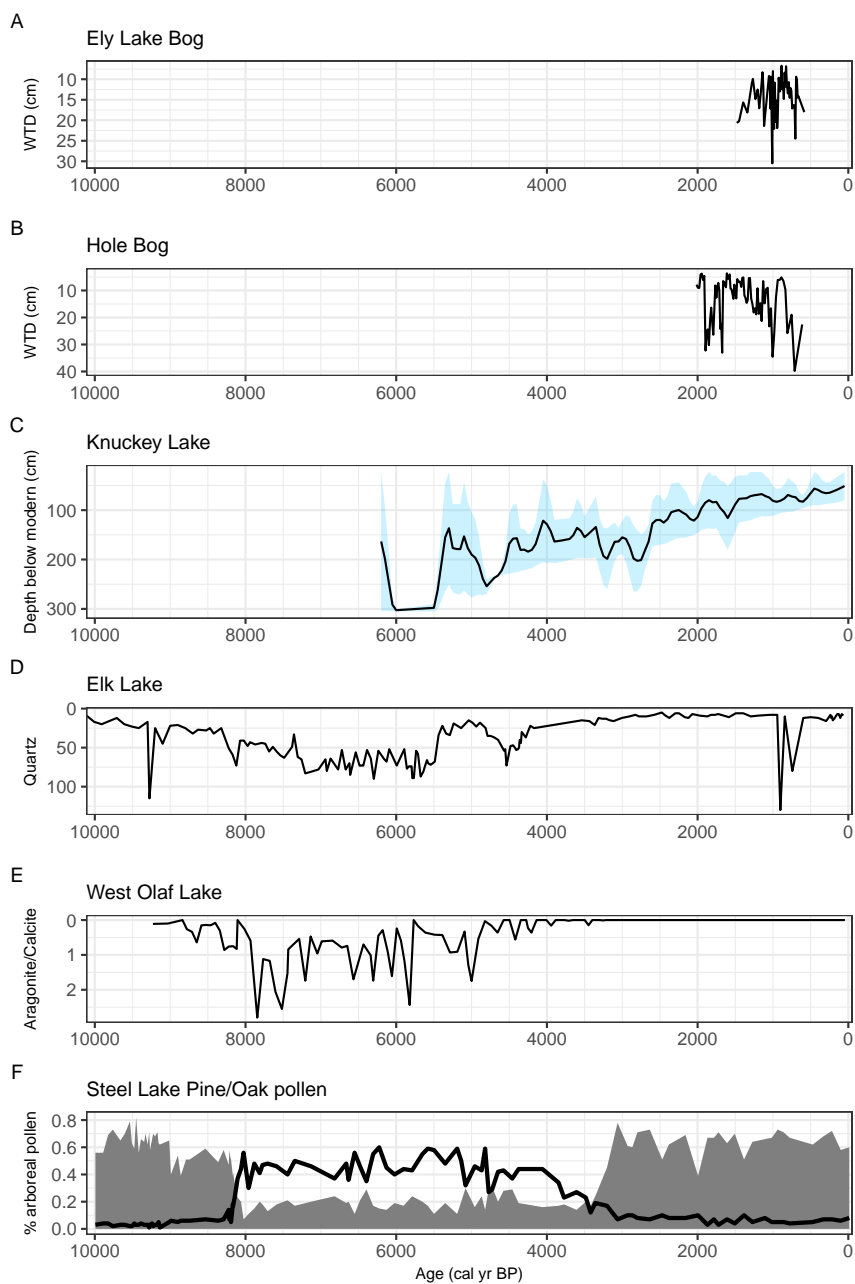


FIGURE E.3. Records of hydroclimate from Minnesota over the Holocene. A: Testate amoebae-inferred water-table depth from Ely Lake Bog. B: Testate amoebae-inferred water-table depth Hole Bog (Booth et al., 2006). C: Lake-level record from Knuckey Lake. D: Quartz record, a proxy for dustiness, from Elk Lake (Dean, 1993). E: Aragonite:Calcite ratio from West Olaf Lake (Nelson and Hu, 2008). F: Pine (shaded) and Oak (black line) pollen from Steel Lake (Nelson et al., 2004).

*Ely Lake Bog water-table depth and charcoal record* The 1000-year record from 1500-500 BP of testate amoeba-inferred water-table depth from Ely Lake Bog and the microcharcoal shows the characteristic decadal to multi-decadal variability that is common in bog water-table depth records (Figure E.2C-D). The deep, dry event centered at 1010 BP is associated with a 20-cm drop in water-table and a large peak in charcoal influx. This began a 75-year period of variable dry conditions, including another dry event with a large charcoal peak at 950 BP. Dry conditions and high charcoal also returned from 750-700 BP.

#### E.4.2. New records in Minnesota context

Figure E.3 shows the new records from Ely Lake Bog and Knuckey Lake in the context of the last 10,000 years. This context sharply contrasts the smooth centennial-to-millennial scale variability of the lake-level reconstruction compared to the decadal-scale variability of the bog water-table depth records (Figure E.3A-C).

The standard sequence of Holocene changes in Minnesota is represented in the Elk Lake Quartz record (Dean, 1993), West Olaf Lake aragonite-to-calcite ratio record (Nelson and Hu, 2008), and Steel Lake pine versus oak pollen record (Nelson et al., 2004)(Figure E.3D-F). All three of these records suggest increased aridity from 8 to 6 ka with some differences in the magnitude and variability between sites. After 6 ka, the three records show the same general trends towards wetter conditions, but the timing and magnitude differs. Differences could be related to spatial differences (e.g., West Olaf Lake is further west in Minnesota than Elk Lake and Steel Lake (Figure E.1) and thus the end of the mid-Holocene warm and

