2010

Climate Change, Gene Flow, and the Legendary Synchrony of Snowshoe Hares

Ellen Cheng
The University of Montana

CLIMATE CHANGE, GENE FLOW, AND THE LEGENDARY SYNCHRONY OF SNOWSHOE HARES

By

ELLEN CHENG

B.S. Biological Sciences, Cornell University, Ithaca, New York, 1995
M.S. Wildlife Ecology, Utah State University, Logan, Utah, 2000

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in Fish and Wildlife Biology

December 2010

Approved by:

Perry Brown, Associate Provost for Graduate Education
Graduate School

Dr. L. Scott Mills, Co-Chair
Wildlife Biology

Dr. Karen E. Hodges, Co-Chair
University of British Columbia Okanagan

Dr. Jon Graham
Mathematical Sciences

Dr. Winsor H. Lowe
Wildlife Biology

Dr. Michael K. Schwartz
Wildlife Biology

Cheng, Ellen, Ph.D., Fall 2010 *Fish and Wildlife Biology*

Climate change, gene flow, and the legendary synchrony of snowshoe hares

Co-Chairperson: Dr. L. Scott Mills
Co-Chairperson: Dr. Karen E. Hodges

In recent decades, climate change has been invoked in the apparent collapse of some of the best-known examples of cyclic and synchronous population dynamics among boreal species. Simultaneously, some studies have predicted that as species' ranges shift poleward and southern habitats fragment in response to climate change, we will lose the southern glacial refugial populations that have historically harbored species' highest genetic diversity and uniqueness. I investigated how climate change and habitat fragmentation may impact genetic and population dynamic processes for the snowshoe hare *(Lepus americanus),* a species historically recognized as a key driver of North American boreal community dynamics.

I collected >1000 genetic samples and >300 time series from 175 cooperators in 30 U.S. states and Canadian provinces and territories. Based on analyses of nuclear and mitochondrial DNA, I identified three highly divergent groups of snowshoe hares in the Boreal, Pacific Northwest, and Southern Rockies regions of North America. I found high genetic diversity in mid-range (Boreal) hare populations, and high genetic uniqueness but lower diversity in the species' southern range (Pacific Northwest and Rockies). If southern populations decline due to climate change, snowshoe hares may still retain high genetic diversity, but will lose many alleles currently unique to southern populations.

In a simulation study comparing five synchrony metrics, I found the Kendall metric performed best with short, noisy time series similar to those available for snowshoe hares. I used this metric in partial Mantel tests, modified correlograms, and shifting window analyses of hare synchrony patterns. Confirming long-held but previously untested assumptions, I found northern hare populations are significantly synchronized at distances up to several thousand kilometers, while southern populations are not significantly synchronized at any of the distance classes evaluated. I found that historical patterns of synchrony still persist for snowshoe hares, in contrast to reports for some other synchronous species. Hare synchrony patterns clustered into groups defined according to genetic criteria—but not ecoregions or climatic regions—highlighting the importance of dispersal and population connectivity in snowshoe hare synchrony.

ACKNOWLEDGEMENTS

I wish I could individually thank the hundreds of people who have contributed to this project.

This work resulted from a major collaborative effort built on the foresight of those who recognized early the value of long-term wildlife monitoring. It also depended on the kindness of the many strangers who volunteered their time and efforts to increase our understanding of this phenomenal biological system.

I am especially grateful to the open generosity of the hunting and trapping community for their enthusiasm in supporting this study. Almost 100 individuals responded to cold calls, flyers, and magazine requests for snowshoe hare genetic samples. Over the years, many also recruited their friends and relatives to help as well. Becky McIntosh deserves special mention. I am lucky to have met Becky in the early years of this project. She invited me to publish articles in trapper magazines, encouraged me to attend my first trapper's convention, and instilled in me a respect for the experience, passion and scientific and observational ability of these conservationists. She and others opened doors for me that I will always value.

I would like to thank the many agency biologists and researchers who donated snowshoe hare genetic samples. In particular, thanks to Jake Ivan and Nate Berg—fellow graduate students who collected genetic samples for this project while also conducting their own graduate research.

The long-term time series compiled for synchrony analysis represent decades of work by many individuals I have never met. I am sincerely thankful to all who generously

shared their hard-earned data for this study. I benefited greatly from the excellent work and helpful insights of Charles Smith, who very kindly sent me boxes of data from his prior analysis of hare synchrony. I am also grateful to the many agency biologists who helped me and my remarkable field assistant, Matthew Strauser, with permitting and field logistics during several seasons of live-trapping hares. A special heartfelt thanks goes to Kelly Wolcott and Scott Fisher for opening their homes to ragged field biologists—their generosity will not be forgotten.

I am deeply indebted to my advisors, L. Scott Mills and Karen Hodges, for their encouragement and support for so many years. They gave me the independence I needed to find my place in this field, but were always there to guide me when I faltered.

Also, I owe a lot to my committee and other mentors at the University of Montana. Jon Graham has given so much of his time, effort, and statistical expertise to this work—I can't thank him enough for all he has done. I am incredibly grateful to Mike Schwartz for continually motivating and challenging me to think critically and become a better scientist. I appreciate the useful insights and constructive feedback that Winsor Lowe and Elizabeth Crone have provided on all aspects of my research. Paulo Alves and Jose Melo Ferreira were tremendous collaborators on this project—they hosted me at their laboratory in Portugal and trained me on genetic techniques and analyses. Joel Berger has taught me that there is an important place for people and compassion in conservation work. He is my role model for the international career I will soon embark on. Dan Pletscher has been a wonderful source of inspiration for me—he brings

leadership and excellence to University of Montana's Wildlife Biology Program, and a wonderful Dave Letterman smile that cheers everyone around him.

This project was only possible through funding from the University of Montana, the National Science Foundation, the U.S. National Park Service, and Animal Welfare Institute. I am grateful to these groups for their dedication to advancing science and conservation. I also deeply appreciate the contributions of many undergraduates and recent graduates who assisted this project in the field, laboratory, and especially in front of the computer.

I would like to thank my friends and colleagues in Missoula, especially the Mills and Berger labs, for being my scientific sounding boards and my family away from home. I also want to give a special thanks to Dave Wager and Jemma for being my closest companions and an important part of my life for most of the past decade.

My family has been my greatest source of support, stability, and encouragement throughout my life. This work is dedicated to them.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF APPENDICES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 1

INTRODUCTION TO THE DISSERTATION

Two central tenets of conservation biology are to maintain the biodiversity we have and to restore what we have lost. Conserving biodiversity means protecting tangibles including landscapes, species, and populations. It also means understanding and protecting intangibles such as genetic diversity and the large-scale ecological processes (e.g., wildfires, predator-prey dynamics, and long-distance migrations) that sustain functioning ecosystems.

Climate change and habitat fragmentation are well-recognized threats to the current and future persistence of species biodiversity (Pereira et al. 2010). How do these anthropogenic disturbances impact ecological processes and other intangible components of biodiversity? My dissertation addresses this question, by investigating patterns and mechanisms of genetic diversity and synchronized population dynamics for the snowshoe hare *(Lepus americanus)*, and how these dynamics may change in the face of future climate change and habitat fragmentation.

IMPETUS FOR RESEARCH

Current patterns of genetic diversity may primarily reflect effects of historical climate cycles (a phylogeography mechanism; Hewitt 1996; Hewitt 2000) or current distribution and dispersal (a core-periphery mechanism; reviewed in Eckert et al. 2008). For snowshoe hares and other boreal species these competing, but not mutually

exclusive, hypotheses generate different predictions for how genetic diversity and uniqueness should be distributed across a latitudinal gradient.

Under many scenarios of climate change, southern populations of boreal species are at great risk of decline and fragmentation over the next century (Iverson and Prasad 1998; IPCC 2001; IPCC 2007). Losing these southern populations could mean losing a large portion of species' total genetic diversities, potentially compromising their abilities to persist in a rapidly changing world. Knowing how genetic diversity is distributed and the processes leading to these patterns could help us pre-empt genetic losses, by identifying conservation challenges and management options.

Similar to genetic diversity, the phenomenon of synchronized population dynamics—the simultaneous rise and fall of populations over large spatial scales—is an intangible, but important, component of ecosystem biodiversity and function that may be threatened by climate change and habitat fragmentation. Over the past two decades, several studies have reported collapses of population cycles and synchrony in voles, grouse, and forest insect species (Williams et al. 2004; Bierman et al. 2006; reviewed in Ims et al. 2008; Kausrud et al. 2008). The apparent shift in the historically predictable dynamics of these species, and concerns about potential loss of associated ecosystem functions (e.g., large-scale pulsed resource flows and disturbances), raise urgencies for increased research attention on mechanisms of population synchrony, and potential dampening effects of climate change and habitat fragmentation on cycles and synchrony (Ims et al. 2008).

STUDY SYSTEM

The snowshoe hare provides an excellent model system for this study because it is abundant and at least patchily distributed across northern latitudes where intense climate-induced habitat change is predicted to first occur (Soja et al. 2006). The snowshoe hare has also historically been recognized as a key component of North America's boreal food web, making up 48% of the vertebrate biomass in a vast ecosystem spanning almost six million square kilometers (Krebs 2011). Snowshoe hares undergo dramatic ten-year population cycles, with repercussions for predators, other herbivores, and plants. These large fluctuations are synchronized across much of the species range, magnifying the spatial scale at which snowshoe hare population cycles drive boreal ecosystem dynamics (Bulmer 1974; Krebs et al. 2001). Thus, snowshoe hares play a critical role in boreal communities: their persistence should be a primary focus of ecological research and conservation efforts.

OVERVIEW OF DISSERTATION

As manifestations of large-scale, complex ecological processes, the study of genetic structure and population synchrony face common challenges in their necessity for large amounts of spatially distributed data and the difficulties of inferring processes from patterns. I approached the genetic structure and population synchrony of snowshoe hares in a similar framework. I cast a wide net across the North American range of snowshoe hares to compile genetic and time series data from hunters, trappers, agency scientists, and academic researchers. In total, I collected >1000 genetic

samples and >300 time series from 175 cooperators in 30 U.S. states and Canadian provinces and territories. From these data I quantified range-wide patterns of genetic diversity and population synchrony. I compared observed to expected patterns under competing, but not mutually exclusive, mechanistic hypotheses.

In my analysis of snowshoe hare genetic structure (Chapter 2), I asked if observed patterns of genetic diversity and uniqueness were more consistent with a core-periphery mechanism (lower genetic diversity in southern and northern peripheral populations compared to core populations) or with a southern refugia phylogeography mechanism (greatest genetic diversity and uniqueness in species' southern range). I found three genetically distinct evolutionary lineages of snowshoe hares, connected by limited gene flow—an extensive Boreal lineage comprising Canada, Alaska, and the northeastern U.S.; and two small, southern lineages occurring in the U.S. Pacific Northwest and southern Rockies. Genetic patterns supported the core-periphery model and revealed a more complex phylogeographic history than the southern refugia hypothesis would suggest—post-glacial recolonization apparently occurred from multiple source refugia, with secondary contact between refugial groups. This complex history, combined with current core-periphery dynamics, generated high genetic diversity and hotspots of genetic uniqueness in core (mid-range) populations, and high genetic uniqueness but low diversity in the species' southern peripheral populations.

What do these results mean for the future genetic diversity of snowshoe hares, in the face of climate change? The prognosis is mixed. If southern hare populations are lost or further fragmented due to climate change, the species will still maintain high

genetic diversity. However, this genetic diversity would represent a reduced subset of current genetic variation in snowshoe hares—many locally unique alleles currently found in the southern genetic groups (Pacific NW & Rockies populations) may be lost as southern populations decline.

While analyzing data on snowshoe hare genetic structure, I discovered that snowshoe hares in the Pacific Northwest of the U.S. have mitochondrial genes more closely related to that found in black-tailed jackrabbits *(Lepus californicus)* than in snowshoe hares. Preliminary results on this topic are presented in the appendix for the snowshoe hare genetics chapter. I will continue this study after completing my dissertation.

The snowshoe hare genetic results fed into my analysis of range-wide synchrony. If hare dispersal is a primary mechanism synchronizing dynamics, patterns of snowshoe hare genetic structure should correspond with patterns of synchrony. Therefore, although I present my research on snowshoe hare genetic structure and population synchrony as separate dissertation components, I use results from my investigation of genetic structure (Chapter 2) to determine if dispersal may be synchronizing range-wide hare dynamics (Chapter 4).

Collecting the geographically distributed, long-term time series data necessary to quantify synchrony patterns at a large spatial scale is a difficult task. The metrics used to quantify synchrony patterns have typically been evaluated in simulation studies assuming minimum time series lengths of 100 years and obvious cyclicity in the data. For snowshoe hares and many other synchronous species, available time series are

short, noisy, and range from barely fluctuating to highly cyclic. Prior to analyzing snowshoe hare synchrony patterns, I compared the performance of five synchrony metrics with simulated data representing typical snowshoe hare time series of varying lengths and different degrees of error (Chapter 3). I found that a metric based on the Kendall rank correlation coefficient exhibited the lowest variance and bias and the highest statistical power when applied to relatively short, noisy time series similar to those available for snowshoe hares. I used the findings from this simulation study to inform my analysis of snowshoe hare synchrony.

In my range-wide evaluation of snowshoe hare synchrony (Chapter 4), I determined if snowshoe hare populations are more synchronous in the northern compared to southern range—a common assumption that had not previously been tested. For northern hare populations, I asked if current synchrony patterns are similar to those reported from the early 1900's, given the reported loss of cycles and synchrony in some other systems. As an initial step in elucidating mechanisms of hare synchrony, I also asked if range-wide hare synchrony patterns correspond with ecoregions, climatic regions, or with the three genetic regions (Boreal, Pacific NW, and Rockies) identified in Chapter 2.

Based on time series primarily covering the past three decades, I found that northern hare populations are highly synchronous and synchrony 'travels' across the boreal ecosystem in a pattern similar to that observed for snowshoe hares 70 years ago. Thus, while historical cycles and synchrony have apparently collapsed for some species, I found no evidence of major dynamic shifts for snowshoe hares.

In contrast to northern populations, southern hare populations are not significantly synchronous at the large spatial scales (thousands of kilometers) examined. This result confirms long-held assumptions of a latitudinal gradient in hare population dynamics, similar to that reported for other taxa, including voles in northern Europe and Japan (Bjornstad et al. 1995, Saitoh et al. 1998, Tkadlec and Stenseth 2001), muskrats (Erb et al. 2000), autumnal moths (*Epirrita autumnata;* Klemola et al. 2002), and several species of grouse (Cattadori and Hudson 1999).

From my comparison of observed and expected patterns of synchrony, only genetically-defined regions significantly corresponded with observed patterns of hare synchrony. Patterns of snowshoe hare synchrony did not significantly correspond with ecoregions or climatic regions, counter to high-profile patterns suggested for a major predator of snowshoe hares, Canada lynx (*Lynx canadensis*) (Stenseth et al. 1999, 2004). Results do not preclude the possibility that climate influences large-scale hare dynamics, but three lines of evidence from this study suggest dispersal of hares or their major predators plays an important role in synchronizing northern hare dynamics: 1) the observed traveling wave of synchrony among northern hare populations is a common signature of dispersal-mediated synchrony (Haydon and Greenwood 2000; Bjornstad et al. 2002; Blasius et al. 1999); 2) high gene flow and synchrony in the relatively homogeneous northern boreal range compared to the naturally fragmented southern range suggest population connectivity promotes synchrony for hares; and 3) higher synchrony of populations within, compared to between, genetic groups (when

confounding effects of geographic distance are controlled for) is consistent with a dispersal mechanism for synchrony.

CONCLUDING THOUGHTS

Throughout my graduate career, a prevailing focus of my work has been the question of how can we monitor wildlife and distill the critical information needed to manage and conserve species? These were, after all, the objectives of most of the biologists and trappers who generously provided me with time series or genetic data, and are the goals of conservation biologists worldwide. My dissertation employed computer modeling and genetic analysis tools to assay key questions for population dynamics, while concurrent side projects on hare-habitat relations in Glacier National Park and the cost-efficiency of non-invasive genetic versus traditional estimates of snowshoe hare abundance gave me experience collecting and interpreting the field data crucial to these evaluations of population dynamics. Results of these side projects will be prepared for publication separately from the dissertation. However, the field work provided an important foundation for my dissertation, presenting me the tools to link field data on distribution, abundance, and genetic structure to the important, though less tangible, biodiversity processes of population synchrony and range-wide genetic structure.

LITERATURE CITED

Bierman, S. M., J. P. Fairbairn, et al. (2006). "Changes over time in the spatiotemporal dynamics of cyclic populations of field voles (Microtus agrestis L.)." American Naturalist **167**(4): 583-590.

Bjornstad, O.N., W. Falck, et al. (1995). “Geographic gradient in small rodent density-fluctuations--a statistical modeling approach.” Proceedings of the Royal Society of London. Series B **262:** 127-133.

Bulmer, M. G. (1974). "A statistical analysis of the 10-year cycle in Canada." Journal of Animal Ecology **43**(3): 701-718.

Cattadori, I.M. and Hudson, P.J. (1999). “Temporal dynamics of grouse at the southern edge of their distribution.” Ecography **22:** 373-374.

Eckert, C. G., K. E. Samis & S. C. Lougheed (2008). “Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond.” Molecular Ecology **17**: 1170.

Erb, J., N.C. Stenseth, et al. (2000). Geographic variation in population cycles of Canadian muskrats (Ondatra zibethicus). Canadian Journal of Zoology **78**: 1009-1016.

Hewitt, G. M. (1996). "Some genetic consequences of ice ages, and their role in divergence and speciation." Biological Journal of the Linnean Society **58**(3): 247-276.

Hewitt, G. M. (2000). "The genetic legacy of the Quaternary ice ages." Nature **405**(6789): 907-913.

Ims, R. A., J. A. Henden, et al. (2008). "Collapsing population cycles." Trends in Ecology & Evolution **23**(2): 79-86.

IPCC (2001). Climate change 2001: the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. New York, USA, Cambridge University Press.

IPCC (2007). Climate Change 2007: Impacts, Adaptation, and Vulnerability. Cambridge, UK, Cambridge University Press.

Iverson, L. R. and A. M. Prasad (1998). "Predicting abundance of 80 tree species following climate change in the eastern United States." Ecological Monographs **68**(4): 465-485.

Kausrud, K. L., A. Mysterud, et al. (2008). "Linking climate change to lemming cycles." Nature **456**(7218): 93-97.

Klemola, T., M. Tanhuanpaa, et al. (2002). “Specialist and generalist natural enemies as an explanation for geographical gradients in population cycles of northern herbivores.” Oikos **99:** 83-94.

Krebs, C. J. (2011). “Of lemmings and snowshoe hares: the ecology of northern Canada.” Proceedings of the Royal Society of London, Series B. Online early. DOI: 10.1098/rspb.2010.1992.

Krebs, C. J., R. Boonstra, et al. (2001). "What drives the 10-year cycle of snowshoe hares?" BioScience **51**(1): 25-35.

Pereira, H. M., P. W. Leadley, et al. (2010). “Scenarios for global biodiversity in the 21st century.” Science **330:** 1496-1501.

Saitoh, T., N.C. Stenseth, et al. (1998). “The population dynamics of the grey- sided vole in Hokkaido, Japan.” Researches on Population Ecology **40:**61-76.

Soja, A. J., N. M. Tchebakova, et al. (2006). "Climate-induced boreal forest change: Predictions versus current observations." Global and Planetary Change **56**(3-4): 274-296.

Stenseth, N. C., K. Chan, et al. (1999). "Common dynamic structure of Canada lynx populations within three climatic regions." Science **285**(5430): 1071-1073.

Stenseth, N. C., D. Ehrich, et al. (2004). "The effect of climatic forcing on population synchrony and genetic structuring of the Canadian lynx." Proceedings of the National Academy of Sciences **01**(16): 6056-6061.

Tkadlec, E., and Stenseth, N.C. (2001). “A new geographical gradient in vole population dynamics.” Proceedings of the Royal Society of London, Series B – Biological Sciences **268:** 1547-1552.

Williams, C. K., A. R. Ives, et al. (2004). "The collapse of cycles in the dynamics of North American grouse populations." Ecology Letters **7**(12): 1135-1142.

CHAPTER 2

SNOWSHOE HARES AND CLIMATE CHANGE:

HOW MUCH DOES THE SOUTHERN EDGE MATTER?

ABSTRACT

With climate change, the ranges of many boreal and northern temperate species are predicted to shift poleward, with increasing fragmentation and loss of populations in their southern ranges. Two competing models of genetic diversity make different predictions about the relative importance of these southern range populations as storehouses of species' genetic diversity, with implications for the effects of climate-related habitat loss on adaptive potential. Under a southern refugia phylogeography model, highest genetic diversity and uniqueness is found in a species' southern range, where populations have persisted and evolved through past ice ages. Under a core-periphery model, northern and southern peripheral populations harbor low genetic diversity due to low gene flow and chronic genetic drift characteristic of populations at the edge of a species' range. I conducted a range-wide study of snowshoe hare *(Lepus americanus)* genetic structure, to test these hypotheses on patterns of boreal genetic diversity and to improve our understanding of potential climate change impacts on boreal species. Trappers, hunters, agency biologists, and other researchers contributed to a large-scale sampling effort that yielded >1000 snowshoe hare genetic samples from 16 U.S. states and 12 Canadian provinces and territories. Based on analyses of mitochondrial DNA (cyt b and d-loop) and eight microsatellite loci, I identified three

highly divergent genetic groups of snowshoe hares: a Boreal group covering Canada, Alaska, and the northeastern U.S.; and two smaller groups in the Pacific Northwest and Rockies regions of the hare's southern range. Genetic patterns supported the core-periphery model and revealed a more complex phylogeographic history than the southern refugia hypothesis would suggest—post-glacial recolonization apparently occurred from multiple source refugia, with secondary contact between refugial groups. This complex history, combined with current core-periphery dynamics, generated high genetic diversity and hotspots of genetic uniqueness in core populations, and high genetic uniqueness but lower diversity in the species' southern range. If southern populations are further lost due to climate change, snowshoe hares may still retain high genetic diversity, but this genetic diversity would represent a reduced subset of current genetic variation in snowshoe hares due to loss of alleles currently unique to southern populations.

INTRODUCTION

Over the next century, North America's boreal forests are predicted to undergo major fragmentation and loss in their southern range due to global warming and human activities (1998; IPCC 2001; IPCC 2007). Current patterns of species genetic diversity may primarily reflect effects of historical climate cycles (a phylogeography hypothesis, Hewitt 1996; Hewitt 2000), current geography (a core-periphery hypothesis; reviewed in Eckert et al. 2008), or an interaction of these forces. These competing, but not mutually exclusive, hypotheses generate different predictions for how genetic diversity should be

distributed across the species range, and how future climate-associated habitat losses may impact species genetic diversity and adaptive potential. In a range-wide study of snowshoe hare *(Lepus americanus)* genetic structure, I tested two hypotheses of boreal patterns of genetic diversity, to improve our understanding of potential climate change impacts on boreal species.

According to one long-standing phylogeography model ('southern refugia'; Hewitt 1996; Hewitt 2000), genetic diversity and uniqueness should be greatest at the southern edge of the range for boreal and other northern hemisphere species, where populations persisted in glacial refugia through recurrent 90,000-year glacial periods of the Quaternary (~2.6 mya–present). During relatively brief 10,000-year interglacial periods, previously frozen habitats of the northern hemisphere were recolonized by refugial populations (Pielou 1991; Hewitt 1996; Hewitt 2000; Shafer et al. 2010). By this mechanism, recolonization via successive founding events generates a pattern of decreasing genetic diversity with increasing distance from southern refugia.

The core-periphery model—an alternative to the southern refugia phylogeography model—states that populations in the geographic core of a species' range may be more genetically variable than populations in the periphery. Peripheral populations often occur as small populations in marginal habitats, with low gene flow and chronic genetic drift (Lawton 1993; Vucetich and Waite 2003). The southern range for many North American boreal species consists of narrow peninsular range extensions into high latitude montane habitats of the USA, including the Olympic Mountains, Coast and Cascade Ranges, Sierra Nevada, Rocky Mountains, and Appalachian Mountains.

These southern peninsular populations tend to be smaller and more fragmented than core populations. Consistent with the core-periphery model, studies on fishers (Wisely et al. 2004), wolverines (Kyle and Strobeck 2002; Schwartz et al. 2007), and brown bears (Paetkau et al. 1998) have demonstrated greater genetic diversity in core populations compared to southern peripheral populations.

In the face of certain climate change and southern habitat loss for boreal species, the conservation implications of the southern refugia phylogeography and the core-periphery models of genetic diversity are very different, so it is important to distinguish the relative influence of these alternative mechanisms in distributing current genetic diversity and uniqueness. If patterns of boreal genetic diversity are primarily driven by phylogeographic history and a founding effect of northern populations since the Last Glacial Maximum (LGM, approximately 18,000 years ago) we would expect highest diversity below the southernmost extent of the LGM (45N – 49N, just below the U.S.-Canada border). Under many scenarios of climate change, much of the boreal and subalpine forests that currently support boreal species south of this delineation will decline over the next century (Iverson and Prasad 1998; IPCC 2001; IPCC 2007). In this case, we would lose a large portion of boreal species' total genetic diversities, potentially compromising their abilities to persist and adapt in a rapidly changing world.

Alternatively, if patterns of genetic diversity are driven primarily by drift in fragmented populations, we would expect lowest diversity in the southernmost montane extensions of the species range and highest diversity in mid-range core habitats, with diversity decreasing again toward the northern range periphery. In this

case, further loss of southern habitats due to climate change may have a relatively small impact on total species genetic diversity.

The snowshoe hare is an ideal species for a range-wide study of patterns of boreal genetic diversity. The hare has a distribution shared by many other North American boreal species. It is generally abundant and is a popular game animal in many parts of its range, so genetic samples are relatively easy collect. Snowshoe hares are also critical prey for Canada lynx *(Lynx canadensis)* and many other boreal carnivores and raptors (Keith 1963; Bulmer 1974; Finerty 1980; Hodges 2000).

Research to date lends mixed support for hypothesized patterns of snowshoe hare genetic diversity. Burton et al. (2002) reported high genetic diversity for snowshoe hares in Yukon, Canada, and two populations in Alaska and northern Montana, USA. The Montana population was highly genetically differentiated (F_{ST} = 0.20) and had several unique alleles, suggesting low gene flow between Montana and more northern hare populations. Morphological evidence also indicates the Pacific Northwest may divide into genetically differentiated populations—out of fifteen recognized subspecies of snowshoe hare, ten occur in British Columbia, Washington, Oregon, and Idaho (Dalquest 1942).

A few Late Wisconsin (20 kya – 10 kya) snowshoe hare fossils from Alaska, Yukon, Alberta, British Columbia, and Ontario suggest snowshoe hares may have survived the Last Glacial Maximum in the northern refugium of Beringia, and persisted near or quickly recolonized an ice-free corridor that opened between the Laurentide

and Cordilleran ice sheets (roughly along the British Columbia-Alberta border) during initial phases of glacial contraction (FAUNMAP Working Group 1994; Appendix 2.1).

I conducted an extensive range-wide nuclear and mitochondrial DNA (mtDNA) analysis to quantify snowshoe hare genetic structure and test these two mechanistic hypotheses of genetic diversity. I specifically asked if snowshoe hare genetic diversity and uniqueness are higher in southern populations (southern refugia phylogeography model) or mid-range populations (core-periphery model).

Gene flow, especially between southern and northern populations, may reduce genetic losses associated with climate change. I therefore used cluster analyses and several measures of genetic differentiation (F_{ST}, Nei's D, and Pi_{XY}) to assess patterns and levels of gene flow within the species range. In this study, I used the maximum southern extent of the LGM as an initial rough delineation for the southern species range because this region separates post-LGM recolonized populations from likely southern glacial refugia. Based on observed patterns of genetic structure, I discuss potential ecological implications of global warming and opportunities for prioritizing conservation efforts for snowshoe hares and other boreal species.

METHODS

Sample Collection

I collected 1014 snowshoe hare genetic samples from 16 U.S. states and 12 Canadian provinces and territories within the species' current range (Appendix 2.2). Ten samples were fecal pellet sets collected from individual hare tracks on Isle Royale,

Michigan, in winter 2009. Eleven samples were University of British Columbia Cowan Museum tissue samples from specimens collected near Vancouver, British Columbia, from 1929 to 1970. The remaining samples were snowshoe hare tissues collected from road kill, game harvests, and live-trapping during 1989–2010. Hunters and trappers, agency biologists, and other researchers donated many of the samples used in this study. Some samples came from on-going demographic studies in Montana and Wyoming (Mills et al. 2005; Hodges et al. 2009). I also conducted additional live-trapping in 2008–2010 to increase sample sizes from the western United States. Samples were stored in silica dessicant and maintained in cool, dry conditions prior to DNA extraction.

DNA Extraction and Mitochondrial DNA Sequencing

Most samples collected from the United States were extracted and amplified in our laboratory at The University of Montana (Missoula, MT), and submitted to The University of Washington's Hi-Throughput Sequencing Solutions laboratory (UWHTSEQ; Seattle, WA) for ExoSAP purification and bi-directional sequencing. A subset of U.S. samples was amplified at University of Porto's Research Center in Biodiversity and Genetic Resources (CIBIO; Vairão, Portugal) and sequenced at the Center of Molecular Analysis (CTM; Vairão, Portugal). Laboratory work was primarily conducted by the same author (E. Cheng) at both facilities. Samples collected from Canada were extracted and bi-directionally sequenced at University of British Columbia Okanagan's Fragment Analysis and DNA Sequencing Services laboratory (FADSS; Kelowna, BC).

I used Qiagen DNeasy Blood & Tissue Kits, following the manufacturer's protocol, to extract DNA from smaller tissue samples (e.g., 3 mm samples from live-trapped

hares). Larger tissue samples were digested in a detergent-based cell lysis buffer and purified by isopropyl alcohol precipitation. Pellet samples were extracted with the QIAamp DNA Stool Mini Kit in a laboratory at The University of Montana designated exclusively for samples collected non-invasively. I initially incubated single pellets in 1.6 mL Qiagen ASL buffer for 20 minutes at 54°C. The DNA-containing supernatant from each pellet was then used for extraction, following the manufacturer's protocol. Tissue or pellet samples that did not successfully extract after two attempts were omitted from analyses.

A 468 bp segment of the mitochondrial control region (D-loop) was amplified with primers LCRSEQ (Melo-Ferreira et al. 2007) and LepD2H (Pierpaoli et al. 1999). The 15 µL PCR reaction volume comprised 15 ng template DNA, 1X PCR buffer, 1.7 mmol $MgCl_2$, 0.13 mmol each dNTP, 0.13 µmol each primer, and 0.4 units Invitrogen Platinum Taq. I used a PCR program of 92°C for 2 minutes; 35 cycles of 92°C for 30 seconds, 52°C for 30 seconds, and 72°C for 30 seconds; followed by 72°C for 2 minutes. PCR products were visualized on a 1% agarose gel stained with ethidium bromide. Prior to purification and sequencing at UWHTSEQ, tissue PCR products were diluted 1:1 with DNA-grade water.

The mitochondrial cytochrome b (cyt b) gene has a slower mutation rate and lower tendency than D-loop for homoplasy over long time scales. Therefore, for 80 geographically distributed genetic samples I amplified a 633 bp segment of cyt b, to confirm D-loop groupings of haplotype lineages and to calculate deep divergence times between lineages. I used primers LGCYF (Alves et al. 2003) and LCYTBR (Melo-Ferreira et

al. 2005); reaction mixtures and PCR conditions were the same as described for the D-loop gene.

For both D-loop and cyt b, sequences were run on either an ABI 3730 or ABI 3130xl Genetic Analyzer (Applied Biosystems Inc., Foster City, CA). Sequences with unresolved bases were re-amplified and re-sequenced once; those still unresolved after two attempts were omitted from analyses. I aligned sequences using CodonCode Aligner v. 3.5.4 (CodonCode Corporation, Dedham, MA). Final sequences will be deposited in GenBank.

Microsatellite Genotyping

Microsatellite analyses were conducted at The University of Montana. I genotyped samples at 8 loci originally developed in the European rabbit, *Oryctolagus cuniculus*, and successfully used with snowshoe hares (Burton et al. 2002; Schwartz et al. 2007): 7L1D3 (Korstanje et al. 2003); SAT02, SAT12, SAT13, SAT16 (Mougel et al. 1997); SOL08, SOL30 (added "GTGTCTT" tail) (Rico et al. 1994); and SOL33 (Surridge et al. 1997). PCR amplifications were conducted as two multiplex reactions, each with 10 µL total volume. Primer ratios for Multiplex 1 were 9:12:12 for SAT13 : SOL08 : SOL30tailed. For Multiplex 2, primer ratios were 11:6:7:3.5:4 for SAT02 : SAT16 : 7L1D3 : SAT12 : SOL33. Each multiplex consisted of 0.2 µL 10X primer mix (forward primer was dye-labeled), 1X QIAmultiplex, and 1.5 ng template DNA.

I used two different touchdown PCR profiles. The program for Multiplex 1 was 95°C for 15 minutes; 24 cycles of 94°C for 30 seconds, 60-50°C for 3 minutes (step down 0.5°C/cycle for first 20 cycles), and 72°C for 1 minute; followed by 60°C for 30 minutes.

Multiplex 2 used an identical program but with a 58-48°C step down temperature cycle. PCR amplifications were run on an ABI 3130xl Genetic Analyzer at the Murdoch DNA Sequencing Facility (Missoula, MT) and scored with GeneMapper v. 3.7 (Applied Biosystems Inc., Foster City, CA).

I manually checked all microsatellite genotypes to confirm allele calls. A genotype was accepted if confirmed with at least two clear chromatograms. A homozygote genotype was not accepted for a locus if any amplification of the sample yielded a clear heterozygote genotype at that locus. If a heterozygote genotype could not be confirmed with a second clear chromatogram, the sample-locus was recorded as missing data in the final set. If one locus in a multiplex was not confirmed, the multiplex for that sample was re-amplified and re-genotyped up to three times. All samples in the final data set had at least seven confirmed loci.

Delimitation of Populations for Analyses

Samples were initially grouped into populations based on geographic location. I subsequently redefined some populations for which preliminary analyses identified genetic substructure (Wahlund effect). I used two geographic criteria: 1) no major landscape features such as large lakes, mountain ranges, or non-forested regions bisecting samples in a population; and 2) a maximum distance of 260 km between any two samples in a population. The second criterion (260 km between populations) was based on the spatial clustering of samples collected in this study, and Burton et al.'s (2002) finding of high gene flow among snowshoe hares separated by >600 km.

Next, for each population separately, I used Genepop v. 4.0.11 (Rousset 2008) to test for Hardy Weinberg Equilibrium (HWE) in all loci, and linkage disequilibrium in all pairs of loci. Markov chain parameters for exact tests were set at 10,000 dememorizations, 100 batches, and 5000 iterations per batch (Raymond and Rousset 1995). I corrected for multiple significance testing Type I error by using the false discovery rate (FDR) approach described by Benjamini and Hochberg (1995) and executed by the R software package "fdrtool" (Strimmer 2008; http://cran.r-project.org/). Potential null alleles and scoring errors due to stuttering and allelic drop-out were identified using Monte Carlo simulation in Micro-Checker v. 2.2.3 (Van Oosterhout et al. 2004). Based on these genetic criteria, I redefined population designations for genetic analyses described below. Populations with fewer than seven samples were not analyzed.

Mitochondrial DNA Analyses

I identified broad-scale genetic groups based on D-loop haplotypes, to distinguish snowshoe hare lineages with different evolutionary histories. I used the simulated annealing procedure in SAMOVA v. 1.0 (Dupanloup et al. 2002) to assign populations to genetic groups for K (number of groups) from 2 to 10. Likely group partitions are associated with high values of Φ_{CT}, which measures proportion of total genetic variation attributed to differences among groups. I ran SAMOVA with 500 initial population partitions and 10,000 iterations for each K. Significance of variance components was evaluated by 1000 permutations of populations among groups.

D-loop haplotypes as identified in DnaSP v. 5.10 (Librado and Rozas 2009) were used in NETWORK v. 4.5.1.6 (http://www.fluxus-engineering.com/) to generate an unrooted median-joining network for inferring ancestral haplotypes that might reveal source populations for post-glacial expansions, and for visualizing finer-scale evolutionary relationships. Nucleotide transitions were 19 times more likely to occur than transversions in my data (see RAxML analysis below), so I weighted transversions three times as high as transitions, as recommended by NETWORK for human mtDNA regions with a similar transitions : transversions rate. I compared the genetic clusters identified by SAMOVA and NETWORK to define the most likely groups (hereafter referred to as haplotype lineages) for subsequent lineage-based mtDNA analyses.

I estimated lineage divergence and expansion times based on mtDNA data by using RAxML 7.0.3 (Stamatakis 2006) to calculate parameters of a nucleotide substitution model. Two measures of genetic diversity—haplotype and nucleotide diversities averaged across loci for each population—were calculated in ARLEQUIN v. 3.5.1.2 (Excoffier et al. 2005). Haplotype diversity is the probability that two individuals randomly selected from a population have the same mitochondrial haplotype. Nucleotide diversity is the average number of nucleotide differences (per base pair) between individuals. For all populations combined, I conducted a quadratic regression analysis between diversity and latitude, to determine if geographic patterns of within-population genetic diversity are consistent with the core-periphery model of genetic drift at range edges. For all populations combined and for each haplotype lineage separately when sample sizes permitted, I also tested for significant linear correlations

between diversity and latitude and longitude, to identify geographic patterns that might reflect post-glacial expansion patterns from southern refugia (southern refugia phylogeography model).

To evaluate genetic differentiation between populations, I calculated pairwise F_{ST} and Pi_{XY} for all population pairs. F_{ST} is a measure of genetic differentiation (between populations) due to drift and fixation, whereas Pi_{XY} is the average number of pairwise differences between individuals from different populations. Both measures increase as gene flow decreases and populations diverge. Significance of tests was determined with 10,000 bootstraps and FDR control for multiple comparisons (ARLEQUIN v. 3.5.1.2).

Post-glacial expansion from refugia can generate signals of major spatial expansion from an initial equilibrium state to a final equilibrium state, which can follow a stepping-stone model of gene flow (e.g., Ray et al's 2003 simulations of spatial expansions occurring over a span of 400–4000 generations). I used ARLEQUIN v. 3.5.1.2 to test the D-loop mismatch distribution for each haplotype lineage against a model of spatial expansion. A mismatch distribution is a frequency histogram tabulating the number of pairwise differences among DNA sequences. Major spatial expansions generate unimodal mismatch distributions from which expansion parameters (time since expansion, pre-expansion and post-expansion effective population sizes) can be estimated (Excoffier 2004). I assessed data fit to expansion models and generated parameter confidence intervals with 20,000 coalescent simulations of the expansion process.

Microsatellite Analyses

Analyses of microsatellite data paralleled tests with mitochondrial DNA, allowing me to compare the long-term historical structure revealed by mtDNA with the more contemporary structure identified by independent microsatellite analyses. I used an individual-based Bayesian clustering program, STRUCTURE v. 2.3.3 (Pritchard et al. 2000), to identify the most likely number of microsatellite genetic clusters in samples, and to assign individuals to their probable cluster of origin. For each model run I used an admixture model with allele frequencies correlated among populations. Parameters were set for a burn-in period of 20,000 generations and 100,000 MCMC iterations after burn-in, as recommended by Pritchard et al. (2000). I used STRUCTURE-generated summary statistics (e.g., α, F, divergence distances among populations) and parameter plots to confirm the burn-in phase was adequate for the Markov chain to converge and parameters to reach equilibrium. To ensure MCMC replicates were sufficient, I checked for consistent results across twenty independent runs for each of K (number of clusters) from 2 to 10. I plotted the log-likelihood of each run against K. The smallest value of K at which the log-likelihood values begin to asymptote indicates an optimal partitioning of the genetic data. I compared results from STRUCTURE with those from SAMOVA on microsatellite data (model parameters: K = 2 to 10; 500 initial partitions and 10,000 iterations for each value of K). The microsatellite genetic clusters jointly defined by these two programs were used for subsequent cluster-based analyses.

I used GENALEX v. 6.3 to calculate summary statistics of genetic diversity averaged across all loci for each population: number of alleles, effective number of

alleles, observed and unbiased expected heterozygosity (Nei 1978), and the population inbreeding coefficient F_{IS}. For all pairs of populations I estimated two measures of genetic differentiation: Nei's D (Nei 1972) and Weir and Cockerham's (1984) F_{ST}. Nei's D is analogous to Pi_{XY} as a measure of the actual differentiation in allele frequencies between populations, whether due to drift or mutation. F_{ST} was calculated in ARLEQUIN v. 3.5.1.2 and significance was determined with 1000 permutations of samples among populations, with FDR correction for multiple comparisons.

Allelic richness (AR) is an important indicator of a population's genetic diversity and long-term evolutionary potential (Hill and Rabash 1986). I used the rarefaction procedure implemented in HP-RARE v. 1.0 (Kalinowski 2005) to calculate the average number of alleles per locus for each population, standardized to the smallest sample size (7 individuals) used in this study. The same program, in combination with GENALEX v. 6.3, allowed me to calculate a variety of statistics associated with the private allelic richness of each population and of each microsatellite genetic cluster.

Private allelic richness (PAR), the average number of alleles unique to a population, is often used to identify populations for priority conservation and to pinpoint regions within a species' current range that may have served as source refugia during glacial maxima. Populations must be relatively isolated for a sufficiently long time to accumulate private alleles. However, estimates of PAR can be biased by uneven density and distribution of sampling effort. In this study, northeastern states were more densely sampled than were some other portions of the species range. If private alleles were evenly distributed across the species range, PAR estimates for the northeast U.S.

could be lower simply due to sampling bias— less distance separates sampled populations in this region, so it would be more difficult to find alleles unique to only one sampled population.

To identify and account for these potential sampling biases in PAR estimates, I calculated the traditional standardized (to a sample of 7 individuals) PAR estimate as well as two "buffered" PAR estimates for each population: PAR_{350} is standardized PAR for each population when all sampled populations within a 350 km radius are excluded from calculation, and PAR_{500} is the same calculation with a 500 km buffer.

I distinguished between the various population PAR estimates (described above) and cluster PAR, which I estimated individually for each population as the average number of alleles private to the genetic cluster but potentially shared by multiple populations within the cluster. Cluster PAR (PAR_{clus}) was calculated as Kalinowski's (2004) hierarchical measure of PAR standardized to one population per genetic cluster and 7 individuals per population. To further identify patterns in the distribution of cluster private alleles I also calculated $\%PA_{clus}$, the proportion of alleles in each population that were cluster private alleles. As in mitochondrial DNA analyses, I conducted quadratic regression and linear correlation tests between genetic metrics (unbiased expected heterozygosity, allelic richness, PAR) and latitude and longitude to identify geographic patterns of population expansion or genetic drift at range periphery.

RESULTS

Mitochondrial DNA Data

I obtained D-loop sequences for 1006 samples and cytochrome b sequences for 80 samples. After exclusion of populations with fewer than seven complete sequences, the final data set for D-loop analyses comprised 39 populations with 893 samples represented by 365 haplotypes. Cyt b data were used only for confirming haplotype lineages and estimating lineage divergence times; I used the full 80-sample data set, which was represented by 43 haplotypes.

SAMOVA identified three major snowshoe hare haplotype groups with striking geographic correspondences: (i) a Pacific Northwest lineage including California, western Oregon, and western Washington but excluding the Olympic peninsula; (ii) a Rockies lineage comprising only Utah and Wyoming; and (iii) a Boreal lineage that covered the entire northern and eastern range of the species. In addition to the three major lineages, three smaller but also evolutionarily distinct lineages were associated with individual sampled populations: (i) WA3 in Olympic National Park, USA; (ii) OR2 in Malheur National Forest, eastern Oregon, USA; and (iii) CO1 (Colorado). (Fig. 2.1; Table 2.1)

The Pacific NW and WA3 snowshoe hares were highly diverged from all other lineages. A BLAST-search of their haplotypes (Altschul et al. 1990) showed all individuals sampled from these populations were more closely related to black-tailed jackrabbits *(Lepus californicus)* than to snowshoe hares. This finding was further confirmed by

phylogenetic and network analyses that located black-tailed jackrabbits basal to these introgressed snowshoe hare lineages (Appendix 2.3).

Because the introgressed hares were so genetically distinct from all other snowshoe hares, the proportion of genetic variation explained by differences among lineages (Φ_{CT}) was very high and differed by only 1% for K = 2 to 10 in SAMOVA analysis (Table 2.1A). To better identify an optimal K, I reran SAMOVA with the introgressed hares excluded (Table 2.1B). In this subset analysis K = 4 identified the same non-introgressed lineages defined by K = 6 in the full SAMOVA run. The introgressed lineages did not further subdivide in the full SAMOVA until K = 9. Thus, the best-supported groups were derived from K = 6 in the full SAMOVA. Both SAMOVA analyses indicated strong differentiation of these lineages, which were also clearly defined in the haplotype networks for the D-loop (Fig. 2.2) and cyt b (not shown).

No haplotypes were shared by populations in different lineages (Appendix 2.4). However, two populations (WA4 and MT1) with primarily Boreal haplotypes also had haplotypes associated with a non-Boreal lineage as defined in the D-loop haplotype network (Fig. 2.3). These mixed-lineage populations appear to be points of contact and limited gene flow between otherwise highly distinct snowshoe hare lineages. D-loop haplotypes were shared among populations in the Boreal lineage. A few haplotypes, such as Hap2 found in British Columbia (BC1) and Quebec (QC3), were shared by distant populations. The most ancestral haplotypes within the Boreal lineage occurred in the eastern species range (New Brunswick /New England region). West Virginia was the only Boreal population for which I did not find shared haplotypes. This population has been

relatively isolated from more northern populations since the early 1900's (Brooks 1955). Within the Pacific Northwest introgressed lineage, haplotypes were shared only among the three southernmost populations sampled: Oregon (OR1) and California (CA 1 and CA2). Although UT1 (Utah) and WY1 (Wyoming) were combined in a single lineage, I did not find any shared haplotypes between these populations. In summary, the haplotype network confirmed SAMOVA results that snowshoe hare populations are divided among several highly differentiated evolutionary lineages. Populations in the Boreal lineage share many haplotypes, due to high gene flow and/or incomplete lineage sorting. Most populations in other lineages retain unique haplotype sets.

Most lineages and populations were characterized by high haplotype and nucleotide diversities (Table 2.2; Appendices 2.5 and 2.6), suggesting historical accumulations of genetic variation and/or secondary admixture of dispersers from different refugial populations. In general, the UT-WY lineage had the lowest diversities of all. Within the Boreal lineage, BC1 (Vancouver, British Columbia) had unusually low diversity (only 2 haplotypes among 17 samples), which may reflect relative geographic isolation or recent glacial recession. Island and southern range peninsular populations also had low diversities, especially noticeable for IR1 (Isle Royale island, Michigan), WV1 (West Virginia, at the southern edge of the Boreal lineage range), and CA1 and CA2 (California, at the southern edge of the Pacific NW lineage range). Most other populations had haplotype diversities >0.85 and correspondingly high nucleotide diversities. The unusually high nucleotide diversity for WA4 ($\pi = 0.0681$) can be

attributed to the joint occurrence of haplotypes from introgressed and non-introgressed lineages in this population (Figs. 2.2 and 2.3).

For analyses of geographic patterns in genetic diversity, I omitted BC1 because its unusually low diversity could have obscured large-scale patterns. I also excluded the two mixed-lineage populations, WA4 and MT1. With all other populations combined, haplotype diversity ($r^2 = 0.35$, $p = 0.0003$; Fig. 2.4) and nucleotide diversity ($r^2 = 0.20$, $p = 0.009$) had significant quadratic relationships with latitude. I did not find significant correlations between longitude and any measure of genetic diversity (all $p > 0.17$).

The two measures of genetic differentiation, F_{ST} and Pi_{XY},were highly correlated ($r = 0.85$, $p < 0.001$). Average pairwise differences across lineages (Pi_{XY}) were high, especially between introgressed and non-introgressed lineages (Appendix 2.7). Within the Boreal lineage, low values of F_{ST} and Pi_{XY} between populations suggested moderate to high gene flow. Other lineages showed strong genetic structure and little gene flow.

I used the subset of cyt b data to calculate time since divergence of snowshoe hare lineages. Two studies of *Lepus* phylogeny have placed the divergence between snowshoe hares and black-tailed jackrabbits at 4.79 mya (95% CI 4.03–5.90 mya; Matthee et al. 2004) and 5.649 mya (95% CI 3.504–8.097 mya; Wu et al. 2005). I obtained an average divergence between black-tailed jackrabbits and non-introgressed snowshoe hares of Dxy(JC) = 0.10232 in this study (Dxy calculated in DnaSP v. 5.10 with Jukes-Cantor correction; cyt b data included five black-tailed jackrabbit samples). Using this average species divergence and the published divergence times, I calculated a cyt b divergence rate ranging from 1.26% to 2.92% per million years. I used this range of

divergence rates to determine that snowshoe hare haplotype lineages likely separated during the Pleistocene, prior to the LGM. Most recently, CO1 separated from the UT-WY lineage approximately 40 kya (60 kya–26 kya; Table 2.3). The oldest divergence was probably between the Boreal lineage and all other non-introgressed lineages ~ 1 mya (1.5 mya–65 kya), although there is considerable overlap between the range of divergence times for many lineage separations. Post-introgression divergence of Pacific NW and WA3 snowshoe hare populations from black-tailed jackrabbits seems to have occurred around the same time as divergence of Boreal snowshoe hares from all other lineages. Although there can be considerable error associated with divergence estimates from genetic data, it seems most snowshoe hare haplotype lineages separated from each other at some time in the early- to mid-Pleistocene, with onset of the major Quaternary glacial cycles. Thus, recolonization of recently unglaciated northern territories after the LGM (~18,000 years ago) must have occurred primarily by species expansion from Boreal refugia, rather than from refugia of more southern lineages.

The data suggest the Boreal and Pacific NW lineages underwent spatial expansions during the late Pleistocene but well before the LGM. The pairwise mismatch distributions of D-loop haplotypes for these lineages fit expected distributions from a spatial expansion model (Fig. 2.5). I used the estimated lineage divergence times for CO1 / UT-WY (0.60–0.26 mya) to calibrate a D-loop mutation rate that could be used to estimate time since expansion of these lineages. A divergence of Dxy(JC) = 0.038 for CO1 / UT-WY yielded a mutation rate of 3.2% to 7.3% per million years. For a 460 bp D-loop segment (indels excluded), this rate translates to 14.5 to 33.5 mutations per million

years. Based on this mutation rate and mismatch model parameters for tau (units of mutational time since expansion), the Boreal and Pacific NW lineages may have undergone spatial expansions around the same time, between 211 kya and 90 kya.

Microsatellite Data

I obtained final microsatellite genotypes for 922 samples. After excluding populations with fewer than 7 confirmed genotypes, I had 853 samples in 39 populations. Four percent of 343 population-loci combinations were significantly out of Hardy Weinberg Equilibrium after FDR-control for multiple comparisons. Almost half (N = 6 of 14) of the significant results were for two populations, ME1 (Maine) and NY1 (New York). Deviations from HWE were not associated with any particular loci. Significant results were generally associated with positive F_{IS} values, indicating a deficit of heterozygotes. Slightly over 5% of 1026 tests for linkage disequilibrium (covering all loci pairs for all populations) were significant after FDR-control. Most of the significant results (46 out of 53) were again associated with ME1 and NY1. Micro-checker identified potential null alleles in 7.6% of 328 population-loci tests. Null alleles and deviations from HWE were not associated more frequently with any particular loci, and genotypic disequilibrium was not consistently attributed to a particular locus pair. Burton et al's (2002) study also did not report problems with any of the microsatellite markers used in common with my research. I therefore retained all loci for analyses. I re-examined ME1, NY1, and other populations with high F_{IS} and significant deviation from HWE, to determine if heterozygote deficits could be due to clustered sampling within

populations. I redefined one population (MB1 in Manitoba) for which elimination of sample clusters restored population equilibrium.

The microsatellite data largely confirmed the independent findings from mitochondrial data regarding patterns of snowshoe hare genetic structure and diversity. To distinguish between microsatellite-based genetic groups and mtDNA-based haplotype lineages, I will refer to the former as genetic clusters. Microsatellite data identified three genetic clusters that broadly corresponded with the three major haplotype lineages from mtDNA analyses: Boreal, Pacific NW, and Rockies (Fig. 2.1; Table 2.4). However, with microsatellite analyses the Pacific NW cluster was larger—it included all introgressed hare populations as well as populations from eastern Oregon, eastern Washington, southern British Columbia, and northern Montana. Microsatellite data also grouped CO1 with the Rockies cluster. Group composition was identical at K = 3 in both SAMOVA and STRUCTURE, and the log-likelihood plot for STRUCTURE identified K = 3 as the best-supported division for microsatellite data (Appendix 2.8). With K = 3, the estimated membership of every population in its most likely genetic cluster was > 90%, when averaged across all individuals in the population. Differences among the three genetic clusters accounted for 16% of total genetic variation.

Most snowshoe hare populations were characterized by high allelic richness and high expected heterozygosities (Table 2.5). These measures of genetic diversity were highly correlated ($r = 0.94$, $p < 0.001$). Among Boreal populations I found low genetic diversities in Alaska, Yukon, and British Columbia (Figs. 2.6 and 2.7). This region of western North America was the last to recover from the most recent ice age (Fulton

1991), although some areas of central Alaska and far western Yukon (eastern Beringia) remained ice-free through the glaciations. Consistent with mitochondrial data and with the effects of drift on genetic diversity, I also found low diversities in the Rockies microsatellite cluster and in southern range peninsular populations: PA1 and PA2 (Pennsylvania) and WV1 (West Virginia) of the Boreal cluster, and CA1 and CA2 (California) in the Pacific NW.

BC1 grouped with the Pacific NW cluster but its geographic position and its very low genetic diversity suggest it was recolonized within the past 6,000 years, when the southern remnant of the Cordilleran glacier finally disappeared. Alternatively or in addition, genetic drift may explain its low diversity. As in mitochondrial analyses, I omitted this population from my examination of large-scale geographic trends in genetic diversity. Genetic diversity was weakly correlated with longitude for the Boreal cluster (allelic richness, $r = 0.32$, $p = 0.10$; heterozygosity, $r = 0.56$, $p = 0.003$) and for all populations combined (allelic richness, $r = 0.32$, $p = 0.05$; heterozygosity, $r = 0.45$, $p = 0.005$). With all populations combined, genetic diversity had a significant quadratic relationship with latitude ($r^2 = 0.35 – 0.37$, $p < 0.0002$). Diversity was also highly linearly correlated with latitude for the Pacific NW cluster (allelic richness, $r = 0.74$, $p = 0.04$; heterozygosity, $r = 0.77$, $p = 0.03$; Table 2.6; Figs. 2.8 and 2.9).

The two measures of genetic differentiation, F_{ST} and Nei's D, were highly correlated across all populations ($r = 0.93$, $p < 0.001$). An F_{ST} of 0.20 means approximately 20% of genetic variation is distributed among populations and 80% within populations, when variation is measured as expected heterozygosity. Under

simplifying assumptions, an F_{ST} of 0.20 translates to one migrant per generation, a level of gene flow adequate for minimizing inbreeding depression while maintaining adaptive variation (Wright 1931; Mills and Allendorf 1996). High pairwise F_{ST} (usually > 0.20) and Nei's D supported my earlier findings of little gene flow between genetic clusters and between populations in the Pacific NW and Rockies clusters (Table 2.7). In contrast, gene flow was moderate to high (F_{ST} ~ 0.03 to 0.15) between populations of the Boreal cluster.

I found high private allelic richness (PAR) in many Rockies and northern Pacific NW populations, consistent with their high F_{ST} and Nei's D estimates and long-term refugial status (Table 2.5; Fig. 2.10). However, Wyoming and California had notably low population PAR. It is difficult to pinpoint causes, but genetic drift or recent bottlenecks can reduce PAR through loss of rare private alleles (Nei et al. 1975). In the Boreal cluster, several populations stood out for relatively high PAR values: AK6 (southeast Alaska), AB1 and AB2 (southern Alberta), SK1 (Saskatchewan), MI1 (Michigan), and NB1 (New Brunswick). An equal number of Boreal populations had very low PAR values. When I buffered PAR estimates to minimize biases due to uneven spatial sampling densities (PAR_{350} and PAR_{500}), the same populations as above were identified for high PAR. Estimates for northern Pacific NW populations increased sharply with a 350 km PAR buffer, suggesting many alleles were shared by nearby sampled populations but not by more distant populations. Results changed little when the PAR buffer was increased to 500 km (not shown).

PAR_{clus} and $\%PA_{clus}$ estimates give a different picture with useful insights on the distribution of alleles among sampled populations. High PAR_{clus} and $\%PA_{clus}$ values for most Boreal populations and especially for the eastern Boreal range reveal a high diversity of microsatellite alleles that are found only in the Boreal cluster ("cluster private alleles"), but are well-distributed among Boreal populations (Fig. 11). On average, cluster private alleles constituted 15% of total allelic richness in Boreal populations but only 2 - 4% of total allelic richness in Pacific NW and Rockies populations (Table 2.5).

DISCUSSION

Large-scale Genetic Structure

Based on multiple genetic markers analyzed in over 1000 individuals sampled across the entire range of snowshoe hares, I detected three major genetic groups of snowshoe hares in the Boreal, Rockies, and Pacific Northwest regions of North America. MtDNA and microsatellite analyses differed slightly in delineation of haplotype lineages and genetic clusters. I use "groups" to broadly reference the three genetically distinct snowshoe hare assemblages identified by both types of markers, and "lineages" when specifically referring to delineations based on mtDNA analyses.

The Boreal group includes the entire northern, core, and eastern range of the snowshoe hare. Much of this area was covered by glaciers during the LGM and prior glacial advances. This large group is characterized by extensive gene flow across its 6000 km span. The Rockies and Pacific NW groups occur in the southwest species range,

which was largely ice-free during the Pleistocene. Both groups are characterized by low gene flow and highly differentiated populations.

The Pacific Northwest has previously been identified for exceptionally high levels of genetic differentiation (both within the Pacific NW and between the Pacific NW and other regions) for many boreal and temperate species, including Oregon spotted frogs (*Rana pretiosa)*, Pacific giant salamanders *(Dicamptodon tenebrosus),* American black bear *(Ursus americanus)*, various genera of squirrels, voles, and several species of birds (Arbogast 1999; Arbogast and Kenagy 2001). The Pacific NW group of snowshoe hares is unique in its nuclear DNA and in its inclusion of snowshoe hare populations apparently completely fixed for introgressed black-tailed jackrabbit mitochondrial DNA from historical hybridization. After introgression, divergence of Pacific NW hares from black-tailed jackrabbits is estimated to have occurred ~1 mya (1.5 mya–65 kya), coincident with early- to mid-Pleistocene glacial cycles and also with divergence of Boreal snowshoe hares from the Rockies group.

Very little recent or historical gene flow connects hares among these groups since their divergence. Lines of evidence supporting this conclusion include similarities in the major genetic groups determined by mtDNA and microsatellite analyses, the large number of mutations separating the groups in a haplotype network, a complete absence of haplotypes shared between groups, a large number of microsatellite alleles and mtDNA haplotypes private to each group, and high estimates of genetic differentiation (F_{ST}, Pi_{XY} and Nei's D) between groups. Given the deep divergences for these groups, recolonization of newly available Boreal habitats after the LGM must have occurred

primarily from refugial populations within the Boreal group rather than from the Pacific NW or Rockies groups.

Latitudinal Patterns of Genetic Diversity

A primary objective of this study was to test competing hypotheses about latitudinal patterns of genetic diversity in snowshoe hares. With all groups combined, I found a significant quadratic relationship between all measures of genetic diversity and latitude. Consistent with the core-periphery model, diversity was lowest in southern range peninsular populations, highest for core populations at mid-range latitudes, and declined again for populations in the northern range.

Recolonization of the Boreal group likely did not occur by rapid, successive founding from a primary southern refugium. The data do not support this scenario, which should have generated a pattern of southern richness and northern purity. Instead, high genetic diversities, complex patterns of haplotype sharing among populations, and prior fossil evidence (FAUNMAP Working Group 1994) all point to multiple-source recolonization of the Boreal with high levels of secondary admixture. Therefore, both current geography and a complex phylogeographic history may have contributed to observed patterns of genetic diversity for snowshoe hares, supporting a more nuanced view of mechanisms generating and distributing genetic diversity (Stewart and Lister 2001; Petit et al. 2003; Shafer et al. 2010).

Spatial Expansions

My assessment of nucleotide mismatch distributions determined that major spatial expansions of snowshoe hare mitochondrial lineages pre-dated the LGM. Using

best-fit parameters to expansion models and a calibrated D-loop mutation rate, I estimated that spatial expansions occurred within both the Boreal and Pacific NW groups approximately 211 kya – 90 kya. These estimates are approximations because they are based on several assumptions about homogeneity of mutation rates in mtDNA. However, it seems reasonable to conclude at a minimum that major expansions in these lineages occurred prior to the LGM. This finding corresponds with the high genetic diversities currently observed in Boreal populations. The haplotype network also supports this scenario. Within the Boreal D-loop network, several star-shaped phylogenies indicate episodes of rapid demographic growth, but multiple mutations had accumulated on the phylogenies since the expansions, suggesting they occurred well before the LGM.

Minor population bottlenecks can be obscured by the strong signals of an initial major expansion from an equilibrium population (Rogers 1995). The data suggest refugial populations during the LGM were relatively numerous, and recolonizing individuals may have moved north in moderately large numbers. I would therefore not expect to find a post-LGM signal of major expansion. Instead, the most recent major expansion of the Boreal and Pacific NW lineages may have coincided with the end of the Illinoian 'great glaciation' (300 kya – 130 kya) that preceded the Wisconsin glaciation. The Illinoian glaciers extended farther south in the hare range than did many prior glaciations and also compared to the LGM. It is plausible this glacial cycle severely restricted Boreal hare populations, such that post-Illinoian recolonization of the Boreal range generated the observed signal of expansion in the Boreal group. Effects on the

Pacific NW group may have been similar. An Oregon population (OR1) was basal within the Pacific NW D-loop network, suggesting spatial expansion may have occurred north and south from here.

Private Allelic Richness

Private allelic richness (PAR) is a component of genetic diversity, but I consider it separately as an important measure of the genetic uniqueness of populations, for prioritizing conservation actions and for identifying potential source populations of post-LGM range expansion. Glacial refugial populations are often characterized by high PAR. Within the Boreal group I identified several populations with relatively high PAR, supporting earlier results consistent with a hypothesis of multiple refugia within the Boreal, and post-glacial recolonization originating from many fronts. New Brunswick has high PAR and is also basal within the Boreal haplotype network, along with other eastern populations (primarily Maine and Quebec). Eastern populations were probably a major refugial source for post-glacial expansion in the Boreal. This idea is also supported by a weak but significant longitudinal gradient in expected heterozygosity, with higher diversity in the eastern Boreal. Eastern forests expanded across North America five times faster than western forests as glaciers contracted after the LGM, allowing early expansion from eastern refugia (Delcourt and Delcourt 1987; Williams et al. 1993).

Population PAR was very high for Michigan and Saskatchewan. In addition, moderately high PAR estimates for southern Alberta and southeast Alaska were consistent with the fossil evidence and paleogeographic reconstructions of the most recent ice age, which point to a possible refugium for snowshoe hares in eastern

Beringia (central Alaska and western Yukon). Any relict snowshoe hare populations in Beringia were probably small and isolated— by most accounts Beringia was primarily tundra, although small stands of poplar may have persisted (Pielou 1991). Given the earlier described biases of PAR estimates, high PAR for a sampled population should not be considered proof of refugial status, but simply an indicator of potential refugia.

As expected, most populations in the Rockies and Pacific NW groups had high population PAR. The relative stability of these southern refugial zones through past glacial cycles, combined with low gene flow among populations, allowed them to accumulate private alleles over long time periods. Populations within these groups are unique at small spatial scales. Most studies of North American boreal and north temperate species have similarly found high differentiation of populations at southern latitudes compared to the boreal north (Green et al. 1996). This pattern is not surprising, given the long-term stability of southern populations and the fragmented nature of their montane habitats.

Although most boreal populations were characterized by low population PAR, many had a high frequency of alleles private to the Boreal group (PAR_{clus}). In addition, these cluster private alleles made up a much larger proportion of total alleles in Boreal (15%) versus Pacific NW (4%) and Rockies (2%) populations. These findings are congruent with a scenario of multiple source refugia and high gene flow among newly recolonized Boreal populations. Multiple refugial populations could independently accumulate private alleles during glacial periods. High gene flow during post-glacial expansion would then ensure many of these private alleles would be widely distributed

among Boreal populations. Thus, most alleles private to the Boreal group would be shared among populations within the group, consistent with the data. These dynamics promote high genetic diversity in Boreal populations despite repeated continental-scale range contractions and expansions.

Gene Flow

Gene flow was very low among genetic groups and within the Pacific NW and Rockies groups, but moderate to high within the Boreal group. I observed very low levels of admixture for K=3 in STRUCTURE analysis, confirming limited contemporary gene flow among the three genetic groups. However, with mitochondrial data, I found two populations near the border between Boreal and non-Boreal groups that appeared to be points of contact between genetic groups. These populations in northern Washington and northern Montana each had Boreal D-loop haplotypes as well as a small proportion of haplotypes from a non-Boreal group. My classification of haplotypes to groups was based on visual assessment of their positions within the haplotype network. While this approach is somewhat subjective, haplotype groups were well-differentiated by many mutations in the network. My ability to detect contact populations between genetic groups was limited by my sampling scheme. The two contact populations I identified were the only populations I sampled at the border between Boreal and non-Boreal groups. Greater sampling of populations at this border probably would have revealed greater gene flow between adjacent Boreal and non-Boreal genetic groups. Nevertheless, it seems the zones of contact between Boreal and non-Boreal groups are narrow. My inability to detect contact populations between any

two non-Boreal genetic groups may also reflect limited sampling. However, given the very low gene flow among non-Boreal populations, it would probably be difficult to locate a zone of contact between two non-Boreal groups.

In addition to the three major haplotype lineages, mitochondrial analyses identified three sublineages (Washington's Olympic Peninsula, eastern Oregon, and Colorado) within the Rockies and Pacific Northwest populations that diverged from the main lineages over 20,000 years ago. Gene flow within and among these western USA groups is very limited. Except for the southernmost California population I sampled (CA2), populations in these groups did not share mitochondrial haplotypes. As found with other species (Payseur et al. 2004; Sequeira et al. 2005; Brito 2007), the discrepancy between microsatellite and mtDNA markers in delineation of the contact zone between the Boreal and Pacific NW groups may be related to different effective population sizes for mtDNA versus nuclear markers, differences in marker evolutionary rates or modes of inheritance, or marker-related differences in selection associated with introgression of black-tailed jackrabbit mtDNA in Pacific NW snowshoe hare populations.

Conservation and Management Implications

The data revealed three deeply diverged major lineages and three sublineages of snowshoe hares, equivalent to six distinct Evolutionarily Significant Units (ESU's), as defined by Moritz (1994). I found high genetic diversity throughout most of the species range, and high differentiation among populations in the U.S. southern Rockies and Pacific Northwest. The Boreal north is a dynamic source of genetic diversity for

snowshoe hares, in large part due to high levels of gene flow among populations and a large geographic coverage that has allowed persistence of multiple refugial populations during glacial periods. In contrast to the low diversity characteristic of successive founding from an individual source population, the high diversity of the Boreal appears to derive from expansion and secondary admixture in large numbers primarily from southern and eastern refugia but also potentially from Beringia. In terms of remaining genetic diversity, the genetic consequences of global warming for snowshoe hares, whereby southern populations are lost but northern ones persist, may not be large because of this genetic pattern.

Although results indicated genetic variation would be retained under a future global warming scenario, it would represent a reduced subset of current genetic variation in snowshoe hares. A large proportion of the species' PAR is found in the southern Rockies and Pacific NW, where long-term relative stability and isolation have allowed populations to accumulate mutations and unique genetic constitutions. Based on estimates of population PAR, loss of a population in the Rockies or Pacific NW would have 50% greater impact on total allelic richness of the species than would loss of a population in the Boreal group. In contrast to the highly connected Boreal populations, loss of a population in the Rockies and Pacific NW could mean complete loss of many alleles unique to that population.

If conserving overall species genetic diversity is the goal, my findings support Hampe and Petit's (2005) call for prioritizing conservation of rear edge populations. For snowshoe hares and many other boreal species these rear edge populations are already

losing genetic diversity due to natural (and possibly anthropogenic) habitat fragmentation and genetic drift (Paetkau et al. 1998, Kyle and Strobeck 2002, Wisely et al. 2004, Schwartz et al. 2007). The topographic complexity of these southern range extensions means species may be able to track warming-related habitat shifts to higher altitudes (Hewitt 2000; Guralnick 2007), but at a cost of further fragmenting populations. Because they are highly differentiated at a relatively small spatial scale, it would be difficult to find genetically similar sources to supplement declining populations in the Rockies and Pacific NW. Furthermore, past attempts at translocating and supplementing snowshoe hares have been largely unsuccessful (Murray 2003), with the notable exception of their introduction to Newfoundland.

Loss of southern hare habitats due to global warming does not necessarily mean loss of snowshoe hare populations and their genes. It is possible snowshoe hares may adapt to habitat changes associated with global warming; additionally, populations currently persisting in warmer environments may expand their range. In addition, the data suggest the border region between British Columbia and Alberta and Washington, Idaho, and Montana may be an important contact zone for several snowshoe hare ESU's. Gene flow is still limited in this region, but for genes with a strong selective advantage (selection coefficient > 5%) immigration rates as low as one individual per 10 generations may be sufficient to maintain adaptive connectivity across a species range (Rieseberg and Burke 2001; Lowe and Allendorf 2010). Conservation of these points of contact may facilitate adaptively beneficial gene flow among genetic groups.

Neutral genetic variation, as measured in this study, is often used as a surrogate for adaptive genetic variation, which is the real metric of interest when evaluating adaptive potential and for ranking populations by long-term conservation value. This study provides important insights on how loss of southern hare populations may affect overall species genetic diversity. However, a more targeted question that should be addressed is, "How much would loss of southern hare populations impact the species' ability to adapt to global warming?" Such studies would require evaluation of quantitative genetic trait variations directly linked to traits with adaptive value for warmer environments.

In the face of certain climate change with uncertain impacts it is difficult to predict how species conservation efforts can best be prioritized to maximize long-term persistence. Predator-prey dynamics may change as the species composition of ecosystems are reshuffled, routes of population connectivity and gene flow may shift, and entire ecosystems may temporarily imbalance as new ecological thresholds are breached (Running and Mills 2009). We cannot anticipate the unforeseen, but we can use our understanding of the present to heed the advice of geneticist Otto Frankel (1974) that "...at this point of decision-making it may be our evolutionary responsibility to keep evolutionary options open so far as we can..." Given the important role of snowshoe hares in boreal and montane forest ecosystems, maintaining the genetic diversity and future adaptive potential of snowshoe is a conservation objective with widespread impact.

LITERATURE CITED

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." Journal of Molecular Biology **215**: 403-410.

