ECOSYSTEM ORGANIZATION AND TRANSFORMATION: THE ROLE OF NETWORK ARCHITECTURE IN THE DEVELOPMENT OF ENVIRON INDIRECT EFFECTS

by

STUART ROSS BORRETT

(Under the Direction of Bernard C. Patten)

ABSTRACT

Indirect effects appear to play a significant role in the organization and transformation of complex adaptive systems (CAS). In ecosystems, one type of CAS, organisms and their environments are coupled by an intricate network of energy, matter, and information exchanges. This makes it possible for one species to affect the distribution, abundance, and behavior of other species without direct contact. Because we do not fully comprehend factors controlling the development and propagation of indirect effects through the interaction networks of CAS, we cannot predict their importance *a priori*.

The overarching goal of the research presented in this dissertation is to develop a better understanding of processes that create, constrain, and sustain ecological systems. The research focuses on the role of ecosystem architecture (structure and function) in the development and propagation of environ indirect effects, which result from energy–matter transactions. In Chapters 2 and 3 we distinguish the role of biodiversity and the number and pattern of interactions in creating indirect pathways. In the process, we reveal a form of hierarchical organization based on strongly connected components. In Chapters 4 and 5, we examine how the distribution of energy–matter flux across a fixed structure influences the magnitude of environ indirect effects in models of nitrogen in the Neuse River Estuary and phosphorus in Lake Sidney Lanier, respectively. Finally, in Chapter 6, we assess the relative role of structure and function in the development of environ indirect effects in 20 ecosystem models of energy flow. We find that while certain structural elements are important (e.g., cycles) they are insufficient to determine the magnitude of environ indirect effects; the distribution of boundary inputs and internal flows is critical.

This research adds to our growing understanding of the causes and consequences of indirect effects in ecosystem organization and transformation. It characterizes new aspects of the role ecosystem architecture plays in the development and propagation of indirect effects, and expands the methodology of Network Environ Analysis, an environmental extension of economic Input-Output Analysis. These fundamental developments will provide a firmer foundation for critical environmental management concepts including ecosystem health, integrity, and sustainability.

INDEX WORDS: Network Environ Analysis; indirect effects; network analysis; ecological modeling; pathway proliferation; strongly connected components; food web; cycles; structure; flow analysis; ecosystem; indirect flow index; systems ecology

ECOSYSTEM ORGANIZATION AND TRANSFORMATION: THE ROLE OF NETWORK ARCHITECTURE IN THE DEVELOPMENT OF ENVIRON INDIRECT EFFECTS

by

STUART ROSS BORRETT

B.A., Austin College, 1997

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

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2005

ECOSYSTEM ORGANIZATION AND TRANSFORMATION: THE ROLE OF NETWORK ARCHITECTURE IN THE DEVELOPMENT OF ENVIRON INDIRECT EFFECTS

by

STUART ROSS BORRETT

Major Professor:	Bernard C. Patten
Committee:	M. Bruce Beck David K. Gattie H. Ronald Pulliam Peter G. Verity

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2005

DEDICATION

To Stephanie Lee Wert Borrett and Caitlin McGuire Borrett

ACKNOWLEDGEMENTS

Many people have been integral to the research presented here. This work would not have been possible without the support and encouragement of Dr. B.C. Patten, who took a chance on an aspiring graduate student that knew nothing about systems ecology, ecological modeling, or even the mathematics required. Further, this research builds on the previous work of Dr. Patten and his many collaborators that have constructed the Holoecology Research Program. I also want to thank members of my advisory committee, past and present, including Dr. M.B. Beck, Dr. G.M. Cowie, Dr. D.K. Gattie, Dr. J.L. Meyer, Dr. H.R. Pulliam, Dr. R.G. Wiegert, and Dr. P.G. Verity. In addition, I am grateful for the efforts of my chapter collaborators including Dr. B.D. Fath, Dr. S.J. Whipple, Dr. R.R. Christian, Dr. O.O. Osidele.

Graduate students are a core strength of the Institute of Ecology. I am deeply indebted to my colleagues who have joined me in an exploration of the natural world, ecology, and science in general. I would especially like to thank Mitch Pavao-Zuckerman and Stephanie Madson for their continued support and collaborations. The Systems Ecology Laboratory, now known as the Systems Ecology and Engineering group, have supported and critiqued much of the work that appears in this dissertation – as well as many elements that were left out. I value discussions with David Gattie, Stuart Whipple, John Schramski, Seth Bata, Brian Fath, and Jill Goldstein.

I am very appreciative of the love and support of my family. Their assistance has been critical. My wife Stephanie Wert Borrett spent many hours reading and editing my class papers, manuscripts and presentations, asking me tough questions that cut to the heart of the research, and tolerating my incessant working, all while maintaining a full-time job, completing a masters degree and keeping the household going. Our daughter Caitlin is a recent arrival, but she has provided motivation to finish and a temperament that made it possible. Linda Wert's support over the last 6 months has been indispensable: from caring for Caitlin to delivering food while I was sequestered to write. Finally, I am indebted to my parents Doug and Sally Borrett who wouldn't let me get away without learning to read (despite my young protests), who taught me be to be curious and ask questions, and consistently supported my intellectual, emotional, and physical development.

Thank you all.

TABLE OF CONTENTS

Page

APPENDIX

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

“…the goodness and the excellence of the piper or sculptor, or the practiser of any art, and generally those who have any function or business to do, lies in that function…,” Aristotle in Nichomachean Ethics

“Form follows function,” Louis H. Sullivan

“Form and function are one,” Frank Lloyd Wright

1.1 FORM AND FUNCTION

Understanding the dialectic between form and function is a classic problem that pervades the arts and sciences. Form is defined as “the shape and structure of something as distinguished from its material” (Merriam-Webster Online Dictionary, 2005a); function is “the action for which a person or thing is specially fitted or used or for which a thing exists” (Merriam-Webster Online Dictionary, 2005b). Plato emphasized the value of the true (but he believed not empirically observable) underlying form of an object or idea (Kraut, 2004), while Aristotle claimed the highest value derived from an entity’s function (Ferris, 1988; Kraut, 2001). The significance of form and function continues to challenge our disciplinary studies today (Krakauer, 2003). Archeologists divine the function of unearthed human relics by studying their structure and associations. Architecture embodies both form and function, although the opening quotes by Wright and his mentor Sullivan demonstrate that architects do not necessarily agree on which element takes precedence. Wright suggests the distinction between form and function is artificial; the two are convolved and co-implicating. Biology is steeped in this idea. Biologists routinely study the interplay of living forms and their functions. For example, the anatomical forms of vertebrate bones (e.g., their shapes and sizes) are routinely interpreted by how they relate to where muscles and connective tissues attach and how this enables organism movement.

DNA sequencing is the first step toward uncovering structural patterns expected to reveal the structure and ultimately the function of genes (Collins and Patrinos, 1998).

Ecologists have been investigating the processes that create and maintain ecosystem structure and function for over a quarter century (e.g., Tansley, 1935; Lindeman, 1942; Odum, 1953; Patten, 1959; Teal, 1959; Odum, 1960; Odum, 1962, 1969). Ecosystems are open, dissipative, far-from-equilibrium thermodynamic systems composed of biotic and abiotic elements transferring and transforming energy, matter, and information. Ecosystems receive high-exergy flows of energy and matter at their boundaries, perform organizational work of aggradation (departure from equilibrium), and produce heat and degraded byproducts which are dissipated to the environment (Jørgensen et al., 1992; Patten, 1998a; Jørgensen et al., 2000). Like other kinds of complex adaptive hierarchical systems (Levin, 1998; Patten et al., 2002), ecosystems appear to self-organize in response to thermodynamic gradients (Schneider and Kay, 1994; Capra, 1996; Müller, 1996, 1998; Patten, 1998b).

Beyond general curiosity, ecosystem research is motivated by a desire to translate this knowledge into sound environmental management (Christensen, 1996; Reichman and Pulliam, 1996); ecosystems provide the natural capital and services that support human endeavors (Costanza et al., 1997; Daily, 1997). Today, ecologists are grappling with concepts like ecosystem health (Rapport et al., 1998; Costanza and Mageau, 1999), ecosystem integrity (Kay, 1991; Westra and Lemons, 1995), and sustainability (Patten, 1988; Kay et al., 1999; Holling, 2001), trying to adequately define them and determine their use. The challenge, however, is that we still lack fundamental understanding regarding the causes and consequences of ecosystem organization and transformation (Reichman and Pulliam, 1996; Jørgensen, 2002). Advances have been made on this topic (e.g., Odum, 1969; Hannon, 1973; Finn, 1976; Carpenter et al., 1985; Ulanowicz, 1986, 1997; Gunderson and Holling, 2002; Jørgensen, 2002), but the problems are complex and refractory, leaving us with much uncertainty.

The overarching goal of the research presented in this dissertation is to deepen our knowledge of the causes and consequences of ecosystem organization (form and function). My aim is to improve the theoretical understanding of processes that create, constrain, and sustain ecosystems. Ultimately, I expect this work will provide a firmer foundation for concepts like ecosystem health, integrity, and sustainability. In this chapter, I introduce my conceptual framework and focal research question. I conclude with an overview of the dissertation chapters and discuss the ties that bind them.

1.2 INDIRECT EFFECTS

Indirect effects, the ability of one species or system element to influence another without directly interacting with it, appear to have a significant role in the organization and transformation of biological systems. They have long been recognized by biologists (Darwin, 1959), and the challenge they pose for environmental management helped spark the environmental movement (Carson, 1962). Though arduous to investigate, ecologists are increasingly recognizing the significance of indirect effects as components of ecological interactions (Patten, 1983; Andrewartha and Birch, 1984; Patten, 1984; Miller and Kerfoot, 1987; Strauss, 1991; Wootton, 1994, 2002) and evolution (Miller and Travis, 1996; Laland et al., 1999; Odling-Smee et al., 2003). Wootton (2002) recently argued that indirect effects are "… a fundamental cause of ecosystem complexity."

Part of the difficulty of analyzing indirect effects is that there appear to be multiple types (Miller and Kerfoot, 1987; Strauss, 1991; Wootton, 1994). Wootton (1993; 1994; 2002) initially partitioned indirect effects into two types: interaction chains and interaction modifications.

Interaction chains occur when one species impacts another by affecting a third. This occurs by linking two or more direct interactions together. For example, in the chain A → B → C, A indirectly influences C by directly influencing B. As ecologists are often interested in the abundance or density of species, these have also been termed ‘density mediated interactions’ (Abrams et al., 1996). According to Wootton (2002), several commonly studied ecological interactions arise from this type of indirect interaction, including keystone predation, trophic cascades, apparent competition, indirect mutualisms, and exploitative competition. Interaction modifications, on the other hand, occur when a species modifies the interaction between two additional species. At least two types of interaction modifications have been identified. ‘Trait–mediated indirect effects’ (Abrams et al., 1996) occur when one species changes the traits or behavior of a second that alters how it interacts with a third. For example, Pacific killifish (*Fundulus parvipinnis*) parasitized by the trematode *Euhaporchis californiensis* tend to exhibit conspicuous behavior making them more susceptible to predation by avian predators than unparasitized killifish (Lafferty and Morris, 1996). ‘Environment–mediated interaction modifications’ occur when one species changes the environmental context in which two species interact (Wootton, 2002). Ecosystem engineers, such as beavers that construct ponds providing habitat for aquatic organisms, are one example (Jones et al., 1997).

Again, organisms and elements of their abiotic environments are coupled together through an intricate network of energy-matter exchanges (Patten et al., 1976; Ulanowicz, 1986; Higashi and Burns, 1991). While observation and analysis of these transaction networks cannot identify specific mechanisms or depict all types of indirect interactions (Loehle, 1990; Wootton, 1994), they do capture many types of indirect interactions reflected in trophic dynamics and biogeochemistry (Patten, 1990; Higashi and Burns, 1991). These include some indirect interactions from each of the three broad categories identified earlier. Therefore, analysis of these flow–storage networks may illuminate the consequences of particular system organizations, particularly the indirect effects mediated by this transaction network.

Through *Network Environ Analysis* (NEA), Patten and colleagues (Patten, 1981, 1983, 1984, 1985b; Higashi and Patten, 1986, 1989; Patten et al., 1990; Patten, 1991) have begun to characterize one type of indirect effects, termed *environ indirect effects*, in ecosystem flow–storage models. One way these effects are characterized at the whole-system level in NEA is by the ratio of indirect-to-direct flows (Indirect/Direct). This ratio indicates the significance of indirect flows within a system's internal *environs*, which are within system, compartment-specific, input and output oriented environments (Patten, 1978, 1981, 1982, 1992). Previous results suggest that Indirect/Direct tends to be greater than unity in model ecosystems, implying that indirect flows are dominant (Patten, 1981, 1983, 1985b; Higashi and Patten, 1986, 1989; Patten et al., 1990; Patten, 1991; Fath, 1998; Fath, 2004). These results led Patten (in prep.) to hypothesize that the natural world is unified through indirect effects, which is one of the cardinal hypotheses of Holoecology (see also Chapter 2).

Past work indicated that both system structure and function influence the degree of indirect effects. For example, Higashi showed algebraically how an increase in system size (number of species, n), proportion of direct links (L) connected or connectivity ($C = L/n^2$), strength of feedback and direct relations, looping or storage, and cycling tend to increase Indirect/Direct (Patten et al., 1990; Patten, 1991). The relative importance of structure and function in the development of environ indirect effects, however, is unknown. This knowledge is crucial to deepening our understanding of how indirect effects bind together ecological systems and alter how systems respond to environmental impacts. This is the specific objective of this research.

1.3 NETWORK PERSPECTIVE

Throughout this dissertation I employ a network perspective (Ponstein, 1966; Patten et al., 1976; Ulanowicz, 1986; Higashi and Burns, 1991; Margalef, 1991; Albert and Barabási, 2002; Newman, 2003; Patten, in prep.). From this perspective, complex systems like ecosystems are modeled as a graph or directed graph in which agents are represented as graph nodes (n), non-directed interactions are edges (e), and directed interactions are arcs or links (L). In an ecosystem network model, species, functional groups, and abiotic compartments are nodes while energy–matter transactions and transformations between the nodes are links. The ecosystem *structure* (form) is then captured by the arrangement or topology of these nodes and interactions. Food-webs are a classic ecological network in which only one type of transactive relation – feeding – is typically identified. Ecosystem models also include information regarding the magnitude of energy–matter flow and storage, or ecosystem *function*, along each connection. Mathematically, the network model becomes a weighted digraph. This perspective is useful because it allows analysts to utilize graph, combinatoric, and linear algebra mathematical tools to holistically characterize and quantify system structure and function. The research presented in this dissertation uses and extends one type of ecosystem network analysis: Network Environ Analysis (NEA). NEA is an environmental application and extension of economic Input–Output Analysis (Leontief, 1965, 1966). It is used to holistically investigate environmental systems. It operates like a macroscope (rather than a microscope) to characterize whole-system organization by describing, quantifying and analyzing the component-level environments or environs that compose a system. Fath and Patten (1999) review the foundations of NEA. Methodological details are well described in extant literature (Barber et al., 1979; Matis and Patten, 1981; Patten and Matis, 1982; Hippe, 1983; Fath, 1998; Fath and Patten, 1999; Patten, in prep.; Patten et al., in prep.).

1.4 DISSERTATION OVERVIEW

Chapters 2-6 form the dissertation core. They were written as stand-alone publishable units, but collectively they pivot around the focal question: what is the role of ecosystem architecture (form and function) in the development of environ indirect effects? In Chapters 2 and 3, my coauthors and I focus on structure. We develop and clarify the *pathway proliferation* concept introduced by Patten and colleagues (Patten et al., 1982; Patten, 1985a), provide two new measures for the rate of pathway proliferation, and investigate factors determining these rates. In the process, we uncover the presence of multiple *strongly connected components* in food webs, which introduce a form of modularity (hierarchical subdivisions into more or less interacting subsystems) into these models. Both pathway proliferation and strongly connected components are expected to influence the development of indirect effects. In Chapter 4 and 5, we explore how flow and storage variability impact measures of indirect effects developed in NEA while model structure remains constant. Flow–storage variability in the Neuse River Estuary model of nitrogen flux, which we inspect in Chapter 4, is created by a temporal sequence of models. The variability is generated by model uncertainty due in part to limited empirical data in the model of Lake Sidney Lanier we investigate in Chapter 5. In Chapter 6, we perform a comparative NEA of twenty ecosystem models of energy flux. Both structure and function vary in these models, allowing us to assess the relative significance of the two elements on the magnitude of environ indirect effects. The final chapter of the dissertation provides a summary of key findings in the dissertation and a discussion of their larger significance. The appendices describe a MATLAB® function (NEA.m) created with Brian Fath to facilitate the application of NEA. Modifications of this function were used to perform analyses throughout this dissertation.