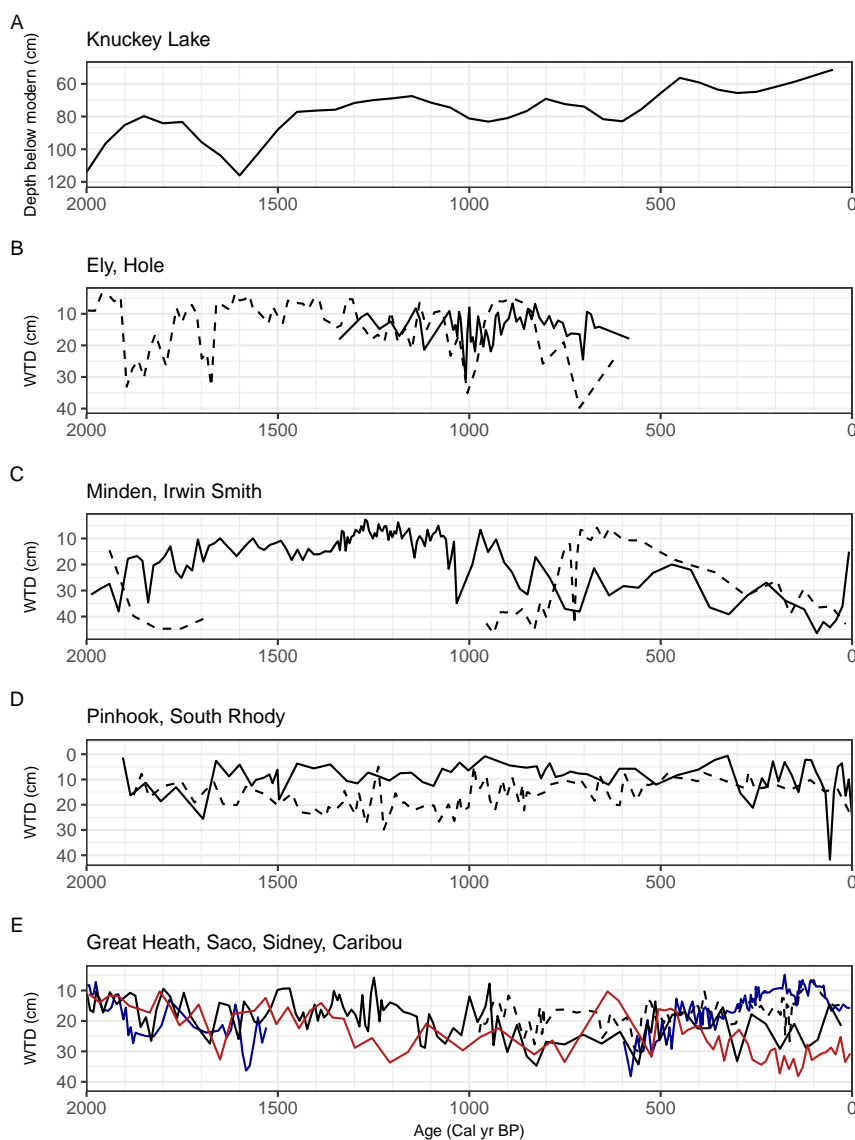


FIGURE E.4. Last 2000 years of hydrologic variability from Knuckey Lake and a network of bog water-table depths from the eastern United States. A: Lake-level reconstruction from Knuckey Lake. B: Water-table depth reconstructions from Ely Lake Bog (solid line; this study) and Hole Bog (dashed line; Booth et al. (2006)). C: Water-table depth reconstructions from Minden Bog (solid line; Booth et al. (2006)) and Irwin Smith Bog (dashed line; Booth et al. (2012)). D: Water-table depth reconstructions from Pinhook Bog (solid line; Booth et al. (2012)) and South Rhody Bog (dashed line; Booth et al. (2012)). E. Water-table depth reconstructions from Sidney Bog (solid line; Clifford and Booth (2013)), Great Heath Bog (dashed line; Clifford and Booth (2013)), Saco Bog (blue line; Clifford and Booth (2013)), and Caribou Bog (red line; Nolan, et al. 2019).

dry conditions may have occurred earlier). Differences may also be related to the nature of the proxies. For example, the complex relationship between changes in lake depth and changes in geochemistry (e.g., Donovan et al. (2002); Smith et al. (2002)) could complicate direct comparisons.

The timing of the beginning of the Knuckey Lake lake-level record and the multi-centennial variability from 5.5 to 4.5 ka is consistent with existing records, especially Elk Lake. Further, the timing of the decrease in oak and return of pine at Steel Lake, around 4.5 to 3 ka, is consistent with the trend towards wetter conditions at Knuckey Lake.

In the last 2000 years, major dry events are recorded at Ely Lake Bog and Hole Bog at 1000 yr BP and 750 yr BP (Figure E.4B). These events appear as minor (potentially insignificant relative to uncertainties) lake-level drawdowns at Knuckey Lake (Figure E.4A). Another dry period in Hole Bog at 1700 yr BP also appears as a drawdown in the Knuckey Lake reconstruction.

#### E.4.3. Bog water-table depth records from the eastern United States

The dry events at 1000 and 750 yr BP in Minnesota bogs and at Knuckey Lake also clearly appear in the record from Minden Bog in Michigan (Figure E.4A-C). Irwin Smith Bog, also in Michigan, does not record the event at 1000 yr BP due to a hiatus from 1700 to 955 yr BP, but does clearly record a drought event at 750 yr BP. Dry events from 1900 to 1650 yr BP at Hole Bog do not appear in the Michigan bogs.

Pinhook Bog (northern Indiana) and South Rhody Bog (Upper Peninsula, MI) do not

record dry events at 1000 and 750 yr BP (Figure E.4D). However, these bogs differ from Minden and Irwin Smith (also Ely and Hole) in two key ways: Pinhook and South Rhody are kettle peatlands and they experience significant lake-effect snow in winter, whereas Minden and Irwin Smith are raised, ombrotrophic peatlands and experience less lake-effect snow (Booth et al., 2012).

Bog water-table depth records from Maine document a dry event at 550 yr BP and possible dry events at 1850, 1650 to 1550, and 850 yr BP (Figure E.4E). The Upper Midwest dry events at 1000 and 750 yr BP do not appear to extend into Maine.

A principal component analyses (PCA) on the Upper Midwest subset and the Maine subset of bog records supports these events as significant shared variations Figure F.1. In the Upper Midwest, the first principal component (PC) of a PCA using Ely Lake Bog, Hole Bog, and Minden Bog over 620 to 1340 yr BP (their shared period of overlap) represents 53% of the variance and includes expressions of the dry events at 1000 and 750 yr BP.

In Maine, the first PC of a PCA using Saco Bog, Sidney Bog, Great Heath Bog, and Caribou Bog from 0 to 600 yr BP represents 56% of the variance and contains the shared dry event at 500 yr BP. And the first PC of a PCA using Saco, Sidney, and Caribou from 1550 to 2000 yr BP represents 55% of the shared variance and contains dry events at 1600 and 1800 yr BP.