Alves, P. C., N. Ferrand, et al. (2003). "Ancient introgression of Lepus timidus mtDNA into L. granatensis and L. europaeus in the Iberian Peninsula." Molecular Phylogenetics and Evolution **27**: 70-80.

Arbogast, B. S. (1999). "Mitochondrial DNA phylogeography of the new world flying squirrels (Glaucomys): Implications for pleistocene biogeography." Journal of Mammalogy **80**(1): 142-155.

Arbogast, B. S. and G. J. Kenagy (2001). "Comparative phylogeography as an integrative approach to historical biogeography." Journal of Biogeography **28**(7): 819-825.

Benjamini, Y. and Y. Hochberg (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing." Journal of the Royal Statistical Society, Series B **57**(1): 289-300.

Brito, P. H. (2007). "Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of tawny owls (Strix aluco)." Molecular Ecology **16**: 3423-3437.

Brooks, M. (1955). "An Isolated Population of the Virginia Varying Hare " The Journal of Wildlife Management **19**(1): 54-61.

Bulmer, M. G. (1974). "A statistical analysis of the 10-year cycle in Canada." Journal of Animal Ecology **43**(3): 701-718.

Burton, C., C. J. Krebs, et al. (2002). "Population genetic structure of the cyclic snowshoe hare (Lepus americanus) in southwestern Yukon, Canada." Molecular Ecology **11**: 1689-1701.

Conroy, C. J. and J. A. Cook (2000). "Phylogeography of a post-glacialcolonizer: Microtus longicaudus (Rodentia : Muridae)." Molecular Ecology **9**(2): 165-175.

Crowley, T. J. (1990). "Are there any satisfactory geologic analogs for a future greenhouse warming?" Journal of Climate **3**: 1282-1292.

Dalen, L., V. Nystrom, et al. (2007). "Ancient DNA reveals lack of postglacial habitat tracking in the Arctic fox." Proceedings of the National Academy of Sciences **104**: 6726-6729.

Dalquest, W. W. (1942). "Geographic vairation in northwestern snowshoe hares." Journal of Mammalogy: 166 - 183.

Delcourt, P. A. and H. Delcourt (1987). Long term forest dynamics: a case study of Late-Quaternary forests in eastern North America. New York, Springer-Verlag.

Dupanloup, I., S. Schneider, et al. (2002). "A simulated annealing approach to define the genetic structure of populations." Molecular Ecology **11**: 2571-2581.

Eckert, C. G., K. E. Samis & S. C. Lougheed (2008). “Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond.” Molecular Ecology **17**: 1170.

Excoffier, L. (2004). "Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model." Molecular Ecology **13**(4): 853-864.

Excoffier, L., G. Laval, et al. (2005). "Arlequin ver. 3.0: An integrated software package for population genetics data analysis." Evolutionary Bioinformatics Online **1**: 47-50.

FAUNMAP Working Group (1994). "FAUNMAP: a database documenting late Quaternary distributions of mammal species in the United States." Illinois State Museum Scientific Papers **25**(1-2): 1-690.

Finerty, J. P. (1980). The population ecology of cycles in small mammals. New Haven, CT, Yale University Press.

Fisher, R. A. (1930). The genetical theory of natural selection. Oxford, U.K., Clarendon Press.

Frankel, O. H. (1974). "Genetic conservation: our evolutionary responsibility." Genetics **48**: 53-65.

Fulton, F. J. (1991). "A conceptual model for growth and decay of the Cordilleran ice sheet." Geographie Physique Et Quaternaire **45**(3): 281-286.

Green, D. M., T. F. Sharbel, et al. (1996). "Postglacial range fluctuation, genetic subdivision and speciation in the western North American spotted frog complex, Rana pretiosa." Evolution **50**(1): 374-390.

Guralnick, R. (2007). "Differential effects of past climate warming on mountain and flatland species distributions: a multispecies North American mammal assessment." Global Ecology and Biogeography **16**(1): 14-23.

Hampe, A. and R. J. Petit (2005). "Conserving biodiversity under climate change: the rear edge matters." Ecology Letters **8**(5): 461-467.

Hewitt, G. (2000). "The genetic legacy of the Quaternary ice ages." Nature **405**(6789): 907-913.

Hewitt, G. M. (1996). "Some genetic consequences of ice ages, and their role in divergence and speciation." Biological Journal of the Linnean Society **58**(3): 247-276.

Hill, W. G. and J. Rabash (1986). "Models of long term artificial selection in finite populations." Genetical Research **48**: 41-50.

Hodges, K. E. (2000). The Ecology of Snowshoe Hares in Northern Boreal Forests. Ecology and Conservation of Lynx in the United States. L. F. Ruggiero, K. B. Aubry, S. W. Buskirket al. Boulder, CO, USA, University Press of Colorado**:** 117-161.

Hodges, K. E., L. S. Mills, et al. (2009). "Snowshoe hare distribution and abundance in Yellowstone National Park." Journal of Mammalogy **90:** 870-878.

IPCC (2001). Climate change 2001: the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. New York, USA, Cambridge University Press.

IPCC (2007). Climate Change 2007: Impacts, Adaptation, and Vulnerability. Cambridge, UK, Cambridge University Press.

Iverson, L. R. and A. M. Prasad (1998). "Predicting abundance of 80 tree species following climate change in the eastern United States." Ecological Monographs **68**(4): 465-485.

James, J. W. (1971). "The founder effect and response to artificial selection." Genetical Research **12**: 249-266.

Kalinowski, S. T. (2004). "Counting alleles with rarefaction: Private alleles and hierarchical sampling designs." Conservation Genetics **5**: 539-543.

Kalinowski, S. T. (2005). "HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness." Molecular Ecology Notes **5**(1): 187-189.

Keith, L. B. (1963). Wildlife's ten year cycle. Madison, WI, University of Wisconsin Press.

Korstanje, R., G. F. Gillissen, et al. (2003). "Mapping of rabbit microsatellite markers using chromosome-specific libraries." Journal of Heredity **94**: 161-169.

Kyle, C. J. and C. Strobeck (2002). "Connectivity of peripheral and core populations of north american wolverines." Journal of Mammalogy **83**(4): 1141-1150.

Lawton, J. H. (1993). "Range, population abundance and conservation." Trends in Ecology & Evolution **8**(11): 409-413.

Lessa, E. P., J. A. Cook, et al. (2003). "Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary." Proceedings of the National Academy of Sciences **100**(18): 10331-10334.

Lowe, W. H. and F. W. Allendorf (2010). "What can genetics tell us about population connectivity?" Molecular Ecology **19**(15): 3038-3051.

Librado, P. and J. Rozas (2009). "DnaSP v5: A software for comprehensive analysis of DNA polymorphism data." Bioinformatics **25**: 1451-1452.

Matthee, C. A., B. J. van Vuuren, et al. (2004). "A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene." Systematic Biology **53**(3): 433-447.

Melo-Ferreira, J., P. Boursot, et al. (2007). "The rise and fall of the mountain hare (Lepus timidus) during Pleistocene glaciations: expansion and retreat with hybridization in the Iberian Peninsula." Molecular Ecology **16**: 605-618.

Melo-Ferreira, J., P. Boursot, et al. (2005). "Invasion from the cold past: extensive introgression of mountain hare (Lepus timidus) mitochondrial DNA into three other hare species in northern Iberia." Molecular Ecology **14**: 2459-2464.

Mills, L. S., and F. W. Allendorf. (1996). "The one-migrant-per-generation rule in conservation and management." Conservation Biology **10:** 1509-1518.

Mills, L. S., P. C. Griffin, et al. (2005). "Pellet count indices compared to mark-recapture estimates for evaluating snowshoe hare density." Journal of Wildlife Management **69**(3): 1053-1062.

Moritz, C. (1994). "Defining evolutionarily significant units for conservation." Trends in Ecology & Evolution **9**: 373-375.

Mougel, F., J. C. Mounolou, et al. (1997). "Nine polymorphic microsatellite loci in the rabbit, *Oryctolagus cuniculus*." Animal Genetics **28**: 58-71.

Murray, D. L. (2003). Snowshoe hares and other hares. Wild Mammals of North America, Vol. II. G. A. Feldhamer and B. C. Thompson. Baltimore, MD, Johns Hopkins University Press.

Nei, M. (1972). "Genetic distance between populations." The American Naturalist **106**(949): 283-292.

Nei, M. (1978). "Estimation of average heterozygosity and genetic distance from a small number of individuals." Genetics **89**: 583-590.

Nei, M., T. Maruyama, et al. (1975). “The bottleneck effect and genetic variability in populations.” Evolution **29:** 1-10.

Paetkau, D., L. P. Waits, et al. (1998). "Variation in genetic diversity across the range of North American brown bears." Conservation Biology **12**(2): 418-429.

Payseur, B. A., J. G. Krenz, et al. (2004). "Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice." Evolution **58**: 2064-2078.

Petit, R. J., I. Aguinagalde, et al. (2003). "Glacial refugia: Hotspots but not melting pots of genetic diversity." Science **300**(5625): 1563-1565.

Pielou, E. C. (1991). After the ice age: the return of life to glaciated North America. Chicago, University of Chicago Press.

Pierpaoli, M., F. Riga, et al. (1999). "Species distinction and evolutionary relationships of the Italian hare (Lepus corsicanus) as described by mitochondrial DNA sequencing." Molecular Ecology **8**(11): 1805-1817.

Pritchard, J. K., M. Stephens, et al. (2000). "Inference of population structure using multilocus genotype data." Genetics **155**: 945-959.

Pujol, B. and J. R. Pannell (2008). "Reduced responses to selection after species range expansion." Science **321**(5885): 96-96.

Ray, N., M. Currat, et al. (2003). "Intra-deme molecular diversity in spatially expanding populations." Molecular Biology and Evolution **20**(1): 76-86.

Raymond, M. and F. Rousset (1995). "An exact test for population differentiation." Evolution **49**: 1283-1286.

Rico, C., I. Rico, et al. (1994). "Four polymorphic microsatellite loci for the European wild rabbit, *Oryctolagus cuniculus*." Animal Genetics **25**: 367.

Rieseberg, L. H. and J. M. Burke (2001). "A genic view of species integration." Journal of Evolutionary Biology **14**: 883-886.

Rogers, A. R. (1995). "Genetic evidence for a Pleistocene population explosion." Evolution **49**(4): 608-615.

Rousset, F. (2008). "Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux." Molecular Ecology Resources **8**: 103-106.

Running, S. W. and L. S. Mills (2009). Managing for resilience: how natural ecosystems can cope with climate change, Resources for the Future.

Schwartz, M. K., K. B. Aubry, et al. (2007). "Inferring geographic isolation of wolverines in California using historical DNA." Journal of Wildlife Management **71**(7): 2170-2179.

Schwartz, M. K., K. L. Pilgrim, et al. (2007). "DNA markers for identifying individual snowshoe hares using field-collected pellets." Northwest Science **81**(4): 316-322.

Sequeira, F., J. Alexandrino, et al. (2005). "Genetic exchange across a hybrid zone within the Iberian endemic golden-striped salamander, Chioglossa lusitanica." Molecular Ecology **14**: 245-254.

Shafer, A. B. A., C. I. Cullingham, et al. (2010). “Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America.” Molecular Ecology **19**: 4589-4621.

Soltis, D. E., M. A. Gitzendanner, et al. (1997). "Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America." Plant Systematics and Evolution **206**(1-4): 353-373.

Stamatakis, A. (2006). "RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models." Bioinformatics **22**(21): 2688-2690.

Stewart, J. R. and A. M. Lister (2001). "Cryptic northern refugia and the origins of the modern biota." Trends in Ecology & Evolution 16(11): 608-613.

Strimmer, K. (2008). "fdrtool: a versatile R package for estimating local and tail area-based false discovery rates." Bioinformatics.

Surridge, A. K., D. J. Bell, et al. (1997). "Polymorphic microsatellite loci in the European rabbit *(Oryctolagus cuniculus)* are also amplified in other lagomorph species." Animal Genetics **28**: 302-305.

Swindell, W. R. and J. L. Bouzat (2005). "Modeling the adaptive potential of isolated populations: Experimental simulations using Drosophila." Evolution **59**(10): 2159-2169.

Tamura, K., J. Dudley, et al. (2007). "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0." Molecular Biology and Evolution **24**: 1596-1599.

Van Oosterhout, C., W. P. Hutchinson, et al. (2004). "MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data." Molecular Ecology Notes **4**: 535-538.

Vucetich, J. A. and T. A. Waite (2003). "Spatial patterns of demography and genetic processes across the species' range: Null hypotheses for landscape conservation genetics." Conservation Genetics **4**: 639-645.

Weir, B. S. and C. C. Cockerham (1984). "Estimating F-statistics for the analysis population structure." Evolution **38**(6): 1358-1370.

Williams, M. A. J., D. L. Dunkerley, et al. (1993). Quaternary Environments. New York, Edward Arnold.

Wisely, S. M., S. W. Buskirk, et al. (2004). "Genetic diversity and structure of the fisher (martes pennanti) in a peninsular and peripheral metapopulation." Journal of Mammalogy **85**(4): 640-648.

Wright, S. (1931). "Evolution in Mendelian populations." Genetics **16**: 97-259.

Wu, C. H., J. P. Wu, et al. (2005). "Molecular phylogenetics and biogeography of Lepus in Eastern Asia based on mitochondrial DNA sequences." Molecular Phylogenetics and Evolution **37**(1): 45-61.

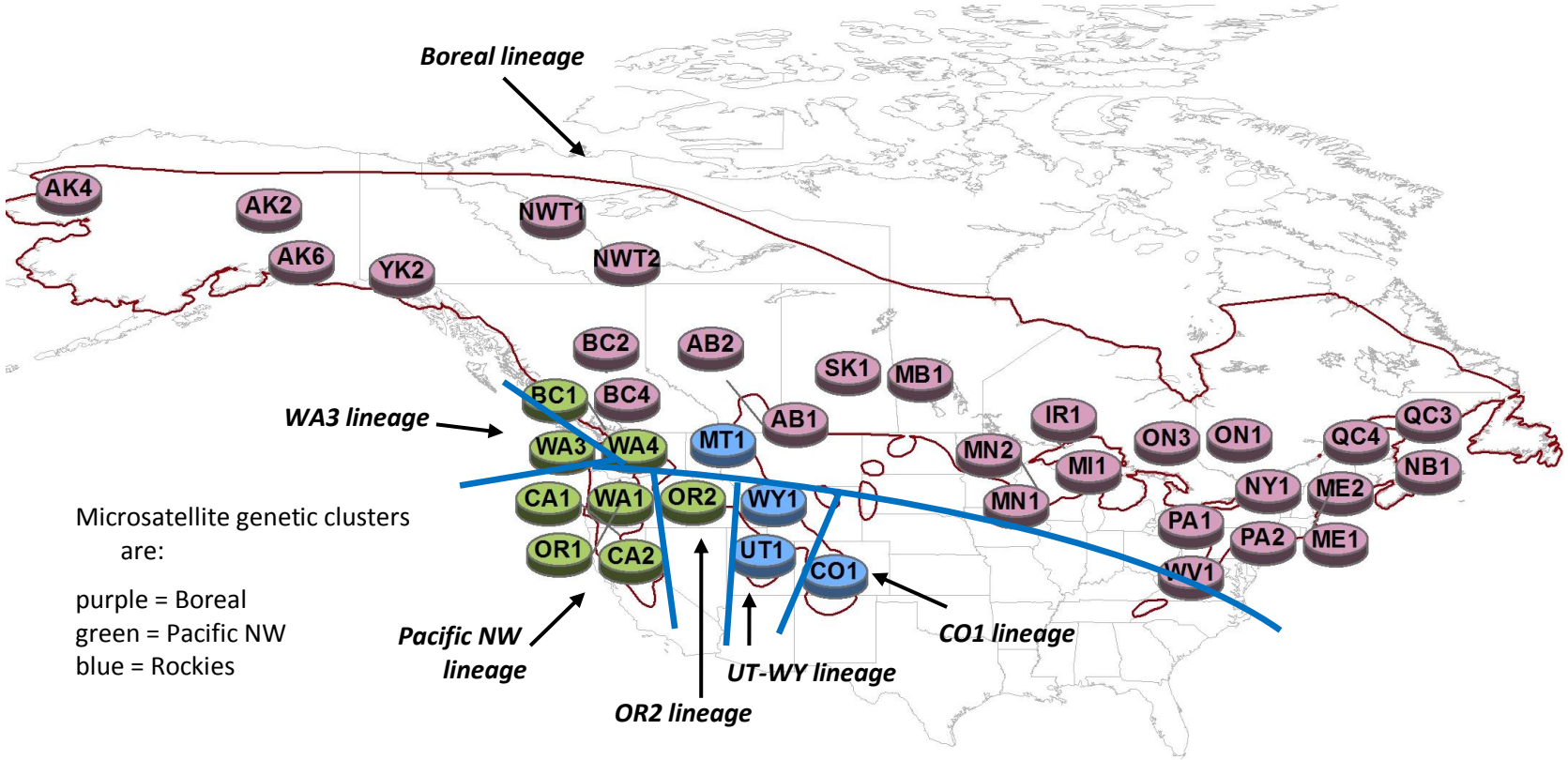


Figure 2.1
Sampling locations and geographic distribution of major genetic groups. The range of the snowshoe hare, as determined by Erxleben (1777), is outlined in red. Population names are indicated in circles. Color of circle indicates membership in one of three microsatellite genetic clusters, as defined by STRUCTURE. Thick blue lines separate six haplotype lineages (names italicized) based on SAMOVA analysis of D-loop haplotypes.

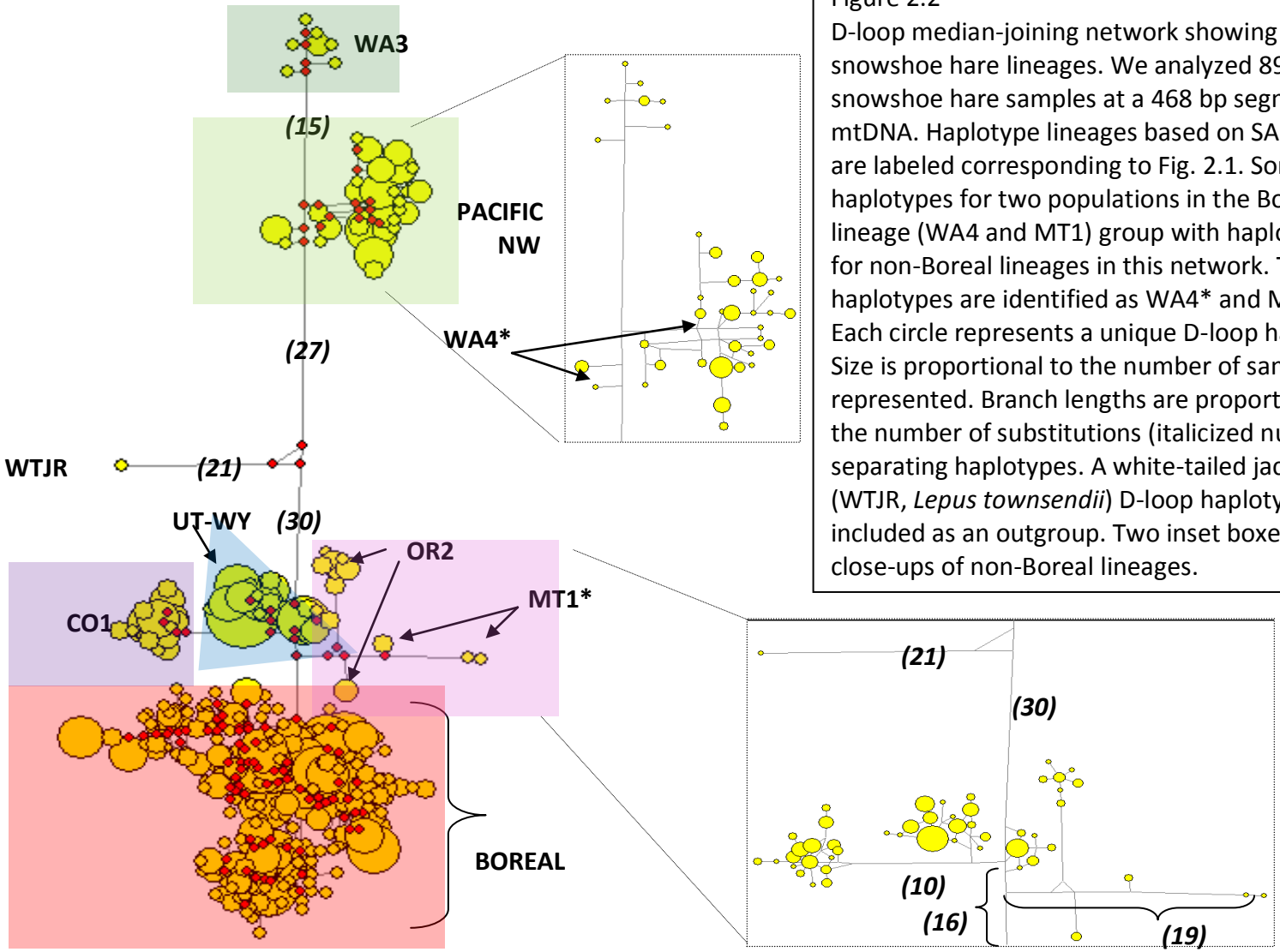


Figure 2.2
D-loop median-joining network showing six snowshoe hare lineages. We analyzed 893 snowshoe hare samples at a 468 bp segment of mtDNA. Haplotype lineages based on SAMOVA are labeled corresponding to Fig. 2.1. Some haplotypes for two populations in the Boreal lineage (WA4 and MT1) group with haplotypes for non-Boreal lineages in this network. These haplotypes are identified as WA4* and MT1*. Each circle represents a unique D-loop haplotype. Size is proportional to the number of samples represented. Branch lengths are proportional to the number of substitutions (italicized number) separating haplotypes. A white-tailed jackrabbit (WTJR, *Lepus townsendii*) D-loop haplotype is included as an outgroup. Two inset boxes show close-ups of non-Boreal lineages.

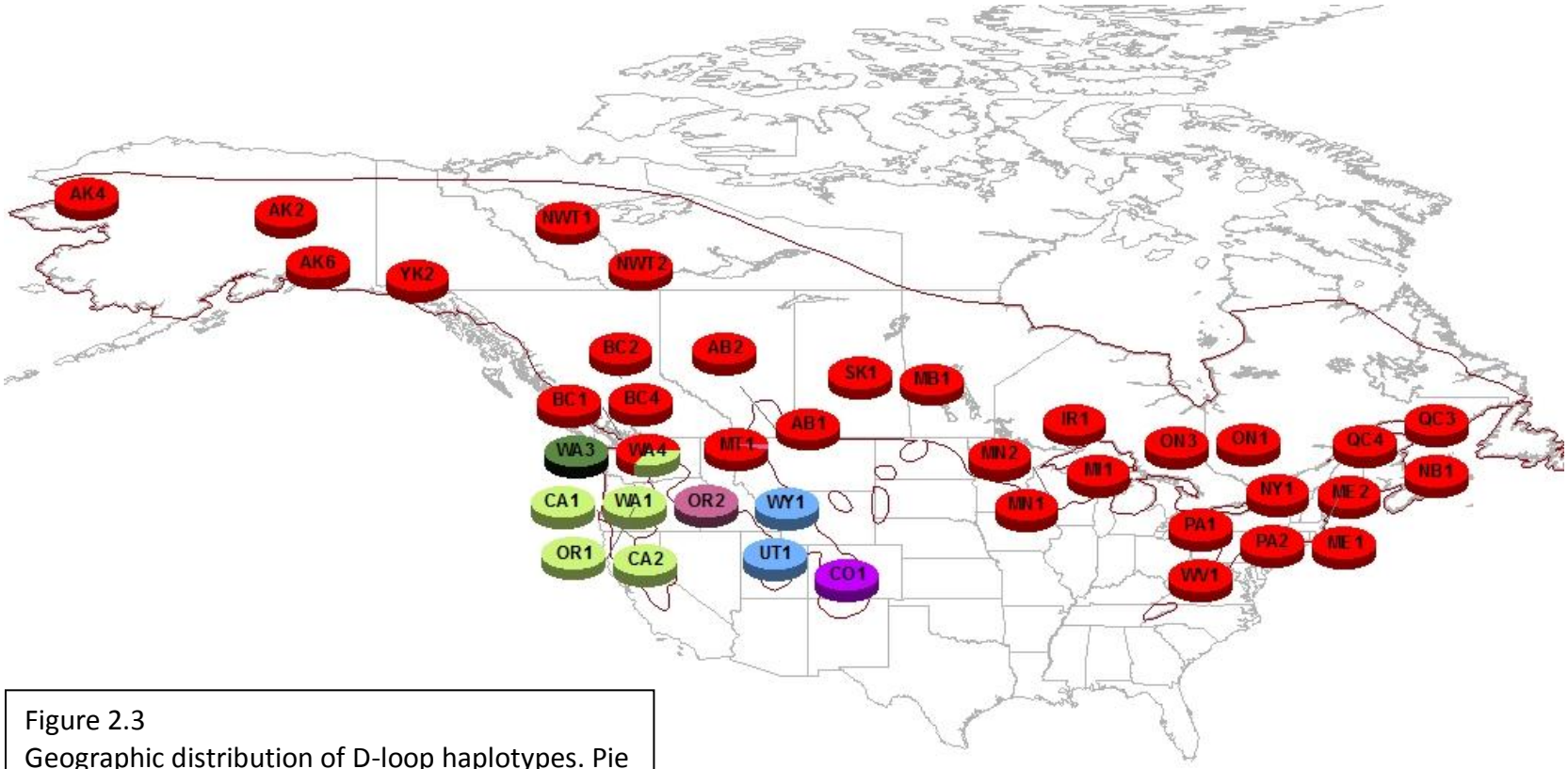


Figure 2.3
Geographic distribution of D-loop haplotypes. Pie charts indicate the proportion of each population represented by different haplotype groups, corresponding to overlay colors in Fig. 2.2.

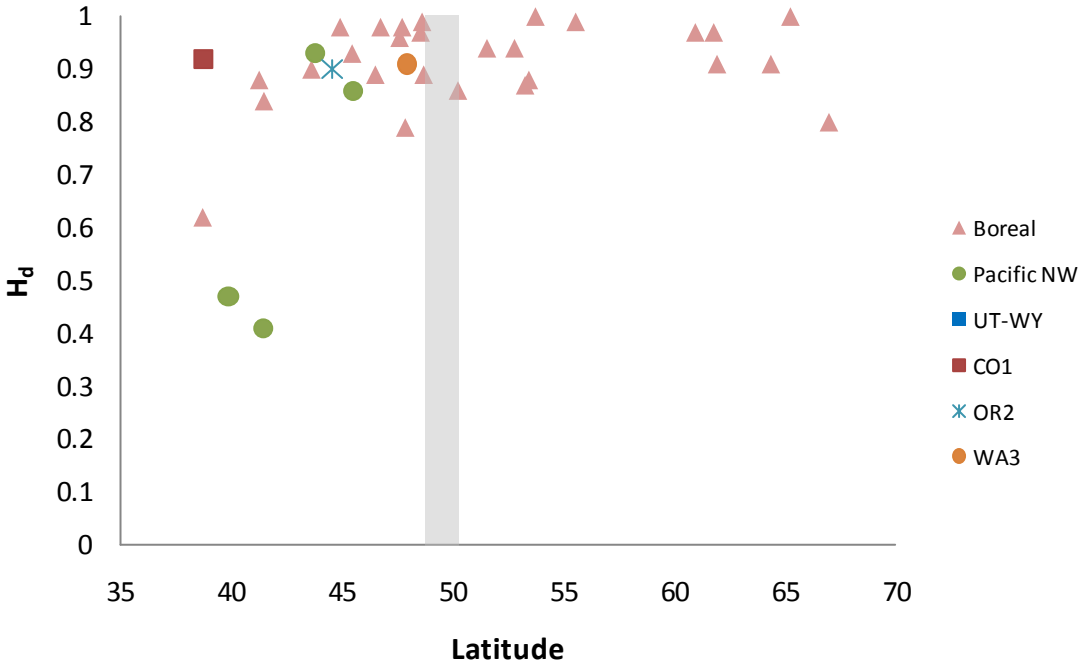
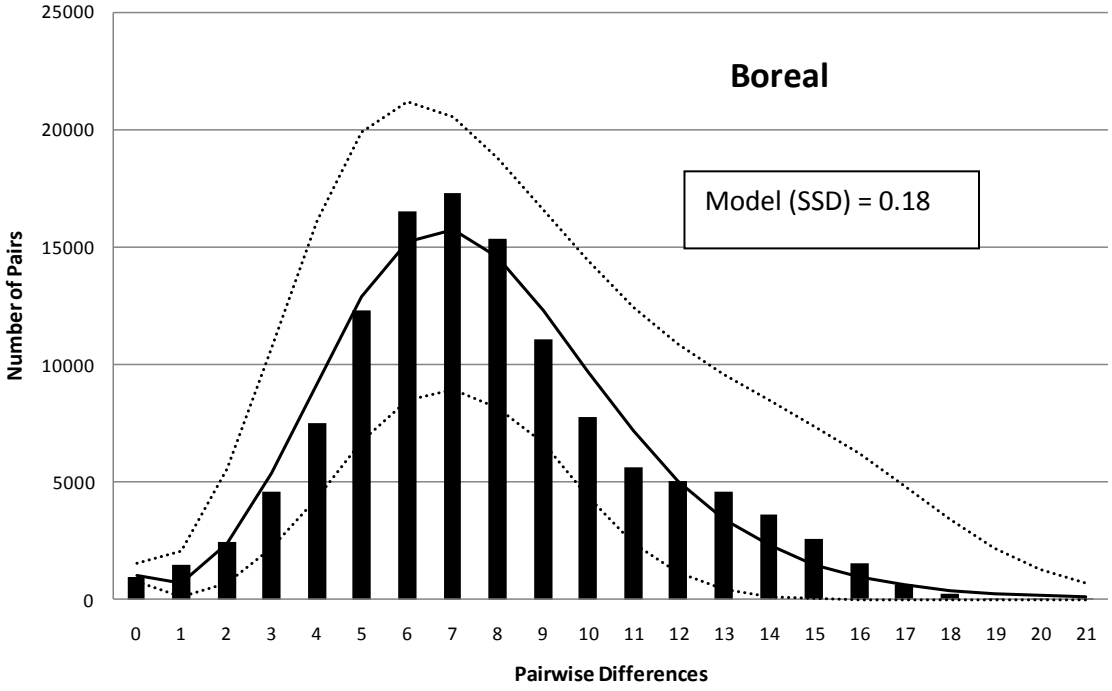


Figure 2.4
For each sampled population, haplotype diversity (H_d) plotted against latitude. Populations are color-coded by haplotype lineage. The significant quadratic relationship for all lineages combined ($r^2 = 0.35$, $p = 0.0003$) is largely driven by the low haplotype diversities ($H_d < 0.70$) of three southern range peninsular populations (two in California, one in West Virginia). The southernmost extent of the LGM is identified by the gray vertical bar.



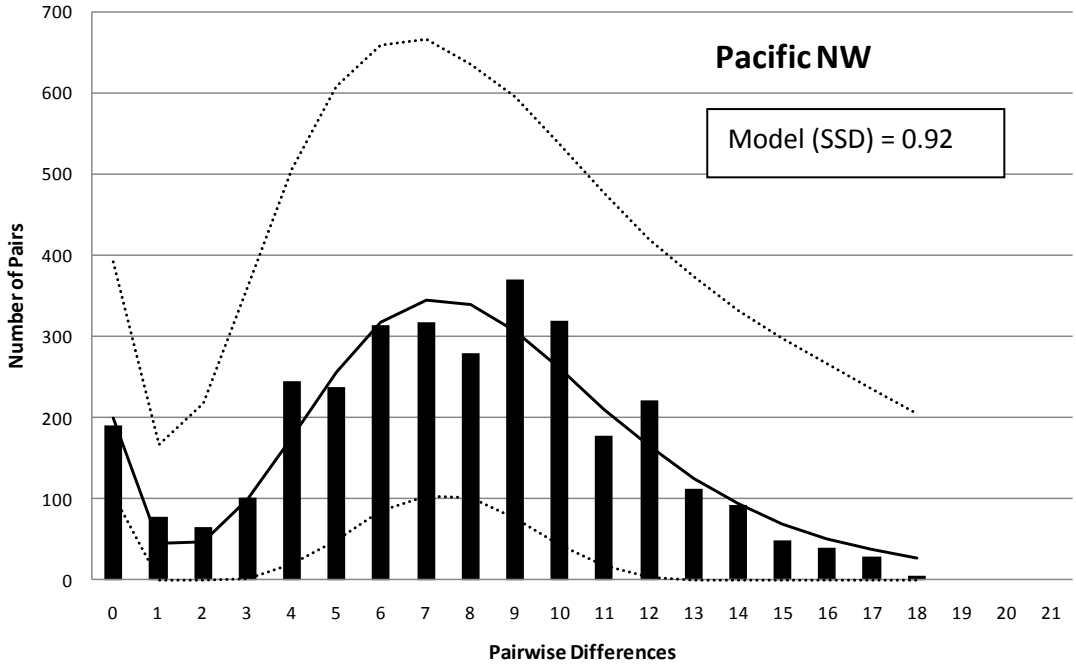


Figure 2.5
For Boreal and Pacific NW lineages, D-loop mismatch distributions (solid bars) plotted against expected distributions under a model of spatial expansion. Expected distribution is drawn as a solid line with 95% confidence intervals (dotted lines). The null for Model(SSD) is an expansion model; therefore, a non-significant Model(SSD) indicates the expansion model cannot be statistically rejected.

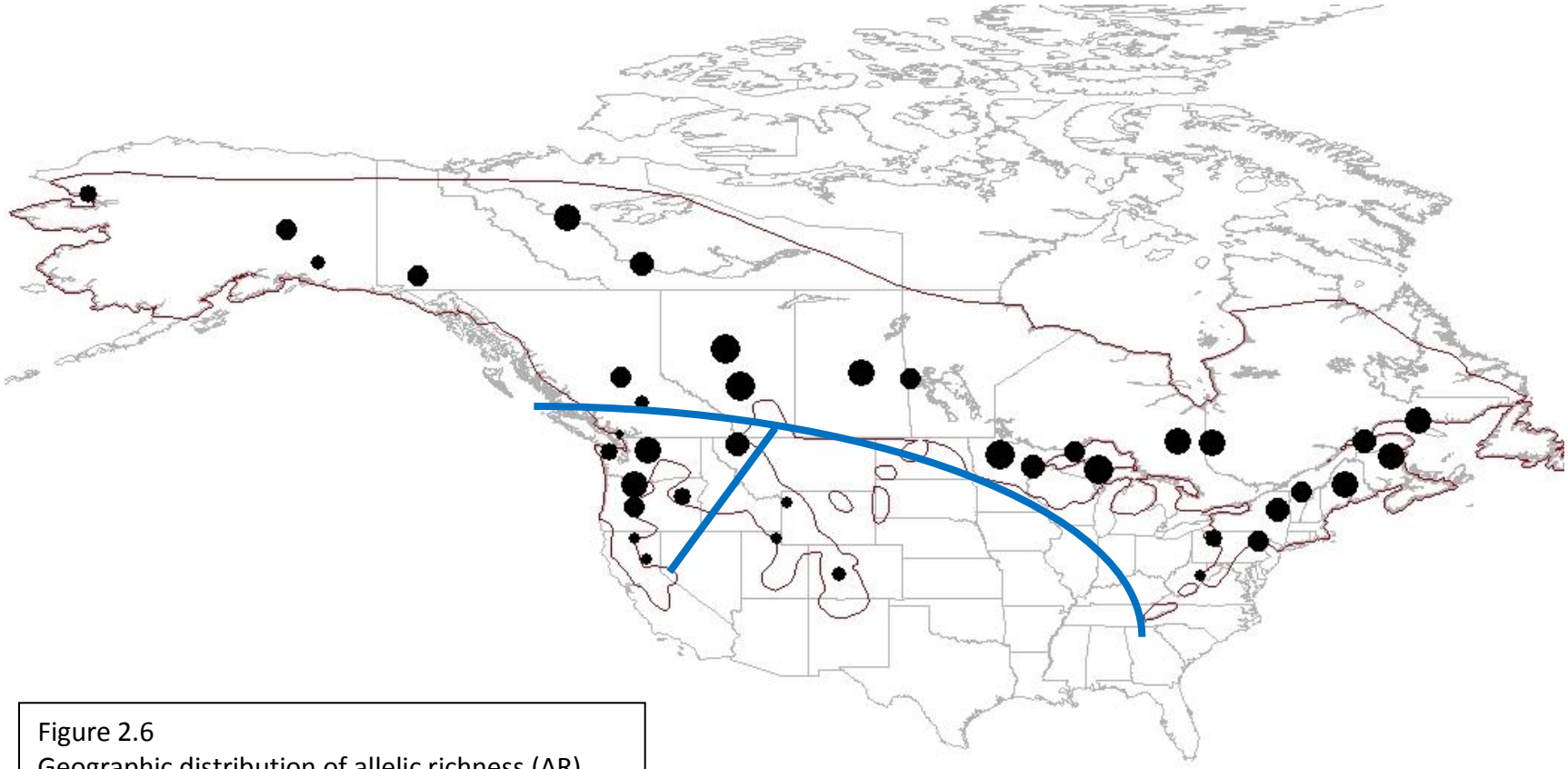


Figure 2.6
Geographic distribution of allelic richness (AR) scaled to a sample size of seven individuals. AR estimates for sampled populations range from 3.63 to 6.20. Circles are proportional to AR divided into 10 evenly spaced categories. Thick blue lines separate three microsatellite genetic clusters identified by STRUCTURE.

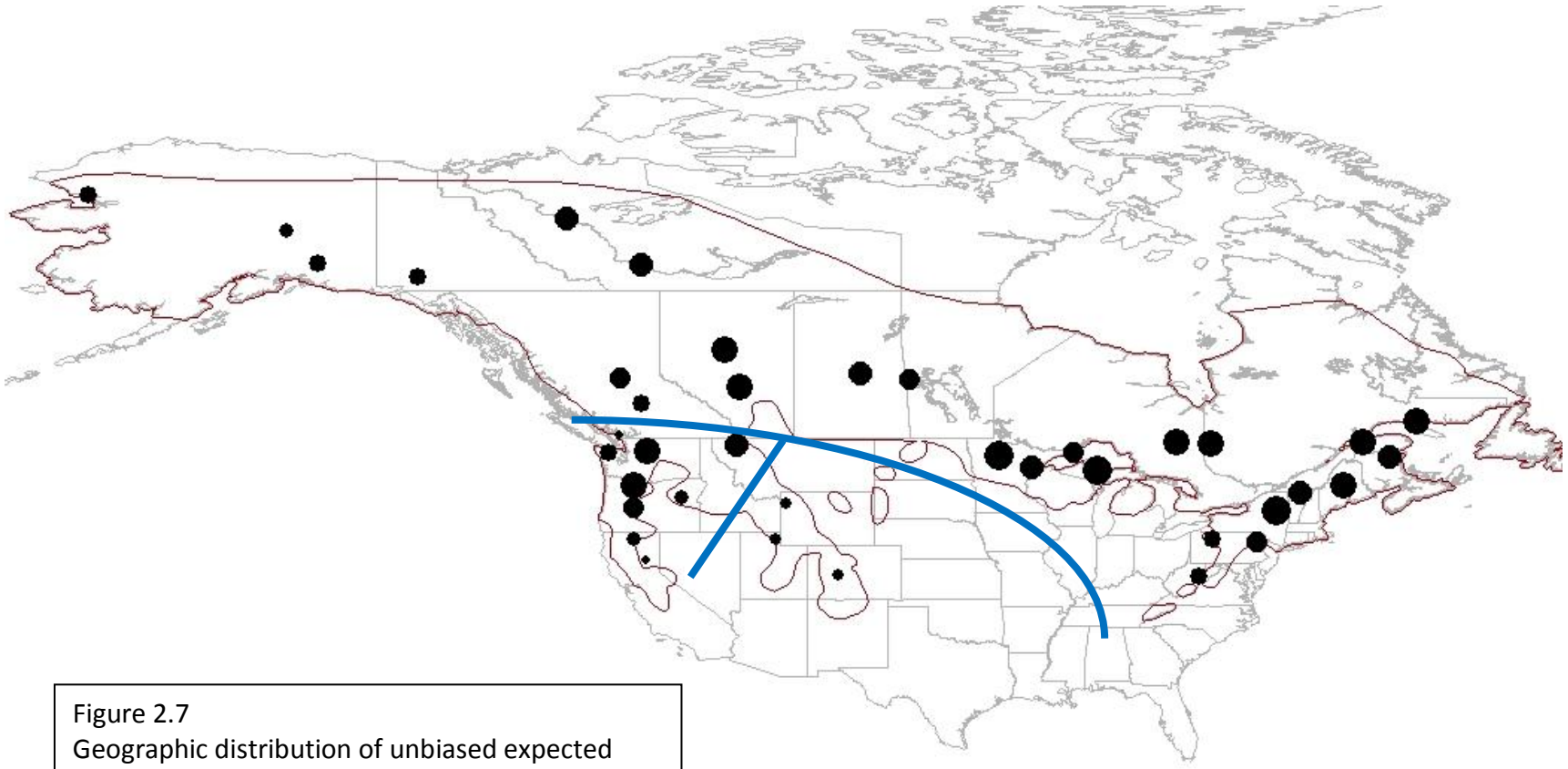


Figure 2.7
Geographic distribution of unbiased expected heterozygosities (H_e). Heterozygosity estimates for sampled populations range from 0.51 to 0.80. Circles are proportional to H_e divided into 10 evenly spaced categories. Thick blue lines separate three microsatellite genetic clusters identified by STRUCTURE.

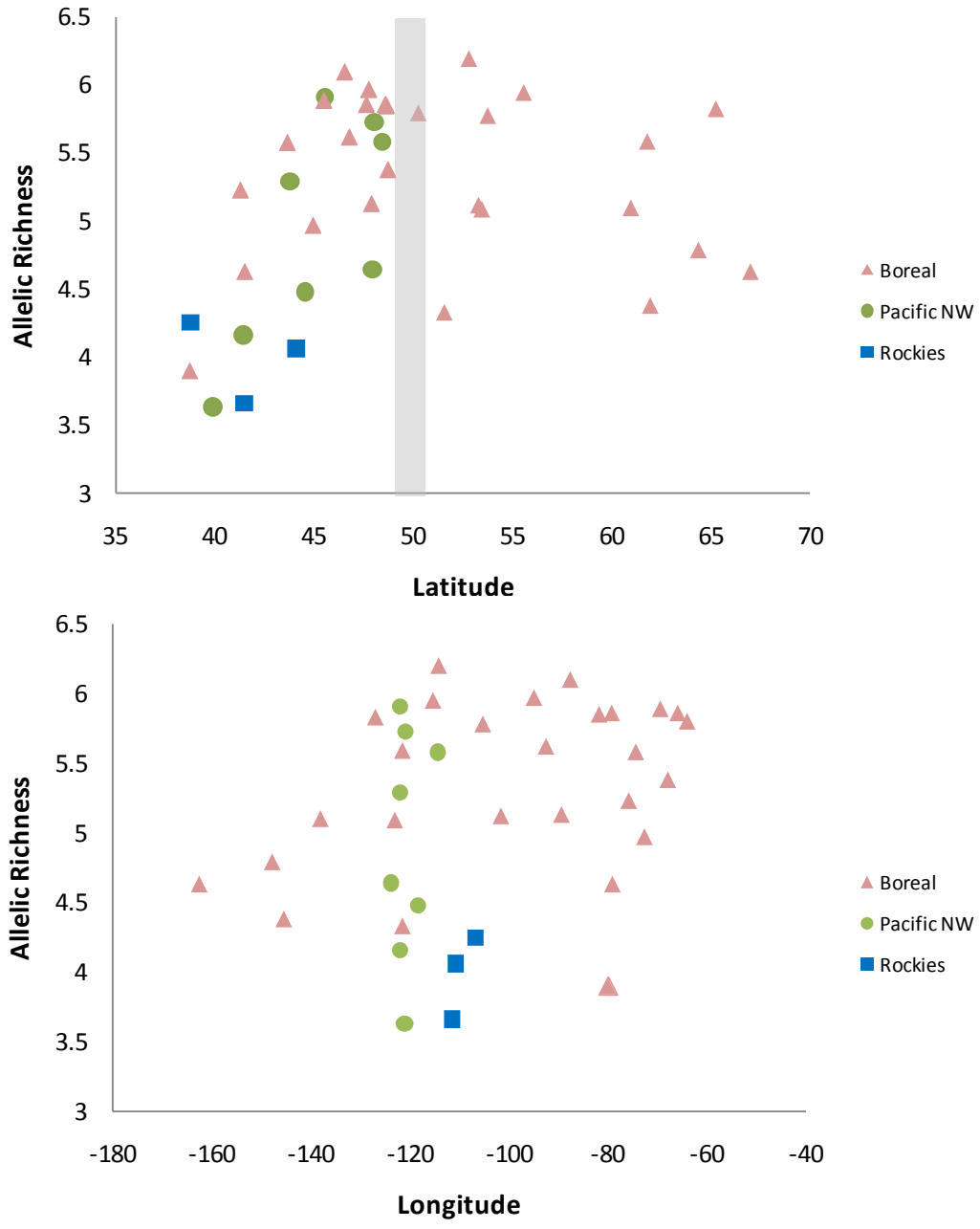


Figure 2.8
For each sampled population, allelic richness (AR) plotted against latitude (top) and longitude (bottom). Populations are color-coded by genetic cluster, as identified by STRUCTURE and SAMOVA. Allelic richness is significantly correlated with latitude for Pacific NW ($r = 0.42$, $p = 0.4$). With all lineages combined, allelic richness has a significant quadratic relationship with latitude ($r^2 = 0.37$, $p = 0.0001$). The southernmost extent of the LGM is marked by the gray vertical bar.

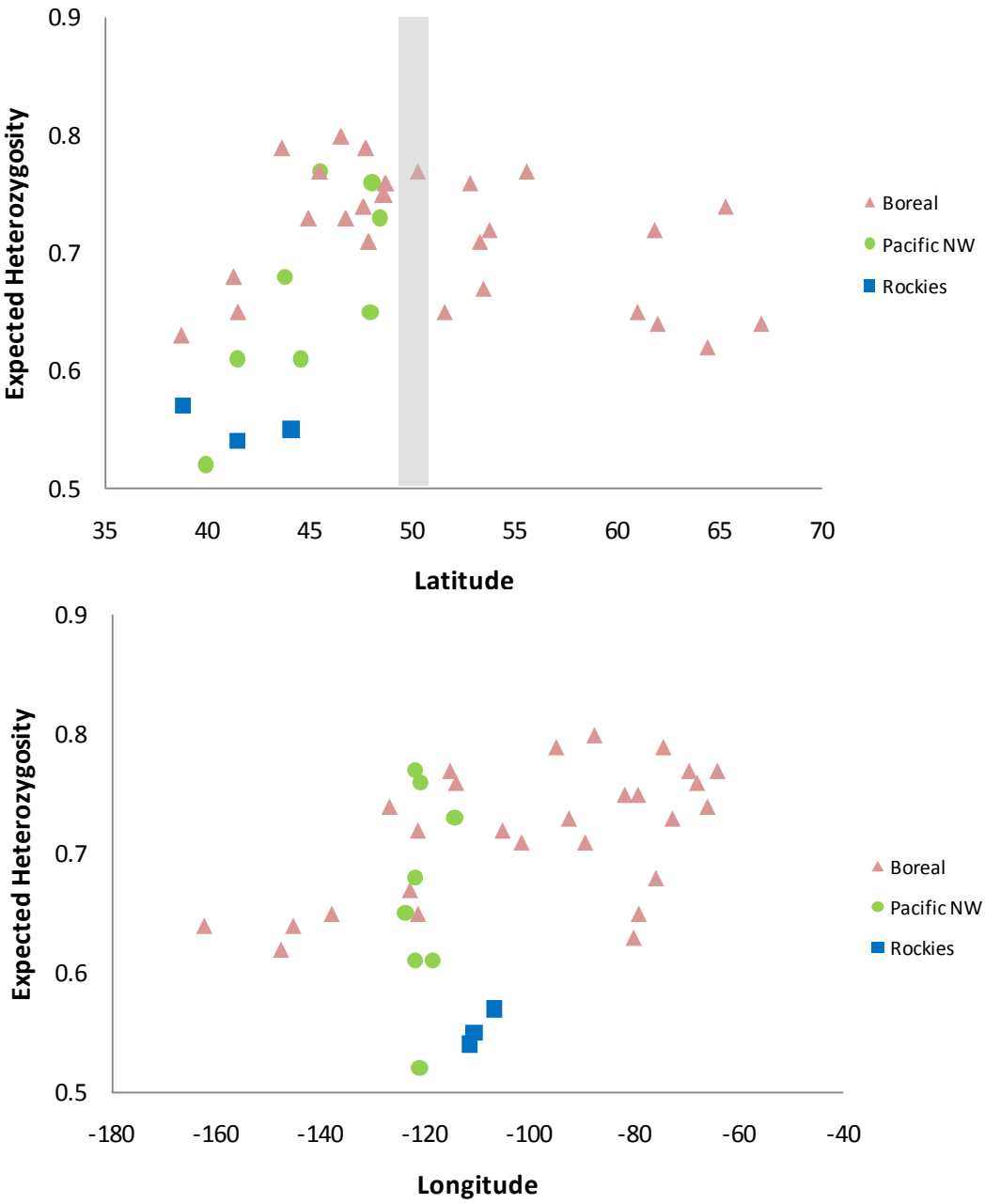


Figure 2.9
For each sampled population, unbiased heterozygosity (H_e) plotted against latitude (top) and longitude (bottom). Populations are color-coded by STRUCTURE genetic cluster. Heterozygosity is significantly correlated with latitude for Pacific NW ($r = 0.77$, $p = 0.03$). With all lineages combined, heterozygosity has a significant quadratic relationship with latitude ($r = 0.35$, $p = .0003$). The southernmost extent of the LGM is marked by the gray vertical bar. Heterozygosity is correlated with longitude for the Boreal cluster ($r = 0.56$, $p = 0.003$).

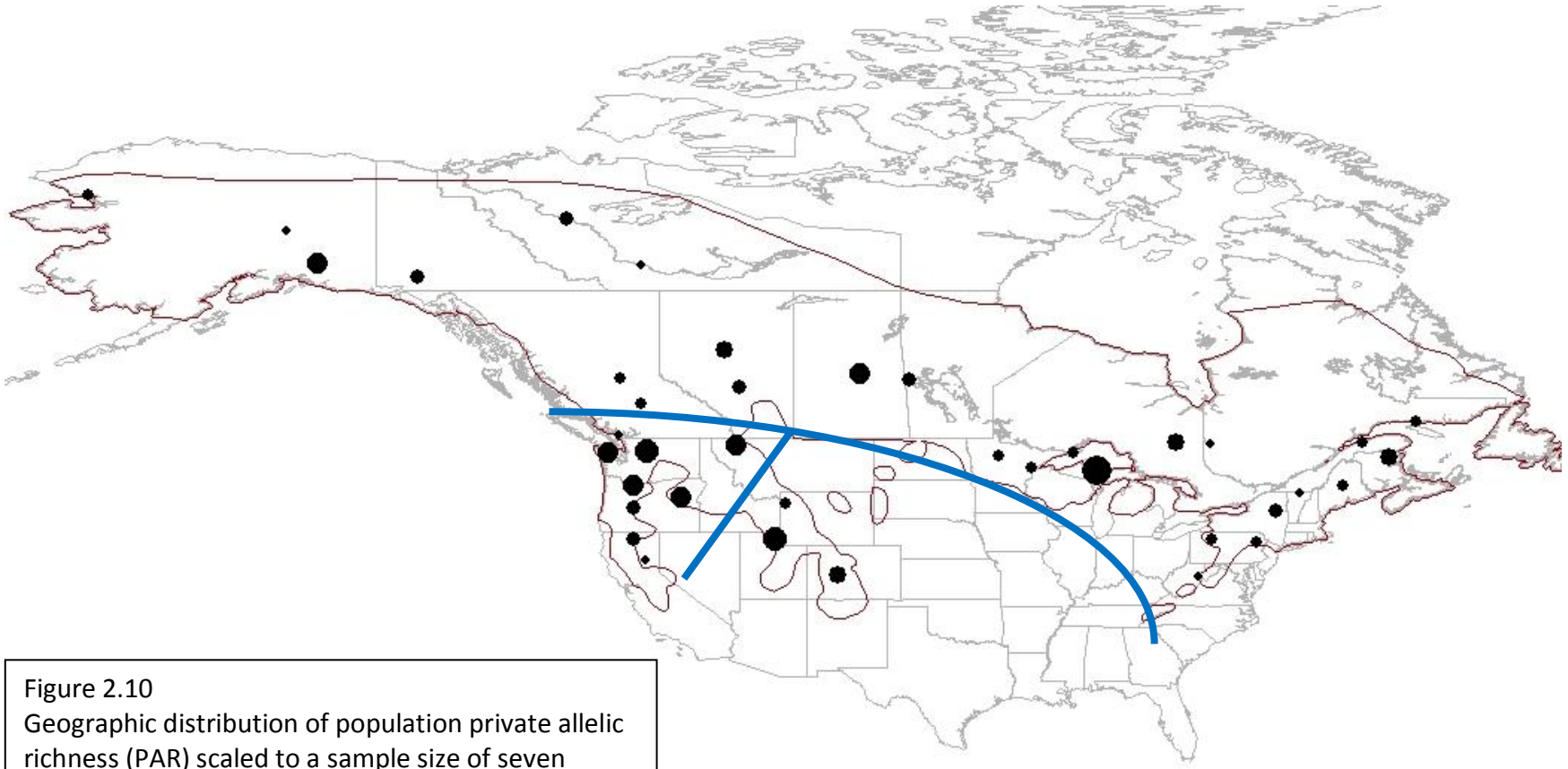


Figure 2.10
Geographic distribution of population private allelic richness (PAR) scaled to a sample size of seven individuals. PAR estimates for sampled populations range from 0.00 to 0.33. Circles are proportional to PAR divided into 10 evenly spaced categories. Thick blue lines separate three microsatellite genetic clusters identified by STRUCTURE.

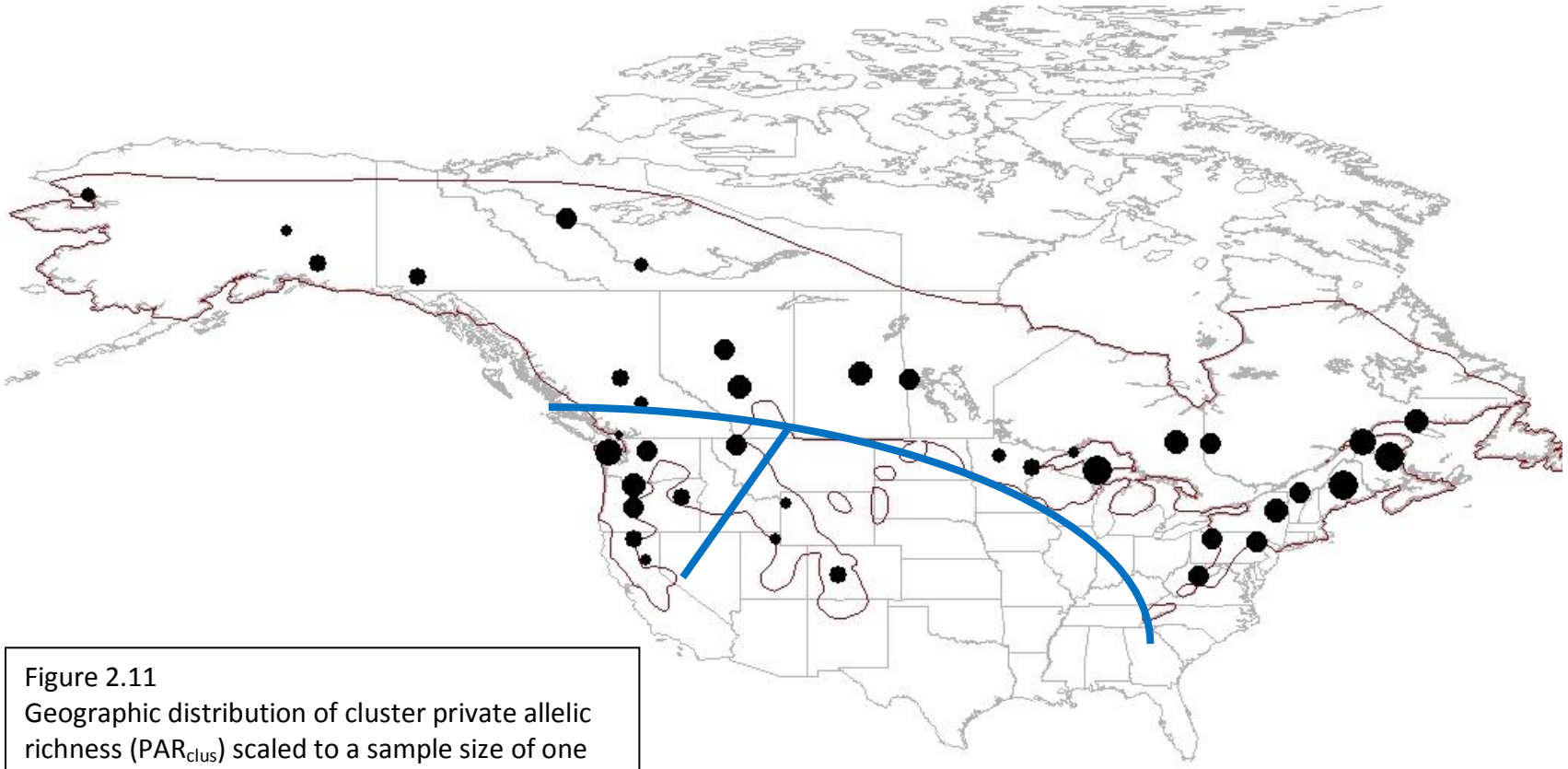
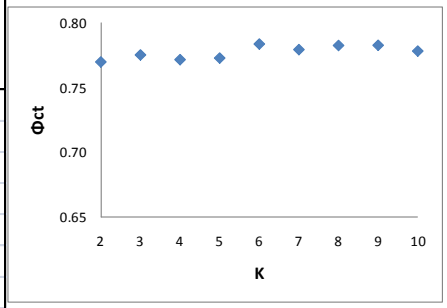


Figure 2.11
Geographic distribution of cluster private allelic richness (PAR_{clus}) scaled to a sample size of one population per genetic cluster and seven individuals per population. PAR estimates for sampled populations range from 0.00 to 0.33. Circles are proportional to PAR divided into 10 evenly spaced categories. Thick blue lines separate three microsatellite genetic clusters identified by STRUCTURE.

(A)

K	SAMOVA Groups	% of Variation		
		Among Groups (Φct)	Among Populations	Within Populations
2	(CA1, CA2, OR1, WA1, WA3) (all others)	77.0	13.3	9.7
3	(CA1, CA2, OR1, WA1) (WA3) (all others)	77.6	12.9	9.6
4	(CA1, OR1, WA1) (CA2) (WA3) (all others)	77.2	13.1	9.7
5	(CA1, CA2, OR1, WA1, WA3) (CO1) (OR2) (UT1, WY1) (all others)	77.3	6.7	16.0
6	**(CA1, CA2, OR1, WA1) (WA3) (CO1) (OR2) (UT1, WY1) (all others)**	**78.4**	**5.7**	**15.9**
7	(CA1, CA2, OR1, WA1) (WA3) (CO1) (OR2) (UT1, WY1) (MI1) (all others)	78.0	5.8	16.2
8	(CA1, CA2, OR1, WA1) (WA3) (CO1) (OR2) (UT1, WY1) (MI1) (WA4) (all others)	78.3	4.6	17.1
9	(CA1, CA2) (OR1, WA1) (WA3) (CO1) (OR2) (UT1, WY1) (WA4) (AK6) (all others)	78.3	4.5	17.2
10	(CA1, CA2) (OR1, WA1) (WA3) (CO1) (OR2) (UT1, WY1) (WA4) (BC1) (NL1) (all others)	77.9	4.3	17.9



(B)

K	SAMOVA Groups	% of Variation		
		Among Groups (Φct)	Among Populations	Within Populations
2	(OR2, CO1, UT1, WY1) (all others)	56.7	16.5	26.7
3	(CO1) (OR2, UT1, WY1) (all others)	62.8	11.5	25.8
4	(CO1) (OR2) (UT1, WY1) (all others)	66.1	8.9	25.1
5	(CO1) (OR2) (UT1, WY1) (WA4) (all others)	66.7	6.9	26.4
6	(CO1) (OR2) (UT1, WY1) (WA4) (BC1) (all others)	66.1	6.5	27.4
7	(CO1) (OR2) (UT1) (WY1) (WA4) (NL2) (all others)	66.1	6.3	27.6
8	(CO1) (OR2) (UT1) (WY1) (WA4) (BC1) (MI1) (all others)	66.5	5.6	27.9
9	(CO1) (OR2) (UT1) (WY1) (WA4) (BC1) (AK6) (IR1) (all others)	65.9	5.6	28.5
10	(CO1) (OR2) (UT1) (WY1) (WA4) (IR1) (ME2) (MI1) (PA1) (all others)	64.7	6.5	28.8

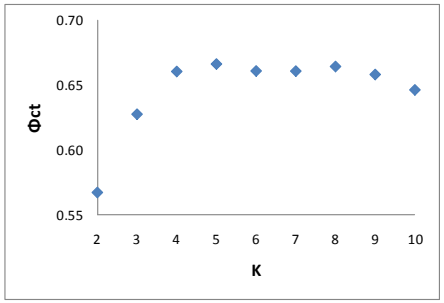


Table 2.1
Best groupings based on mitochondrial D-loop sequences for K= 2 to 10, using SAMOVA. Significance levels were evaluated by 1,000 permutations of populations among groups. All results are significant at $p < 0.01$. (A) SAMOVA analysis with all populations included; (B) SAMOVA analysis with Pacific NW populations omitted. Bolded row (K = 6 in Table 2.1A) identifies the haplotype grouping used in mtDNA analyses. Accompanying figures show Φ_{CT} (proportion of total genetic variation attributed to differences among groups) plotted against K (number of groups).

Pop	N	Haplotypes	H_d	π
AB1	9	7	0.94 (0.07)	0.0091 (0.0056)
AB2	18	17	0.99 (0.02)	0.0193 (0.0104)
AK2	28	13	0.91 (0.03)	0.0081 (0.0047)
AK4	15	6	0.80 (0.08)	0.0126 (0.0072)
AK6	10	7	0.91 (0.08)	0.0092 (0.0056)
BC1	17	2	0.12 (0.10)	0.0003 (0.0005)
BC2	25	14	0.88 (0.05)	0.0096 (0.0055)
BC4	9	7	0.94 (0.07)	0.0144 (0.0085)
IR1	8	4	0.79 (0.11)	0.0166 (0.0099)
MB1	13	7	0.87 (0.07)	0.0136 (0.0078)
ME1	40	16	0.93 (0.02)	0.0141 (0.0076)
ME2	10	9	0.98 (0.05)	0.0130 (0.0077)
MI1	8	6	0.89 (0.11)	0.0092 (0.0058)
MN1	35	28	0.98 (0.01)	0.0164 (0.0087)
MN2	12	11	0.98 (0.04)	0.0187 (0.0104)
MT1	127	31	0.94 (0.01)	0.0195 (0.0100)
NB1	20	13	0.96 (0.03)	0.0129 (0.0072)
NWT1	9	9	1.00 (0.05)	0.0181 (0.0105)
NWT2	18	14	0.97 (0.03)	0.0174 (0.0095)
NY1	13	8	0.90 (0.06)	0.0127 (0.0073)
ON1	14	12	0.97 (0.03)	0.0183 (0.0101)
ON3	19	18	0.99 (0.02)	0.0199 (0.0106)
PA1	10	6	0.84 (0.10)	0.0093 (0.0057)
PA2	13	8	0.88 (0.07)	0.0163 (0.0091)
QC3	19	10	0.86 (0.07)	0.0153 (0.0084)
QC4	17	9	0.89 (0.05)	0.0118 (0.0067)
SK1	9	9	1.00 (0.05)	0.0184 (0.0107)
WA4	29	10	0.87 (0.04)	0.0681 (0.0341)
WV1	13	3	0.62 (0.07)	0.0109 (0.0064)
YK2	30	20	0.97 (0.02)	0.0133 (0.0072)
BOREAL	***20.57 (21.78) Total N= 617***	***11.13 (6.63)***	***0.89 (0.17)***	***0.0155 (0.0108)***
OR2	***17***	***8***	***0.90 (0.04)***	***0.0147 (0.0081)***
CO1	***64***	***17***	***0.92 (0.01)***	***0.0058 (0.0035)***
UT1	25	6	0.62 (0.10)	0.0024 (0.0019)
WY1	80	15	0.84 (0.02)	0.0048 (0.0030)
UT-WY	***52.50 (38.89) Total N= 105***	***10.50 (6.36)***	***0.73 (0.16)***	***0.0036 (0.0017)***
CA1	12	2	0.41 (0.13)	0.0009 (0.0010)
CA2	7	2	0.47 (0.17)	0.0052 (0.0036)
OR1	32	15	0.93 (0.02)	0.0160 (0.0086)
WA1	30	11	0.86 (0.04)	0.0164 (0.0087)
PACIFIC NW	***20.25 (12.61) Total N= 81***	***7.50 (6.56)***	***0.67 (0.27)***	***0.0096 (0.0078)***
WA3	***9***	***7***	***0.91 (0.09)***	***0.0084 (0.0053)***

Table 2.2
D-loop haplotype (H_d) and nucleotide (π) diversities for each of 39 sampled populations. Populations are grouped into six haplotype lineages identified by SAMOVA, with lineage averages and total sample size for each lineage in bold italics. N = number of individuals; Haplotypes = number of unique haplotypes per population. Cluster standard deviations are in parentheses.

Divergence Event	Matthee	Wu	LOW	HIGH
CO1 / UT-WY	0.353	0.416	0.258	0.597
OR2 / UT-WY	0.653	0.770	0.478	1.104
PacNW / WA3	0.769	0.907	0.562	1.299
Boreal / (UT-WY, CO1, OR2)	0.888	1.047	0.650	1.501
Pacific NW-WA3 / BTJR	0.995	1.174	0.728	1.682

Table 2.3
Estimated divergence times for haplotype lineages, in millions of years before present. Results are ordered from most recent (top row) to oldest (bottom row) divergence. In first column, a forward slash ("/") separates the lineages or groups of lineages for which divergence times are estimated. Lineages are as identified by SAMOVA. Divergences between introgressed and non-introgressed lineages are not calculated because timing of lineage divergence would have been obscured by subsequent introgression. Four estimates are presented for each lineage divergence, reflecting uncertainties in timing of snowshoe hare / black-tailed jackrabbit divergence time. This species divergence estimate was used to calibrate divergence rate (% divergence per million years) for the cyt b segment used in this study. Cyt b divergence rate was subsequently used to estimate lineage divergence times. Matthee et al. (2004) estimated SSH-BTJR divergence at 4.79 mya (95% CI: 4.03–5.90 mya). Wu et al. (2005) estimated SSH-BTJR divergence at 5.65 mya (95% CI: 3.50–8.10 mya). The columns in this table represent lineage divergence times calculated using (Col 2) Matthee mean estimate of 4.79 mya; (Col 3) Wu mean estimate of 5.65 mya; (Col 4) Wu 95% LOW estimate of 3.50 mya; and (Col 5) Wu 95% HIGH estimate of 8.10 mya for SSH-BTJR divergence time.

K	SAMOVA Groups	% of Variation		
		Among Groups (Φct)	Among Populations	Within Populations
2	(UT1, WY1) (all others)	14.6	13.6	71.8
3	**(CO1, UT1, WY1) (BC1, CA1, CA2, MT1, OR1, OR2, WA1, WA3, WA4) (all others)**	**15.7**	**7.9**	**76.4**
4	(CO1) (UT1, WY1) (BC1, CA1, CA2, MT1, OR1, OR2, WA1, WA3, WA4) (all others)	17.0	6.8	76.2
5	(CO1) (UT1, WY1) (CA1, CA2) (BC1, MT1, OR1, OR2, WA1, WA3, WA4) (all others)	17.4	6.3	76.3
6	(CO1) (UT1, WY1) (CA1, CA2) (BC1) (MT1, OR1, OR2, WA1, WA3, WA4) (all others)	17.7	6.0	76.3
7	(CO1) (UT1, WY1) (CA1, CA2) (BC1) (MT1, OR2, WA4) (OR1, WA1, WA3) (all others)	17.9	5.5	76.7
8	(CO1) (UT1, WY1) (CA1) (CA2) (BC1) (MT1, OR2, WA4) (OR1, WA1, WA3) (all others)	18.1	5.3	76.6
9	(CO1) (UT1, WY1) (CA1) (CA2) (BC1) (MT1, OR2, WA4) (OR1, WA1, WA3) (NL1, NL2) (all others)	18.2	4.7	77.1
10	(CO1) (UT1, WY1) (CA1) (CA2) (BC1) (MT1, WA4) (OR2) (OR1, WA1, WA3) (NL1, NL2) (all others)	18.3	4.5	77.2

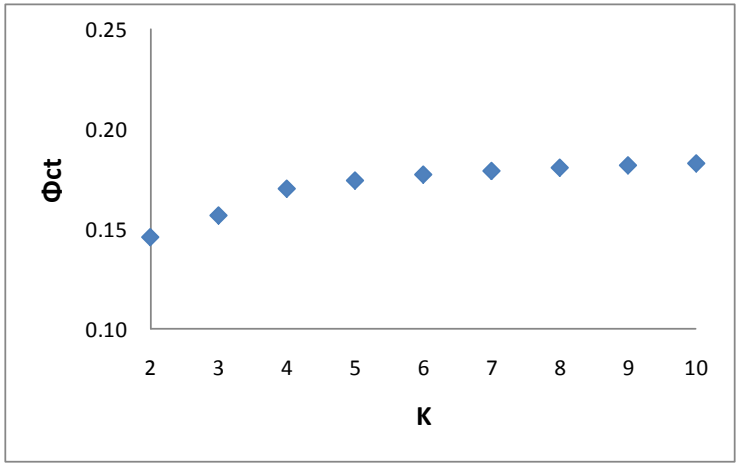


Table 2.4
Best groupings based on eight microsatellite loci for K= 2 to 10, using SAMOVA. Significance levels were evaluated by 1,000 permutations of populations among groups. All results are significant at $p < 0.01$. Bolded row (K = 3) identifies the genetic grouping that corresponds with STRUCTURE cluster results and was therefore used in microsatellite analyses. Accompanying figure shows Φct (proportion of total genetic variation attributed to differences among groups) plotted against K (number of groups).