1.5 REFERENCES

Abrams, P.A., Menge, B.A., Mittlebach, G.G., Spiller, D., and Yodzis, P., 1996. The role of indirect effects in food webs. In: G. Polis and K.O. Winemiller (Editors), Food webs: dynamics and structure, Chapman and Hall, New York, pp 371-395.

Albert, R., and Barabási, A.-L., 2002. Statistical mechanics of complex networks. Rev. Mod. Phys., **74**:47-97.

Andrewartha, H.G., and Birch, L.C., 1984. The ecological web: more on the distribution and abundance of animals. University of Chicago Press, Chicago.

Barber, M.C., Patten, B.C., and Finn, J.T., 1979. Review and evaluation of input-output flow analysis for ecological applications. In: J.H. Matis, B.C. Patten, and G.C. White (Editors), Compartmental Analysis of Ecosystem Models, International Co-operative Publishing House, Fairland, Maryland, pp 43-72.

Capra, F., 1996. The web of life: a new scientific understanding of living systems, 1st Anchor Books edition. Anchor Books, New York.

Carpenter, S.R., Kitchell, J.F., and Hodgson, J.R., 1985. Cascading trophic interactions and lake productivity: fish predation and herbivory can regulate lake ecosystems. Bioscience, **35**:634-639.

Carson, R., 1962. Silent spring. Houghton Mifflin, Boston.

Christensen, N.L.B., Ann M.; Brown, James H.; Carpenter, Stephen; D'Antonio, Carla; Francis, Rober; Franklin, Jerry F.; MacMahon, James A.; Noss, Reed F.; Parsons, David J.; Peterson, Charles H.; Turner, Monica G.; Woodmansee, Robert G., 1996. The report of the Ecological Society of America Committee on the Scientific Basis for Ecosystem Management. Ecological Applications, **6**:665-691.

Collins, F.S., and Patrinos, A., 1998. New goals for the U.S. human genome project: 1998-2003. Science, **282**:682-689.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R.V., Paruelo, J., Raskin, R.G., Sutton, P., and van den Belt, M., 1997. The value of the world's ecosystem services and natural capital. Nature, **387**:253-260.

Costanza, R., and Mageau, M., 1999. What is a healthy ecosystem? Aquatic Ecology, **33**:105-115.

Daily, G.C., Editor. 1997. Nature's Services. Island Press, Washington, D.C.

Darwin, C., 1959. On the origin of species by means of natural selection. John Murray, London.

Fath, B.D., 1998. Network analysis: foundations, extensions, and applications of a systems theory of the environment. Ph.D. Dissertation. University of Georgia, Athens, Georgia.

Fath, B.D., 2004. Network analysis applied to large-scale cyber-ecosystems. Ecol. Model., **171**:329-337.

Fath, B.D., and Patten, B.C., 1999. Review of the foundations of network environ analysis. Ecosystems, **2**:167-179.

Ferris, T., 1988. Coming of age in the Milky Way. Morrow, New York.

Finn, J.T., 1976. Measures of ecosystem structure and function derived from analysis of flows. J. Theor. Biol., **56**:363-380.

Gunderson, L.H., and Holling, C.S., 2002. Panarchy: understanding transformations in human and natural systems. Island Press, Washington, DC.

Hannon, B., 1973. The structure of ecosystems. J. Theor. Biol., **41**:535-546.

Higashi, M., and Burns, T.P., 1991. Theoretical studies of ecosystems: the network perspective. Cambridge University Press, Cambridge.

Higashi, M., and Patten, B.C., 1986. Further aspects of the analysis of indirect effects in ecosystems. Ecol. Model., **31**:69-77.

Higashi, M., and Patten, B.C., 1989. Dominance of indirect causality in ecosystems. Am. Nat., **133**:288-302.

Hippe, P.W., 1983. Environ analysis of linear compartmental systems: the dynamic, time-invariant case. Ecol. Model., **19**:1-26.

Holling, C.S., 2001. Understanding the complexity of economic, ecological, and social systems. Ecosystems, **4**:390-405.

Jones, C.G., Lawton, J.H., and Shachak, M., 1997. Positive and negative effects of organisms as physical ecosystem engineers. Ecology, **78**:1946-1957.

Jørgensen, S.E., 2002. Integration of ecosystem theories: A pattern, 3rd edition. Kluwer Academic Publishers, Dordrecht; Boston.

Jørgensen, S.E., Patten, B.C., and Straškraba, M., 1992. Ecosystems emerging: toward an ecology of complex systems in a complex future. Ecol. Model., **62**:1-27.

Jørgensen, S.E., Patten, B.C., and Straškraba, M., 2000. Ecosystems emerging: 4. Growth. Ecol. Model., **126**:249-284.

Kay, J.J., 1991. A nonequilibrium thermodynamic framework for discussing ecosystem integrity. Environmental Management, **15**:483-495.

Kay, J.J., Regier, H.A., Boyle, M., and Francis, G., 1999. An ecosystem approach for sustainability: addressing the challenge of complexity. Futures, **31**:721-742.

Krakauer, D., 2003. An introduction to form and fuction. **(March 20, 2005)**:[online] http://discuss.santafe.edu/formandfunction/introduction.

Kraut, R., 2001. Aristotle's Ethics. In: E.N. Zalta (Editor), The Stanford encyclopedia of philosphy, [online] http://plato.stanford.edu/archives/sum2001/entries/aristotle-ethics/

Kraut, R., 2004. Plato. In: E.N. Zalta (Editor), The Stanford encyclopedia of philosophy, [online] http://plato.stanford.edu/archives/sum2004/entries/plato/

Lafferty, K.D., and Morris, A.K., 1996. Altered behavior of parasitized killifish increases susceptibility to predation by bird final hosts. Ecology, **77**:1390-1397.

Laland, K.N., Odling-Smee, F.J., and Feldman, M.W., 1999. Evolutionary consequences of niche construction and their implications for ecology. Proc. Natl. Acad. Sci. USA, **96**:10242-10247.

Leontief, W.W., 1965. The structure of the American economy. Sci. Am., **212**:25-35.

Leontief, W.W., 1966. Input-output economics. Oxford University Press, New York.

Levin, S.A., 1998. Ecosystems and the biosphere as complex adaptive systems. Ecosystems, **1**:431-436.

Lindeman, R.L., 1942. The trophic-dynamic aspect of ecology. Ecology, **23**:399-418.

Loehle, C., 1990. Indirect effects - a critique and alternate methods. Ecology, **71**:2382-2386.

Margalef, R., 1991. Networks in ecology. In: M. Higashi and T.P. Burns (Editors), Theoretical studies of ecosystems: the network perspective, Cambridge University Press, Cambridge, pp 41-57.

Matis, J.H., and Patten, B.C., 1981. Environ analysis of linear compartmental systems: the static, time invariant case. Bull. Int. Stat. Inst., **48**:527-565.

Merriam-Webster Online Dictionary, 2005a. Form. **(March 22, 2005)**:http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=form.

Merriam-Webster Online Dictionary, 2005b. Function. **(March 22, 2005)**:http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=function.

Miller, T.E., and Kerfoot, W.C., 1987. Redefining indirect effects. In: W.C. Kerfoot and A. Sih (Editors), Predation: direct and indirect impacts on aquatic communities, University Press of New England, London, pp 33-37.

Miller, T.E., and Travis, J., 1996. The evolutionary role of indirect effects in communities. Ecology, **77**:1329-1335.

Müller, F., 1996. Emergent properties of ecosystems - consequences of self-organizing processes? Senckenbergiana maritima, **27**:151-168.

Müller, F., 1998. Gradients in ecological systems. Ecol. Model., **108**:3-21.

Newman, M.E.J., 2003. The structure and function of complex networks. Siam Review, **45**:167-256.

Odling-Smee, F.J., Laland, K.N., and Feldman, M.W., 2003. Niche construction: the neglected process in evolution. Princeton University Press, Princeton, N.J.

Odum, E.P., 1953. Fundamentals of ecology. Saunders, Philadelphia,.

Odum, E.P., 1962. Relationships between structure and function in ecosystems. Japanese Journal of Ecology, **12**:108-118.

Odum, E.P., 1969. The strategy of ecosystem development. Science, **164**:262-270.

Odum, H.T., 1960. Ecological potential and analogue circuits for the ecosystem. Am. Scientist, **48**:1-8.

Patten, B.C., 1959. An introduction to the cybernetics of the ecosystem: the trophic dynamic aspect. Ecology, **40**:221-231.

Patten, B.C., 1978. Systems approach to the concept of environment. Ohio J. of Sci., **78**:206-222.

Patten, B.C., 1981. Environs: the superniches of ecosystems. Am. Zool., **21**:845-852.

Patten, B.C., 1982. Environs: relativistic elementary particles for ecology. Am. Nat., **119**:179-219.

Patten, B.C., 1983. On the quantitative dominance of indirect effects in ecosystems. In: W.K. Lauenroth, G.V. Skogerboe, and M. Flug (Editors), Analysis of ecological systems: state-of-the-art in ecological modelling, Elsevier, Amsterdam, pp 27-37.

Patten, B.C., 1984. Further developments toward a theory of the quantitative importance of indirect effects in ecosystems. Verh. Gesellschaft für Ökologie, **13**:271-284.

Patten, B.C., 1985a. Energy cycling in the ecosystem. Ecol. Model., **28**:1-71.

Patten, B.C., 1985b. Energy cycling, length of food chains, and direct versus indirect effects in ecosystems. Can. Bull. Fish. Aqu. Sci., **213**:119-138.

Patten, B.C., 1988. Ecology's AWFUL theorem: sustaining sustainability. Ecol. Model., **108**:97-105.

Patten, B.C., 1990. Environ theory and indirect effects: a reply to Loehle. Ecology, **71**:2386-2393.

Patten, B.C., 1991. Network ecology: indirect determination of the life-environment relationship in ecosystems. In: M. Higashi and T. Burns (Editors), Theoretical Studies of Ecosystems: the network perspective, Cambridge University Press, New York, New York, pp 288-351.

Patten, B.C., 1992. Energy, emergy and environs. Ecol. Model., **62**:29-69.

Patten, B.C., 1998a. Network orientors: Steps toward a cosmography of ecosystems: Orientors for directional development, self-organization, and autoevolution. In: F. Müler and M. Leupelt (Editors), Eco Targets, Goal Functions, and Orientors, Springer, Berlin, pp 137-160.

Patten, B.C., 1998b. Steps toward a cosmography of ecosystems: 20 remarkable properties of life in environment. In: F. Mueller and M. Leupelt (Editors), Eco Targets, Goal Functions, and Orientors, Springer-Verlag, New York, New York

Patten, B.C., in prep. Holoecology: the unification of nature by network indirect effects. Columbia University Press, New York.

Patten, B.C., Bosserman, R.W., Finn, J.T., and Cale, W.G., 1976. Propagation of cause in ecosystems. In: B.C. Patten (Editor), Systems Analysis and Simulation in Ecology, Vol. IV, Academic Press, New York, pp 457-579.

Patten, B.C., Fath, B.D., Choi, J.S., Bastianoni, S., Borrett, S.R., Brandt-Williams, S., Debeljak, M., Fonseca, J., Grant, W.E., Karnawati, D., Marques, J.C., Moser, A., Müller, F., Pahl-Wostl, C., Seppelt, R., Seinborn, W.H., and Svirezhev, Y.M., 2002. Complex adaptive hierarchical systems. In: R. Costanza and S.E. Jørgensen (Editors), Understanding and solving environmental problems in the 21st century: Toward a new, integrated hard problem science, Elsevier Science, Ltd, Oxford, pp 41-87.

Patten, B.C., Gattie, D.K., Bata, S., Borrett, S.R., Fath, B.D., Schramski, J.R., and Whipple, S.J., in prep. Environs and network environ analysis: introduction and overview. Ecol. Model.

Patten, B.C., Higashi, M., and Burns, T.P., 1990. Trophic dynamics in ecosystem networks: significance of cycles and storage. Ecol. Model., **51**:1-28.

Patten, B.C., and Matis, J.H., 1982. The water environs of the Okefenokee swamp: an application of static linear environ analysis. Ecol. Model., **16**:1-50.

Patten, B.C., Richardson, T.H., and Barber, M.C., 1982. Path analysis of a reservoir ecosystem model. Can. Water Resourc. J., **7**:252-282.

Ponstein, J., 1966. Matrices in graph and network theory. Van Gorcum & Comp., Assen,.

Rapport, D., Costanza, R., Epstein, P.R., Gaudet, C., and Levins, R., Editors. 1998. Ecosystem health. Blackwell Science, Oxford.

Reichman, O.J., and Pulliam, H.R., 1996. The scientific basis for ecosystem management. Ecological Applications, **6**:694-696.

Schneider, E.D., and Kay, J.J., 1994. Life as a manifestation of the second law of thermodynamics. Mathematical and Computer Modelling, **19**:25-48.

Strauss, S.Y., 1991. Indirect effects in community ecology - their definition, study and importance. Trends Ecol. Evol., **6**:206-210.

Tansley, A.G., 1935. The use and abuse of vegetational concepts and terms. Ecology, **16**:284-307.

Teal, J.M., 1959. Energy flow in the salt marsh ecosystem in Georgia. Anat. Rec., **134**:647-647.

Ulanowicz, R.E., 1986. Growth and development: ecosystems phenomenology. Springer-Verlag, New York.

Ulanowicz, R.E., 1997. Ecology, the ascendent perspective. Columbia University Press, New York.

Westra, L., and Lemons, J., Editors. 1995. Perspectives on ecological integrity. Kluwer Academic Publishers, Dordrecht, Netherlands.

Wootton, J.T., 1993. Indirect effects and habitat use in an intertidal community: interaction chains and interaction modifications. Am. Nat., **141**:71-89.

Wootton, J.T., 1994. The nature and consequences of indirect effects in ecological communities. Ann. Rev. Ecol. Syst., **25**:443-466.

Wootton, J.T., 2002. Indirect effects in complex ecosystems: recent progress and future challenges. Journal of Sea Research, **48**:157-172.

CHAPTER 2

STRUCTURE OF PATHWAYS IN ECOLOGICAL NETWORKS: RELATIONSHIPS BETWEEN LENGTH AND NUMBER[1]

---

[1] Borrett, S.R. and B.C. Patten. 2003. Ecological Modelling 170: 173-184.
Reprinted here with permission of publisher

ABSTRACT

In ecosystems, network structure determines adjacent (direct) and non-adjacent (indirect) pathways over which energy, matter, and information can flow. The more pathways the more possible ways conservative substance can move in zero-sum transactions between network nodes that the pathways interconnect, and the more possible non-conservative, nonzero-sum relations can be secondarily derived from these. Structural analysis is a tool we employ, from a family of input–output methods for exploring zero- and nonzero-sum attributes of environmental networks, to count pathways of varying length between network nodes. In this paper we examine the relationship between pathway length (k) and number (P_k) as determined by system size (n, number of nodes) and extent and pattern of connectance (C). We develop a measure (m_a) of pathway growth in numbers with increasing length, and then normalize this to the maximum rate possible (m_a/m_c) for a given system size. These measures apply to two pathway types—*paths*, $m_a(0)$ and $m_a(0)/m_c(0)$, which forbid adjacent node repetitions, and *walks*, $m_a(1)$ and $m_a(1)/m_c(1)$, which allow them. We find that network size has a curvilinear effect on the pathway number vs. length relationship, and extent and pattern of connectance are convolved. Values computed for the paths and walks of three ecosystem models (oyster reef, freshwater marsh, and reservoir cove) are used to compare their pathway structure.

2.1 INTRODUCTION

Environ analysis (Patten, 1978a, 1982; Matis and Patten, 1981) is a family of input–output methods of network analysis descended from Leontief (1965, 1966) and introduced into ecology by Hannon (1973). Environ methodology is employed in the holistic study of ecological networks to describe, quantify, and analyze component-level environments (environs) within systems. These environments are of two types, afferent (input environs) which are generated by history, and efferent (output environs) which are propagated to the future. Environ methods include analyses for structure, flows, utilities, and control within systems, and are a form of complex-systems analysis as described by Allen and Hoekstra (1992, p. *xiv*). In this paper, we elucidate several attributes of environ network structure, which underlies much of the environ analysis methodology.