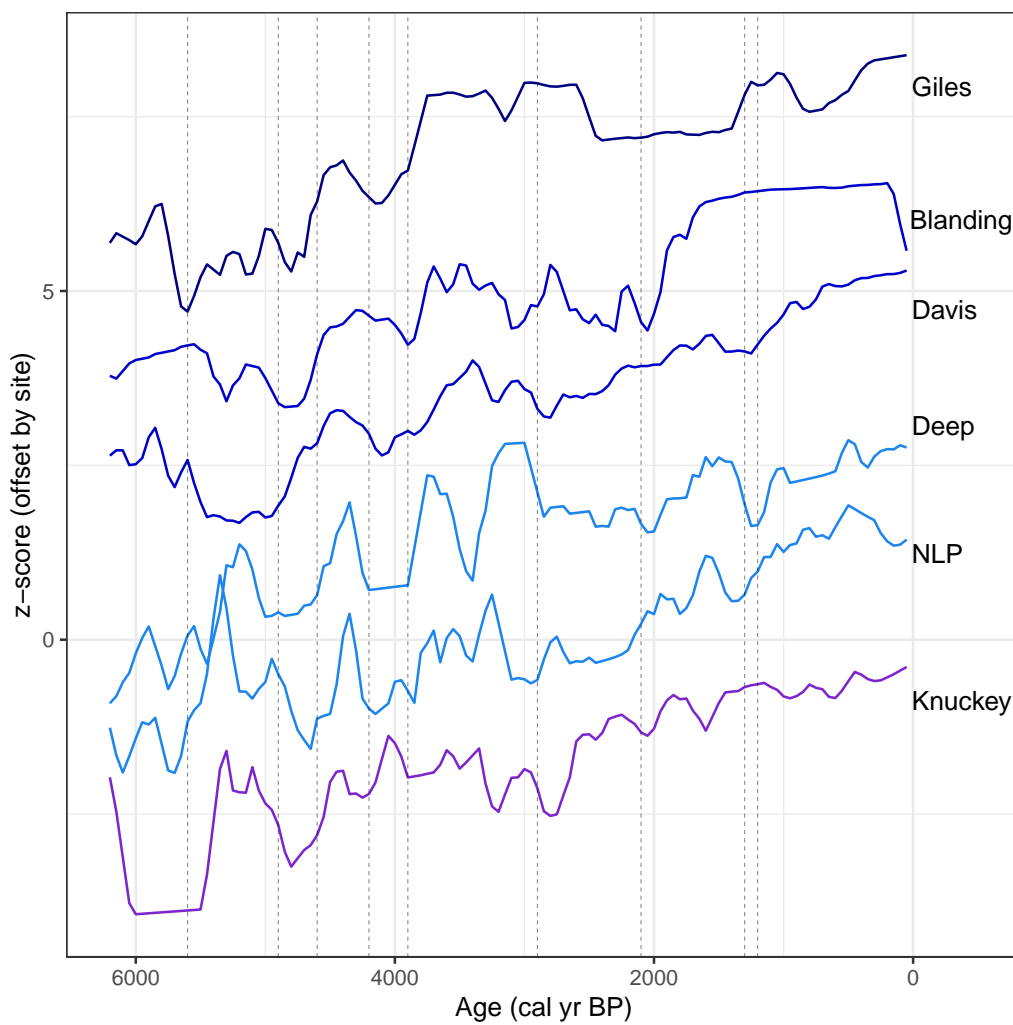


FIGURE E.5. Z-Scores of lake-level records from the eastern US over the past 6200 years. The vertical lines indicate the periods of coherent multi-centennial variability among Massachusetts lake-level records identified by Newby et al. (2014).



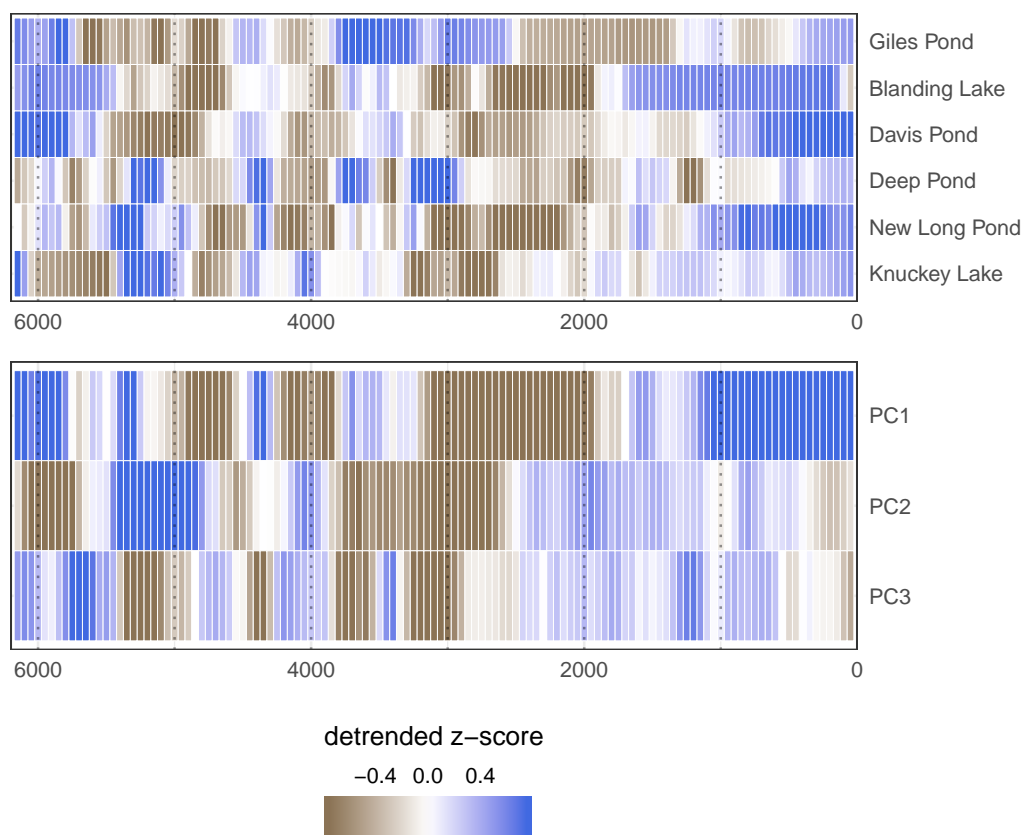


FIGURE E.6. Top: Mid-latitude North American lake-level records over the last 6200 years converted to z-scores and detrended. Bottom: Principal components of the detrended z-scores of the eastern US lake-level records. PC1 explains 43.5% of the variance, PC2 explains 21.1%, and PC3 explains 14.8%.

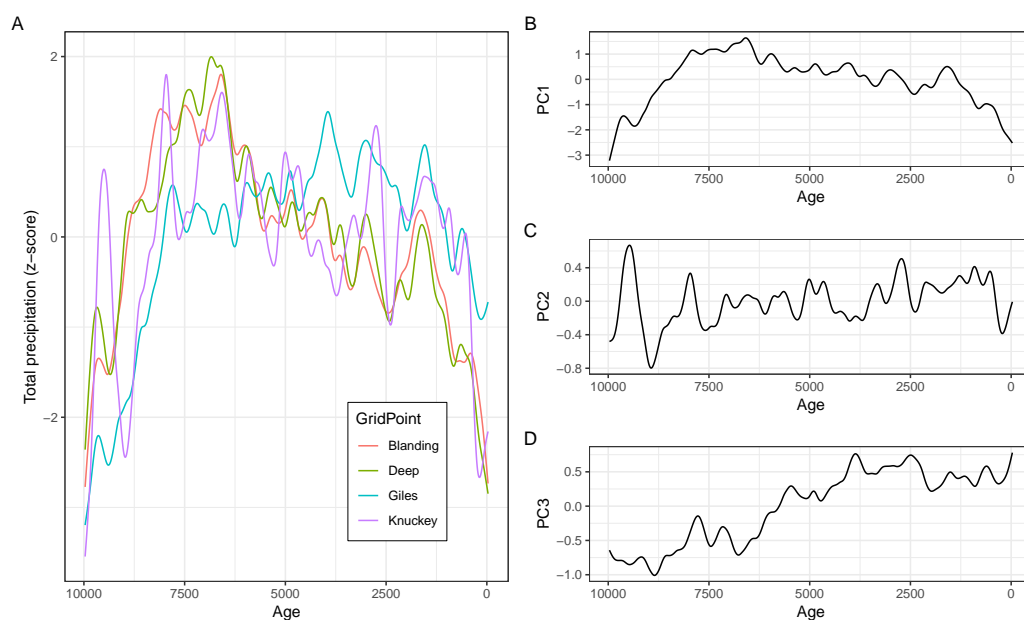


FIGURE E.7. A. Smoothed z-scores of total precipitation for the grid points in the TraCE-21ka simulation with lake-level records used in this study. B. First principal component (59% of the variance) of the TraCE-21ka precipitation data for the four grid points in A. Second principal component (22% of the variance) of the TraCE-21ka precipitation data for the four grid points in A. C. Third principal component (16% of the variance) of the TraCE-21ka precipitation data for the four grid points in A.