POP	N	A	A_{eff}	AR	H_o	H_e	F_{IS}	PAR	PAR_{350}	PAR_{clus}	$\%PA_{clus}$
AB1	9	7 (0.94)	4.46 (0.93)	6.2	0.69 (0.09)	0.76 (0.04)	0.06 (0.08)	0.13	0.13	3.14	21%
AB2	18	8.63 (1.39)	5.32 (1.18)	5.95	0.69 (0.06)	0.77 (0.05)	0.08 (0.04)	0.14	0.14	2.96	12%
AK2	28	7.5 (1.6)	4.06 (1.2)	4.79	0.56 (0.07)	0.62 (0.08)	0.06 (0.04)	0.01	0.02	2.28	9%
AK4	15	5.75 (1.33)	3.9 (0.89)	4.63	0.56 (0.08)	0.64 (0.09)	0.09 (0.07)	0.07	0.07	2.22	10%
AK6	9	4.75 (0.82)	3.32 (0.75)	4.38	0.61 (0.08)	0.64 (0.07)	0 (0.08)	0.18	0.19	2.09	14%
BC2	25	7.88 (2.07)	5.09 (1.51)	5.09	0.57 (0.09)	0.67 (0.08)	0.14 (0.05)	0.07	0.07	2.55	12%
BC4	9	4.75 (0.67)	2.96 (0.4)	4.33	0.56 (0.07)	0.65 (0.05)	0.09 (0.1)	0.08	0.08	2.30	14%
IR1	10	5.88 (1.06)	3.99 (0.85)	5.13	0.66 (0.07)	0.71 (0.06)	0.01 (0.07)	0.05	0.05	2.40	13%
MB1	13	6.5 (0.96)	3.96 (0.73)	5.12	0.59 (0.08)	0.71 (0.05)	0.16 (0.07)	0.10	0.10	2.52	15%
ME1	40	9.75 (1.83)	5.68 (1.47)	5.89	0.72 (0.06)	0.77 (0.04)	0.05 (0.05)	0.05	0.06	3.32	20%
ME2	10	5.63 (1.12)	4.12 (0.96)	4.97	0.7 (0.08)	0.73 (0.05)	0 (0.09)	0.03	0.03	2.55	15%
MI1	8	6.5 (0.87)	4.47 (0.57)	6.1	0.72 (0.06)	0.8 (0.04)	0.03 (0.07)	0.33	0.33	3.16	17%
MN1	34	9.63 (2)	5.64 (1.52)	5.62	0.58 (0.06)	0.73 (0.07)	0.2 (0.05)	0.07	0.07	2.68	11%
MN2	12	7.5 (1.3)	5.25 (1.13)	5.97	0.7 (0.03)	0.79 (0.04)	0.07 (0.04)	0.04	0.04	2.78	9%
NB1	20	9 (1.68)	4.99 (1.52)	5.86	0.68 (0.06)	0.74 (0.04)	0.06 (0.06)	0.16	0.17	3.25	23%
NWT1	9	6.5 (1.13)	4.61 (0.99)	5.83	0.69 (0.06)	0.74 (0.06)	0.01 (0.06)	0.11	0.11	3.04	13%
NWT2	18	7.63 (1.55)	5.17 (1.31)	5.59	0.67 (0.07)	0.72 (0.07)	0.02 (0.08)	0.01	0.01	2.75	13%
NY1	13	6.75 (0.77)	4.5 (0.5)	5.58	0.7 (0.04)	0.79 (0.03)	0.07 (0.05)	0.10	0.10	3.03	16%
ON1	11	7 (1.41)	4.96 (1.34)	5.86	0.67 (0.07)	0.75 (0.06)	0.05 (0.1)	0.01	0.02	3.06	14%
ON3	19	9.25 (1.66)	5.18 (1.33)	5.85	0.66 (0.04)	0.75 (0.05)	0.08 (0.04)	0.15	0.15	3.01	15%
PA1	10	5.25 (0.98)	3.4 (0.66)	4.63	0.63 (0.07)	0.65 (0.07)	-0.03 (0.06)	0.07	0.15	2.57	18%
PA2	13	6.63 (1.22)	4.18 (1.02)	5.23	0.7 (0.07)	0.68 (0.07)	-0.07 (0.04)	0.07	0.07	2.70	15%
QC3	20	8.5 (1.61)	5.34 (1.29)	5.8	0.7 (0.07)	0.77 (0.05)	0.06 (0.07)	0.09	0.09	2.87	13%
QC4	17	7 (1.24)	4.48 (0.84)	5.38	0.71 (0.07)	0.76 (0.04)	0.05 (0.06)	0.04	0.05	2.99	23%
SK1	8	6.13 (1.14)	4.42 (1.09)	5.78	0.56 (0.11)	0.72 (0.07)	0.2 (0.12)	0.20	0.20	3.09	17%
WV1	14	4.5 (0.6)	2.77 (0.33)	3.9	0.5 (0.08)	0.63 (0.05)	0.2 (0.08)	0.02	0.10	2.25	19%
YK2	30	8.5 (2.13)	4.23 (1.15)	5.1	0.62 (0.05)	0.65 (0.07)	0.02 (0.04)	0.11	0.11	2.52	16%
BOREAL	***Total N= 442***	***7.05 (1.50)***	***4.46 (0.78)***	***5.35 (0.61)***	***0.64 (0.06)***	***0.72 (0.06)***	***0.06 (0.07)***	***0.09 (0.07)***	***0.10 (0.07)***	***2.74 (0.35)***	***15% (4%)***
CO1	58	7.75 (1.94)	3.61 (0.88)	4.25	0.55 (0.1)	0.57 (0.11)	0 (0.05)	0.14	0.14	1.98	1%
UT1	25	4.88 (1.03)	2.78 (0.49)	3.66	0.5 (0.09)	0.54 (0.09)	0.1 (0.07)	0.24	0.27	1.60	4%
WY1	77	7.25 (1.57)	3.43 (0.92)	4.06	0.53 (0.1)	0.55 (0.1)	0.03 (0.03)	0.07	0.10	1.62	1%
ROCKIES	***Total N= 160***	***6.63 (1.53)***	***3.27 (0.44)***	***3.99 (0.30)***	***0.53 (0.02)***	***0.55 (0.02)***	***0.04 (0.05)***	***0.15 (0.08)***	***0.17 (0.09)***	***1.73 (0.21)***	***2% (1%)***
BC1	15	4.25 (0.8)	2.23 (0.27)	3.3	0.46 (0.08)	0.51 (0.08)	0.05 (0.09)	0.00	0.02	1.02	0%
CA1	12	4.88 (1.08)	3.06 (0.66)	4.16	0.59 (0.1)	0.61 (0.07)	-0.01 (0.1)	0.12	0.13	1.79	1%
CA2	7	3.63 (0.63)	2.58 (0.48)	3.63	0.48 (0.11)	0.52 (0.12)	-0.03 (0.06)	0.01	0.01	1.42	2%
MT1	100	12.13 (3.36)	5.88 (1.83)	5.58	0.7 (0.07)	0.73 (0.06)	0.04 (0.03)	0.17	0.17	2.13	4%
OR1	32	9 (2.15)	5.39 (1.75)	5.29	0.64 (0.09)	0.68 (0.09)	0.05 (0.03)	0.11	0.23	2.37	5%
OR2	17	5.88 (1.23)	3.66 (0.88)	4.48	0.62 (0.11)	0.61 (0.09)	-0.01 (0.09)	0.19	0.23	1.73	5%
WA1	30	9.25 (2.1)	5.88 (1.2)	5.91	0.71 (0.06)	0.77 (0.06)	0.05 (0.03)	0.17	0.38	2.58	6%
WA3	9	5.13 (1.16)	3.76 (0.76)	4.64	0.71 (0.11)	0.65 (0.11)	-0.17 (0.06)	0.19	0.29	2.34	10%
WA4	29	9.38 (2.37)	5.96 (1.47)	5.73	0.67 (0.08)	0.76 (0.06)	0.12 (0.05)	0.25	0.35	2.32	4%
PACIFIC NW	***Total N= 251***	***7.06 (2.94)***	***4.27 (1.52)***	***4.75 (0.94)***	***0.62 (0.09)***	***0.65 (0.10)***	***0.01 (0.08)***	***0.13 (0.08)***	***0.20 (0.13)***	***1.97 (0.51)***	***4% (3%)***

Table 2.5

Microsatellite diversities averaged across 8 polymorphic loci for each of 39 sampled populations. Populations are grouped into three genetic clusters identified by STRUCTURE and SAMOVA, with cluster averages and total sample size for each cluster in bold italics. N = number of individuals; A = number of different alleles; A_{eff} = effective number of alleles, AR = allelic richness standardized to seven samples per population; H_o= observed heterozygosity; H_e= unbiased expected heterozygosity; F_{IS} = inbreeding coefficient; PAR = Population private allelic richness standardized to seven samples per population; PAR_{350} = Population private allelic richness, with all sampled populations within a 350 km radius excluded; PAR_{clus} = Kalinowski's (2004) hierarchical measure of PAR standardized to one population per cluster and seven individuals per population; $\%PA_{clus}$ = proportion of alleles in each population that are cluster private alleles. Standard deviations are in parentheses.

	COMBINED_Quad	COMBINED	Boreal	Pacific NW
	LATITUDINAL TREND			
AR	**0.37****	0.26	-0.05	**0.74***
H_e	**0.35****	0.17	-0.3	**0.77***
PAR	0.001	-0.07	0.05	**0.85****
PAR_{350}	0.008	-0.19	-0.07	**0.73***
PAR_{clus}	0.08	0.21	**-0.46***	0.64
	LONGITUDINAL TREND			
AR		**0.32***	0.32	0.22
H_e		**0.45****	**0.56****	0.17
PAR		-0.15	-0.05	0.13
PAR_{350}		-0.18	0.02	-0.22
PAR_{clus}		**0.51****	**0.53****	-0.31

Table 2.6
Column 1 shows the quadratic relationship (r^2) between microsatellite genetic diversity and latitude for all lineages combined. Columns 2-4 show Pearson's correlation (r) between microsatellite genetic diversity and latitude / longitude for all lineages combined, Boreal and Pacific NW lineages each separately. The Rockies lineage was not evaluated separately due to small sample sizes. Significant results are marked (*$p < 0.05$; **$p < 0.01$).

	AB1	AB2	AK2	AK4	AK6	BC2	BC4	IR1	MB1	ME1	ME2	MI1	MN1	MN2	NB1	NWT1	NWT2	NY1	ON1	ON3	PA1	PA2	QC3	QC4	SK1	WV1	YK2	CO1	UT1	WY1	BC1	CA1	CA2	MT1	OR1	OR2	WA1	WA3	WA4	
AB1		0.38	0.37	0.44	0.39	0.27	0.43	0.49	0.31	0.22	0.35	0.37	0.33	0.27	0.25	0.23	0.24	0.31	0.23	0.29	0.34	0.29	0.40	0.38	0.30	0.46	0.39	1.41	1.33	1.49	1.02	0.98	0.74	0.74	0.78	0.79	0.75	0.87	0.64	AB1
AB2	0.07		0.18	0.17	0.19	0.14	0.21	0.26	0.20	0.17	0.22	0.32	0.08	0.15	0.16	0.15	0.11	0.25	0.18	0.12	0.37	0.20	0.22	0.16	0.16	0.37	0.16	0.71	0.80	0.78	0.84	1.09	0.98	0.74	0.99	0.93	0.72	1.13	0.69	AB2
AK2	0.11	0.06		0.11	0.12	0.10	0.18	0.40	0.22	0.24	0.28	0.52	0.13	0.22	0.22	0.14	0.15	0.29	0.19	0.24	0.40	0.18	0.31	0.33	0.14	0.43	0.05	1.02	1.32	1.48	1.06	1.14	1.09	0.82	1.20	0.83	0.96	1.47	0.79	AK2
AK4	0.12	0.04	0.04		0.16	0.15	0.29	0.27	0.25	0.25	0.30	0.55	0.18	0.28	0.25	0.23	0.16	0.33	0.26	0.21	0.55	0.24	0.35	0.32	0.21	0.55	0.11	0.97	1.18	1.19	0.70	1.12	0.97	0.73	1.16	1.10	0.83	1.42	0.68	AK4
AK6	0.09	0.04	0.03	0.04		0.14	0.27	0.36	0.24	0.29	0.32	0.50	0.19	0.25	0.29	0.17	0.15	0.37	0.26	0.21	0.49	0.26	0.33	0.39	0.20	0.55	0.12	1.16	1.34	1.49	0.94	0.93	0.92	0.84	1.09	0.94	0.92	1.20	0.83	AK6
BC2	0.07	0.04	0.03	0.05	0.03		0.13	0.25	0.15	0.15	0.20	0.36	0.11	0.22	0.16	0.08	0.10	0.26	0.11	0.15	0.27	0.14	0.30	0.30	0.13	0.30	0.11	1.12	1.26	1.32	0.87	0.82	0.78	0.77	0.89	0.77	0.74	1.17	0.68	BC2
BC4	0.10	0.04	0.06	0.09	0.08	0.03		0.44	0.24	0.28	0.29	0.46	0.17	0.23	0.23	0.17	0.21	0.36	0.23	0.21	0.39	0.24	0.32	0.39	0.22	0.35	0.22	1.23	1.59	1.54	1.17	1.15	1.04	0.93	1.00	0.75	0.86	1.26	0.89	BC4
IR1	0.09	0.04	0.12	0.07	0.09	0.06	0.11		0.35	0.23	0.32	0.46	0.28	0.42	0.30	0.30	0.25	0.39	0.33	0.21	0.51	0.37	0.45	0.31	0.37	0.46	0.36	0.98	0.84	0.87	1.07	0.91	0.94	1.20	1.41	1.72	1.20	1.97	1.00	IR1
MB1	0.06	0.04	0.08	0.07	0.06	0.04	0.06	0.07		0.27	0.28	0.34	0.18	0.20	0.29	0.17	0.20	0.34	0.23	0.17	0.42	0.26	0.29	0.34	0.24	0.46	0.21	1.03	1.02	1.01	0.91	1.03	0.93	0.76	0.89	0.74	0.72	1.11	0.74	MB1
ME1	0.03	0.03	0.08	0.07	0.08	0.04	0.07	0.04	0.06		0.15	0.31	0.17	0.23	0.08	0.14	0.14	0.16	0.10	0.15	0.26	0.12	0.27	0.17	0.18	0.27	0.23	1.11	1.14	1.16	1.00	1.00	0.93	0.94	1.11	1.25	0.92	1.21	0.83	ME1
ME2	0.06	0.04	0.09	0.08	0.08	0.05	0.07	0.06	0.06	0.02		0.39	0.23	0.28	0.19	0.27	0.24	0.25	0.22	0.22	0.37	0.23	0.35	0.25	0.26	0.45	0.31	1.08	1.18	1.17	1.06	0.99	0.85	0.96	0.99	1.11	0.83	1.20	0.83	ME2
MI1	0.05	0.04	0.14	0.13	0.11	0.08	0.10	0.07	0.06	0.05	0.06		0.33	0.26	0.35	0.30	0.33	0.33	0.36	0.31	0.44	0.43	0.39	0.43	0.42	0.55	0.48	0.98	1.22	1.13	0.98	1.10	1.17	0.91	0.80	0.97	0.66	0.95	0.79	MI1
MN1	0.07	0.01	0.05	0.05	0.05	0.03	0.04	0.05	0.04	0.04	0.05	0.06		0.14	0.16	0.12	0.09	0.26	0.11	0.12	0.26	0.16	0.19	0.17	0.12	0.34	0.13	0.68	0.89	0.93	0.98	1.29	1.20	0.72	1.21	0.80	0.89	1.29	0.73	MN1
MN2	0.03	0.01	0.07	0.07	0.05	0.05	0.05	0.07	0.03	0.04	0.04	0.02	0.02		0.26	0.20	0.16	0.24	0.21	0.15	0.43	0.28	0.21	0.29	0.28	0.47	0.26	0.83	0.93	0.97	1.04	1.11	1.13	0.81	0.92	0.81	0.73	1.05	0.79	MN2
NB1	0.04	0.03	0.08	0.07	0.08	0.04	0.06	0.06	0.07	0.01	0.03	0.06	0.04	0.04		0.14	0.14	0.16	0.13	0.15	0.29	0.11	0.27	0.15	0.15	0.29	0.22	1.12	1.31	1.32	1.04	1.20	1.09	0.91	1.20	1.09	0.97	1.45	0.88	NB1
NWT1	0.02	0.01	0.04	0.06	0.03	0.00	0.02	0.05	0.02	0.01	0.04	0.03	0.01	0.02	0.01		0.08	0.27	0.11	0.15	0.23	0.16	0.26	0.26	0.12	0.30	0.13	1.03	1.19	1.25	1.04	0.96	0.93	0.83	0.99	0.84	0.81	1.14	0.75	NWT1
NWT2	0.04	0.01	0.05	0.04	0.03	0.02	0.05	0.04	0.04	0.03	0.05	0.06	0.01	0.02	0.02	-0.01		0.25	0.13	0.12	0.30	0.15	0.29	0.21	0.11	0.35	0.14	0.77	1.09	1.15	0.86	0.91	0.99	0.81	1.09	0.99	0.79	1.32	0.69	NWT2
NY1	0.04	0.04	0.09	0.09	0.09	0.07	0.08	0.06	0.07	0.02	0.04	0.04	0.05	0.02	0.02	0.04	0.05		0.22	0.22	0.44	0.26	0.25	0.21	0.31	0.37	0.31	1.22	1.20	1.23	1.12	1.12	1.12	0.95	1.05	1.26	0.91	1.18	0.92	NY1
ON1	0.03	0.02	0.06	0.07	0.06	0.02	0.05	0.06	0.04	0.00	0.03	0.05	0.01	0.02	0.01	-0.01	0.01	0.03		0.16	0.24	0.12	0.24	0.20	0.13	0.29	0.19	0.90	1.07	1.14	1.17	1.16	1.08	0.84	1.20	0.94	0.98	1.40	0.80	ON1
ON3	0.05	0.01	0.08	0.06	0.05	0.04	0.05	0.03	0.03	0.03	0.04	0.04	0.02	0.01	0.03	0.01	0.02	0.03	0.02		0.38	0.24	0.16	0.22	0.20	0.45	0.25	0.94	1.01	0.95	0.92	1.01	0.90	0.84	1.03	0.99	0.81	1.16	0.82	ON3
PA1	0.08	0.09	0.15	0.17	0.15	0.09	0.12	0.13	0.12	0.07	0.10	0.10	0.07	0.10	0.08	0.05	0.08	0.10	0.05	0.10		0.25	0.51	0.39	0.30	0.24	0.42	1.51	1.76	1.91	1.46	1.35	1.29	1.12	1.40	1.00	1.23	1.46	1.10	PA1
PA2	0.06	0.04	0.06	0.08	0.07	0.04	0.06	0.08	0.07	0.02	0.05	0.09	0.04	0.06	0.02	0.02	0.03	0.05	0.01	0.06	0.07		0.34	0.23	0.16	0.24	0.19	1.17	1.39	1.57	0.93	1.07	1.04	0.85	1.23	1.05	0.96	1.51	0.79	PA2
QC3	0.07	0.04	0.10	0.10	0.09	0.08	0.08	0.08	0.06	0.05	0.07	0.06	0.04	0.03	0.06	0.04	0.06	0.04	0.04	0.02	0.12	0.08		0.25	0.31	0.55	0.30	1.11	1.03	1.05	1.07	1.51	1.34	0.83	1.17	0.99	0.97	1.06	0.93	QC3
QC4	0.07	0.02	0.11	0.09	0.10	0.08	0.10	0.05	0.07	0.03	0.04	0.07	0.03	0.04	0.02	0.04	0.04	0.03	0.03	0.04	0.10	0.05	0.04		0.26	0.31	0.31	0.84	0.84	0.88	1.14	1.59	1.47	0.92	1.51	1.31	1.15	1.39	0.97	QC4
SK1	0.05	0.02	0.04	0.05	0.04	0.02	0.04	0.07	0.05	0.03	0.04	0.07	0.01	0.04	0.02	-0.01	0.00	0.05	0.00	0.03	0.07	0.02	0.06	0.04		0.35	0.11	0.94	1.27	1.34	1.05	1.12	1.04	0.79	1.20	0.91	0.88	1.31	0.75	SK1
WV1	0.12	0.10	0.16	0.18	0.18	0.10	0.12	0.13	0.14	0.08	0.13	0.13	0.10	0.12	0.09	0.08	0.10	0.10	0.08	0.13	0.08	0.07	0.14	0.09	0.10		0.48	1.51	1.64	1.83	1.52	1.42	1.57	1.27	1.62	1.28	1.39	1.61	1.23	WV1
YK2	0.11	0.05	0.01	0.03	0.03	0.04	0.07	0.10	0.06	0.07	0.09	0.12	0.04	0.07	0.07	0.03	0.04	0.08	0.05	0.07	0.14	0.06	0.09	0.09	0.02	0.16		0.94	1.19	1.30	0.96	1.08	1.08	0.80	1.21	0.94	0.93	1.41	0.76	YK2
CO1	0.30	0.21	0.30	0.29	0.31	0.30	0.32	0.26	0.28	0.25	0.28	0.25	0.21	0.23	0.27	0.27	0.24	0.27	0.25	0.25	0.34	0.30	0.26	0.24	0.26	0.34	0.28		0.49	0.44	1.47	1.65	1.90	0.97	1.87	1.22	1.19	2.18	0.94	CO1
UT1	0.30	0.23	0.34	0.32	0.34	0.31	0.35	0.26	0.28	0.25	0.29	0.28	0.24	0.24	0.29	0.29	0.28	0.27	0.28	0.26	0.36	0.32	0.25	0.24	0.31	0.36	0.31	0.23		0.10	1.68	1.12	1.69	1.03	1.51	1.54	1.23	1.48	0.93	UT1
WY1	0.32	0.24	0.36	0.33	0.36	0.33	0.36	0.27	0.29	0.28	0.31	0.28	0.27	0.27	0.31	0.31	0.30	0.29	0.30	0.27	0.38	0.35	0.28	0.26	0.33	0.38	0.33	0.22	0.06		1.51	1.03	1.52	0.96	1.25	1.60	0.95	1.36	0.87	WY1
BC1	0.27	0.23	0.32	0.26	0.30	0.27	0.33	0.29	0.27	0.24	0.28	0.26	0.26	0.26	0.27	0.28	0.25	0.27	0.29	0.25	0.36	0.28	0.26	0.27	0.29	0.37	0.29	0.38	0.41	0.40		0.99	0.66	0.33	0.59	0.76	0.37	0.77	0.29	BC1
CA1	0.22	0.22	0.30	0.28	0.26	0.23	0.28	0.22	0.24	0.21	0.23	0.22	0.25	0.22	0.24	0.22	0.22	0.22	0.24	0.22	0.30	0.26	0.25	0.26	0.25	0.32	0.27	0.36	0.33	0.33	0.33		0.41	0.88	0.35	1.12	0.44	0.75	0.51	CA1
CA2	0.21	0.24	0.32	0.29	0.29	0.25	0.30	0.26	0.26	0.22	0.24	0.26	0.27	0.25	0.26	0.25	0.26	0.25	0.27	0.23	0.33	0.28	0.27	0.28	0.28	0.36	0.30	0.40	0.41	0.40	0.30	0.18		0.53	0.37	0.63	0.42	0.62	0.40	CA2
MT1	0.14	0.15	0.20	0.18	0.19	0.18	0.20	0.19	0.16	0.17	0.18	0.15	0.16	0.15	0.17	0.16	0.17	0.16	0.16	0.16	0.22	0.18	0.16	0.17	0.16	0.24	0.19	0.24	0.25	0.26	0.13	0.21	0.17		0.41	0.21	0.25	0.56	0.14	MT1
OR1	0.17	0.19	0.27	0.25	0.24	0.22	0.23	0.23	0.20	0.20	0.20	0.16	0.22	0.17	0.22	0.20	0.22	0.19	0.22	0.20	0.26	0.24	0.21	0.23	0.23	0.29	0.25	0.34	0.32	0.32	0.21	0.12	0.14	0.12		0.48	0.11	0.25	0.21	OR1
OR2	0.20	0.21	0.25	0.28	0.26	0.22	0.22	0.29	0.20	0.23	0.25	0.21	0.20	0.19	0.24	0.21	0.23	0.24	0.22	0.22	0.26	0.26	0.22	0.25	0.23	0.30	0.25	0.33	0.36	0.37	0.29	0.30	0.24	0.07	0.16		0.44	0.66	0.37	OR2
WA1	0.13	0.13	0.21	0.18	0.19	0.17	0.18	0.18	0.14	0.15	0.15	0.10	0.16	0.12	0.16	0.14	0.15	0.14	0.16	0.14	0.21	0.18	0.15	0.17	0.16	0.23	0.20	0.27	0.27	0.26	0.14	0.13	0.14	0.06	0.03	0.13		0.29	0.13	WA1
WA3	0.18	0.20	0.31	0.29	0.27	0.26	0.27	0.27	0.23	0.21	0.23	0.18	0.23	0.19	0.24	0.22	0.24	0.20	0.24	0.21	0.28	0.27	0.20	0.23	0.24	0.31	0.28	0.37	0.35	0.35	0.27	0.23	0.23	0.14	0.08	0.21	0.07		0.42	WA3
WA4	0.12	0.12	0.20	0.16	0.18	0.16	0.18	0.16	0.15	0.14	0.15	0.12	0.14	0.13	0.16	0.13	0.14	0.14	0.14	0.15	0.20	0.16	0.15	0.16	0.14	0.22	0.18	0.24	0.24	0.25	0.12	0.14	0.14	0.04	0.06	0.12	0.02	0.11		WA4
	AB1	AB2	AK2	AK4	AK6	BC2	BC4	IR1	MB1	ME1	ME2	MI1	MN1	MN2	NB1	NWT1	NWT2	NY1	ON1	ON3	PA1	PA2	QC3	QC4	SK1	WV1	YK2	CO1	UT1	WY1	BC1	CA1	CA2	MT1	OR1	OR2	WA1	WA3	WA4	

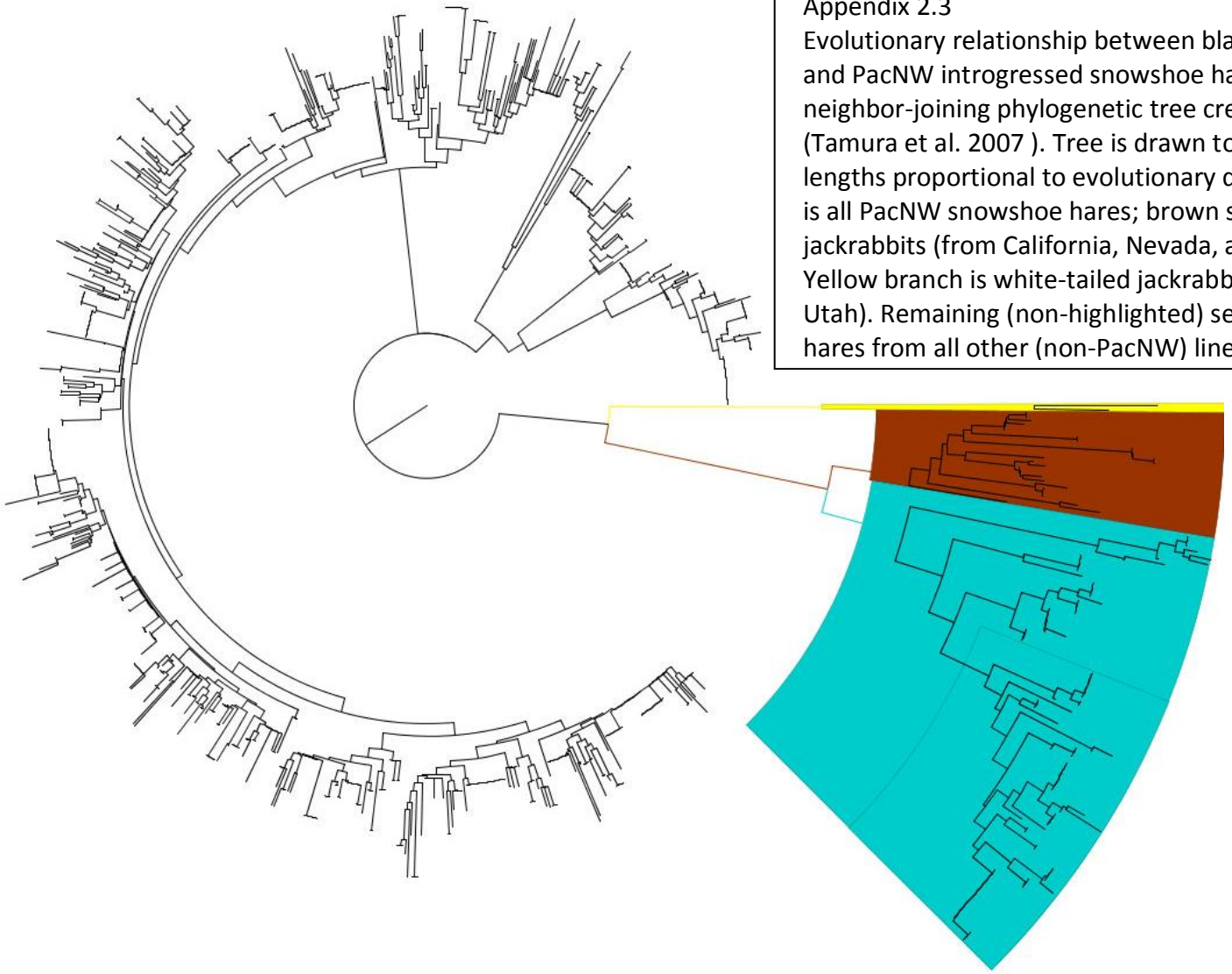
Table 2.7
F_{ST} (below diagonal) and Nei's D (above diagonal) for each population pair. Populations are grouped into three genetic clusters identified by STRUCTURE.

Site Name	State/ Prov/ Territory	Latitude	Longitude	MIN	MAX
North Cove	NE	400700	992200	10000	10000
North Cove	NE	400700	992200	10000	10000
North Cove	NE	400700	992200	10000	10000
North Cove	NE	400700	992200	10000	10000
North Cove	NE	400700	992200	10000	10000
Itasca	MN	470700	950700	10000	10000
Charlie Lake Cave	BC	561635	1205615	10000	10000
Charlie Lake Cave	BC	561635	1205615	10000	10000
Charlie Lake Cave	BC	561635	1205615	10000	10000
Layser Cave	WA	462200	1214500	10000	10000
Layser Cave	WA	462200	1214500	10000	10000
Layser Cave	WA	462200	1214500	10000	10000
Judd Peak South	WA	463000	1215200	10000	10000
Kelso Cave	ON	433000	795500	10000	10000
Boss Hill	AB	523500	1124500	10000	10000
Broken Mammoth	AK	641600	1460700	10000	10000
Broken Mammoth	AK	641600	1460700	10000	10000
Broken Mammoth	AK	641600	1460700	10000	10000
Broken Mammoth	AK	641600	1460700	10000	10000
Lindenmeier	CO	405200	1050000	10000	10000
Moonshiner	ID	432200	1123700	10000	10000
Little Box Elder Cave	WY	423700	1053700	10000	10000
New Paris #4	PA	400500	783900	10000	10000
Mummy Cave	WY	442700	1094500	10000	10000
Mummy Cave	WY	442700	1094500	10000	10000
Frankstown Cave	PA	402200	781500	10000	10000
Welsh Cave	KY	375225	844450	10000	10000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
January Cave	AB	501118	1143106	10000	20000
Blue Fish Caves I-III	YU	670800	1404700	10000	20000
Little Box Elder Cave	WY	423700	1053700	10000	20000
Clark's Cave	VA	380510	793925	10000	20000
Crystal Ball Cave	UT	390000	1130000	10000	20000
Jaguar Cave	ID	441653	1125455	10000	40000
Natural Chimneys	VA	382200	780500	10000	40000
Bat Cave	MO	375200	921500	10000	40000
Back Creek Cave #2	VA	380400	795330	10000	40000

Appendix 2.1
Locations and estimated ages of snowshoe hare fossils from Late Wisconsin period, approximately coinciding with LGM. MIN and MAX are estimated years before present. Data are from FAUNMAP Working Group (1994).

State/ Province/ Territory	Genetic Samples Donors
Alaska	Dubois, Golden, Kielland, Johnson, Morton, Paragi
Alberta	Donahue, Grandjambe, Myers, TEvans, Walde
British Columbia	Boland, Doyle, Gautron, Pellerin, Polfus, Ransome
California	Cheng, Cheyne, Simons
Colorado	Ivan, Wait
Maine	Blood, Gomm, Harrison, Jakubas
Manitoba	Duncan, Koshel, Olson, Roberts
Michigan	Peterson, Quigley, Spreeman
Minnesota	Erb, Mortensen, Palas, Schrage
Montana	Cheng, Hodges, Mills, Tyers
New Brunswick	Lavigne, Wheelwright
New York	Cucharale, Fuhs, Huston, Platoni
Newfoundland & Labrador	Crowe, Jennings, Lane, McGrath
Northwest Territories	Carriere
Nova Scotia	OBrien
Ontario	Chartrand, Thompson
Oregon	Cheng, Hennings, Strauser
Pennsylvania	Allabaugh, Bodenhorn, Kerner, Mills, Pawelski
Prince Edward Island	Dibblee
Quebec	Jolicoeur, Sirois
Saskatchewan	Gordon, Luthi, Weber
Utah	Olsen, Strauser
Vermont	Richardson
Washington	Gremel, Griffin, MacCracken, Strauser, Walker
West Virginia	Tolin
Wisconsin	Janak, Meyer
Wyoming	Berg, Hodges, Mills
Yukon	Doyle, ODonoghue, Sealy, Sheriff

Appendix 2.2
Donors of snowshoe hare genetic samples used in this study, listed alphabetically by state/ province/ territory of collection.



Appendix 2.3
Evolutionary relationship between black-tailed jackrabbits and PacNW introgressed snowshoe hares, based on neighbor-joining phylogenetic tree created in MEGA v. 4.0 (Tamura et al. 2007). Tree is drawn to scale, with branch lengths proportional to evolutionary distance. Blue segment is all PacNW snowshoe hares; brown segment is black-tailed jackrabbits (from California, Nevada, and New Mexico); Yellow branch is white-tailed jackrabbits (from Colorado and Utah). Remaining (non-highlighted) segments are snowshoe hares from all other (non-PacNW) lineages.

Appendix 2.4

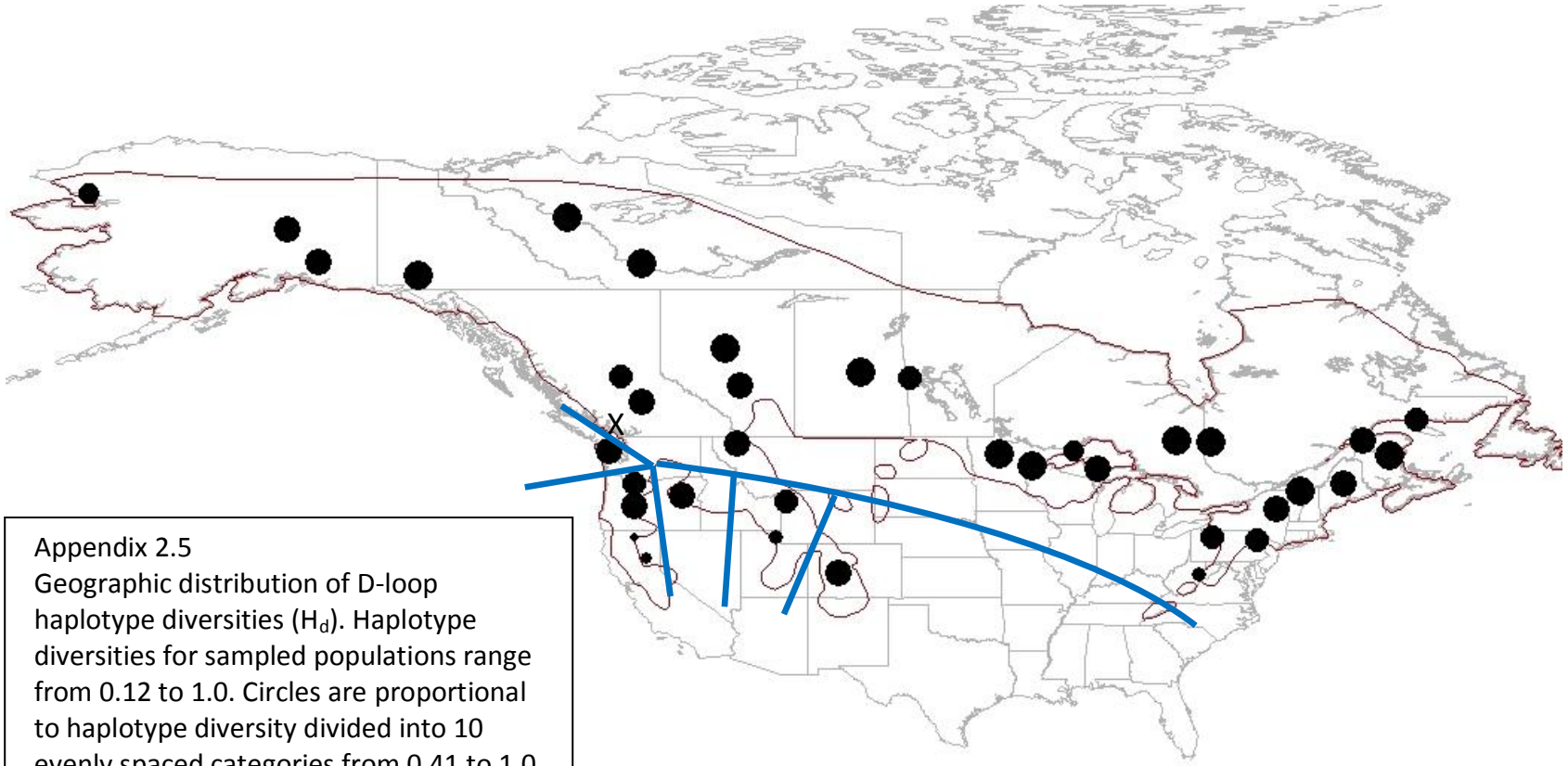
Distribution of D-loop haplotypes among populations. Populations are grouped into six haplotype lineages identified by SAMOVA. For each population, N= number of individuals. For each haplotype, H_N = number of sampled individuals carrying that haplotype; Unique Pops= number of different populations carrying that haplotype.

			BOREAL																										OR2	CO1	UT-WY		PACNW				WA3
			AB1	AB2	AK2	AK4	AK6	BC1	BC2	BC4	IR1	MB1	ME1	ME2	MI1	MN1	MN2	MT1	NB1	NWT1	NWT2	NY1	ON1	ON3	PA1	PA2	QC3	QC4	SK1	WA4	WV1	YK2	OR2	CO1	UT1	WY1	CA1	CA2	OR1	WA1	WA3
N			9	18	28	15	10	17	25	9	8	13	40	10	8	35	12	127	20	9	18	13	14	19	10	13	19	17	9	29	13	30	17	64	25	80	12	7	32	30	9
Haplotype Name	H_N	Unique Pops																																							
Hap1	2	1																			2																				
Hap2	17	2						16																			1														
Hap3	1	1						1																																	
Hap4	3	1																			3																				
Hap5	1	1																			1																				
Hap6	1	1																			1																				
Hap7	7	4							2							2		2			1																				
Hap9	1	1																	1																						
Hap10	7	2											3													4															
Hap11	1	1																			1																				
Hap12	1	1																														1									
Hap13	1	1																														1									
Hap14	2	1																														2									
Hap15	1	1																			1																				
Hap16	10	3	2	2	6																																				
Hap17	1	1		1																																				
Hap18	2	2		1																									1											
Hap19	1	1		1																																				
Hap20	2	2		1																			1																	
Hap21	1	1		1																																				
Hap22	1	1		1																																				
Hap23	1	1		1																																				
Hap24	1	1		1																																				
Hap25	1	1																			1																				
Hap26	1	1		1																																				
Hap27	1	1		1																																				
Hap28	6	4		1								3			1									1																	
Hap29	2	2		1								1																													
Hap30	1	1		1																																				
Hap31	2	2											1						1																						
Hap32	2	1																	2																						
Hap33	3	1																	3																						
Hap34	1	1																	1																						
Hap35	1	1																			1																				
Hap36	15	6											6	1					2			4			1			1													
Hap37	2	1																	2																						
Hap38	3	2											2						1																						
Hap39	2	1																	2																						
Hap40	1	1																	1																						
Hap41	1	1																	1																						
Hap42	1	1																			1																				
Hap43	2	1																	2																						
Hap44	3	2												2					1																						
Hap45	1	1																			1																				
Hap46	9	2							8	1																															
Hap47	1	1							1																																
Hap48	1	1							1																																
Hap49	1	1																			1																				
Hap50	3	2							1	2																															
Hap51	1	1							1																																
Hap52	2	2							1	1																															
Hap53	1	1								1																															
Hap54	1	1								1																															
Hap55	2	1								2																															
Hap56	1	1																			1																				
Hap57	1	1								1																															
Hap58	2	1										2																													
Hap59	1	1										1																													
Hap60	2	1																			2																				
Hap63	1	1																		1																					
Hap64	1	1																		1																					
Hap65	1	1																		1																					
Hap66	1	1																		1																					
Hap67	1	1																		1																					
Hap68	1	1																		1																					
Hap69	5	2										4								1																					
Hap70	1	1																		1																					
Hap71	1	1																		1																					
Hap72	4	1							4																																
Hap73	1	1																														1									
Hap74	5	3			1		1																									3									
Hap75	3	1																														3									
Hap76	1	1		1																																				
Hap77	1	1		1																																				
Hap78	1	1		1																																				
Hap79	1	1	1																																					
Hap80	2	1																														2									
Hap81	1	1	1																																					
Hap82	1	1	1																																					
Hap83	2	1	2																																					
Hap84	1	1																						1																	
Hap86	14	3			5	6																										3									

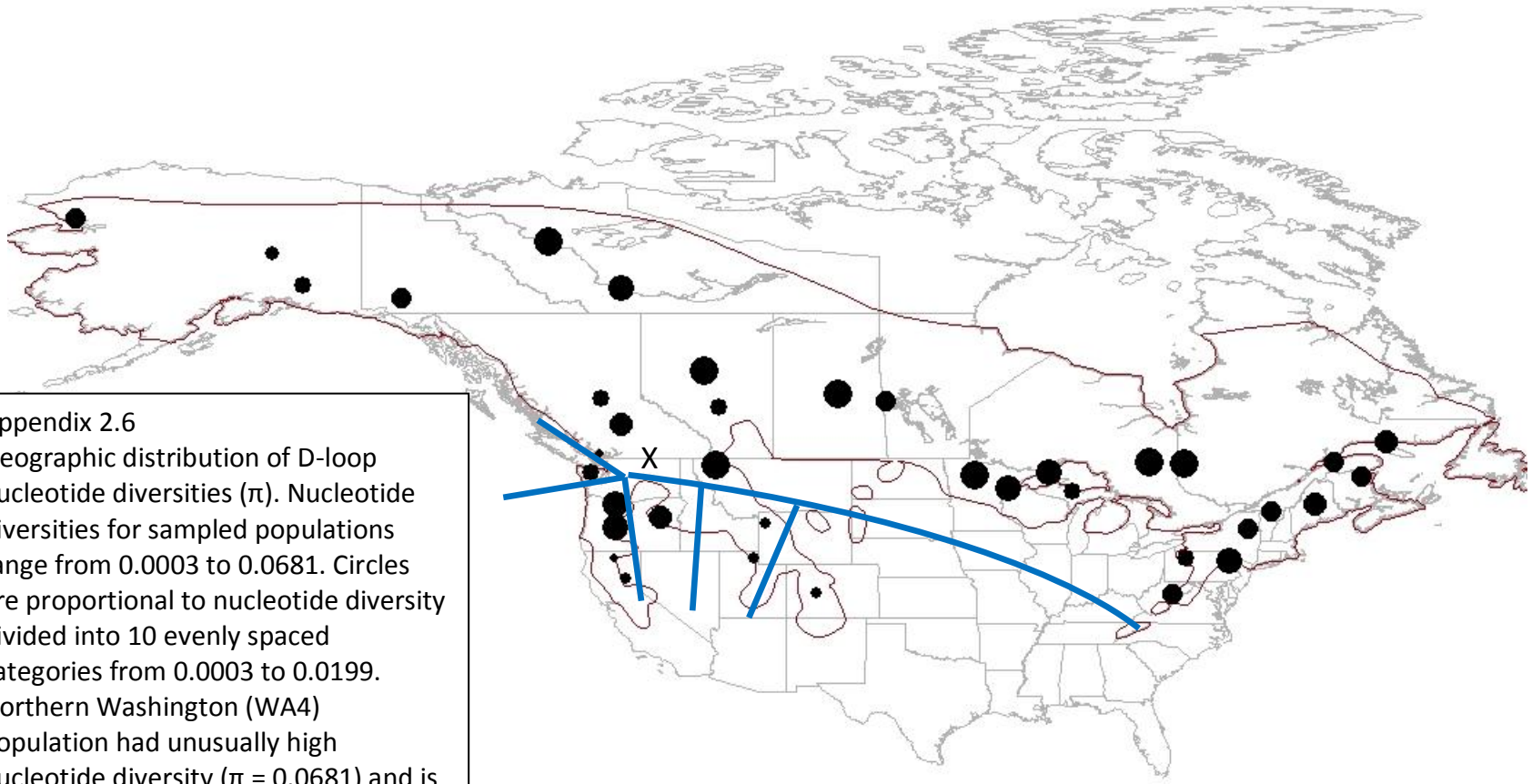
			BOREAL																													OR2	CO1	UT-WY		PACNW				WA3
			AB1	AB2	AK2	AK4	AK6	BC1	BC2	BC4	IR1	MB1	ME1	ME2	MI1	MN1	MN2	MT1	NB1	NWT1	NWT2	NY1	ON1	ON3	PA1	PA2	QC3	QC4	SK1	WA4	WY1	YK2	OR2	CO1	UT1	WY1	CA1	CA2	OR1	WA1	WA3
Hap87	1	1																					1																	
Hap88	1	1																					1																	
Hap89	1	1																					1																	
Hap90	1	1																					1																	
Hap91	3	3														1							1						1											
Hap92	3	2																					2				1													
Hap93	1	1																					1																	
Hap94	1	1																														1								
Hap95	1	1																					1																	
Hap96	2	1																					2																	
Hap97	1	1																					1																	
Hap98	1	1																					1																	
Hap99	1	1																					1																	
Hap100	1	1																														1								
Hap101	1	1	1																																					
Hap102	3	2	1													2																									
Hap103	1	1							1																															
Hap104	1	1							1																															
Hap105	1	1							1																															
Hap106	1	1							1																															
Hap107	1	1							1																															
Hap108	1	1							1																															
Hap109	2	2			1																											1							
Hap110	1	1																														1								
Hap111	1	1																									1													
Hap112	1	1																									1													
Hap113	7	1																									7													
Hap114	2	1																									2													
Hap115	1	1																									1													
Hap116	2	1																									2													
Hap117	6	2												1														5												
Hap118	2	1																									2													
Hap119	1	1																									1													
Hap120	1	1																											1											
Hap121	1	1																											1											
Hap122	1	1																											1											
Hap123	1	1																											1											
Hap124	1	1																						1																
Hap125	3	1																										3												
Hap126	1	1																						1																
Hap127	1	1																						1																
Hap128	1	1																						1																
Hap129	4	2					3																									1							
Hap130	5	2												4														1												
Hap131	2	1																										2												
Hap132	3	2												1														2												
Hap133	1	1																										1												
Hap134	1	1																											1											
Hap135	2	2															1												1											
Hap136	2	2															1												1											
Hap137	1	1																						1																
Hap138	2	1																						2																
Hap139	6	3																				2		1		3														
Hap140	1	1																						1																
Hap141	1	1																						1																
Hap142	1	1																						1																
Hap143	1	1																						1																
Hap144	1	1																						1																
Hap145	1	1																						1																
Hap146	2	2									1													1																
Hap147	1	1																						1																
Hap148	1	1										1																												
Hap149	1	1										1																												
Hap150	1	1																										1											
Hap151	1	1																										1											
Hap152	5	2											4	1																										
Hap153	27	1																																	27				
Hap154	2	1											2																										
Hap155	7	1																																	7				
Hap156	3	1													3																								
Hap157	1	1													1																								
Hap158	1	1													1																								
Hap159	4	2													1	3																							
Hap160	2	2													1	1																							
Hap161	1	1												1																									
Hap162	1	1												1																									
Hap163	1	1												1																									
Hap164	1	1												1																									
Hap165	1	1												1																									
Hap166	3	1														3																							
Hap167	1	1														1																							
Hap168	1	1														1																							
Hap169	1	1														1																							
Hap170	1	1					1																																
Hap171	1	1														1																							
Hap172	1	1														1																							
Hap173	1	1														1																							
Hap174	1	1														1																							
Hap175	1	1														1																							
Hap176	2	2				1	1																																
Hap177	6	1																												6									
Hap178	6	1																												6									
Hap179	1	1					1																																
Hap180	1	1					1																																
Hap181	2	1					2																																

			BOREAL																														OR2	CO1	UT-WY		PACNW				WA3
			AB1	AB2	AK2	AK4	AK6	BC1	BC2	BC4	IR1	MB1	ME1	ME2	MI1	MN1	MN2	MT1	NB1	NWT1	NWT2	NY1	ON1	ON3	PA1	PA2	QC3	QC4	SK1	WA4	WY1	YK2	OR2	CO1	UT1	WY1	CA1	CA2	OR1	WA1	WA3
Hap182	6	1											6																											
Hap183	4	1											4																											
Hap184	1	1											1																											
Hap185	2	1																				2																		
Hap186	1	1																				1																		
Hap187	6	1																																6						
Hap188	8	1																																8						
Hap189	2	1																																2						
Hap190	6	1																																6						
Hap191	1	1											1																											
Hap192	5	1																																5						
Hap193	3	1																																3						
Hap194	3	1																																3						
Hap195	9	1																																9						
Hap196	4	1			4																																			
Hap197	2	1			2																																			
Hap198	1	1			1																																			
Hap199	1	1															1																							
Hap200	1	1															1																							
Hap201	2	1															2																							
Hap202	1	1															1																							
Hap203	1	1															1																							
Hap204	1	1															1																							
Hap205	1	1															1																							
Hap206	1	1															1																							
Hap207	1	1															1																							
Hap208	9	1																																		9				
Hap209	1	1																																		1				
Hap210	1	1																																		1				
Hap211	10	1																																		10				
Hap212	3	1																																		3				
Hap213	1	1																																		1				
Hap214	1	1																																		1				
Hap215	8	1																																		8				
Hap216	3	1																																3						
Hap217	1	1																																1						
Hap218	2	1																																2						
Hap219	1	1																																		1				
Hap220	9	1																																9						
Hap221	1	1			1																																			
Hap222	2	1			2																																			
Hap223	1	1			1																																			
Hap224	2	1			2																																			
Hap225	1	1																							1															
Hap226	2	2																				1			1															
Hap227	1	1																							1															
Hap228	1	1																							1															
Hap229	1	1																							1															
Hap230	1	1																							1															
Hap231	15	1																																	15					
Hap232	2	1																																	2					
Hap233	1	1																																	1					
Hap234	4	1																																	4					
Hap235	2	1																																	2					
Hap236	1	1																																	1					
Hap237	10	2																2												8											
Hap238	1	1																																1						
Hap239	4	1																						4																
Hap240	1	1																						1																
Hap241	1	1																						1																
Hap242	2	1																						2																
Hap243	1	1																																1						
Hap244	1	1																												1										
Hap245	1	1																														1								
Hap246	1	1																				1																		
Hap247	1	1														1																								
Hap248	1	1																														1								
Hap249	2	1																														2								
Hap250	1	1																														1								
Hap251	2	1																														2								
Hap252	1	1																														1								
Hap253	1	1																														1								
Hap254	1	1																				1																		
Hap255	1	1																				1																		
Hap256	1	1											1																											
Hap257	2	1											2																											
Hap258	14	1																14																						
Hap259	10	1																10																						
Hap260	7	1																7																						
Hap261	9	1																9																						
Hap262	6	1																																		6				
Hap263	15	1																15																						
Hap264	2	1																2																						
Hap265	3	1																3																						
Hap266	5	1																5																						
Hap267	3	1																												3										
Hap268	4	1																												4										
Hap269	1	1																												1										
Hap270	1	1																												1										
Hap271	13	1																13																						
Hap272	2	1																2																						
Hap273	1	1																1																						
Hap274	1	1																1																						
Hap275	2	1																2																						
Hap276	1	1																1																						

			BOREAL																															OR2	CO1	UT-WY		PACNW				WA3
			AB1	AB2	AK2	AK4	AK6	BC1	BC2	BC4	IR1	MB1	ME1	ME2	MI1	MN1	MN2	MT1	NB1	NWT1	NWT2	NY1	ON1	ON3	PA1	PA2	QC3	QC4	SK1	WA4	WY1	YK2	OR2	CO1	UT1	WY1	CA1	CA2	OR1	WA1	WA3
Hap277	1	1																1																							
Hap278	9	1																9																							
Hap279	1	1																1																							
Hap280	1	1																1																							
Hap281	7	1																7																							
Hap282	1	1																1																							
Hap283	1	1																1																							
Hap284	1	1											1																												
Hap285	2	1																																2							
Hap286	1	1																																1							
Hap287	1	1											1																												
Hap288	1	1																							1																
Hap289	1	1			1																																				
Hap290	1	1														1																									
Hap291	1	1														1																									
Hap292	1	1														1																									
Hap293	3	1																3																							
Hap294	2	1																2																							
Hap295	6	2														1		5																							
Hap296	1	1																1																							
Hap297	2	1																2																							
Hap298	1	1																1																							
Hap299	2	1																2																							
Hap300	1	1																1																							
Hap301	1	1																																		1					
Hap302	2	1																																		2					
Hap303	2	1																																		2					
Hap304	3	1				3																																			
Hap305	3	1				3																																			
Hap306	1	1				1																																			
Hap307	1	1				1																																			
Hap308	1	1														1																									
Hap309	1	1			1																																				
Hap310	3	1									3																															
Hap311	3	1									3																															
Hap312	1	1									1																															
Hap313	1	1														1																									
Hap314	1	1														1																									
Hap315	1	1														1																									
Hap316	1	1														1																									
Hap317	2	1														2																									
Hap318	1	1														1																									
Hap319	1	1														1																									
Hap320	1	1														1																									
Hap321	2	1																															2								
Hap322	1	1																															1								
Hap323	1	1																											1												
Hap324	3	1																															3								
Hap325	4	1																															4								
Hap326	1	1																															1								
Hap327	3	1																															3								
Hap328	2	1																															2								
Hap329	1	1																															1								
Hap330	1	1																												1											
Hap331	2	1																																2							
Hap332	14	2																																			9	5				
Hap333	3	1																																			3					
Hap334	3	1																																							3
Hap335	1	1																																							1
Hap336	1	1																																							1
Hap337	1	1																																							1
Hap338	3	2																																				2	1			
Hap339	7	1																																						7	
Hap340	1	1																																						1	
Hap341	2	1																																						2	
Hap342	3	1																																						3	
Hap343	5	1																											5												
Hap344	1	1																																						1	
Hap345	4	1																											4												
Hap346	1	1																																							1
Hap347	8	1																																						8	
Hap348	4	1																																						4	
Hap349	1	1																																						1	
Hap350	1	1																																							1
Hap351	1	1																																							1
Hap352	1	1																											1												
Hap353	1	1																																						1	
Hap354	1	1																																						1	
Hap355	1	1																																						1	
Hap356	1	1																																					1		
Hap357	1	1																																					1		
Hap358	6	1																																					6		
Hap359	4	1																																					4		
Hap360	1	1																																					1		
Hap361	4	1																																					4		
Hap362	1	1																																					1		
Hap363	3	1																																					3		
Hap364	2	1																																					2		
Hap365	3	1																																					3		
Hap366	2	1																																					2		
Hap367	1	1																																					1		
Hap368	1	1																																					1		
Hap369	1	1																																					1		



Appendix 2.5
Geographic distribution of D-loop haplotype diversities (H_d). Haplotype diversities for sampled populations range from 0.12 to 1.0. Circles are proportional to haplotype diversity divided into 10 evenly spaced categories from 0.41 to 1.0. Vancouver, British Columbia (BC1) population had unusually low haplotype diversity (H_d = 0.12) and is represented with "X"—the next highest haplotype diversity (0.41) was used to set the lower limit for haplotype diversity categories in this figure. Thick blue lines separate six mtDNA haplotype lineages identified by SAMOVA.

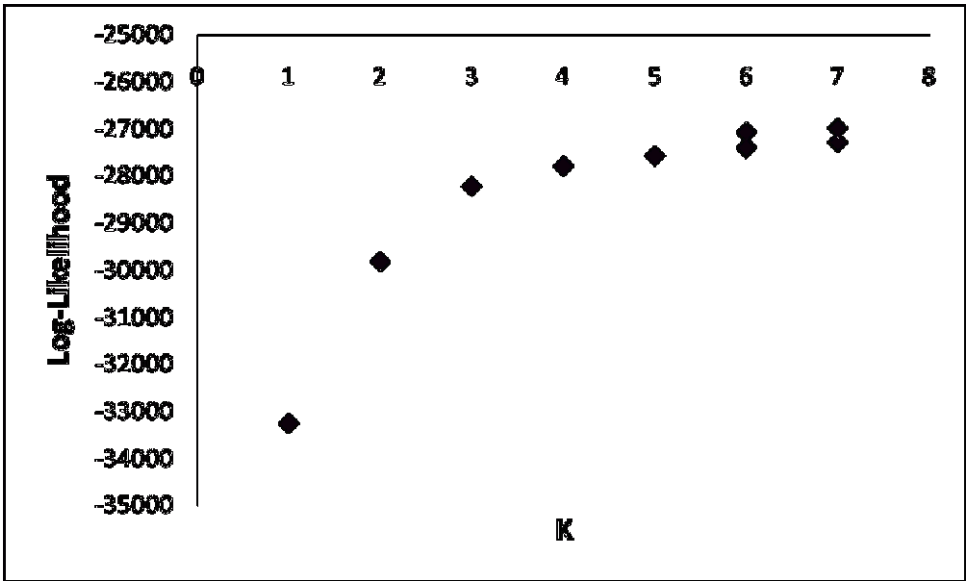


Appendix 2.6
Geographic distribution of D-loop nucleotide diversities (π). Nucleotide diversities for sampled populations range from 0.0003 to 0.0681. Circles are proportional to nucleotide diversity divided into 10 evenly spaced categories from 0.0003 to 0.0199. Northern Washington (WA4) population had unusually high nucleotide diversity (π = 0.0681) and is represented with "X"—the next lowest nucleotide diversity (π = 0.0199) was used to set the upper limit for nucleotide diversity categories in this figure. Thick blue lines separate six mtDNA haplotype lineages identified by SAMOVA.

	AB1	AB2	AK2	AK4	AK6	BC1	BC2	BC4	IR1	MB1	ME1	ME2	MI1	MN1	MN2	MT1	NB1	NWT1	NWT2	NY1	ON1	ON3	PA1	PA2	QC3	QC4	SK1	WA4	WV1	YK2	OR2	CO1	UT1	WY1	CA1	CA2	OR1	WA1	WA3	
AB1	4.27	0.91	0.84	1.25	1.20	4.37	1.93	1.87	2.31	0.48	2.88	2.06	0.39	0.55	0.81	1.08	2.15	0.57	0.81	1.55	1.22	1.02	2.07	1.75	1.59	2.93	1.06	9.08	1.81	1.49	23.30	25.09	20.79	22.48	62.05	61.11	58.03	59.49	63.53	AB1
AB2	0.09	9.03	1.63	1.34	2.18	4.22	0.79	0.99	0.46	0.34	1.23	0.97	1.26	0.14	0.37	0.48	0.76	0.38	0.31	0.57	0.37	0.02	1.68	0.83	0.53	1.17	0.00	7.08	1.02	1.87	21.21	23.91	19.94	22.51	60.30	59.08	54.79	56.60	60.39	AB2
AK2	0.18	0.22	3.81	0.56	0.34	5.01	2.50	2.11	3.04	1.14	3.33	2.62	1.97	1.70	2.11	1.43	2.78	0.46	1.15	2.40	2.31	1.81	3.46	2.78	2.03	3.25	1.84	10.06	2.70	0.21	23.90	25.38	21.99	23.80	64.16	63.17	59.79	61.31	65.17	AK2
AK4	0.19	0.15	0.12	5.88	0.89	3.82	1.29	1.14	2.25	0.96	2.74	1.91	2.01	1.68	2.04	1.31	2.08	0.45	1.07	1.81	2.23	1.63	3.46	2.34	1.06	2.50	1.55	8.77	2.34	0.86	23.27	25.45	22.13	24.64	64.16	62.99	58.74	60.65	64.04	AK4
AK6	0.22	0.22	0.08	0.14	4.32	4.95	3.08	2.45	3.90	1.48	4.15	3.09	2.19	2.17	2.54	1.98	3.28	0.88	1.61	2.94	2.83	2.27	4.13	3.35	2.51	3.79	2.39	10.52	3.41	0.50	24.64	25.82	22.95	24.91	63.45	62.60	58.84	60.64	65.00	AK6
BC1	0.75	0.47	0.67	0.58	0.76	0.12	3.41	3.62	5.43	4.44	6.71	5.83	4.90	4.87	5.31	4.66	5.93	4.26	4.35	5.53	5.35	4.68	6.38	5.98	3.19	6.61	4.79	9.70	5.80	5.36	24.97	27.21	23.54	25.24	61.43	60.49	57.28	58.59	62.83	BC1
BC2	0.30	0.11	0.38	0.21	0.41	0.55	4.51	0.25	0.86	1.00	1.87	1.35	2.30	1.53	1.60	1.13	1.51	0.99	0.96	1.24	1.73	1.22	3.35	1.67	0.24	1.91	1.06	6.51	1.71	2.76	22.00	25.65	20.81	23.70	61.20	59.79	55.26	57.32	60.79	BC2
BC4	0.25	0.10	0.33	0.16	0.31	0.62	0.06	6.70	0.81	1.00	2.23	1.54	2.31	1.67	1.79	1.39	1.54	0.47	0.76	1.55	2.16	1.44	3.74	1.93	0.87	1.99	1.37	6.78	2.08	2.46	22.46	25.01	21.54	23.62	59.62	58.55	54.18	56.40	60.56	BC4
IR1	0.28	0.05	0.41	0.26	0.40	0.70	0.16	0.10	7.77	0.81	1.25	0.90	2.20	1.05	0.87	1.17	0.69	0.72	0.78	0.72	1.24	0.87	2.57	1.00	1.14	1.22	0.48	7.96	1.38	3.32	21.27	25.09	20.40	22.78	63.00	61.63	56.75	58.88	61.90	IR1
MB1	0.07	0.04	0.21	0.14	0.21	0.62	0.17	0.13	0.11	6.36	1.75	0.95	0.56	0.25	0.03	0.55	1.09	0.06	0.12	0.65	0.77	0.29	1.81	0.84	1.09	1.63	0.13	8.54	1.11	1.67	22.29	24.84	20.43	22.18	63.49	62.34	58.33	60.15	64.01	MB1
ME1	0.31	0.15	0.38	0.30	0.40	0.58	0.24	0.25	0.16	0.21	6.59	0.06	3.01	1.73	1.88	1.62	0.26	1.72	1.65	0.22	0.99	1.14	3.08	0.40	1.38	0.23	1.41	8.28	0.93	3.42	20.09	22.81	19.43	22.81	63.77	62.28	56.33	59.32	62.85	ME1
ME2	0.28	0.10	0.38	0.24	0.37	0.72	0.22	0.20	0.12	0.13	0.01	6.10	1.97	1.12	1.11	1.25	0.00	1.15	1.04	0.00	0.86	0.76	2.68	0.10	1.15	0.09	0.95	8.29	0.56	2.97	21.19	22.76	19.95	23.08	64.14	62.79	56.89	59.97	64.23	ME2
MI1	0.08	0.12	0.34	0.27	0.34	0.78	0.34	0.29	0.27	0.08	0.32	0.27	4.32	0.47	0.43	1.59	2.14	1.25	1.18	1.45	1.11	0.90	1.87	1.41	2.23	2.78	1.04	9.88	1.84	2.67	23.66	24.83	22.12	23.74	63.39	62.56	59.07	60.56	64.91	MI1
MN1	0.06	0.02	0.22	0.19	0.23	0.48	0.19	0.18	0.12	0.03	0.20	0.13	0.04	7.65	0.00	0.67	1.06	0.65	0.39	0.62	0.41	0.00	1.54	0.61	1.37	1.55	0.00	8.74	1.03	2.10	21.59	24.31	20.52	22.61	62.69	61.72	57.72	59.40	63.09	MN1
MN2	0.10	0.04	0.30	0.22	0.27	0.60	0.22	0.18	0.09	0.00	0.22	0.13	0.05	0.00	8.73	1.06	1.24	0.80	0.58	0.66	0.35	0.19	1.58	0.55	1.69	1.70	0.00	8.82	0.96	2.55	22.29	24.23	20.67	22.34	62.54	61.53	57.55	59.37	63.28	MN2
MT1	0.09	0.05	0.14	0.12	0.17	0.35	0.11	0.12	0.11	0.05	0.16	0.11	0.13	0.07	0.10	9.09	1.22	0.28	0.30	0.91	1.08	0.38	2.43	1.16	0.90	1.65	0.25	7.58	1.60	1.74	20.43	24.19	19.05	21.98	61.78	60.46	56.17	57.87	61.24	MT1
NB1	0.28	0.09	0.37	0.26	0.37	0.64	0.23	0.20	0.11	0.15	0.04	0.00	0.27	0.13	0.16	0.12	6.02	1.05	0.97	0.00	0.80	0.66	2.35	0.32	1.10	0.02	0.79	8.34	0.69	2.96	19.92	22.67	19.02	21.83	63.57	62.23	56.32	59.34	62.80	NB1
NWT1	0.08	0.04	0.12	0.07	0.13	0.61	0.17	0.06	0.08	0.01	0.21	0.14	0.16	0.08	0.08	0.03	0.14	8.47	0.00	0.92	1.26	0.60	2.58	1.32	0.93	1.61	0.37	7.71	1.40	0.89	20.94	24.02	19.44	21.72	61.29	60.14	55.98	57.83	61.78	NWT1
NWT2	0.09	0.03	0.18	0.13	0.18	0.50	0.14	0.08	0.09	0.01	0.19	0.12	0.13	0.05	0.07	0.03	0.12	0.00	8.16	0.82	1.08	0.26	2.03	0.99	0.95	1.49	0.09	7.97	1.34	1.47	21.37	24.32	20.46	22.14	62.22	61.12	57.04	58.75	62.42	NWT2
NY1	0.22	0.06	0.35	0.23	0.36	0.68	0.20	0.20	0.10	0.10	0.03	0.00	0.21	0.07	0.08	0.08	0.00	0.12	0.10	5.94	0.41	0.34	1.83	0.00	1.02	0.24	0.47	8.38	0.44	2.62	20.44	23.50	19.83	22.76	64.05	62.74	57.28	60.00	63.46	NY1
ON1	0.14	0.04	0.31	0.24	0.29	0.58	0.23	0.21	0.13	0.09	0.13	0.10	0.12	0.05	0.04	0.11	0.11	0.13	0.11	0.05	8.54	0.13	1.19	0.40	1.07	1.11	0.53	8.10	0.65	2.38	20.37	22.06	18.43	21.41	60.46	59.32	55.12	56.99	60.96	ON1
ON3	0.10	0.00	0.24	0.17	0.22	0.48	0.16	0.14	0.08	0.03	0.14	0.08	0.08	0.00	0.02	0.04	0.08	0.06	0.03	0.04	0.01	9.27	1.44	0.36	0.97	0.95	0.00	7.69	0.82	1.95	20.43	23.25	19.20	21.31	60.83	59.79	55.61	57.18	61.47	ON3
PA1	0.32	0.17	0.47	0.39	0.49	0.80	0.43	0.41	0.31	0.24	0.33	0.34	0.30	0.17	0.19	0.20	0.30	0.29	0.22	0.26	0.14	0.15	4.36	2.00	3.01	3.03	1.79	10.42	2.12	3.70	23.25	25.67	21.42	23.30	63.70	62.68	59.19	60.68	64.06	PA1
PA2	0.21	0.09	0.36	0.26	0.35	0.65	0.24	0.21	0.12	0.11	0.06	0.01	0.17	0.07	0.06	0.11	0.05	0.14	0.11	0.00	0.05	0.04	0.24	7.60	1.55	0.51	0.52	8.93	0.65	2.94	20.68	23.80	20.32	22.87	64.11	62.94	57.60	60.27	63.75	PA2
QC3	0.19	0.06	0.29	0.14	0.28	0.45	0.04	0.11	0.14	0.14	0.17	0.14	0.25	0.15	0.18	0.09	0.14	0.11	0.11	0.13	0.12	0.11	0.32	0.18	7.13	1.50	0.98	6.36	1.57	2.14	21.30	24.13	19.41	22.80	60.23	58.70	54.10	56.00	59.52	QC3
QC4	0.36	0.14	0.42	0.31	0.42	0.70	0.28	0.25	0.17	0.22	0.03	0.02	0.35	0.18	0.20	0.15	0.00	0.21	0.18	0.04	0.14	0.11	0.37	0.08	0.19	5.54	1.22	8.62	0.83	3.29	21.27	22.63	19.74	22.95	64.23	62.81	56.29	59.89	63.90	QC4
SK1	0.14	0.00	0.29	0.19	0.28	0.63	0.18	0.15	0.05	0.02	0.18	0.12	0.13	0.00	0.00	0.02	0.11	0.04	0.01	0.07	0.06	0.00	0.22	0.06	0.12	0.17	8.60	7.87	1.13	2.12	20.68	24.88	20.00	22.48	62.21	61.01	56.52	58.34	61.84	SK1
WA4	0.24	0.23	0.36	0.26	0.28	0.31	0.25	0.18	0.20	0.25	0.33	0.23	0.25	0.32	0.25	0.37	0.28	0.20	0.25	0.25	0.24	0.25	0.27	0.26	0.22	0.27	0.21	31.81	8.47	9.83	23.46	24.30	19.35	23.56	29.39	27.76	24.69	25.96	34.22	WA4
WV1	0.27	0.12	0.39	0.30	0.42	0.72	0.27	0.27	0.19	0.16	0.12	0.09	0.28	0.12	0.12	0.14	0.11	0.19	0.16	0.07	0.09	0.09	0.31	0.09	0.20	0.13	0.15	0.25	5.08	2.95	22.35	23.06	19.83	22.41	63.29	62.09	57.09	59.90	63.30	WV1
YK2	0.20	0.21	0.04	0.12	0.07	0.57	0.34	0.28	0.34	0.21	0.35	0.32	0.31	0.23	0.27	0.17	0.33	0.13	0.18	0.30	0.26	0.21	0.39	0.31	0.25	0.35	0.25	0.34	0.33	6.22	23.43	25.66	21.94	23.96	62.71	61.67	58.18	59.58	63.55	YK2
OR2	0.89	0.84	0.89	0.88	0.89	0.92	0.87	0.88	0.87	0.87	0.83	0.87	0.89	0.83	0.86	0.75	0.85	0.86	0.85	0.86	0.85	0.83	0.89	0.86	0.85	0.87	0.86	0.67	0.88	0.86	2.71	19.16	14.04	15.94	71.04	68.68	63.04	64.96	62.46	OR2
CO1	0.81	0.75	0.84	0.80	0.81	0.89	0.82	0.79	0.78	0.79	0.77	0.78	0.80	0.77	0.76	0.73	0.78	0.76	0.76	0.78	0.74	0.74	0.81	0.77	0.78	0.78	0.77	0.51	0.79	0.80	0.84	6.86	18.03	18.13	64.85	63.98	59.49	62.55	64.22	CO1
UT1	0.92	0.82	0.90	0.88	0.92	0.97	0.88	0.90	0.89	0.88	0.81	0.89	0.92	0.80	0.85	0.71	0.85	0.87	0.83	0.88	0.83	0.81	0.91	0.86	0.84	0.87	0.87	0.52	0.89	0.85	0.86	0.84	1.16	5.19	59.23	57.55	53.09	55.68	57.83	UT1
WY1	0.90	0.87	0.90	0.90	0.91	0.93	0.90	0.90	0.90	0.89	0.86	0.90	0.91	0.85	0.88	0.77	0.88	0.89	0.87	0.89	0.87	0.86	0.90	0.89	0.88	0.89	0.89	0.71	0.90	0.88	0.87	0.86	0.72	2.26	62.09	61.67	59.83	60.36	63.53	WY1
CA1	0.97	0.91	0.96	0.95	0.97	1.00	0.95	0.95	0.95	0.95	0.92	0.96	0.97	0.91	0.93	0.88	0.94	0.94	0.92	0.95	0.93	0.91	0.97	0.94	0.93	0.95	0.94	0.55	0.96	0.93	0.97	0.94	0.98	0.97	0.42	0.29	4.26	2.20	23.44	CA1
CA2	0.95	0.89	0.95	0.93	0.95	0.99	0.94	0.92	0.92	0.92	0.91	0.93	0.95	0.90	0.90	0.87	0.92	0.91	0.90	0.93	0.90	0.89	0.95	0.91	0.91	0.93	0.91	0.50	0.94	0.92	0.96	0.92	0.98	0.96	0.24	2.45	2.58	1.06	21.19	CA2
OR1	0.89	0.87	0.91	0.89	0.90	0.92	0.90	0.88	0.88	0.89	0.89	0.89	0.89	0.88	0.88	0.86	0.89	0.88	0.88	0.89	0.88	0.87	0.90	0.88	0.88	0.89	0.88	0.57	0.89	0.89	0.94	0.89	0.92	0.94	0.42	0.26	7.50	2.36	18.24	OR1
WA1	0.90	0.87	0.91	0.90	0.90	0.92	0.90	0.88	0.88	0.89	0.89	0.89	0.90	0.89	0.88	0.87	0.89	0.88	0.88	0.89	0.88	0.87	0.90	0.89	0.88	0.90	0.88	0.57	0.90	0.90	0.94	0.89	0.92	0.94	0.26	0.10	0.24	7.67	19.19	WA1
WA3	0.94	0.89	0.94	0.93	0.94	0.98	0.93	0.92	0.92	0.92	0.91	0.93	0.94	0.90	0.90	0.87	0.92	0.91	0.90	0.93	0.90	0.89	0.94	0.91	0.91	0.93	0.91	0.56	0.93	0.92	0.96	0.92	0.97	0.96	0.93	0.87	0.73	0.74	3.94	WA3
	AB1	AB2	AK2	AK4	AK6	BC1	BC2	BC4	IR1	MB1	ME1	ME2	MI1	MN1	MN2	MT1	NB1	NWT1	NWT2	NY1	ON1	ON3	PA1	PA2	QC3	QC4	SK1	WA4	WV1	YK2	OR2	CO1	UT1	WY1	CA1	CA2	OR1	WA1	WA3	

Appendix 2.7

F_{ST} (below diagonal) and Pi_{XY} (corrected average pairwise differences, above diagonal) for each population pair. Populations are grouped into six haplotype lineages identified by SAMOVA.