Network *structure* establishes the pathways (direct and indirect) over which conserved material can flow between entities in connected systems (Fath and Patten, 1999). The relationship between pathway numbers and length is a significant component of this structure because it describes how quickly indirect pathways increase, and with this also, proportional opportunities for substance to move and for positive “nonzero-sumness” (Wright, 2000) to be realized. In a well-connected system, the number of pathways in a network increases without bound as pathway length increases (Fath, 1998; Patten, 1985). The rate of this increase, however, varies (Borrett, 2001; Fath, 1998). Here, we investigate the cause of this variation by developing measures to examine the proposition that the relationship between pathway length and number is determined by three variables: system size, connectance degree, and connectance pattern. We apply our measures to three ecosystem models and then consider the results.

2.1.1 TRANSACTIONS AND RELATIONS

The network perspective (Margalef, 1963; Higashi and Burns, 1991) conceptualizes an ecosystem as a reticulum of interactions. These are of two types, transactions and relations (Fath and Patten, 1998).

Transactions are physical (ontic, electromagnetic) and have primacy. They are zero-sum exchanges of conservative substances between adjacent network nodes representing processes and storages (slow processes). Zero-sumness means what is gained (+) by one member of an interacting node pair identically equals what is lost (–) by the other. The quantities being equal, their signs opposite, and the substance transferred conservative means the interactive sum is zero.

Relations are phenomenal (epistemic, semiotic) interactions as experienced and interpreted by participants and observers. What is transferred between non-adjacent nodes is conservative but nonzero-sum (Wright, 2000), and in the latter property begins to have attributes of non-conservative information. Competition, predation, and mutualism are classic examples of relations in ecology. In general, relations have their basis in transactions but need not in themselves involve direct energy or matter interchange, *e.g.*, predation does, but competition and mutualism do not.

By conceptualizing complex systems like ecosystems as networks, we can analyze them holistically for properties arising from the transactions and relations interconnecting component parts.

2.1.2 PATHWAYS AND CYCLES

Network structure can be represented by digraphs (directed graphs) or isomorphic matrices (Margalef, 1963; Hannon, 1973; Patten, 1985). In digraphs a *node* is a point representing a system component and an *arc* is an arrow denoting a directed transaction (energy or matter flow) or relation (Figure 2.1A). A *pathway* is a sequence of nodes and arcs over which a substance flows (in transactions) or relationships emerge (in relations) from node j to node i. Pathway *length* (k) is the number of arcs from an initial to a terminal node. A *direct pathway* between two nodes has only one arc, $k = 1$; an *indirect pathway* has $k > 1$ arcs. Pathways can be *simple*, where there are no repeated nodes, or *compound*, with repeated nodes. In the Fig. 2.1 network $j \rightarrow f \rightarrow g \rightarrow h \rightarrow i$ is a simple pathway from j to i while $j \rightarrow f \rightarrow g \rightarrow f \rightarrow i$ is a compound pathway. Both are indirect and of length $k = 4$.

Cycles are an important type of pathway in network structures (Patten, *et al.*, 1990). A *cycle* is a pathway starting and ending at the same node, as in the embedded pathway segment … $\rightarrow f \rightarrow g \rightarrow f \rightarrow$ … in our example (Figure 2.1A). Like other pathways, cycles can be simple (first time around) or compound. A cycle of length $k = 1$ is a *self-loop*; this can be used to represent node storage in a network (e.g., …$\rightarrow i \rightarrow i \rightarrow i \rightarrow$ …).

2.1.3 NETWORK STRUCTURAL MEASURES

The s*ize* (n) of a network equals its number of nodes. *Connectance* (C) is the proportion of realized direct ($k = 1$) pathways, often termed *links* (L) in food-web literature (Cohen *et al.,* 1990; Martinez, 1994; Bersier *et al.*, 1999; Williams and Martinez, 2000), divided by the number of possible links, n^2: $C = L/n^2$ (see Margalef, 1991; Newman, 2002 for alternative formulations).

Digraphs have isomorphic representations as adjacency matrices, $\mathrm{A}_{n \times n} = (a_{ij})_{n \times n}$ (Figure 2.1B). An *adjacency matrix* is a square matrix whose rows and columns are indexed by the node vector ($i, j = 1, 2, \ldots, n$) and whose elements take values $a_{ij} = 1$ if and only if a link directed from column j to row i exists, otherwise $a_{ij} = 0$.

Each column denotes the presence (ones) or absence (zeros) of links directed from node *j* to node *i* in the corresponding digraph, that is, links in the *j*'th output environ of the system. In Figure 2.1B, for example, column *j* has ones at rows *f* and *i* denoting links $j \rightarrow f$ and $j \rightarrow i$ in Figure 2.1A. The orientation from columns (initial nodes) to rows (terminal nodes) is the same as that of the digraph arrows, reflecting the fact that output environs extend forward through the system from starting to ending nodes.

Each row of A denotes the presence or absence of links directed to node *i* from row *j* in the system's *i*'th input environ. Thus, in Figure 2.1B row *i* has ones at columns *j*, *f* and *h* denoting links $i \leftarrow j$, $i \leftarrow f$, and $i \leftarrow h$ in Figure 2.1A. The orientation to rows (terminal nodes) from columns (initial nodes) is opposite that of the digraph arrows, reflecting the fact that input environs extend backward through the system from ending to starting nodes.

This reverse orientation can be treated separately in the columns-to-rows format of A matrices by performing the operations of this paper on corresponding transposes $A^{T}_{n \times n} = (a_{ji})_{n \times n}$, mentally reversing the arrow directions. Numerical results would differ since the pathway structure directed from initial to terminal nodes differs from that in the reverse direction, but the principles elaborated would be the same. Therefore, for present purposes we will limit further attention to output environs whose structure is generated by matrices A, not their transposes.

We will employ two forms of the adjacency matrix. A(0) will denote a matrix with zero principal diagonal entries, $a_{jj} = 0$, signifying no node storage, and A(1) will be a matrix with diagonal ones, $a_{jj} = 1$, representing digraph self-loops denoting storage. The former pathways are termed *paths* and the latter *walks* in subsequent usage. *Connectance pattern* is the network geometry (topology) of nodes and arcs as reflected in a digraph or its corresponding adjacency matrix.

A power matrix, A^k can be used to compute the number of pathways of length $k \geq 0$ directed from j to i in a network (Harary, 1969; Hill, 1981); $k = 0$ denotes self-loops. The total number of pathways (P_k) of different lengths (k) from all nodes to all other nodes is given by the scalar measure:

$$P_k = \sum_{j=1}^{n} \sum_{i=1}^{n} (a_{ij})^k .$$

The number of pathways of all lengths from each j to each i is generated by the matrix power series:

$$I + A + A^2 + A^3 + \ldots + A^k + \ldots$$

The first term $A^0 = I_{n \times n}$ corresponding to $k = 0$ signifies the network's n nodes as initial in pathway propagation. The next term A denotes adjacency – direct links ($k = 1$) defined by zero-sum transactions directed from j's to i's. The remaining terms A^k, $k > 1$, denote pathways that are non-adjacent, indirect, relational, and carry nonzero-sum flows. In a well-connected system, one with cyclic feedback and a maximum real eigenvalue greater than one, the series diverges because sums of powers of the A-matrix entries grow as k increases (Fath, 1998). This is consistent with the fact that ecological systems are dissipative; energy and matter pass through them and ultimately exit as a limit process. This implies $k \rightarrow \infty$ and therefore $A^k \rightarrow \infty$. Also, as k increases, P_k increases approximately geometrically. By log-transforming P_k we generate a line of slope m_a (index a referring to the digraph corresponding to A) that can be used to describe the rate of growth of P_k with k.

2.1.4 NETWORK STRUCTURE IN ENVIRON ANALYSIS

The methods of environ analysis (Patten 2003, Chapter 4) can be grouped in pairs:

(1 & 2) Input and output environ *structure*, which is our focus in this paper.

(3 & 4) Input and output environ *flows to throughflows.* Throughflows (T_i) are sums of flows at nodes ($i = 1, \ldots, n$) in storage-and-flow networks. The input environ case corresponds to original Leontief (1965, 1966) input–output analysis.

(5 & 6) Input and output environ flows to node *storage* (x_i, $i = 1, \ldots, n$). From the network perspective, storage is interpreted as flow impedance, like capacitance in electrical networks.

(7 & 8) Throughflow- and storage-based *utilities*, $u(T_i)$ and $u(x_i)$ respectively, generated by net input- and output-environ direct flows between node pairs (Patten 1991). Utility measures direct and indirect values of throughflow and storage conferred by the particulars of network organization.

(9 & 10) Throughflow- (Patten, 1978b) and storage-based (Patten, 2003) control, $c(T_i)$ and $c(x_i)$ respectively, exerted over network distances, k. Direct and indirect control are assessed and quantified between each node pair in a system.

Each of these methods generates a unique perspective on relationships inherent in ecological networks. These derive from one overarching result from analyses 3–6 above: flows $f(\mathrm{A}^k)$ over pathways A^k of lengths $k > 1$ in the adjacency matrix power series often exceed in aggregate those associated with the direct links in the adjacency matrix A itself. That is, $\sum_{k>1} f(\mathrm{A}^k) > f(\mathrm{A})$. The reason is because although conservative quantities dissipate with transfers and transformations in networks, this happens more slowly than the rates at which pathways are generated in well-connected networks (Patten, 1985). Each pathway of whatever length carries some quantity of transferred substance until lengths k_d are reached where materials have been effectively dissipated (dissipation is a limit process) such that flows over those pathways are zero. This fact allows power series of non-dimensional flows to converge, which is what enables input–output methods to calculate integral flows carried by all pathways P_k of all lengths $k \geq 0$ in a system. Small, even infinitesimal, flows over large, approaching astronomical, numbers of pathways generated before lengths k_d are reached can add up to significant total flow over paths P_k of lengths $k > 1$. This is in fact the essential input–output network dynamic. Its consequence is that indirect effects dominate (Patten, 1984; Higashi and Patten, 1989) and make holistic determination (Patten *et al.,* 1976) the principal mode of causality in nature (Patten, 2003).

Higashi (see Patten, 1991) developed algebraic formulations to show that certain network properties increase the dominance of indirect over direct effects. These are system size (n), connectivity (C), network looping (storage), network cycling, feedback cycling (return to initial nodes), and strength of direct arc flows. All but the last are structural properties, as considered here. In this paper we investigate the relationship between pathway length (k) and number (P_k) as determined by the first two of Higashi's properties, n and C.

As pathway structure defines how conserved substances can flow in connected networks, the more pathways there are the more possible ways exist for energy and matter to move between nodes. Structural analysis, however, does not consider how much substance actually flows along each pathway, it only enumerates the pathways. To determine flow relationships requires other analyses such as those in the environ set of methods. As stated previously, here our topic is structural analysis.

2.2 MATERIALS AND METHODS

We will investigate how three properties of networks—(1) system size, (2) connectance degree, and (3) connectance pattern—contribute to the rate of pathway proliferation. Below we describe methods to test each, and introduce model ecosystems we will analyze as examples.

2.2.1 SYSTEM SIZE

To determine if there is a relationship between system size and pathway length and number, we analyzed digraphs of varying sizes ($n = 1$ to $n = 200$). Both arbitrary digraphs (a) and their complete counterparts (c) were studied. Complete digraphs have a link to and from every node ($a_{ij} = 1, \forall i,j$); their analyses quantify walks due to principal diagonal values of one, $a_{jj} = 1$. For paths, the same digraphs with self-loops excluded ($a_{ij} = 1\ \forall i \neq j$, and $a_{jj} = 0$) were analyzed. For a given network a (or c) we determined m_a (or m_c), the slope of a line describing the relationship between pathway length and the logarithm of pathway number. Slopes were calculated as differences between an arbitrarily selected pair of adjacent P_k values; the values used were $k = 50$ and 51: m_a (or m_c) $= \log_{10} (P_{51}) - \log_{10} (P_{50})$. If size does not affect the relationship slopes should be roughly equivalent between networks. Use of complete graphs neutralized the effects of degree and pattern of connectance, and also generated the maximum number of possible pathways.

2.2.2 EXTENT OF CONNECTANCE

If there were a relationship between system size and pathway length and number, the size effect would have to be removed in order to examine the connectance of systems of different sizes. To accommodate this we created a normalized response variable by taking ratios of slopes generated by actual digraphs, m_a, to those for complete graphs, m_c, of the same size. For paths this measure is $m_a(0)/m_c(0)$ and for walks $m_a(1)/m_c(1)$. The second part of our investigation was to determine if degree of connectance in adjacency matrices altered m_a/m_c ratios. We held system size constant and then systematically added connections (Figure 2.2, networks A1–A4). If connectivity does not affect the relationship, then m_a/m_c should remain constant. It became clear that augmenting connectivity typically alters the pattern of connectance as well (next section). The two attributes are convolved, such that if varying the pattern changes m_a/m_c, the effects of degree cannot be distinguished from pattern.

2.2.3 PATTERN OF CONNECTANCE

To investigate the effects of connectivity pattern we held network size and degree of connectance constant while varying the pattern (Figure 2.2, networks A5–A8). Changes were made so as to maintain cyclic feedback. Each test network had $n = 9$ and $C = 0.12$. In A5 there is a chain of links spanning the network, with two links in the upper right corner providing the possibility for cycle development. Matrix A6 is similar to A5 but the chain is slightly altered. In A7, the chain is replaced with a block of connections in the lower left corner. The final manipulation (A8) was to aggregate the links in the top right and bottom left corners. Structural analysis was performed on each of these networks to determine m_a/m_c for paths and walks. We would consider connectance pattern as having no effect if these measures remained constant as topology changed.

In addition to actually changing connections (Figure 2.2, A5–A8) it is also possible to change the perceived linkage pattern. For example, renumbering nodes in a digraph would leave the pattern of connections unchanged, but the display in the corresponding adjacency matrix would (typically) be altered. To test if this kind of change would be reflected in our measures we analyzed alternative representations of the same network created by interchanging two or more columns and then making an identical change in the corresponding rows (Figure 2.2, networks A9–12; the A9 matrix functions as a control). In A10 column (row) 5 (of A9) was placed between columns (rows) 1 and 2, column (row) 7 between columns (rows) 2 and 3, and column (row) 8 between columns (rows) 3 and 4. Matrix A11 was created by moving column (row) 4 into the first column (row) position and column (row) 9 between columns (rows) 6 and 7. Matrix A12 was generated by moving columns (rows) 5 and 6 into the first two positions. These row and column interchanges clearly changed perceived patterns in the adjacency matrices, but without altering actual network topology. These manipulations are equivalent to symmetric permutations of the adjacency matrix. If m_a/m_c ratios changed under this kind of manipulation we would conclude that the measure is invalid for network structural analysis. A proper measure would yield unique m_a/m_c ratios for a network irrespective of its representation.

2.2.4 EXAMPLE NETWORKS

We applied structural network analysis to three static, steady state, ecosystem models to compare their direct and indirect pathway structure. The models differ in size, connectivity, connectance pattern, and type of ecosystem represented. Our goal was to determine the degree of similarity in these models' network structures. The models were for (1) an intertidal oyster reef in coastal South Carolina (Dame and Patten 1981), (2) a marsh called Little Cooter Prairie in Okefenokee Swamp (Whipple 1995), and (3) a reservoir cove in Lake Texoma (Patten *et al.* 1975). We will refer to these as oyster, marsh, and cove models, respectively. The oyster model has six compartments and connectivity $C = 0.50$. Patten (1985) described the direct and indirect pathway structure of this network, and the model has been used further to develop ecological network theory (*e. g.*, Patten *et al.* 1990, Higashi and Burns 1991). The marsh model has 42 compartments and a connectivity of $C = 0.20$. The cove model has 33 compartments and connectivity $C = 0.30$. Like the oyster model, the cove model has been used before in theory development and many of its specific pathways have been identified (*e. g.*, Patten *et al.,* 1982).

2.3 RESULTS

Experimental results showed network size, and extent and pattern of connectance all affect the relationship between path numbers (P_k) and length (k). The effect of connectance degree was inconclusive because this is convolved with pattern. Relabeling digraph nodes without altering real topology changed the adjacency matrix pattern, but did not alter m_a/m_c ratios.

2.3.1 SYSTEM SIZE

Figure 2.3 shows results of determining if size (n) of fully connected networks increases the rate (m_c) at which higher order pathways are generated. As network size increases the rate of growth of pathway numbers increases and the difference between m_c values for corresponding paths and walks decreases (Figure 2.3A). The reason is that the proportion of self-loops (n) in a fully connected network relative to all arcs (n^2) decreases by the factor $n/n^2 = 1/n$. This causes the relationship between system size and rate of pathway increase to be curvilinear (Figure 2.3B). In fully connected networks size influences the growth of system structure most in smaller networks where the ratio of n to n^2 is greatest.