#### E.4.4. Continental-scale lake-level synthesis

Knuckey Lake represents a linchpin connecting a network of New England lake-level records 1500 km to the east and sites 1300 km to the west in the Central Rockies (Figure E.1) that documents centennial to multi-millennial hydrologic trends and variability over the past 6,200 years (Figure E.5). In this paper we focus on the eastern US network. Most multi-century events appear among subsets of sites in the network. At 5.6 ka, Giles, in Maine, and Blanding and Davis, in inland southern New England, all decreased and remained low until 4.9 ka. In contrast, Deep and New Long, in coastal Massachusetts, and Knuckey, in Minnesota, all increased over the same interval. All of the sites decreased from 4.9 to 4.6 ka and then increased from 4.6 to 4.2 ka. From 4.2 to 3.9 ka, all of the lake-levels, except Knuckey, decreased.

All eastern US lake-level records increased after 3.9 ka. Giles Pond maintained a high lake level until 2.6 ka, but many other lakes began to dry earlier, around 2.9 ka. Deep Pond recorded a major, transient dry event at 3.5 ka that was also weakly expressed at New Long Pond and Knuckey Lake. Some late-Holocene variability at Knuckey Lake appears anti-phased with southern New England sites (e.g., dry events at 3.2, 2.8, 2.0, and 1.5 ka). A dry event from 1.3 to 1.2 was recorded at Davis, Deep, and New Long. Giles Pond was wet during this period and Blanding had lost sensitivity due to its outlet stream (Shuman and Burrell, 2017). All of the records indicated high lake levels in the most recent part of the Holocene.

To isolate the coherent multicentennial variability we removed the insolation trend (Berger and Loutre, 1991) from the six eastern US lake-level records and performed principal components analysis (Figure E.6). PC1 explains 43.5% of the variance and features many of the multi-century events described above, including: the sequence of dry to wet to dry from 4.9 to 3.9 ka, the dry period from 2.9 to 2.1 ka, and the recent wet conditions. PC2 and PC3 explain 21.1% and 14.8% of the variance, respectively, and modulate the broad-scale features from PC1. All of the sites load positively onto the first empirical orthogonal function (EOF), but EOF2 and EOF3 have both positive and negative loadings. Giles loads strongly negatively on EOF2 while New Long (coastal MA) and Knuckey (MN) load positively. On EOF3, Deep Pond (coastal MA) loads negatively while Blanding and Davis (inland southern New England) load positively (Figure F.2).

#### E.4.5. Multi-centennial precipitation variability in the TraCE model simulation

The TraCE precipitation simulations over the past 10,000 years for the four grid points with lake-level records contain large decadal-scale variability modulated by multi-centennial to millennial scale variations (Figure E.7, F.3, F.4). The simulated precipitation time series generally lack the Holocene-length trends of the lake-level records because those trends are largely driven the effects of changing summer temperatures on water-balance (Shuman and Plank, 2011).

The multi-centennial to millennial variability in the simulated precipitation is comparable to that of the lake-level records. This suggests that this variability is a real part of

the physical climate system. The TraCE simulation can also quantify the ratio of decadal variability to the multi-centennial variability. Decadal variability is on the order of 50-100 mm/yr precipitation while the multi-centennial variability represents 15-30 mm/yr variations. These relatively small variations become important when averaged over long time periods (Shuman et al., 2018).

A principal components analysis on the TraCE precipitation data reveals a dominant first PC representing 59% of the variance with all positive EOF loadings. The second PC (22% variance) is dominated by loading from the Minnesota grid cell and looks similar to the Knuckey Lake lake-level reconstruction. The third PC (15% variance) is dominated by loading from the Maine grid cell and looks very similar to the Giles Pond lake-level reconstruction.

## E.5. Discussion

### E.5.1. Multi-scale Holocene hydrologic variability in Minnesota

Lake-level and bog water-table depth records complement one another because they are sensitive to environmental variability at markedly different temporal scales. The dry events at 1000 and 750 yr BP represent an opportunity for direct comparison of lake and bog responses. Sharp declines in the bog water-table depth appear as muted drawdowns in the lake-level record. Future work should monitor modern lake-level and bog water-table depth changes in the same region to better understand their relative scaling.

The Knuckey Lake record, in the context of the the warm and dry mid-Holocene in Minnesota, allows for estimation of the magnitude of precipitation changes necessary to drive lake-level changes. Quantitative estimates from pollen and lake-level modeling suggest that the maximum mid-Holocene dry period was characterized by precipitation decreases of up to 30% and temperature increases of 2°C (Bartlein et al., 1984; Almendinger, 1993). Thus, the increase in the Knuckey Lake lake-level from 6.2 ka to present (approximately 2.5 meters) represents an increase in precipitation of at most 30%. Multi-centennial variability at Knuckey Lake is characterized by around 0.5 m of lake-level change which could represent a precipitation change of up to 5%. This estimate is also supported by the analysis of the TraCE precipitation data in which the magnitude of multi-centennial variability is approximately 20% of the magnitude of decadal variability in simulated precipitation (Figure F.3).

Co-located lake-level and bog water-table depth records together also illustrate the smoothed, time-averaged version of past climate variability and change represented by lake-level reconstructions. Decadal, interannual, annual, seasonal variability are superimposed on the centennial to millennial scale variability in the lake-level reconstruction. Using different proxies and multiple archives, each with their own unique sensitivities provides a more realistic view of past climate variability, but analyzing networks with records from multiple different proxies requires careful analysis of each record and understanding of the unique climate-sensing properties of each proxy (Shuman et al., 2018; Marlon et al., 2017; Jackson, 2012; Evans et al., 2013).

### E.5.2. Bog water-table depth records resolve regional hydroclimatic variability

Water-table depth reconstructions from bogs in Minnesota and Michigan record dry events at 1000 and 750 yr BP at multiple sites (Booth et al., 2006). Bogs in Maine record dry events at 500 yr BP and 1800 yr BP (Clifford and Booth, 2015), but do not register the Upper Midwest events at 1000 and 750 yr BP. These events are also confirmed in principal component analyses. The lack of coherence between the Upper Midwest and northern New England could be a function of spatial covariance versus time-scale in the climate system (Mitchell, 1976). That is, on the decadal to multi-decadal time scales of the bog records the spatial covariance of hydroclimate may generally be limited to the regional level.

Sub-continental scale synthesis suggests that nearby testate amoebae-inferred water-table depth records document similar hydroclimatic variability. Individual records may have site-specific responses, but given adequate replication, a network of bog water-table depth records could accurately record spatial patterns of decadal to multi-decadal hydroclimatic variability over the past 2000 to 3000 years.

### E.5.3. Coherent multi-centennial lake-level variability across mid-latitude eastern North America

Our analysis of the lake-level network from mid-latitude North America shows clear, coherent multi-centennial hydroclimate variability in the eastern US over the past 6000 years. Some of these events are evident in all or nearly all records (e.g., the sequence

of wet to dry to wet from 4.9 to 3.9 ka), while others only appear in a subset (e.g., dry conditions at Giles, Blanding, and Davis from 5.5 to 4.9 ka, but wet conditions at Deep, New Long and Knuckey over the same interval). This indicates spatial patterns of past hydroclimate variability. We know from instrumental records and tree-ring based drought atlases that different spatial patterns of droughts and pluvials are possible. Comparison of patterns in the paleo record with patterns from the observational record and climate models allows for identification of underlying climate dynamical mechanisms (Woodhouse et al., 2009; Coats et al., 2015). Detailed investigation of the climate dynamical drivers of multi-centennial variability will require more Holocene-length climate model simulations.