Appendix 2.8
Program STRUCTURE results on microsatellite data. Log-likelihood of genetic clusters plotted against K, number of clusters. The most likely number of genetic clusters is K = 3, where graph reaches an inflection. With K = 3, the estimated membership of every population in its most likely genetic cluster is > 90%, when averaged across all individuals in the population.

CHAPTER 3

IS SYNCHRONY DETECTABLE WHEN TIME SERIES ARE SHORT AND DATA LESS THAN PERFECT?

ABSTRACT

Synchrony is a widespread phenomenon that has been documented for species across many taxa. Identifying patterns and causes of synchrony is fundamental to understanding the factors that influence animal numbers and large-scale population processes, with broad implications for conserving and managing ecological systems. Two statistical issues hamper current synchrony research. First, different metrics define synchrony in subtly different ways and therefore generate different numeric or qualitative estimates when applied to the same data, making it challenging to compare results across studies. Second, we have little understanding of how real-world data limitations, such as short time series length and sampling error, interact with data smoothing and cyclicity to affect synchrony estimates. I conducted a simulation study of five commonly used synchrony metrics (Pearson, Spearman, Kendall, Percent Match, and Symbolic) to address these issues. Synchrony estimates were highly correlated under most simulated scenarios. The Symbolic metric was monotonically, but not linearly, correlated with the other metrics, suggesting results can be at least qualitatively compared across studies applying these metrics. In comparisons of the four most similar metrics (i.e., excluding the Symbolic metric), the Kendall metric exhibited

the lowest standard deviation and bias, and had the highest statistical power, regardless of time series length (15 – 100 years), sampling error (σ=0-30% of mean abundance), or data smoothing, for both cyclic and non-cyclic time series. For all metrics, increasing time series from 15 to 25 years considerably reduced variability in synchrony estimates. Data smoothing greatly reduced bias in synchrony estimates, especially for cyclic time series, when sampling error was high (15-30%). However, smoothing also increased variability of estimates in all cases, most noticeably for time series < 50 years in length. Overall, the correspondence of the metrics is good news for comparing studies that apply different synchrony metrics. For the types of dynamics and measures of performance examined in this study, the Kendall metric performed best under most data scenarios, suggesting it should perform well in future studies of synchrony.

INTRODUCTION

Population ecology has been marked by periods of rapid advances in theory, inspired and grounded by empirical observations and experimental research. In a classic example, Charles Elton's (1924) pioneering paper on the cycles of Norwegian lemmings *(Lemmus lemmus)*—nearly coincident with Lotka's (1925) and Volterra's (1926) mathematical formulations for predator-prey cycles—initiated substantial advances in our understanding of factors that regulate animal populations. Subsequent interest in population cycles spurred not only vigorous scientific debates, but also relevant

laboratory and field experiments (e.g., Nicholson 1954; Krebs et al. 1995) and initiation of many long-term studies on cyclic populations (Turchin 2003).

Many species that cycle also exhibit synchrony in these cycles over large spatial scales, i.e., populations increase and decline simultaneously (or with a consistent time lag) across much of the species range. Although this phenomenon of synchrony did not receive the level of attention that population cycles did in the early 1900's, it had long been recognized and anecdotally reported among trappers and fur traders. In 1953, Moran published one of the first statistical analyses and mechanistic explanations of this phenomenon (later dubbed the 'Moran effect'), based on Canada-wide lynx *(Lynx canadensis)* fur return data kept by the Hudson's Bay Company. Interest in synchrony resurged in the 1990's, with growing interest in landscape-level ecological processes, and the recognition that large-scale spatial correspondence in population dynamics characterizes many species of high conservation and economic interest. In the two decades since, the ecological literature has been rife with examinations of synchronous dynamics in forest insects, marine fishes, oaks and other trees, viruses, birds, small and medium-sized boreal and arctic mammals, and large mammals including caribou (*Rangiffer tarandus)* and muskox (*Ovibos moschatus*)(reviewed in Liebhold et al. 2004a).

Despite theoretical, computational, and statistical advances in synchrony research, a substantial gap persists between development of synchrony theory and its application to real populations. A primary obstacle to application of this theory is the

absence of a standard operational definition of what constitutes synchrony. Most synchrony metrics identify two time series as fully synchronized when they are identical (but see Cazelles 2004 for the Symbolic metric). However, when time series are not identical, metrics may generate different estimates of synchrony when applied to the same data, depending on how each metric defines synchrony (Buonaccorsi et al. 2001).

Synchrony metrics can be broadly categorized as those that measure correlation synchrony, those that measure phase synchrony, and hybrid metrics. The Pearson correlation is a commonly applied measure of correlation synchrony that quantifies the linear relationship between two time series. Two time series correlated in relative, but not absolute, yearly abundances may exhibit low synchrony as measured by the Pearson correlation, but they may be highly synchronized based on nonparametric measures of correlation synchrony (e.g., Spearman or Kendall correlations) that simply ask if years rank in a similar order when sorted by abundance (Zar 1999).

Hybrid metrics employ a more general definition of synchrony than metrics of correlation synchrony. For example, Buonaccorsi et al. (2001) proposed a hybrid metric that quantifies the proportion of times two populations increase or decrease simultaneously, regardless of degree of change. Studies have also defined synchrony as the proportion of times two cyclic populations peak or trough simultaneously (e.g., Krebs et al. 2002).

Within the past decade, a class of synchrony metrics based on the concept of phase synchrony has emerged in the ecological literature, appropriate for (but not

restricted to) quantifying synchrony of systems with dynamics characterized as Uniform Phase evolution with Chaotic Amplitudes (UPCA; Blasius and Stone 2000; Grenfell et al. 2001; Cazelles and Stone 2003; Cazelles 2004; Liebhold et al. 2004a). This type of oscillation has been observed in many cyclic systems, including Canada lynx (Ranta et al. 1997), ruffed grouse (*Bonasa umbellus*) and snowshoe hares (*Lepus americanus*; Keith 1963), and several species of voles and lemmings (Krebs and Myers 1942), which cycle with relatively constant frequency but erratic peak abundances among populations. Phase synchrony metrics broadly define synchrony as oscillation in rhythm, with or without a time lag, and regardless of the amplitude of peaks. For example, two populations with a consistent time lag between peaks and very different peak amplitudes each cycle could exhibit the same degree of phase synchrony as two populations that always peak simultaneously with equal peak amplitude. In contrast, correlation or hybrid synchrony metrics would consider the latter time series more synchronized than the former.

Choice of synchrony metric primarily depends on how a researcher chooses to define synchrony, but may also be limited by the type and quality of time series data. For example, a researcher interested in whether forest insect outbreaks co-occur but not interested in correspondence of population abundances between major outbreaks could estimate synchrony as the proportion of times outbreaks occur simultaneously. However, if the data consist of a few 25-year time series with an average of 2–3 outbreaks per population, data may be insufficient for applying this metric. Frequently,

different researchers apply subtly different metrics to studies of the same species. In these cases, it may be unclear how to interpret conflicting estimates among metrics or how to compare results across studies.

In addition to differing interpretations of what constitutes “synchrony”, we have little information on how real-world data limitations affect metric performance. Most theoretical synchrony studies have been based on a minimum of 100 years of simulated time series data (but see Cazelles 2004, who used 50 years)—much longer than is available for most species (Cazelles and Stone 2003). Further, it is not clear how sampling error, the common practice of smoothing data prior to analyses, or the level of cyclicity in time series data interact with time series length to affect synchrony estimates and our ability to determine when two populations are significantly synchronized. Addressing such uncertainties would improve our ability to quantify synchrony patterns in real data and, ultimately, our ability to infer synchrony mechanisms from these patterns.

The gap between synchrony theory and application would be reduced through attention to three areas of synchrony research:

1) How comparable are synchrony estimates calculated by different metrics?
2) How do data quality (time series length and sampling error), data smoothing, and cyclicity in time series affect synchrony estimates?
3) How do data quality (time series length and sampling error), data smoothing, and cyclicity in time series affect statistical power of synchrony metrics?

I address these questions in a simulation analysis of five commonly used synchrony metrics. To ground the simulations in natural history, I mimic time series representing the approximately 10-year population cycles of snowshoe hares in their northern species range, and the more erratic fluctuating time series representative of hares in their southern range. A simulation study based on the dynamics of snowshoe hares is widely applicable to a large number of mammalian and avian species that exhibit 10-year cyclic and synchronous dynamics in at least part of their range, including Canada lynx (Ranta et al. 1997), mink (Mustela vison; Holmengen et al. 2008), muskrat (Ondatra zibethicus; Erb et al. 2000), marten (Martes americana; Cowan 1938), rock ptarmigan (Lagopus muta; Gundmundsson 1958), and ruffed and red grouse (Lagopus lagopus scoticus; Moss et al. 1996). I develop recommendations for comparing synchrony estimates across studies that use different metrics and for interpreting results in light of known data issues.

METHODS

Measuring Synchrony

Synchrony analysis begins with the collection of population data on a regular (usually, yearly) basis from multiple sites distributed across a study region. Populations can be synchronized in abundance, reproductive timing, or mortality. A single measure of synchrony can be calculated across all sites. However, for identification of geographic patterns in synchrony and causal mechanisms, pairwise synchrony estimates are usually

calculated. The spatial scale of synchrony and the locations of geographic 'breaks' in synchrony are compared against expectations for hypothesized mechanisms. My simulations focus on methods for estimating pairwise synchrony in population abundance, which is the most common form of synchrony examined in animals.

I identified five synchrony metrics for comparison, including the most frequently used correlation metrics (Pearson, Spearman, and Kendall correlation), a hybrid metric (Percent Match), and a phase synchrony metric (Symbolic; Cazelles and Stone 2003). Attributes of each metric are outlined in Table 3.1.

Synchrony in population abundances can be measured on raw abundance counts, but is more typically applied to first-differenced log-transformed abundances, or to residuals after fitting a linear regression or autoregressive model to the data. Synchrony on first-differenced log-transformed abundances [$\log(N_t) - \log(N_{t-1})$, where N_t = abundance at time t] measures synchrony in population growth rates rather than in abundances themselves. In this study I was interested in examining synchrony in population growth rates. Therefore, I applied the standard correlation measures of synchrony (Pearson, Spearman, and Kendall correlation) to first-differenced log-transformed abundances. The Percent Match and Symbolic metrics transform data to growth or cyclic phase categories prior to analyses, so were applied to raw abundance data.

Data Simulations

Each of the five metrics applies a different operational definition of synchrony. Because my goal was to compare synchrony estimates based on different metrics, and their sensitivities to quality issues common in real time series data, I did not compare synchrony metrics against a 'standard' definition of synchrony that might favor some metrics over others. Instead, I compared the performance of each metric against a baseline of its own performance on 'ideal' data—1000-year data simulated without sampling error and without data smoothing.

I selected four actual snowshoe hare time series representative of northern (cyclic) and southern (non-cyclic) populations to serve as base models for generating time series for simulations (Fig. 3.1). For each of the four simulations (representing two cyclic and two non-cyclic data sets) and five synchrony metrics, I calculated the standard deviation and bias of synchrony estimates (relative to baseline ideal synchrony estimates) for all combinations of these factor-levels:

1) Time series length: 15, 25, 35, 50, 75, 100 years
2) Sampling error: 0, 15%, 30% error
3) Data treatment: smoothed vs. unsmoothed prior to synchrony analysis

For each base model, separate simulations were conducted in R (http://cran.r-project.org/; Appendix 3.8) as follows:

1) Data smoothing is a subjective visualization technique commonly used in time series and synchrony analyses to minimize potential effects of anticipated but

unknown sampling error in data. I applied loess (locally weighted regression) smoothing to the base data, using a smoothing parameter of 7 surrounding data points (span = 7 / length of time series) with normally distributed weighting, to arrive at the smoothed estimate for each year [R function: loess()]. On average, this smoothing parameter seemed to best capture most fluctuations in the time series data while smoothing over minor potential 'noise'. I used the same loess parameters for all data smoothing in these simulations.

2) I fit an autoregressive (AR) time series model to the smoothed base data, using the Akaike Information Criterion to choose the order of the AR model [R function: ar()].
3) One thousand pairs of 1000-year time series were simulated from an AR model using the same model coefficients determined for the base data [R function: arima.sim()]. These time series pairs were the null time series for testing performance of metrics in a scenario of 'no synchrony'. The paired time series were generated independently from the same AR model, to mimic the level of synchrony that might be observed due to chance between populations with similar endogenous dynamics, but with no specific synchronizing mechanism acting on the populations.
4) I used a Cholesky decomposition to correlate the 1000 time series pairs at a range of values, from Pearson's r = 0.05 to 0.95 [R function: chol()]. Although Pearson's correlation is only one way to measure synchrony between time

series, it seemed reasonable that time series pairs spanning the full range of correlation values as measured by Pearson's coefficient would also span a wide range of correlation values as measured by each of the other synchrony metrics examined.

5) The 1000 time series pairs correlated by Cholesky decomposition represented an approximately uniform distribution of Pearson synchrony estimates. However, when a different synchrony metric (e.g., Percent Match) is applied to these same data they produce a different distribution of synchrony estimates, because different metrics calculate synchrony in different ways. To evaluate the performance of each metric using data representing a uniform distribution of synchrony values, I selected metric-specific subsets of 500 time series pairs (from the full set of 1000 time series pairs) evenly spanning the range from no synchrony to maximum synchrony, as calculated by each metric. For example, the Percent Match subset of data had approximately 100 time series pairs in each of five categories of Percent Match synchrony estimates ranging from 0.5 (no synchrony) to 1.0 (maximum synchrony). For each metric, the metric-specific subset of 500 time series pairs with error = 0 and data treatment = unsmoothed were considered 'baseline' data for evaluating metric performance.
6) Multiple levels of simulated error were applied to the simulated time series, to generate data for evaluating the influence of sampling error on metric performance. To each data point in each time series, I added an error value

randomly selected from a normal distribution with mean = 0 and standard deviation = y × mean abundance of the time series, where y represented the three levels of simulated sampling error: 0, 0.15, and 0.30.

7) I also evaluated the influence of data smoothing and time series length on metric performance. Data smoothing was simulated by applying a loess smooth as described earlier. The effects of time series length were simulated by truncating the 1000-year time series to specified lengths (15, 25, 35, 50, 75, 100 years), resmoothing the time series for these shorter lengths (when data treatment = smoothed), and recalculating synchrony estimates.
8) Finally, I evaluated the performance of each metric under a scenario of 'no synchrony', by calculating synchrony on the null time series data generated in Step #3. I generated error, applied data smoothing, and truncated time series lengths on the null data pairs as described for the correlated time series pairs.

Evaluation of Results

To determine if synchrony estimates calculated by different metrics are comparable, I examined scatterplots of synchrony estimates for each pair of metrics applied to the same time series data. I calculated Pearson's r on these synchrony estimates, to quantify the linear association for each pair of metrics. The data I used for this analysis were the 1000 pairs of correlated 1000-year time series (Step #4) from which metric-specific data subsets were subsequently chosen. I compared synchrony

estimates for three error levels (0, 0.15, 0.30) and two data treatments (smoothed vs. unsmoothed), for cyclic and non-cyclic time series.

Performance of each metric for different combinations of data quality (time series length and sampling error), data treatment (smoothing versus no smoothing), and cyclicity (cyclic versus non-cyclic) are reported as the standard deviation and bias of synchrony estimates relative to baseline synchrony estimates. For example, for the Kendall metric I evaluated the influence of short (25-year) time series length and moderate (0.15) sampling error on synchrony of unsmoothed, cyclic data by estimating synchrony on the 500 pairs of cyclic time series in the Kendall subset under this scenario. I compared these synchrony estimates to those obtained on the corresponding baseline (1000-year, error = 0, data treatment = unsmoothed) Kendall time series. In this way, the performance of each metric under different data scenarios was compared against its performance on 'ideal' baseline data, rather than against a specific definition of synchrony.

In this study, I define statistical power as the ability of a metric to distinguish (as significant) time series synchronized at a specified level (measured in Step #5 above) versus independently generated null time series (Step #3 above). To evaluate statistical power of each metric under different data scenarios, I used a 10% moving window analysis to calculate the proportion of time series pairs that were significant at $\alpha = 0.05$ for the range of synchrony estimates from 'no synchrony' to 'full synchrony'. To provide a reference level of statistical power for comparison across metrics and data scenarios, I

calculated the minimum baseline synchrony at which 50% of estimated synchrony values were significant at $\alpha = 0.05$. A higher baseline 50% significance level translates to lower statistical power. For example, I applied the following steps to evaluate statistical power of the Spearman metric for unsmoothed, cyclic 25-year time series with sampling error = 0.15:

1) I truncated cyclic, null time series from Step #3 to a length of 25 years and applied a simulated error level of 0.15.
2) I calculated the Spearman synchrony estimate for these 1000 pairs of modified null time series data. I identified the 95% upper CI bound for null data under this data scenario.
3) I identified all correlated time series pairs (Step #5 above) with baseline Spearman synchrony ranging from r = 0 – 0.10 (a 10% window). Because the subset of 500 time series pairs for each metric was chosen to achieve a uniform set of correlations, on average each 10% moving window of baseline synchrony estimates included 50 time series pairs.
4) Within the range of baseline synchrony estimates covered by the 10% window, I calculated the proportion of time series pairs with an estimated synchrony higher than the 95% upper CI bound for null data, for the specific combination of factors evaluated (in this case, unsmoothed, cyclic 25-year time series with sampling error = 0.15). This proportion approximately represented the power of

the Spearman metric when baseline Spearman synchrony was r = 0.05 (the middle value in the 10% window range), for this data scenario.

5) I shifted the moving window by 1% (to a baseline synchrony range of r = 0.01 – 0.11) to calculate power when baseline Spearman synchrony was r = 0.06. I repeated this moving window calculation for the entire range of Spearman synchrony values.

RESULTS

Data Simulations

The simulated time series were visually comparable to the four snowshoe hare time series (two cyclic, two non-cyclic) from which they were modeled (Fig. 3.1). Results were similar using the two cyclic data sets and the two non-cyclic data sets. Therefore, I present results from one cyclic simulation and one non-cyclic simulation.

For all metrics except for the Symbolic metric, Cholesky decomposition was able to simulate correlated baseline time series pairs equally spanning the full range of synchrony from no synchrony to maximum synchrony. For the Symbolic metric, the maximum synchrony of 2.0 is obtained only when time series are identical and each of the four cycle phases (increase, peak, decrease, and trough) occurs equally, i.e., 25% of the time. In this study, the Symbolic synchrony of most Cholesky decomposed baseline time series pairs was less than 1.0 (out of a maximum synchrony of 2.0). As shown in Fig. 3.2 and Appendix 3.1, baseline time series pairs with high Pearson correlation (close to r

= 1.0) had Symbolic synchrony estimates close to 1.0. My inability to generate baseline time series with Symbolic synchrony much greater than 1.0 was likely due to the fact that simulated time series did not have approximately equal representation of the cycle phases. Ten-year cyclic time series are approximately represented by 10% peak phase, 10% trough phase, 40% increase phase, and 40% decrease phase cycles. Therefore, 10-year cyclic time series cannot achieve maximum synchrony, as defined by the Symbolic metric.

Correspondence in Synchrony Estimates Across Metrics

Can synchrony estimates be compared across studies that apply different metrics? To gain a general sense of similarities among metrics, I examined scatterplots and calculated Pearson's correlation on synchrony estimates calculated by different metrics, under various data scenarios. For each comparison in this analysis, synchrony was calculated by two different metrics on the same set of 1000 simulated long (1000-year) time series pairs. High correspondence between estimates for 1000-year data does not necessarily mean high correspondence for shorter time series.

I found very high correlation in synchrony estimates calculated using the three standard correlation metrics (Pearson, Spearman, Kendall), regardless of error rate, data treatment (smoothing versus no smoothing), and cyclicity (cyclic versus non-cyclic) (Pearson's r = 0.98 – 1.00; Table 3.2; Appendix 3.2). The Percent Match metric was also highly correlated with the standard correlation metrics under most scenarios (Pearson's r = 0.83 – 1.00). The Symbolic metric showed moderate to high correspondence in

synchrony estimates with other metrics (Pearson's r = 0.73 – 0.96). Fig. 3.2 and Appendix 3.1 show that the relationship between the Symbolic and other metrics is monotonic, but not linear, under some data scenarios. Percent Match and Symbolic metric correspondence with other metrics dropped quickly as error rate increased from 0.15 to 0.30 for unsmoothed (especially cyclic) data. Data smoothing maintained high levels of correspondence among metrics, regardless of error rate (Pearson's r = 0.90 – 1.00).

Influence of Data Quality, Data Treatment, and Cyclicity on Synchrony Estimates

Among the correlation (Pearson, Spearman, Kendall) and hybrid (Percent Match) synchrony metrics, the Kendall metric exhibited the lowest standard deviations and bias (Fig. 3.3; Table 3.3; Appendix 3.3) and the highest statistical power (Figs. 3.6 and 3.7; Table 3.5; Appendices 3.5, 3.6 and 3.7), under most data scenarios. The Symbolic metric applies a very different definition of synchrony from the other metrics evaluated in this study. It is difficult to interpret what it means for this metric to exhibit larger or smaller standard deviation or bias in its estimates compared to the other metrics in this study. Therefore, I do not compare these qualities of the Symbolic metric against other metrics; I only evaluate the influence of data quality, data treatment, and cyclicity on Symbolic synchrony estimates relative to baseline Symbolic synchrony estimates (below).

Time series length greatly affected the standard deviation of synchrony estimates compared to baseline synchrony (i.e., 'ideal', 1000-year time series with no

error and no data smoothing) for all metrics (Fig. 3.3; Table 3.3; Appendix 3.3). As time series length increased, standard deviation decreased. For all metrics, increasing time series from 15 to 25 years resulted in a relatively large reduction in standard deviation of synchrony estimates. The effects of time series length on bias varied by metric, but in general synchrony estimates for longer time series exhibited lower or similar bias compared to shorter time series. For the Pearson and Kendall metrics, time series length had little effect on synchrony bias. For the Spearman and Percent Match metric, with smoothed data, bias decreased moderately as time series length increased. For the Symbolic metric, bias was greatly influenced by time series length. The Symbolic metric consistently overestimated synchrony with shorter time series (Fig. 3.3 and 3.4; Table 3.3; Appendix 3.3), and this bias was less pronounced for longer time series.

As error increased, standard deviation and bias also increased (Figs. 3.3 and 3.5; Table 3.3; Appendix 3.3). Increasing error led to underestimation of synchrony, particularly at higher synchrony values. Error essentially flattened out the relationship between actual (baseline) and estimated synchrony.

Under all scenarios, smoothing considerably reduced bias when error was present and did not have much effect when error was not present (Figs. 3.3 and 3.5; Table 3.3; Appendix 3.3). Smoothing increased the standard deviation of synchrony estimates with lower error rates and shorter time series (Table 3.3). However, when error rates were high or time series long, smoothing reduced standard deviation of synchrony estimates.

In most cases, standard deviation and bias of synchrony estimates were lower for non-cyclic compared to cyclic time series (Figs. 3.3 and 3.5; Table 3.3; Appendix 3.3). Smoothing was much more effective at reducing standard deviation and bias of synchrony estimates for cyclic compared to non-cyclic data. Bias and variance of the Symbolic metric were similar for cyclic and non-cyclic data.

Influence of Data Quality, Data Treatment, and Cyclicity on Statistical Power

In evaluations of statistical power, the Kendall metric performed best among the correlation and hybrid metrics compared in this study (Figs. 3.6 and 3.7; Table 3.5; Appendices 3.5–3.7). The higher statistical power of this metric across all simulated scenarios is likely due to its lower 95% CI bounds rather than to higher synchrony estimates (Table 3.4; Appendix 3.4). Because the Symbolic metric applies a very different definition of synchrony compared to other metrics in this study, I do not compare statistical power of the Symbolic metric against other metrics. However, I evaluate the influence of data quality, data treatment, and cyclicity on Symbolic statistical power (below).

For all metrics, the effects of time series length, error, and data treatment on statistical power were opposite to the effects of these factors on standard deviation of synchrony estimates. This result makes sense because, across different data scenarios, the statistical power of a metric is negatively correlated with its standard deviation when synchrony = 0. In general, statistical power was positively correlated with time series length and negatively correlated with error. Smoothing reduced statistical power

(especially for cyclic time series) when error rates were low and time series were short, and increased statistical power when error rates were high and time series were long. Cyclic time series exhibited lower statistical power than non-cyclic time series.

DISCUSSION

Identifying patterns and causes of synchrony is fundamental to applied population ecology because it addresses the age-old question of what determines animal numbers. Two persistent questions in synchrony analyses are how to choose among metrics that calculate synchrony in subtly different ways, and how to compare estimates across studies that apply different metrics. This simulation provides insights for choosing and interpreting synchrony metrics.

Correspondence in Synchrony Estimates Across Metrics

Except when error rates were high and data unsmoothed, I found generally high correspondence in the synchrony estimates for the metrics compared, suggesting it may be reasonable to qualitatively compare results across studies employing these different metrics. Although Symbolic synchrony estimates were non-linearly related to estimates of the other metrics in this study, the relationship was monotonic, i.e., the relative ranking of synchrony estimates was consistent across metrics.

Influence of Data Quality, Data Treatment, and Cyclicity on Synchrony Estimates

High correspondence among metrics does not translate to equal performance, when performance is measured as variance and bias in synchrony estimates compared

to baseline ('true') synchrony. Results suggest that in studies examining correlation synchrony of 10-year cyclic species, the Kendall and Percent Match metrics would perform well compared to the Pearson and Spearman metrics, for a broad range of time series lengths, sampling error rates, and data cyclicity. Among the correlation and hybrid synchrony metrics analyzed in this study, the Kendall metric exhibited the lowest bias and standard deviation relative to baseline ('true') synchrony estimates, and had the highest statistical power, under most simulated scenarios. By these same measures, the Percent Match metric typically performed better than or comparable to the Pearson and Synchrony metrics. Among the metrics compared, the Percent Match metric has the added advantage of being applicable to a wide range of data collection methods, including surveys in which respondents are simply asked if population numbers increased or decreased from previous years (e.g., Canada's national Wildlife Enquiry Surveys; Chitty 1948). Therefore, the Percent Match metric may be particularly suitable for synchrony studies that combine qualitative and quantitative time series data.

Both time series length and sampling error greatly affect synchrony estimates. The shortest time series length (15 years) and highest error level (0.30) used in this study produced synchrony estimates that were largely incapable of distinguishing between high and low synchrony. When time series were short, the variability around the actual (baseline) synchrony estimate was very high for all metrics. When error was high, estimated synchrony was always biased low regardless of the actual (baseline)

synchrony value. Thus, both time series length and error affect our ability to distinguish high from low synchrony, but through different effects on the data.

While time series length did not affect bias for most synchrony metrics evaluated, the Symbolic metric overestimated synchrony compared to baseline synchrony for short time series. It is unclear why this bias occurs (Fig. 3.4; Table 3.3). However, Cazelles (2004) emphasized that the appropriate question to consider with the Symbolic metric is not "What is the level of synchrony between two time series?" but is rather "Is this level of synchrony *significant* or not?" Given the observed bias in Symbolic synchrony estimates for short time series, my study supports Cazelles' (2004) suggestion that this metric may be more appropriate for distinguishing significant from non-significant synchrony, than than for comparing actual synchrony estimates.

Increasing data length from 15 to 25 years substantively reduced the variability of synchrony estimates. Given the 10-year cycles simulated in this study, this finding roughly concurred with the rule-of-thumb for analyzing cyclic data (minimum of three cycle lengths; Turchin 2003). This finding also held for the non-cyclic data, but to a lesser degree. Although increasing the length of time series data from 15 to 25 years decreases variability of synchrony estimates, it requires extensive additional money, effort, and elapsed time. In some cases, a more cost-efficient option for improving data for synchrony analyses might be to reduce sampling error in the time series abundance data.

Data smoothing greatly reduced error-associated bias in these simulations, but it came at the cost of increasing variability in synchrony estimates, especially for the shorter data lengths (< 50 years) that are typically available with real data. Depending on how data are collected, sampling error may be reduced in the data by increasing the per-area sampling effort, averaging results over multiple sampling sessions or sites, or better standardizing data collection efforts over time and space. At this time it is unclear how or if standard errors on abundance estimates can be used to 'correct' the underestimation bias of error in synchrony estimates (Buonaccorsi et al. 2001).

When error rates are high, the choice of whether or not to smooth data depends in part on how bias and variance of pairwise synchrony estimates may influence interpretation of results. When pairwise synchrony estimates are used to evaluate geographic patterns in synchrony data, high variance may reduce statistical power to detect a significant geographic pattern in the data. On the other hand, bias in pairwise synchrony estimates may lead to incorrect conclusions about geographic patterns if the bias is not randomly distributed with respect to geography. For example, all metrics in this study underestimated synchrony when simulated sampling error was high. If sampling error rates are generally higher in one part of the species range (e.g., due to reduced sampling efforts), we may incorrectly conclude that populations are less synchronized in that region. In this case, the benefits of smoothing the data to reduce bias may outweigh the costs of increased variability in synchrony estimates.

Cyclicity interacted with data quality and data treatment to influence variability and bias in synchrony estimates. I found that the standard deviation and bias of synchrony estimates (relative to baseline synchrony) was typically lower for non-cyclic compared to cyclic time series. When error was present, data smoothing was much more effective at reducing both bias and variability in synchrony estimates for cyclic compared to non-cyclic data. This finding makes sense—with cyclic data, smoothing was more likely to 'remove' error but retain the underlying periodic cycles. With non-cyclic data, the distinction between error and the underlying population dynamics was not as clear, and smoothing was less effective in reducing noise. Thus, smoothing to remove anticipated but unknown error in time series data would be more useful with cyclic data.

Influence of Data Quality, Data Treatment, and Cyclicity on Statistical Power

The Kendall metric exhibited higher statistical power than other metrics, under the data scenarios simulated in this study. However, for short time series (15 years) and in the presence of sampling error (error = 0.15 and 0.30), all metrics had little power to distinguish synchrony from independently generated null time series. Smoothing data greatly increased statistical power, consistent with the finding that smoothing reduces bias when error is high.

Study Limitations and Suggestions for Future Research

This study was designed to examine the relative advantages and limitations of five synchrony metrics when faced with data quality issues common in real data. I

simulated the very short time series (as little as 15 years) that are frequently available for synchrony studies. However, I employed only one data smoothing method (loess) with a single set of smoothing parameters. It is possible that other smoothing methods would be more effective at reducing bias and variability of synchrony estimates. Likewise, the levels of sampling error I chose (0, 0.15, and 0.3) fell within the range of errors observed in actual snowshoe hare time series data, when sampling error could be calculated. However, the method of randomly selecting independent error values from a normal distribution represents only one form of error that may be present in real data.

It is difficult to predict how results might differ for dynamics different from the 10-year cyclic and non-cyclic time series simulated in this study. For example, several species of oak cycle with 2-4 year periods (Sork et al. 1993; Liebhold et al. 2004b), voles and lemmings exhibit 3-5 year cycles (Korpimaki and Krebs 1996), and some grouse populations exhibit 4-6 year cycles (Williams 1985). With shorter cycles, it may be possible to generate time series with higher estimates of Symbolic synchrony because these shorter cycles (especially 4-year cycles) should be much better able to satisfy this metric's criterion of equal representation of cycle phase within a time series. Thus, it may possible to evaluate the performance of the Symbolic metric at a greater range of values than in this study. The general relationship between time series length, bias, and standard deviation may not change with shorter cycles, but low standard deviation may be achieved with shorter time series when cycles are shorter, i.e., the number of cycles may be as important as the number of years in reducing standard deviation.

For forest insect species with sudden, large outbreaks in abundance interspersed with periods of relative dormancy (e.g., some gypsy moth populations, *Lymantria dispar*; Williams and Liebhold 1995; Liebhold et al. 2000), data smoothing may lead to correspondence among synchrony metrics even greater than observed in this study because estimates for all synchrony metrics should be highly influenced by timing of peak outbreaks, and little influenced by subtle changes in population numbers between outbreaks.

Calculating pairwise synchrony is the first step in synchrony analysis. After pairwise estimates are calculated, they are typically incorporated in Mantel tests, cluster tests, and a variety of other statistical treatments for identifying spatial patterns in synchrony and, ultimately, proposing likely mechanisms. It would be useful to examine how variability and bias in this early step of pairwise estimates of synchrony influences identification of spatial patterns in synchrony. Further, it may be possible to use simulation-based results on variability in synchrony estimates to inform analysis of spatial synchrony patterns. For example, in Mantel tests, weighting synchrony estimates based on time series length (shorter time series = lower weight) may improve our ability to identify distance-synchrony relationships.

In this study, I evaluated statistical power by the proportion of significant time series when baseline data exhibited synchrony above zero (for Percent Match, this 'no synchrony' threshold was 0.5). In reality, most researchers probably would not consider very low levels of synchrony (e.g., Kendall synchrony of 0.05 out of a maximum 1.0) to

be biologically significant. Future investigations might address the question, “What is a biologically derived threshold value for synchrony?” so we can consider the ability of metrics to distinguish *biologically significant* synchrony from ‘no synchrony’.

LITERATURE CITED

Blasius, B. and L. Stone (2000). “Chaos and phase synchronization in ecological systems.” International Journal of Bifurcation and Chaos **10**(10): 2361-2380.

Buonaccorsi, J. P., J. S. Elkinton, et al. (2001). "Measuring and testing for spatial synchrony." Ecology **82**(6): 1668-1679.

Cazelles, B. (2004). "Symbolic dynamics for identifying similarity between rhythms of ecological time series." Ecology Letters **7**(9): 755-763.

Cazelles, B. and L. Stone (2003). "Detection of imperfect population synchrony in an uncertain world." Journal of Animal Ecology **72**(6): 953-968.

Chitty, H. (1948). "The snowshoe rabbit enquiry, 1943-46." Journal of Animal Ecology **17**(1): 39-44.

Cowan, I. M. (1938). “The fur trade and the fur cycle: 1825-1857.” B.C. Historical Quarterly **2:** 19-30.

Descamps, S., N.G. Yoccoz, et al. (2010). “Detecting population heterogeneity in effects of North Atlantic Oscillations on seabird body condition: get into the rhythm.” Oikos **1991:** 1526-1536.

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

Erb, J., N. C. Stenseth, et al. (2000). “Geographic variation in population cycles of Canadian muskrats (Ondatra zibethicus).” Canadian Journal of Zoology **78:** 1009-1016.

Gouhier, T. C., F. Guichard, et al. (2010). "Synchrony and Stability of Food Webs in Metacommunities." American Naturalist **175**(2): E16-E34.

Grenfell, B. T., O. N. Bjornstad, et al. (2001). "Travelling waves and spatial hierarchies in measles epidemics." Nature **414**(6865): 716-723.

Gudmundsson, F. (1958). “Some reflections on ptarmigan cycles.” Paper given at 12th International Ornithological Congress, Helsinki, June 1958.

Ims, R. A. and H. P. Andreassen (2000). "Spatial synchronization of vole population dynamics by predatory birds." Nature **408**(6809): 194-196.

Ims, R. A. and H. P. Andreassen (2005). "Density-dependent dispersal and spatial population dynamics." Proceedings of the Royal Society of London. Series B **272**(1566): 913-918.

Holmengen, N., K. L. Seip, et al. (2009). "Predator-prey coupling: interaction between mink Mustela vison and muskrat Ondatra zibethicus across Canada." Oikos **118**(3): 440-448.

Johnson, D. M., A. M. Liebhold, et al. (2006). “Geographical variation in the periodicity of gypsy moth outbreaks.” Ecography **29:** 367-374

Keith, L. B. (1963). Wildlife's ten year cycle. Madison, WI, University of Wisconsin Press.

Korpimaki, E. and C. J. Krebs (1996). "Predation and population cycles of small mammals - A reassessment of the predation hypothesis." BioScience **46**(10): 754-764.

Krebs, C. J., S. Boutin, et al. (1995). "Impact of food and predation on the snowshoe hare cycle." Science **269**(5227): 1112-1115.

Krebs, C.J., S. Boutin, et al. (2001). Ecosystem Dynamics of the Boreal Forest: the Kluane Project , Oxford University Press.

Krebs, C. J., A. J. Kenney, et al. (2002). "Synchrony in lemming and vole populations in the Canadian Arctic." Canadian Journal of Zoology **80**: 1323-1333.

Krebs, C.J. and J.H. Myers. (1974). “Population cycles in small mammals.” Advances in Ecological Research **8:** 267-399.

Liebhold, A., J. Elkinton, et al. (2000). "What causes outbreaks of the gypsy moth in North America?" Population Ecology **42**(3): 257-266.

Liebhold, A., W. D. Koenig, et al. (2004a). "Spatial synchrony in population dynamics." Annual Review of Ecology Evolution and Systematics **35**: 467-490.

Liebhold, A., V. Sork, et al. (2004b). "Within-population spatial synchrony in mast seeding of North American oaks." Oikos **104**(1): 156-164.

Lotka, A. J. (1925). Elements of Physical Biology. Baltimore, Williams & Wilkins Co.

Menard, F., F. Marsca, et al. (2007). “Climatic oscillations and tuna catch rates in the Indian Ocean: a wavelet approach to time series analysis.” Fisheries Oceanography **16**(1): 95-104.

Moran, P. A. P. (1953). "The statistical analysis of the Canadian lynx cycle, I. Structure and prediction." Australian Journal of Zoology **1**(2): 163-173.

Moss, R., A. Watson, et al. (1996). "Experimental prevention of a population cycle in red grouse." Ecology **77**(5): 1512-1530.

Nicholson, A. J. (1954). "An outline of the dynamics of animal populations." Australian Journal of Zoology **2**: 9-65.

Noakes, D. J. (2009). "Synchrony of marine fish catches and climate and ocean regime shifts in the North Pacific Ocean." Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science **1**: 155-168.

Orell, P., J. Erkinaro, et al. (2007). "Synchrony in the downstream migration of smolts and upstream migration of adult Atlantic salmon in the subarctic River Utsjoki." Journal of Fish Biology **71**: 1735-1750.

Ranta, E., V. Kaitala, et al. (1997). "The spatial dimension in population fluctuations." Science **278**: 1621-1623.

Shanker, K. and R. Sukumar (1999). "Synchrony in small mammal populations of montane forest patches in southern India." Journal of Animal Ecology **68**: 50-59.

Sork V. L., J. E. Bramble, et al. (1993). “Ecology of mast fruiting in three species of North American deciduous oaks.” Ecology **74:** 528-541.

Turchin, P. (2003). Complex Population Dynamics: a Theoretical / Empirical Synthesis, Princeton University Press.

Volterra, V. (1926). "Variazioni e fluttuazioni del numero d'individui in specie animali conviventi." Mem. R. Accad. Naz. dei Lincei **2**: 31-113.

Williams, D.W. and A. M. Liebhold (1995). “Influence of weather on the synchrony of gypsy moth (Lepidoptera: Lymantriidae) outbreaks in New England.” Environmental Entomology **24:** 987-995.

Williams, J. (1985). “Statistical analysis of fluctuations in red grouse bag data.” Oecologia **65:** 269-272.

Zar, J. H. (1999). Biostatistical analysis. Upper Saddle River, NJ, Prentice-Hall.

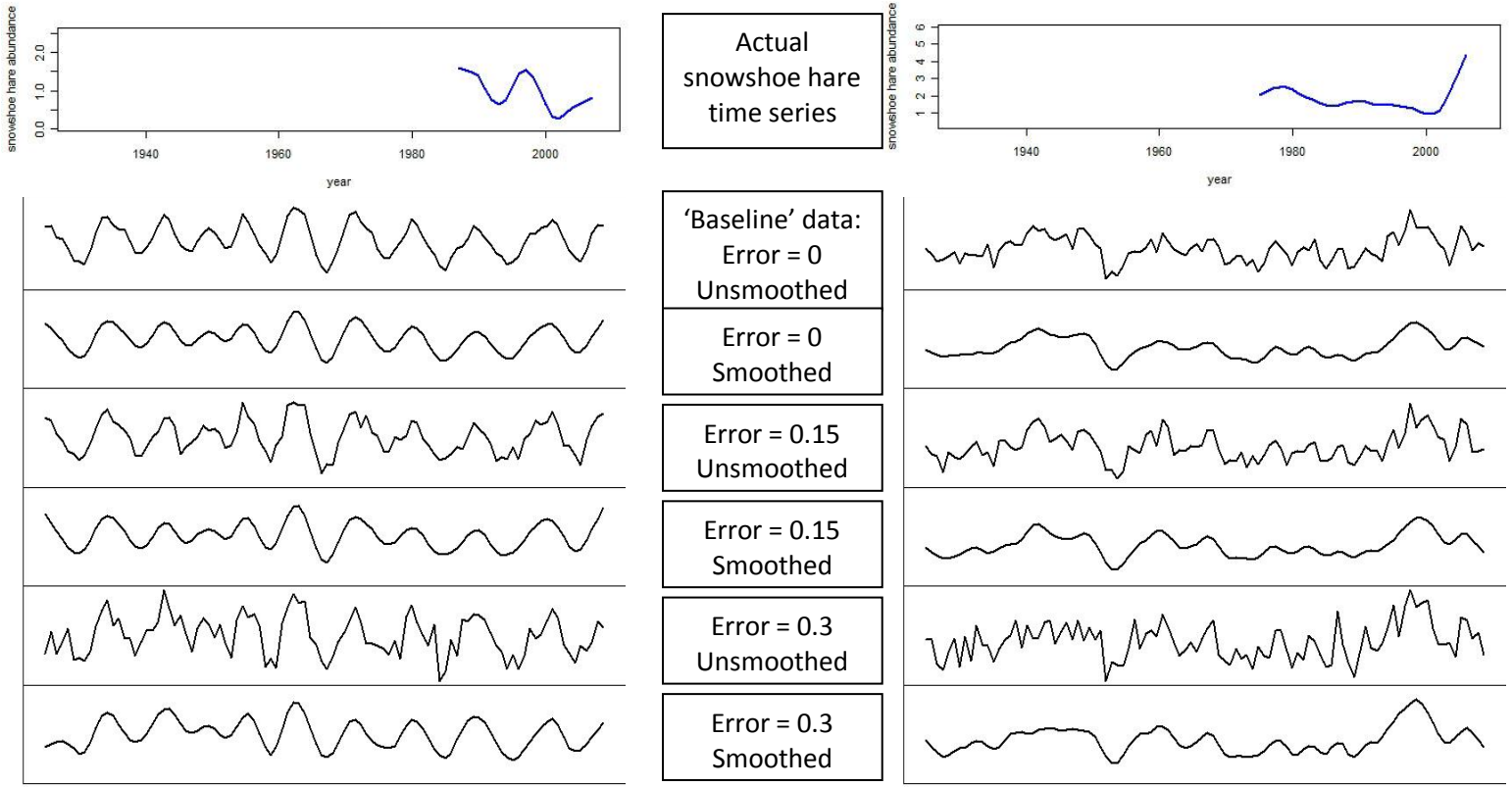


Figure 3.1
Actual and simulated snowshoe hare time series data. (A) Cyclic time series. Top panel: Actual snowshoe hare live-trap data collected in Yukon (courtesy of C.J. Krebs and the Kluane Boreal Forest Ecosystem Project; Krebs et al. 2001). Bottom panel: Representative 100-year subsets for six forms of simulated time series generated from an autoregressive model based on the Yukon data. (B) Non-cyclic time series. Top panel: Actual snowshoe hare harvest data collected in Utah (courtesy of Utah Division of Wildlife Resources). Bottom panel: Representative 100-year subsets for six forms of simulated time series generated from an autoregressive model based on the Utah data.

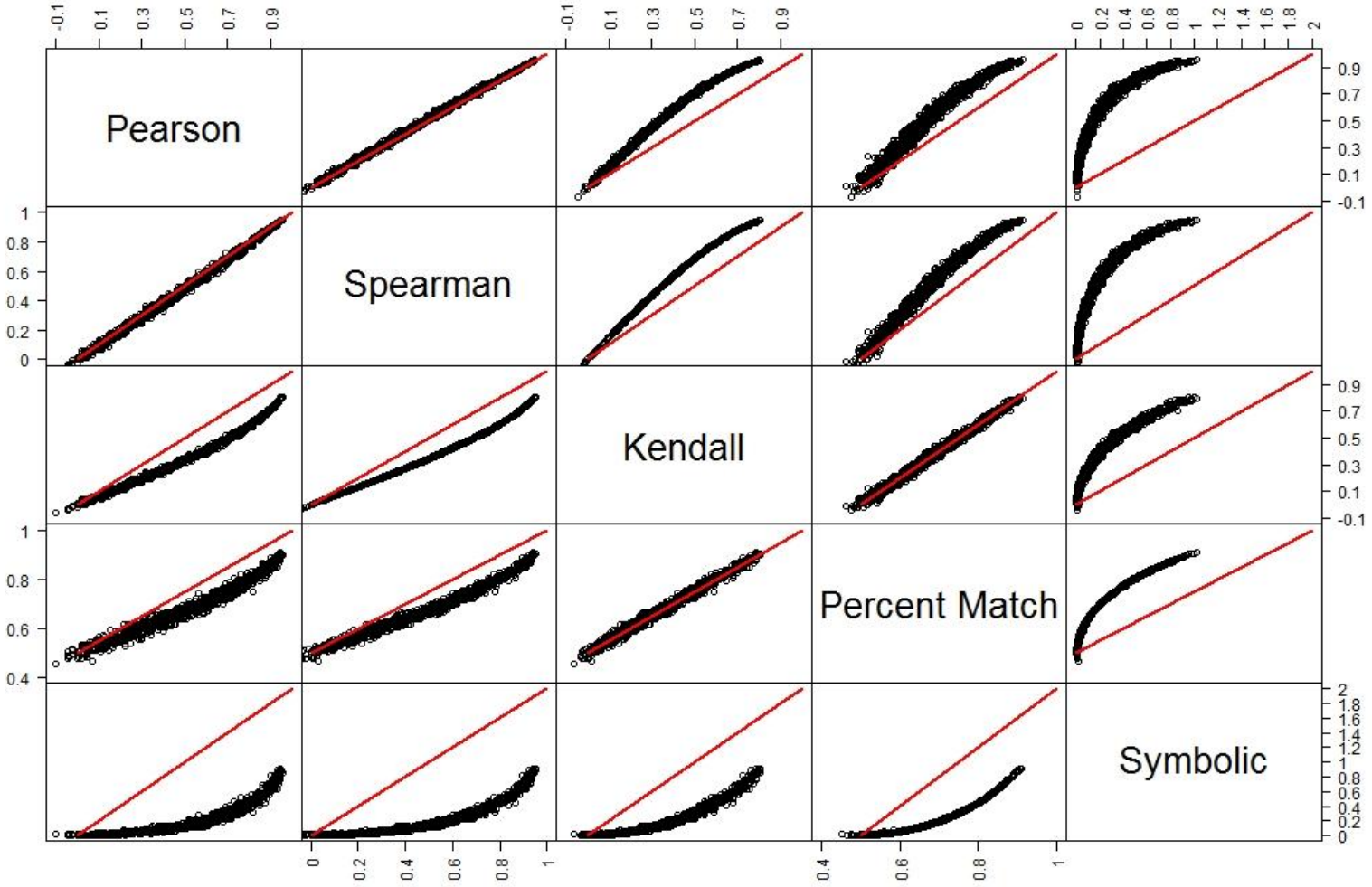


Figure 3.2 [CYCLIC TIME SERIES, ERROR = 0]
Scatterplots comparing synchrony estimates for each pair of metrics. The red line is the expected relationship for perfect correlation between metrics. Results for unsmoothed time series are above the diagonal; smoothed time series, below diagonal. The same time series data pairs were used for all scatterplots.

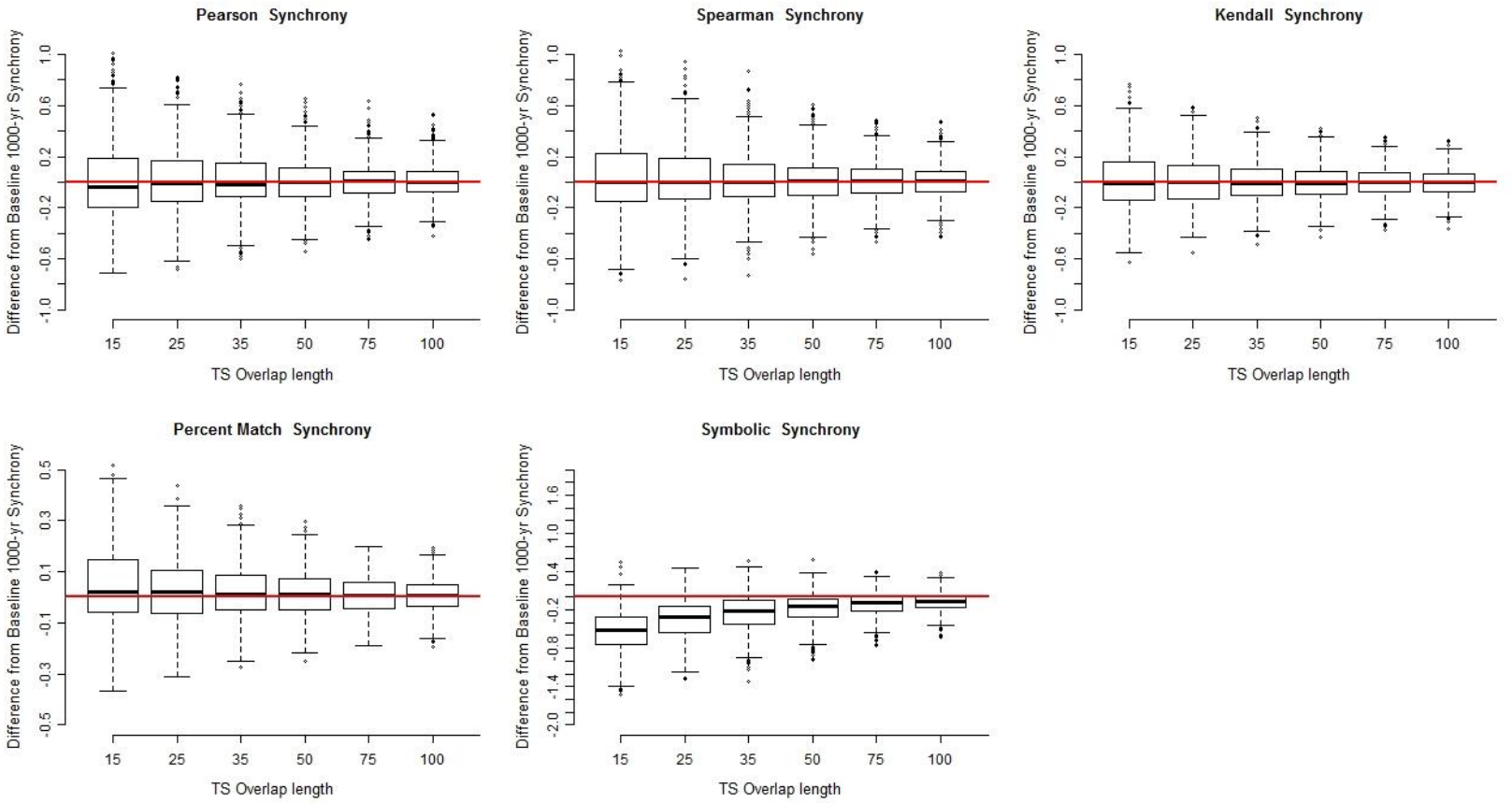


Figure 3.3 [CYCLIC, UNSMOOTHED TIME SERIES, ERROR = 0]
Boxplots showing difference in synchrony from baseline (1000-yr, error = 0, unsmoothed) synchrony. For each metric, results are shown for time series ranging in length from 15 to 100 years (x-axis). The red lines are the expected relationships for no difference between estimated and baseline synchrony. Values below the line (as for Symbolic metric) indicate synchrony is overestimated compared to baseline synchrony. In each boxplot, the center line is the median value, the box encloses the first to third quartiles of data, and box whiskers extend to 1.5 times the interquartile range. Data points exceeding this range are shown as circles.

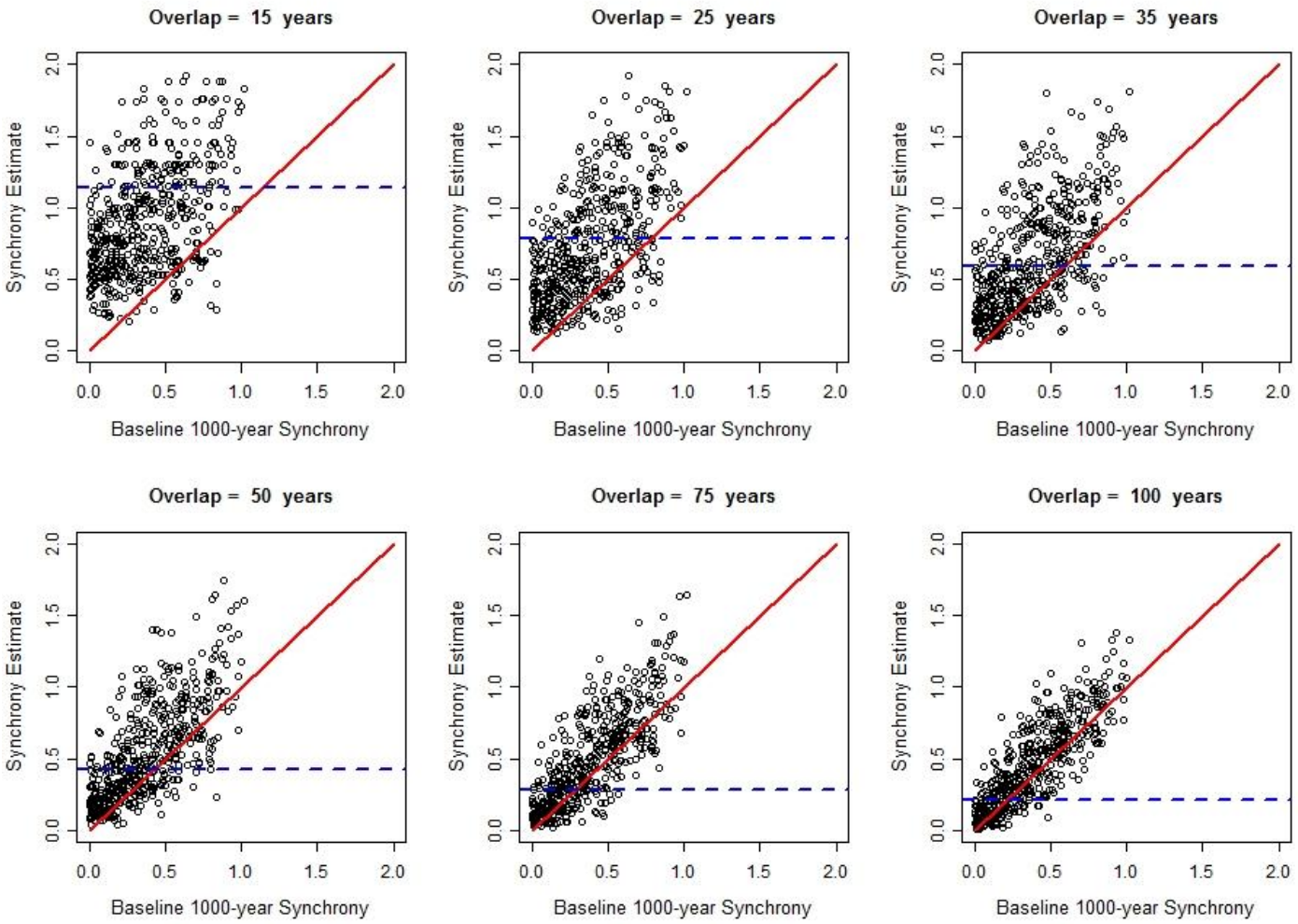


Figure 3.4 [CYCLIC, UNSMOOTHED TIME SERIES, ERROR = 0]
For the Symbolic metric, scatterplots of estimated synchrony (y-axis) vs. baseline synchrony (x-axis; baseline means 1000-yr, error = 0, unsmoothed) for six time series lengths. Red lines are expected relationships for no difference between estimated and baseline synchrony. For shorter time series, a positive bias in estimated synchrony is evident. Dashed blue line shows the 95% upper confidence interval for 'no synchrony', i.e., synchrony estimates above this line are considered significant.

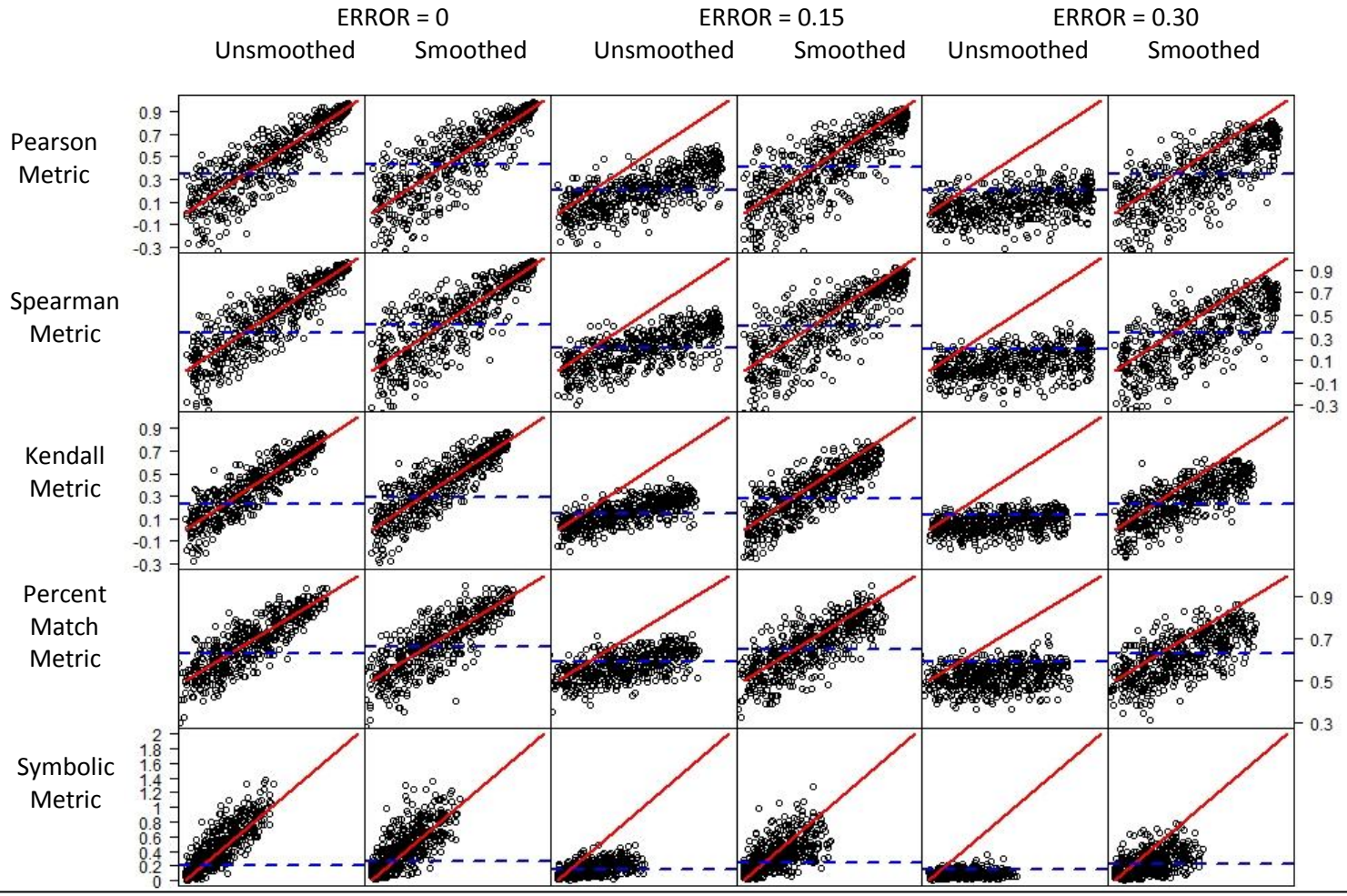


Figure 3.5 [CYCLIC, 100-YEAR TIME SERIES]
For each metric (rows), scatterplots of estimated synchrony (y-axis) vs. baseline synchrony (x-axis; baseline means 1000-yr, error = 0, unsmoothed) under various scenarios of error (0, 0.15, 0.30) and data treatment (smoothed vs. unsmoothed). Red lines are expected relationships for no difference between estimated and baseline synchrony. Dashed blue lines show the 95% upper confidence interval for 'no synchrony'.

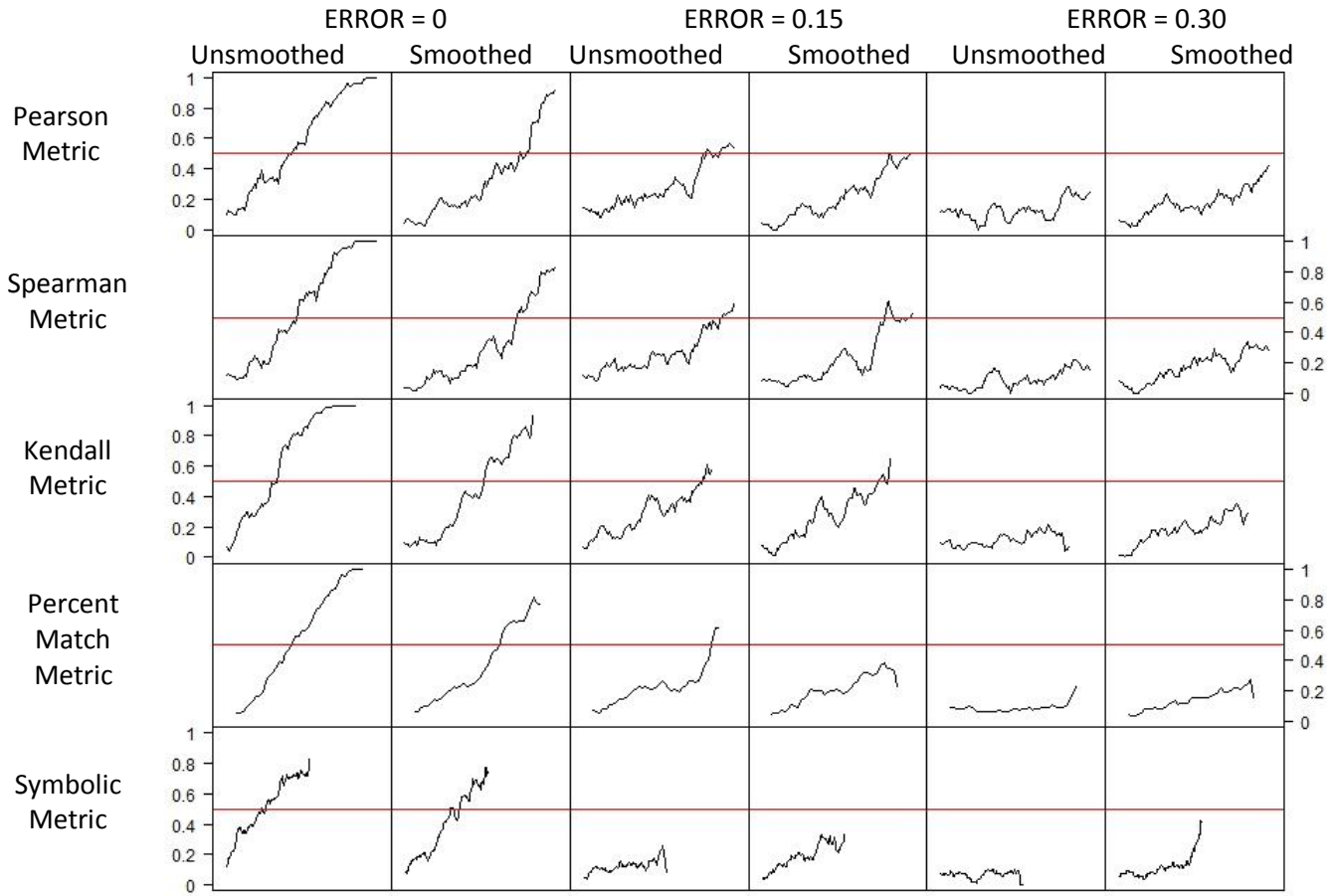


Figure 3.6 [CYCLIC, 15-YEAR TIME SERIES]
For each metric (rows), proportion of time series pairs significant at α = 0.05 (y-axis) vs. baseline synchrony (x-axis; baseline means 1000-yr, error = 0, unsmoothed) under various scenarios of error (0, 0.15, 0.30) and data treatment (smoothed vs. unsmoothed). For Pearson, Spearman, and Kendall metrics, baseline synchrony ranges 0 – 1.0; for Percent Match metric, 0.5 – 1.0; for Symbolic metric, 0 – 2.0. Red lines indicate 50% of time series pairs are significant. This figure corresponds with data in Table 3.5.

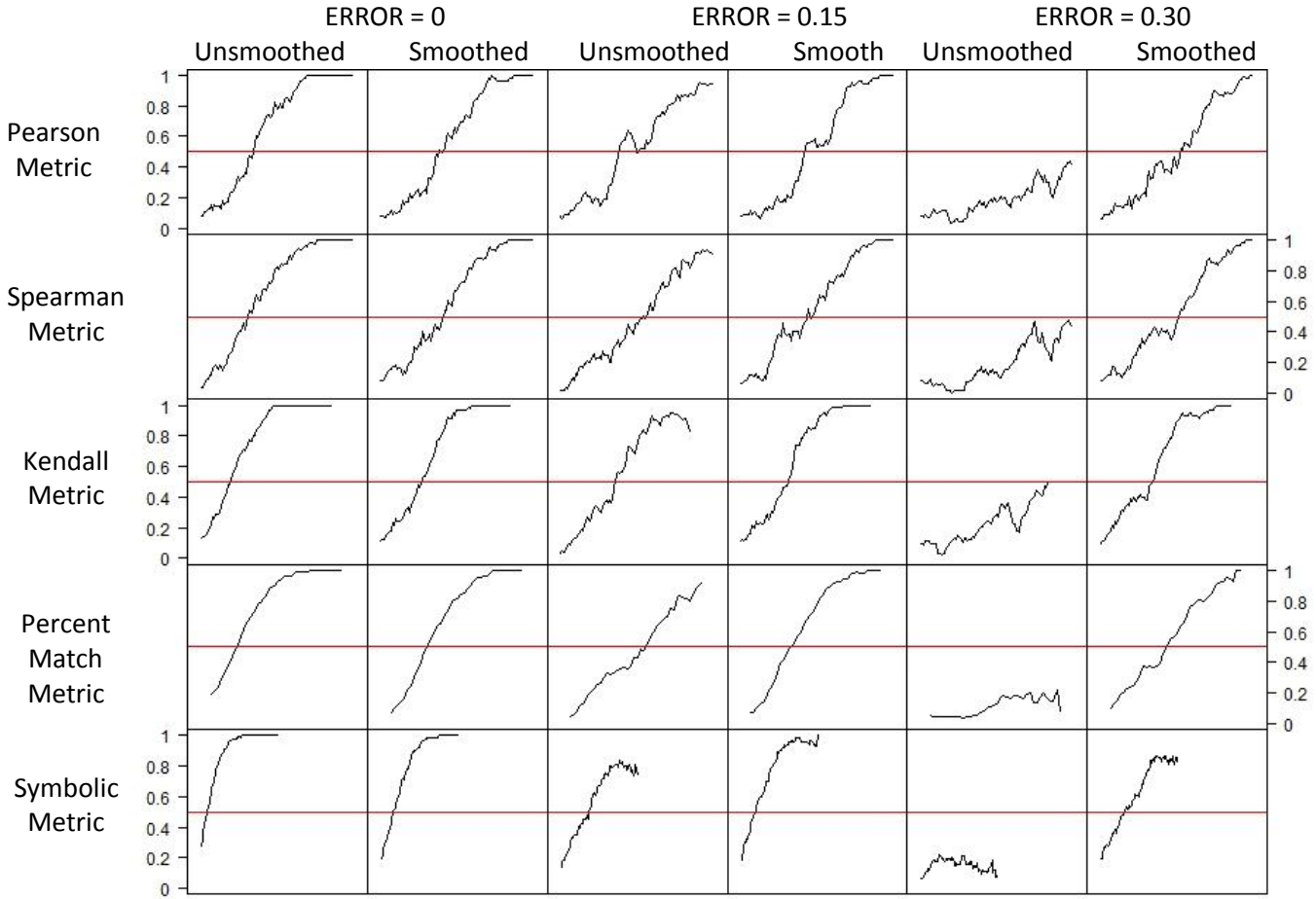


Figure 3.7 [CYCLIC, 100-YEAR TIME SERIES]
For each metric (rows), proportion of time series pairs significant at α = 0.05 (y-axis) vs. baseline synchrony (x-axis; baseline means 1000-yr, error = 0, unsmoothed) under various scenarios of error (0, 0.15, 0.30) and data treatment (smoothed vs. unsmoothed). For Pearson, Spearman, and Kendall metrics, baseline synchrony ranges 0 – 1.0; for Percent Match metric, 0.5 – 1.0; for Symbolic metric, 0 – 2.0. Red lines indicate 50% of time series pairs are significant. This figure corresponds with data in Table 3.5.