2.3.2 EXTENT OF CONNECTANCE

To evaluate degree of connectance in non-complete digraphs requires use of a normalized metric like m_a/m_c. Given $C < 1$ (incomplete connectance) one can visualize slopes of actual networks (m_a) being less than those of corresponding complete networks (m_c). Then, in Figure 2.3A format, m_a values would generate lines of lesser slope than m_c and curve forms like those of Figure 2.3B would rise more slowly. Connectance changes made in the experimental networks A1–A4 of Figure 2.2 changed the values of m_a/m_c (Table 2.1). These ratios increased with connectance but not consistently. For example, network A8 in Table 2.1 has lower connectance than A1, but its m_a/m_c values are larger. This shows that extent and pattern of connectance are interwoven quantities, not easily separated by simple measures.

2.3.3 PATTERN OF CONNECTANCE

Changing the pattern of connectance clearly alters the relationship between number (P_k) and length (k) of pathways. In Table 2.1 network A5 has $m_a(0)/m_c(0) = 0.0413$ and $m_a(1)/m_c(1) = 0.3345$. The small pattern change in A6 vs. A5 causes $m_a(0)/m_c(0)$ to increase to 0.1024 and $m_a(1)/m_c(1)$ to 0.3668. The pattern in A7 generates $m_a(0)/m_c(0) = 0.3211$ and $m_a(1)/m_c(1) = 0.4923$, while A8 has an $m_a(0)/m_c(0)$ ratio of 0.3333 and an $m_a(1)/m_c(1)$ value of 0.5000. As connections become more concentrated in the lower left and top right corners of the matrix, reflecting an increase in the number of cycles, m_a/m_c increases.

Reordering network nodes in the adjacency matrix under the constraint that order is consistent for both rows and columns, does not affect m_a/m_c. In all four representations of a fixed topology (A9–A12), the same path and walk ratios were obtained (Table 2.1): $m_a(0)/m_c(0) = 0.3406$ and $m_a(1)/m_c(1) = 0.5046$. This result is not unexpected and is quite general because symmetric permutations of adjacency matrices generate isomorphic graphs, which have identical structural properties (Logofet 1993; Bondy and Murty 1976).

2.3.4 COMPARISON OF EXAMPLE MODELS

In non-complete networks of the oyster, marsh, and cove models pathway numbers (P_k) increased without bound as length (k) increased (Figure 2.3). As in the Figure 2.2 experimental networks, walks increased faster than paths and differences in their rates of increase decreased with system size. The rate (m_a) at which number of pathways increases in the oyster model is less than that of the marsh and cove models, which between themselves are very similar. Rates $m_a(0)$ and $m_a(1)$ of, respectively, path and walk generation in the three models indicate that the difference between marsh and cove models is small and that between these two and the oyster model larger (Table 2.2).

The normalized measure m_a/m_c indicates somewhat different relationships than those based on non-normalized slopes. In Table 2.2, $m_a(0)/m_c(0)$ is least for the oyster model (0.4750) and greatest for the cove model (0.6826), whereas $m_a(1)/m_c(1)$ is least for the marsh model (0.6321) and greatest (0.7027) for the cove model. The ranking changes between paths and walks. Faster walk than path development in the oyster vs. the other two models is a size ($n/n^2 = 1/n$) effect, as previously described. Adding n self-loops to $< n^2$ arcs (in incomplete graphs) has a greater effect when n is small. It is also evident from higher values of $m_a(0)/m_c(0)$ and $m_a(1)/m_c(1)$ for the cove model (Table 2.2) that more of the potential ways to move energy and matter are realized in this model than in the oyster and marsh models.

2.4 DISCUSSION

In ecosystems virtual structure is defined by direct (adjacent) zero-sum transactions of conservative energy and matter exchanged between organisms and their environments. The transactions set up more diverse and wide-ranging informational relations that are indirect (non-adjacent) and nonzero-sum. Relations make life in the ecosphere infinitely varied and for the most part positively experienced by organisms (Patten, 1991) over their time-lines (Jørgensen *et al.*, 1992).

Network structure establishes pathways over which conserved material can flow between entities in connected systems. The relationship between pathway numbers and length is a significant network attribute because it describes how quickly indirect pathways increase, and thus, the volume of pathways available for transactions and relations. In this paper we have demonstrated, consistent with other findings from environ analysis, that network size (n) and connectance degree (C) and pattern contribute to shaping of the pathway length (k) and number (P_k) relationship. System size was shown to have a curvilinear effect. Connectance extent and pattern were found to jointly influence this relationship, though in ways not easily separated. Our main conclusions are:

1. In well-connected networks, pathways increase geometrically with length regardless of system size, as measured by m_a;
2. System size has a curvilinear effect on the relationship between pathway length and number;
3. Irrespective of system size, connectance can be compared utilizing the measure m_a/m_c; and
4. While connectance degree is entangled with pattern, connectance pattern has a clear role in determining the rate of pathway increase, as measured by m_a.

Our measure of path growth, m_a, indicates an increase rate based on the total number of pathways P_k of selected lengths k. A less empirical (and less heuristic) metric is the dominant eigenvalue of the adjacency matrix, $max|\lambda|$. This analytically determines the asymptotic growth rate of paths between any node pair (Fath 1998, Hill unpublished manuscript). This rate is identical for all node pairs and represents the overall growth rate of pathways for that structure. Experimental trials using the maximum eigenvalue measure (Tables 1 and 2) are consistent with those reflected in m_a.

By eliminating the effect of size, our normalized measures m_a/m_c for paths and walks characterize the combined effect of connectance and topology on the indirect network structure.

The measures can be used to compare indirect structure between systems of different sizes, and the next logical step in this research would be to try to establish a relationship between particular connectance patterns and associated variation in m_a/m_c values. Teasing apart the effect of magnitude from the pattern of connectance itself we consider important to making inferences from m_a/m_c back to network structure. Other research questions also present themselves, such as the relationship between m_a and m_a/m_c and network properties such as the cycling index (Finn, 1976), or some of the eleven cardinal properties of "holoecology" (Patten, 2003) derived from the environ approach:

#1. *Network proliferation*—increase in pathway numbers (P_k) with length (k), which is geometric increase as described herein;

#2. *Network nonlocality*—dominance of indirect effects as carried by transactive flows (f), $\sum_{k>1} f(\mathrm{A}^k) > f(\mathrm{A})$;

#3. *Network holocontrol*—dominance of indirect control (c), $\sum_{k>1} \mathrm{c}(\mathrm{A}^k) > \mathrm{c}(\mathrm{A})$;

#4. *Network homogenization*—tendency to uniformly distribute causality;

#5. *Network amplification*—obtaining more than face value from boundary inputs $f(a_{ij}^{(0)})$: $\sum_{k>0} f(a_{ij}^{(k)}) > f(a_{ij}^{(0)})$;

#6. *Network unfolding*—proliferation of transfer levels as a limit process, which in food webs leads to trophic pyramids with an indefinite number of trophic levels;

#7. *Network synergism*—dominance of indirect over direct utility (u), $\sum_{k>1} u(\mathrm{A}^k) > u(\mathrm{A})$;

#8 *Network mutualism*—indirect utilities becoming more positive than direct;

#9. *Network aggradation*—internal order exceeding generated disorder;

#10. *Network enfolding*—recursive incorporation of indirect into direct causes such that $f(\mathrm{A})$ is really $f(\mathrm{A}(\sum_{k\geq 1} f(\mathrm{A}^k)))$; and

#11. *Network holoevolution*—the coevolution of wholes and parts together.

The classical concept of "pattern and process" in ecology comes to the fore in this list. Pathway notations ($\mathrm{A}^k = (a_{ij}^{(k)})$, $k \geq 0$) introduced into some of the definitions show how central network structure is to the principal results of environ analysis. That network properties reflecting function ("process") can be depicted (in functional notation) as functions of structure ("pattern") underscores how closely the two are intertwined. In fact, as previously stated, structure is virtual and follows function (transactions), which is actual. From this it is reasonable to expect that structural measures such as m_a and m_a/m_c can be useful in predicting or bounding measures of function such as the cycling index or some of the eleven above. The logic here is that labyrinthine transactions define and implicate, respectively, the direct and indirect pathway structure within systems, including cycles. In fact the implication is mutual. Structure and function, or pattern and process, go together and methods like those in the set of environ analyses are going to be required to distinguish one from the other.

Scale and modeling issues will be ever present in this. The oyster reef, swamp marsh, and reservoir cove ecosystems were each modeled at particular scales not necessarily alike. Network size and connectivity in models reflect the resolution at which systems are studied. A network representation of the interior of a cell can be far more intricate than that for an ecosystem (*e.g.,* Guet *et al.,* 2002; Jeong *et al.,* 2000; Maslov and Sneppen, 2002). At their own scales it is true that both can be equally complex, or the smaller perhaps more complex than the larger. In general, finer grain of observation leads to more nodes and links, which by Higashi's algebraic relations translates into greater dominance of indirect effects. However, recalling the definition of connectance ($C = L/n^2$, L the number of links), it is clear that finer resolution of nodes (n) will cause n^2 to grow faster than n, rarefying C. At very large system sizes adjacency matrices can be expected to become extremely sparse though number of links is astronomical. A system with a billion nodes ($n = 10^9$) and a trillion transactions interconnecting them ($L = 10^{12}$) will be only one-millionth connected ($L/n^2 = 10^{-6}$), a sparse universe indeed in the present experience of ecological modeling. How to balance practical needs to describe and predict with the companion need to understand principles in order to better describe and predict, will present a continuing challenge.

2.5 ACKNOWLEDGMENTS

This manuscript benefited from reviews by Brian D. Fath, Stuart J. Whipple, Mitchell Pavao-Zuckerman, and anonymous reviewers. SRB acknowledges financial support from EPA-NSF STAR Water and Watersheds Grant #R825758.

2.6 REFERENCES

Allen, T.F.H., Hoekstra, T.W. 1992. Toward a unified ecology. Columbia University Press, New York.

Bersier, L.F., Dixon, P., Sugihara, G. 1999. Scale-invariant or scale-dependent behavior of the link density property in food webs: A matter of sampling effort? Am. Nat. 153, 676-682.

Bondy, J.A., Murty, U.S.R. 1976. Graph Theory with Applications. North-Holland, New York.

Borrett, S.R. 2001. Sources of system complexity in the Lake Lanier ecosystem, GA, USA. Paper read at Ecological Society of America 86th Annual Meeting, August, at Madison, WI.

Cohen, J.E., Briand, F., Newman, C.M. 1990. Community food webs: data and theory. Springer-Verlag, New York.

Dame, R.F., Patten, B.C. 1981. Analysis of energy flows in an intertidal oyster reef. Mar. Ecol. Progr. Ser. 5, 115-124.

Fath, B.D. 1998. Network analysis: foundations, extensions, and applications of a systems theory of the environment. (Ph.D. Thesis). Institute of Ecology, University of Georgia, Athens, Georgia.

Fath, B.D., Patten, B.C. 1998. Network synergism: emergence of positive relations in ecological systems. Ecol. Model. 107, 127-143.

Fath, B.D., Patten, B.C. 1999. Review of the foundations of network environ analysis. Ecosystems 2, 167-179.

Finn, J.T. 1976. Measures of ecosystem structure and function derived from analysis of flows. J. Theor. Biol. 56, 363-380.

Guet, C.C., Elowitz, M.B., Hsing, W., Leibler, S. 2002. Combinatorial synthesis of genetic networks. Science 296, 1466-1470.

Hannon, B. 1973. The structure of ecosystems. J. Theor. Biol. 41, 535-546.

Harary, F. 1969. Graph theory. Addison-Wesley Pub. Co., Reading, Massachusetts.

Higashi, M., Burns, T.P. 1991. Theoretical Studies of Ecosystems: The Network Perspective. Cambridge University Press, Cambridge.

Higashi, M., Patten, B.C. 1989. Dominance of indirect causality in ecosystems. Am. Nat. 133, 288-302.

Hill IV, J. 1981. Influence: a theory of the structural organization of systems. (Ph.D. Thesis), University of Georgia, Athens, Georgia, USA.

Hill IV, J. On the largest real root of a polynomial. Unpublished manuscript.

Jeong, H., Tomnor, B., Oltvai, Z.N., Barabási, A.-L. 2000. The large-scale organization of metabolic networks. Nature 407, 651-654.

Jørgensen, S.E., Patten, B.C., Straškraba, M. 1992. Ecosystems emerging: toward an ecology of complex systems in a complex future. Ecol. Model. 62, 1-27.

Leontief, W.W. 1965. The structure of the American economy. Sci. Am. 212, 25-35.

Leontief, W.W. 1966. Input-Output Economics. Oxford University Press, New York.

Logofet, D. O. 1993. Matrices and Graphs: Stability Problems in Mathematical Ecology. CRC Press, Boca Raton.

Margalef, R. 1963. Certain unifying principles in ecology. Am. Nat. 97, 357-374.

Margalef, R. 1991. Networks in ecology. In: Higashi, M., Burns, T. P. (eds.). Theoretical studies of ecosystems: the network perspective. Cambridge University Press, Cambridge, pp. 41-57.

Martinez, N.D. 1994. Scale-dependent constraints on food-web structure. Am. Nat. 144, 935-953.

Maslov, S., Sneppen, K. 2002. Specificity and stability in topology of protein networks. Science 296, 910-913.

Matis, J.H., Patten, B.C. 1981. Environ analysis of linear compartmental systems: the static, time invariant case. Proc. 42nd Session, International Statistical Institute, December 4-14, 1979, Manila, Philippines. Bull. Int. Stat. Inst. 48: 527-565.

Newman, M.E.J. 2002. Random graphs as models of networks. In: Bornholdt, S., Schuster, H.G. (eds.). Handbook of graphs and networks. Wiley-VCH, Berlin. preprint [online] http://arxiv.org/abs/cond-mat/0202208

Patten, B.C. 1978a. Systems approach to the concept of environment. Ohio J. of Sci. 78, 206-222.

Patten, B.C. 1978b. Energy environments in ecosystems. In: Fazzolare, R.A., Smith, C.B. (eds.). Energy Use Management. Pergamon Press, New York, New York, pp. 853-857.

Patten, B.C. 1982. Environs: Relativistic Elementary Particles for Ecology. Am. Nat. 119, 179-219.

Patten, B.C. 1984. Toward a theory of the quantitative dominance of indirect effects in ecosystems. Verh. Gesellschaft für Ökologie. 13, 271-284.

Patten, B.C. 1985. Energy cycling in the ecosystem. Ecol. Model. 28, 1-71.

Patten, B.C. 2003. Holoecology: The Unification of Nature by Network Indirect Effects. Complexity in Ecological Systems Series. Columbia University Press, New York (in preparation).

Patten, B.C., Bosserman, R.W., Finn, J.T., Cale, W.G. 1976. Propagation of cause in ecosystems. In: Patten, B.C. (ed.) Systems Analysis and Simulation in Ecology, Vol. IV. Academic Press, New York, pp. 457-579.

Patten, B.C., Egloff, D.A., Richardson, T.H., and 38 coauthors. 1975. Total ecosystem model for a cove in Lake Texoma. In: Patten, B.C. (ed.) Systems Analysis and Simulation in Ecology, Vol. III. Academic Press, New York, pp. 205-421.

Patten, B.C., Higashi, M., Burns, T.P. 1990. Trophic dynamics in ecosystem networks: significance of cycles and storage. Ecol. Model. 51, 1-28.

Patten, B.C., Richardson, T.H., Barber, M.C. 1982b. Path analysis of a reservoir ecosystem model. Can. Water Resourc. J. 7, 252-282.

Whipple, S.J. 1995. Systems analysis of the path and trophic structure on an ecosystem model of a macrophyte marsh in the Okefenokee Swamp. (Ph.D. Thesis). Institute of Ecology, University of Georgia, Athens, Georgia, USA.

Williams, R.J., Martinez, N.D. 2000. Simple rules yield complex food webs. Nature 404, 180-183.

Wright, R. 2000. NonZero: the Logic of Human Destiny. Pantheon, New York.