Modes of climate variability that affect interannual hydroclimate variability in the modern system, such as the North Atlantic Oscillation (NAO) and Atlantic Multidecadal Variability (AMV), continue to drive interannual variability over the Holocene, but are unlikely to be the dominant causes of multi-centennial variability observed in the lake-level network. Analyses of the TraCE-21ka simulation suggests that multi-centennial variability is present in this long transient climate model run. The exact sequences of events vary, suggesting multi-centennial variability is primarily a result of unforced internal atmospheric variability.

The network of lake-level records contains an extensive set of correlated multi-centennial variability. Each lake has its own idiosyncrasies, but most major events are replicated at other nearby lakes. Principal components analysis further elucidates shared multi-centennial patterns. This suggests that a more extensive network of lake-level records across eastern



North America yields promise for generating a spatially-resolved reconstruction of multi-centennial hydroclimate variability over the last 6,000 to 8,000 years.

#### E.6. Acknowledgments

This research was supported by the National Science Foundation Macrosystems Biology Program (DEB-1241851). Fieldwork was supported by LacCore. Some data were obtained from the Neotoma Paleoecology Database (<http://www.neotomadb.org>). We thank Michael Grunst for field assistance. John Tipton, Jonathan King, and Kevin Anchukaitis provided useful comments that improved the manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

## REFERENCES

- Almendinger JE (1993) A groundwater model to explain past lake levels at Parkers Prairie, Minnesota, USA. *The Holocene* 3(2): 105–115.
- Ault TR, Cole JE, Overpeck JT, Pederson GT, St George S, Otto-Bliesner B, Woodhouse CA and Deser C (2013) The continuum of hydroclimate variability in western North America during the last millennium. *Journal of Climate* 26(16): 5863–5878.
- Bartlein PJ, Webb TI and Fleri E (1984) Holocene climatic change in the northern Midwest: pollen-derived estimates. *Quaternary Research* 22(3): 361–374.
- Berger A and Loutre MF (1991) Insolation values for the climate of the last 10 million years. *Quaternary Science Reviews* 10(4): 297–317.
- Booth RK (2010) Testing the climate sensitivity of peat-based paleoclimate reconstructions in mid-continental North America. *Quaternary Science Reviews* 29(5-6): 720–731.
- Booth RK, Jackson ST, Sousa VA, Sullivan ME, Minckley TA and Clifford MJ (2012) Multi-decadal drought and amplified moisture variability drove rapid forest community change in a humid region. *Ecology* 93(2): 219–226.
- Booth RK, Notaro M, Jackson ST and Kutzbach JE (2006) Widespread drought episodes in the western Great Lakes region during the past 2000 years: geographic extent and potential mechanisms. *Earth and Planetary Science Letters* 242(3-4): 415–427.
- Brugam RB, Grimm EC and Eyster-Smith NM (1988) Holocene environmental changes in Lily Lake, Minnesota inferred from fossil diatom and pollen assemblages. *Quaternary Research* 30(1): 53–66.
- Büntgen U, Franke J, Frank D, Wilson R, González-Rouco F and Esper J (2010) Assessing the spatial signature of European climate reconstructions. *Climate Research* 41(2): 125–130.
- Charman DJ, Hendon D, Woodland WA, Association QR et al. (2000) *The identification of testate amoebae (Protozoa: Rhizopoda) in peats*. Quaternary Research Association.
- Clifford MJ and Booth RK (2013) Increased probability of fire during late Holocene droughts in northern New England. *Climatic change* 119(3-4): 693–704.
- Clifford MJ and Booth RK (2015) Late-Holocene drought and fire drove a widespread change in forest community composition in eastern North America. *The Holocene* 25(7): 1102–1110.

- Coats S, Cook BI, Smerdon JE and Seager R (2015) North American pancontinental droughts in model simulations of the last millennium. *Journal of Climate* 28(5): 2025–2043.
- Coats S, Smerdon JE, Cook BI and Seager R (2013) Stationarity of the tropical pacific teleconnection to North America in CMIP5/PMIP3 model simulations. *Geophysical Research Letters* 40(18): 49274932.
- Coats S, Smerdon JE, Cook BI, Seager R, Cook ER and Anchukaitis KJ (2016) Internal ocean-atmosphere variability drives megadroughts in western North America. *Geophysical Research Letters* 43(18): 98869894.
- Cook E, Anchukaitis K, Buckley B, D'Arrigo R, Jacoby G and Wright W (2010) Asian monsoon failure and megadrought during the last millennium. *Science* 328(5977): 486489.
- Cook ER (2004) Long-term aridity changes in the western United States. *Science* 306(5698): 10151018.
- Cook ER, Seager R, Kushnir Y, Briffa KR, Büntgen U, Frank D, Krusic PJ, Tegel W, van der Schrier G, Andreu-Hayles L et al. (2015) Old World megadroughts and pluvials during the Common Era. *Science Advances* 1(10): e1500561.
- Dean WE (1993) Physical properties, mineralogy, and geochemistry of Holocene varved sediments from Elk Lake, Minnesota. *SPECIAL PAPERS-GEOLOGICAL SOCIETY OF AMERICA* : 135–135.
- Digerfeldt G, Almendinger JE and Björck S (1992) Reconstruction of past lake levels and their relation to groundwater hydrology in the Parkers prairie sandplain, west-central Minnesota. *Palaeogeography, Palaeoclimatology, Palaeoecology* 94(1-4): 99–118.
- DiNezio PN and Tierney JE (2013) The effect of sea level on glacial Indo-Pacific climate. *Nature Geoscience* 6(6): 485.
- Donovan JJ, Smith AJ, Panek VA, Engstrom DR and Ito E (2002) Climate-driven hydrologic transients in lake sediment records: calibration of groundwater conditions using 20th century drought. *Quaternary Science Reviews* 21(4-6): 605–624.
- Evans MN, Tolwinski-Ward SE, Thompson DM and Anchukaitis KJ (2013) Applications of proxy system modeling in high resolution paleoclimatology. *Quaternary science reviews* 76: 16–28.
- Fawcett PJ, Werne JP, Anderson RS, Heikoop JM, Brown ET, Berke MA, Smith SJ, Goff F, Donohoo-Hurley L, Cisneros-Dozal LM et al. (2011) Extended megadroughts in the southwestern United States during Pleistocene interglacials. *Nature* 470(7335): 518.