Table 3.1. Synchrony metrics compared in this study

Synchrony Metric	Description of Metric	Minimum Value	No Synchrony	Maximum Value
Pearson product moment correlation	Pearson correlation measures the linear dependence between two variables. Often, the relationship between abundances of two 'synchronized' populations can be much more complex than a simple linear relation. For example, two populations may cycle with the same frequency, peaking simultaneously. If the amplitudes of these cyclic peaks are not correlated (e.g., in a year when one population has a high peak, the other population has a relatively low peak), Pearson synchrony between these two populations may be low, despite the visual appearance that these populations are oscillating in synchrony. Examples: Ranta et al. 1997; Ims and Andreassen 2000	-1 (perfect negative correlation)	0	1
Spearman rank correlation coefficient	Spearman correlation measures how well the relationship between two variables can be described using a monotonic function. Spearman synchrony estimates are not affected by logarithmic transformations of the data. A perfect Spearman correlation of 1 results when two time series are monotonically related, even if their relationship is not linear. Spearman correlation is often thought of as a Pearson correlation between ranked variables. Examples: Shanker and Sukumar 1999; Orell et al. 2007	-1 (perfect negative correlation)	0	1
Kendall rank correlation coefficient	Proponents of Kendall correlation in statistical analyses have argued that the Kendall coefficient has an intuitively simple and direct interpretation, in constrast to the Spearman coefficient. The Kendall coefficient is the difference between P(same) and P(different), where P(same) is the probability for population numbers at two randomly selected time points to move in the same direction. For example, if abundances for Population A in Years 2 and 6 are 12 hares and 18 hares respectively, abundance is "increasing" between these time points. If abundances for Population B in the same years are also increasing, the population numbers are said to move in the same direction for these randomly selected time points. Examples: Noakes 2009; Gouhier et al. 2010	-1 (perfect negative correlation)	0	1

Synchrony Metric	Description of Metric	Minimum Value	No Synchrony	Maximum Value
Percent Match	Percent Match is a simple, straightforward metric proposed by Buonaccorsi et al. (2001) as a measure of how two time series change together. It is calculated as: (number of times series i & j move in the same direction)/(T-1), where T = the number of data points in a time series. The first step in estimating Percent Match is converting each time series to a string of variables ("I"ncrease and "D"ecrease) describing the growth or decline in each time series between successive time points. Percent Match is similar to Kendall correlation except the former uses only adjacent time points whereas Kendall correlation uses all possible pairs of points. This metric can be applied to a wide range of data, including survey data in which respondents report only if population numbers for a species increased or decreased from the previous year. Examples: Ims and Andreassen 2005	0 (perfect negative correlation)	0.5	1
Symbolic	The Symbolic metric, as proposed by Cazelles (2004), measures the degree of phase synchrony between two time series. Phase synchrony occurs when two time series oscillate at the same 'pace', rising and falling with the same rhythm. The amplitude of population cycles between the time series does not need to be correlated, nor does the timing of cyclic peaks need to match. A population that consistently peaks one year after another population would be as synchronized as two populations that always peak simultaneously. The first step in estimating the Symbolic metric is converting each time series to a string of variables ("I"ncrease, "P"eak, "D"ecrease, and "T"rough) based on examining the abundance at a time point relative to abundance in the previous and subsequent time points. A unique characteristic of this metric is that it is a function of the amount of the "mutual information" between two time series. Consequently, two time series that increase every year from Year 1 to Year 10 (whether with the same or different growth rates) would have a Symbolic synchrony of zero. Similarly, two identical time series that change direction (in population growth) five times in 20 years would be more synchronized than would two identical time series that change direction only two times in 20 years—despite the fact that in both scenarios the two time series are identical. For the Symbolic metric, the maximum synchrony of 2.0 is obtained only when two conditions are satisfied: 1) the four cycle phases are equally represented in both time series, i.e., the symbolic sequence for each time series comprises 25% increase, 25% peak, 25% decrease, and 25% trough phases; and 2) there is exact phase correspondence between the two time series, i.e., given the symbolic sequence of one time series, the sequence of the paired time series can be exactly predicted. The Symbolic metric does not distinguish between negative and positive correlation, just as it does not distinguish between time series that peak simultaneously versus with a consistent lag. Examples: Menard et al. 2007; Descamps et al. 2010	0 (no synchrony)	0	2

ERROR: 0

	Pearson	Spearman	Kendall	Percent Match	Symbolic
Pearson		1.00	0.99	0.98	0.90
Spearman	1.00		1.00	0.99	0.91
Kendall	0.99	1.00		0.99	0.95
Percent Match	0.98	0.99	0.99		0.95
Symbolic	0.90	0.91	0.94	0.95	

ERROR: 0.15

	Pearson	Spearman	Kendall	Percent Match	Symbolic
Pearson		1.00	1.00	0.95	0.92
Spearman	1.00		1.00	0.97	0.93
Kendall	1.00	1.00		0.97	0.94
Percent Match	0.98	0.99	0.99		0.94
Symbolic	0.92	0.93	0.95	0.95	

ERROR: 0.30

	Pearson	Spearman	Kendall	Percent Match	Symbolic
Pearson		0.98	0.98	0.83	0.73
Spearman	1.00		1.00	0.88	0.76
Kendall	1.00	1.00		0.88	0.76
Percent Match	0.98	0.99	0.99		0.79
Symbolic	0.92	0.93	0.94	0.95	

Table 3.2 [CYCLIC TIME SERIES]
Correlation (Pearson's r) in synchrony estimates for each pair of metrics, under various scenarios of error (0, 0.15, 0.30). Correlations for unsmoothed time series are above the diagonal; smoothed time series, below diagonal. These results correspond with Figure 3.2.

		PEARSON					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.33	0.27	0.23	0.19	0.17	0.15
	0.15	0.33	0.28	0.25	0.23	0.21	0.20
	0.3	0.37	0.32	0.30	0.28	0.27	0.26
SMOOTH	0	0.44	0.35	0.29	0.24	0.21	0.18
	0.15	0.47	0.36	0.29	0.24	0.21	0.19
	0.3	0.45	0.35	0.31	0.27	0.23	0.20

		PERCENT MATCH **scaled					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.30	0.25	0.21	0.18	0.15	0.13
	0.15	0.32	0.26	0.23	0.20	0.18	0.18
	0.3	0.33	0.29	0.26	0.24	0.23	0.22
SMOOTH	0	0.37	0.31	0.25	0.22	0.18	0.16
	0.15	0.37	0.30	0.25	0.21	0.18	0.16
	0.3	0.38	0.31	0.26	0.22	0.20	0.17

		SPEARMAN					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.32	0.27	0.23	0.20	0.17	0.15
	0.15	0.33	0.28	0.25	0.23	0.21	0.20
	0.3	0.38	0.33	0.30	0.28	0.27	0.26
SMOOTH	0	0.42	0.34	0.29	0.24	0.20	0.18
	0.15	0.44	0.36	0.31	0.25	0.21	0.19
	0.3	0.43	0.34	0.32	0.27	0.22	0.20

		SYMBOLIC (synchrony range 0 - 2.0)					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.35	0.32	0.29	0.24	0.20	0.16
	0.15	0.32	0.30	0.27	0.24	0.23	0.22
	0.3	0.32	0.29	0.27	0.26	0.25	0.25
SMOOTH	0	0.40	0.36	0.32	0.27	0.23	0.20
	0.15	0.39	0.35	0.32	0.27	0.23	0.21
	0.3	0.37	0.32	0.29	0.25	0.23	0.22

		KENDALL					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.24	0.20	0.16	0.13	0.12	0.11
	0.15	0.26	0.22	0.20	0.18	0.17	0.16
	0.3	0.29	0.26	0.24	0.23	0.22	0.21
SMOOTH	0	0.32	0.25	0.21	0.17	0.15	0.13
	0.15	0.33	0.25	0.21	0.17	0.15	0.13
	0.3	0.33	0.26	0.22	0.19	0.17	0.15

** Percent Match metric is scaled as SD/range, where range of synchrony values for Percent Match = 0.5. This metric is scaled to facilitate comparison of metric performance with the correlation synchrony metrics. The Symbolic metric is not scaled because it applies a definition of synchrony that cannot be easily compared with other metrics in this study.

Table 3.3A [CYCLIC TIME SERIES]
Standard deviation of the difference between estimated and baseline (1000-yr, error = 0, unsmoothed) synchrony. For each metric, upper panel = unsmoothed time series and lower panel = smoothed time series. Blue shading indicates smoothing increased standard deviation of synchrony estimates (compared to unsmoothed data); yellow shading indicates smoothing reduced standard deviation of synchrony estimates.

	ERROR	PEARSON 15	25	35	50	75	100
UNSMOOTH	0	0.02	0.02	0.01	0.01	0.01	0.01
	0.15	0.30	0.30	0.29	0.29	0.29	0.29
	0.3	0.44	0.43	0.42	0.42	0.42	0.42
SMOOTH	0	0.04	0.03	0.01	0.01	0.01	0.01
	0.15	0.12	0.09	0.08	0.07	0.06	0.06
	0.3	0.17	0.16	0.16	0.15	0.15	0.15

	ERROR	PERCENT MATCH **scaled 15	25	35	50	75	100
UNSMOOTH	0	0.08	0.05	0.03	0.02	0.01	0.01
	0.15	0.30	0.27	0.26	0.25	0.25	0.24
	0.3	0.41	0.38	0.37	0.35	0.34	0.34
SMOOTH	0	0.09	0.05	0.04	0.03	0.02	0.02
	0.15	0.13	0.10	0.09	0.08	0.06	0.06
	0.3	0.22	0.17	0.16	0.15	0.14	0.14

	ERROR	SPEARMAN 15	25	35	50	75	100
UNSMOOTH	0	0.04	0.03	0.02	0.01	0.01	0.01
	0.15	0.30	0.29	0.29	0.29	0.29	0.29
	0.3	0.44	0.43	0.42	0.41	0.41	0.41
SMOOTH	0	0.06	0.03	0.02	0.01	0.02	0.02
	0.15	0.12	0.10	0.09	0.07	0.07	0.06
	0.3	0.21	0.18	0.17	0.15	0.15	0.15

	ERROR	SYMBOLIC (synchrony range 0 - 2.0) 15	25	35	50	75	100
UNSMOOTH	0	-0.54	-0.35	-0.25	-0.17	-0.11	-0.07
	0.15	-0.33	-0.08	0.04	0.12	0.19	0.22
	0.3	-0.25	0.01	0.12	0.20	0.26	0.29
SMOOTH	0	-0.51	-0.36	-0.24	-0.16	-0.09	-0.06
	0.15	-0.44	-0.26	-0.15	-0.06	0.01	0.04
	0.3	-0.35	-0.14	-0.04	0.04	0.11	0.14

	ERROR	KENDALL 15	25	35	50	75	100
UNSMOOTH	0	0.01	0.01	0.00	-0.01	0.00	0.00
	0.15	0.24	0.24	0.24	0.24	0.25	0.25
	0.3	0.36	0.35	0.35	0.34	0.34	0.34
SMOOTH	0	0.00	0.01	0.00	-0.01	0.00	0.00
	0.15	0.08	0.07	0.06	0.05	0.05	0.05
	0.3	0.14	0.14	0.13	0.13	0.13	0.13

** Percent Match metric is scaled as SD/range, where range of synchrony values for Percent Match = 0.5. This metric is scaled to facilitate comparison of metric performance with the correlation synchrony metrics. The Symbolic metric is not scaled because it applies a definition of synchrony that cannot be easily compared with other metrics in this study.

Table 3.3B [CYCLIC TIME SERIES]
Bias of estimated synchrony compared to baseline (1000-yr, error = 0, unsmoothed) synchrony. For each metric, upper panel = unsmoothed time series and lower panel = smoothed time series. Blue shading indicates smoothing increased absolute bias of synchrony estimates (compared to unsmoothed data); yellow shading indicates smoothing reduced absolute bias of synchrony estimates.

	ERROR	PEARSON 15	25	35	50	75	100
UNSMOOTH	0	0.21	0.16	0.13	0.11	0.08	0.07
	0.15	0.30	0.25	0.22	0.19	0.17	0.16
	0.3	0.38	0.33	0.31	0.28	0.26	0.25
SMOOTH	0	0.38	0.27	0.22	0.18	0.14	0.12
	0.15	0.40	0.29	0.24	0.20	0.16	0.14
	0.3	0.45	0.33	0.29	0.26	0.22	0.20

	ERROR	PERCENT MATCH **scaled 15	25	35	50	75	100
UNSMOOTH	0	0.23	0.18	0.15	0.12	0.10	0.09
	0.15	0.28	0.23	0.20	0.18	0.16	0.16
	0.3	0.33	0.29	0.27	0.24	0.22	0.21
SMOOTH	0	0.31	0.26	0.21	0.17	0.14	0.12
	0.15	0.33	0.27	0.24	0.20	0.16	0.14
	0.3	0.38	0.29	0.26	0.22	0.19	0.17

	ERROR	SPEARMAN 15	25	35	50	75	100
UNSMOOTH	0	0.22	0.16	0.14	0.12	0.09	0.08
	0.15	0.31	0.25	0.22	0.19	0.18	0.17
	0.3	0.38	0.34	0.31	0.28	0.27	0.26
SMOOTH	0	0.39	0.28	0.22	0.18	0.14	0.13
	0.15	0.39	0.31	0.26	0.21	0.17	0.15
	0.3	0.46	0.35	0.31	0.27	0.23	0.21

	ERROR	SYMBOLIC (synchrony range 0 - 2.0) 15	25	35	50	75	100
UNSMOOTH	0	0.34	0.28	0.25	0.20	0.16	0.13
	0.15	0.36	0.30	0.27	0.25	0.25	0.24
	0.3	0.38	0.32	0.31	0.29	0.29	0.28
SMOOTH	0	0.38	0.30	0.28	0.23	0.17	0.15
	0.15	0.38	0.31	0.29	0.26	0.22	0.21
	0.3	0.38	0.33	0.31	0.29	0.27	0.25

	ERROR	KENDALL 15	25	35	50	75	100
UNSMOOTH	0	0.16	0.12	0.10	0.08	0.06	0.05
	0.15	0.24	0.20	0.17	0.16	0.15	0.14
	0.3	0.29	0.25	0.23	0.21	0.21	0.20
SMOOTH	0	0.29	0.21	0.17	0.13	0.10	0.09
	0.15	0.30	0.23	0.19	0.16	0.13	0.11
	0.3	0.34	0.26	0.23	0.20	0.17	0.16

** Percent Match metric is scaled as SD/range, where range of synchrony values for Percent Match = 0.5. This metric is scaled to facilitate comparison of metric performance with the correlation synchrony metrics. The Symbolic metric is not scaled because it applies a definition of synchrony that cannot be easily compared with other metrics in this study.

Table 3.3C [NON-CYCLIC TIME SERIES]
Standard deviation of the difference between estimated and baseline (1000-yr, error = 0, unsmoothed) synchrony. For each metric, upper panel = unsmoothed time series and lower panel = smoothed time series. Blue shading indicates smoothing increased standard deviation of synchrony estimates (compared to unsmoothed data); yellow shading indicates smoothing reduced standard deviation of synchrony estimates.

	ERROR	15	25	35	50	75	100
	PEARSON						
	0	0.01	0.01	0.01	0.01	0.00	0.01
	0.15	0.22	0.22	0.22	0.22	0.22	0.22
	0.3	0.38	0.37	0.39	0.38	0.39	0.39
	0	0.03	0.00	0.01	0.01	0.00	0.01
	0.15	0.12	0.08	0.09	0.09	0.09	0.09
	0.3	0.27	0.23	0.24	0.24	0.24	0.24

	ERROR	15	25	35	50	75	100
	SPEARMAN						
	0	0.02	0.02	0.01	0.01	0.00	0.00
	0.15	0.24	0.23	0.23	0.23	0.23	0.23
	0.3	0.39	0.39	0.39	0.39	0.39	0.39
	0	0.07	0.04	0.02	0.01	0.00	0.00
	0.15	0.16	0.13	0.12	0.11	0.09	0.09
	0.3	0.30	0.27	0.26	0.25	0.24	0.24

	ERROR	15	25	35	50	75	100
	KENDALL						
	0	-0.01	-0.01	0.00	0.00	0.00	0.00
	0.15	0.19	0.19	0.20	0.20	0.20	0.20
	0.3	0.31	0.31	0.32	0.32	0.32	0.32
	0	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
	0.15	0.08	0.08	0.09	0.08	0.08	0.09
	0.3	0.23	0.20	0.20	0.20	0.20	0.21

	ERROR	15	25	35	50	75	100
	PERCENT MATCH **scaled						
UNSMOOTH	**0**	0.09	0.05	0.04	0.03	0.02	0.02
UNSMOOTH	**0.15**	0.26	0.23	0.22	0.21	0.19	0.19
UNSMOOTH	**0.3**	0.34	0.32	0.32	0.32	0.31	0.31
SMOOTH	**0**	0.11	0.08	0.07	0.07	0.05	0.04
SMOOTH	**0.15**	0.20	0.17	0.16	0.14	0.13	0.12
SMOOTH	**0.3**	0.29	0.25	0.25	0.24	0.23	0.23

	ERROR	15	25	35	50	75	100
	SYMBOLIC (synchrony range 0 - 2.0)						
UNSMOOTH	**0**	-0.51	-0.34	-0.24	-0.16	-0.10	-0.07
UNSMOOTH	**0.15**	-0.27	-0.03	0.08	0.16	0.22	0.24
UNSMOOTH	**0.3**	-0.21	0.04	0.15	0.24	0.30	0.33
SMOOTH	**0**	-0.41	-0.26	-0.17	-0.10	-0.03	0.00
SMOOTH	**0.15**	-0.31	-0.15	-0.04	0.04	0.11	0.14
SMOOTH	**0.3**	-0.26	-0.04	0.07	0.16	0.23	0.26

** Percent Match metric is scaled as SD/range, where range of synchrony values for Percent Match = 0.5. This metric is scaled to facilitate comparison of metric performance with the correlation synchrony metrics. The Symbolic metric is not scaled because it applies a definition of synchrony that cannot be easily compared with other metrics in this study.

Table 3.3D [NON-CYCLIC TIME SERIES]
Bias of estimated synchrony compared to baseline (1000-yr, error = 0, unsmoothed) synchrony. For each metric, upper panel = unsmoothed time series and lower panel = smoothed time series. Blue shading indicates smoothing increased absolute bias of synchrony estimates (compared to unsmoothed data); yellow shading indicates smoothing reduced absolute bias of synchrony estimates.

		PEARSON					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.70	0.60	0.52	0.45	0.39	0.35
	0.15	0.51	0.40	0.33	0.28	0.24	0.20
	0.3	0.49	0.40	0.33	0.26	0.22	0.20
SMOOTH	0	0.86	0.73	0.66	0.56	0.47	0.43
	0.15	0.85	0.72	0.63	0.53	0.46	0.41
	0.3	0.79	0.63	0.56	0.47	0.40	0.35

		PERCENT MATCH **scaled					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.47	0.44	0.43	0.32	0.28	0.26
	0.15	0.33	0.28	0.26	0.24	0.20	0.18
	0.3	0.33	0.28	0.26	0.24	0.20	0.18
SMOOTH	0	0.73	0.60	0.54	0.44	0.33	0.32
	0.15	0.73	0.52	0.49	0.40	0.33	0.30
	0.3	0.60	0.52	0.43	0.36	0.28	0.26

		SPEARMAN					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.71	0.60	0.53	0.45	0.37	0.35
	0.15	0.50	0.40	0.32	0.29	0.24	0.21
	0.3	0.48	0.39	0.33	0.27	0.23	0.19
SMOOTH	0	0.82	0.72	0.65	0.54	0.46	0.42
	0.15	0.82	0.69	0.64	0.56	0.45	0.40
	0.3	0.77	0.62	0.54	0.48	0.39	0.34

		SYMBOLIC (synchrony range 0 - 2.0)					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	1.14	0.79	0.59	0.42	0.29	0.21
	0.15	1.02	0.70	0.48	0.33	0.20	0.15
	0.3	0.99	0.63	0.47	0.32	0.20	0.14
SMOOTH	0	1.30	0.95	0.71	0.53	0.35	0.27
	0.15	1.30	0.92	0.67	0.47	0.33	0.25
	0.3	1.24	0.85	0.61	0.42	0.29	0.22

		KENDALL					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.54	0.42	0.37	0.30	0.26	0.23
	0.15	0.36	0.28	0.22	0.20	0.16	0.14
	0.3	0.34	0.28	0.24	0.18	0.15	0.13
SMOOTH	0	0.67	0.54	0.47	0.38	0.32	0.29
	0.15	0.65	0.51	0.47	0.38	0.31	0.28
	0.3	0.60	0.46	0.39	0.32	0.27	0.23

** Percent Match metric is scaled as (x-0.50)/range, where range of synchrony values for Percent Match = 0.5. This metric is scaled to facilitate comparison of metric performance with the correlation synchrony metrics. The Symbolic metric is not scaled because it applies a definition of synchrony that cannot be easily compared with other metrics in this study.

Table 3.4 [CYCLIC TIME SERIES]
The 95% upper confidence limit for 'no synchrony, calculated on independent (uncorrelated) time series. These values correspond with the blue dotted lines in Figs. 3.4 and 3.5. For each metric, upper panel = unsmoothed time series and lower panel = smoothed time series. For Percent Match and Symbolic metrics, data are not scaled, but the range of synchrony values for these metrics is provided in parentheses because they differ from the 0 – 1 range of other metrics.

		PEARSON					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.69	0.57	0.44	0.38	0.38	0.37
	0.15	NA	0.86	0.82	0.65	0.61	0.40
	0.3	NA	NA	NA	NA	NA	NA
SMOOTH	0	0.74	0.71	0.63	0.49	0.40	0.40
	0.15	0.86	0.72	0.70	0.66	0.45	0.44
	0.3	NA	0.84	0.72	0.66	0.60	0.53

		PERCENT MATCH **scaled					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.62	0.56	0.52	0.34	0.30	0.26
	0.15	NA	NA	NA	0.70	0.60	0.56
	0.3	NA	NA	NA	NA	NA	NA
SMOOTH	0	NA	0.70	0.58	0.50	0.32	0.32
	0.15	NA	0.80	0.62	0.50	0.42	0.36
	0.3	NA	NA	0.76	0.60	0.48	0.44

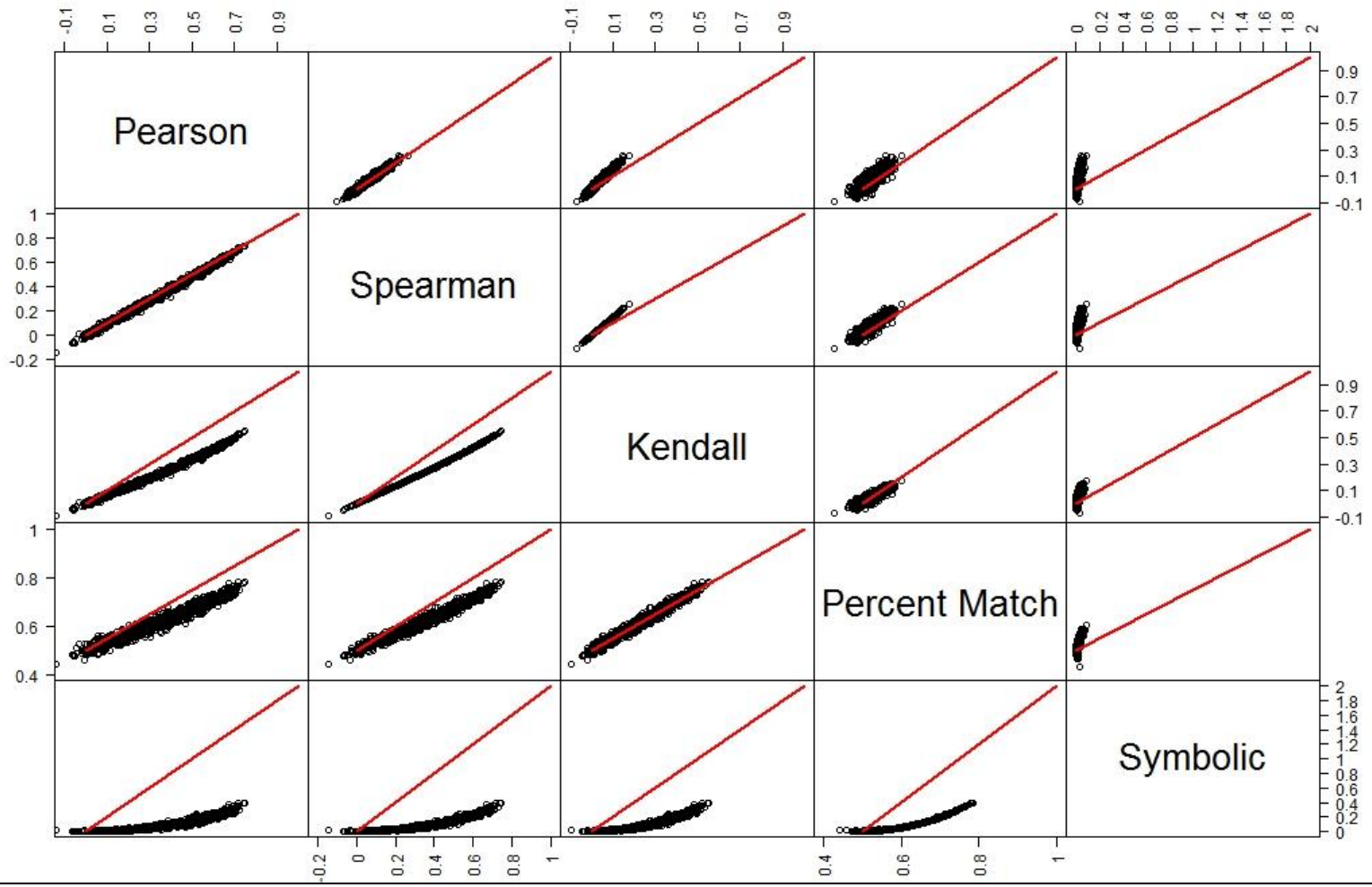
		SPEARMAN					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.73	0.52	0.52	0.39	0.36	0.34
	0.15	NA	0.85	0.78	0.70	0.65	0.54
	0.3	NA	NA	NA	NA	NA	NA
SMOOTH	0	0.84	0.65	0.61	0.45	0.45	0.43
	0.15	NA	0.70	0.72	0.68	0.50	0.45
	0.3	NA	0.84	0.74	0.67	0.58	0.52

		SYMBOLIC (synchrony range 0 - 2.0)					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.76	0.45	0.40	0.30	0.22	0.18
	0.15	NA	NA	1.01	0.89	0.52	0.44
	0.3	NA	NA	NA	NA	NA	NA
SMOOTH	0	0.95	0.71	0.43	0.37	0.29	0.25
	0.15	NA	NA	0.86	0.45	0.38	0.27
	0.3	NA	NA	NA	0.69	0.56	0.38

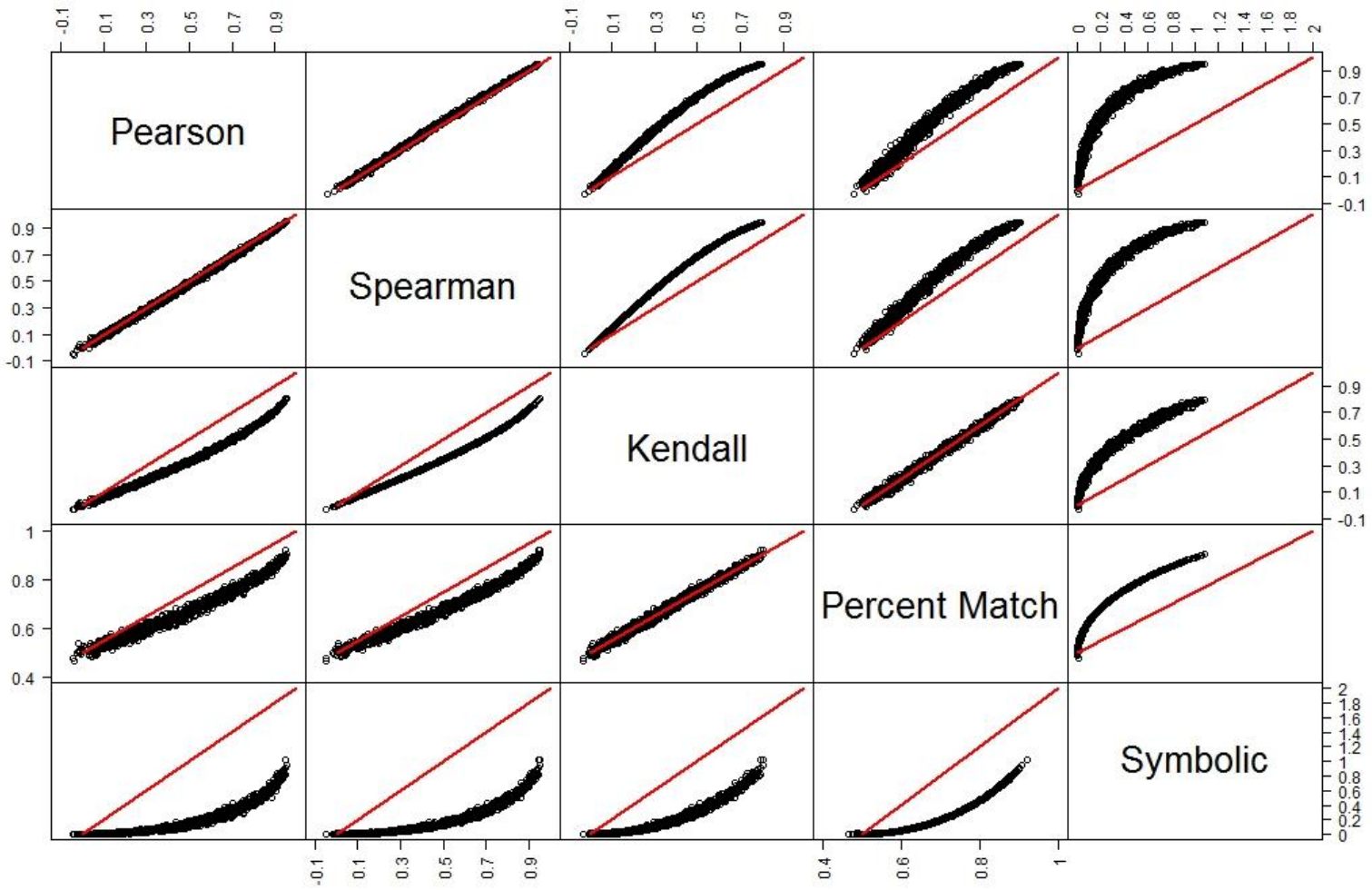
		KENDALL					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.54	0.40	0.39	0.29	0.27	0.23
	0.15	NA	0.66	0.46	0.53	0.38	0.38
	0.3	NA	NA	NA	NA	NA	NA
SMOOTH	0	0.67	0.47	0.45	0.39	0.33	0.30
	0.15	0.75	0.55	0.53	0.44	0.34	0.34
	0.3	NA	0.66	0.51	0.47	0.43	0.36

** Percent Match metric is scaled as (x-0.50)/range, where range of synchrony values for Percent Match = 0.5. This metric is scaled to facilitate comparison of metric performance with the correlation synchrony metrics. The Symbolic metric is not scaled because it applies a definition of synchrony that cannot be easily compared with other metrics in this study.

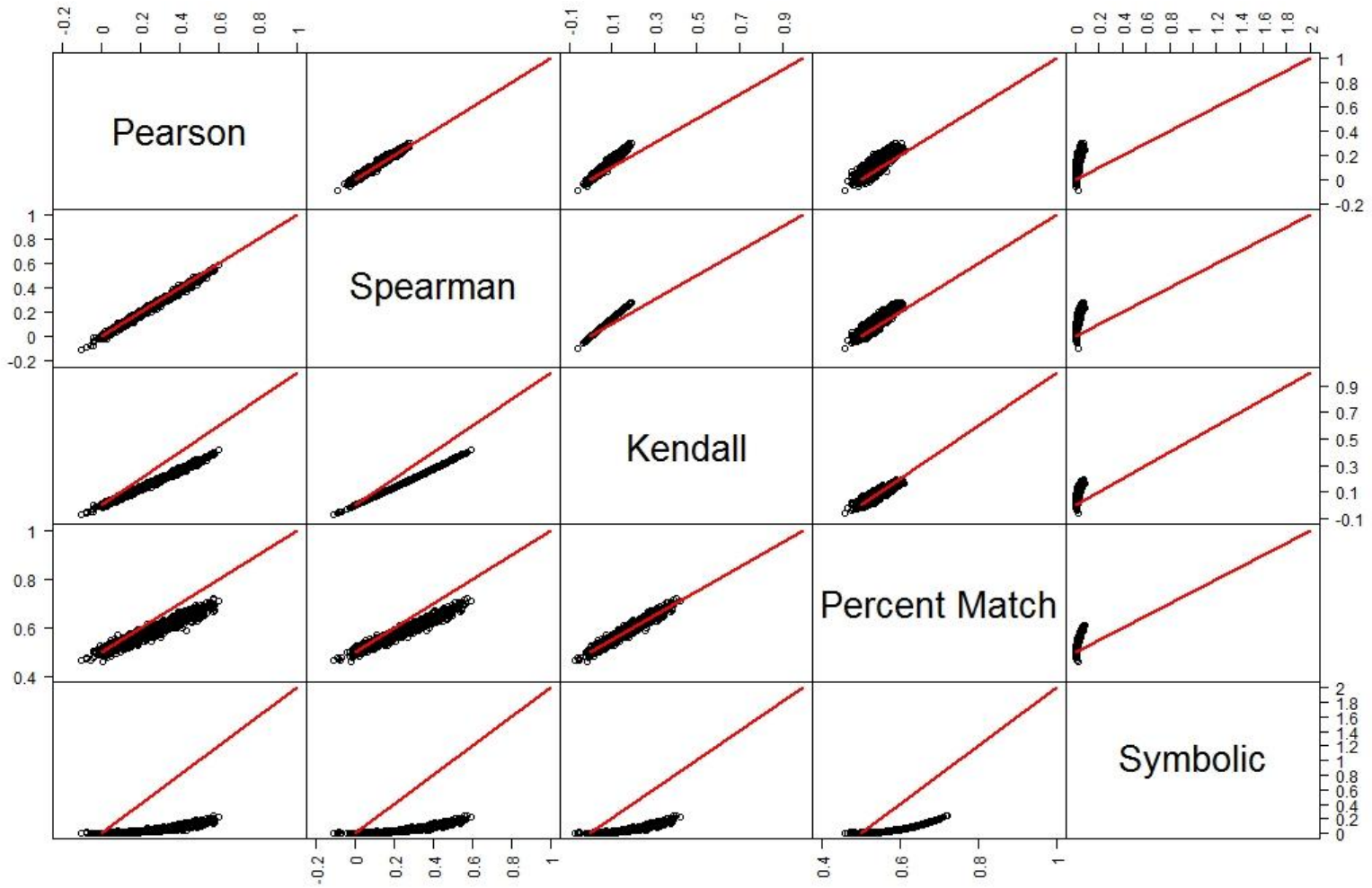
Table 3.5 [CYCLIC TIME SERIES]
The 50% significance categories for each metric. This number indicates the minimum baseline synchrony value for which at least 50% of estimated synchrony values are significant at $\alpha = 0.05$. "NA" indicates no baseline synchrony value had at least 50% of estimated synchrony values significant at $\alpha = 0.05$.



Appendix 3.1 [CYCLIC TIME SERIES, ERROR = 0.30]
Scatterplots comparing synchrony estimates for each pair of metrics. The red line is the expected relationship for perfect correlation between metrics. Results for unsmoothed time series are above the diagonal; smoothed time series, below diagonal. The same time series data pairs were used for all scatterplots.



Appendix 3.1 [NON-CYCLIC TIME SERIES, ERROR = 0]
Scatterplots comparing synchrony estimates for each pair of metrics. The red line is the expected relationship for perfect correlation between metrics. Results for unsmoothed time series are above the diagonal; smoothed time series, below diagonal. The same time series data pairs were used for all scatterplots.



Appendix 3.1 [NON-CYCLIC TIME SERIES, ERROR = 0.30]
Scatterplots comparing synchrony estimates for each pair of metrics. The red line is the expected relationship for perfect correlation between metrics. Results for unsmoothed time series are above the diagonal; smoothed time series, below diagonal. The same time series data pairs were used for all scatterplots.

ERROR: 0

	Pearson	Spearman	Kendall	Percent Match	Symbolic
Pearson		1.00	0.99	0.99	0.92
Spearman	1.00		1.00	0.99	0.92
Kendall	0.99	1.00		1.00	0.95
Percent Match	0.99	0.99	0.99		0.96
Symbolic	0.91	0.92	0.95	0.95	

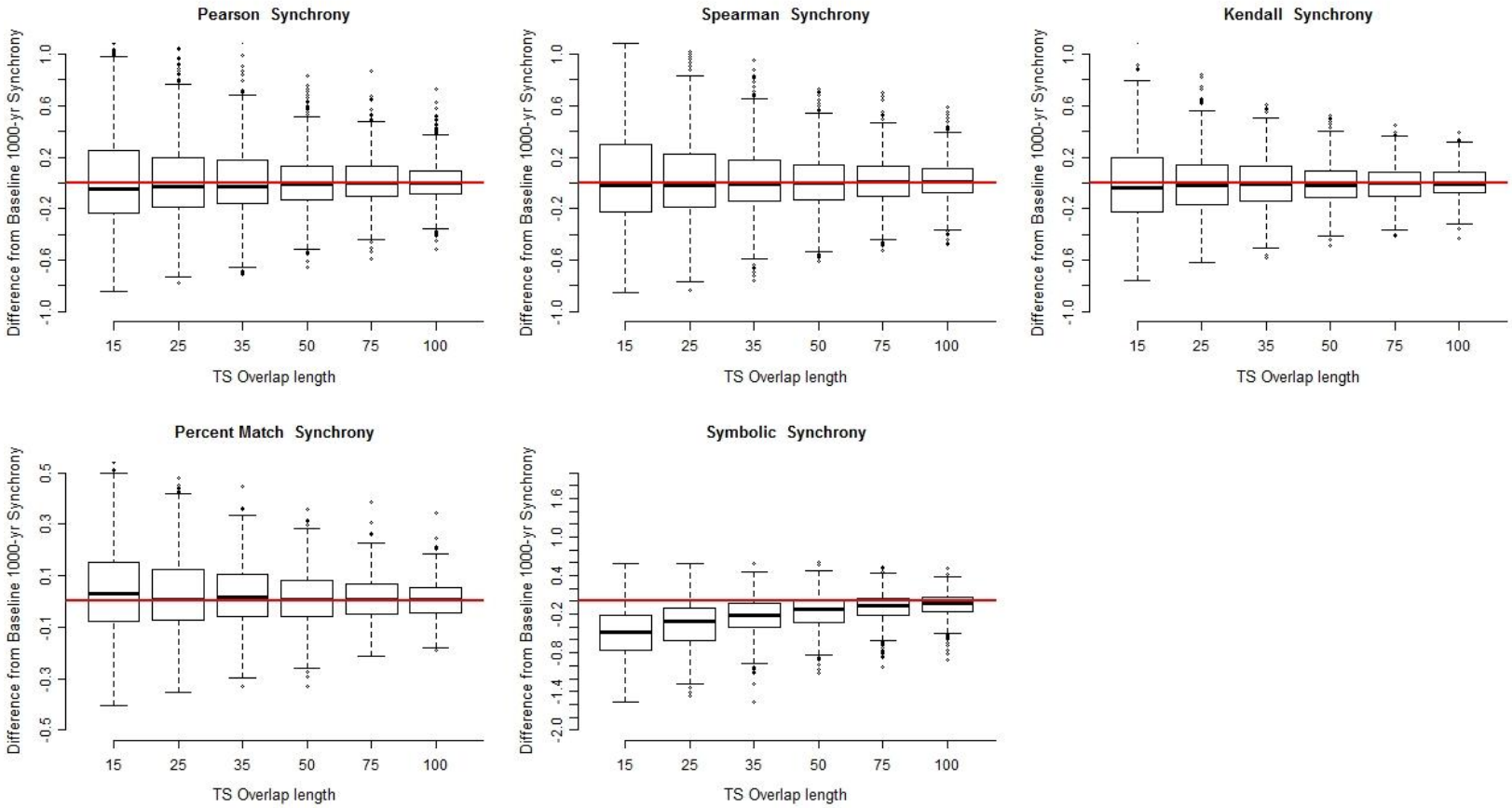
ERROR: 0.15

	Pearson	Spearman	Kendall	Percent Match	Symbolic
Pearson		1.00	1.00	0.97	0.92
Spearman	1.00		1.00	0.98	0.93
Kendall	1.00	1.00		0.98	0.94
Percent Match	0.98	0.99	0.99		0.95
Symbolic	0.93	0.94	0.95	0.96	

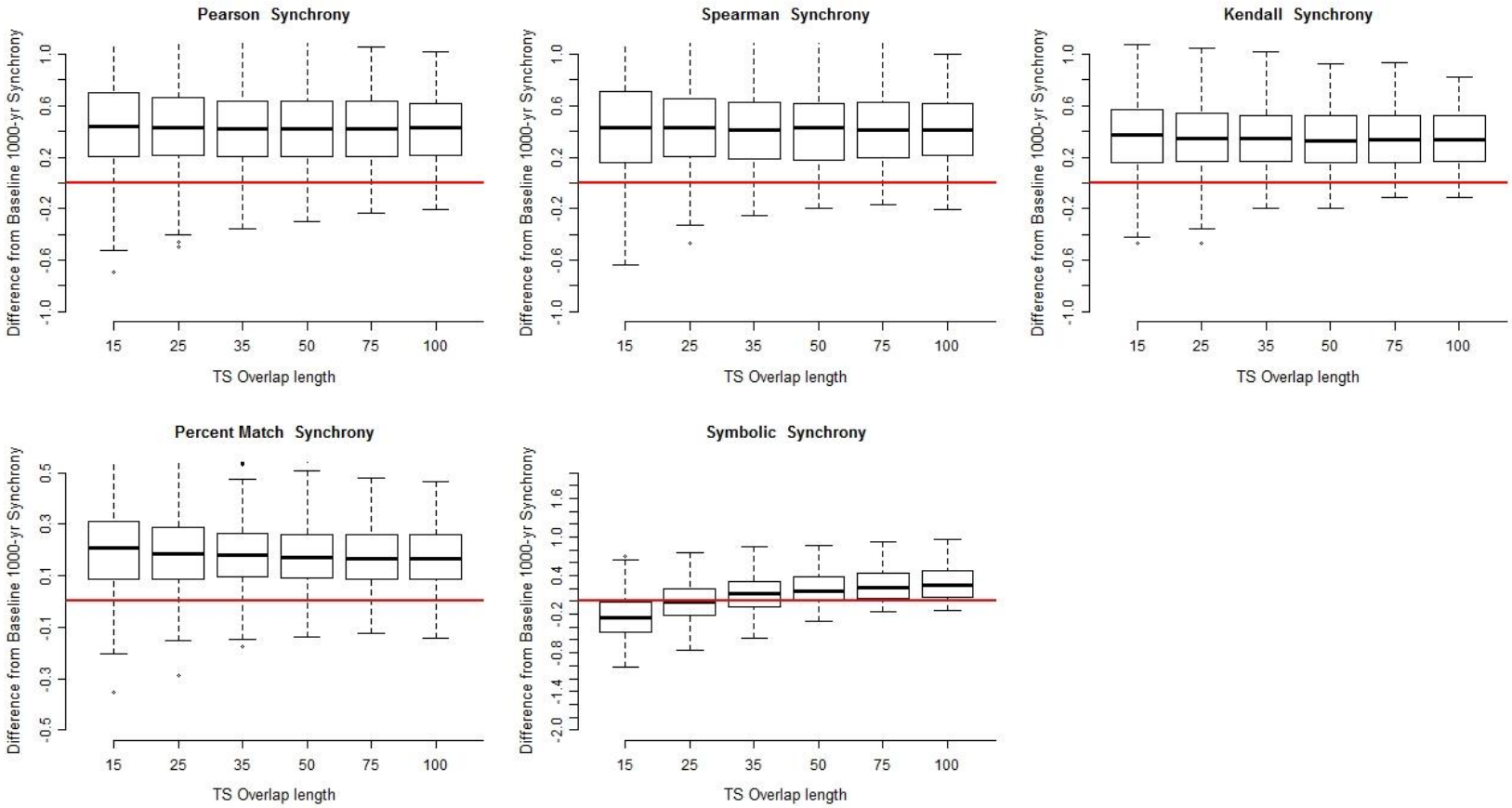
ERROR: 0.30

	Pearson	Spearman	Kendall	Percent Match	Symbolic
Pearson		0.98	0.98	0.87	0.78
Spearman	1.00		1.00	0.91	0.82
Kendall	1.00	1.00		0.91	0.82
Percent Match	0.96	0.98	0.98		0.88
Symbolic	0.91	0.92	0.92	0.95	

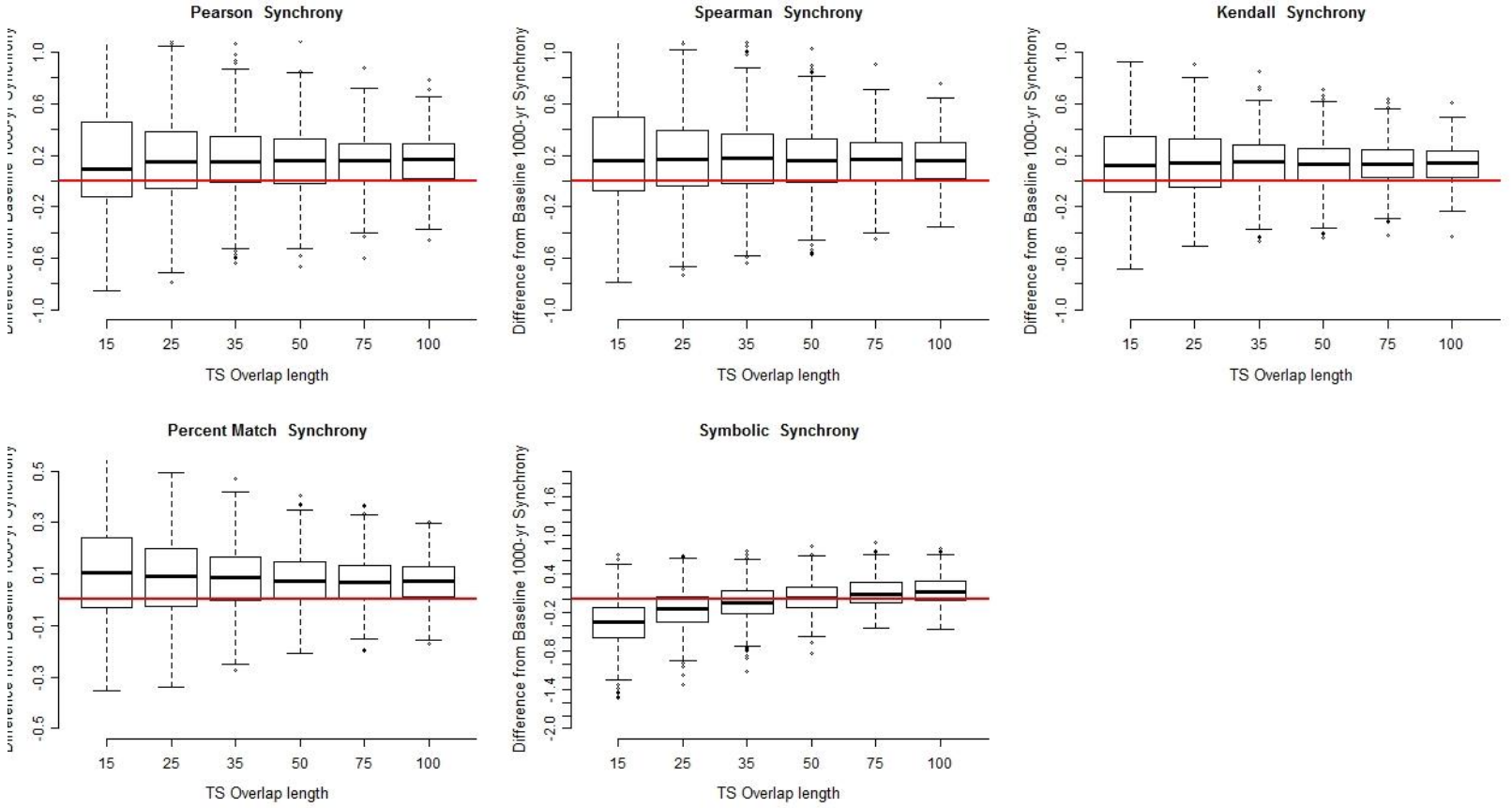
Appendix 3.2 [NON-CYCLIC TIME SERIES]
Correlation (Pearson's r) in synchrony estimates for each pair of metrics, under various scenarios of error (0, 0.15, 0.30). Correlations for unsmoothed time series are above the diagonal; smoothed time series, below diagonal.



Appendix 3.3 [CYCLIC, SMOOTHED TIME SERIES, ERROR = 0]
Boxplots showing difference in synchrony from baseline (1000-yr, error = 0, unsmoothed) synchrony. For each metric, results are shown for time series ranging in length from 15 to 100 years (x-axis). The red lines are the expected relationships for no difference between estimated and baseline synchrony. Values below the line (as for Symbolic metric) indicate synchrony is overestimated compared to baseline synchrony. In each boxplot, the center line is the median value, the box encloses the first to third quartiles of data, and box whiskers extend to 1.5 times the interquartile range. Data points exceeding this range are shown as circles.



Appendix 3.3 [CYCLIC, UNSMOOTHED TIME SERIES, ERROR = 0.30]
Boxplots showing difference in synchrony from baseline (1000-yr, error = 0, unsmoothed) synchrony. For each metric, results are shown for time series ranging in length from 15 to 100 years (x-axis). The red lines are the expected relationships for no difference between estimated and baseline synchrony. Values below the line (as for Symbolic metric) indicate synchrony is overestimated compared to baseline synchrony. In each boxplot, the center line is the median value, the box encloses the first to third quartiles of data, and box whiskers extend to 1.5 times the interquartile range. Data points exceeding this range are shown as circles.



Appendix 3.3 [CYCLIC, SMOOTHED TIME SERIES, ERROR = 0.30]
Boxplots showing difference in synchrony from baseline (1000-yr, error = 0, unsmoothed) synchrony. For each metric, results are shown for time series ranging in length from 15 to 100 years (x-axis). The red lines are the expected relationships for no difference between estimated and baseline synchrony. Values below the line (as for Symbolic metric) indicate synchrony is overestimated compared to baseline synchrony. In each boxplot, the center line is the median value, the box encloses the first to third quartiles of data, and box whiskers extend to 1.5 times the interquartile range. Data points exceeding this range are shown as circles.

		PEARSON					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.45	0.35	0.29	0.23	0.19	0.16
	0.15	0.50	0.35	0.31	0.25	0.20	0.18
	0.3	0.50	0.36	0.32	0.28	0.22	0.18
SMOOTH	0	0.78	0.57	0.49	0.41	0.32	0.28
	0.15	0.78	0.58	0.48	0.40	0.33	0.28
	0.3	0.71	0.55	0.45	0.38	0.29	0.25

		PERCENT MATCH **scaled					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.33	0.28	0.26	0.20	0.17	0.16
	0.15	0.33	0.28	0.26	0.24	0.17	0.16
	0.3	0.33	0.28	0.26	0.24	0.17	0.16
SMOOTH	0	0.47	0.44	0.37	0.28	0.20	0.18
	0.15	0.47	0.44	0.31	0.28	0.20	0.20
	0.3	0.47	0.36	0.31	0.28	0.20	0.18

		SPEARMAN					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.46	0.35	0.29	0.23	0.18	0.16
	0.15	0.49	0.34	0.29	0.24	0.20	0.17
	0.3	0.50	0.37	0.32	0.28	0.22	0.18
SMOOTH	0	0.76	0.59	0.49	0.41	0.32	0.27
	0.15	0.73	0.56	0.47	0.39	0.31	0.27
	0.3	0.68	0.53	0.45	0.38	0.28	0.26

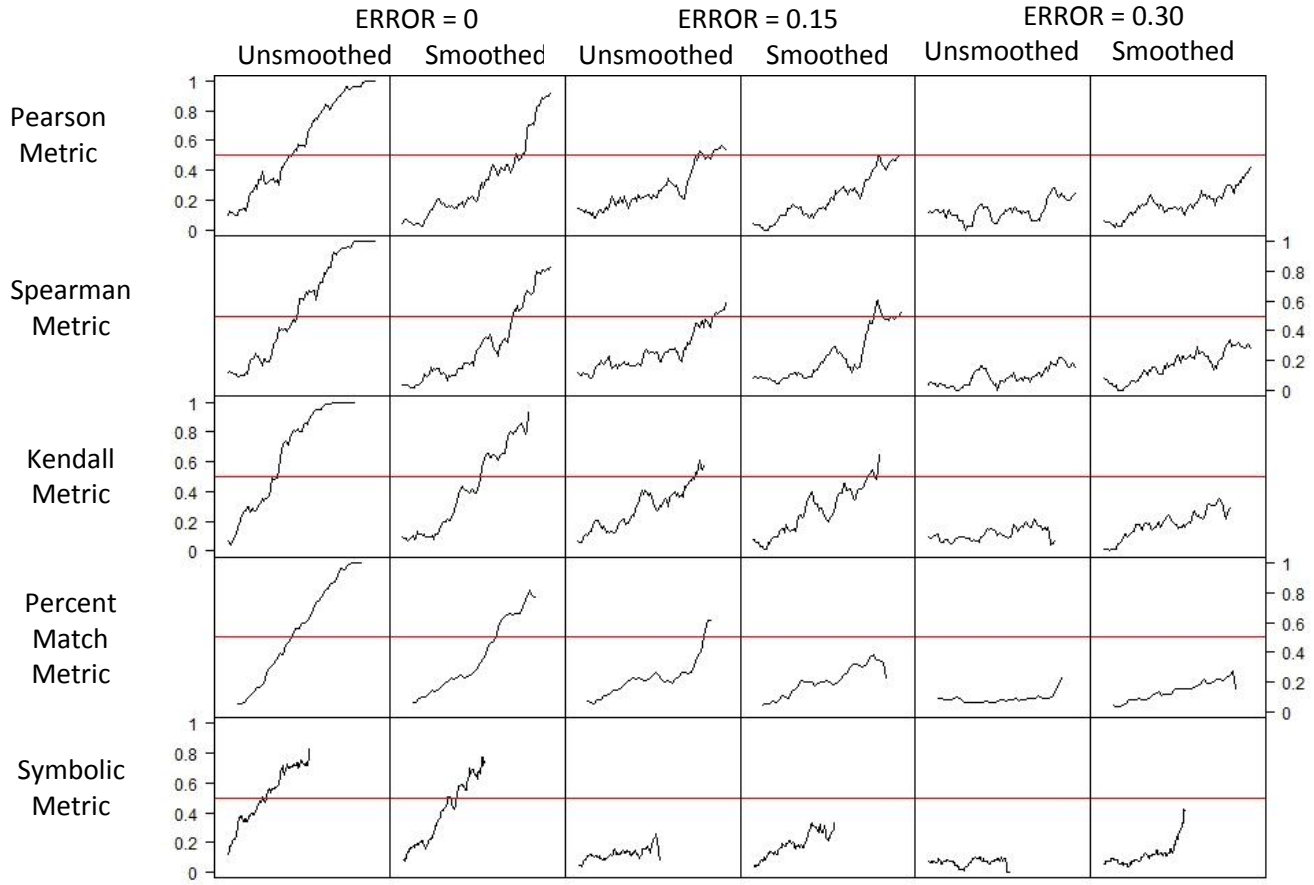
		SYMBOLIC (synchrony range 0 - 2.0)					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	1.00	0.64	0.46	0.31	0.20	0.13
	0.15	1.03	0.63	0.45	0.32	0.20	0.15
	0.3	1.00	0.63	0.47	0.30	0.20	0.15
SMOOTH	0	1.02	0.70	0.52	0.36	0.23	0.17
	0.15	1.09	0.70	0.51	0.35	0.23	0.17
	0.3	1.09	0.70	0.51	0.34	0.23	0.17

		KENDALL					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.34	0.25	0.20	0.16	0.12	0.11
	0.15	0.36	0.25	0.20	0.17	0.14	0.12
	0.3	0.36	0.26	0.23	0.19	0.15	0.13
SMOOTH	0	0.58	0.42	0.33	0.29	0.21	0.18
	0.15	0.56	0.41	0.33	0.26	0.21	0.18
	0.3	0.52	0.37	0.30	0.26	0.19	0.18

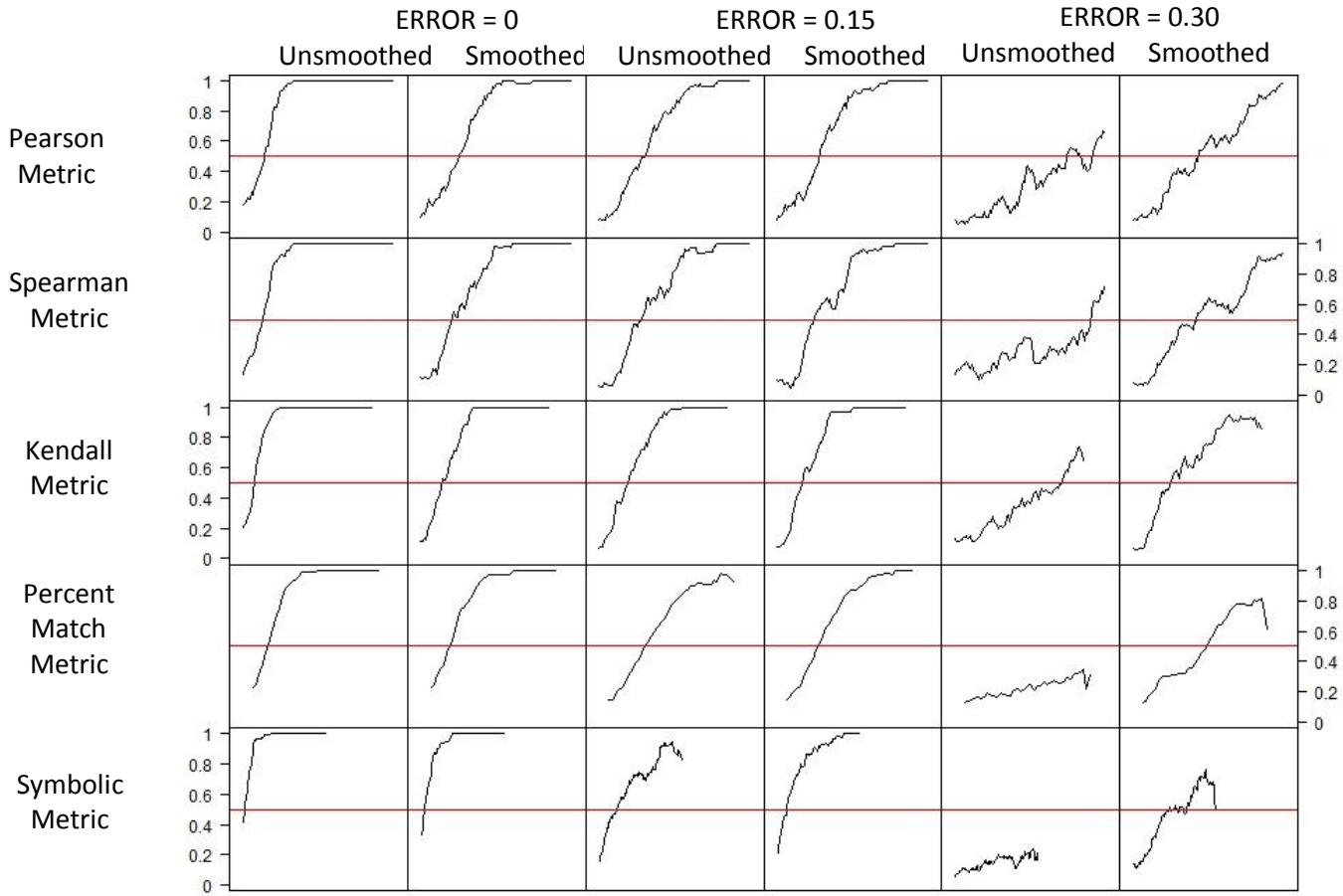
** Percent Match metric is scaled as (x-0.50)/range, where range of synchrony values for Percent Match = 0.5. This metric is scaled to facilitate comparison of metric performance with the correlation synchrony metrics. The Symbolic metric is not scaled because it applies a definition of synchrony that cannot be easily compared with the other metrics in this study.

Appendix 3.4 [NON-CYCLIC TIME SERIES]

The 95% upper confidence limit for 'no synchrony, calculated on independent (uncorrelated) time series. For each metric, upper panel = unsmoothed time series and lower panel = smoothed time series. For Percent Match and Symbolic metrics, data are not scaled, but the range of synchrony values for these metrics is provided in parentheses because they differ from the 0 – 1 range of other metrics.



Appendix 3.5 [NON-CYCLIC, 15-YEAR TIME SERIES]
For each metric (rows), proportion of time series pairs significant at α = 0.05 (y-axis) vs. baseline synchrony (x-axis; baseline means 1000-yr, error = 0, unsmoothed) under various scenarios of error (0, 0.15, 0.30) and data treatment (smoothed vs. unsmoothed). For Pearson, Spearman, and Kendall metrics, baseline synchrony ranges 0 – 1.0; for Percent Match metric, 0.5 – 1.0; for Symbolic metric, 0 – 2.0. Red lines indicate 50% of time series pairs are significant. This figure corresponds with data in Table 3.5.



Appendix 3.6 [NON-CYCLIC, 100-YEAR TIME SERIES]
For each metric (rows), proportion of time series pairs significant at α = 0.05 (y-axis) vs. baseline synchrony (x-axis; baseline means 1000-yr, error = 0, unsmoothed) under various scenarios of error (0, 0.15, 0.30) and data treatment (smoothed vs. unsmoothed). For Pearson, Spearman, and Kendall metrics, baseline synchrony ranges 0 – 1.0; for Percent Match metric, 0.5 – 1.0; for Symbolic metric, 0 – 2.0. Red lines indicate 50% of time series pairs are significant. This figure corresponds with data in Table 3.5.

		PEARSON					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.45	0.28	0.23	0.23	0.19	0.19
	0.15	0.79	0.57	0.55	0.49	0.32	0.33
	0.3	NA	NA	NA	NA	NA	0.73
SMOOTH	0	0.74	0.52	0.43	0.39	0.30	0.28
	0.15	NA	0.67	0.55	0.48	0.35	0.31
	0.3	NA	0.94	0.75	0.74	0.57	0.45

		PERCENT MATCH **scaled					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0.00	0.44	0.36	0.30	0.26	0.24	0.20
	0.15	0.84	0.78	0.68	0.52	0.36	0.34
	0.30	NA	NA	NA	NA	NA	NA
SMOOTH	0.00	0.62	0.58	0.48	0.36	0.28	0.22
	0.15	NA	NA	0.58	0.46	0.30	0.30
	0.30	NA	NA	NA	0.78	0.52	0.50

		SPEARMAN					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.48	0.39	0.33	0.21	0.17	0.18
	0.15	0.89	0.70	0.50	0.47	0.32	0.31
	0.3	NA	NA	NA	NA	NA	0.87
SMOOTH	0	0.72	0.62	0.48	0.34	0.29	0.24
	0.15	0.80	0.58	0.50	0.47	0.33	0.28
	0.3	NA	0.95	0.86	0.73	0.50	0.44

		SYMBOLIC (synchrony range 0 - 2.0)					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0.00	0.52	0.30	0.23	0.16	0.13	0.12
	0.15	NA	NA	0.93	0.88	0.43	0.32
	0.30	NA	NA	NA	NA	NA	NA
SMOOTH	0.00	0.64	0.55	0.44	0.22	0.16	0.14
	0.15	NA	0.78	0.77	0.43	0.30	0.21
	0.30	NA	NA	NA	1.05	0.89	0.53

		KENDALL					
	ERROR	15	25	35	50	75	100
UNSMOOTH	0	0.36	0.24	0.20	0.16	0.13	0.12
	0.15	0.77	0.44	0.37	0.35	0.27	0.23
	0.3	NA	NA	NA	NA	0.79	0.70
SMOOTH	0	0.53	0.39	0.33	0.29	0.18	0.18
	0.15	0.76	0.48	0.36	0.33	0.29	0.21
	0.3	NA	0.75	0.47	0.50	0.35	0.28

** Percent Match metric is scaled as (x-0.50)/range, where range of synchrony values for Percent Match = 0.5. This metric is scaled to facilitate comparison of metric performance with the correlation synchrony metrics. The Symbolic metric is not scaled because it applies a definition of synchrony that cannot be easily compared with other metrics in this study.

Appendix 3.7 [NON-CYCLIC TIME SERIES]

The 50% significance categories for each metric. This number indicates the minimum baseline synchrony value for which at least 50% of estimated synchrony values are significant at $\alpha = 0.05$. "NA" indicates no baseline synchrony value had at least 50% of estimated synchrony values significant at $\alpha = 0.05$.