2.7 TABLES

Table 2.1: Effect of connectance degree and pattern on measures of the rate of pathway proliferation for walks ($m_a(1)/m_c(1)$), paths ($m_a(0)/m_c(0)$), and the asymptotic rate (max|λ|); Fath, 1998; Hill, unpublished manuscript)

Matrix	n	C	$m_a(1)/m_c(1)$	$m_a(0)/m_c(0)$	max|λ|
Connectance					
A1	9	0.20	0.2688	0.4602	1.7489
A2	9	0.25	0.3406	0.5046	2.0306
A3	9	0.41	0.6449	0.7161	3.8229
A4	9	0.62	0.8352	0.8642	5.6786
Pattern of Connectance					
A5	9	0.12	0.0413	0.3345	1.0851
A6	9	0.12	0.1024	0.3668	1.2388
A7	9	0.12	0.3211	0.4923	1.9498
A8	9	0.12	0.3333	0.5000	2.0000
Network Representation					
A9	9	0.25	0.3406	0.5046	2.0306
A10	9	0.25	0.3406	0.5046	2.0306
A11	9	0.25	0.3406	0.5046	2.0306
A12	9	0.25	0.3406	0.5046	2.0306

Matrix number corresponds to experimental networks shown in Figure 2.2

Table 2.2: Comparison of pathway proliferation rates in three ecosystem models

Model	n	C	$m_a(0)$	$m_c(1)$	$m_a(0)/m_c(0)$	$m_a(1)/m_c(1)$	max|λ|
Oyster	6	0.50	0.33	0.50	0.4750	0.6400	2.1479
Marsh	42	0.20	0.98	1.03	0.6095	0.6321	9.6170
Cove	33	0.30	1.02	1.06	0.6826	0.7027	10.4217

2.8 FIGURE LEGENDS

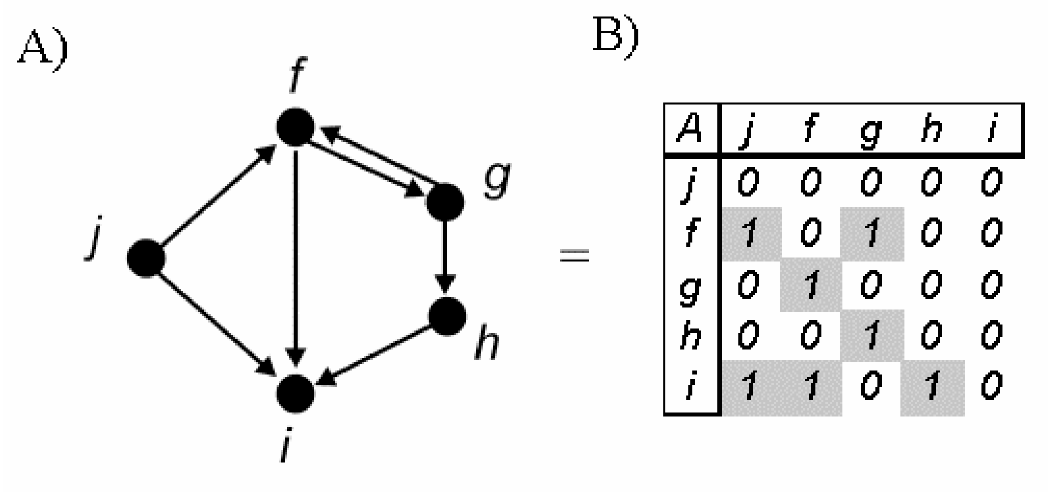
Figure 2.1: Example Network. Two isomorphic representations of an example network with $n = 5$ and $C = 0.28$. A) Digraph, B) Adjacency matrix. Note how the pattern of connections in the digraph is represented in the matrix.

Figure 2.2: Experimental Networks. Twelve experimental networks were used to investigate degree and pattern of connectance. A1–A4 are the same size and have increasing connectivity. A5–A8 have the same size and connectance, but the pattern of linkage is altered. In A9–A12 size, connectance, and pattern of connection remain constant but reordering the node vector alters the adjacency matrix representation.

Figure 2.3: Relationships between number of pathways (P_k) and pathway length (k) as network size (n) increases from 1 to 200. A) Selected networks to demonstrate relationships. Note how differences between paths, A(0), and walks, A(1), decrease as size increases. B) Using the slope (m_a) of the relationship to describe it, we see that size has a curvilinear affect on m_a.

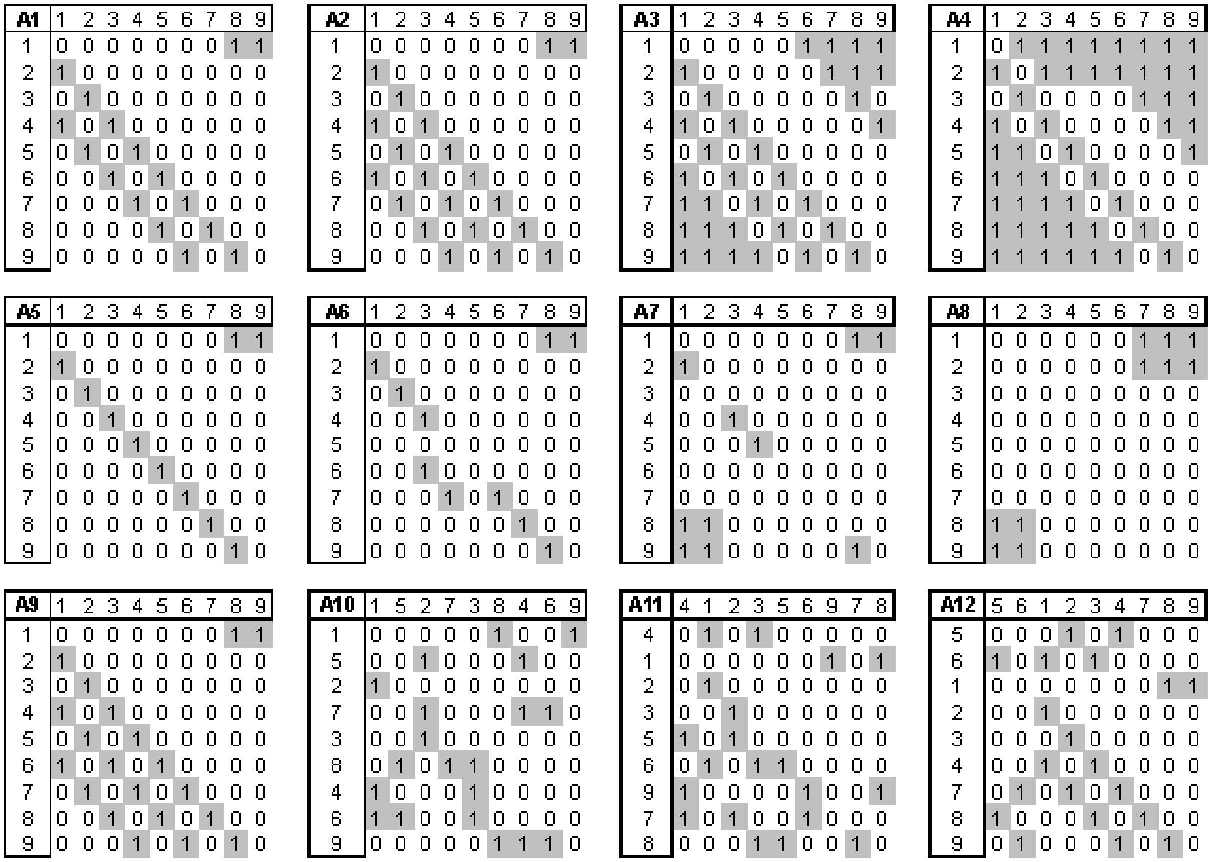
Figure 2.4: Structural comparison of three model ecosystems: A) oyster reef (Dame and Patten 1981), B) Okefenokee marsh (Whipple 1995) and C) Lake Texoma cove (Patten *et al.,* 1975).

Figure 2.1



A	j	f	g	h	i
j	0	0	0	0	0
f	1	0	1	0	0
g	0	1	0	0	0
h	0	0	1	0	0
i	1	1	0	1	0

Figure 2.2



A1	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	0	1	1
2	1	0	0	0	0	0	0	0	0
3	0	1	0	0	0	0	0	0	0
4	1	0	1	0	0	0	0	0	0
5	0	1	0	1	0	0	0	0	0
6	0	0	1	0	1	0	0	0	0
7	0	0	0	1	0	1	0	0	0
8	0	0	0	0	1	0	1	0	0
9	0	0	0	0	0	1	0	1	0

A2	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	0	1	1
2	1	0	0	0	0	0	0	0	0
3	0	1	0	0	0	0	0	0	0
4	1	0	1	0	0	0	0	0	0
5	0	1	0	1	0	0	0	0	0
6	1	0	1	0	1	0	0	0	0
7	0	1	0	1	0	1	0	0	0
8	0	0	1	0	1	0	1	0	0
9	0	0	0	1	0	1	0	1	0

A3	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	1	1	1	1
2	1	0	0	0	0	0	1	1	1
3	0	1	0	0	0	0	0	1	0
4	1	0	1	0	0	0	0	0	1
5	0	1	0	1	0	0	0	0	0
6	1	0	1	0	1	0	0	0	0
7	1	1	0	1	0	1	0	0	0
8	1	1	1	0	1	0	1	0	0
9	1	1	1	1	0	1	0	1	0

A4	1	2	3	4	5	6	7	8	9
1	0	1	1	1	1	1	1	1	1
2	1	0	1	1	1	1	1	1	1
3	0	1	0	0	0	0	1	1	1
4	1	0	1	0	0	0	0	1	1
5	1	1	0	1	0	0	0	0	1
6	1	1	1	0	1	0	0	0	0
7	1	1	1	1	0	1	0	0	0
8	1	1	1	1	1	0	1	0	0
9	1	1	1	1	1	1	0	1	0

A5	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	0	1	1
2	1	0	0	0	0	0	0	0	0
3	0	1	0	0	0	0	0	0	0
4	0	0	1	0	0	0	0	0	0
5	0	0	0	1	0	0	0	0	0
6	0	0	0	0	1	0	0	0	0
7	0	0	0	0	0	1	0	0	0
8	0	0	0	0	0	0	1	0	0
9	0	0	0	0	0	0	0	1	0

A6	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	0	1	1
2	1	0	0	0	0	0	0	0	0
3	0	1	0	0	0	0	0	0	0
4	0	0	1	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0
6	0	0	1	0	0	0	0	0	0
7	0	0	0	1	0	1	0	0	0
8	0	0	0	0	0	0	1	0	0
9	0	0	0	0	0	0	0	1	0

A7	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	0	1	1
2	1	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
4	0	0	1	0	0	0	0	0	0
5	0	0	0	1	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	1	1	0	0	0	0	0	0	0
9	1	1	0	0	0	0	0	1	0

A8	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	1	1	1
2	0	0	0	0	0	0	1	1	1
3	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	1	1	0	0	0	0	0	0	0
9	1	1	0	0	0	0	0	0	0

A9	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	0	1	1
2	1	0	0	0	0	0	0	0	0
3	0	1	0	0	0	0	0	0	0
4	1	0	1	0	0	0	0	0	0
5	0	1	0	1	0	0	0	0	0
6	1	0	1	0	1	0	0	0	0
7	0	1	0	1	0	1	0	0	0
8	0	0	1	0	1	0	1	0	0
9	0	0	0	1	0	1	0	1	0

A10	1	5	2	7	3	8	4	6	9
1	0	0	0	0	0	1	0	0	1
5	0	0	1	0	0	0	1	0	0
2	1	0	0	0	0	0	0	0	0
7	0	0	1	0	0	0	1	1	0
3	0	0	1	0	0	0	0	0	0
8	0	1	0	1	1	0	0	0	0
4	1	0	0	0	1	0	0	0	0
6	1	1	0	0	1	0	0	0	0
9	0	0	0	0	0	1	1	1	0

A11	4	1	2	3	5	6	9	7	8
4	0	1	0	1	0	0	0	0	0
1	0	0	0	0	0	0	1	0	1
2	0	1	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0
5	1	0	1	0	0	0	0	0	0
6	0	1	0	1	1	0	0	0	0
9	1	0	0	0	0	1	0	0	1
7	1	0	1	0	0	1	0	0	0
8	0	0	0	1	1	0	0	1	0

A12	5	6	1	2	3	4	7	8	9
5	0	0	0	1	0	1	0	0	0
6	1	0	1	0	1	0	0	0	0
1	0	0	0	0	0	0	0	1	1
2	0	0	1	0	0	0	0	0	0
3	0	0	0	1	0	0	0	0	0
4	0	0	1	0	1	0	0	0	0
7	0	1	0	1	0	1	0	0	0
8	1	0	0	0	1	0	1	0	0
9	0	1	0	0	0	1	0	1	0

Figure 2.3

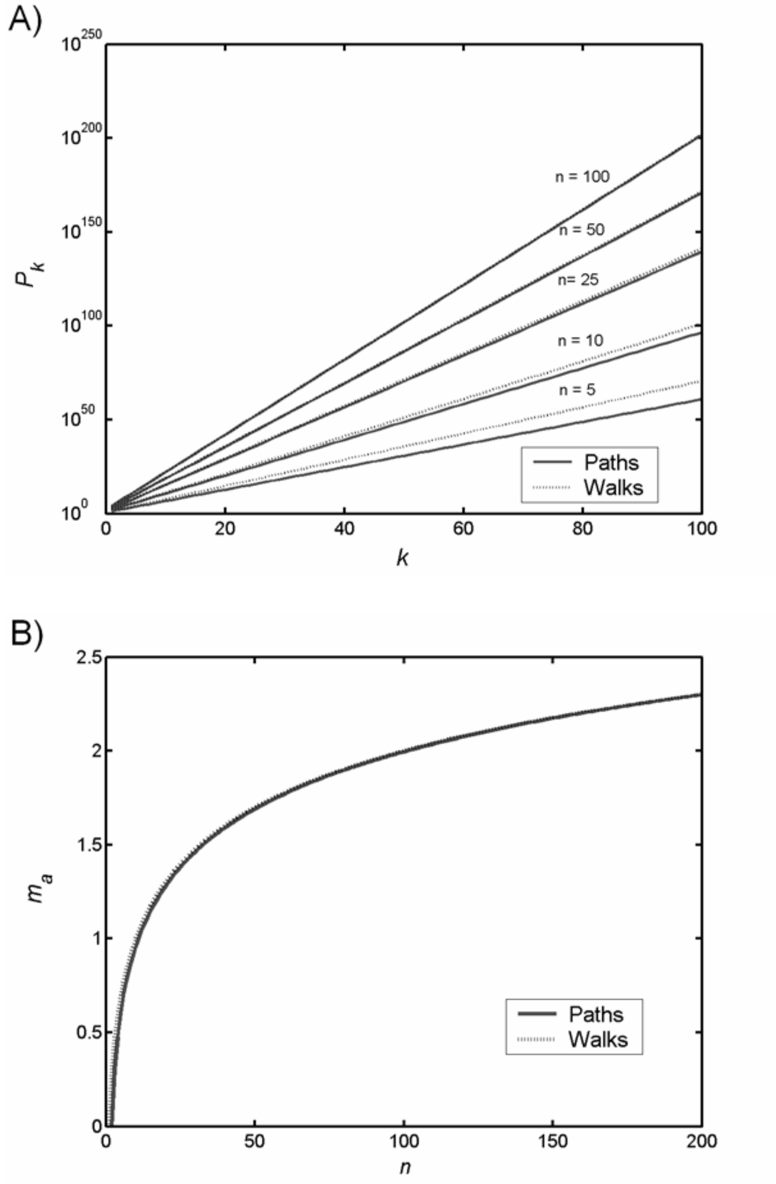
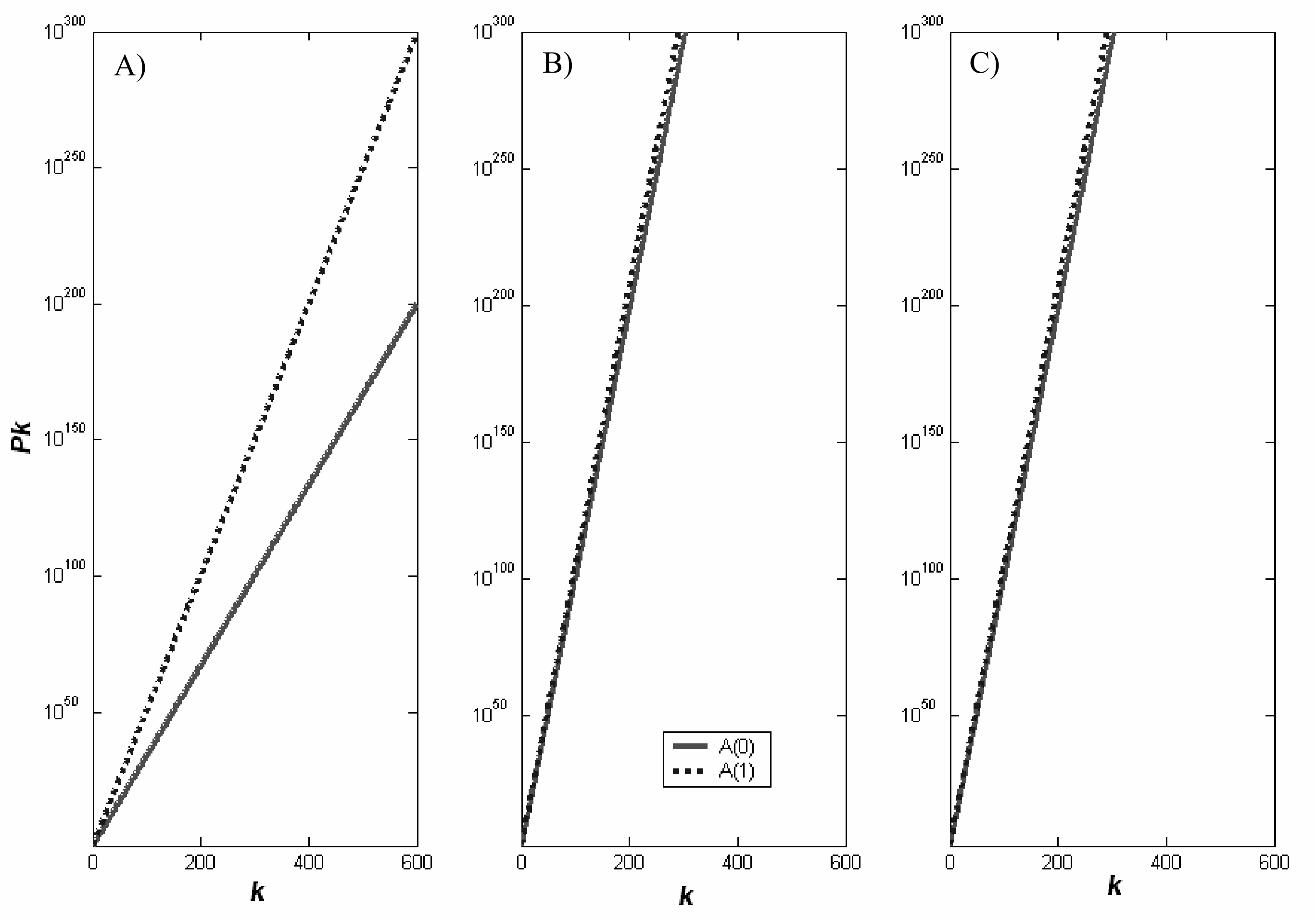