- Fritz SC, Ito E, Yu Z, Laird KR and Engstrom DR (2000) Hydrologic variation in the northern Great Plains during the last two millennia. *Quaternary Research* 53(2): 175–184.
- Gomez-Navarro JJ, Werner J, Wagner S, Luterbacher J and Zorita E (2015) Establishing the skill of climate field reconstruction techniques for precipitation with pseudoproxy experiments. *Climate dynamics* 45(5-6): 1395–1413.
- Hartmann D, Klein Tank A, Rusticucci M, Alexander L, Brönnimann S, Charabi Y, Dentener F, Dlugokencky E, Easterling D, Kaplan A, Soden B, Thorne P, Wild M and Zhai P (2013) *Observations: Atmosphere and Surface*, book section 2. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. ISBN ISBN 978-1-107-66182-0, p. 159254. DOI:10.1017/CBO9781107415324.008. URL [www.climatechange2013.org](http://www.climatechange2013.org).
- He F (2011) *Simulating transient climate evolution of the last deglaciation with CCSM 3*, volume 72. University of Wisconsin.
- Henderson AK, Nelson DM, Hu FS, Huang Y, Shuman BN and Williams JW (2010) Holocene precipitation seasonality captured by a dual hydrogen and oxygen isotope approach at Steel Lake, Minnesota. *Earth and Planetary Science Letters* 300(3-4): 205–214.
- Innes J, Blackford J and Simmons I (2004) Testing the integrity of fine spatial resolution palaeoecological records: microcharcoal data from near-duplicate peat profiles from the North York Moors, UK. *Palaeogeography, Palaeoclimatology, Palaeoecology* 214(4): 295–307.
- Jackson ST (2012) Representation of flora and vegetation in Quaternary fossil assemblages: known and unknown knowns and unknowns. *Quaternary Science Reviews* 49: 1–15.
- Jacques JMS, Cumming BF and Smol JP (2008) A 900-year pollen-inferred temperature and effective moisture record from varved Lake Mina, west-central Minnesota, USA. *Quaternary Science Reviews* 27(7-8): 781–796.
- Keen KL and Shane LC (1990) A continuous record of Holocene eolian activity and vegetation change at Lake Ann, east-central Minnesota. *Geological Society of America Bulletin* 102(12): 1646–1657.
- Liu Z, Otto-Bliesner B, He F, Brady E, Tomas R, Clark P, Carlson A, Lynch-Stieglitz J, Curry W, Brook E et al. (2009) Transient simulation of last deglaciation with a new mechanism for Bølling-Allerød warming. *Science* 325(5938): 310–314.
- Marcott SA, Shakun JD, Clark PU and Mix AC (2013) A reconstruction of regional and global temperature for the past 11,300 years. *Science* 339(6124): 1198–1201.

- Marlon JR, Pederson N, Nolan C, Goring S, Shuman B, Robertson A, Booth R, Bartlein PJ, Berke MA, Clifford M et al. (2017) Climatic history of the northeastern United States during the past 3000 years. *Climate of the Past* 13(10): 1355.
- Masson-Delmotte V, Schulz M, Abe-Ouchi A, Beer J, Ganopolski A, Gonzalez Ruoco J, Jansen E, Lambeck K, Luterbacher J, Naish T, Osborn T, Otto-Bliesner B, Quinn T, Ramesh R, Rojas M, Shao X and Timmermann A (2013) Information from Paleoclimate Archives. In: Intergovernmental PoCC (ed.) *Climate Change 2013 - The Physical Science Basis*. Cambridge: Cambridge University Press, pp. 383–464.
- Mayewski PA, Rohling EE, Stager JC, Karlén W, Maasch KA, Meeker LD, Meyerson EA, Gasse F, van Kreveld S, Holmgren K et al. (2004) Holocene climate variability. *Quaternary Research* 62(3): 243–255.
- Meko DM, Woodhouse CA, Baisan CA, Knight T, Lukas JJ, Hughes MK and Salzer MW (2007) Medieval drought in the upper Colorado river basin. *Geophysical Research Letters* 34(10).
- Mitchell JM (1976) An overview of climatic variability and its causal mechanisms. *Quaternary Research* 6(4): 481–493.
- Nelson DM and Hu FS (2008) Patterns and drivers of Holocene vegetational change near the prairie–forest ecotone in Minnesota: revisiting McAndrews transect. *New Phytologist* 179(2): 449–459.
- Nelson DM, Hu FS, Tian J, Stefanova I and Brown TA (2004) Response of C3 and C4 plants to middle-Holocene climatic variation near the prairie–forest ecotone of Minnesota. *Proceedings of the National Academy of Sciences* 101(2): 562–567.
- Newby PE, Shuman BN, Donnelly JP, Karnauskas KB and Marsicek J (2014) Centennial-to-millennial hydrologic trends and variability along the North Atlantic Coast, USA, during the Holocene. *Geophysical Research Letters* 41(12): 4300–4307.
- Nolan C, Booth RK, Shuman BN and Jackson ST (2019) Using co-located lake-level and bog water-table depth records to understand Holocene climate and vegetation changes in Maine. *Quaternary Science Reviews, in prep*.
- PAGES HC (2017) Comparing proxy and model estimates of hydroclimate variability and change over the Common Era. *Climate of the Past* 13(12): 1851–1900.
- PAGES kC (2013) Continental-scale temperature variability during the past two millennia. *Nature Geoscience* 6(5): 339–346.
- Parnell A (2014) Bchron: Radiocarbon dating, age-depth modelling, relative sea level rate estimation, and non-parametric phase modelling. *R package version* 4(1).

- Pribyl P and Shuman BN (2014) A computational approach to Quaternary lake-level reconstruction applied in the central Rocky Mountains, Wyoming, USA. *Quaternary Research* 82(1): 249–259.
- Rodysill JR, Anderson L, Cronin TM, Jones MC, Thompson RS, Wahl DB, Willard DA, Addison JA, Alder JR, Anderson KH, Anderson L, Barron JA, Bernhardt CE, Hostetler SW, Kehrwald NM, Khan NS, Richey JN, Starratt SW, Strickland LE, Toomey MR, Treat CC and Wingard GL (2018) A North American hydroclimate synthesis (NAHS) of the Common Era. *Global and Planetary Change* 162: 175–198.
- Saros JE, Fritz SC and Smith AJ (2000) Shifts in mid-to late-holocene anion composition in Elk Lake (Grant County, Minnesota): comparison of diatom and ostracode inferences. *Quaternary International* 67(1): 37–46.
- Shakun J, Clark P, He F, Marcott S, Mix A, Liu Z, Otto-Bliesner B, Schmittner A and Bard E (2012) Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature* 484(7392): 4954.
- Shuman B (2003) Controls on loss-on-ignition variation in cores from two shallow lakes in the northeastern United States. *Journal of Paleolimnology* 30(4): 371–385.
- Shuman B (2012) Patterns, processes, and impacts of abrupt climate change in a warm world: the past 11,700 years. *Wiley Interdisciplinary Reviews: Climate Change* 3(1): 19–43.
- Shuman B, Henderson AK, Plank C, Stefanova I and Ziegler SS (2009) Woodland-to-forest transition during prolonged drought in Minnesota after ca. AD 1300. *Ecology* 90(10): 2792–2807.
- Shuman B and Plank C (2011) Orbital, ice sheet, and possible solar controls on holocene moisture trends in the North Atlantic drainage basin. *Geology* 39(2): 151–154.
- Shuman BN and Burrell SA (2017) Centennial to millennial hydroclimatic fluctuations in the humid northeast United States during the Holocene. *Quaternary Research* 88(3): 514–524.
- Shuman BN, Routson C, McKay N, Fritz S, Kaufman D, Kirby ME, Nolan C, Pederson GT and St-Jacques JM (2018) Placing the Common Era in a Holocene context: Millennial-to-centennial patterns and trends in the hydroclimate of North America over the past 2000 years. *Clim. Past* .
- Shuman BN and Serravezza M (2017) Patterns of hydroclimatic change in the Rocky Mountains and surrounding regions since the last glacial maximum. *Quaternary Science Reviews* 173: 58–77.