Appendix 3.8
R code for the main components of this simulation study. To minimize repetition, some segments of code are presented only for the Pearson metric. These segments, which are identified by the text '----- REPEAT FOR EACH METRIC -----', typically require minor modifications in parameter names or values prior to application with other synchrony metrics.

```
#####################################################################
# CHAPTER 3--SYNCHRONY METRICS SIMULATION
# APPENDIX R CODE
#
# NOTES:
# 1) Code is presented in this order: a) simulation functions, b) code for the main simulation, c) graphing functions.
# 2) Code is a subset of the original code and presents only the main simulation components. In several areas I present the code for Pearson
# metric, then indicate the code needs to be repeated (usually with slight modifications in parameter names) for the other simulation metrics.
# 3) Cholesky decomposition equation generates an error if asked to correlate time series to r = 0 or r = 1. Therefore, I correlate time series
#  at r = 0.05 to 0.95. The actual range of correlation values after Cholesky decomposition is close to r = 0 to 1.0.
#####################################################################

#>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>> CODE FOR SIMULATION FUNCTIONS
#####################################################################
#
# FUNC.FIRST.LAST:  FUNCTION TO GENERATE A VECTOR OF FIRST AND LAST DATA COLUMNS & TOTAL NUMBER OF DATA COLUMNS
#
# Input
#         [x]       Matrix of time series data
#
# Output
#         [z]       Matrix where col1=column of first TS value, col2=column of last TS value, col3=length of TS
#
# Notes
#         [1]       This function is called by FUNC.SMOOTH
```

```
#
#########################################################################

Func.first.last<-function(x)
        {
        ##### INITIALIZE MATRIX #####
        z<-matrix(NA,nrow(x),3)

        ##### GENERATE FIRST-LAST MATRIX #####
        for (i in 1:nrow(x))
                {
                z[i,1]<-min(which(!is.na(x[i,])))
                z[i,2]<-max(which(!is.na(x[i,])))
                z[i,3]<-z[i,2]-z[i,1]+1
                }                       # END FOR-I
        return(z)
        }                               # END FUNC.FIRST.LAST


#########################################################################
#
# FUNC.SMOOTH:  FUNCTION TO SMOOTH TIME SERIES
#
# Input
#       [x]             Matrix or vector of (raw) time series data
#       [yr.vec]        X-axis vector of years for time series data
#       [prec]          Vector or value of data precision
#       [win.span]      Approximate number of neighboring datapoints used to smooth each point
#       [rows]          Number of rows in 'x' (if 'x' is a vector, rows=1)
#
```

```
# Output
#       [z]     Matrix or vector of smoothed time series data, same dimensions as 'x'
#
# Notes
#       [1]     This function code has some components that are not necessary for this simulation, but are included
#               because they are necessary for other simulations that call this function.
#       [2]     Loess is a nearest neighbor smoother, using (100 x span) % of the data to estimate at each point.
#               Default span is 7/(length of TS). This default was chosen by eyeballing the graphs of smoothed TS.
#               This span smooths TS in a way that seems, on average, to best capture the primary fluctuations in
#               TS and smooth over potential 'noise'. It is equivalent to using 7 surrounding datapoints (weighting
#               normally distributed) to smooth each point.
#
#########################################################################

Func.smooth<-function(x,yr.vec,prec,win.span,rows)
        {
        if(rows==1){x<-t(as.matrix(x))}                 # if x is a vector, convert x to a matrix

        ##### INITIALIZE MATRICES #####
        z<-matrix(NA,rows,ncol(x))                              # initialize output matrix
        first.last<-matrix(NA,rows,3)                           # initialize first-last matrix

        ##### SMOOTH THE DATA #####
        first.last<-Func.first.last(x)
        for (i in 1:rows)
                {
                x[i,]<-x[i,]*(1/prec[i])                # multiply every value in the time series by 1/precision
                len<-first.last[i,3]
                z[i,first.last[i,1]:first.last[i,2]]<-loess.smooth(yr.vec,x[i,],span=(win.span/len),family="gaussian",
                evaluation=len,na.rm=TRUE)$y
```

```
                                                                    # loess.smooth returns a list ($x) of years and a list ($y) of estimates
                z[i,]<-z[i,]*prec[i]
                z[i,]<-ifelse((z[i,]%%prec[i])>(prec[i]/2),(trunc(z[i,]/prec[i])+1)*prec[i],(trunc(z[i,]/prec[i])*prec[i]))
                                                                    # restore time series to original precision
                z[i,]<-ifelse(z[i,]<prec[i],prec[i],z[i,])      # replace zeroes in smoothed TS with lowest possible value
                }                                                   # END FOR-I
        return(z)
        }                                                           # END FUNC.SMOOTH


#########################################################################
#
# FUNC.CHOLESKY:  FUNCTION TO CORRELATE TIME SERIES BY CHOLESKY DECOMPOSITION
#
# Input
#       [dat]   Matrix with 2 rows, each row has a separate time series
#       [r]     Desired correlation level
#
# Output
#       [dat2]  Matrix with 2 rows, each row has a separate time series--they are correlated as specified
#
#########################################################################

Func.Cholesky<-function(dat,r)
        {
        ##### INITIALIZE MATRICES #####
        dat2<-matrix(NA,2,ncol(dat))
        R<-matrix(NA,2,2)

        ##### CORRELATE TIME SERIES #####
```

```
        R<-matrix(c(1,r,r,1),2,2)
        ch<-chol(R)

        dat2<-t(ch)%*%dat                       # dat2 matrix now has the two time series correlated with specified rho

        return(dat2)
        }                                       # END FUNC.CHOLESKY


#########################################################################
#
# FUNC.PERCMATCH.CATEG:  FUNCTION TO ASSIGN PECENT MATCH CATEGORIES TO A TIME SERIES MATRIX
#
# Input
#       [x]     Matrix or vector of time series data
#       [rows]  Number of rows in 'x' (if 'x' is a vector, rows=1)
#
# Output
#       [z]     Matrix or vector of PercMatch-categorized time series data (-1,0,1)
#
#########################################################################

Func.PercMatch.categ<-function(x,rows)
        {
        if(rows==1){x<-t(as.matrix(x))}                 # if x is a vector, convert x to a matrix

        ##### INITIALIZE VECTORS #####
        a.TS<-matrix(NA,rows,ncol(x))                   # pre-TS matrix for categorizing A.synchrony
        b.TS<-matrix(NA,rows,ncol(x))                   # mid-TS matrix for categorizing A.synchrony
```

```
        diff<-matrix(NA,rows,ncol(x))                   # matrix of b.TS - a.TS
        z<-matrix(NA,rows,ncol(x))                      # matrix of TS data transformed to A.synchrony categories

        ##### ASSIGN MATRICES #####
        a.TS[,2:ncol(x)]<-x[,1:(ncol(x)-1)]
        b.TS[,2:ncol(x)]<-x[,2:ncol(x)]

        diff<-b.TS-a.TS
        z<-diff/abs(diff)                               # convert diff to logicals -1,0,1
        z[is.nan(z)]<-0                                 # convert NaN's to 0's
        return(z)
        }                                               # END FUNC.PERCMATCH.CATEG


#########################################################################
#
# FUNC.PHASE.CATEG:  FUNCTION TO ASSIGN PHASE CATEGORIES TO TIME SERIES
#
# Input
#       [x]     Matrix or vector of time series data
#       [rows]  Number of rows in 'x' (if 'x' is a vector, rows=1)
#
# Output
#       [z]     Matrix or vector of phase-categorized time series data
#
# Notes
#       [1]     In some cases, the time series has consecutive numbers of the same value (i.e., TS not increasing
#               or decreasing). For these, I determined phase category by looking forward or backward to the first
#               unequal value. For example, 4-3-3-3-3-2 would be categorized as D-D-D-D-D and 4-3-3-3-5-7 would be
#               categorized as T-T-T-I (where D=Decrease, T=Trough, I=Increase). Sometimes this causes “problems”
```

```
#                when the string of consecutive numbers comes at the beginning or end of a time series. For example,
#                if the time series begins with 4-4-4-4-5-7 I would categorize this as NA-NA-NA-I because, without
#                knowing what number precedes the string of 4's I can't tell if the 4's are trough or increasing
#                years. This is a modification of Cazelles (2004), which did not specify how to categorize such trends.
#
#########################################################################

Func.phase.categ<-function(x,rows)
        {
        if(rows==1){x<-t(as.matrix(x))}                 # if x is a vector, convert x to a matrix
        yrs<-ncol(x)

        ##### INITIALIZE VECTORS #####
        a.TS<-matrix(NA,rows,ncol(x))                   # pre-TS matrix for categorizing phase synchrony
        b.TS<-matrix(NA,rows,ncol(x))                   # mid-TS matrix for categorizing phase synchrony
        c.TS<-matrix(NA,rows,ncol(x))                   # post-TS matrix for categorizing phase synchrony

        pre.diff<-matrix(NA,rows,ncol(x))               # matrix of b.TS - a.TS
        post.diff<-matrix(NA,rows,ncol(x))              # matrix of c.TS - b.TS

        conv.pre.diff<-matrix(NA,rows,ncol(x))          # pre.diff matrix converted to logicals (-1,0,1) and *2
        conv.post.diff<-matrix(NA,rows,ncol(x))         # post.diff matrix converted to logicals (-1,0,1) and *3

        numeric.cat<-matrix(NA,rows,ncol(x))            # matrix symbolizing TS data as numeric growth categories
        z<-matrix(NA,rows,ncol(x))                      # matrix of TS data transformed to phase categories

        ##### ASSIGN MATRICES #####
        a.TS[,2:(ncol(x)-1)]<-x[,1:(ncol(x)-2)]
        b.TS[,2:(ncol(x)-1)]<-x[,2:(ncol(x)-1)]
        c.TS[,2:(ncol(x)-1)]<-x[,3:ncol(x)]
```

```
num.list<-c(-5,-1,1,5)                          # numeric equivalent of growth categories
cat.list<-c("D","P","T","I")                    # category equivalent of growth categories

##### CATEGORIZE DATA AS I(ncrease),P(eak),D(ecrease),T(rough) #####
pre.diff<-b.TS-a.TS
post.diff<-c.TS-b.TS

conv.pre.diff<-2*(pre.diff/abs(pre.diff))       # convert pre.diff to logicals -1,0,1 then multiply by 2
conv.post.diff<-3*(post.diff/abs(post.diff))    # convert post.diff to logicals -1,0,1 then multiply by 3

for (i in 1:rows)
        {
        which.pre.nan<-which(is.nan(conv.pre.diff[i,]))         # which pre.diff columns are FLAT
        pre.lag<-1                                              # number of values to go back for flats
        conv.pre.NA<-which((which.pre.nan-pre.lag)<1)           # avoid 'subscript out of bounds' error
        conv.pre.diff[i,which.pre.nan[conv.pre.NA]]<-NA
        if(length(conv.pre.NA)>0){which.pre.nan<-which.pre.nan[-conv.pre.NA]}

        while(length(which.pre.nan)>0)                          # while there are still flat pre.diff columns
                {                                               # keep looking back one step to decide appropriate category
                replace.pre.nan<-b.TS[i,which.pre.nan]-a.TS[i,(which.pre.nan-pre.lag)]          # substitute pre.diff vector
                conv.pre.diff[i,which.pre.nan]<-2*(replace.pre.nan/abs(replace.pre.nan))
                which.pre.nan<-which(is.nan(conv.pre.diff[i,]))
                pre.lag<-pre.lag+1                              # keep going back one TS value

                conv.pre.NA<-which((which.pre.nan-pre.lag)<1)           # avoid 'subscript out of bounds' error
                conv.pre.diff[i,which.pre.nan[conv.pre.NA]]<-NA
                if(length(conv.pre.NA)>0){which.pre.nan<-which.pre.nan[-conv.pre.NA]}
                }                                               # END WHILE-LENGTH
```

```
        which.post.nan<-which(is.nan(conv.post.diff[i,]))       # which post.diff columns are FLAT
        post.lag<-1                                             # number of values to go forward for flats
        conv.post.NA<-which((which.post.nan+post.lag)>yrs)      # avoid 'subscript out of bounds' error
        conv.post.diff[i,which.post.nan[conv.post.NA]]<-NA
        if(length(conv.post.NA)>0){which.post.nan<-which.post.nan[-conv.post.NA]}

        while(length(which.post.nan)>0)                         # while there are still flat post.diff columns
                {                                               # keep looking forward one step to decide appropriate category
                replace.post.nan<-c.TS[i,(which.post.nan+post.lag)]-b.TS[i,which.post.nan]      # substitute post.diff vector
                conv.post.diff[i,which.post.nan]<-3*(replace.post.nan/abs(replace.post.nan))
                which.post.nan<-which(is.nan(conv.post.diff[i,]))
                post.lag<-post.lag+1                            # keep going forward one TS value

                conv.post.NA<-which((which.post.nan+post.lag)>yrs)# avoid 'subscript out of bounds' error
                conv.post.diff[i,which.post.nan[conv.post.NA]]<-NA
                if(length(conv.post.NA)>0){which.post.nan<-which.post.nan[-conv.post.NA]}
                }                                               # END WHILE-LENGTH
        }                                                       # END FOR-I

numeric.cat<-conv.pre.diff+conv.post.diff                       # sum of pre&post.diff; represents TS vector as numeric categories
z<-numeric.cat

for(i in 1:length(num.list))
        {
        z<-replace(z,which(numeric.cat==num.list[i]),cat.list[i])  # matrix of growth categories
        }                                                       # END FOR-I
return(z)
}                                                               # END FUNC.PHASE.CATEG
```

```
#########################################################################
#
# FUNC.PHASE.ENTROPY:  FUNCTION TO CALCULATE ENTROPY OF PHASE-CATEGORIZED TIME SERIES
#
# Input
#         [x.vec]   Vector of a phase-categorized time series
#
# Output
#         [z.val]   Value representing entropy of 'x'
#########################################################################

Func.phase.entropy<-function(x.vec)
        {
        if(all(is.na(x.vec))){z.val<-NA}else                    # if vector all NA's, entropy value is NA (not zero!)
                {
                cat.prob<-table(x.vec)/sum(!is.na(x.vec))       # for each TS, probability of each category
                log2.cat.prob<-ifelse(cat.prob==0,0,log2(cat.prob))     # if prob=0, log2=0
                z.val<-(-1)*sum(cat.prob*log2.cat.prob)
                }                                               # END IF-ELSE
        return(z.val)
        }                                                       # END FUNC.PHASE.ENTROPY


#########################################################################
#
# FUNC.JOINT.ENTROPY:  FUNCTION TO CALCULATE JOINT ENTROPY OF TWO PHASE-CATEGORIZED TIME SERIES
#
# Input
#         [x,y]     Two phase-categorized time series matrices or vectors
```

```
#       [rows]  Number of rows in 'x' & 'y' (if vectors, rows=1)
#
# Output
#       [z]     Joint entropy(ies) for the time series
#
#############################################################################

Func.joint.entropy<-function(x,y,rows)
        {
        if(rows==1)             # if x & y are vectors, convert them to matrices
                {
                x<-t(as.matrix(x))
                y<-t(as.matrix(y))
                }               # END IF-ROWS
        joint.data<-rowSums(!is.na(x)&!is.na(y))        # per row, sums number of cells of x that have joint data with y
        cum.entropy<-0
        for (k in c("D","I","P","T"))                   # for each category in x
                {
                for (l in c("D","I","P","T"))           # for each category in y
                        {
                        cat.pair.count<-rowSums((x==k)&(y==l),na.rm=TRUE)                # per row, number of cells that meet criteria
                        cat.pair.prob<-cat.pair.count/joint.data                         # probability vector of specified cat.pair
                        log2.cat.pair.prob<-ifelse(cat.pair.prob==0,0,log2(cat.pair.prob))       # if prob=0, log2=0
                        cum.entropy<-cum.entropy+cat.pair.prob*log2.cat.pair.prob
                        }                                                       # END FOR-L
                }                                                               # END FOR-K
                z<-(-1)*cum.entropy                                             # joint entropy for TS pair
                return(z)
        }                                                                       # END FUNC.JOINT.ENTROPY
```

```
#>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>> CODE FOR MAIN SIMULATION

######################################################################
#
# INITIALIZE VARIABLES FOR SIMULATION
#
######################################################################

# Simulation parameters selected by user
Base.TS.vec<-[CAN READ IN TIME SERIES FROM AN EXTERNAL FILE]    # Actual time series data (abundance estimates) that the simulated time series will be based on
Smooth.def<-7                                   # Span parameter for loess smoothing
Tot.sub<-1000                                   # Number of surrogate time series pairs to generate for each Base TS
Tot.pick<-500                                   # Number of surrogate time series pairs to select for each metric (multiple of 10)
Max.TS.length<-1000                             # Maximum TS length simulated
Overlap<-c(15,25,35,50,75,100,1000)             # TS Overlap lengths to examine
Err.SD<-c(0,.15,.3)                             # Data error rates to simulate
Meth<-c("Pearson","Spearman","Kendall","PercMatch","Symbolic")        # Synchrony metrics
Treat<-c("U","Sm")                                                    # Simulation data treatment (smooth vs. not smooth the data)
Base.prec<-0.01                 # Specify the precision of time series data (e.g., 0.1, 0.01) so simulated smoothed data will have same precision

# Array to store synchrony estimates for all 1000-yr TS pairs. This array will be used by Func.Plot.Metrics to compare metrics.
Full.dat<-array(NA,dim=c(length(Err.SD),length(Treat),length(Meth),Tot.sub))
dimnames(Full.dat)<-list(Err.SD,Treat,Meth,1:Tot.sub)

# Arrays to store synchrony estimates for each metric ----- REPEAT FOR EACH METRIC -----
Sim.Pearson.ary<-array(NA,dim=c(length(Err.SD),length(Treat),length(Overlap),Tot.pick))
dimnames(Sim.Pearson.ary)<-list(Err.SD,Treat,paste("Overlap.",Overlap,sep=""),1:Tot.pick)
```

```
# Arrays to store 'no synchrony' (i.e., pre-Cholesky) estimates for each metric ----- REPEAT FOR EACH METRIC -----
N.Sim.Pearson.ary<-array(NA,dim=c(length(Err.SD),length(Treat),length(Overlap),Tot.sub))
dimnames(N.Sim.Pearson.ary)<-list(Err.SD,Treat,paste("Overlap.",Overlap,sep=""),1:Tot.sub)

# Arrays to store difference between synchrony and baseline synchrony estimates ----- REPEAT FOR EACH METRIC
D.Sim.Pearson.ary<-array(NA,dim=c(length(Err.SD),length(Treat),length(Overlap),Tot.pick))
dimnames(D.Sim.Pearson.ary)<-list(Err.SD,Treat,paste("Overlap.",Overlap,sep=""),1:Tot.pick)

# Arrays for power analysis ----- REPEAT FOR EACH METRIC -----
Power.Pearson.ary<-array(NA,dim=c(length(Err.SD),length(Treat),length(Overlap)-1,200))
dimnames(Power.Pearson.ary)<-list(Err.SD,Treat,paste("Overlap.",c(15,25,35,50,75,100),sep=""),1:200)

# Arrays for power analysis calculation of baseline synchrony estimates with at least 50% significant values ----- REPEAT FOR EACH METRIC -----
Sig50.Pearson.ary<-array(NA,dim=c(length(Err.SD),length(Treat),length(Overlap)-1))
dimnames(Sig50.Pearson.ary)<-list(Err.SD,Treat,paste("Overlap.",c(15,25,35,50,75,100),sep=""))

# Interim arrays for calculating correlation synchrony (Pearson, Spearman, Kendall)
a.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
b.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
a.diff.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
b.diff.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))

N.a.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
N.b.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
N.a.diff.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
N.b.diff.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))

# Interim arrays for calculating Percent Match synchrony
a.PercMatch.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
```

```
b.PercMatch.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
check.match<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))

N.a.PercMatch.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
N.b.PercMatch.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
N.check.match<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))

# Interim arrays for calculating Symbolic synchrony
a.phase.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
b.phase.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))

HS.a<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub))
HS.b<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub))
HSU.sim<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub))

N.a.phase.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))
N.b.phase.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length))

N.HS.a<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub))
N.HS.b<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub))
N.HSU.sim<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub))

# Other arrays
Temp.sync.categ<-matrix(NA,length(Meth),Tot.sub)                # Temporarily store baseline synchrony decile categories for all metrics
rownames(Temp.sync.categ)<-Meth

Big.ary<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length,2))              # Store the post-Cholesky TS data
dimnames(Big.ary)<-list(Err.SD,Treat,1:Tot.sub,1:Max.TS.length,1:2)

N.Big.ary<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Max.TS.length,2)) # Store the pre-Cholesky TS data to evaluate 'no synchrony'
```

```
dimnames(N.Big.ary)<-list(Err.SD,Treat,1:Tot.sub,1:Max.TS.length,1:2)

Select.col.mat<-matrix(NA,5,Tot.pick)                    # Store the index numbers representing the 500 selected TS pairs for each metric
rownames(Select.col.mat)<-Meth

Met.Pearson<-array(NA,dim=c(length(Err.SD),length(Treat),length(Meth),length(Meth)))          # Array of Pearson correlation btwn metrics
dimnames(Met.Pearson)<-list(Err.SD,Treat,Meth,Meth)

No.sync.summ<-array(NA,dim=c(length(Meth),length(Err.SD),length(Treat),length(Overlap)))     #Array to store 95% upper CI bound
dimnames(No.sync.summ)<-list(Meth,Err.SD,Treat,Overlap)

SD.1000UE.ary<-array(NA,dim=c(length(Err.SD),length(Treat),length(Meth),length(Overlap)))
        #  Array of SD of synchrony difference from baseline
dimnames(SD.1000UE.ary)<-list(Err.SD,Treat,Meth,paste("Overlap.",Overlap,sep=""))

Bias.1000UE.ary<-array(NA,dim=c(length(Err.SD),length(Treat),length(Meth),length(Overlap)))
        #  Array of bias of synchrony difference from baseline
dimnames(Bias.1000UE.ary)<-list(Err.SD,Treat,Meth,paste("Overlap.",Overlap,sep=""))


#########################################################################
#
# BEGIN SIMULATION
#
#########################################################################

##### GENERATE PAIRS OF ERROR = 0, UNSMOOTHED TIME SERIES, BASED ON AR COEFFICIENTS FOR THE BASE TIME SERIES
for (i in 1:Tot.sub)
        # The arima parameters for generating surrogate TS should be determined separately for each Base TS--replace in arima.sim eqn
        {
```

```
        N.Big.ary[1,1,i,,1]<-arima.sim(n = Max.TS.length, list(ar =.6835, ma = c(0,0)),sd = sqrt(.3236))
        N.Big.ary[1,1,i,,2]<-arima.sim(n = Max.TS.length, list(ar =.6835, ma = c(0,0)),sd = sqrt(.3236))
        }                               # END FOR-I

##### APPLY ERROR AND DATA SMOOTHING TO 'NO SYNCHRONY' (PRE-CHOLESKY) TIME SERIES
# Generate error in 'no synchrony'(pre-Cholesky) time series as Err.SD X half the range of each surr TS pair
for (e in 2:length(Err.SD))             # Omit first Err.SD, which which error = 0
        {
        for (i in 1:Tot.sub)
                {
                range.half<-(max(N.Big.ary[1,1,i,,])-min(N.Big.ary[1,1,i,,]))/2
                N.Big.ary[e,1,i,,1]<-N.Big.ary[1,1,i,,1]+rnorm(Max.TS.length,0,Err.SD[e]*range.half)
                N.Big.ary[e,1,i,,2]<-N.Big.ary[1,1,i,,2]+rnorm(Max.TS.length,0,Err.SD[e]*range.half)
                }                       # END FOR-I
        }                               # END FOR-E

# Call FUNC.SMOOTH to smooth 'no synchrony' (pre-Cholesky) time series
for (e in 1:length(Err.SD))
        {
        N.Big.ary[e,2,,,1]<-Func.smooth(N.Big.ary[e,1,,,1],1:Max.TS.length,rep(Base.prec,Tot.sub),Smooth.def,Tot.sub)
        N.Big.ary[e,2,,,2]<-Func.smooth(N.Big.ary[e,1,,,2],1:Max.TS.length,rep(Base.prec,Tot.sub),Smooth.def,Tot.sub)
        }                               # END FOR-E

##### CHOLESKY DECOMPOSITION TO CORRELATE ERROR =0, UNSMOOTHED TIME SERIES
Rho.vec<-.01*seq(5,95,(90/Tot.sub))             # Vector of Pearson correlations to simulate, from r = 0.05 to 0.95

for (i in 1:Tot.sub)
        {
        Big.ary[1,1,i,,]<-t(Func.Cholesky(t(N.Big.ary[1,1,i,,]),Rho.vec[i]))
        }                               # END FOR-I
```

```
##### APPLY ERROR AND DATA SMOOTHING TO CORRELATED (POST-CHOLESKY) TIME SERIES
# Generate error in correlated (post-Cholesky) time series as Err.SD X half the range of each surr TS pair
for (e in 2:length(Err.SD))
        {
        for (i in 1:Tot.sub)
                {
                range.half<-(max(Big.ary[1,1,i,,])-min(Big.ary[1,1,i,,]))/2
                Big.ary[e,1,i,,1]<-Big.ary[1,1,i,,1]+rnorm(Max.TS.length,0,Err.SD[e]*range.half)
                Big.ary[e,1,i,,2]<-Big.ary[1,1,i,,2]+rnorm(Max.TS.length,0,Err.SD[e]*range.half)
                }                               # END FOR-I
        }                                       # END FOR-E

# Call FUNC.SMOOTH to smooth correlated (post-Cholesky) time series
for (e in 1:length(Err.SD))
        {
        Big.ary[e,2,,,1]<-Func.smooth(Big.ary[e,1,,,1],1:Max.TS.length,rep(Base.prec,Tot.sub),Smooth.def,Tot.sub)
        Big.ary[e,2,,,2]<-Func.smooth(Big.ary[e,1,,,2],1:Max.TS.length,rep(Base.prec,Tot.sub),Smooth.def,Tot.sub)
        }           # END FOR-E

##### FOR EACH METRIC, SELECT A SUBSET OF TIME SERIES PAIRS REPRESENTING AN APPROXIMATELY UNIFORM DISTRIBUTION OF
#    SYNCHRONY VALUES FROM NO SYNCHRONY TO FULL SYNCHRONY. APPLY THESE STEPS TO ERROR = 0, UNSMOOTHED, CORRELATED TIME
#    SERIES (I.E., CORRELATED BASELINE TIME SERIES):
#       1) CALCULATE METRIC-SPECIFIC SYNCHRONY OF ALL TIME SERIES PAIRS
#       2) CATEGORIZE BY SYNCHRONY ESTIMATE
#       3) SELECT SUBSET OF TIME SERIES
#    IN THIS SECTION, CALCULATE SYNCHRONY OF 'NO SYNCHRONY' BASELINE TIME SERIES FOR GENERATING 'NO SYNCHRONY' DISTRIBUTIONS
#####

##### THIS IS STEP 1:  CALCULATE METRIC-SPECIFIC SYNCHRONY OF ALL TIME SERIES PAIRS
```

```
for (e in 1:length(Err.SD))                                          # For each error rate
        {
        for (t in 1:length(Treat))                                   # For each simulation type
                {
                ##### FOR PEARSON, SPEARMAN & KENDALL METRICS, CALCULATE SYNCHRONY OF ALL BASELINE TIME SERIES PAIRS
                # Find first differenced natural logs for correlation metrics
                a.ln[e,t,,]<-log(Big.ary[e,t,,,1])                   # Natural log of TS data for TS1 of each pair
                b.ln[e,t,,]<-log(Big.ary[e,t,,,2])                   # Natural log of TS data for TS2 of each pair
                a.diff.ln[e,t,,2:Max.TS.length]<-t(apply(a.ln[e,t,,],1,diff))
                        # Calculate difference btwn ln values of consecutive years & assign to latter year
                b.diff.ln[e,t,,2:Max.TS.length]<-t(apply(b.ln[e,t,,],1,diff))

                # Repeat for the 'no synchrony' time series
                N.a.ln[e,t,,]<-log(N.Big.ary[e,t,,,1])               # Natural log of TS data for TS1 of each pair
                N.b.ln[e,t,,]<-log(N.Big.ary[e,t,,,2])               # Natural log of TS data for TS2 of each pair
                N.a.diff.ln[e,t,,2:Max.TS.length]<-t(apply(N.a.ln[e,t,,],1,diff))
                        # Calculate difference btwn ln values of consecutive years & assign to latter yr
                N.b.diff.ln[e,t,,2:Max.TS.length]<-t(apply(N.b.ln[e,t,,],1,diff))

                # Calculate synchrony for correlation metrics
                for (i in 1:Tot.sub)
                        {
                        Full.dat[e,t,"Pearson",i]<-cor(a.diff.ln[e,t,i,],b.diff.ln[e,t,i,],method="pearson",use="na.or.complete")
                        Full.dat[e,t,"Spearman",i]<-cor(a.diff.ln[e,t,i,],b.diff.ln[e,t,i,],method="spearman",use="na.or.complete")
                        Full.dat[e,t,"Kendall",i]<-cor(a.diff.ln[e,t,i,],b.diff.ln[e,t,i,],method="kendall",use="na.or.complete")

                        N.Sim.Pearson.ary[e,t,length(Overlap),i]<-
cor(N.a.diff.ln[e,t,i,],N.b.diff.ln[e,t,i,],method="pearson",use="na.or.complete")
                        N.Sim.Spearman.ary[e,t,length(Overlap),i]<-
cor(N.a.diff.ln[e,t,i,],N.b.diff.ln[e,t,i,],method="spearman",use="na.or.complete")
```

```
        N.Sim.Kendall.ary[e,t,length(Overlap),i]<-cor(N.a.diff.ln[e,t,i,],N.b.diff.ln[e,t,i,],method="kendall",use="na.or.complete")
        }                   # END FOR-I

##### FOR PERCENT MATCH METRIC, CALCULATE SYNCHRONY OF ALL BASELINE TIME SERIES PAIRS
# Call FUNC.PERCMATCH.CATEG to find matching trends (between time series in a pair) for Percent Match metric
# Categorize matrix of time series data as increase, decrease, or no change
a.PercMatch.cat[e,t,,]<-Func.PercMatch.categ(Big.ary[e,t,,,1],Tot.sub)
b.PercMatch.cat[e,t,,]<-Func.PercMatch.categ(Big.ary[e,t,,,2],Tot.sub)
check.match[e,t,,]<-b.PercMatch.cat[e,t,,]-a.PercMatch.cat[e,t,,]          # Puts zero in every cell where series dxn matches

# Categorize matrix of time series data as increase, decrease, or no change
N.a.PercMatch.cat[e,t,,]<-Func.PercMatch.categ(N.Big.ary[e,t,,,1],Tot.sub)
N.b.PercMatch.cat[e,t,,]<-Func.PercMatch.categ(N.Big.ary[e,t,,,2],Tot.sub)
N.check.match[e,t,,]<-N.b.PercMatch.cat[e,t,,]-N.a.PercMatch.cat[e,t,,]

# Calculate synchrony for Percent Match metric
Full.dat[e,t,"PercMatch",]<-apply(check.match[e,t,,]==0,1,sum,na.rm=TRUE)/Max.TS.length        # proportion of matches
N.Sim.PercMatch.ary[e,t,length(Overlap),]<-apply(N.check.match[e,t,,]==0,1,sum,na.rm=TRUE)/Max.TS.length
        # proportion of matches

##### FOR SYMBOLIC METRIC, CALCULATE SYNCHRONY OF ALL BASELINE TIME SERIES PAIRS
# Call FUNC.PHASE.CATEG to calculate synchrony for Symbolic metric
# Categorize matrix of time series data as increase, decrease, peak, or trough
a.phase.cat[e,t,,]<-Func.phase.categ(Big.ary[e,t,,,1],Tot.sub)
b.phase.cat[e,t,,]<-Func.phase.categ(Big.ary[e,t,,,2],Tot.sub)

# Categorize matrix of time series data as increase, decrease, peak, or trough
 N.a.phase.cat[e,t,,]<-Func.phase.categ(N.Big.ary[e,t,,,1],Tot.sub)
N.b.phase.cat[e,t,,]<-Func.phase.categ(N.Big.ary[e,t,,,2],Tot.sub)
```

```
            # Call FUNC.PHASE.ENTROPY to calculate entropy of phase-categorized time series
            HS.a[e,t,]<-apply(a.phase.cat[e,t,,],1,Func.phase.entropy)
            HS.b[e,t,]<-apply(b.phase.cat[e,t,,],1,Func.phase.entropy)

            # Call FUNC.JOINT.ENTROPY to calculate joint entropy of paired phase-categorized time series
            HSU.sim[e,t,]<-Func.joint.entropy(a.phase.cat[e,t,,],b.phase.cat[e,t,,],Tot.sub)
            Full.dat[e,t,"Symbolic",]<-HS.a[e,t,]+HS.b[e,t,]-HSU.sim[e,t,]

            # Repeat steps for 'no synchrony' pre-Cholesky time series
            N.HS.a[e,t,]<-apply(N.a.phase.cat[e,t,,],1,Func.phase.entropy)
            N.HS.b[e,t,]<-apply(N.b.phase.cat[e,t,,],1,Func.phase.entropy)
            N.HSU.sim[e,t,]<-Func.joint.entropy(N.a.phase.cat[e,t,,],N.b.phase.cat[e,t,,],Tot.sub)
            N.Sim.Symbolic.ary[e,t,length(Overlap),]<-N.HS.a[e,t,]+N.HS.b[e,t,]-N.HSU.sim[e,t,]
            }                                                                    # END FOR-T
        }                                                                        # END FOR-E

##### THESE ARE STEPS 2 & 3: CATEGORIZE BASELINE TIME SERIES BY SYNCHRONY ESTIMATE & SELECT SUBSET OF TIME SERIES
Temp.sync.categ<-ceiling(Full.dat[1,1,,]*10)          # Categorizes synchrony values 1-10, with 1=synchrony up to .1, etc.
Temp.sync.categ[Temp.sync.categ<=0]<-NA               # Time series pairs with synchrony <0 are excluded
Temp.sync.categ[Temp.sync.categ[4,]<=.5]<-NA          # For Percent Match, time series pairs with synchrony <0.5 are excluded
Tab.sync.categ<-apply(Temp.sync.categ,1,table)        # For each metric, tabulate the number of synchrony estimates in each category

# THE SECTION BELOW IS FOR PEARSON SYNCHRONY METRIC, BUT SHOULD BE REPEATED FOR EACH METRIC
# ----- START SECTION TO BE REPEATED FOR EACH METRIC -----
Index.temp<-rep(NA,Tot.pick)
Temp.vec<-sort(unique(Temp.sync.categ["Pearson",]))   # Vector of all synchrony categories for the metric
Pick.per.cat<-floor(Tot.pick/length(Temp.vec))        # Number of time series pairs to select per synchrony category
Low.cat<-which(Tab.sync.categ$Pearson<Pick.per.cat)   # Identifies which synchrony categories have <Pick.per.cat values
Low.cat.OK<-length(Low.cat)>0                 # TRUE if there are synchrony categories with insufficient time series pairs to choose from
```

```
# For categories with insufficient time series pairs to choose from, take all available pairs
if(Low.cat.OK)
        {
        Index.temp<-which(Temp.sync.categ["Pearson",]%in%Temp.vec[Low.cat])   # Identifies index number of estimates in the low categories
        Cat.left<-Temp.vec[-Low.cat]
        } else
        {
        Cat.left<-Temp.vec
        }
Cum.count<-length(na.omit(Index.temp))                  # Running count of the number of time series pairs selected so far (from low categories)
Diff.count<-Tot.pick-Cum.count                          # The number of time series pairs remaining to select
Full.vec<-1:Tot.sub
Pick.per.cat<-floor(Diff.count/length(Cat.left))        # Recalculate Pick.per.cat

for (j in Cat.left)                 # From each remaining synchrony category, randomly select the specified number of time series pairs
        {
        if (length(which(Temp.sync.categ["Pearson",]==j))<Pick.per.cat)
                {
                Index.temp<-c(na.omit(Index.temp),which(Temp.sync.categ["Pearson",]==j))
                } else
                {
                Index.temp<-c(na.omit(Index.temp),sample(which(Temp.sync.categ["Pearson",]==j),Pick.per.cat))
                }                       # END IF-THEN-ELSE
        }                               # END FOR-J
Diff.last<-Tot.pick-length(Index.temp)                                  # Determine how many values left to pick
skip.na<-which(is.na(Temp.sync.categ["Pearson",]))
Full.vec[skip.na]<-NA
Remain.vec<-Full.vec[-Index.temp]
Remain.vec.whole<-na.omit(Remain.vec)
if(Diff.last>0){Index.temp<-c(Index.temp,sample(Remain.vec.whole,Diff.last))}    # Randomly select remaining values from remaining indices
```

```
Select.col.mat["Pearson",]<-Index.temp
# ----- END SECTION TO BE REPEATED FOR EACH METRIC -----

# Assign baseline synchrony estimates to metrics for the longest overlap length
for (e in 1:length(Err.SD))                                  # For each error rate
        {
        for (t in 1:length(Treat))                           # For each data treatment
                {
                Sim.Pearson.ary[e,t,length(Overlap),]<-Full.dat[e,t,"Pearson",Select.col.mat["Pearson",]]    # ----- REPEAT FOR EACH METRIC -----
                }                                    # END FOR-T
        }                                            # END FOR-E

##### SO FAR, SYNCHRONY HAS ONLY BEEN CALCULATED FOR THE LONGEST TIME SERIES LENGTH. THE NEXT STEP IS TO CALCULATE
#    SYNCHRONY FOR SHORTER TIME SERIES LENGTHS, ONLY USING THE SUBSET OF TIME SERIES PAIRS SELECTED FOR EACH METRIC. FOR
#    SYNCHRONY OF SMOOTHED TIME SERIES,SMOOTHING HAS TO BE RE-APPLIED TO THE SHORTER TIME SERIES RATHER THAN TRUNCATING
#    A SUBSET OF THE SMOOTHED LONGEST TIME SERIES (RESULTS WOULD BE DIFFERENT). MUCH OF THIS CODE IS A REPEAT OF EARLIER
#    CODE FOR CALCULATING SYNCHRONY, WITH MINOR CHANGES.
#####

for (L in 1:(length(Overlap)-1))                                     # For each of the remaining overlap lengths to calculate
        {
        Smooth.ary<-array(NA,dim=c(length(Err.SD),Tot.sub,Overlap[L],2))        # Arrays to hold the shortened, smoothed time series data
        N.Smooth.ary<-array(NA,dim=c(length(Err.SD),Tot.sub,Overlap[L],2))

        ##### CALL FUNC.SMOOTH TO SMOOTH THE SHORTER TIME SERIES
        for (e in 1:length(Err.SD))
                {
                Smooth.ary[e,,,1]<-Func.smooth(Big.ary[e,,1:Overlap[L],1],1:Overlap[L],rep(Base.prec,Tot.sub),Smooth.def,Tot.sub)
                Smooth.ary[e,,,2]<-Func.smooth(Big.ary[e,,1:Overlap[L],2],1:Overlap[L],rep(Base.prec,Tot.sub),Smooth.def,Tot.sub)
```

```
        N.Smooth.ary[e,,,1]<-Func.smooth(N.Big.ary[e,,1:Overlap[L],1],1:Overlap[L],rep(Base.prec,Tot.sub),Smooth.def,Tot.sub)
        N.Smooth.ary[e,,,2]<-Func.smooth(N.Big.ary[e,,1:Overlap[L],2],1:Overlap[L],rep(Base.prec,Tot.sub),Smooth.def,Tot.sub)
        }                   # END FOR-E

##### INITIALIZE AND ASSIGN TEMPORARY ARRAYS FOR CALCULATING CORRELATION SYNCHRONY ON SHORTER TIME SERIES
# For unsmoothed data, can simply truncate the original data
sub.a.diff.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Overlap[L]))
sub.b.diff.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Overlap[L]))

sub.a.diff.ln[,1,,]<-a.diff.ln[,,1:Overlap[L]]
sub.b.diff.ln[,1,,]<-b.diff.ln[,,1:Overlap[L]]

N.sub.a.diff.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Overlap[L]))
N.sub.b.diff.ln<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Overlap[L]))

N.sub.a.diff.ln[,1,,1:Overlap[L]]<-N.a.diff.ln[,,1:Overlap[L]]
N.sub.b.diff.ln[,1,,1:Overlap[L]]<-N.b.diff.ln[,,1:Overlap[L]]

# For smoothed data, need to recalculate synchrony
sub.a.ln<-array(NA,dim=c(length(Err.SD),Tot.sub,Overlap[L]))
sub.b.ln<-array(NA,dim=c(length(Err.SD),Tot.sub,Overlap[L]))

N.sub.a.ln<-array(NA,dim=c(length(Err.SD),Tot.sub,Overlap[L]))
N.sub.b.ln<-array(NA,dim=c(length(Err.SD),Tot.sub,Overlap[L]))

# Find first differenced natural logs for correlation metrics
for (e in 1:length(Err.SD))
        {
        sub.a.ln[e,,]<-log(Smooth.ary[e,,,1])
        sub.b.ln[e,,]<-log(Smooth.ary[e,,,2])
```

```
        sub.a.diff.ln[e,2,,2:Overlap[L]]<-t(apply(sub.a.ln[e,,],1,diff))
        sub.b.diff.ln[e,2,,2:Overlap[L]]<-t(apply(sub.b.ln[e,,],1,diff))

        N.sub.a.ln[e,,]<-log(N.Smooth.ary[e,,,1])
        N.sub.b.ln[e,,]<-log(N.Smooth.ary[e,,,2])
        N.sub.a.diff.ln[e,2,,2:Overlap[L]]<-t(apply(N.sub.a.ln[e,,],1,diff))
        N.sub.b.diff.ln[e,2,,2:Overlap[L]]<-t(apply(N.sub.b.ln[e,,],1,diff))
        }                                   # END FOR-E

##### INITIALIZE AND ASSIGN TEMPORARY ARRAYS FOR CALCULATING PERCENT MATCH SYNCHRONY ON SHORTER TIME SERIES
# For unsmoothed data, can simply truncate the original data
sub.check.match<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Overlap[L]))
sub.check.match[,1,,]<-check.match[,,1:Overlap[L]]

N.sub.check.match<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Overlap[L]))
N.sub.check.match[,1,,]<-N.check.match[,,1:Overlap[L]]

# For smoothed data, need to recalculate synchrony
sub.a.PercMatch.cat<-array(NA,dim=c(length(Err.SD),Tot.sub,Overlap[L]))
sub.b.PercMatch.cat<-array(NA,dim=c(length(Err.SD),Tot.sub,Overlap[L]))

N.sub.a.PercMatch.cat<-array(NA,dim=c(length(Err.SD),Tot.sub,Overlap[L]))
N.sub.b.PercMatch.cat<-array(NA,dim=c(length(Err.SD),Tot.sub,Overlap[L]))

# Call FUNC.PERCMATCH.CATEG to find matching trends (between time series in a pair) for Percent Match metric
for (e in 1:length(Err.SD))
        {
        # Categorize matrix of time series data as increase, decrease, or no change
        sub.a.PercMatch.cat[e,,]<-Func.PercMatch.categ(Smooth.ary[e,,,1],Tot.sub)
        sub.b.PercMatch.cat[e,,]<-Func.PercMatch.categ(Smooth.ary[e,,,2],Tot.sub)
```

```
sub.check.match[e,2,,]<-sub.b.PercMatch.cat[e,,]-sub.a.PercMatch.cat[e,,]
        # Puts zero in every cell where series dxn matches

# Categorize matrix of time series data as increase, decrease, or no change
N.sub.a.PercMatch.cat[e,,]<-Func.PercMatch.categ(N.Smooth.ary[e,,,1],Tot.sub)
N.sub.b.PercMatch.cat[e,,]<-Func.PercMatch.categ(N.Smooth.ary[e,,,2],Tot.sub)
N.sub.check.match[e,2,,]<-N.sub.b.PercMatch.cat[e,,]-N.sub.a.PercMatch.cat[e,,]
        # Puts zero in every cell where series dxn matches
}                               # END FOR-E

##### INITIALIZE AND ASSIGN TEMPORARY ARRAYS FOR CALCULATING SYMBOLIC SYNCHRONY ON SHORTER TIME SERIES
# For unsmoothed data, can simply truncate the original data; for smoothed data, need to recalculate synchrony
sub.a.phase.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Overlap[L]))
sub.b.phase.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Overlap[L]))

sub.a.phase.cat[,1,,]<-a.phase.cat[,,1:Overlap[L]]
sub.a.phase.cat[,1,,Overlap[L]]<-NA                     # The last value should be NA
sub.b.phase.cat[,1,,]<-b.phase.cat[,,1:Overlap[L]]
sub.b.phase.cat[,1,,Overlap[L]]<-NA                     # The last value should be NA

sub.HS.a<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.pick))
sub.HS.b<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.pick))
sub.HSU.sim<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.pick))

N.sub.a.phase.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Overlap[L]))
N.sub.b.phase.cat<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub,Overlap[L]))

N.sub.a.phase.cat[,1,,]<-N.a.phase.cat[,,1:Overlap[L]]
N.sub.a.phase.cat[,1,,Overlap[L]]<-NA                   # The last value should be NA
N.sub.b.phase.cat[,1,,]<-N.b.phase.cat[,,1:Overlap[L]]
```

```
N.sub.b.phase.cat[,1,,Overlap[L]]<-NA                                   # The last value should be NA

N.sub.HS.a<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub))
N.sub.HS.b<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub))
N.sub.HSU.sim<-array(NA,dim=c(length(Err.SD),length(Treat),Tot.sub))

# Call FUNC.PHASE.CATEG to calculate synchrony for Symbolic metric
# Categorize matrix of time series data as increase, decrease, peak, or trough
for (e in 1:length(Err.SD))
        {
        sub.a.phase.cat[e,2,,]<-Func.phase.categ(Smooth.ary[e,,,1],Tot.sub)
        sub.b.phase.cat[e,2,,]<-Func.phase.categ(Smooth.ary[e,,,2],Tot.sub)

        N.sub.a.phase.cat[e,2,,]<-Func.phase.categ(N.Smooth.ary[e,,,1],Tot.sub)
        N.sub.b.phase.cat[e,2,,]<-Func.phase.categ(N.Smooth.ary[e,,,2],Tot.sub)
        }                                       # END FOR-E

##### CALCULATE SYNCHRONY
for (e in 1:length(Err.SD))                                             # For each error rate
        {
        for (t in 1:length(Treat))                                      # For each data treatment
                {
                # For Pearson, Spearman, and Kendall metrics
                for (i in 1:Tot.pick)
                        {
                        Sim.Pearson.ary[e,t,L,i]<-
cor(sub.a.diff.ln[e,t,Select.col.mat["Pearson",i],],sub.b.diff.ln[e,t,Select.col.mat["Pearson",i],],
                                method="pearson",use="na.or.complete")
                        Sim.Spearman.ary[e,t,L,i]<-
cor(sub.a.diff.ln[e,t,Select.col.mat["Spearman",i],],sub.b.diff.ln[e,t,Select.col.mat["Spearman",i],],
```

```
                                        method="spearman",use="na.or.complete")
                                    Sim.Kendall.ary[e,t,L,i]<-
cor(sub.a.diff.ln[e,t,Select.col.mat["Kendall",i],],sub.b.diff.ln[e,t,Select.col.mat["Kendall",i],],
                                        method="kendall",use="na.or.complete")
                                    }                               # END FOR-I

                                for (i in 1:Tot.sub)
                                        {
                                        N.Sim.Pearson.ary[e,t,L,i]<-
cor(N.sub.a.diff.ln[e,t,i,],N.sub.b.diff.ln[e,t,i,],method="pearson",use="na.or.complete")
                                        N.Sim.Spearman.ary[e,t,L,i]<-
cor(N.sub.a.diff.ln[e,t,i,],N.sub.b.diff.ln[e,t,i,],method="spearman",use="na.or.complete")
                                        N.Sim.Kendall.ary[e,t,L,i]<-
cor(N.sub.a.diff.ln[e,t,i,],N.sub.b.diff.ln[e,t,i,],method="kendall",use="na.or.complete")
                                         }                               # END FOR-I

                                # For Percent Match metric
                                Sim.PercMatch.ary[e,t,L,]<-apply(sub.check.match[e,t,Select.col.mat["PercMatch",],]==0,1,sum,na.rm=TRUE)/Overlap[L]
                                N.Sim.PercMatch.ary[e,t,L,]<-apply(N.sub.check.match[e,t,,]==0,1,sum,na.rm=TRUE)/Overlap[L]

                                # For Symbolic metric
                                # Call FUNC.PHASE.ENTROPY to calculate entropy of phase-categorized time series
                                sub.HS.a[e,t,]<-apply(sub.a.phase.cat[e,t,Select.col.mat["Symbolic",],],1,Func.phase.entropy)
                                sub.HS.b[e,t,]<-apply(sub.b.phase.cat[e,t,Select.col.mat["Symbolic",],],1,Func.phase.entropy)

                                # Call FUNC.JOINT.ENTROPY to calculate joint entropy of paired phase-categorized time series
                                sub.HSU.sim[e,t,]<-Func.joint.entropy(sub.a.phase.cat[e,t,Select.col.mat["Symbolic",],],
                                sub.b.phase.cat[e,t,Select.col.mat["Symbolic",],],Tot.pick)
                                Sim.Symbolic.ary[e,t,L,]<-sub.HS.a[e,t,]+sub.HS.b[e,t,]-sub.HSU.sim[e,t,]
```

```
                        # Repeat for 'no synchrony' pre-Cholesky time series
                        N.sub.HS.a[e,t,]<-apply(N.sub.a.phase.cat[e,t,,],1,Func.phase.entropy)
                        N.sub.HS.b[e,t,]<-apply(N.sub.b.phase.cat[e,t,,],1,Func.phase.entropy)
                        N.sub.HSU.sim[e,t,]<-Func.joint.entropy(N.sub.a.phase.cat[e,t,,],N.sub.b.phase.cat[e,t,,],Tot.sub)
                        N.Sim.Symbolic.ary[e,t,L,]<-N.sub.HS.a[e,t,]+N.sub.HS.b[e,t,]-N.sub.HSU.sim[e,t,]
                        }                               # END FOR-T
                }                                       # END FOR-E
        }                                               # END FOR-L

##### CALCULATE DIFFERENCE BETWEEN SYNCHRONY AND BASELINE SYNCHRONY ESTIMATES
for (e in 1:length(Err.SD))
        {
        for (t in 1:length(Treat))
                {
                for (L in 1:length(Overlap))
                        {
                        D.Sim.Pearson.ary[e,t,L,]<-Sim.Pearson.ary[1,1,length(Overlap),]-Sim.Pearson.ary[e,t,L,]
                                #----- REPEAT FOR EACH METRIC -----
                        }                                       # END FOR-L
                }                                               # END FOR-T
        }                                                       # END FOR-E

##### INITIALIZE LIST FOR PLOTTING GRAPHS AND ALSO FOR FILLING STANDARD DEVIATION & BIAS ARRAYS  ----- REPEAT FOR EACH METRIC -----
Methods.list<-list()
Methods.list$Pearson<-list(ary=Sim.Pearson.ary,d.ary=D.Sim.Pearson.ary,n.ary=N.Sim.Pearson.ary,Meth.min=0,Meth.max=1,
        dmax=1,Type.min=-.3,Type.max=1,pow=Power.Pearson.ary,Pow.min=0,Pow.max=100,sig50=Sig50.Pearson.ary)

##### CALCULATE STANDARD DEVIATION, BIAS, AND 95% UPPER CI BOUND
for (e in 1:length(Err.SD))
        {
```

```
        for (t in 1:length(Treat))
                {
                for (m in 1:length(Meth))
                        {
                        for (L in 1:length(Overlap))
                                {
                                SD.1000UE.ary[e,t,m,L]<-sd(Methods.list[[m]]$d.ary[e,t,L,])
                                        # Calculate standard deviation of difference from baseline
                                Bias.1000UE.ary[e,t,m,L]<-mean(Methods.list[[m]]$d.ary[e,t,L,])
                                        # Calculate bias of difference from baseline

                                Sort.sync<-sort(Methods.list[[m]]$n.ary[e,t,ov,],na.last=NA)            # Sort the 'no synchrony' estimates
                                No.sync.summ[m,e,t,ov]<-Sort.sync[floor(.95*length(Sort.sync))]             # Identify the 95% upper CI bound
                                 }                              # END FOR-L
                        }                               # END FOR-M
                }                               # END FOR-T
         }                                      # END FOR-E

##### CALCULATE PEARSON'S CORRELATION BETWEEN SYNCHRONY METRICS
for (e in 1:length(Err.SD))
        {
        for (t in 1:length(Treat))
                {
                for (i in 1:length(Meth))
                        {
                        for (j in 1:length(Meth))
                                {
                                Met.Pearson[e,t,i,j]<-cor(Full.dat[e,t,i,],Full.dat[e,t,j,],method="pearson")
                                }                                               # END FOR-J
                        }                                               # END FOR-I
```

```
        }                                                   # END FOR-T
    }                                                       # END FOR-E

##### THE POWER ANALYSIS BELOW IS FOR PEARSON, SPEARMAN AND KENDALL SYNCHRONY METRICS, BUT SHOULD BE REPEATED FOR
#  PERCENT MATCH AND # SYMBOLIC METRICS REMEMBER TO CHANGE THE RANGE OF SYNCHRONY VALUES FOR PERCENT MATCH AND
#  SYMBOLIC METRICS
#####

# ----- START SECTION TO BE REPEATED FOR PERCENT MATCH AND SYMBOLIC METRICS -----
for (m in 1:3)
    {
    for (e in 1:length(Err.SD))
        {
        for (t in 1:length(Treat))
            {
            for (ov in 1:length(Overlap)-1)      # Omit the longest (1000-yr) time series length for power analysis
                {
                 for (x in 5:95)
                    # For all synchrony estimates (need to change this range for Percent Match and Symbolic metrics)
                     {
                    win.min<-x/100-.05
                    # Window size is 0.10 (e.g., r = 0 - 0.10) (need to change for Percent Match and Symbolic metrics)
                    win.max<-x/100+.05
                    # Identify which synchrony estimates fall within the window
                     pick.list<-which(Methods.list[[m]]$ary[1,1,7,]>=win.min &  Methods.list[[m]]$ary[1,1,7,]<win.max)
                    if(length(pick.list)>10)     # Only calculate if there are at least 10 values within the window range
                        {
                        # Calculate the proportion of estimates greater than the 95% CI for 'no synchrony'
                        Methods.list[[m]]$pow[e,t,ov,x]<-sum(Methods.list[[m]]$ary[e,t,ov,pick.list]>
                            No.sync.summ[m,e,t,ov])/length(pick.list)
```

```
                }                       # END IF-LENGTH
            }                           # END FOR-X
    # If there are any baseline synchrony values with at least 50% significant values in this category (of error-treatment-overlap)
            if(length(which(Methods.list[[m]]$pow[e,t,ov,]>.5,arr.ind=TRUE))>0)
                {
                # Calculate the minimum baseline synchrony value at which at least 50% of values are significant
                Methods.list[[m]]$sig50[e,t,ov]<-
min(which(Methods.list[[m]]$pow[e,t,ov,]>.5,arr.ind=TRUE),na.rm=TRUE)
                }                   # END IF-LENGTH
            }                       # END FOR-OV
        }                           # END FOR-T
    }                               # END FOR-E
}                                   # END FOR-M
# ----- END SECTION TO BE REPEATED FOR PERCENT MATCH AND SYMBOLIC METRICS -----

save.image(file=paste("ARMA1000_",substr(as.character(Sys.time()),1,10),".RData",sep=""))# save workspace--name includes date



#>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>> CODE FOR GRAPHING FUNCTIONS

#####################################################################
#
# FUNC.PLOT.SIMTYPE:  FUNCTION TO PLOT THE VARIOUS SIMULATION TYPES (DIFFERENT ERROR & TREATMENT LEVELS) FOR A BASE
# TIME SERIES
#
# Input
#       [TS.vec] List of Big.ary vector values
#
#####################################################################
```

```
Func.plot.simtype<-function(TS.vec)
        {
        Yr<-1:Max.TS.length
        for (i in 1:length(TS.vec))                             # For every time series
                {
                par(mfrow=c(6,1))                                       # Display six graphs per page
                par(mar=c(0,0,0,0))                             # Pack graphs tightly
                par(oma=c(2,2,0,0))                             # Create outer margin for labels for (t in 1:4)
                for (e in 1:length(Err.SD))                     # For each error rate
                        {
                        for(t in 1:length(Treat))           # For each data treatment
                                {
                                lo<-min(Big.ary[,,TS.vec[i],,1],na.rm=TRUE)
                                hi<-max(Big.ary[,,TS.vec[i],,1],na.rm=TRUE)
                                plot(Big.ary[e,t,TS.vec[i],,1]~Yr,lwd=2,type='l',xaxt="n",yaxt="n",ylim=c(lo,hi))
                                }                                       # END FOR-T
                        title(xlab="Year",ylab="Density",cex.lab=1.2,outer=T)
                        }                                               # END FOR-E
                }                                                       # END FOR-I
        }                                                               # END FUNC.PLOT.SIMTYPE


#########################################################################
#
# FUNC.PLOT.OVERLAP.ACTUAL:  FUNCTION TO PLOT RELATIONSHIP BTWN SYNCHRONY ESTIMATE & 1000-YR UNERRORED
# SYNCHRONY ESTIMATE
#
# Input
#       [typ]           Treat:  U, Sm
```

```
#       [err]           Error:  0,.25,.3
#       [m]             Metric: Pearson, Spearman, Kendall, PercMatch, Symbolic
#
#########################################################################

Func.plot.overlap.actual<-function(typ,err,m)
        {
        t<-which(Treat==typ)
        e<-which(Err.SD==err)
        x.lo<-Methods.list[[m]]$Meth.min
        x.hi<-Methods.list[[m]]$Meth.max
        y.lo<-.1*floor(min(Methods.list[[m]]$ary[e,t,,],na.rm=T)*10)
        y.hi<-.1*ceiling(max(Methods.list[[m]]$ary[e,t,,],na.rm=T)*10)

        par(mfrow=c(2,3))

        for (ov in 1:(length(Overlap)-1))
                {
                plot(Methods.list[[m]]$ary[e,t,ov,]~Methods.list[[m]]$ary[1,1,length(Overlap),],xlab='Baseline 1000-year Synchrony',
                xlim=c(x.lo,x.hi),ylab='Synchrony Estimate',main=paste("Overlap = ",Overlap[ov]," years"),type='p',pch=1,
                        ylim=c(y.lo,y.hi),cex.lab=1.2,cex.axis=1.1)
                segments(x.lo,x.lo,x.hi,x.hi,col="red",lwd=2)
                abline(h=No.sync.summ[which(Meth==m),e,t,ov],col="blue",lwd=2.5,lty="dashed")                   # 95% upper CI line
                }                               # END FOR-OV
        }                                       # END FUNC.PLOT.OVERLAP.ACTUAL



#########################################################################
#
```

```
# FUNC.PLOT.OVERLAP.DIFF.B:  FUNCTION TO PLOT THE BOXPLOT RELATIONSHIP BTWN TIME SERIES OVERLAP &
# SYNCHRONY DIFFERENCE FROM BASELINE
#
# Input
#       [typ]           Treat:  U, Sm
#       [err]           Error:  0,.25,.30
#
#########################################################################

Func.plot.overlap.diff.b<-function(typ,err)
        {
        t<-which(Treat==typ)
        e<-which(Err.SD==err)

        ov<-1:(length(Overlap)-1)
        par(mfrow=c(2,3))
        for(m in 1:5)
                {
                boxplot(Methods.list[[m]]$d.ary[e,t,ov,]~ov,xlab='TS Overlap length',ylab='Difference from Baseline 1000-yr Synchrony',
                        ylim=c(-1*Methods.list[[m]]$dmax,Methods.list[[m]]$dmax),axes=FALSE,main=paste(Meth[m],"
                        Synchrony"),cex.lab=1.2,cex.axis=1.1)
                axis(1,at=1:length(Overlap),labels=Overlap,cex.lab=1.2,cex.axis=1.1)
                axis(2,seq(-1*Methods.list[[m]]$dmax,Methods.list[[m]]$dmax,.2),cex.axis=1.1)
                abline(h=0,col="red",lwd=2)
                }                                                       # END FOR-M
        }                                                               # END FUNC.PLOT.OVERLAP.DIFF.B
```

```
#########################################################################
#
# FUNC.PLOT.METRICS:  FUNCTION TO PLOT THE RELATIONSHIP BTWN 5 METRICS' SYNCHRONY ESTIMATES, for 1000-yr TIME SERIES ONLY
#
# Input
#         [err]               Error:  0,.25,.3
#
#########################################################################

Func.plot.metrics<-function(err)
        {
        e<-which(Err.SD==err)

        par(mfrow=c(5,5))
        par(mar=c(0,0,0,0))
        par(oma=c(3,10,3,10))

        for(ro in 1:5)     # For its row, the metric is the y-axis
                {
                real.y.lo<-Methods.list[[ro]]$Meth.min
                real.y.hi<-Methods.list[[ro]]$Meth.max
                # Min & max are determined across both smoothed & unsmoothed types. Min and max may extend beyond normal range.
                y.lo<-min(real.y.lo,.1*floor(min(Methods.list[[ro]]$ary[e,,length(Overlap),],na.rm=T)*10))
                y.hi<-max(real.y.hi,.1*ceiling(max(Methods.list[[ro]]$ary[e,,length(Overlap),],na.rm=T)*10))

                for(co in 1:5)     # For its column, the metric is the x-axis
                        {
                        real.x.lo<-Methods.list[[co]]$Meth.min
                        real.x.hi<-Methods.list[[co]]$Meth.max
                        x.lo<-min(real.x.lo,.1*floor(min(Methods.list[[co]]$ary[e,,length(Overlap),],na.rm=T)*10))
```

```
                x.hi<-max(real.x.hi,.1*ceiling(max(Methods.list[[co]]$ary[e,,length(Overlap),],na.rm=T)*10))

                if(ro<co){t<-1} else {t<-2}                     # Upper triangle will be unsmoothed, lower is smoothed

                plot(Full.dat[e,t,ro,]~Full.dat[e,t,co,],type=ifelse(ro==co, "n", "p"),xlim=c(x.lo,x.hi),ylim=c(y.lo,y.hi),axes=FALSE,pch=1)

                box(lty="solid")        # Make a solid box around each plot
                if(ro==co) text(mean(c(x.lo,x.hi)),mean(c(y.lo,y.hi)),Meth[ro],cex=2.5)

                if(ro==5&&even(co))
                        {axis(1,at=seq(x.lo,x.hi,.2),labels=seq(x.lo,x.hi,.2),cex.lab=1.2,cex.axis=1.1,las=3,xpd=TRUE)}
                                #x axis on bottom for type even
                if(ro==1&&odd(co))
                        {axis(3,at=seq(x.lo,x.hi,.2),labels=seq(x.lo,x.hi,.2),cex.lab=1.2,cex.axis=1.1,las=3,xpd=TRUE)}
                                #x axis on top for type odd

                if(co==1&&even(ro))
                        {axis(2,at=seq(y.lo,y.hi,.2),labels=seq(y.lo,y.hi,.2),cex.lab=1.2,cex.axis=1.1,las=1,xpd=TRUE)}
                                #y axis on left for type even
                if(co==5&&odd(ro))
                        {axis(4,at=seq(y.lo,y.hi,.2),labels=seq(y.lo,y.hi,.2),cex.lab=1.2,cex.axis=1.1,las=1,xpd=TRUE)}
                                #y axis on right for type odd
                if(ro!=co) segments(real.x.lo,real.y.lo,real.x.hi,real.y.hi,col="red",lwd=2)
                }                       # END FOR-CO
        }                               # END FOR-RO
}                                       # END FUNC.PLOT.METRICS


#########################################################################
#
```

```
# FUNC.PLOT.TYPE:  FUNCTION TO PLOT THE RELATIONSHIP BETWEEN SYNCHRONY ESTIMATE AND BASELINE SYNCHRONY, IN ONE
# LARGE MATRIX
#
# Input
#         [ov.num]                    Overlap 1:6
#
#########################################################################

Func.plot.type<-function(ov.num)
        {
        par(mfrow=c(5,6))
        par(mar=c(0,0,0,0))
        par(oma=c(3,10,3,10))

        for(m in 1:5)   # For each metric (one metric per row)
                {
                x.lo<-Methods.list[[m]]$Meth.min
                x.hi<-Methods.list[[m]]$Meth.max
                y.lo<-Methods.list[[m]]$Type.min
                y.hi<-Methods.list[[m]]$Type.max
                for (e in 1:length(Err.SD))                                     # For each error type
                        {
                        for(t in 1:length(Treat))                               # For each data treatment
                                {

        plot(Methods.list[[m]]$ary[e,t,ov.num,]~Methods.list[[m]]$ary[1,1,length(Overlap),],type='p',pch=1,axes=FALSE,xlim=c(x.lo,x.hi),
                                        ylim=c(y.lo,y.hi))
                                segments(x.lo,x.lo,x.hi,x.hi,col='red',lwd=2)
                                abline(h=No.sync.summ[m,e,t,ov.num],col="blue",lwd=2.5,lty="dashed")          # 95% upper CI line
```

```
                box(lty="solid")                                    # Make a solid box around each plot

                # Last column for even-numbered metrics should have axes; odd, metric names
                if(e==length(Err.SD)&& t==length(Treat)&& even(m))

                        {axis(4,at=seq(y.lo,y.hi,.2),labels=seq(y.lo,y.hi,.2),cex.lab=1.2,cex.axis=1.1,las=1,xpd=TRUE)}
                if(e==length(Err.SD)&& t==length(Treat)&& odd(m))
                        {axis(4,yaxt="n",title(ylab=Meth[m]),cex.lab=1.2)}

                if(e==1 && t==1 && even(m))     # First column for even-numbered metrics should have names; odd, axes

                        {axis(2,yaxt="n",ylab=Meth[m],cex.lab=1.2)}
                if(e==1 && t==1 && odd(m))
                        {axis(2,at=seq(y.lo,y.hi,.2),labels=seq(y.lo,y.hi,.2),cex.lab=1.2,cex.axis=1.1,las=1,xpd=TRUE)}
                }               # END FOR-T
        }                       # END FOR-E
    }                           # END FOR-M
}                               # END FUNC.PLOT.TYPE


#########################################################################
#
# FUNC.PLOT.POWER:  FUNCTION TO PLOT %SIGNIFICANT (AT ALPHA=0.05) SYNCHRONY ESTIMATES AGAINST BASELINE SYNCHRONY
#
# Input
#       [ov.num]                Overlap 1:6
#########################################################################

Func.plot.power<-function(ov.num)
        {
```

```
par(mfrow=c(5,6))
par(mar=c(0,0,0,0))
par(oma=c(3,10,3,10))

for(m in 1:5)   # For each metric (one metric per row)
        {
        x.lo<-Methods.list[[m]]$Pow.min
        x.hi<-Methods.list[[m]]$Pow.max
        y.lo<-0
        y.hi<-1

        for (e in 1:length(Err.SD))                                     # For each error type
                {
                for(t in 1:length(Treat))                               # For each data treatment
                        {
                        plot(Methods.list[[m]]$pow[e,t,ov.num,x.lo:x.hi],type='l',pch=1,axes=FALSE,ylim=c(y.lo,y.hi))
                        abline(h=.5,col="red")
                        box(lty="solid")                                # Make a solid box around each plot

                        if(e==length(Err.SD)&& t==length(Treat)&& even(m))# Last column for even-numbered metrics should have axes

                                {axis(4,at=seq(y.lo,y.hi,.2),labels=seq(y.lo,y.hi,.2),cex.lab=1.2,cex.axis=1.1,las=1,xpd=TRUE)}

                        if(e==1 && t==1 && odd(m))      # First column for odd-numbered metrics should have axes
                                {axis(2,at=seq(y.lo,y.hi,.2),labels=seq(y.lo,y.hi,.2),cex.lab=1.2,cex.axis=1.1,las=1,xpd=TRUE)}
                        }               # END FOR-T
                }                       # END FOR-E
         }                              # END FOR-M
}                                       # END FUNC.PLOT.POWER
```

CHAPTER 4

REVISITING THE LEGENDARY SYNCHRONY OF SNOWSHOE HARES

ABSTRACT

In recent decades, alarming collapses in the dynamics of several historically cyclic and synchronous species have underscored the profound effects these phenomena can have on trophic webs and ecosystem function. Identifying the mechanisms that synchronize population cycles could help us understand the role of anthropogenic disturbances, such as habitat fragmentation and climate change, in these large-scale processes. In the first range-wide analysis of snowshoe hare *(Lepus americanus)* dynamics, I used partial Mantel tests, modified correlograms, and shifting window analyses to quantify patterns and test hypothesized mechanisms of synchrony. Confirming long-held but previously untested assumptions, I found northern hare populations (boreal forests of Alaska and Canada) were significantly synchronized at distances up to several thousand kilometers, while southern populations (montane forests in the contiguous U.S.) were not significantly synchronized at any of the distance classes evaluated. I found that historical patterns of synchrony still persist for snowshoe hares, in contrast to reports for some other synchronous species. Hare synchrony patterns clustered into groups defined according to genetic criteria—but not ecoregions or climatic regions—suggesting an important role for dispersal in synchronizing northern snowshoe hare populations.

INTRODUCTION

Since its origins in the early 1900's, the discipline of population ecology has been concerned with answering the fundamental question: "How and why do population numbers change in time and space?" (Royama 1992; Kingsland 1995). At the core of this effort has been a quest to unravel the puzzle of population cycles and the related phenomenon of synchrony, the simultaneous rise and fall of populations over large spatial scales, sometimes across millions of square kilometers (Turchin 2003; Liebhold et al. 2004).

Population cycles and synchronous dynamics, famously exemplified by snowshoe hares (*Lepus americanus*) and Canada lynx (*Lynx canadensis*), are widespread phenomena reported for many taxa (reviewed in Liebhold et al. 2004). Understanding mechanisms of cycles and synchrony in ecological systems is not only relevant to conservation of threatened and endangered species, but also has economically important applications in the management of forest insect outbreaks (Peltonen et al. 2002) and regional fisheries (Fréon et al. 2003). For example, studies have shown that synchrony can affect long-term species persistence: climate-mediated synchrony can increase the chance of subpopulations simultaneously declining to extinction during a period of harsh weather, whereas dispersal-mediated synchrony can enhance population persistence via the regular movement of individuals from higher to lower density populations (Allen et al. 1993; Heino et al. 1997; Palmqvist and Lundberg 1998; Lloyd and May 1999; Matter 2001; Johst and Drechsler 2003; Roy et al. 2005).

Ecologically, population cycles and synchrony can also have profound effects on trophic dynamics and the functioning of ecosystems, as has been demonstrated with lemmings (*Dicrostonyx* and *Lemus* spp.) and snowshoe hares in boreal forests (Keith 1963; Bulmer 1974; Ims and Fuglei 2005; Krebs 2011), voles in Fennoscandia (*Microtus* and *Clethrionomys* spp.; Hanski et al. 2001; Hanski et al. 2006), various oaks (*Quercus* spp.; Koenig and Knops 2001; Haynes et al. 2009), and gypsy moths *(Lymantria dispar)* and other forest insects (Klemola et al. 2002; Stone 2004).

Over the past two decades, studies have reported alarming collapses of population cycles or synchrony in several species of voles, grouse, and forest insects (Bierman et al. 2006; reviewed in Ims et al. 2008; Kausrud et al. 2008). The significance and persistence of these recent shifts in population dynamics have yet to be determined. The restoration of 'lost' cycles and synchronous dynamics has already been reported for some species (Brommer et al. 2010). Nevertheless, the possible regime changes reported in the dynamics of historically cyclic and synchronous species, and concerns about potential loss of ecosystem functions (e.g., pulsed flows of resources and disturbances), raise urgencies for increased research attention on mechanisms driving these ecological processes. Identifying mechanisms would help us understand the potential role of anthropogenic disturbances, such as habitat fragmentation and climate change, in dampening cyclic and synchronous population dynamics.

Cyclic populations have a greater tendency than non-cyclic populations to synchronize over substantial distances, through a process known as "phase locking"

(Jansen 1999). We have a reasonably good understanding of the delayed density dependent processes that generate population cycles for many species (reviewed in Turchin 2003). However, the factors generating cycles are typically distinct from the mechanisms synchronizing these cycles across a species' range (Turchin and Hall 2003). Researchers have identified three primary mechanisms for synchrony: 1) dispersal from higher to lower density populations; 2) trophic interactions with predators or other species that are themselves synchronized or highly mobile; and 3) strong weather impacts on population growth combined with climatic similarities over broad geographic regions. Any combination of these factors can act in tandem (Liebhold et al. 2004a). Geographic patterns, such as the spatial scale of synchrony or correspondence between synchrony patterns and ecological gradients, can give us clues regarding the major mechanisms operating on a particular species (Koenig 1999).

For example, cyclic dynamics and spatial synchrony may break down at lower latitudes, a pattern suggested for snowshoe hares (Smith 1983; Hodges 2000b; Murray 2000), voles (Bjornstad et al. 1995; Saitoh et al. 1998; Tkadlec and Stenseth 2001), muskrats (*Ondatra zibethicus;* Erb et al. 2000), autumnal moths (*Epirrita autumnata;* Klemola et al. 2002), and grouse (Cattadori and Hudson 1999). For many of these boreal and northern temperate species, dampened cycles and synchrony appear to be driven by source-sink dynamics in the naturally heterogeneous landscapes prevalent in their southern range, whereby movement to and mortality in open-canopy stands is capable

of depressing landscape-level population growth (Howell 1923; Bjornstad et al. 1995; Griffin and Mills 2009).

The snowshoe hare has historically been recognized as a key driver of boreal community dynamics in North America (Keith 1963; Bulmer 1974; Keith 1983; Krebs et al. 2001; Ruesink et al. 2002). The large-scale synchrony of its high-amplitude population cycles has rippling impacts at multiple trophic levels and across almost 6 million square kilometers of Canada's and Alaska's boreal forests. While the mechanisms driving hare cycles have been intensively studied (e.g., MacLulich 1937; Keith and Windberg 1978; Hodges et al. 1999; Krebs et al. 2001; Sinclair et al. 2003), we have limited understanding of the processes synchronizing these cycles across large areas. Likewise, for hares and lynx, a range-wide north-south break in synchrony has not been empirically tested.

For snowshoe hares, as with many other species, the long-term, geographically extensive time series data required for synchrony analysis are difficult to come by (Bulmer 1974; Koenig 1999). The only quantitative analyses to date of snowshoe hare synchrony over a large part of the species range are based on 17 years of questionnaire data collected across Canada and a few northern U.S. states during the 1930's and 1940's, an effort known as the Canadian Snowshoe Rabbit Enquiry (e.g., Chitty 1948). These data were analyzed in two separate studies to reveal large-scale synchrony across much of Canada, with population trends in eastern and western Canada lagging those of central Canada by approximately three years (Smith 1983; Ranta et al. 1997a). A

traveling wave of synchrony has similarly been reported for Canada lynx, based on fur return data for the past two centuries (Elton and Nicholson 1942; Smith and Davis 1981; Ranta et al. 1997b; Haydon and Greenwood 2000).

It has been a decade since these attempts to evaluate southern snowshoe hare synchrony patterns, and reported patterns of northern hare synchrony are based on data collected 70 years ago. Over the next century, North America's boreal forests are predicted to undergo major fragmentation and habitat shifts due to global warming and human activities (IPCC 2001; IPCC 2007). Given recent apparent collapses in cycles and synchrony for other species and indications that natural habitat fragmentation may depress large-scale dynamics, how have patterns of hare synchrony changed in the past century, and what changes can we predict for the future?