A)
10^{250}
10^{200}
10^{150}
10^{100}
10^{50}
10^{0}
P_k
n = 100
n = 50
n= 25
n = 10
n = 5
Paths
Walks
0
20
40
60
80
100
k
B)
2.5
2
1.5
1
0.5
0
m_a
Paths
Walks
0
50
100
150
200
n

Figure 2.4


A)
B)
C)
Pk
k
k
k
10^300
10^250
10^200
10^150
10^100
10^50
0
200
400
600
A(0)
A(1)

CHAPTER 3

PATHWAY PROLIFERATION AND MODULARITY IN ECOLOGICAL NETWORKS[1]

---

[1] Borrett, S.R, B.D. Fath, and B.C. Patten, To be submitted to *Journal of Theoretical Biology*

ABSTRACT

Large-scale structural patterns have been found in network models of complex systems including a skewed node degree distribution and small-world topology. These patterns suggest common organizational constraints and similar functional consequences. In this paper, we investigate a structural pattern termed *pathway proliferation*. Previous research enumerating pathways linking species determined that as pathway length increases, the number of pathways increases without bound at differing rates. Variable rates translate into important differences in the flow of energy, matter, and information in ecosystems. In this paper, we further clarify the pathway proliferation concept, describe factors influencing the node-node proliferation rate, and characterize pathway proliferation rates in 17 large empirical food-webs. We show that pathway proliferation 1) occurs if there is at least one cycle or feedback in the graph, 2) the proliferation rate, indicated by the dominant eigenvalue $\lambda_1(A)$, is identical among nodes within a *strongly connected component*, but can vary between components within a network, and 3) though topology can be a factor, the proliferation rate is heavily influenced by the number of nodes (compartments, species) and number of direct links (flows). Strongly connected components introduce a form of modularity in food-webs. Our analysis of 17 large empirical food-webs revealed that ten contained at least one non-trivial ($n>1$) strongly connected component. Thus, over half the food-webs contained at least one simple cycle, establishing feedback loops. The number of trophospecies in one or more a strongly connected component ranged from 2% to 60% of the total species. Pathway proliferation rates ranged from 2 to 10.25 (pathways per length) in these modules, and half appeared to be determined by trophospecies and trophic relation richness; food-web topology was not a factor. However, in half of the modules ecological processes appeared to have constructed structure uncommon in randomized

components. We conclude that the presence of strongly connected components and pathway proliferation in ecological networks reveals subgroups of species that will be functionally integrated by cyclic indirect effects.

3.1 INTRODUCTION

Large-scale structural patterns have been uncovered in network models of complex systems, suggesting the possibility of common organizational constraints and similar functional consequences. Network models are mathematical graphs composed of nodes and undirected edges or directed arcs that connect the nodes. For example, in a social network nodes might represent individuals in a community and the edges or links could represent a social relationship between the individuals such as collaboration (Newman, 2001a, b). In a model of the World Wide Web, web pages are nodes connected by hyperlinks (Albert et al., 1999; Barabási and Albert, 1999). Ecologists use network models in many ways, including representing trophic relations in food-webs and more generally energy–matter flux in ecosystems. In these networks, species or functional groups form the node set while the presence of energy and matter transfers and transformations are represented by links.

Traditionally, complex systems have been modeled using random graphs (Erdös and Rényi, 1959, 1960; Gardner and Ashby, 1970; May, 1972). Ecologists recognize, however, that random graphs are inadequate models of ecological systems (DeAngelis, 1975; Lawler, 1978); food-web and ecosystem models often contain structures not commonly found in random graphs (DeAngelis, 1975; Pimm, 1979, 1982; Cohen et al., 1990; Pimm et al., 1991). In food-webs, these hypothesized structures have included short food-chain lengths (Pimm and Lawton, 1977; Post, 2002) and little or no cycling (Cohen et al., 1990). Several forms of modularity (hierarchic compartmentalization into subsystems) have been hypothesized for food-webs and ecosystems (May, 1972; Pimm, 1979; Pimm and Lawton, 1980; Allen and Starr, 1982; Yodzis, 1982). Ulanowicz and Wolff (1991) demonstrated that random networks (based on the Poisson, uniform, Gaussian, negative exponential, and log-normal probability distribution functions) were inadequate representations of real ecosystems. The organization of ecological systems is shaped and constrained by thermodynamic laws and natural history (Lawler, 1978; Jørgensen et al., 1992; Müller and Leupelt, 1998; Williams and Martinez, 2000).

Likewise, investigations of other types of complex systems have recently identified a number of distinctive patterns common in complex systems not found in purely random graphs (Albert and Barabási, 2002; Newman, 2002, 2003). For example, the distribution of node degree (i.e., the number of edges (links) incident to a node) is often skewed in models of complex systems, following an exponential distribution or a power-law distribution rather than the Poisson distribution of random graphs. The power-law distribution was found in the World Wide Web (Barabási and Albert, 1999), metabolic networks (Jeong et al., 2000), and some but not all food-webs (Dunne et al., 2002b; Montoya and Sole, 2002). The power-law degree distribution implies that there are a large number of nodes with very few connections, while a few nodes have a large number of connections (Barabási, 2002). This topology tends to increase network robustness to random node or edge deletion, while making it more sensitive to targeted attacks (Albert et al., 2000; Dunne et al., 2002b). The small-world pattern is another commonly found topology (Watts and Strogatz, 1998; Watts, 1999). In small-world networks, the degree of node clustering is higher and the maximum distance (where distance is the shortest path between two nodes) is lower than expected from random graphs. This arrangement tends to increase the transmission speed of diseases, energy, matter, and information through networks. The largest distance in food-web graphs tends to be small, but the degree of clustering varies (Dunne et al., 2002a).

Pathway proliferation is another large-scale topological characteristic of networks, with implications for energy, matter, and information transmission. It is the tendency for the number of pathways in a network to increase geometrically without bound as pathway length increases. Patten and colleagues (Patten et al., 1982; Patten, 1985a, b) first observed this tendency in small, well-connected ecosystem models during the early development of ecosystem network analysis. The rate of pathway proliferation is variable among networks (Fath, 1998; Borrett and Patten, 2003). This is significant because the pathway proliferation rate characterizes how quickly the number of indirect pathways increases, and thus, the number of pathways available for interactions. Often food-web investigations emphasize the shortest pathway, assuming that most significant interactions occur over this route (e.g. Caldarelli et al., 1998; Post et al., 2000). However, previous results from Network Environ Analysis, an environmental application and extension of economic Input-Output Analysis, indicate that flows over longer indirect pathways can be significant or even dominant constituents of total system throughflow (Patten, 1983; Higashi and Patten, 1986, 1989). This has important implications for trophodynamics (Patten et al., 1990; Burns et al., 1991; Whipple, 1998) and biogeochemical cycling in ecosystems (Borrett et al., submitted, see Chapter 4). Given the possible significance of indirect pathways in network models of conservative transport systems like ecosystems, it is critical to understand the network characteristics influencing the pathway proliferation rate.

In this paper, we clarify the pathway proliferation concept, describe factors influencing the proliferation rate, and characterize pathway proliferation rates in 17 large empirical food-webs. In Section 3.2 we review relevant mathematics to build a better understanding of pathway proliferation. In Section 3.3 we apply this understanding to 17 food-web models drawn from the literature. This analysis reveals a type of modularity in some of the food-webs, lending support to the hypothesis that food-webs have a modular structure (May, 1972; Pimm and Lawton, 1980;

Yodzis, 1982; Krause et al., 2003). We conclude by summarizing our findings and suggesting next steps along this research path.

3.2 PATHWAY PROLIFERATION

Though pathway proliferation was introduced into the ecological literature over two decades ago (Patten et al., 1982; Patten, 1985a, b), it is not well understood. Here, we synthesize mathematical results from graph theory and matrix algebra to 1) determine a method for quantifying the node-node pathway proliferation rate, 2) identify the bounds and expected value of the rate, and 3) uncover the possibility of differing rates of pathway proliferation for modules within a network.

Network models of complex systems are mathematically graphs which can be directed or weighted (Ponstein, 1966; Bang-Jensen and Gutin, 2001). A graph G is specified by a set of n nodes and e unoriented edges ($0 \leq e \leq \frac{n(n-1)}{2} + n$), where edges indicate an undirected relationship between two nodes. A directed graph (digraph) D is also specified by a set of n nodes, but instead of edges it has L oriented arcs or links ($0 \leq L \leq n^2$). Edges and links can be assigned weights to represent the relationship strength. In this paper, we focus on simple unweighted digraphs, where simple implies no more than one link from one node to any other. We do this for several reasons. First, directed graphs are often appropriate for ecological applications as many ecological processes are oriented (e.g., predation and excretion generate energy and matter flows from one ecosystem element to another). Second, while the mathematics described in this paper may apply to non-simple and weighted graphs, our interest here is primarily network structure as it is a necessary if not sufficient element of understanding ecosystem organization. Digraph structure is partially characterized by two connectivity measures, connectance $C = L/n^2$ and link density L/n. These are common metrics in food-web literature (Cohen et al., 1990; Martinez, 1991).

3.2.1 QUANTIFYING PATHWAY PROLIFERATION

In a directed graph, a pathway is an alternating sequence of nodes and links connecting a starting and terminal node. Pathway length m is the number of links in the pathway. For example, in the directed graph D in Figure 3.1A there is a pathway of length 2 from node 1 to node 3 (e.g., 1 → 2 → 3). Cycles are pathways with the same starting and terminal nodes, and a cycle of length one is a self loop. In our example network, 1 → 2 → 3 → 1 is cycle of length three, and 4 → 4 is a self loop. Pathways with self loops are termed walks, those without are paths (Patten, 1985a). D can alternatively be represented by its associated and isomorphic adjacency matrix $A_{n\times n} = (a_{ij})$, where $a_{ij} = 1$ if there is a link from j to i (note column to row orientation), otherwise $a_{ij} = 0$ (3.1B). The number of direct links terminating or starting at a node is termed the in-degree and out-degree, respectively. These are calculated as $k^{in} = (k_i^{in}) = \sum_{j=1}^{n} a_{ij}$ and $k^{out} = (k_j^{out}) = \sum_{i=1}^{n} a_{ij}$, where k^{in} and k^{out} are $1 \times n$ and $n \times 1$ vectors respectively. Average in- and out-degrees ($<k^{in}>$ or $<k^{out}>$), and degree distributions P(k) are ways of characterizing network structure (Albert and Barabási, 2002; Newman, 2003).

Indirect pathways (m>1) are enumerated by raising the adjacency matrix to the mth power $A^m = (a_{ij})^m$ (Ponstein, 1966). Again, pathway proliferation is the tendency for the number of pathways in a network to increase without bound as a function of increasing pathway length (Figure 3.1C). Borrett and Patten (2003, see Chapter 2) approximated whole system pathway proliferation rate as $\left(\frac{\log\left(\sum\sum a_{ij}^{(m)}\right)}{\log\left(\sum\sum a_{ij}^{(m-1)}\right)}\right)$ where $a_{ij}^{(m)}$ signifies the a_{ij} element of A^m. This estimate holds only for sufficiently large m. Therefore, following Fath (1998), here we will define pathway proliferation in terms the rate at which $a_{ij}^{(m)}$ changes as m → ∞, making pathway proliferation a combinatorial property of A (Seneta, 1973).

Development of A^m as m increases is a combinatorial property of A determined by its characteristic polynomial, $\pi(\lambda) = \det(\lambda I - A)$, where λ is the variable in the polynomial and $I_{n\times n}$ is the identity matrix (Seneta, 1973; Godsil, 1993). The eigenvalues of A, λ_i ($i = 1,\ldots,n$), are roots of the characteristic polynomial, determined as solutions to $\pi(\lambda) = 0$. The set of eigenvalues $\{\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_i \geq \ldots \geq \lambda_n\}$ is termed the spectrum of A, which is studied to deduce structural properties of graphs (Cvetkoviâc et al., 1980). For undirected graphs, A is symmetric and all λ_i are real; in directed graphs λ_i may be complex. Further, these eigenvalues must satisfy $AW = \Lambda W$, where $W_{n\times n} = [w_1|w_2|\ldots|w_i|\ldots|w_n]$ is a composite matrix of the right eigenvectors w_i, and $\Lambda_{n\times n} = \mathrm{diag}(\lambda_i)$ is a matrix with the eigenvalues of A on the principle diagonal and zeros in all other positions. The eigenvalues must also satisfy $VA = V\Lambda$, where $V_{n\times n} = [v_1|v_2|\ldots|v_i|\ldots|v_n]$ is a composite matrix of the left eigenvectors v_i ($i = 1,\ldots,n$). A^m is then determined as

$$A^m = W * \Lambda^m * W^{-1}$$

$$A^m = W * \begin{bmatrix} \lambda_1^m & 0 & \cdots & 0 \\ 0 & \lambda_2^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & \lambda_n^m \end{bmatrix} * W^{-1}, \tag{3.1}$$

where the columns of W are the right eigenvectors of A, and W^{-1} is the matrix inverse of W, which are the transposed complex conjugates of the left eigenvectors V (Caswell, 2001). If we let v_i^* be the i^{th} row of W^{-1}, then we can rewrite (3.1) as

$$A^m = \lambda_1^m w_1 v_1^* + \lambda_2^m w_2 v_2^* + \lambda_3^m w_3 v_3^* + \ldots + \lambda_i^m w_i v_i^* + \ldots + \lambda_n^m w_n v_n^* . \qquad (3.2)$$

This equation is the pathway generating function of A (Godsil, 1993), and illustrates how the development of A^m as m increases depends on the spectrum of A.

Digraphs and matrices can be classified using two different schemes. We introduce these classifications here because they are essential to the application of a theorem that will allow us later to develop a more succinct estimate of the rate of pathway proliferation. First, digraphs can be classified as one of three types: strongly connected, weakly connected, and disconnected. A digraph is strongly connected (strong) if it is possible to reach every node from every other over a pathway of unspecified length. Bang-Jensen and Gutin (2001) define a (sub)digraph with only one node as strong, though this seems trivial for our purposes. A weakly connected (weak) digraph is one in which it is possible to reach any node from any other node if we ignore link orientation, but every node cannot be reached from every other node when following link orientation. Nodes of a weak digraph must have an in-degree or out-degree of at least 1. A disconnected graph is one that contains one or more non-adjacent strong or weak components.

Second, adjacency matrices associated with digraphs are also of different types. All adjacency matrices are non-negative because all $a_{ij} \geq 0$. Adjacency matrices are irreducible if and only if they are associated with strong digraphs, while those associated with weakly connected and disconnected digraphs are reducible (Berman and Plemmons, 1979). Irreducible matrices are further divided into two classes: primitive and imprimitive. A primitive matrix is an irreducible matrix that becomes positive ($a_{ij} > 0$, for all i,j) when raised to a sufficiently large power (Seneta, 1973).