- Smith AJ, Donovan JJ, Ito E, Engstrom DR and Panek VA (2002) Climate-driven hydrologic transients in lake sediment records: multiproxy record of mid-Holocene drought. *Quaternary Science Reviews* 21(4-6): 625–646.
- Stevenson S, Timmermann A, Chikamoto Y, Langford S and DiNezio P (2015) Stochastically generated North American megadroughts. *Journal of Climate* 28(5): 1865–1880.
- Stocker T, Qin D, Plattner GK, Alexander L, Allen S, Bindoff N, Breon FM, Church J, Cubasch U, Emori S, Forster P, Friedlingstein P, Gillett N, Gregory J, Hartmann D, Jansen E, Kirtman B, Knutti R, Krishna Kumar K, Lemke P, Marotzke J, Masson-Delmotte V, Meehl G, Mokhov I, Piao S, Ramaswamy V, Randall D, Rhein M, Rojas M, Sabine C, Shindell D, Talley L, Vaughan D and Xie SP (2013) *Technical Summary*, book section TS. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. ISBN 978-1-107-66182-0, p. 33115. DOI:10.1017/CBO9781107415324.005. URL [www.climatechange2013.org](http://www.climatechange2013.org).
- Thirumalai K, Quinn TM, Okumura Y, Richey JN, Partin JW, Poore RZ and Moreno-Chamorro E (2018) Pronounced centennial-scale Atlantic Ocean climate variability correlated with Western Hemisphere hydroclimate. *Nature communications* 9(1): 392.
- Tinner W and Hu FS (2003) Size parameters, size-class distribution and area-number relationship of microscopic charcoal: relevance for fire reconstruction. *The Holocene* 13(4): 499–505.
- Wanner H, Solomina O, Grosjean M, Ritz SP and Jetel M (2011) Structure and origin of Holocene cold events. *Quaternary Science Reviews* 30(21-22): 3109–3123.
- Webb TI, Cushing EJ and Wright Jr HE (1983) Holocene changes in the vegetation of the Midwest. In: *Quaternary environments of the United States, volume 2, The Holocene*. University of Minnesota Press, pp. 142–165.
- Woodhouse CA and Overpeck JT (1998) 2000 years of drought variability in the central United States. *Bulletin of the American Meteorological Society* 79(12): 2693–2714.
- Woodhouse CA, Russell JL and Cook ER (2009) Two modes of north american drought from instrumental and paleoclimatic data. *Journal of Climate* 22(16): 4336–4347.
- Wright HE, Mann D and Glaser P (1984) Piston corers for peat and lake sediments. *Ecology* 65(2): 657–659.
- Wright Jr HE, Winter TC and Patten HL (1963) Two pollen diagrams from southeastern Minnesota: Problems in the regional Late-Glacial and Postglacial vegetational history. *Geological Society of America Bulletin* 74(11): 1371–1396.

## F. APPENDIX F

SUPPLEMENTAL INFORMATION FOR COHERENT MID-TO-LATE HOLOCENE  
HYDROLOGIC VARIABILITY ACROSS TEMPORAL SCALES IN MID-LATITUDE  
EASTERN NORTH AMERICA



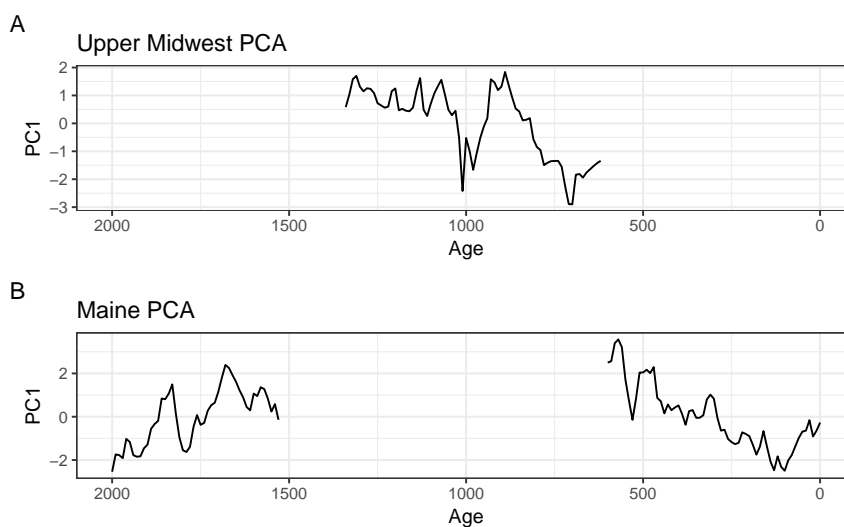


FIGURE F.1. A. First principal component (53% variance) of Upper Midwest bog records: Ely Lake Bog, Hole Bog, and Minden Bog. B. First principal component of two PCAs of Maine bog records. The PCA for present to 600 yr BP uses all four Maine bogs and the first PC represents 56% of the variance. The PCA for 1550 to 2000 yr BP uses Saco Bog, Sidney Bog, and Caribou Bog. The first PC represents 55% of the variance.

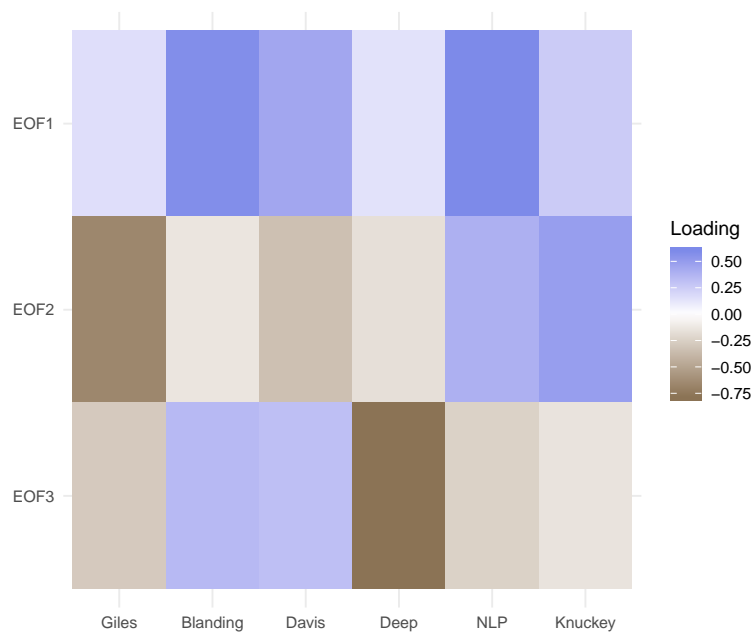


FIGURE F.2. EOF loadings for the six lake-level records from the eastern US.