In the first range-wide analysis of snowshoe hare dynamics, I quantified synchrony patterns and tested hypothesized mechanisms of hare synchrony. I compiled all available snowshoe hare time series from throughout the species distribution to answer three questions relevant to understanding current patterns and likely mechanisms of hare synchrony, necessary precursors to predicting future dynamics: 1) Do synchrony patterns differ between northern and southern hare populations?; 2) Have synchrony patterns changed in recent decades?; and 3) Do range-wide synchrony patterns correspond with ecoregions, climatic regions, or genetic groups?

To address these questions, I calculated synchrony for all time series pairs with at least 12 years of overlapping data. I combined partial Mantel tests and modified

correlograms (Koenig and Knops 1998) to quantify and contrast the geographic scale of synchrony for northern and southern hare populations. Finally, I tested the correspondence between synchrony patterns and three grouping factors (ecoregions, climatic regions, or genetic groups) in partial Mantel tests to evaluate the importance of two external abiotic factors (habitat and weather) and one internal biotic factor (hare dispersal, as measured by gene flow) in synchronizing hare populations.

METHODS

Compiling Time Series

Over the past several decades, many disparate efforts by Canadian provincial/territorial and U.S. federal and state wildlife agencies and independent researchers have accumulated snowshoe hare time series exceeding the 17-year span of the Rabbit Enquiry data. My starting points for data compilation were the collections of snowshoe hare time series reported in Keith (1963) and Hodges (2000a, 2000b). I updated time series for which more recent data were available and sought additional snowshoe hare time series from:

1) Authors of snowshoe hare literature cited in review papers by Hodges (2000a, 2000b), Murray (2003), Ellsworth and Reynolds (2006), and Ferron and St-Laurent (2007)

2) Authors of snowshoe hare literature published since 1995, identified by a Web of Knowledge database search using keywords “Lepus americanus” or “snowshoe hare*”
3) Canadian National Parks
4) Trappers Associations for Canadian provinces and territories
5) Canadian provincial/territorial and U.S. federal and state wildlife agencies
6) Large U.S. National Forests within the range of snowshoe hares

Data ‘Cleaning’

For analyses, I excluded all time series spanning less than 14 years, to ensure at least one full cycle was represented in the data. I did not use time series collected at recent post-clearcut or post-fire sites, to eliminate biases associated with these habitat treatments, as hare numbers are known to decline in the years immediately following these disturbances (Griffin and Mills 2009; Hodges et al. 2009). I also excluded time series that did not extend into the second half of the twentieth century, because my primary focus was on current hare synchrony patterns.

The time series data varied in survey method (harvest, live-trap, sightings, pellets, winter tracks), geographic extent (ranging from 14-ha study sites to entire Canadian provinces/ territories), and temporal overlap. The challenge of this study was to maximize the number and spatial distribution of time series incorporated in a single range-wide analysis of snowshoe hare synchrony, while minimizing potential biases due to combining different types of data. I approached this challenge in four ways.

First, I used Kendall rank correlation (Zar 1999) to identify synchrony patterns in the data, because a previous simulation study (Chapter 3) showed this metric was more robust to short time series and sampling error, and less prone to Type I error than other synchrony metrics evaluated. Because the Kendall metric is a nonparametric test of synchrony, I also expected this metric to be less affected by biases associated with survey method (e.g., potential over-estimation by pellet indices when population sizes are low; Mills et al. 2005) compared to parametric alternatives such as Pearson correlation. Although the Kendall metric had not been previously applied in snowshoe hare synchrony analyses, I demonstrate this metric performed comparably with other methods in identifying the traveling wave pattern of synchrony earlier reported for Canada lynx and snowshoe hares (Appendix 4.1).

Second, I combined time series for some small-scale study sites (<30 ha), so time series would more likely reflect annual changes in population abundance instead of seasonal movements or local source-sink dynamics (Griffin and Mills 2009). Data were combined only when they were collected using the same method and when sites were less than 30 km apart (Appendix 4.2). For the data I compiled, this scale tended to separate population clusters occurring in different habitat types (e.g., mature spruce-fir versus young lodgepole habitats). Within a set of time series, data were sometimes missing for one or more years for a subset of time series. In addition, the mean and variance of population estimates often differed among time series because even at small spatial scales, study sites can differ greatly in habitat quality for snowshoe hares.

Therefore, to ensure the combined abundance estimates for these data would not be biased by which time series were missing data, I scaled time series to a mean of zero and unit variance prior to averaging the data across a set of time series.

Third, I standardized the recorded years of data collection, so data collected at similar times of the year were consistently assigned to the same year for Lag 0 synchrony analyses. For example, most winter harvest data were collected from autumn/winter of one year to early spring of the subsequent year. I consistently recorded winter harvest numbers as estimates for the autumn/winter season (e.g., harvest data collected October 2007 – March 2008 would be assigned to 2007). Data collected between January and April of a year using 'snapshot' methods, such as live-trapping, winter track counts, or sightings, were also recorded as estimates for the previous year. The justification was that these short-term estimates and indices of population sizes overlapped winter harvest seasons recorded for the previous year. 12 out of 49 time series (25%) were pellet counts that averaged pellets deposited over the course of a year, or snapshot methods collected during the summer. For these time series, it was difficult to determine whether estimates should correspond better with winter harvests recorded for that year or for the previous year. Therefore, I analyzed them in two ways. First, I simply used the calendar year, which means that summer counts were included with the harvest data for that same year's fall/winter season. Second, I used the higher of two synchrony estimates (data recorded as current year

versus previous year) for these twelve time series. Hereafter, I refer to the second method as Adjusted Lag 0 synchrony.

Finally, to evaluate the potential validity of combining different data in a single range-wide analysis, I asked if synchrony estimates were correlated between different survey methods (e.g., harvest compared to harvest per hunter, or pellets compared to harvest) for geographically overlapping areas. For this analysis, I calculated Pearson's correlation coefficient between synchrony estimates for each pair of methods. For example, when both harvest and sightings time series were available for the same U.S. state, I calculated synchrony between the harvest time series and the 48 other time series in my analysis, calculated synchrony between the sightings time series and the 48 other time series in my analysis, then determined the correlation between these two sets of synchrony estimates.

Synchrony Analyses

All analyses were conducted in Program R (http://cran.r-project.org/; Appendix 4.7). The first step in identifying synchrony patterns and influencing factors was to estimate Lag 0 and Adjusted Lag 0 Kendall synchrony for every pairwise comparison of snowshoe hare time series with at least 12 years of overlapping data. Although the minimum time series length included in this analysis was 14 years, the number of years a pair of time series actually overlapped depended on which years each time series spanned. For example, a time series spanning 1960-1974 and a time series spanning 1970-1984 overlap in only five years, 1970-1974.

After I calculated pairwise synchrony estimates, I used a partial Mantel test to determine if synchrony patterns differ between northern and southern hare populations, using the 49th parallel to distinguish these groups (similar to lynx; Parker 2001). The partial Mantel test determines correspondence between two pairwise variables while controlling for potential confounding effects of a third pairwise variable (typically, the geographic distance between populations). The matrices used in my partial Mantel analysis were: 1) a matrix of pairwise Kendall synchrony estimates, 2) a categorical matrix coded as "0" if both populations occurred south of the 49th parallel and "1" if both populations occurred north of the 49th parallel, and 3) a matrix of geographic distance between each pair of populations. Significance of partial Mantel test statistics was evaluated with 1000 resamples.

To gain additional insight on how synchrony patterns differ between northern and southern snowshoe hares, I compared the spatial scale of synchrony (i.e., the maximum geographic distance between two time series exhibiting significant synchrony) for these two regions. For this analysis I used the modified correlogram (Koenig and Knops 1998). Unlike the Mantel correlogram, which tests synchrony within each geographic distance category against the overall average synchrony of the entire dataset, the modified correlogram tests whether the synchrony within each geographic distance category is significantly different from zero, thus identifying the spatial scale of synchrony. I conducted separate modified correlogram analyses for northern and southern hare time series and for a range (3–5) of geographic distance categories.

Significance of synchrony within each distance category was evaluated with (the smaller of) 100 resamples or the maximum number of resamples possible given the number of population pairs within a distance category.

To determine if hare synchrony patterns have changed in recent decades, I conducted separate analyses for northern and southern hare populations. Separate analyses were necessary because different data were available for northern and southern populations. For northern hare populations, I determined if the modified correlogram for northern time series I had compiled (primarily spanning the past three decades) exhibited a U-shaped relationship between synchrony and geographic distance qualitatively similar to that demonstrated for the 1931–1947 Canadian Snowshoe Rabbit Enquiry data (Ranta et al. 1997a). This pattern, in which synchrony first declines and then increases with geographic distance, is indicative of a traveling wave of synchrony (Ranta et al. 1997a; Ranta et al. 1997b). Although the similarity of my results with Ranta et al.'s (1997) scatterplot could not be statistically tested, I expected a U-shaped synchrony pattern to be significantly synchronous at least within the smallest and largest geographic distance categories.

For southern hare populations (south of the 49th parallel), no large-scale historical assessment of snowshoe hare synchrony was available for comparison with current synchrony patterns. Therefore, I evaluated long-term trends in synchrony of southern hare populations by conducting a shifting window analysis on the five longest southern hare time series I had compiled. These five time series (NY1, MI3, MN2, PA1,

WI2) covered the Great Lakes and northeast U.S. region and spanned 37–69 years in length. Beginning with a window size of 15 years, for every pairwise comparison of the five time series I estimated synchrony for the first 15 years of data, shifted the analysis window by one year, estimated synchrony for the next 15 years of data, and repeated this procedure to the end of the time series. I plotted successive synchrony estimates against time, calculated the slope of the regression, and averaged this slope over all pairwise comparisons for each window size. To ensure conclusions were not influenced by window size, I repeated this analysis with window sizes of 25 and 35 years.

I used the partial Mantel test to evaluate potential mechanisms of hare synchrony. I asked if snowshoe hare populations were significantly more synchronized within versus between groups defined by five Level I Ecoregions (Eastern Temperate Forests, Marine West Coast Forests, Northern Forests, Northwestern Forested Mountains, and Taiga; CEC 1997), three climatic regions, or three genetic groups, while controlling for geographic distance. For this analysis, a time series pair was coded as "0" if both populations occurred in different (ecoregion, climatic, or genetic) groups and "1" if they occurred in the same group. The climatic regions (Pacific, Continental, and Atlantic) were defined by the spatial influences of the North Atlantic Oscillation (NAO), following Stenseth et al. (1999). To facilitate comparisons with Stenseth et al's (1999) finding that Canada lynx dynamics are similar within NAO climatic regions, I limited my analysis of climatic regions to snowshoe hare time series from the northern species' range, as Stenseth did with lynx. The genetic groups (Boreal, Pacific NW, and Rockies)

were defined in my earlier range-wide analysis (Chapter 2) of snowshoe hare genetic structure.

RESULTS

Time Series and Data 'Cleaning'

Wildlife agencies and individual researchers generously donated 309 snowshoe hare time series at least six years in duration. Approximately half of these time series consisted of pellet counts and nearly a quarter were hunter harvests (Fig. 4.1). Only 45% (139) of the time series met criteria for inclusion in analyses; the others were shorter than 14 years or were collected in post-clearcut or post-burn treatments. After combining time series for small-scale study sites (Appendix 4.2), a final set of 49 time series within the years 1923-2010 and ranging from 14 to 75 years in duration were analyzed in this study (Appendices 4.3 and 4.4).

The final set of time series covered a large portion of the snowshoe hare range, but with notable gaps of data for the Pacific Northwest, British Columbia, and southeastern U.S. states (Fig. 4.2). Few time series were available for central Canada. Most harvest time series occurred in the eastern half of the species range. These were among the longest time series used in this study. Therefore, there was a strong geographic bias in both time series length and survey method. Time series were also more spatially clustered in the southern and eastern species range than in the western range.

Comparisons of different survey methods for harvest data (harvest, harvest/hunter, harvest/hunter-day, and harvest/area) yielded good correlations, with a minimum Pearson's r = 0.76 across ten comparisons (Appendix 4.5). Few data were available for evaluating non-harvest survey methods. My comparisons involving non-harvest data were limited to southern population evaluations of harvest against non-harvest methods that were applied to (often much smaller) geographic subsets of the harvest area. Correlations between synchrony estimates based on harvest and non-harvest methods were poor, ranging from Pearson's r = 0.0 (for comparison of harvest and pellet-based synchrony estimates) to r = 0.43 (for comparison of harvest and track-based synchrony estimates).

Because my analysis comparing survey methods revealed poor correlation between harvest and non-harvest methods, I used the partial Mantel test to further examine methods biases associated with survey methods. In this test I coded a time series pair as "0" if data were collected using different methods, and "1" if they were collected using the same method. Synchrony estimates were not significantly higher for time series pairs with the same compared to different survey methods ($r=0.07$, $p = 0.13$).

Results based on Lag 0 and Adjusted Lag 0 synchrony estimates were qualitatively similar across all tests (not shown), so results are presented only for Lag 0 synchrony.

Synchrony Analyses

Northern hare populations exhibited higher synchrony than southern hare populations ($r = 0.39$, $p<0.001$). All modified correlogram results were qualitatively similar regardless of the number of geographic distance categories (3–5), so I present results for four distance categories, consistent with the four clusters of hare synchrony Ranta et al. (1997a) identified from the Canadian Snowshoe Rabbit Enquiry data. For northern hare populations, synchrony was significant for 3 of 4 geographic distance categories ($p < 0.01$; Fig. 4.3). Southern hare populations were not significantly synchronized with each other within any distance category ($p = 0.11–0.42$; Fig. 4.4). To investigate a potential bias due to survey methods, I re-analyzed southern hare data using harvest estimates only—results were still non-significant in all distance categories ($p = 0.11–0.46$). The distribution of geographic distances for southern versus northern populations were similar (Fig. 4.5), suggesting their different synchrony patterns were not due to greater separation of southern populations within distance categories.

The modified correlogram for northern hare populations exhibited a U-shaped relationship between synchrony and geographic distance (Fig. 4.3). Northern hare populations were significantly synchronized in all distance categories except for intermediate distances (2901–4350 km). In this study, the largest distance category (> 4350 km) comprised Alaska snowshoe hare populations paired with northern Quebec snowshoe hare populations (QC3 and QC7). Synchrony was significant or nearly significant for all pairwise comparisons between Alaska and the northern Quebec

populations, except for comparisons with combined pellet data from five study sites across Kenai National Wildlife Refuge (AK5; Table 4.1). In contrast, synchrony was not significant between Alaska and many southern Quebec populations (Table 4.1).

I did not find a significant long-term trend in synchrony for southern hare populations, based on a shifting window analysis of synchrony for the five longest southern hare time series available (e.g., Appendix 4.6). Averaged across all pairwise population comparisons for each window size, mean slopes when synchrony was plotted over time were < 0.01; mean standard deviations were < 0.02.

Synchrony was not higher for populations within versus between ecoregions ($r = 0\text{-}.05$, $p = 0.27$) or climatic regions ($r = 0.08$, $p = 0.25$), when partial Mantel analyses controlled for the confounding effects of geographic distance. Using the same analysis method, populations occurring in the same genetic group were more synchronized than populations in different genetic groups ($r = 0.17$, $p = 0.02$).

DISCUSSION

In the first range-wide, quantitative analysis of snowshoe hare synchrony, I confirmed the long-held, but previously untested, assumption that northern hare populations are more synchronized than southern hare populations. I found that historical patterns of synchrony still persist for snowshoe hares, in contrast to recent apparent collapses in the dynamics of some other synchronous species. Hare synchrony patterns clustered into groups defined according to genetic criteria, suggesting a likely

role for dispersal (of hares or their predators) in synchronizing snowshoe hare populations.

Northern vs. Southern Hare Dynamics

Based on the partial Mantel test, I found a significant difference in synchrony between northern and southern hare populations. This result was also supported by independent modified correlogram analyses, which identified a large spatial scale of synchrony in the northern hare range (populations separated by several thousand kilometers were significantly synchronized) and a relative lack of synchrony among southern hare populations. These findings are consistent with previous assumptions that snowshoe hares exhibit a latitudinal gradient in cycles and synchrony, similar to several other boreal and north temperate species (Bjornstad et al. 1995, Saitoh et al. 1998, Tkadlec and Stenseth 2001).

Temporal Patterns of Hare Synchrony

While the past few decades have seen dampening population dynamics for several historically cyclic and synchronous species (Bierman et al. 2006; reviewed in Ims et al. 2008; Kausrud et al. 2008), northern snowshoe hare populations appear to have retained historical patterns of synchrony. Time series data based primarily on the past three decades echo the spatial scale and U-shaped pattern of synchrony reported in analyses of northern hare time series from the 1930's and 1940's and Canada lynx data based on the past two centuries (Elton and Nicholson 1942; Smith and Davis 1981; Ranta et al. 1997b; Haydon and Greenwood 2000).

Likewise, the relatively low synchrony among southern hare populations appears to reflect long-term dynamics rather than a recent collapse of formerly synchronous dynamics. These results suggest the factors generating the current latitudinal gradient in hare synchrony have been operating since at least the mid-1900's. This finding does not preclude an anthropogenic influence on hare synchrony patterns. While habitat clearing and fragmentation have increased in recent decades, these and many other anthropogenic activities (e.g., fire suppression) have been influencing the quality of lynx and hare habitat in North America for much of the past century, especially in southern portions of the species' range (Poole 2003).

Mechanisms of Hare Synchrony

Hare synchrony patterns did not significantly correspond with ecoregions or climatic regions in this analysis. This finding contrasts a large number of studies, especially from the past decade, that have identified major climatic influences on species dynamics (e.g., Ottersen et al. 2001; Post and Forchhammer 2002; Stenseth et al. 2004; Grotan et al. 2005) and a primary role for global warming in collapsing cycles and synchrony (Ims et al. 2008; Kausrud et al. 2008; Gilg et al. 2009). Stenseth et al. (1999) had suggested Canada lynx population dynamics are more similar within than between the NAO climatic regions I analyzed in this study, and that lynx were less likely to disperse across climatic region boundaries due to unfamiliarity with different snow conditions in different regions. If lynx were synchronizing snowshoe populations directly or indirectly, I would have expected similar climate-influenced boundaries on snowshoe

hare synchrony as suggested for lynx. I did not find this pattern, which suggests some combination of: 1) the proposed climatic boundary for lynx is incorrect; 2) snowshoe hares are not synchronized by lynx; or 3) the snowshoe hare data were insufficient to detect a climate-influenced boundary on synchrony patterns.

Hare synchrony patterns significantly corresponded with genetic groups. In Chapter 2, I found a north-south pattern in snowshoe hare genetic structure, with gene flow higher at northern latitudes. The finding that synchrony and gene flow patterns are significantly associated for snowshoe hares suggests dispersal may play an important role in synchronizing hare dynamics. The observed traveling wave of synchrony among northern hare populations is also a common signature of dispersal-mediated synchrony (Haydon and Greenwood 2000; Bjornstad et al. 2002; Blasius et al. 1999). These patterns—a latitudinal gradient in synchrony and traveling wave synchrony in northern habitats—are equally consistent with a mechanism of hare dispersal and a mechanism of lynx (or other specialist predator) dispersal (Schwartz et al. 2002) combined with strong predator-prey interactions: more data are needed to determine which mechanisms are in fact operating.

Making Full Use of Disparate Data for Range-wide Analyses

This study tentatively suggests strong patterns of synchrony can be revealed by analyses combining data using different survey methods. Most previous studies of synchrony, including the most frequently cited classic studies, have been based on a limited number of large-scale, long-term time series data collected using standardized

survey methods across all study sites. For example, most synchrony studies of North American mammals have been based on Hudson's Bay Company Canada-wide fur return data (e.g., Elton 1942; Moran 1953; Erb 2000; Haydon 2001); many studies of synchrony in forest insect outbreaks used tree mortality patterns determined from aerial survey maps (e.g., Bjornstad et al. 2002; Aukema et al. 2008; Haynes et al. 2009); synchrony in microtine rodents and their predators in Fennoscandia have been based on trap data (e.g., Steen et al. 1996; Lambin et al. 1998); studies of disease epidemics use nationally collected medical data (Bjornstad 2000; Grenfell 2001); and many bird studies employ large-scale breeding bird survey or harvest data (e.g., Paradis et al. 2000; Bellamy et al. 2003; Kvasnes et al. 2010).

For these studies, it is possible that other time series sources might extend inferences, and for many other species, mixed-source time series are the only data available for studying large-scale population processes such as synchrony. My study takes a different approach to synchrony analysis—I bring together time series data from 30 different sources, representing six survey methods, and ranging from 14 to 75 years in duration.

My results suggest at least partial success in detecting synchrony using mixed-source data. Survey method was not a significant grouping factor in a partial Mantel test, and the large-scale synchrony patterns revealed in this study matched expectations based on historical patterns and expectations. However, I cannot discount the possibility that results were influenced by the combination of multiple data types. Most notably,

when I explicitly compared synchrony estimates from harvest versus non-harvest survey methods, I did not find a strong correlation in these estimates. The lack of correlation may be due to differences in geographic area covered (e.g., track data for the northern third of Wisconsin compared to harvest data for all of Wisconsin), temporal coverage (e.g., harvest estimates collected over a 5-month period versus pellet counts averaged over a year), or differences associated with the survey methods themselves. The data are insufficient for distinguishing among these possibilities. Because all the data for methods comparisons came from southern populations, it is likely the low correlations are due in part to real differences in dynamics for the different geographic areas covered (rather than differences inherent to the methods themselves)—because in the southern range, populations fluctuate relatively independently at small spatial scales.

Anthropogenic Impacts on Snowshoe Hare Synchrony

How might habitat fragmentation and climate change impact future snowshoe hare dynamics? Results suggest that synchrony in hares is greater where populations are more connected. Major fragmentation of North America's relatively homogeneous boreal forests could disrupt the large-scale, synchronized snowshoe hare cycles that have historically driven boreal community dynamics. Climate change has been invoked as the most likely cause of recent declines in cycles and synchrony of other species (Bierman et al. 2006; Ims et al. 2008; Kausrud et al. 2008). For snowshoe hares, there is no evidence that climate change has yet impacted or is likely to dampen hare dynamics in the near future.

LITERATURE CITED

Aukema, B. H., A. L. Carroll, et al. (2008). "Movement of outbreak populations of mountain pine beetle: influences of spatiotemporal patterns and climate." Ecography **31**(3): 348-358.

Bellamy, P. E., P. Rothery, et al. (2003). "Synchrony of woodland bird populations: the effect of landscape structure." Ecography **26**: 338-348.

Bierman, S. M., J. P. Fairbairn, et al. (2006). "Changes over time in the spatiotemporal dynamics of cyclic populations of field voles (Microtus agrestis L.)." American Naturalist **167**(4): 583-590.

Bjornstad, O. N. (2000). "Cycles and synchrony: two historical 'experiments' and one experience." Journal of Animal Ecology **69**(5): 869-873.

Bjornstad, O.N., W. Falck, et al. (1995). “Geographic gradient in small rodent density-fluctuations--a statistical modeling approach.” Proceedings of the Royal Society of London. Series B **262:** 127-133.

Bjornstad, O. N., R. A. Ims, et al. (1999). "Spatial population dynamics: analyzing patterns and processes of population synchrony." Trends in Ecology and Evolution **14**(11): 427-432.

Bjornstad, O. N., M. Peltonen, et al. (2002). "Waves of larch budmoth outbreaks in the European Alps." Science **298**: 1020-1023.

Blasius, B., A. Huppert, et al. (1999). "Complex dynamics and phase synchronization in spatially extended ecological systems." Nature 399(6734): 354-359.

Brommer, J. E., H. Pietiainen, et al. (2010). "The return of the vole cycle in southern Finland refutes the generality of the loss of cycles through 'climatic forcing'." Global Change Biology **16**(2): 577-586.

Bulmer, M. G. (1974). "A statistical analysis of the 10-year cycle in Canada." Journal of Animal Ecology **43**(3): 701-718.

Buonaccorsi, J. P., J. S. Elkinton, et al. (2001). "Measuring and testing for spatial synchrony." Ecology **82**(6): 1668-1679.

Cattadori, I.M. and Hudson, P.J. (1999). “Temporal dynamics of grouse at the southern edge of their distribution.” Ecography **22:** 373-374.

CEC (1997). “Ecological Regions of North America: Toward a Common Perspective.” Montreal, Commission for Environmental Cooperation of North America.

Chitty, D. (1996). Do Lemmings Commit Suicide? Beautiful Hypotheses and Ugly Facts, Oxford University Press, New York.

Chitty, H. (1948). "The snowshoe rabbit enquiry, 1943-46." Journal of Animal Ecology **17**(1): 39-44.

Ellsworth, E. and T. D. Reynolds (2006). Snowshoe hare (Lepus americanus): a technical conservation assessment. R. M. R. USDA Forest Service**:** 67.

Elton, C. and M. Nicholson (1942). "Fluctuations in numbers of the muskrat (Ondatra zibethica) in Canada." Journal of Animal Ecology **11**(1): 96-126.

Elton, C. and M. Nicholson (1942). "The ten year cycle in numbers of the lynx in Canada." Journal of Animal Ecology **11**(2): 215-244.

Erb, J., N.C. Stenseth, et al. (2000). Geographic variation in population cycles of Canadian muskrats (Ondatra zibethicus). Canadian Journal of Zoology **78**: 1009-1016.

Ferron, J. and M. H. St- Laurent (2007). Forest-fire regime: the missing link to understand snowshoe hare population fluctuations?, Springer-Verlag Berlin Heidelberg.

Gilg, O., B. Sittler, et al. (2009). "Climate change and cyclic predator-prey population dynamics in the high Arctic." Global Change Biology 15(11): 2634-2652.

Grenfell, B. T., O. N. Bjornstad, et al. (2001). "Travelling waves and spatial hierarchies in measles epidemics." Nature 414(6865): 716-723.

Grotan, V., B. Saether, et al. (2005). "Climate causes large-scale spatial synchrony in population fluctuations of a temperate herbivore." Ecology 86(6): 1472-1482.

Griffin, P. C. and L. S. Mills (2009). "Sinks without borders: snowshoe hare dynamics in a complex landscape." Oikos **118**(10): 1487-1498.

Hanski, I., L. Hansson, et al. (1991). "Specialist predators, generalist predators, and the microtine rodent cycle." Journal of Animal Ecology **60**(1): 353-367.

Hanski, I., H. Henttonen, et al. (2001). "Small-rodent dynamics and predation." Ecology **82**(6): 1505-1520.

Haydon, D. T. and P. E. Greenwood (2000). "Spatial coupling in cyclic population dynamics: models and data." Theoretical Population Biology **58**: 239-254.

Haydon, D. T., N. C. Stenseth, et al. (2001). "Phase coupling and synchrony in the spatiotemporal dynamics of muskrat and mink populations across Canada." Proceedings of the National Academy of Sciences of the United States of America **98**(23): 13149-13154.

Haynes, K. J., A. M. Liebhold, et al. (2009). "Spatial synchrony propagates through a forest food web via consumer-resource interactions." Ecology **90**(11): 2974-2983.

Hitch, A. T. and Leberg, P. L. (2007) “Breeding distributions of north American bird species moving north as a result of climate change.” Conservation Biology **21:** 534–539.

Hodges, K. E., L.S. Mills, et al. (2009) “Distribution and Abundance of Snowshoe Hares in Yellowstone National Park.” Journal of Mammalogy **90**(4) 870-878.

Hodges, K. E. (2000a). The Ecology of Snowshoe Hares in Northern Boreal Forests. Ecology and Conservation of Lynx in the United States. L. F. Ruggiero, K. B. Aubry, S. W. Buskirk et al. Boulder, CO, USA, University Press of Colorado**:** 117-161.

Hodges, K. E. (2000b). Ecology of Snowshoe Hares in Southern Boreal and Montane Forests. Ecology and Conservation of Lynx in the United States. L. F. Ruggiero, K. B. Aubry, S. W. Buskirk et al. Boulder, CO, USA, University Press of Colorado**:** 163-206.

Hodges, K. E., C. J. Krebs, et al. (1999). "Snowshoe hare demography during a cyclic population low." Journal of Animal Ecology **68**(3): 581-594.

Howell, A.B. (1923). “Periodic fluctuations in the numbers of small mammals.” Journal of Mammalogy **4:** 149-155.

Ims, R. A. and E. Fuglei (2005). "Trophic interaction cycles in tundra ecosystems and the impact of climate change." BioScience **55**(4): 311-322.

Ims, R. A., J. A. Henden, et al. (2008). "Collapsing population cycles." Trends in Ecology & Evolution **23**(2): 79-86.

IPCC (2001). Climate change 2001: the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. New York, USA, Cambridge University Press.

IPCC (2007). Climate Change 2007: Impacts, Adaptation, and Vulnerability. Cambridge, UK, Cambridge University Press.

Kausrud, K. L., A. Mysterud, et al. (2008). "Linking climate change to lemming cycles." Nature **456**(7218): 93-U3.

Keith, L. B. (1963). Wildlife's ten year cycle. Madison, WI, University of Wisconsin Press.

Keith, L. B., J. R. Cary, et al. (1984). "Demography and ecology of a declining snowshoe hare population." Wildlife Monographs **90**: 1-43.

Keith, L. B. and L. A. Windberg (1978). "A demographic analysis of the snowshoe hare cycle." Wildlife Monographs **58**.

King, A. A. and W. M. Schaffer (2001). "The geometry of a population cycle: A mechanistic model of snowshoe hare demography." Ecology **82**(3): 814-830.

Kingsland, S. E. (1995). Modeling Nature: Episodes in the History of Population Ecology, University of Chicago Press, Chicago, Ill.

Klemola, T., M. Tanhuanpaa, et al. (2002). “Specialist and generalist natural enemies as an explanation for geographical gradients in population cycles of northern herbivores.” Oikos **99:** 83-94.

Koenig, W. D. (1999). "Spatial autocorrelation of ecological phenomena." Trends in Ecology and Evolution **14**(1): 22-26.

Koenig, W. D. and J. M. H. Knops (1998). "Testing for spatial autocorrelation in ecological studies." Ecography **21**(4): 423-429.

Koenig, W. D. and J. M. H. Knops. (2001). “Seed-crop size and eruptions of North American boreal seed-eating birds.” Journal of Animal Ecology **70:** 609-620.

Krebs, C. J. (2011). “Of lemmings and snowshoe hares: the ecology of northern Canada” Proceedings of the Royal Society of London, Series B. Advance online article published October 2010.

Krebs, C. J., R. Boonstra, et al. (2001). "What drives the 10-year cycle of snowshoe hares?" BioScience **51**(1): 25-35.

Kvasnes, M. A. J., T. Storaas, et al. (2010). "Spatial dynamics of Norwegian tetraonid populations." Ecological Research **25**(2): 367-374.

Lambin, X., D. A. Elston, et al. (1998). "Spatial asynchrony and periodic travelling waves in cyclic populations of field voles." Proceedings of the Royal Society of London. Series B **265:** 1491-1496.

Liebhold, A., W. D. Koenig, et al. (2004). "Spatial synchrony in population dynamics." Annual Review of Ecology Evolution and Systematics **35**: 467-490.

MacFarlane, R. (1905). "Notes on mammals collected and observed in the northern Mackenzie River district, Northwest Territories of Canada, with remarks on explorers and explorations of the far North." Proceedings of the U.S. National Museum **28**(1405):
673-764.

MacLulich, D. A. (1937). "Fluctuations in numbers of the varying hare (Lepus americanus)." University of Toronto Studies, Biological Series **43**: 1-136.

Mills, L. S., P. C. Griffin, et al. (2005). "Pellet count indices compared to mark-recapture estimates for evaluating snowshoe hare density." Journal of Wildlife Management **69**(3): 1053-1062.

Moran, P. A. P. (1953). "The statistical analysis of the Canadian lynx cycle, II: synchronization and meteorology." Australian Journal of Zoology **1**(3): 291-298.

Murray, D. L. (2000). "A geographic analysis of snowshoe hare population demography." Canadian Journal of Zoology **78**(7): 1207-1217.

Murray, D. L. (2003). Snowshoe hares and other hares. Wild Mammals of North America, Vol. II. G. A. Feldhamer and B. C. Thompson. Baltimore, MD, Johns Hopkins University Press.

Ottersen, G., B. Planque, et al. (2001). "Ecological effects of the North Atlantic Oscillation." Oecologia 128(1): 1-14.

Paradis, E., S. R. Baillie, et al. (2000). "Spatial synchrony in populations of birds: effects of habitat, population trend, and spatial scale." Ecology **81**(8): 2112-2125.

Parker, G. (2001). Status Report on the Canada Lynx in Nova Scotia**:** 1-59.

Ranta, E., V. Kaitala, et al. (1997b). "Dynamics of Canadian lynx populations in space and time." Ecography **20**(5): 454-460.

Ranta, E., J. Lindström, et al. (1997a). "Solar activity and hare dynamics: a cross-continental comparison." The American Naturalist **149**: 765-775.

Royama, T. (1992). Analytical Population Dynamics, Chapman & Hall, London.

Ruesink, J. L., K. E. Hodges, et al. (2002). "Mass-balance analyses of boreal forest population cycles: Merging demographic and ecosystem approaches." Ecosystems **5**(2): 138-158.

Ruggiero, L. F., K. B. Aubry, et al. (2000). Ecology and Conservation of Lynx in the United States. F. S. U.S. Department of Agriculture. Boulder, CO, USA, University Press of Colorado.

Saitoh, T., N.C. Stenseth, et al. (1998). “The population dynamics of the grey- sided vole in Hokkaido, Japan.” Researches on Population Ecology **40:**61-76.

Schwartz, M. K., L. S. Mills, et al. (2002). "DNA reveals high dispersal synchronizing the population dynamics of Canada lynx." Nature **415**: 520-522.

Sinclair, A. R. E., D. Chitty, et al. (2003). "Mammal population cycles: evidence for intrinsic differences during snowshoe hare cycles." Canadian Journal of Zoology **81**(2): 216-220.

Smith, C. H. (1983). "Spatial trends in Canadian snowshoe hare, Lepus americanus, population cycles." Canadian Field-Naturalist **97**: 151-160.

Smith, C. H. and J. M. Davis (1981). "A spatial analysis of wildlife's ten-year cycle." Journal of Biogeography **8**(1): 27-35.

Steen, H., R. A. Ims, et al. (1996). "Spatial and temporal patterns of small-rodent population dynamics at a regional scale." Ecology **77**(8): 2365-2372.

Stenseth, N. C., K. Chan, et al. (1999). "Common dynamic structure of Canada lynx populations within three climatic regions." Science **285**(5430): 1071-1073.

Stenseth, N. C., D. Ehrich, et al. (2004). "The effect of climatic forcing on population synchrony and genetic structuring of the Canadian lynx." Proceedings of the National Academy of Sciences **01**(16): 6056-6061.

Stone, L. (2004). “A three-player solution” Nature **430:** 299-300.

Tkadlec, E., and Stenseth, N.C. (2001). “A new geographical gradient in vole population dynamics.” Proceedings of the Royal Society of London, Series B – Biological Sciences **268:** 1547-1552.

Turchin, P. (2003). Complex Population Dynamics: a Theoretical / Empirical Synthesis, Princeton University Press.

Turchin, P. and T. D. Hall (2003). "Spatial synchrony among and within world-systems: insights from theoretical ecology." Journal of World-Systems Research **9**(1): 37-64.

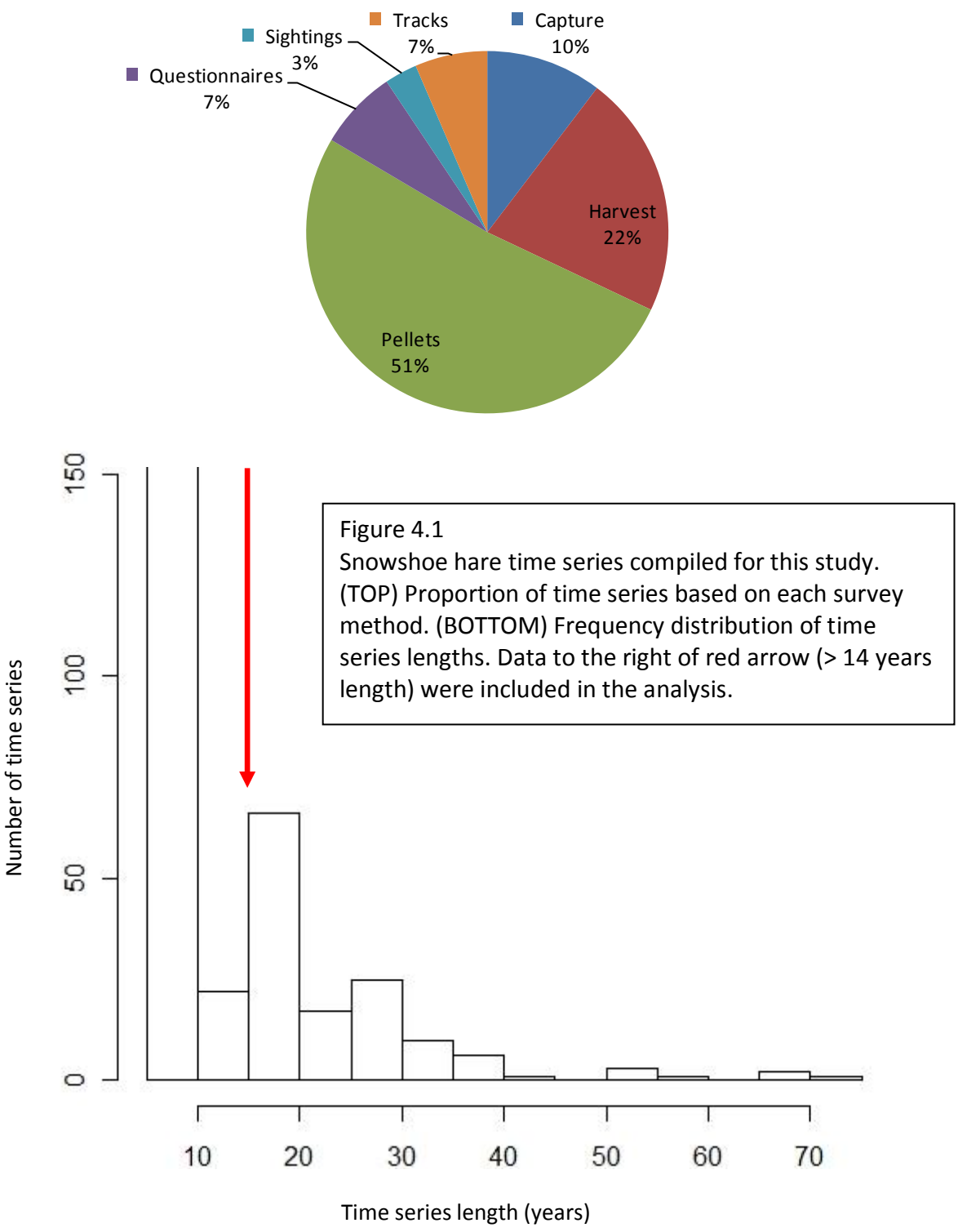


Figure 4.1
Snowshoe hare time series compiled for this study. (TOP) Proportion of time series based on each survey method. (BOTTOM) Frequency distribution of time series lengths. Data to the right of red arrow (> 14 years length) were included in the analysis.

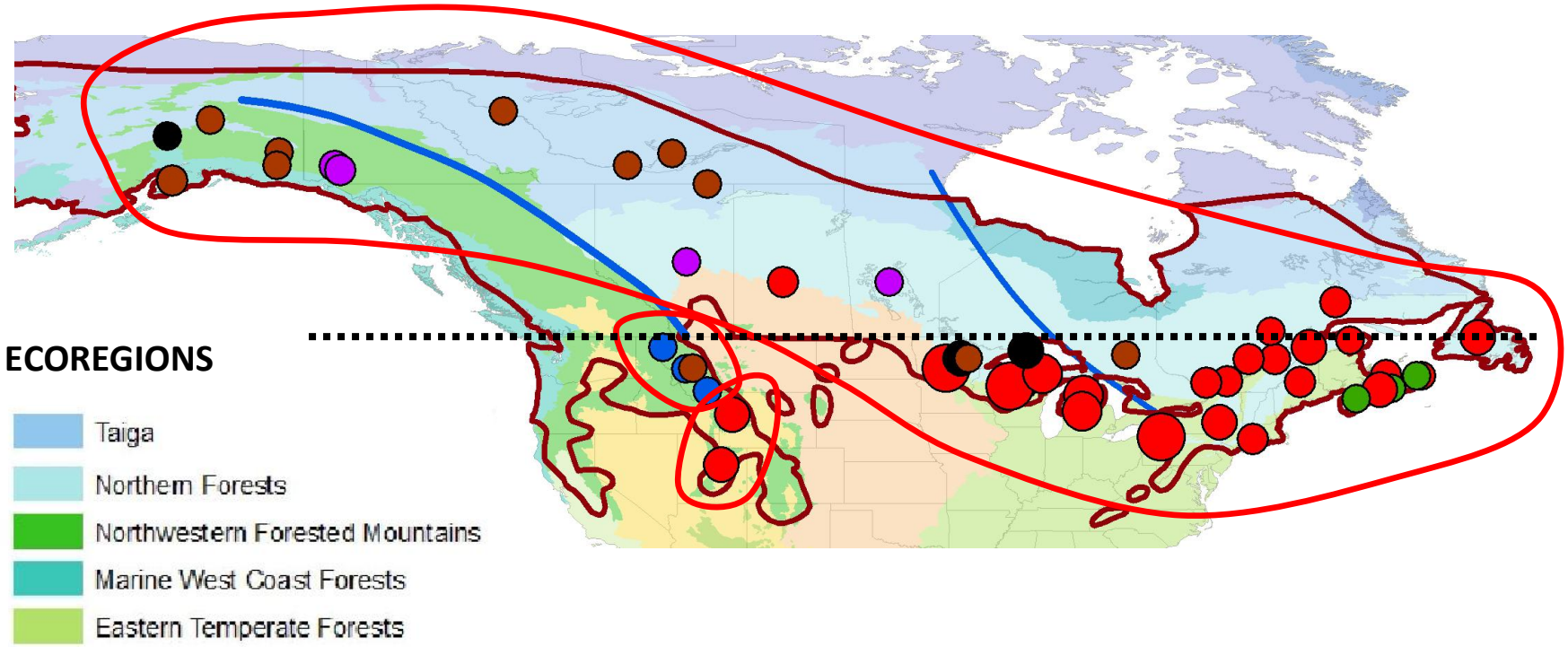


Figure 4.2
Distribution of 49 snowshoe hare time series data analyzed in this study. Size of circle is proportional to length of time series. Color of circle indicates data type: red = harvest, brown = pellets, blue = tracks, black = sightings, green = questionnaires, purple = live-trap. Ecoregions are color-coded as shown in legend. Blue lines separate three NAO climatic regions. Red lines delineate three genetic groups. Black dotted line indicates 49th parallel separating northern from southern hare populations in this study.

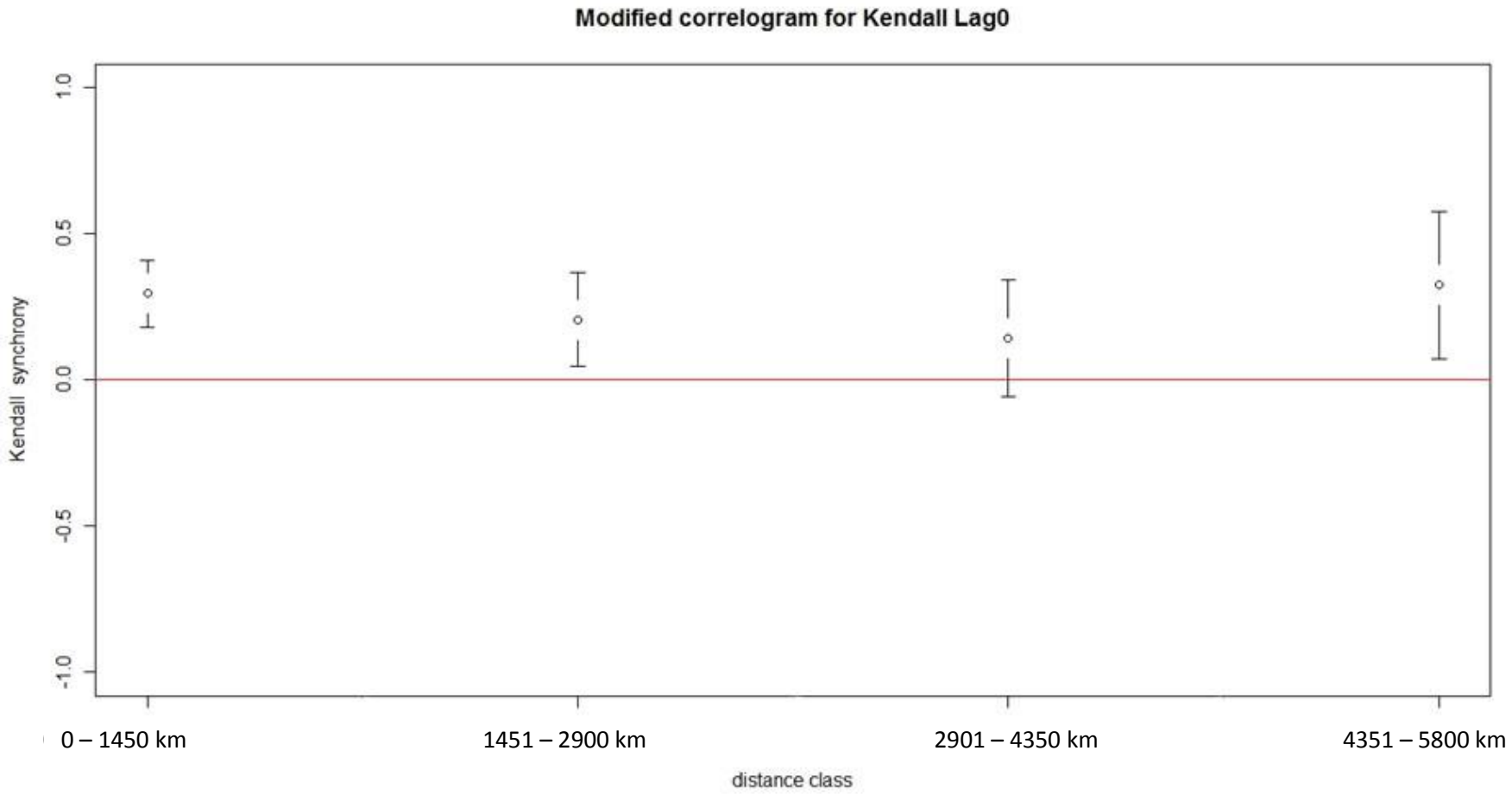


Figure 4.3
Modified correlogram of synchrony against distance for northern snowshoe hares. Error bars represent 95% confidence intervals. The red line indicates zero synchrony.

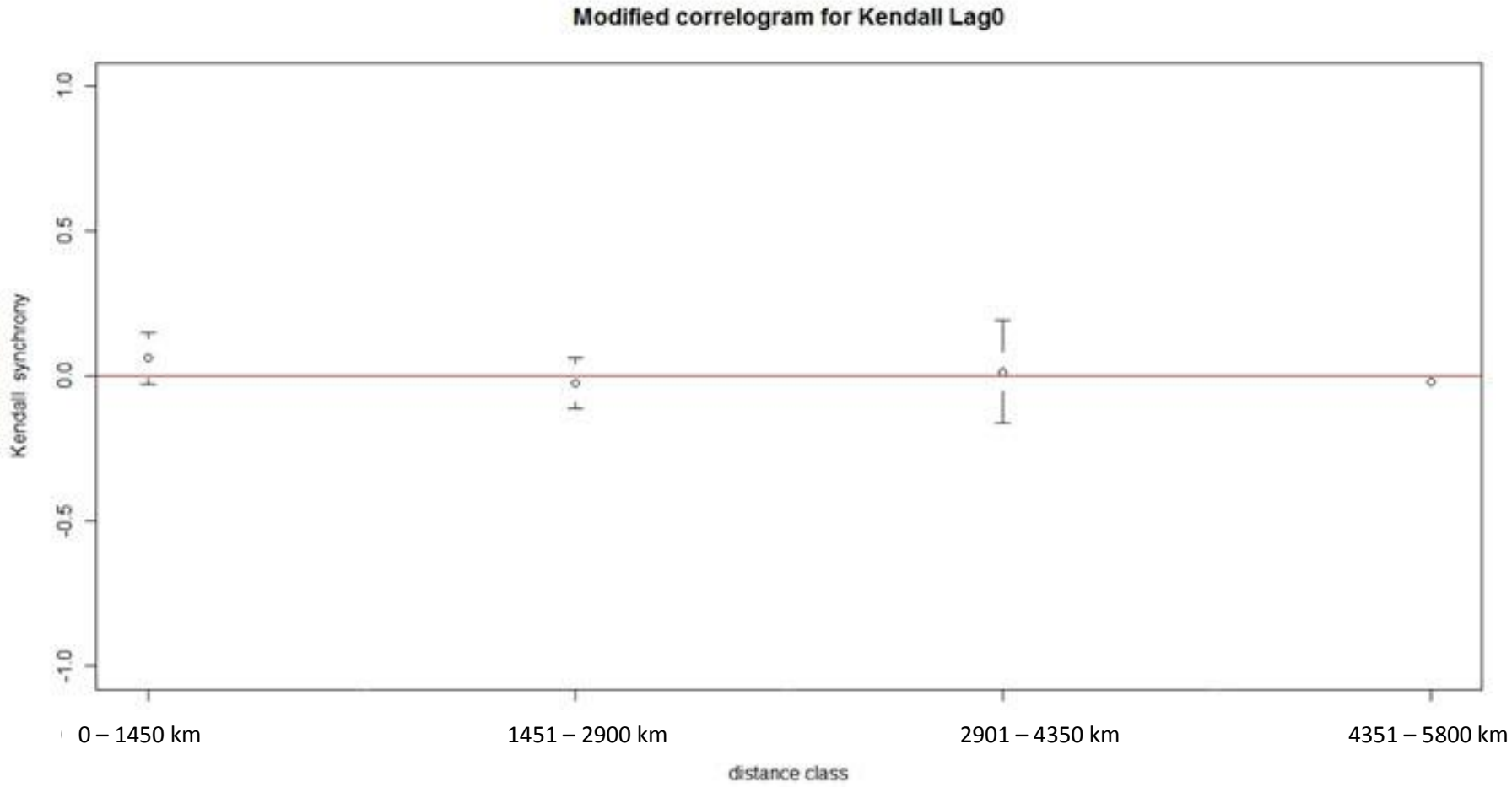


Figure 4.4
Modified correlogram of synchrony against distance for southern snowshoe hares. Error bars represent 95% confidence intervals. The red line indicates zero synchrony.

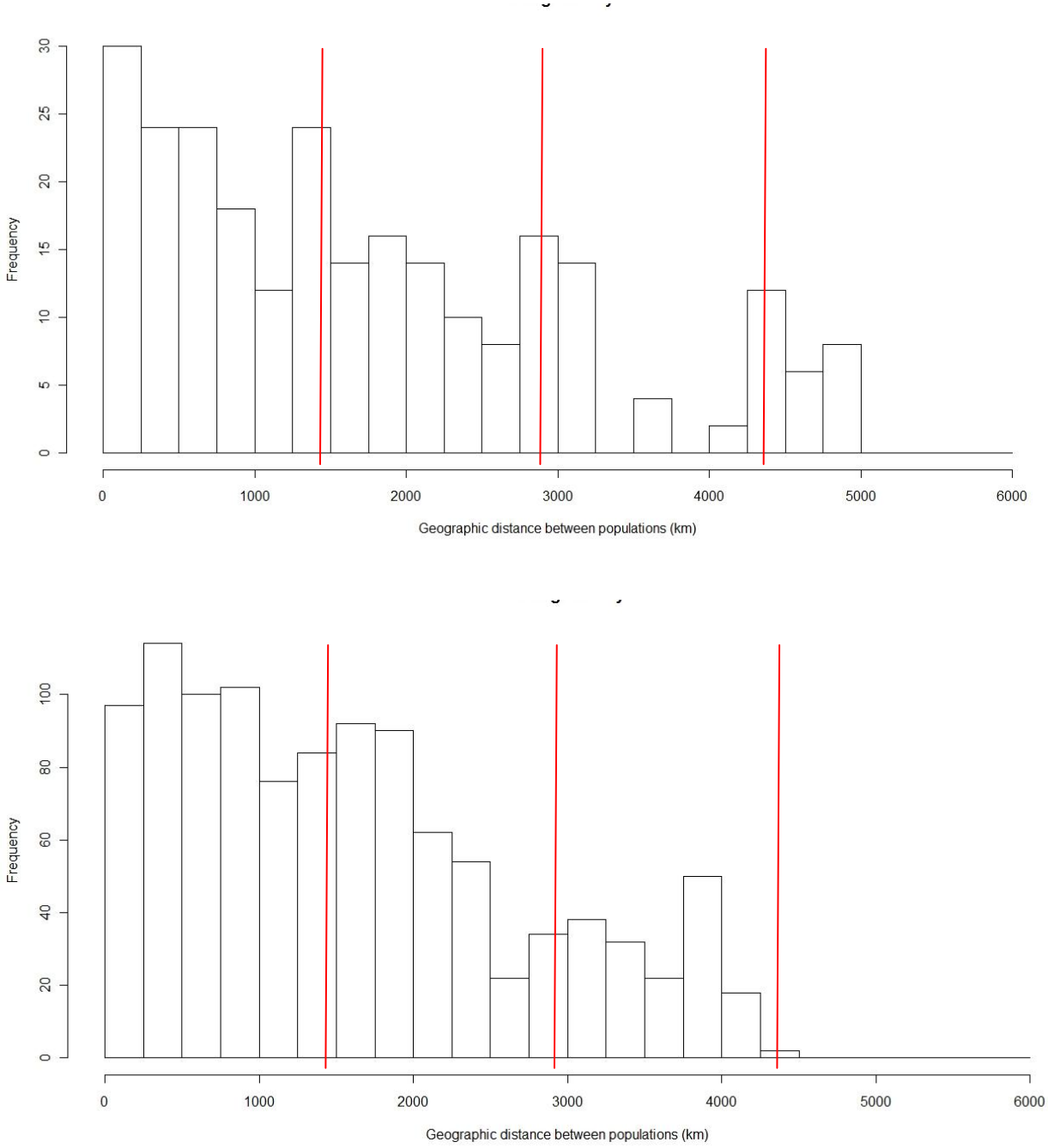


Figure 4.5
Distribution of pairwise geographic distances for (TOP) northern and (BOTTOM) southern snowshoe hare populations. Red lines separate the geographic distance categories used in modified correlogram analyses. All analyzed time series are included.

	QC1	QC2	QC3	QC4	QC5	QC6	QC7	QC8
AK1	0.11	0.30	0.32	0.12	-0.01	0.09	0.23	-0.09
AK2	0.32	0.26	0.40	-0.14	-0.18	0.32	0.30	NA
AK3	0.44	0.33	0.51	0.05	-0.21	0.54	0.49	NA
AK4	0.46	0.33	0.49	-0.05	0.18	0.67	0.56	NA
AK5	0.04	-0.20	-0.08	0.03	-0.05	-0.04	0.03	-0.38

	QC1	QC2	QC3	QC4	QC5	QC6	QC7	QC8
AK1	0.29	0.06	0.05	0.29	0.46	0.29	0.10	0.60
AK2	0.08	0.12	0.04	0.71	0.79	0.05	0.09	NA
AK3	0.01	0.05	0.01	0.39	0.80	0.00	0.02	NA
AK4	0.03	0.10	0.01	0.57	0.21	0.00	0.00	NA
AK5	0.43	0.88	0.63	0.44	0.62	0.55	0.41	0.95

Table 4.1
(TOP) Kendall synchrony estimate between Alaska and Quebec snowshoe hare populations. QC3 and QC7 (highlighted yellow) are northern populations; remainder of Quebec populations are southern. (BOTTOM) P-values for synchrony estimates. “NA” indicates insufficient (<12 years of data overlapping) to estimate synchrony.

Appendix 4.1: Evaluation of Kendall synchrony metric

Justification

Previous studies have taken a variety of approaches to analyzing snowshoe hare and Canada lynx synchrony patterns (e.g., Smith 1983; Ranta et al. 1997a, 1997b; Stenseth et al. 1999) across their northern range. One of the most famous results of these studies has been identification of traveling wave dynamics across large portions of Canada (Smith 1983; Ranta et al. 1997a, 1997b). To my knowledge, the Kendall synchrony metric has not yet been applied to studying lynx and hare synchrony patterns. I chose this metric for my study because a previous simulation study (Chapter 3) showed the Kendall metric was more robust to short time series and sampling error, and less prone to Type I error than other metrics evaluated. These metric qualities were important for my study, which combined data from multiple sources, survey methods, and time periods into a single synchrony analysis.

Prior to using the Kendall metric in this analysis, I wanted to confirm it performed comparably with other methods previously applied to the Canadian Snowshoe Rabbit Enquiry data. However, those data were recorded as categorical variables (“Increase”, “Decrease”, “No Trend”), which could not be analyzed with the Kendall metric. I did not use a similar metric proposed by Buonoccorsi et al. (2001; Percent Match metric) for analysis of categorical data because simulations showed this metric had relatively high Type I error rates (Chapter 3). Therefore, instead of using Snowshoe Rabbit Enquiry data, I applied the Kendall metric to the Canada lynx fur

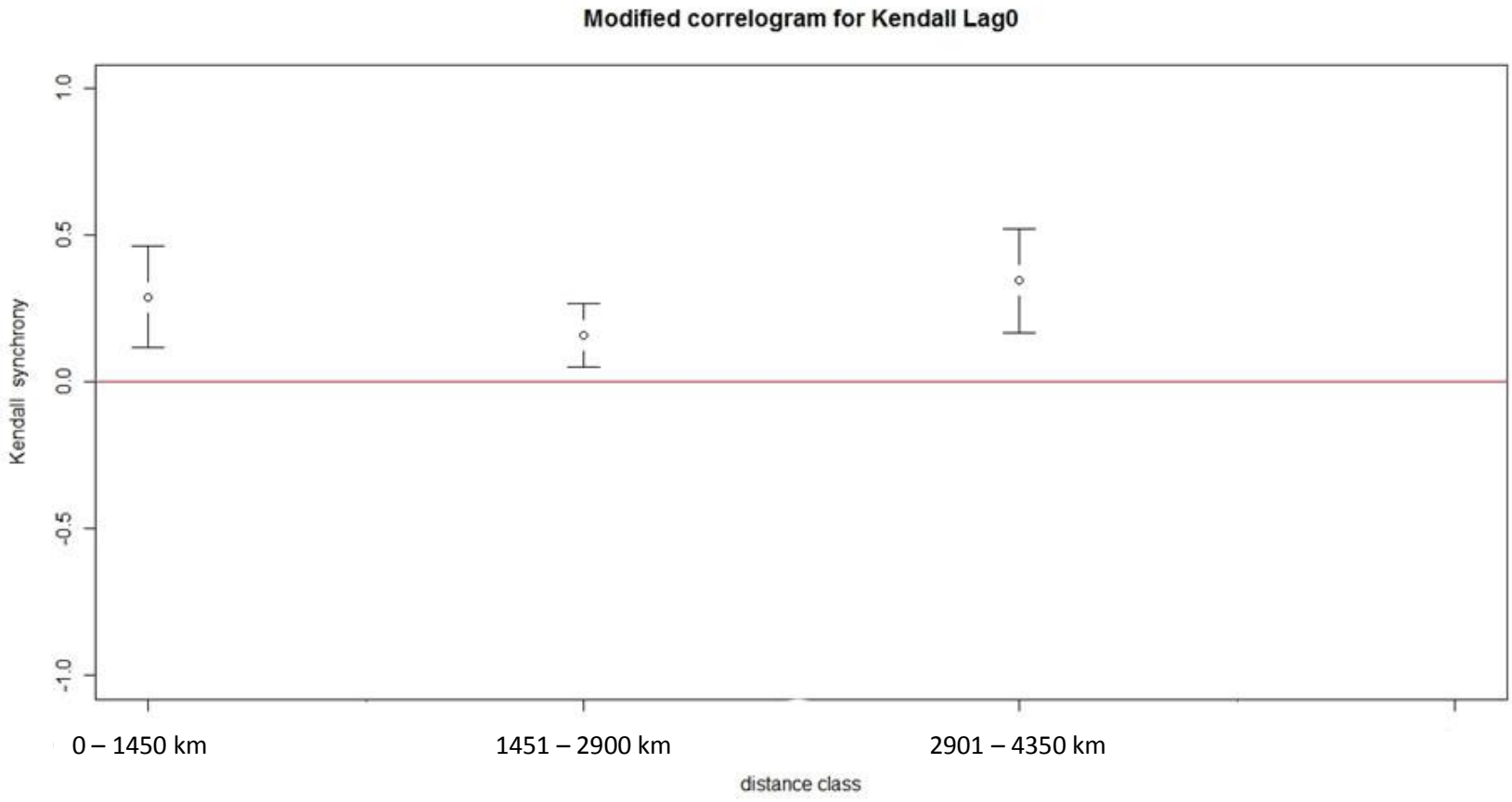
return data collected by Statistics Canada for 1919-1986. Ranta et al. (1997b) had identified a traveling wave pattern of synchrony in these data similar to the pattern observed from the Snowshoe Rabbit Enquiry Data.

Methods

To test the Kendall metric's performance on the lynx data, I calculated Kendall synchrony for every pairwise comparison of the eight time series used in Ranta et al. (1997b). I calculated a modified correlogram (Koenig and Knops 1998) of these data binned into three equal distance categories from 0 to 4350 km, the maximum distance between two time series. If the Kendall metric performed comparably with previous metrics in identifying traveling wave synchrony in the lynx data, I expected to find a U-shaped relationship between synchrony and geographic distance, with highest synchrony in the smallest and largest distance classes. Ranta et al. (1997a; 1997b) have identified this U-shaped pattern as a signature of traveling wave synchrony in previous studies of Canada lynx and snowshoe hares.

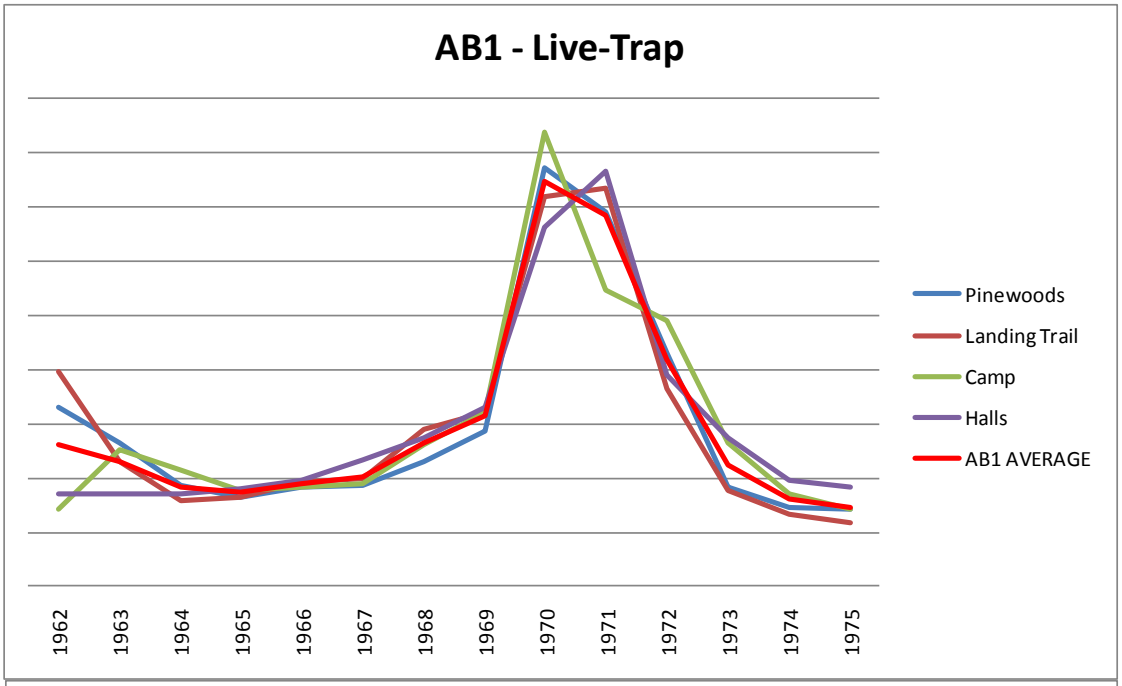
Results

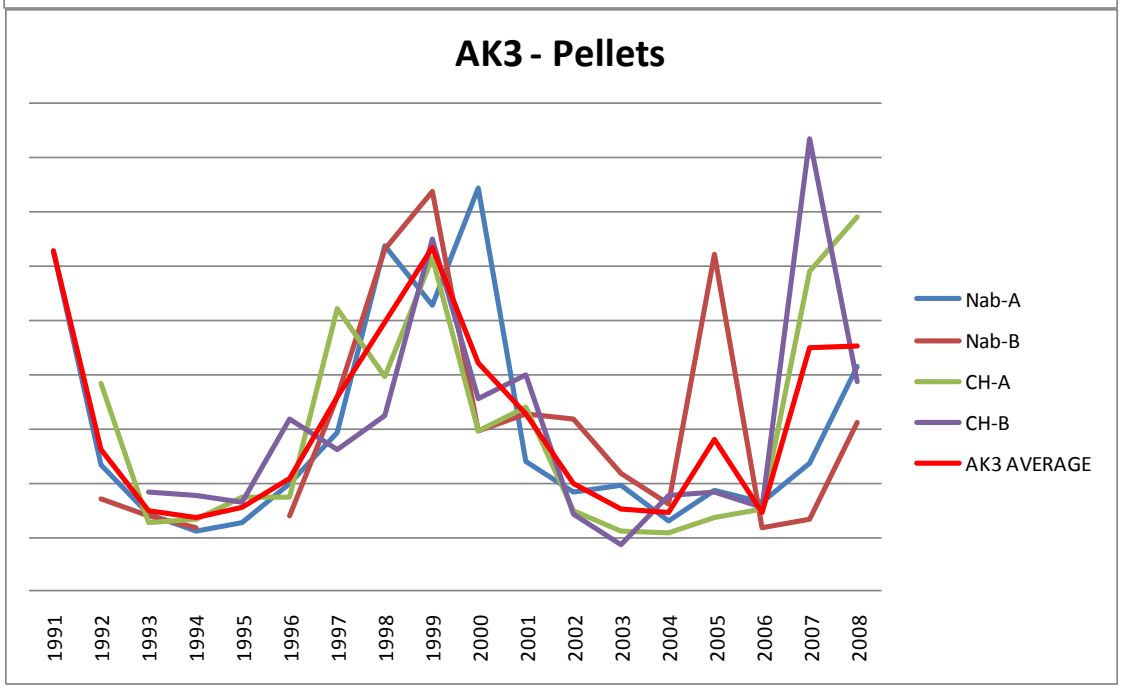
The Kendall metric recovered a U-shaped synchrony-distance relationship qualitatively similar to that reported by Ranta et al. (1997b), using the same eight Canada lynx time series spanning 1919-1986. Synchrony was significant in all distance classes.

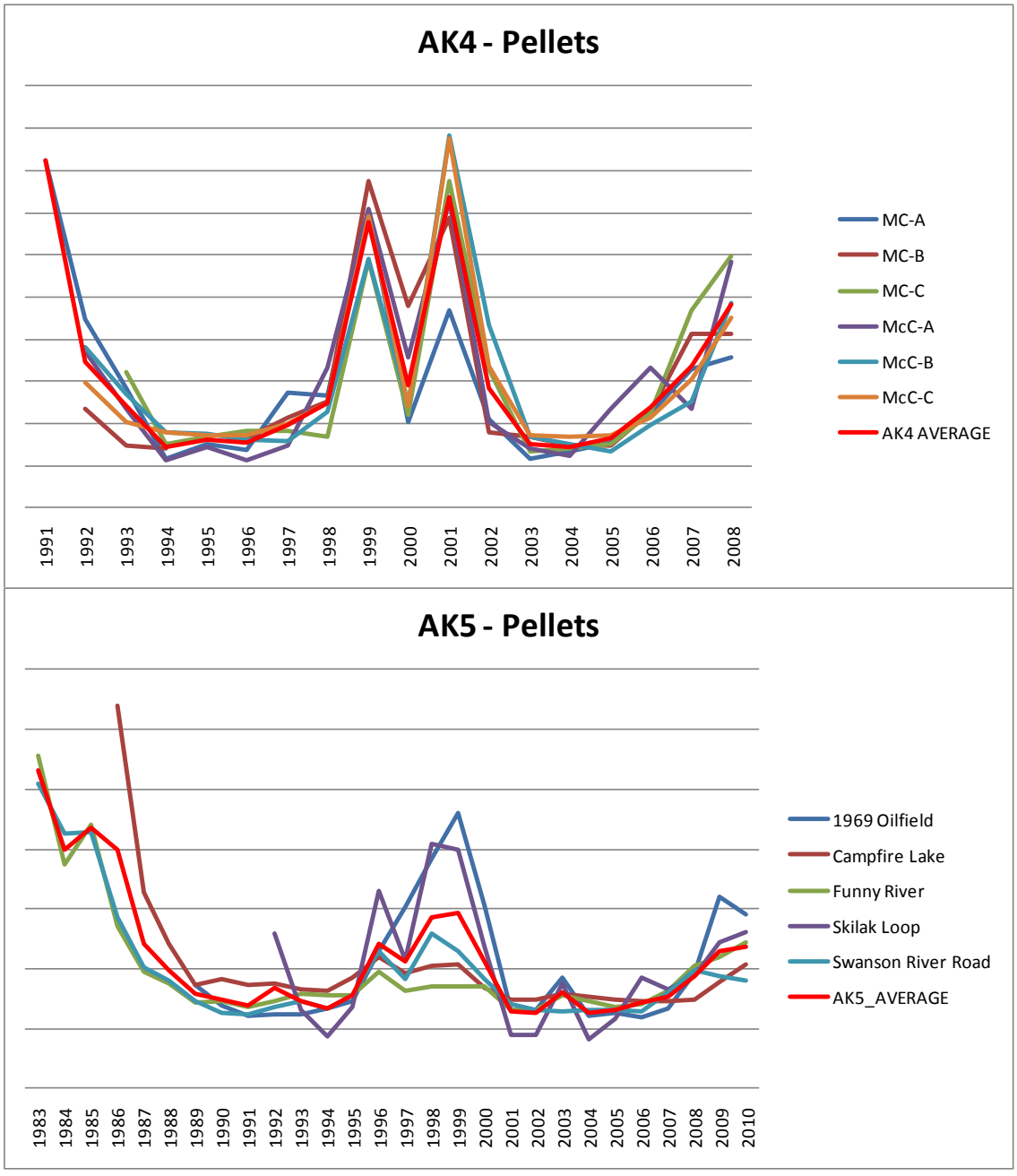


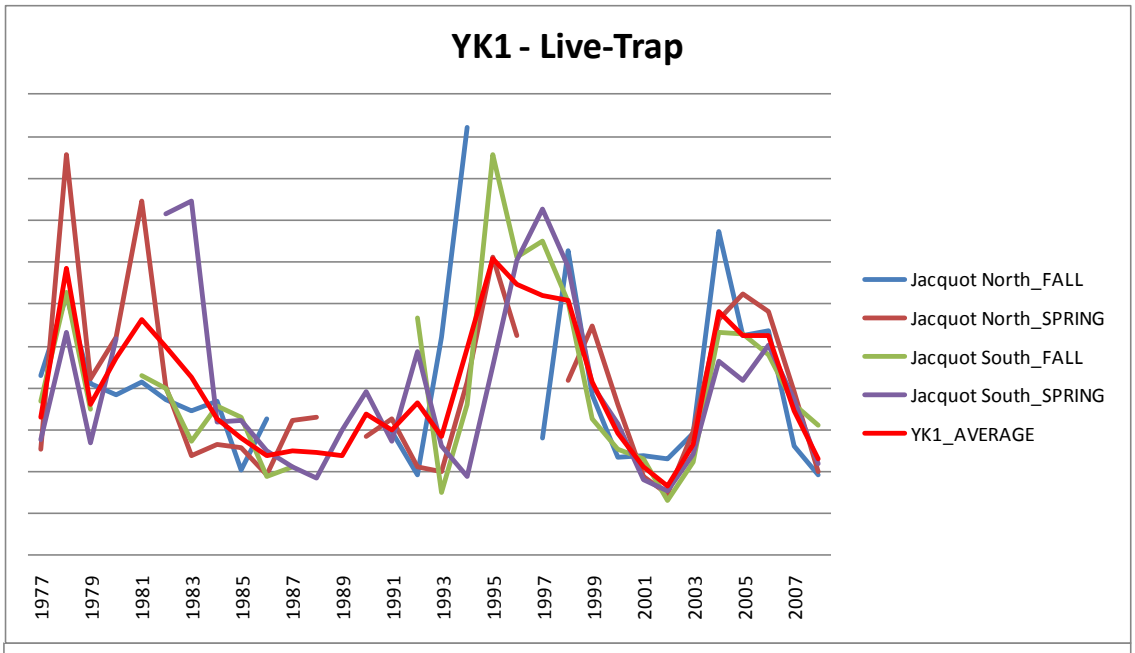
Appendix 4.1 (cont'd)
Modified correlogram of synchrony against distance for Canada lynx, 1919-1986. Error bars represent 95% confidence intervals. The red line indicates zero synchrony.