An induced subdigraph of D is a subset of nodes in D with all links that both start and terminate on the node subset. A maximally induced subdigraph is the largest induced subdigraph that is strong. Weak and disconnected digraphs with associated reducible matrices are decomposable into a unique set of maximally induced strongly connected subdigraphs, termed *strongly connected components* (K_α, $\alpha \leq n$) (Bang-Jensen and Gutin, 2001). This implies that there is at least one simple cycle (no repeated medial nodes) that connects all the nodes in a non-trivial strongly connected component. Adjacency matrices associated with K_α are irreducible. Our example digraph (Figure 3.1A) is disconnected, but contains two connected subdigraphs {1, 2, 3, 4} and {5, 6, 7}. It can be partitioned into five strongly connected components K_1 = {1, 2, 3}, K_2 = {4}, K_3 = {5}, K_4 = {6}, K_5 = {7}. Only K_1 is non-trivial. The adjacency matrix associated with each strongly connected component is irreducible; the adjacency matrices associated with K_1 and K_2 are primitive.

The Perron-Frobenius theorem guarantees there is one real eigenvalue equal or larger than all other eigenvalues, $\lambda_1(A) \geq \lambda_i(A)$ (i = 2,…,n) in irreducible matrices (Seneta, 1973; Berman and Plemmons, 1979). In the literature $\lambda_1(A)$ is alternately referred to as the dominant eigenvalue, the Perron eigenvalue, and the spectral radius.

As shown by Caswell (2001), we can divide both sides of (3.2) by $\lambda_1(A)$,

$$\frac{A^m}{\lambda_1^m} = w_1 v_1 + \frac{\lambda_2^m}{\lambda_1^m} w_2 v_2 + \frac{\lambda_3^m}{\lambda_1^m} w_3 v_3 + \ldots + \frac{\lambda_n^m}{\lambda_1^m} w_n v_n . \tag{3.3}$$

If A is primitive and irreducible, then $\lambda_1(A)$ is strictly larger than $\|\lambda_i\|$ for all i>1, where $\|\bullet\|$ is the norm of $\bullet$ (this is necessary since λ_i may be complex). Taking the limit of both sides of (3.3) as pathway length increases, we find that

$$\lim_{m\to\infty} \frac{A^m}{\lambda_1^m} = w_1 v_1 . \tag{3.4}$$

Thus, while smaller eigenvalues will influence pathway proliferation over shorter path lengths, as path length increases the pathway proliferation rate asymptotically becomes $\lambda_1(A)$: $\frac{A^{m+1}}{A^m} \to \lambda_1(A)$ as m → ∞ (Hill, unpublished ms.). Therefore, $\lambda_1(A)$ is the asymptotic rate of pathway proliferation in a strongly connected graph with a primitive adjacency matrix. The damping ratio $\rho = \frac{\lambda_1(A)}{\|\lambda_2(A)\|}$ characterizes the rate of convergence to $\lambda_1(A)$ (Caswell, 2001).

In strong digraphs with an imprimitive irreducible adjacency matrix there are $d \leq n$ eigenvalues with the same absolute magnitude, and one or more may be complex (Seneta, 1973). The Perron-Frobenius theorem then indicates that the common absolute magnitude of the d eigenvalues will be larger than the other n–d eigenvalues. Thus, as m → ∞ only the d largest eigenvalues will influence pathway proliferation. Caswell (2001) reports that these digraphs generate oscillatory dynamics.

Each strong component K_α of a weak or disconnected digraph will have its own independent rate of pathway proliferation, $\lambda_1(K_\alpha)$ (read "λ_1 of K_α"). Trivial strongly connected components (those with only one node) will have a pathway proliferation rate of 0 or unity depending on whether or not it has a self-loop. The eigenvalues of a reducible matrix are the union set of the eigenvalues of the adjacency matrices associated with strongly connected components (Jain and Krishna, 2003). Thus, the maximum dominant eigenvalue of the strongly connected components will be the dominant eigenvalue of the whole digraph. Further, if a digraph is composed of only trivial strongly connected components without self-loops, pathway proliferation will not occur. This is true of all acyclic digraphs. This suggests an application of the dominant eigenvalue: to detect the presence of cycles (Jain and Krishna, 2003):

1. if $\lambda_1(A) = 0$, then A has no cycles;
2. if $\lambda_1(A) = 1$, then A has at least one cycle and all cycles occur in strongly connected components that have only one simple cycle; and
3. if $\lambda_1(A) > 1$, then A has more than one simple cycle.

Based on an independent development, Fath (1998) interpreted similar results as three classes of feedback: 1) no feedback, 2) simple feedback, and 3) cyclic feedback in strongly connected networks. Notice that a graph with $\lambda_1(A) \geq 1$ could have a reducible or irreducible adjacency matrix, while the adjacency matrix of a graph with $\lambda_1(A) = 0$ is necessarily reducible with n trivial strongly connected components.

Similarly, we can summarize three possibilities for our interpretation of the dominant eigenvalue as a measure of pathway proliferation in digraphs:

1. $\lambda_1(A) = 0$: as $m \rightarrow \infty$ the number of pathways between two nodes declines to zero;
2. $\lambda_1(A) = 1$: as $m \rightarrow \infty$ the number of pathways between nodes in a strongly connected component remains constant; and
3. $\lambda_1(A) > 1$: as $m \rightarrow \infty$ the number of pathways between nodes in at least one strongly connected component (K_α) increases without bound at an asymptotic rate equal to $\lambda_1(K_\alpha)$ where $\max(\lambda_1(K_\alpha)) = \lambda_1(A)$.

3.2.2 BOUNDS AND EXPECTED VALUES OF PATHWAY PROLIFERATION

Given that $\lambda_1(A)$ is the asymptotic rate of pathway proliferation in strongly connected digraphs, it would be useful to know its theoretical bounds and expected value. Matrix theory bounds the dominant eigenvalue of a non-negative matrix by the minimum and maximum column (row) sum, which in the context of directed graphs is the minimum and maximum out-degree (in-degree) where equality holds only if $k^{in} = k^{out}$ (Seneta, 1973; Berman and Plemmons, 1979). Thus, $\max\left(\min(k_j^{in}), \min(k_j^{out})\right) \leq \lambda_1(A) \leq \min\left(\max(k_j^{in}), \max(k_j^{out})\right)$. In a strongly connected digraph with more than one node, all nodes must have at least one and a maximum of n incoming and outgoing links. Therefore, $1 \leq \lambda_1(A) \leq n$ for a strong digraph. As stated previously, a trivial component with no self loops will have $\lambda_1(K_\alpha) = 0$. A complete graph will have $\lambda_1(A) = n$ (allowing self loops). Notice that in the binary matrix A, $\lambda_1(A)$ cannot take values between 0 and 1.

What is the expected value of $\lambda_1(A)$? In undirected random graphs G with $A = (a_{ij})$ where $a_{ij} = a_{ji} = 1$ with probability p ($0 < p < 1$) and $a_{ij} = a_{ji} = 0$ with probability ($1–p$), Juhász proved that $\lim_{n\to\infty} \frac{\lambda_1(A)}{n} = p$ with probability 1 (Cvetkovic and Rowlinson, 1990). This implies that $\lambda_1(A) \sim n^*p$ in the limit of large n; therefore, $\lambda_1(A)$ is an approximation of the number of undirected edges in A. While a random graph is not necessarily connected, a result by Erdős and Rényi (1959; 1960) shows that the fraction of nodes connected in a single component increases rapidly when the average link density exceeds unity. In contrast to random graphs with a Poisson degree distribution, $\lambda_1(A)$ in random graphs with a power-law distribution of node degrees increases with $\sim n^{1/4}$ (Farkas et al., 2001; Goh et al., 2001), and may be further modified by introducing a hierarchic modularity to the network topology (de Aguiar and Bar-Yam, 2005). While random graphs are well-studied, properties of random digraphs are less well-known. Similar to random graphs, the in-degree and out-degree of random digraphs is known to have a Poisson distribution, and when link density is greater than unity the expected size of the largest strongly connected component increases rapidly (Karp, 1990; Luczak, 1990; Barbosa et al., 2003). We are unaware of known results about the spectra of random directed graphs. We numerically verified that $\lambda_1(A) \sim n*p$ remains plausible for random directed graphs by determining the largest eigenvalue in an ensemble of 99,000 random digraphs (50 from each combination of n = {2, 3, …,100} and p={0.05, 0.10,0.15,…,1}). Our results indicate that $\lambda_1(A) \sim n*C = L/n$, where $C = L/n^2$ is an estimate of p (Figure 3.2). As either L increases or n decreases the residual error decreases. We conclude that in random digraphs as in undirected random graphs, $\lambda_1(A)$ is largely determined by the combination of the number of nodes and number of direct connections; pattern of connections has a minor influence. In digraphs with a more structured topology – such as those with power-law in-degree or out-degree distributions or modularity – we might expect $\lambda_1(A)$ to deviate from L/n as it does in undirected graphs, though this remains to be explored.

In ecological networks where n is the number of species (functional groups, etc.) and L is the number of direct transactions, the rate of pathway proliferation will be heavily influenced by species richness and direct link abundance. However, the results of Farkas et al. (2001) and Goh et al. (2001) suggest that if the degree distributions are skewed, as has been demonstrated for some food-webs (Dunne et al., 2002b; Montoya and Sole, 2002; Williams et al., 2002), or the networks contain other types of order such as modularity, then the residual error $d = |\lambda_1(A) - L/n|$ may be larger than expected from random graphs.

In summary, pathway proliferation defined as the tendency for the number of pathways between any two nodes to increase geometrically without bound as pathway length increases 1) occurs if there is at least one cycle or feedback in the graph (i.e., does not occur in acyclic networks), 2) the proliferation rate, indicated by the dominant eigenvalue $\lambda_1(A)$, is identical among nodes within a strongly connected component, but can vary between strongly connected components within a network, and 3) though topology can be a factor, the proliferation rate is heavily influenced by the number of nodes and number of direct links. In addition, networks composed of α non-trivial strongly connected components K_i ($i = 1,\ldots, \alpha$) have a form of structural modularity that may be functionally significant to the complex systems being modeled.

3.3 MODULARITY AND PATHWAY PROLIFERATION IN EXAMPLE FOOD-WEBS

In this section we build on our conceptual and mathematical developments by applying the theory to determine the pathway proliferation rates in 17 of the largest empirical food-webs currently available (Table 3.1). Five of the food-webs are terrestrial including Coachella Valley (Polis, 1991), St. Martin Island Caribbean Island (Goldwasser and Roughgarden, 1993), the El Verde tropical rain forest (Waide and Reagan, 1996), a grassland in the United Kingdom (Martinez et al., 1999), and a food-web associated with the Scotch Broom plant (Memmott et al., 2000). Three food-webs are from freshwater habitats: Skipworth Pond (Warren, 1989), Bridge Brook Lake (Havens, 1992), Little Rock Lake (Martinez, 1991), Canton Creek, and Stony Stream (Townsend et al., 1998). Seven of the food-webs represent marine or oceanic habitats: Chesapeake Bay (Baird and Ulanowicz, 1989), St. Marks Estuary (Christian and Luczkovich, 1999), Ythan Estuary without (1991) and with (1996) its parasite community (Hall and Raffaelli, 1991; Huxham et al., 1996), the Benguela current (Yodzis, 1998), a small Caribbean reef (Opitz, 1996), and the northeastern United States continental shelf (Link, 2002). Following accepted protocol, original food-webs were modified such that any species or functional group with identical predators and prey were grouped as a "trophic species" to reduce methodological bias in the data (Yodzis, 1982; Cohen et al., 1990; Pimm et al., 1991; Yodzis and Winemiller, 1999). These food-webs have been the subject of much recent network analysis (Williams and Martinez, 2000; Dunne et al., 2002b; Dunne et al., 2002a; Williams et al., 2002; Dunne et al., 2004) which has previously reported their number of trophospecies or nodes n, connectance $C = L/n^2$, the proportion of basal species %B ($k_i^{in} = 0$), proportion of intermediate species %I ($k_i^{in}, k_i^{out} > 0$), proportion of top species %T ($k_i^{out} = 0$), and link density L/n. We reproduce this basic network information in Table 3.1 for comparison. In addition, Dunne et al. (2002b; 2002a) showed that several have skewed degree distributions (i.e., power-law, exponential, etc).

We first identified, enumerated, and characterized all strongly connected components (K_α), including their rates of pathway proliferation and damping ratio, in these food-webs. We envisioned three possible outcomes. If food-webs were adequately modeled by random digraphs then we would expect each web to have one giant strongly connected component encompassing most if not all of the nodes with a single pathway proliferation rate close to link density $\lambda_1(A) \sim L/n$. This seemed unlikely given the known skewed degree distributions and evidence that ecological processes construct non-random topologies (Cohen et al., 1990; Williams and Martinez, 2000), despite arguments to the contrary (Kenny and Loehle, 1991). A second possibility is based on the observation that most early food-webs were acyclic (Cohen et al., 1990). Thus, these food-webs might be acyclic, containing no non-trivial strongly connected components. In this case, pathway proliferation would not occur and $\lambda_1(A) = 0$. A final possibility is that the food-webs would tend to have one or more strongly connected components and multiple pathway proliferation rates. This outcome would support the hypothesized modularity of ecological systems which is thought to increase system stability (May, 1972; Pimm and Lawton, 1980; Yodzis, 1982; Krause et al., 2003).

In Section 3.2 we hypothesized that the absolute difference between the dominant eigenvalue and its expected value in random digraphs (L/n) might be a useful indicator of the significance of network topology. To assess this hypothesis, we used Monte Carlo simulations to determine if $d = |\lambda_1(A) - L/n|$ was larger than expected. We had two scales of analysis: whole food-web and non-trivial strongly connected components. For both, we constructed 1001 uniform random digraphs with n nodes and where each possible link was connected with probability p equal to the original network's connectivity ($p = C$). We assessed statistical significance by determining the fraction of random digraphs in which d was equal or greater than observed in our network of interest, Pr(d). Assuming a significance level of $\alpha = 0.05$, $\Pr(d) < 0.05$ implies d is statistically significant.

When applied to the entire food-web, a significant difference with the null model implies that topological factors beyond species and link richness are significant in determining the whole system dominant eigenvalue. This could be the size or frequency of strongly connected components within the network, as suggested by the analysis in Section 3.2, or perhaps a skewed degree distribution. If the food-webs had more than one non-trivial component, we expected the deviation to be large.

As for the whole food-web, a significant deviation of d indicates the significance of network topology in the strongly connected components. However, given that they are irreducible by definition, a significant deviation of d within a component must indicate the significance of another element of topology, such as the degree distribution.

All analyses were conducted using MATLAB 6.5 (The MathWorks, Inc., Release 13).

3.3.1 RESULTS

Food-webs included in this study range in size from 25 to 155 trophospecies and 3% to 32% connectance (Table 3.1). They tend to have a large proportion of intermediate species (i.e., those with $k_j^{in} > 0$ and $k_j^{out} > 0$), although the two stream food-webs are notable exceptions. Ten of the 17 food-webs examined contained at least one non-trivial strongly connected component; six had two. Five of the remaining food-webs had a dominant eigenvalue of unity, implying that at least one node contained a self-loop. Our results reveal that the majority of these food-webs have at least one directed cycle, contrary to earlier food-web theory (Cohen et al., 1990).

While the majority of the food-webs express strongly connected component modularity, the degree of species involvement varies. In food-webs that have a non-trivial strongly connected component, the proportion of the original nodes involved ranges from 2% in the two Ythan Estuary food-webs to 60% in the Caribbean reef model. Recall that the definition of a strongly connected component excludes nodes that have no inputs or no outputs. This effectively excludes all basal ($k_j^{in} > 0$) and top consumer ($k_j^{out} > 0$) species from strongly connected components such that the number of species in strongly connected components is limited by the number of intermediate species. This may be a factor in why the two stream food-webs and the Scotch Broom food-web contain no non-trivial components.

The absolute difference between the dominant eigenvalue of the entire food-web and its expected value based on random digraphs of the appropriate size and connectance ($d = |\lambda_1(A) - L/n|$) ranged from a minimum of 0.5 for the El Verde rainforest to a maximum of 12.9 for the NE US shelf food-web (Table 3.1). In all cases, this difference was significantly different from the random digraph null model, indicating that topology is a significant factor in determining $\lambda_1(A)$. This result is consistent with the presence of one or more small non-trivial components and acyclic digraphs. Our results provide another line of evidence suggesting that the ecological processes that create food-webs lead to more ordered network topologies; random digraphs are not good models for these systems.