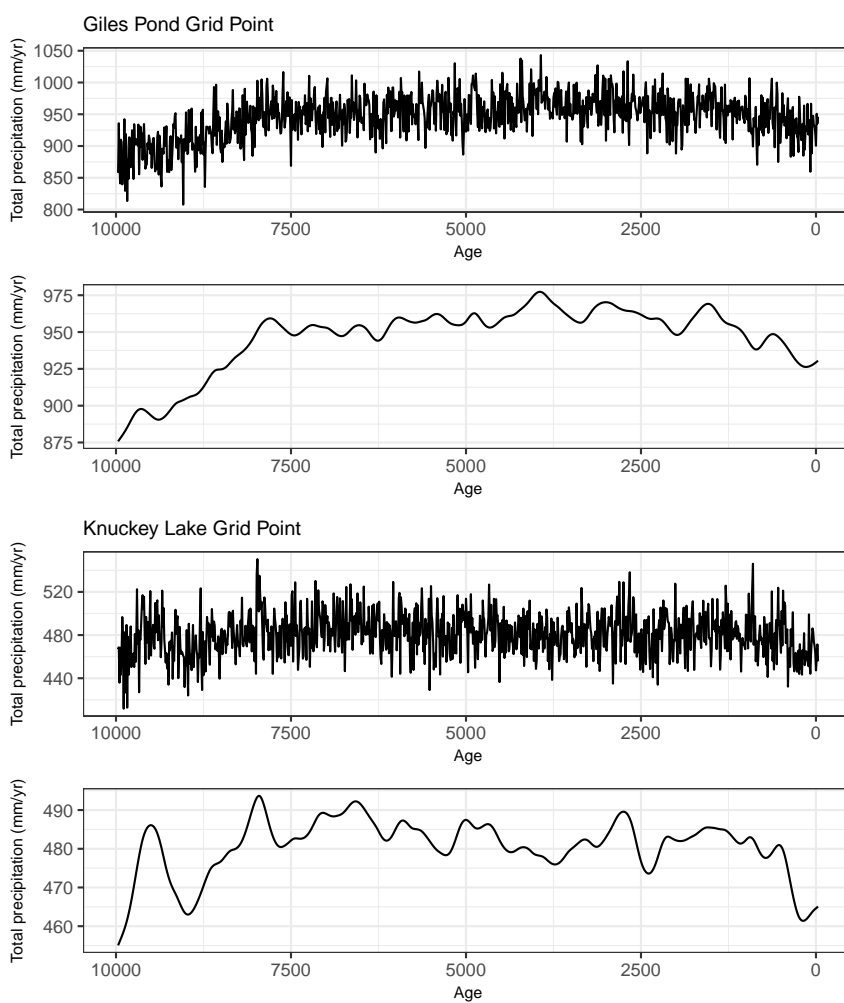


FIGURE F.3. TraCE-21ka total precipitation for the grid points that contain Giles Pond (A) and Knuckey Lake (C) and smoothed with a smooth spline with 40 degrees of freedom (B and D).

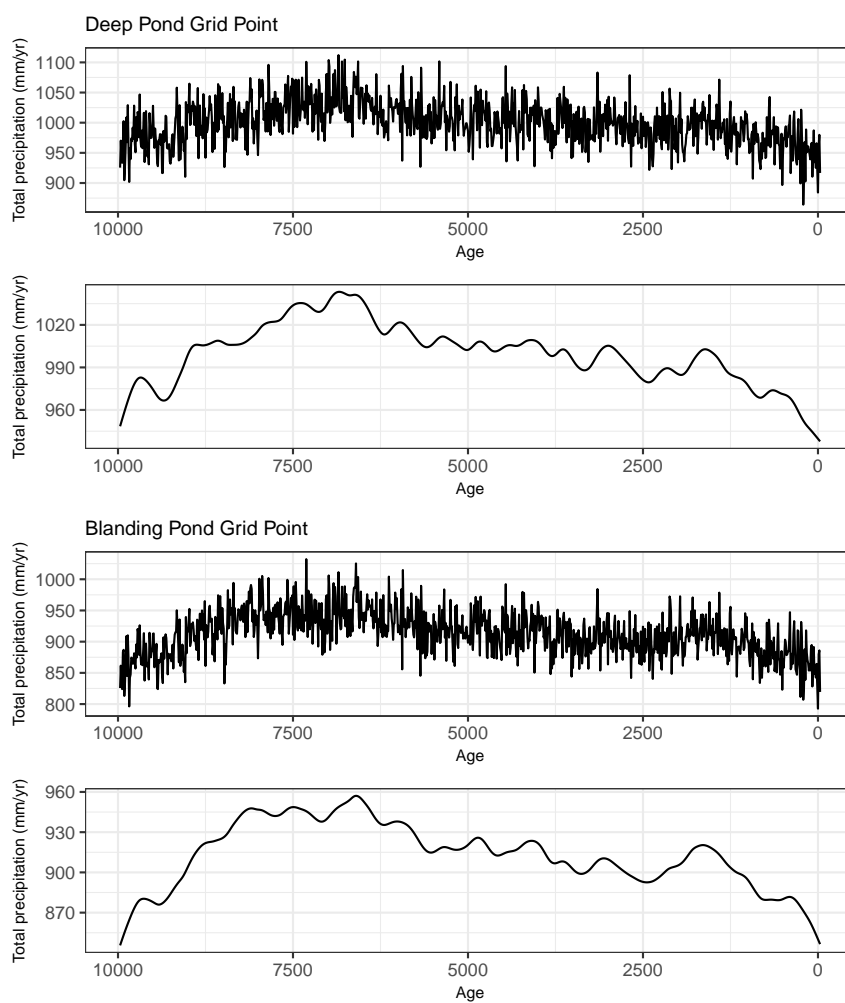


FIGURE F.4. TraCE-21ka total precipitation for the grid points that contain Deep Pond (A) and Blanding Pond (C) and smoothed with a smooth spline with 40 degrees of freedom (B and D).

Sample ID	Core	Depth (cm)	Material dated	<sup>14</sup> C Age	<sup>14</sup> C Error
208729	16A	20	Charcoal pieces	785	20
208730	16A	56	Charcoal pieces	1700	15
208731	16A	93	Charcoal pieces	3040	20
208732	16A	127	Charcoal pieces	3900	20
208733	19A	37	Charcoal pieces	1915	15
208734	19A	88	Charcoal pieces	2330	20
211349	19A	174	Charcoal pieces	4570	20
211351	22A	12	Charcoal pieces	540	20
211352	22A	19	Charcoal pieces	1070	20
208730	22A	28	Charcoal pieces	1110	15
211353	22A	111	Charcoal pieces	3125	15
211354	22A	142	Charcoal pieces	4140	25
211355	22A	180	Charcoal pieces	4495	20
208737	22A	195	Charcoal pieces	4475	20
211356	22A	216	Charcoal pieces	5600	20

TABLE F.1. Radiocarbon dates for Knuckey Lake.

Sample ID	Depth (cm)	Material dated	<sup>14</sup> C Age	<sup>14</sup> C Error
194151	76.5	Sphagnum stems	95	20
211341	87.5	Sphagnum stems	315	15
211342	92.5	Sphagnum stems	360	20
211343	97.5	Sphagnum stems	410	15
194152	102.5	Sphagnum stems	665	20
211344	117.5	Sphagnum stems	920	20
194153	127.5	Sphagnum stems	975	20
211345	137.5	Sphagnum stems	1105	15
194154	165.5	Sphagnum stems	1260	20
211346	171.5	Sphagnum stems	1410	20
211347	177.5	Sphagnum stems	1555	25
211348	182.5	Sphagnum stems	1815	15
194155	187.5	Sphagnum stems	2050	20
194156	192.5	Sphagnum stems	2150	20

TABLE F.2. Radiocarbon dates for Ely Lake Bog.