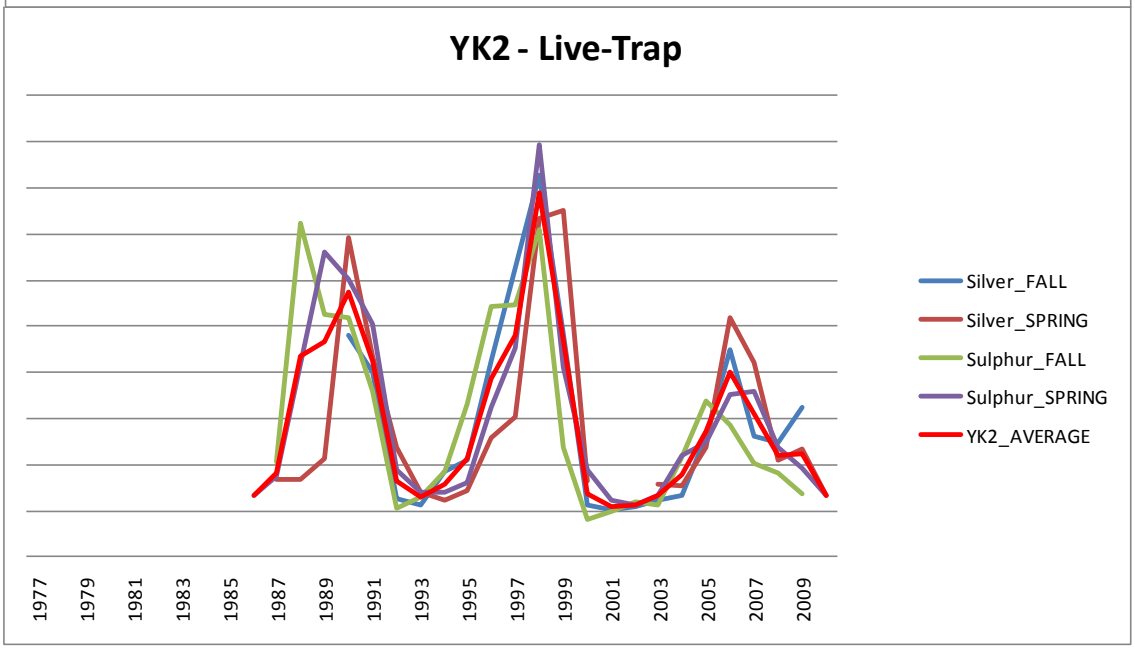
Appendix 4.2
Small-scale time series data combined for synchrony analysis. Each graph represents a set of time series combined for analysis. The red line represents the combined time series.

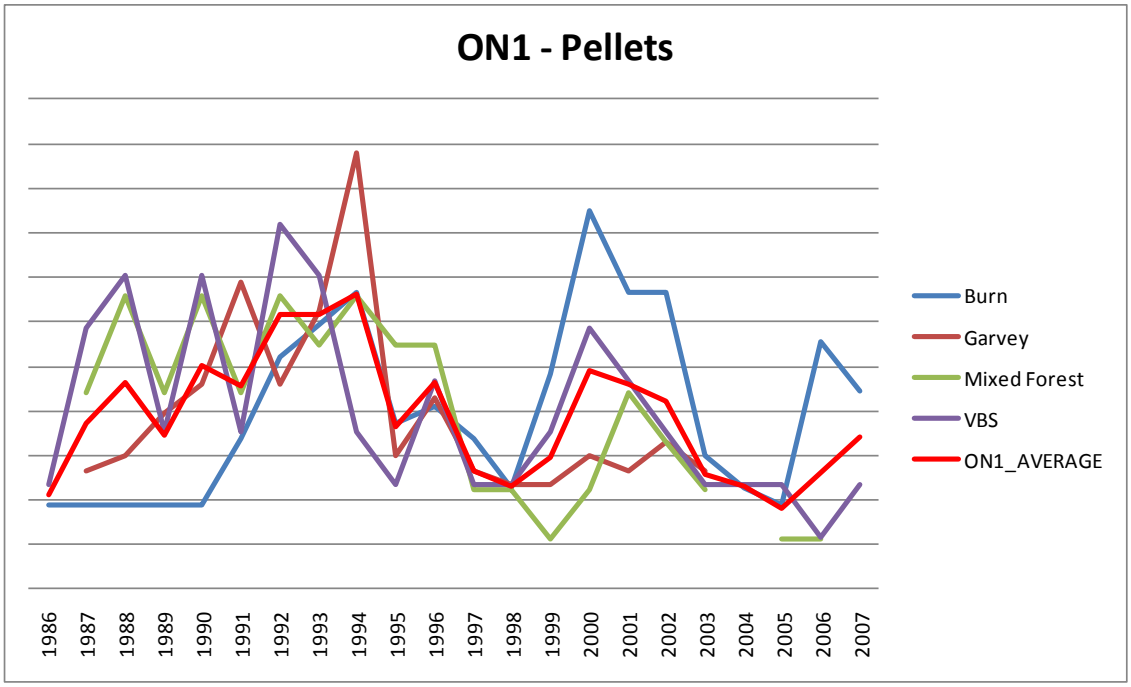




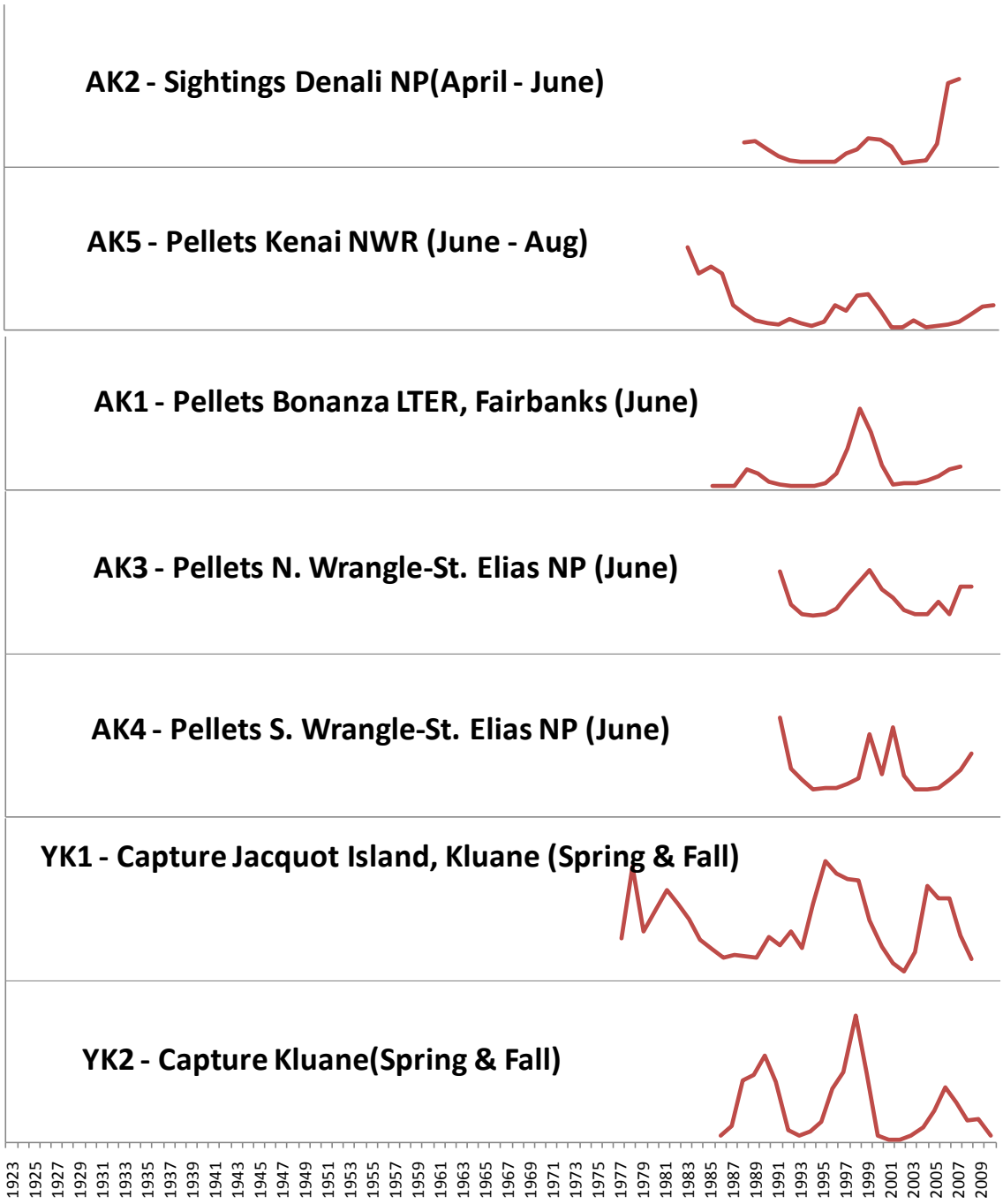

AK4 - Pellets
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
MC-A
MC-B
MC-C
McC-A
McC-B
McC-C
AK4 AVERAGE
AK5 - Pellets
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
1969 Oilfield
Campfire Lake
Funny River
Skilak Loop
Swanson River Road
AK5_AVERAGE

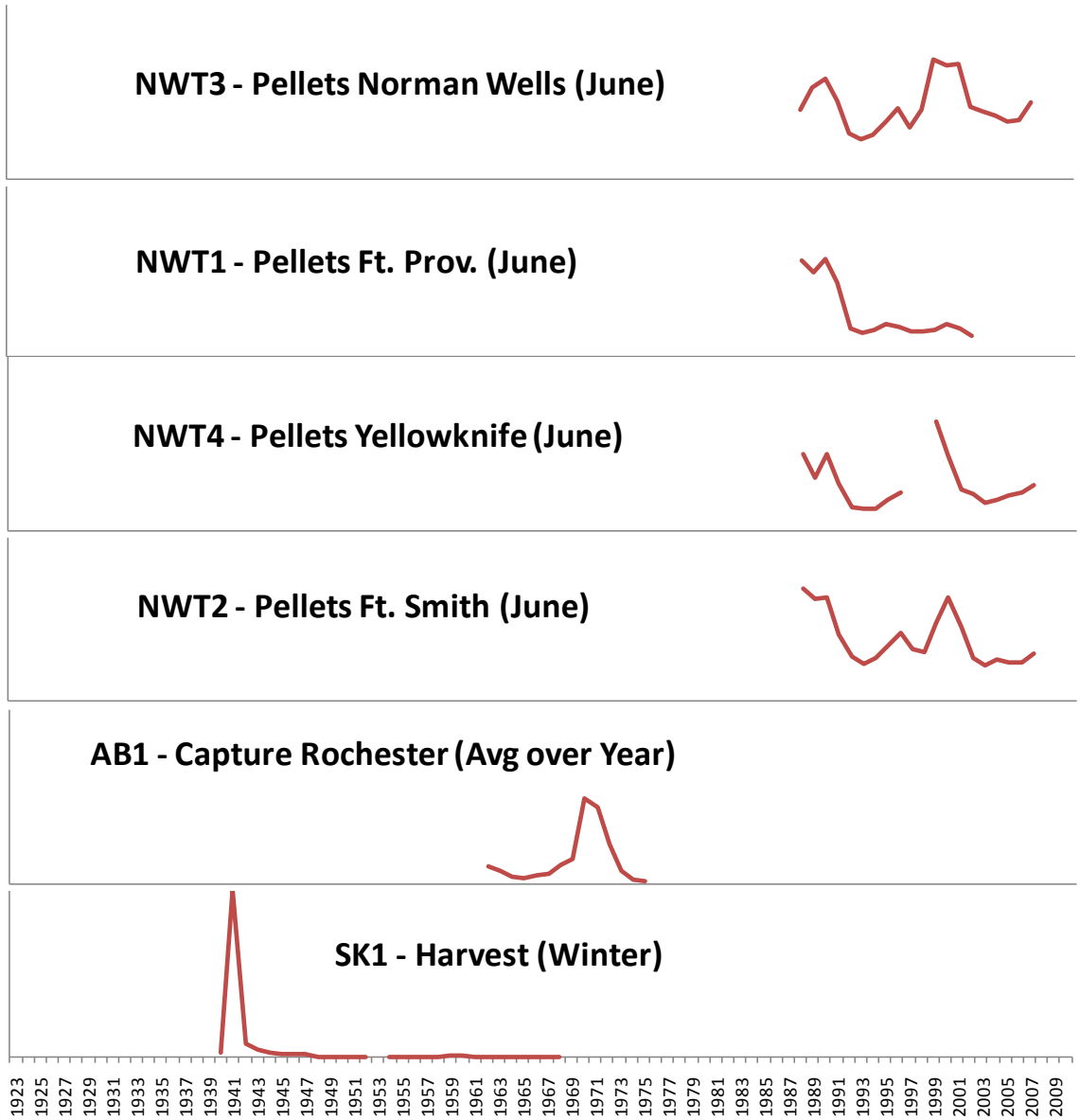

YK1 - Live-Trap
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
Jacquot North_FALL
Jacquot North_SPRING
Jacquot South_FALL
Jacquot South_SPRING
YK1_AVERAGE

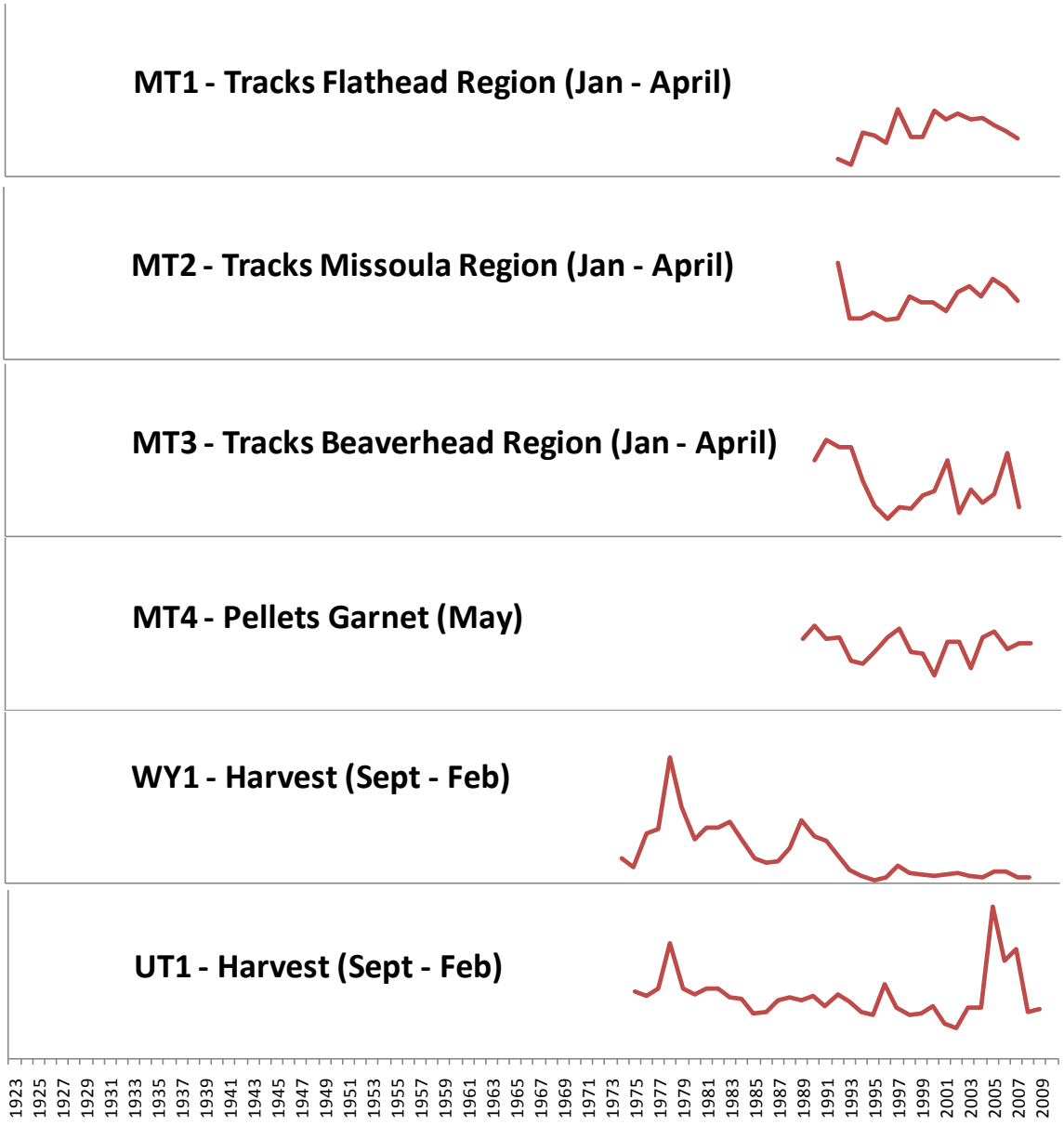

YK2 - Live-Trap
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
Silver_FALL
Silver_SPRING
Sulphur_FALL
Sulphur_SPRING
YK2_AVERAGE

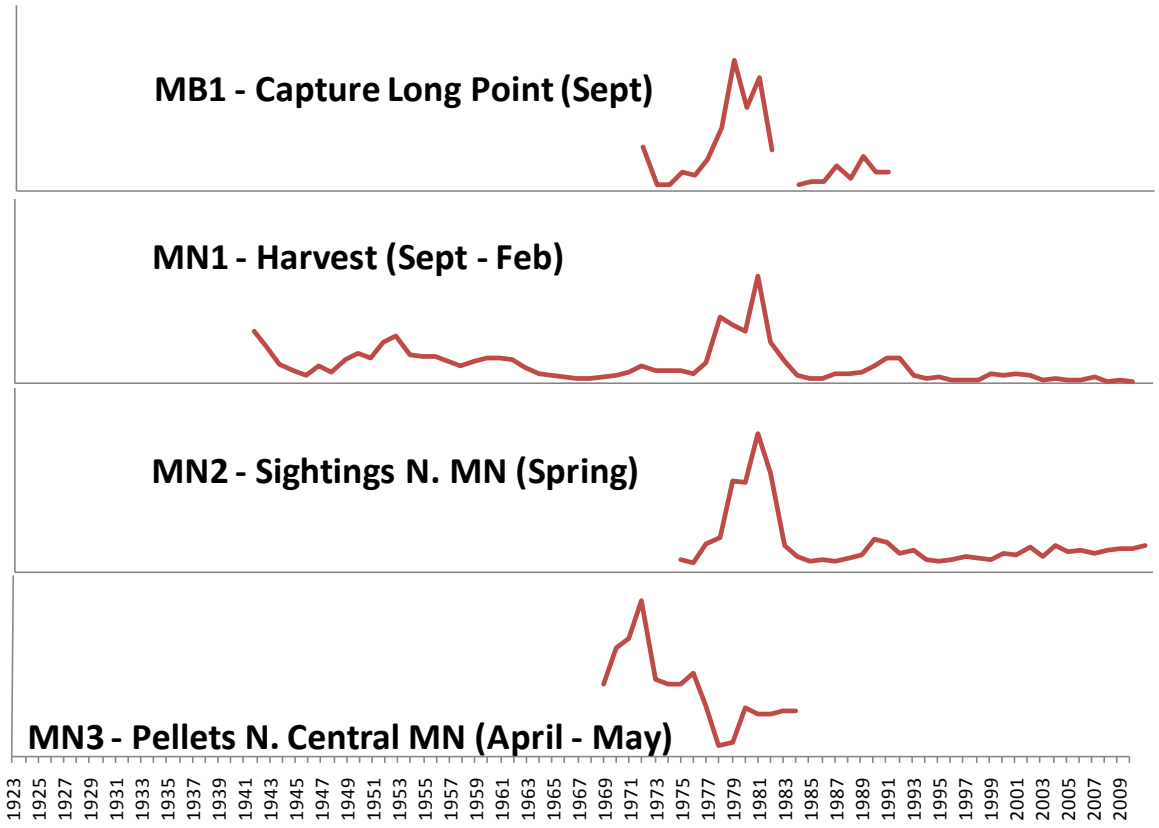

ON1 - Pellets
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
Burn
Garvey
Mixed Forest
VBS
ON1_AVERAGE

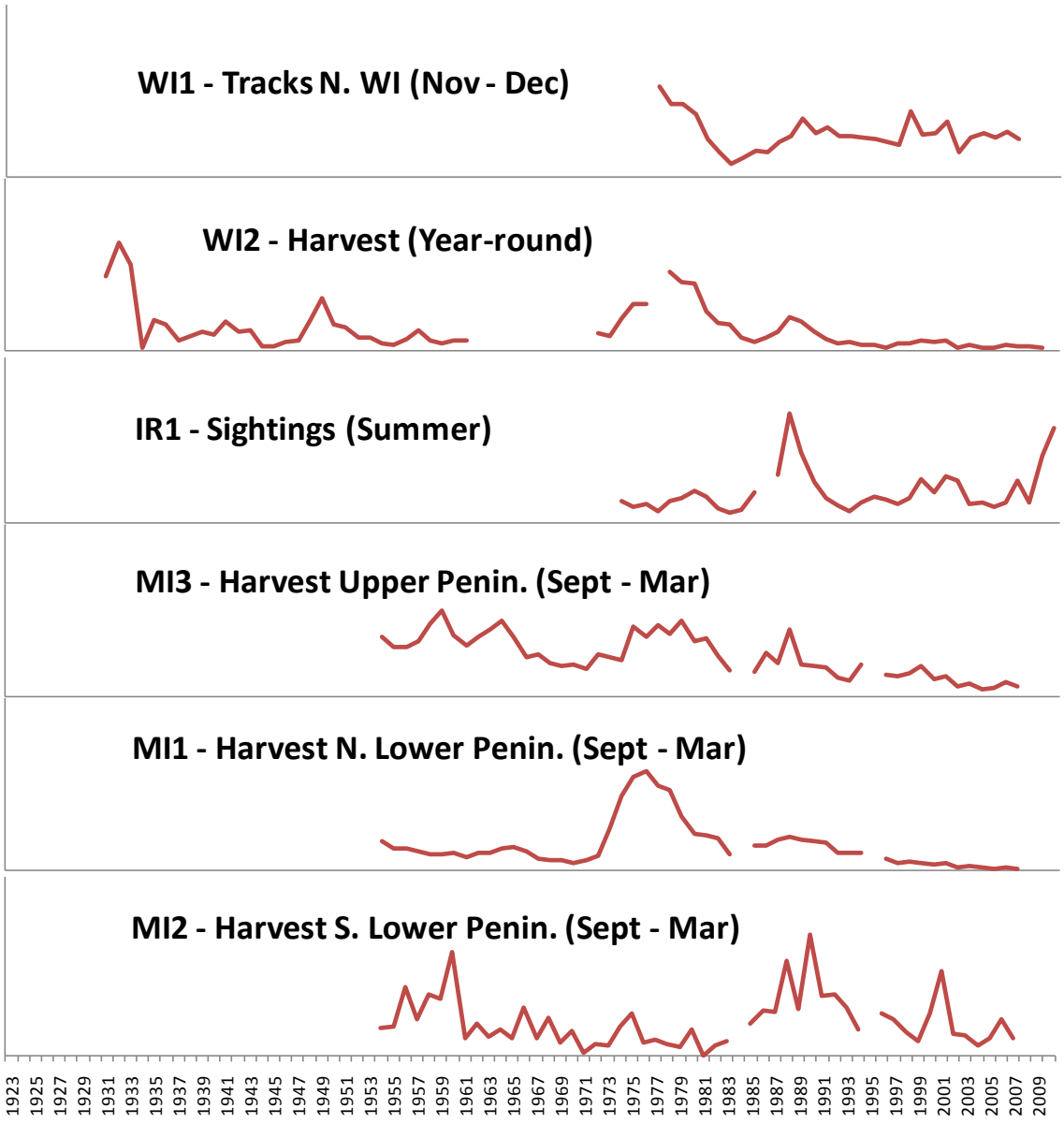
Appendix 4.3
Snowshoe hare time series analyzed in this study, grouped by region.

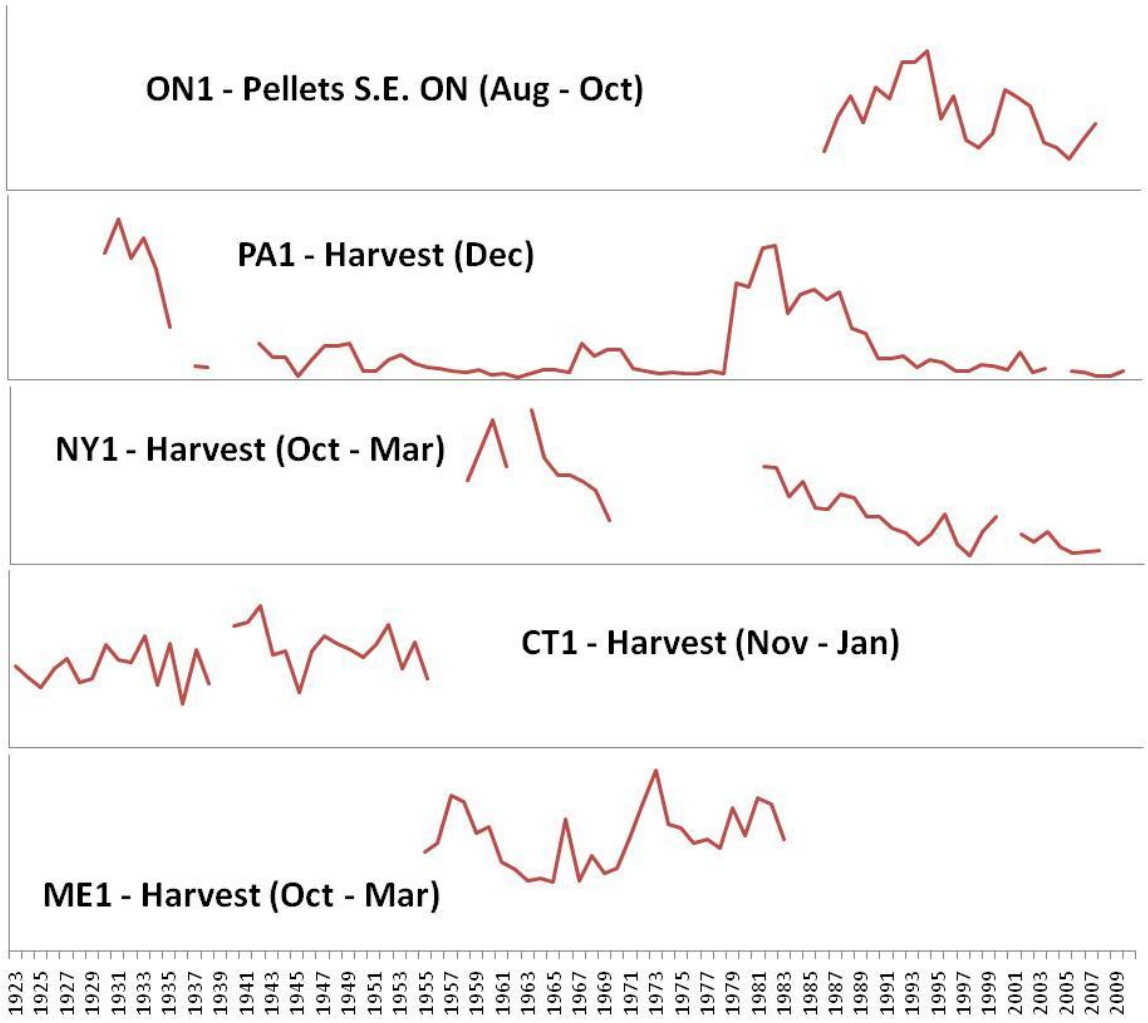


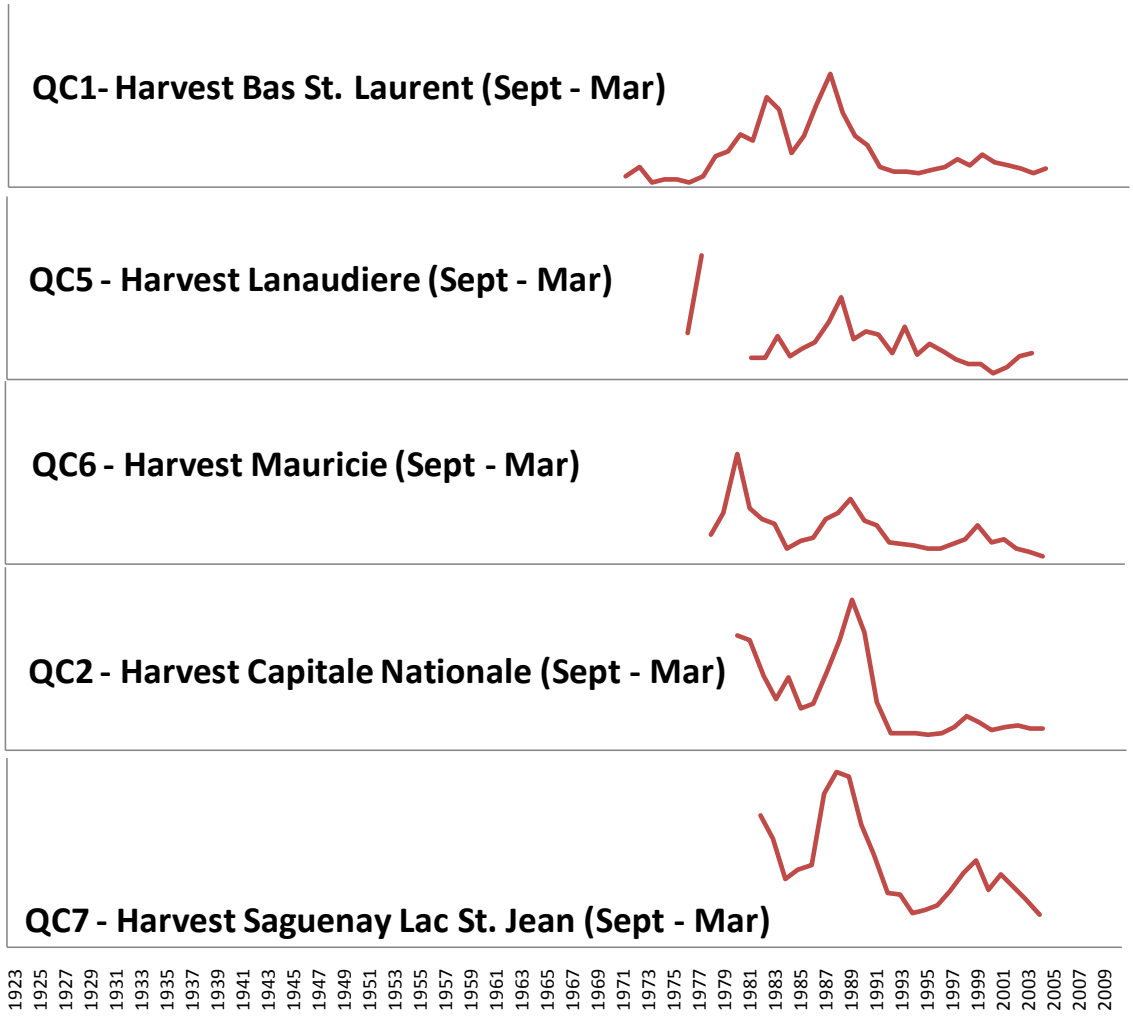

NWT3 - Pellets Norman Wells (June)
NWT1 - Pellets Ft. Prov. (June)
NWT4 - Pellets Yellowknife (June)
NWT2 - Pellets Ft. Smith (June)
AB1 - Capture Rochester (Avg over Year)
SK1 - Harvest (Winter)
1923
1925
1927
1929
1931
1933
1935
1937
1939
1941
1943
1945
1947
1949
1951
1953
1955
1957
1959
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009

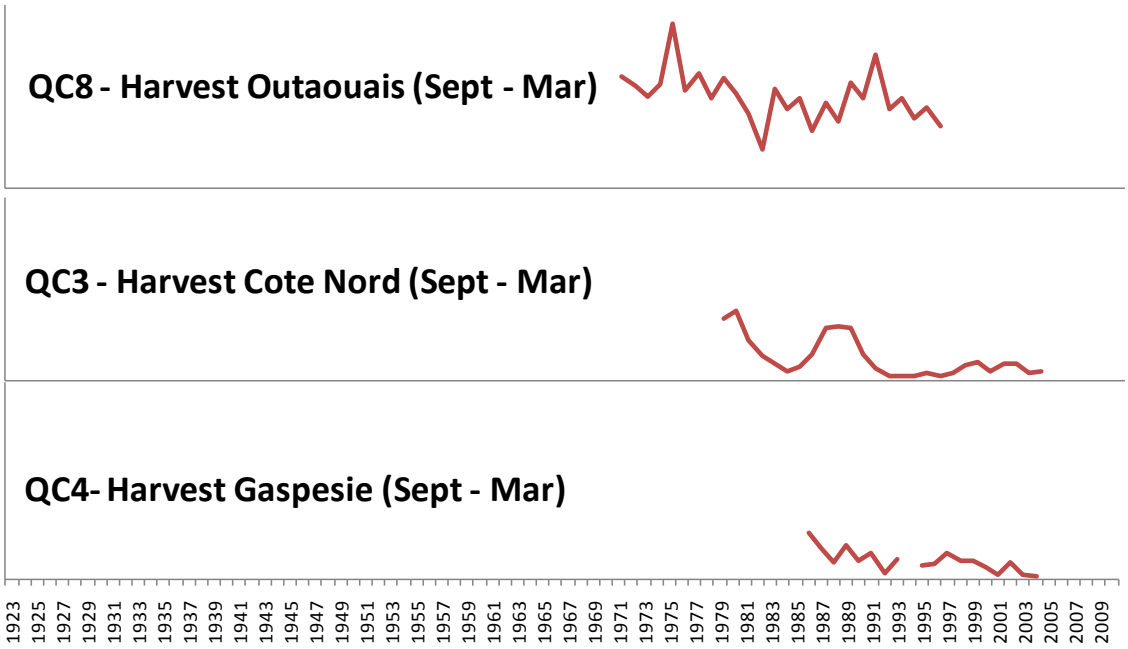

MT1 - Tracks Flathead Region (Jan - April)
MT2 - Tracks Missoula Region (Jan - April)
MT3 - Tracks Beaverhead Region (Jan - April)
MT4 - Pellets Garnet (May)
WY1 - Harvest (Sept - Feb)
UT1 - Harvest (Sept - Feb)
1923
1925
1927
1929
1931
1933
1935
1937
1939
1941
1943
1945
1947
1949
1951
1953
1955
1957
1959
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009

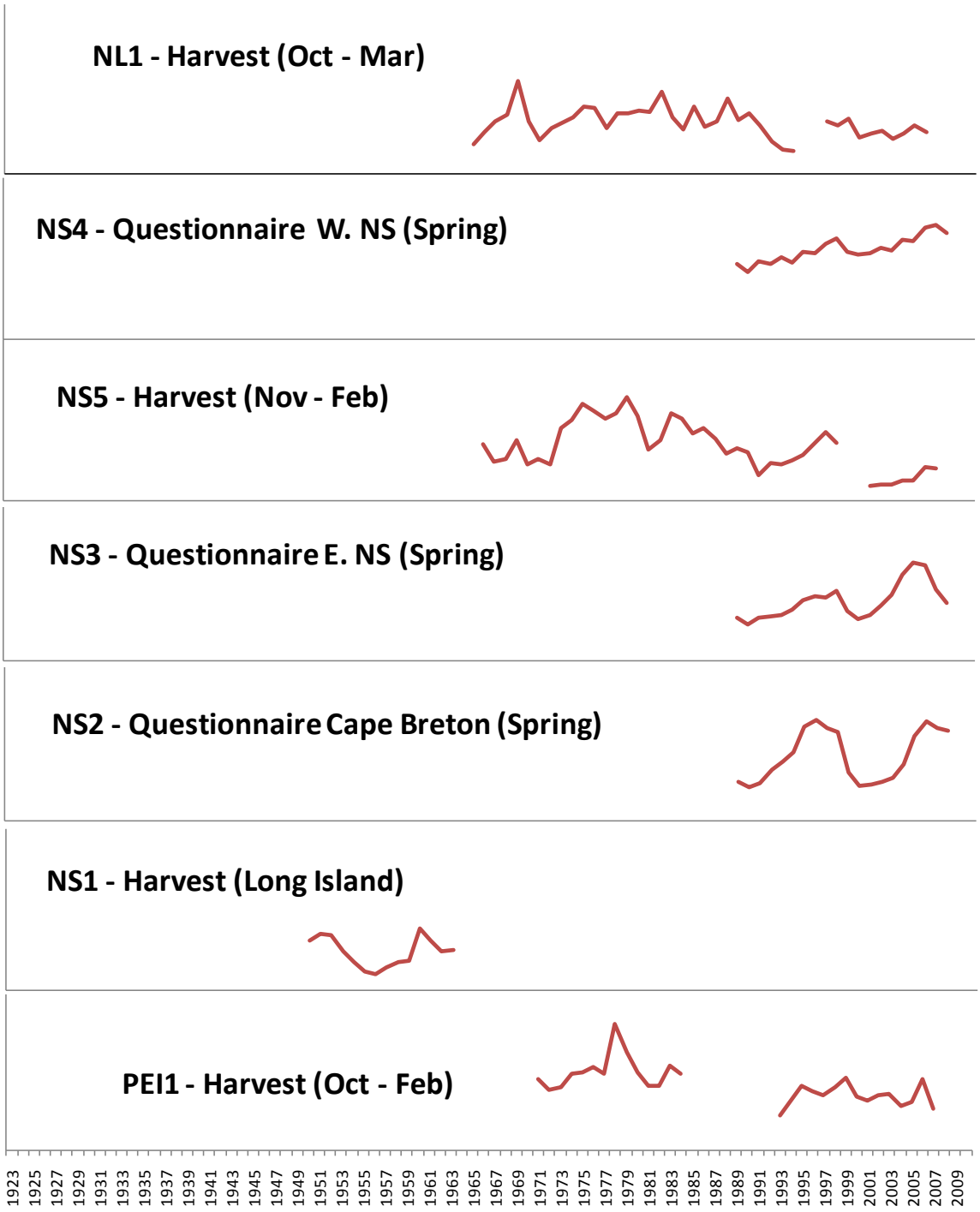

MB1 - Capture Long Point (Sept)
MN1 - Harvest (Sept - Feb)
MN2 - Sightings N. MN (Spring)
MN3 - Pellets N. Central MN (April - May)
1923
1925
1927
1929
1931
1933
1935
1937
1939
1941
1943
1945
1947
1949
1951
1953
1955
1957
1959
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009


WI1 - Tracks N. WI (Nov - Dec)
WI2 - Harvest (Year-round)
IR1 - Sightings (Summer)
MI3 - Harvest Upper Penin. (Sept - Mar)
MI1 - Harvest N. Lower Penin. (Sept - Mar)
MI2 - Harvest S. Lower Penin. (Sept - Mar)
1923
1925
1927
1929
1931
1933
1935
1937
1939
1941
1943
1945
1947
1949
1951
1953
1955
1957
1959
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009


ON1 - Pellets S.E. ON (Aug - Oct)
PA1 - Harvest (Dec)
NY1 - Harvest (Oct - Mar)
CT1 - Harvest (Nov - Jan)
ME1 - Harvest (Oct - Mar)
1923
1925
1927
1929
1931
1933
1935
1937
1939
1941
1943
1945
1947
1949
1951
1953
1955
1957
1959
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009


QC1- Harvest Bas St. Laurent (Sept - Mar)
QC5 - Harvest Lanaudiere (Sept - Mar)
QC6 - Harvest Mauricie (Sept - Mar)
QC2 - Harvest Capitale Nationale (Sept - Mar)
QC7 - Harvest Saguenay Lac St. Jean (Sept - Mar)
1923 1925 1927 1929 1931 1933 1935 1937 1939 1941 1943 1945 1947 1949 1951 1953 1955 1957 1959 1961 1963 1965 1967 1969 1971 1973 1975 1977 1979 1981 1983 1985 1987 1989 1991 1993 1995 1997 1999 2001 2003 2005 2007 2009


QC8 - Harvest Outaouais (Sept - Mar)
QC3 - Harvest Cote Nord (Sept - Mar)
QC4- Harvest Gaspesie (Sept - Mar)
1923
1925
1927
1929
1931
1933
1935
1937
1939
1941
1943
1945
1947
1949
1951
1953
1955
1957
1959
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009


NL1 - Harvest (Oct - Mar)
NS4 - Questionnaire W. NS (Spring)
NS5 - Harvest (Nov - Feb)
NS3 - Questionnaire E. NS (Spring)
NS2 - Questionnaire Cape Breton (Spring)
NS1 - Harvest (Long Island)
PEI1 - Harvest (Oct - Feb)
1923 1925 1927 1929 1931 1933 1935 1937 1939 1941 1943 1945 1947 1949 1951 1953 1955 1957 1959 1961 1963 1965 1967 1969 1971 1973 1975 1977 1979 1981 1983 1985 1987 1989 1991 1993 1995 1997 1999 2001 2003 2005 2007 2009

Population	State / Province / Territory	Data Source	Data Type	Collection Months	Beginning Year	Ending Year	Total Years
AB1	Alberta	L. Keith	Capture	Jan - Dec	1962	1975	14
AK1	Alaska	K. Kielland	Pellets	June	1985	2007	23
AK2	Alaska	C. McIntyre	Sightings	April - June	1988	2007	20
AK3	Alaska	C. Mitchell	Pellets	June	1991	2008	18
AK4	Alaska	C. Mitchell	Pellets	June	1991	2008	18
AK5	Alaska	J. Morton	Pellets	June - Aug.	1983	2010	28
CT1	Connecticut	W. Sondrini	Harvest	Nov -Jan	1923	1955	33
IR1	Isle Royale, Michigan	R. Peterson	Sightings	Summer	1974	2010	36
MB1	Manitoba	W. Koonz	Capture	Sept	1971	1990	19
ME1	Maine	W. Jakubas	Harvest	Oct - Mar	1955	1983	29
MI1	Michigan	G. Karasek	Harvest	Sept - Mar	1954	2007	52
MI2	Michigan	G. Karasek	Harvest	Sept - Mar	1954	2007	52
MI3	Michigan	G. Karasek	Harvest	Sept - Mar	1954	2007	52
MN1	Minnesota	M. Dexter	Harvest	Sept - Feb	1941	2009	69
MN2	Minnesota	J. Erb	Sightings	Spring	1974	2010	37
MN3	Minnesota	T. Fuller	Pellets	April - May	1969	1984	16
MT1	Montana	B. Giddings	Tracks	Jan - April	1992	2007	16
MT2	Montana	B. Giddings	Tracks	Jan - April	1990	2007	17
MT3	Montana	B. Giddings	Tracks	Jan - April	1990	2007	18
MT4	Montana	Anonymous	Pellets	May	1989	2008	20
NL1	Newfoundland	M. McGrath	Harvest	Oct - Mar	1965	2006	40
NS1	Nova Scotia	D. Dodds	Harvest/Area	Nov - Feb	1950	1963	14
NS2	Nova Scotia	M. O'Brien	Questionnaire		1989	2008	20
NS3	Nova Scotia	M. O'Brien	Questionnaire		1989	2008	20
NS4	Nova Scotia	M. O'Brien	Questionnaire		1989	2008	20
NS5	Nova Scotia	M. O'Brien	Harvest	Nov - Feb	1966	2007	40
NWT1	Northwest Territories	S. Carriere	Pellets	June	1988	2002	15
NWT2	Northwest Territories	S. Carriere	Pellets	June	1988	2007	20
NWT3	Northwest Territories	S. Carriere	Pellets	June	1988	2007	20
NWT4	Northwest Territories	S. Carriere	Pellets	June	1988	2007	18
NY1	New York	A. Jacobson	Harvest	Oct - Mar	1958	2007	37
ON1	Ontario	J. Bendell	Pellets	Aug - Oct	1986	2007	22
PA1	Pennsylvania	D. Diefenbach	Harvest	Dec	1930	2009	75
PEI1	Prince Edward Island	R. Dibblee	Harvest/Hunter	Oct - Feb	1971	2007	29
QC1	Quebec	H. Jolicoeur	Harvest/Hunter	Sept - Mar	1971	2004	34
QC2	Quebec	H. Jolicoeur	Harvest/Hunter	Sept - Mar	1980	2004	25
QC3	Quebec	H. Jolicoeur	Harvest/Area	Sept - Mar	1979	2004	26
QC4	Quebec	H. Jolicoeur	Harvest/Hunter	Sept - Mar	1984	2004	21
QC5	Quebec	H. Jolicoeur	Harvest/Hunter	Sept - Mar	1976	2003	26
QC6	Quebec	H. Jolicoeur	Harvest/Hunter	Sept - Mar	1978	2004	27
QC7	Quebec	H. Jolicoeur	Harvest/Hunter	Sept - Mar	1982	2004	23
QC8	Quebec	M. St-Laurent	Harvest/Hunter	Sept - Mar	1971	1996	26
SK1	Saskatchewan	A. Arsenault	Harvest	Winter	1940	1979	30
UT1	Utah	D. Mitchell	Harvest/Hunter	Sept - Feb	1975	2009	35
WI1	Wisconsin	B. Dhuey	Tracks	Nov - Dec	1977	2009	33
WI2	Wisconsin	B. Dhuey	Harvest	Year-round	1931	2009	69
WY1	Wyoming	R. Schilowsky	Harvest	Sept - Feb	1974	2008	35
YK1	Yukon	C. Krebs	Capture	Spring & Fall	1977	2008	32
YK2	Yukon	C. Krebs	Capture	Spring & Fall	1986	2010	25

Appendix 4.4 Summary of snowshoe hare time series analyzed in this study.

POPULATIONS	DATA TYPES	NUMBER OF OVERLAPPING YEARS	RHO
WY1 - WY2	Harvest - Harvest/Hunter	35	0.77
WI1 - WI2	Track[A] - Harvest	33	0.43
WI2 - WI3	Harvest - Harvest/Hunter	27	0.94
UT1 - UT3	Harvest - Harvest/Hunter	32	0.91
UT1 - UT2	Harvest - Harvest/Hunter-Day	32	0.87
QC6 - QC9	Harvest/100 Hunter-Days - Harvest/100km^2	27	0.95
QC6 - QC10	Harvest/100 Hunter-Days - Harvest	27	0.95
PA1 - PA2	Harvest - Harvest/Hunter	28	0.79
NY1 - NY2	Harvest - Harvest/Hunter	37	0.82
NS2 - NS5	Harvest / Questionnaire[B]	20	0.44
NS5 - NS6	Harvest - Harvest/Hunter	28	0.76
MN1 - MN3	Pellets[C] / Harvest	16	0
MN1 - MN4	Harvest - Harvest/Hunter	19	0.87
MN1 - MN2	Harvest / Sightings[D]	37	0.38

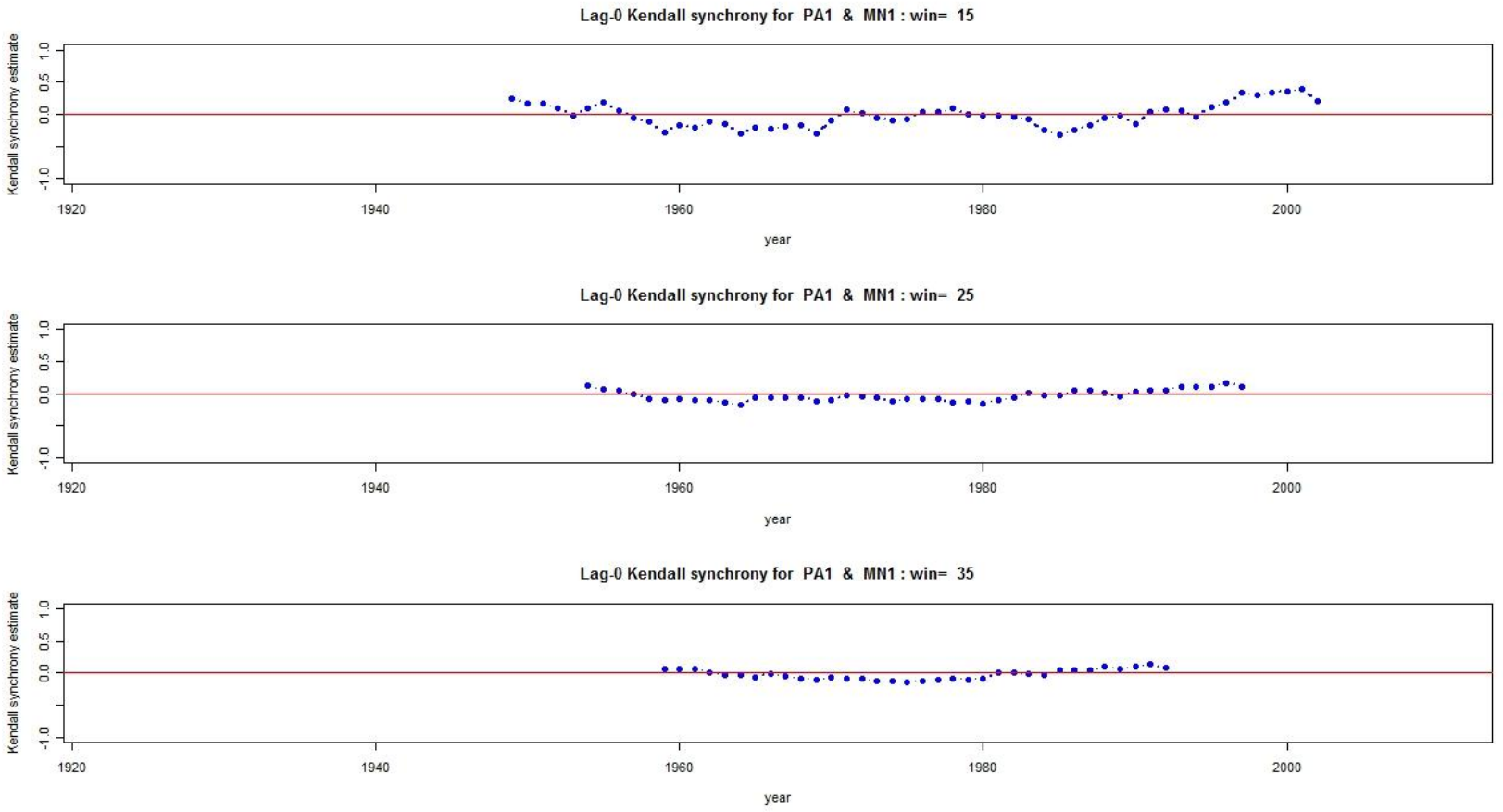
[A] Track data covered the northern 1/3 of Wisconsin; Harvest data covered Wisconsin

[B] Questionnaires covered Cape Breton (a separate island in the eastern 1/4 of Nova Scotia); Harvest data covered Nova Scotia

[C] Pellet data covered 839 km^2 in northeastern Itasca County, Minnesota; Harvest data covered Minnesota

[D] Sightings covered grouse lek sites distributed throughout ~2/3 of eastern and northern Minnesota; Harvest data covered Minnesota

Appendix 4.5
Pearson's correlation (r) between pairwise Kendall synchrony estimates for hare time series data collected using two different survey methods with overlapping, but not necessarily identical, geographic areas. Each pair of survey methods compared is listed in Column 1. Methods covering different geographic areas are superscripted with a letter.



Appendix 4.6
Example of shifting window analysis to examine if synchrony between two time series (PA1 and MN1) exhibits an increasing or decreasing long-term trend. Each point represents the synchrony estimate for the pair of time series calculated on 15, 25, or 35 (window sizes) of time series data centered on the year.

Appendix 4.7
R code for the main components of this analysis.

```
########################################################################
#
# CHAPTER 4--SNOWSHOE HARE SYNCHRONY ANALYSIS
# APPENDIX R CODE
#
# NOTE:
# 1) Code is presented in this order: 1) code for the calculating Kendall synchrony, 2) graphing and analysis functions.
# 2) Data for this analysis should be read in as a .csv file with these columns:
#    - TS_ID: Provides a unique ID for each time series, for identifying time series in graphing functions
#    - Genetic: "B" (Boreal), "P" (Pacific NW), or "R" (Rockies) to designate which genetic group each time series falls within
#    - Climatic: "C" (Continental), "P" (Pacific), or "A" (Atlantic) to designate which climatic group each time series falls within,
#         as definedby Stenseth et al. 1999
#    - Ecoregion: "E" (Eastern Temperate Forests), "M" (Marine West Coast Forests), "N" (Northern Forests), "NW"
#         (Northwestern Forested Mountains), or "T" (Taiga) to designate which of five ecoregions each time series falls within,
#         as defined by CEC 1997.
#    - NorthSouth: "N" (North) or "S" (South) to designate each time series as above or below the 49th parallel
#    - DataType: "H" (Harvest), "C" (Capture), "P" (Pellets), "S" (Sightings), "Q" (Questionnaire), or "T" (Tracks) to identify
#         sampling method for each time series
#    - Remaining columns store the abundance estimate for each year, so column headings should indicate year (e.g., 1923,
#         1924... 2010)
#
########################################################################

#>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>> CODE FOR CALCULATING KENDALL SYNCHRONY
```

```
#########################################################################
#
# INITIALIZE VARIABLES
#
#########################################################################

# Variables corresponding to data
TS.yr.col.start<-13                                    # Column in which TS year data begin
TS.start.yr<-1923                                      # First year in time series data (e.g., 1923)
TS.end.yr<-2010                                        # Last year in time series data (e.g., 2010)
TS.N<-nrow(TS.dat)                                     # Number of time series

# Other variables
min.overlap<-10                                        # Minimum number of overlap years for calculating synchrony

Yr.vector<-c(TS.start.yr:TS.end.yr)                    # Create vector for x-axis of graphs

TS.dat.only<-data.matrix(TS.dat[,TS.yr.col.start:dim(TS.dat)[2]])      # Working submatrix of time series data only
rownames(TS.dat.only)<-TS.dat[,"Source"]

num.yrs<-ncol(TS.dat.only)                             # Maximum time series length

TS.ln<-matrix(NA,TS.N,num.yrs)                         # Matrix of log-transformed time series
rownames(TS.ln)<-TS.dat[,"TS_ID"]
colnames(TS.ln)<-colnames(TS.dat.only)

TS.diff.ln<-matrix(NA,TS.N,num.yrs)                    # Matrix of first difference logged time series
```

```
rownames(TS.diff.ln)<-TS.dat[,"TS_ID"]
colnames(TS.diff.ln)<-colnames(TS.dat.only)

Sync.Kendall.est<-matrix(NA,TS.N,TS.N)                    # Matrix of pairwise Kendall synchrony
rownames(Sync.Kendall.est)<-TS.dat[,"TS_ID"]
colnames(Sync.Kendall.est)<-TS.dat[,"TS_ID"]

overlap<-matrix(NA,TS.N,TS.N)                             # Tracks number of overlapping years for each pair of time series
rownames(overlap)<-TS.dat[,"TS_ID"]
colnames(overlap)<-TS.dat[,"TS_ID"]


#########################################################################
#
# BEGIN SIMULATION
#
#########################################################################

TS.ln<-log(TS.dat.only)                                   # Natural log of time series data
TS.diff.ln[,2:num.yrs]<-t(apply(TS.ln,1,diff))
        # Calculate difference between log values of consecutive years and assign result to later year

# Calculate Kendall synchrony
for(m in 1:TS.N)                                          # For each time series, m
        {
        m.vec<-TS.diff.ln[m,1:num.yrs]                    # Assign first differenced log time series

        for(n in 1:TS.N)                                  # For each time series, n
```

```
                {
                n.vec<-TS.diff.ln[n,1:num.yrs]                          # Assign first differenced log time series

                overlap[m,n]<-sum(!is.na(m.vec)&!is.na(n.vec))          # Calculate number of overlapping years for each time series pair

                # Must meet minimum overlap criterion, and don't calculate synchrony of a time series on itself
                if ((overlap[m,n]>=min.overlap)&&(m!=n))
                        {
                        Sync.Kendall.est[m,n]<-cor(m.vec,n.vec,method="kendall",use="na.or.complete")
                        }                       # END IF-OVERLAP
                }                               # END FOR-N
        }                                       # END FOR-M

save.image(file=paste("SSHsync_",substr(as.character(Sys.time()),1,10),".RData",sep="")) # save workspace--name includes date


#>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>> CODE FOR GRAPHING AND ANALYSIS FUNCTIONS


#####################################################################
#
# FUNC.MODCOR:  FUNCTION TO CALCULATE AND PLOT MODIFIED CORRELOGRAM
#
# Input
#   [sub.mat]    "N", "S" (for northern or southern hares)
#       [num.dist]      Number of distance categories to use (try multiple, to see if qualitative results remain same)
#       [reps]            Number of sample sets to generate distribution of means within each distance class
#
```

```
#########################################################################

Func.modcor<-function(sub.mat="N",num.dist=4,reps=100)
	{
	library(gplots)

	sub.vec<-switch(sub.mat,
		N = which(TS.dat$NorthSouth=="N"),
		S = which(TS.dat$NorthSouth=="S"))

	##### INITIALIZE MATRICES & ARRAYS
	sync.ary<-matrix(NA,length(sub.vec),length(sub.vec))		# Array of base synchrony estimates

	samp.ary<-array(NA,dim=c(num.dist,reps,length(sub.vec)))
		# Array to hold sampled synchrony estimates for each distance class

	dimnames(samp.ary)<-list(paste("dist.",1:num.dist),paste("draw.",1:reps),1:length(sub.vec))

	rep.means<-matrix(NA,num.dist,reps)				# Store mean synchrony for each rep for each distance class

	dist.temp<-matrix(NA,TS.N,TS.N)					# Store geographic distances for each pair of time series
	dist.mat<-matrix(NA,length(sub.vec),length(sub.vec))

	dist.cat.mat<-matrix(NA,length(sub.vec),length(sub.vec))	# Store distance categories

	summary.out<-matrix(NA,num.dist,4)				# Store summary data from modified correlogram analysis
	rownames(summary.out)<-c(paste("dist.class.",1:num.dist))
	colnames(summary.out)<-c("mean","se","p.val","avg.N")
```

```
sync.ary<-Sync.Kendall.est[sub.vec,sub.vec]                    # Subset synchrony matrix for northern or southern hares

dist.temp<-read.table("SSH_geodist_fullmat.txt",header=TRUE)
        # Read in matrix of geographic distance (km) for each pair of time series
dist.mat<-dist.temp[sub.vec,sub.vec]
        # Subset geographic distance matrix for northern or southern hares
dist.mat[lower.tri(dist.mat,diag=TRUE)]<-NA                    # Create upper triangular distance matrix

max.dist<-5800
        # Maximum distance (km) between time series, so x-axis is on same scale for northern and southern hares

##### PLOT HISTOGRAM OF GEOGRAPHIC DISTANCES
x<-as.matrix(dist.mat)                                         # Convert to matrix
y<-as.vector(x)                                                # Then convert matrix to vector, for histogram function
hist(y,breaks=seq(0,6000,by=250),main="Pairwise geographic distances for time series")     # Plot histogram

##### CALCULATE MODIFIED CORRELOGRAM
# Divide synchrony estimates into distance-based categories
dist.space<-max.dist/num.dist                                  # Create equal-spaced distance categories
dist.thresh<-c((0:(num.dist-1))*dist.space)                    # Lower limit of each distance category

# Assign distance category to each time series pair, 1:num.dist
for(k in 1:num.dist)
        {
        if(k==num.dist)
                {dist.cat.mat[which(dist.mat>=dist.thresh[k])]<-k}      # Assign last distance category differently
         else
```

```
                {dist.cat.mat[which((dist.mat>=dist.thresh[k])&(dist.mat<dist.thresh[k+1]))]<-k}
        }                               # END FOR-K

    # For each distance category, randomly select synchrony estimates without duplicating any time series
    # Repeat to get distribution of synchrony estimates

     for (k in 1:num.dist)                                      # For each distance class
        {
        dist.ind<-which(dist.cat.mat==k,arr.ind=TRUE)           # Identify row & column of each time series pair in this
distance category
        dist.ind.len<-1:(length(dist.ind)/2)                    # Vector of original index values to sample from
        ind.track<-matrix(NA,reps,length(dist.ind.len))         # Track indices sampled to remove replicates
        r<-1                                                    # Initialize reps

        while (r<=reps)
            {
            used.vals<-rep(NA,length(sub.vec))          # Vector of time series that can not be further sampled from
            ind.omit<-rep(NA,length(dist.ind.len))          # Vector of indices that can not be further sampled from
            count<-1                                        # Track number of samples
            ind.vec<-rep(NA,length(dist.ind.len))           # Track indices sampled
            ind.remain<-rep(NA,length(dist.ind.len))        # Vector of indices that can still be sampled from

            while(sum(!is.na(ind.omit))<length(dist.ind.len))   # While there are still values that can be sampled
                {
                # Remaining distance indices that can still be sampled from
                if(count==1){ind.remain<-as.vector(dist.ind.len)}else{ind.remain<-as.vector(dist.ind.len[-ind.omit])}

                # Sample one time series pair (can only sample from vectors)
```

```
            if(length(ind.remain)>1){a<-sample(ind.remain,1)}else{a<-ind.remain}
            ind.vec[count]<-a
            # Assign corresponding synchrony values to sample array
            samp.ary[k,r,count]<-sync.ary[dist.ind[a,"row"],dist.ind[a,"col"]]
            # Send row and column of sampled synchrony value to "used" vector
             used.vals[min(which(is.na(used.vals))):(min(which(is.na(used.vals)))+1)]<-as.vector(c(dist.ind[a,]))
             # Identify which distance index values can no longer be sampled from
             ind.omit<-which(dist.ind[,"row"]%in%used.vals|dist.ind[,"col"]%in%used.vals)
            count<-count+1
             }                          # END WHILE-SUM--DONE WITH THIS DRAW
      dup<-FALSE                                  # Set duplicates tally to false
      for(i in 1:(r-1))                           # Check for duplicate sample sets in previous drawings
             {
             if(identical(sort(ind.vec,na.last=TRUE),ind.track[i,]))      # tally up duplicates
                    {dup<-TRUE}                                           # END IF-IDENTICAL
             }                                                            # END FOR-I
        if(!isTRUE(dup))                                                  # If no duplicate sample set exists...
             {
             ind.track[r,]<-sort(ind.vec,na.last=TRUE)                    # Update ind.track with sample set...
             r<-r+1                                                       # and advance r to next drawing
             }                          # END IF-ISTRUE
        else                            # otherwise just advance to next drawing, but don't include this set in results
             {
             samp.ary[k,r,]<-NA
             r<-r+1
             }                          # END IF-ISTRUE-ELSE
        }                               # END WHILE-R-REPS--FINISH ALL DRAWINGS FOR THAT DISTANCE CLASS
summary.out[k,4]<-sum(!is.na(samp.ary[k,,]))/sum(!is.na(ind.track[,1]))
```

```
            }                                   # END FOR-K

        # For each distance category, calculate summary statistics
        rep.means<-apply(samp.ary,c(1,2),mean,na.rm=TRUE)
        summary.out[,1]<-apply(rep.means,1,mean,na.rm=TRUE)                 # Calculate means across reps
        summary.out[,2]<-apply(rep.means,1,sd,na.rm=TRUE)                   # Calculate standard errors
        summary.out[,3]<-1-pnorm(abs(summary.out[,1]/summary.out[,2]))      # Calculate p-values

        # Graph results
        plotCI(x=1:num.dist,y=dist.lag.means[,1],uiw=(1.96*dist.lag.means[,2]),liw=(1.96*dist.lag.means[,2]),err='y',ylim=c(-1,1),
                xlim=c(1,num.dist),ylab=paste(estim," synchrony"),xlab="distance class",main="Modified correlogram for Kendall
                Lag0")
        abline(h=0,col='red')
        }                                                   # END FUNC.MODCOR


#########################################################################
#
#  FUNC.TEMPORAL:  FUNCTION TO PLOT SHIFTING WINDOW ANALYSIS RESULTS
#
#       Input
#               [x.source, y.source]        Two TS_ID to compare
#               [win.vec]                   Vector of 3 window sizes
#
#########################################################################

Func.temporal<-function(x.source,y.source,win.vec)
        {
```

```
##### INITIALIZE MATRICES & ARRAYS
temp.growth.est<-array(NA,dim=c(11,11,length(win.vec),num.yrs))
                # Kendall matrix of synchrony estimate for each window and year
dimnames(temp.growth.est)<-list(TS.dat[1:11,"Source"],TS.dat[1:11,"Source"],win.vec,Yr.vector)

x.num<-which(TS.dat["TS_ID"]==x.source
y.num<-which(TS.dat["TS_ID"]==x.source

##### CALCULATE GROWTH SYNCHRONY FOR EACH 'WINDOW' #####
for (win in win.vec)                                # For each window size
        {
        mid.yr<-floor(win/2+0.5)-1
                # Synchrony estimate will be assigned to mid-yr of window. Subtracted one so can just add
                # mid.yr to starting year to find middle point
        # Initialize variables
        i.vec<-rep(NA,num.yrs)
        j.vec<-rep(NA,num.yrs)

        # Assign variables
        i.vec<-TS.dat.only[x.num,]
        j.vec<-TS.dat.only[y.num,]

        # Next step is to find the actual overlap range
        joint.start<-min(which(!is.na(i.vec)&!is.na(j.vec)))
        joint.end<-max(which(!is.na(i.vec)&!is.na(j.vec)))

        for (z in joint.start:(joint.end-win+1))          # All the TS-i starting points to use
                {
```

```
            # Initialize variables
            i<-rep(NA,win)                              # vector of specified window size
            j<-rep(NA,win)
            ij<-matrix(NA,2,win)
            ij.ln<-matrix(NA,2,win)
            ij.diff.ln<-matrix(NA,2,win)
            ij.a<-matrix(NA,2,win)

            # Identify the time series subset
            i<-i.vec[z:(z+win-1)]
            j<-j.vec[z:(z+win-1)]
            ij<-matrix(c(i,j),nrow=2,byrow=TRUE)

            # Calculate Kendall synchrony
            ij.ln<-log(ij)                              # Natural log of time series data
            ij.diff.ln[,2:ncol(ij.ln)]<-t(apply(ij.ln,1,diff))  # Calculates difference btwn ln values of consecutive years

            temp.growth.overlap<-sum(!is.na(ij.diff.ln[1,])&!is.na(ij.diff.ln[2,]))
            if(temp.growth.overlap>=10)                 # If there are at least 10 overlapping datapoints
                  {
                  temp.growth.est[c,d,which(win.vec==win),(z+mid.yr)]<-
                        cor(ij.diff.ln[1,],ij.diff.ln[2,],method="kendall",use="na.or.complete")
                  }                                     # END IF-TEMP.GROWTH.OVERLAP
            }                                           # END FOR-Z
      }                                                 # END FOR-WIN

##### PLOT RESULTS
par(mfrow=c(3,1))                           # Display three graphs on same page, one for each window size
```

```
        for(w in 1:3)                                           # For each window size
                {
                plot(temp.growth.est[w,]~Yr.vector,lwd=2.5,xlab='year',main=paste("Kendall synchrony for ",
                        TS.dat[x.num,"TS_ID"]," & ",TS.dat[y.num,"TS_ID"],": win= ",win.vec[w]),ylab='Kendall synchrony estimate',
                        type='b',pch=16,col="blue",ylim=c(-1,1))
                }                                               # END FOR-W
        }                                                       # END FUNC.TEMPORAL


#########################################################################
#
# FUNC.MANTEL:  FUNCTION TO CALCULATE PARTIAL MANTEL TEST
#
#########################################################################

Func.mantel<-function()
        {
        library(ncf)

        ##### INITIALIZE VARIABLES
        A.ary<-matrix(NA,TS.N,TS.N)
        A.ary<-Sync.Kendall.est
        A.ary[upper.tri(A.ary,diag=TRUE)]<-NA                   # Lower triangle matrix with Kendall synchrony estimates

        A.Geo.dist<-matrix(NA,TS.N,TS.N)
        A.Geo.dist<-as.matrix(read.table("SSH_geodist_fullmat.txt",header=TRUE))
        A.Geo.dist[upper.tri(A.Geo.dist)]<-NA                   # Lower triangle geographic distance matrix
```

```
A.cat<-matrix(NA,TS.N,TS.N)
x<-TS.dat[,"XXX"]       # REPLACE 'XXX' WITH COLUMN NAME FOR THE CATEGORY ANALYSIS WILL BE BASED ON

##### ASSIGN CATEGORIES TO SYNCHRONY ESTIMATES--WITHIN OR BETWEEN GROUPS
for (i in 1:TS.N)
        {
         for (j in 1:TS.N)
                {
                A.cat[i,j]<-ifelse(x[i]==x[j],1,0)     # '1' if time series are in same group, '0' if time series are in different groups
                }       # END FOR-J
        }               # END FOR-I

A.cat[is.na(A.ary)]<-NA                 # If there is no corresponding synchrony estimate, change category value to NA
A.Geo.dist[is.na(A.ary)]<-NA            # Do same for distance category

partial.mantel.test(A.ary,A.Geo.dist,A.cat)
}                       # END FUNC.MANTEL
```