Inspection of the strongly connected components reveals a diversity of topologies (Table 3.2). The largest strongly connected component, with seventy trophospecies occurs in the El Verde rainforest model; although it is the least well connected (13%), it still has the largest rate of pathway proliferation ($\lambda_1(A) = 10.25$). In contrast, ten of the sixteen strongly connected components have four or fewer species. Five of the strongly connected components only contain two trophospecies, requiring a single simple cycle of path length 2 (e.g., $j \rightarrow i \rightarrow j$). While the two strongly connected components in the Coachella Valley, Skipworth pond, and Benguela are about the same size, one of the two components in Little Rock Lake, Caribbean reef, and NE US shelf is substantially larger than the other. Table 3.3 lists the trophospecies in the two strongly connected components of the Coachella Valley.

All strongly connected component dominant eigenvalues have a multiplicity of one, so the adjacency matrices associated with the component subdigraphs are primitive. Therefore, the dominant eigenvalues represent the strongly connected component asymptotic rates of pathway proliferation. These range from 1.62 in strongly connected components of the two Ythan Estuary food-webs to 10.25 in the large El Verde rainforest strongly connected component and generally increase with link density as would be expected in random digraphs. However, half of the strongly connected components have a statistically significant difference between the dominant eigenvalue and link density including Coachella Valley (K_1 and K_2), El Verde Rainforest (K_1), Little Rock Lake (K_1), Benguela (K_2), Carribian Reef (K_2), and NE US shelf (K_1

and K_2). The topological arrangement of species and links in these three strongly connected components influences their rate of pathway proliferation; the others are largely determined by their species and link richness.

Again, the damping ratio is an index of the speed of convergence to the asymptotic rate of pathway proliferation; a larger ratio indicates faster convergence. Five of the strongly connected components are completely connected. They have a pathway proliferation rate equal to their trophospecies richness and an undefined damping ratio because their second eigenvalues are zero. In these cases the asymptotic rate of pathway proliferation is achieved instantaneously. The other damping ratios range from 1.69 in K_2 of the NE US Shelf to 4.41 in K_1 of the same food-web. Transient dynamics of the pathway proliferation rate, determined by the smaller eigenvalues, are more influential in NE US shelf (K_2). Its pathway proliferation rate does not converge until a pathway length of nearly twenty-two, while in NE US shelf (K_1) the proliferation rate converges by a pathway length of eight.

In summary, the majority of the food-webs we examined contained at least one non-trivial strongly connected component. Six food-webs had two non-trivial strongly connected components; none had more than two. The proportion of species involved in a strongly connected component ranged from 2% to 60%. In all cases, the difference between the dominant eigenvalue of the food-web and the expected value (L/n) in a random network was significant. This difference occurs because the topology of food-webs is non-random; thermodynamic processes and species characteristics (e.g., the species niche) combine to form non-random structure (Chase and Leibold, 2003). Within the strongly connected components, the rate of pathway proliferation ranged from 1.62 to 10.25 and half were indistinguishable from random graphs based on the expected rate of pathway proliferation.

3.4 DISCUSSION

As with any analysis of network models that reveals previously hidden structural patterns, we are left with two basic questions. First, what, if any, significance do these patterns hold for our systems of interest? Are strongly connected components and pathway proliferation simply another pretty pattern, another network or food-web statistic to report, or do they impart some functional significance? The second question cannot be divorced from the first; what system processes might create these structural patterns? Are there ecological processes or forces that might lead to the development of these structures? These are not easy questions to answer, but in this section we attempt to address them for the presence of strongly connected components and pathway proliferation in ecological networks.

Strongly connected components introduce a form of modularity into network models, where modularity is defined as a hierarchical system subdivision into more or less interacting subsystems. Several types of modularity have been proposed in ecological systems. May (1972) hypothesized that ecosystems would have a modular structure based on his investigation of Lyapunov stability in randomly assembled ecosystems. He found that when a system's species were partitioned into blocks of interacting species with few if any connections to other blocks greater overall system stability was more likely. Pimm (1979) termed these blocks of species "compartments", stating that they are "…characterized by strong interactions within compartments, but weak interactions among compartments" (p. 145). Pimm and Lawton (1980) concluded from a study of binary empirical food-webs that, while there was evidence species were grouped into subsystems largely by habitat, compartmentalization as defined by Pimm was an uncommon phenomenon. They noted, however, that a complete test of the hypothesis would require knowledge of the strength of interactions, which was absent in their food-web models.

Yodzis (1982) remarked that modular organization was an old idea in ecology, providing the guild concept (Root, 1967) as an example. He applied the *dominant clique* idea from graph theory to identify another type of modularity in food-webs. He defined "…a clique as a set of species in a given ecosystem with the property that every pair in the set has some food resource in common, and … a dominant clique as a clique which is contained in no other clique" (p. 552). More recently, Krause et al. (2003) used a methodology developed to identify cohesive subgroups in social network analysis to classify another type of modularity in food-webs. They demonstrated that this type of organization increased system stability to species deletion by localizing the effect within a module. The dominant ecological hypothesis is that food-web or ecosystem modularity increases overall system stability by localizing interactions within modules. Given the static, binary, presence-absence information of food-webs in our study we were unable to meaningfully test this hypothesis; stability is inherently a dynamic concept. Known issues with food-web model construction further make this hypothesis difficult to resolve (Cohen et al., 1993; Pollis and Strong, 1996). Empirical food-web models usually do not indicate interaction strength or the temporal and spatial variation of the interactions. These details are expensive to acquire.

Despite the challenge of assessing their effect on system stability, strongly connected components in ecological networks appear to be important functional elements of system organization and provide new insights about species participating in them. By definition (Section 3.2), there is minimally one simple cycle that encircles all nodes in the strongly connected component. This provides at least one channel for cybernetic feedback (positive or negative) to propagate among species in the module (Patten and Odum, 1981; DeAngelis et al., 1986). In food-webs it is reasonable to assume that each predatory species directly benefits by its consumption of prey. This establishes an indirect mutualism spanning the strongly connected component. It also provides the necessary conditions for the strongly connected components of food-webs to function as autocatalytic cycles – systems that catalyze their own production (Maynard Smith and Szathmâry, 1995; Ulanowicz, 1997). Autocatalytic cycles are an essential element of metabolism in chemical and living systems and may have played a role in the origin of life (Maynard Smith and Szathmâry, 1995). Maynard Smith and Szathmâry (1995) described autocatalytic cycles as a force for cooperation among the member species and efficient information integrators. Ulanowicz (1997) identifies several emergent properties autocatalytic cycles may possess, including centripetality, persistence, and autonomy. Centripetality is the tendency of the cycle to attract more of the energy–matter flux of the system. If any member species becomes more efficient at using its resources or better able to acquire new resources such that its population increases, this positive change tends to cascade through the module, collectively benefiting the populations of all species involved. Autocatalytic cycles tend to persist because their general form can be maintained in a system despite fluctuations in the interaction strengths and possible element replacement. In food-webs this implies that when a trophically similar species is introduced to the system, if it is more efficient or in some way ecologically more competitive, it may wholly replace an existing species in the autocatalytic cycle, but the cycle remains. Finally, autocatalytic cycles can establish a degree of autonomy because species in the cycle can actively influence at least a portion of their input environment. In this sense, species in strongly connected components of food-webs are involved in ecosystem engineering (Jones et al., 1997) and niche construction (Laland et al., 1999; Odling-Smee et al., 2003). Ulanowicz (1986; 1997) further argues that the autocatalytic nature of cycles in ecosystems makes them a principal agent in ecosystem growth and development.

Maynard Smith and Szathmâry (1995) remarked that autocatalytic cycles are sensitive to cheaters or parasites that feed off the strongly connected component without participating in the cycle. Top predators feeding on species in a strongly connected component or a downstream strongly connected component might function as parasites in this sense. Perhaps this is why strongly connected components do not occur in all the food-webs analyzed. This and the tendency for centripetality may explain why there are fewer than three strongly connected components in these food webs. We are unable to assess these possibilities with these data, however, because the differences may simply reflect disparities in food-web modeling decisions.

Pathway proliferation rates of strongly connected components provide additional information about the module. Each additional link in a strongly connected component beyond those that form the defining cycle introduces another embedded simple cycle. This lowers the maximal distance between nodes in the module, increases the potential pathways for energy–matter flux, tends to increase the rate of pathway proliferation, and leads to the unbounded growth of pathways as length increases. In some cases, the rate of pathway proliferation will not increase as expected with the number of links. For example, half of the strongly connected components identified in our food-webs had pathway proliferation rates that were significantly different from the expected rate based on random graphs with a Poisson degree distribution. This suggests that module topology differs from what we would expect from a random generating process. As mentioned earlier, we would not necessarily expect the ecological processes structuring food-webs to be random. Species characteristics such as metabolic requirements, food preferences, capture ability and handling time, as well as other niche requirements and natural history constraints will proximately influence the choices of "who eats whom" and how much. The emergent properties of autocatalytic cycles and ecosystems more generally may also provide whole-system constraints. Perhaps the interesting question is not, why the eight strongly connected compartments did not match the random expectation, but why the other half did? Notice that the strongly connected compartments with non-random topologies were the largest modules, while the eight strongly connected components with topologies indistinguishable from random digraphs involved only two or three trophospecies. The universe of possible topologies is much smaller in these small and well-connected ($0.0.75 \leq C \leq 1.0$) modules, making the ecologically created topologies reflected in the food-webs more likely. Five of these modules were completely connected, generating only one possible arrangement pattern. The eight strongly connected components with apparently non-random topologies were less well connected ($0.13 \leq C \leq 0.78$), generating a much larger universe of possible topologies. In six cases the pathway proliferation rate was significantly less than expected, but in the large strongly connected component of the El Verde rainforest and the smaller module of the Caribbean Reef, the pathway proliferation rate was more than expected. At this point, we are unable to provide a satisfactory explanation for these differences.

In our discussion thus far, we have been interchanging food-webs and ecosystems. It is important to recognize, however, that food-webs are a subset of a broader class of ecosystem models of energy–matter flow. Food-webs are generally defined by the relation "who eats whom" that is one process generating energy–matter flux, while ecosystem flow–storage models typically trace a conserved flow unit (e.g., energy, nitrogen, phosphorus) through the system, regardless of the process producing the flow. Thus, non-trophic ecological processes such as excretion and death are captured in flow–storage models, revealing a different picture of ecosystem organization.

Pathway proliferation influences the development and significance of indirect flows in ecosystem flow–storage models. Indirect flows are an important aspect of the ecological significance of the strongly connected components, so here we take a closer look. Indirect flows are derived from two fundamentally distinct types of pathways: chains (e.g., 5 → 6 → 7 in Figure 3.1A) and cycles (e.g., 1 → 3 → 1 in Figure 3.1A). Indirect flows in chains are limited by transfer efficiencies and chain length. In cycles, the number and length of pathways are unlimited such that indirect flows are only limited by transfer efficiencies reflecting energy–matter dissipation and export. As ecosystems are open thermodynamic systems, shorter indirect pathways individually will tend to carry larger indirect flows than longer indirect pathways. A faster rate of pathway proliferation $\lambda_1(A)$ implies that there will be more shorter indirect pathways, increasing the possibility that the magnitude of indirect flows will surpass that of direct flows. Thus, within a strongly connected component the large number of indirect pathways will tend to carry a large fraction of the flow between species (nodes).

More generally, $\lambda_1(A)$ indicates the potential for direct and indirect energy, matter, and information transmission between compartments in a strongly connected component. Realized transmission rates are dependent on the realized use of each pathway. Previous ecosystem network analyses reveal some of the system-level consequences of differential pathway use (Patten, 1985a; Ulanowicz, 1986; Higashi and Patten, 1989; Fath and Patten, 1998; Fath and Patten, 1999; Fath, 2004), but there is much left to learn about this subject. The interplay of this potential and realized network structure is an interesting, important topic for understanding the organization and transformation of complex adaptive systems like ecosystems.

We conclude that the strongly connected components and pathway proliferation are ecologically relevant phenomena because they provide novel insights about the system of interest. Without knowing the strength of interactions or energy–matter flux rates, the presence of these structural features suggests groups of species functionally integrated by indirect effects mediated by autocatalytic cycles. They portend the possibility of integral species relationships that are shifted toward more positive associations and the possibility of the dominance of indirect flows. In some cases, apparent negative interactions such as predation or competition may become more positive through indirect interactions mediated by the autocatalytic cycles of the strongly connected components.

3.5 ACKNOWLEDGEMENTS

We are grateful to Jennifer Dunne for providing the food-web data, and discussions with Stuart Whipple and the University of Georgia Systems and Engineering Ecology group that enhanced this work. The manuscript benefited from reviews by reviews by M.B. Beck, H.R. Pulliam, P.G. Verity, D.K. Gattie, and S. Bata. SRB was supported by National Science Foundation biocomplexity grant (OPP-00-83381) to Skidaway Institute of Oceanography.

3.6 REFERENCES

Albert, R., and Barabási, A.-L., 2002. Statistical mechanics of complex networks. Rev. Mod. Phys., **74**:47-97.

Albert, R., Jeong, H., and Barabasi, A.L., 1999. Diameter of the World-Wide Web. Nature, **401**:130-131.

Albert, R., Jeong, H., and Barabasi, A.L., 2000. Error and attack tolerance of complex networks. Nature, **406**:378-382.

Allen, T.F.H., and Starr, T.B., 1982. Hierarchy: perspectives for ecological complexity. University of Chicago Press, Chicago.

Baird, D., and Ulanowicz, R.E., 1989. The seasonal dynamics of the Chesapeake Bay ecosystem. Ecol. Monogr., **59**:329-364.

Bang-Jensen, J., and Gutin, G., 2001. Digraphs: theory, algorithms, and applications. Springer, London.

Barabási, A.-L., 2002. Linked: the new science of networks. Perseus, Cambridge, Mass.

Barabási, A.-L., and Albert, R., 1999. Emergence of scaling in random networks. Science, **286**:509-512.

Barbosa, V.C., Donangelo, R., and Souza, S.R., 2003. Directed cycles and related structures in random graphs: I - Static properties. Physica A, **321**:381-397.

Berman, A., and Plemmons, R.J., 1979. Nonnegative matrices in the mathematical sciences. Academic Press, New York.

Borrett, S.R., and Patten, B.C., 2003. Structure of pathways in ecological networks: relationships between length and number. Ecol. Model., **170**:173-184.

Borrett, S.R., Whipple, S.J., and Patten, B.C., submitted. Temporal variability of indirect effects in a seven-compartment model of nitrogen flow in the Neuse River Estuary (USA): time series analysis. Ecological Modelling.

Burns, T.P., Higashi, M., Wainright, S.C., and Patten, B.C., 1991. Trophic unfolding of a continental shelf energy-flow network. Ecol. Model., **55**:1-26.

Caldarelli, G., Higgs, P.G., and McKane, A.J., 1998. Modelling coevolution in multispecies communities. J. Theor. Biol., **193**:345-358.

Caswell, H., 2001. Matrix population models: construction, analysis, and interpretation, 2nd edition. Sinauer Associates, Sunderland, Mass.

Chase, J.M., and Leibold, M.A., 2003. Ecological niches: linking classical and contemporary approaches. University of Chicago Press, Chicago.

Christian, R.R., and Luczkovich, J.J., 1999. Organizing and understanding a winter's seagrass foodweb network through effective trophic levels. Ecol. Model., **117**:99-124.

Cohen, J.E., Beaver, R.A., Cousins, S.H., Deangelis, D.L., Goldwasser, L., Heong, K.L., Holt, R.D., Kohn, A.J., Lawton, J.H., Martinez, N., Omalley, R., Page, L.M., Patten, B.C., Pimm, S.L., Polis, G.A., Rejmanek, M., Schoener, T.W., Schoenly, K., Sprules, W.G., Teal, J.M., Ulanowicz, R.E., Warren, P.H., Wilbur, H.M., and Yodzis, P., 1993. Improving Food Webs. Ecology, **74**:252-258.

Cohen, J.E., Briand, F., and Newman, C.M., 1990. Community food webs: data and theory. Springer-Verlag, Berlin ; New York.

Cvetkoviâc, D.M., Doob, M., and Sachs, H., 1980. Spectra of graphs: theory and application. Academic Press, New York.

Cvetkovic, D., and Rowlinson, P., 1990. The largest eigenvalue of a graph: a survey. Linear and Multilinear Algebra, **28**:3-33.

de Aguiar, M.A.M., and Bar-Yam, Y., 2005. Spectral analysis and the dynamic response of complex networks. Physical Review E, **71**:016106.

DeAngelis, D.L., 1975. Stability and connectance in food web models. Ecology, **56**:238-243.

DeAngelis, D.L., Post, W.M., and Travis, C.C., 1986. Positive feedback in natural systems. Springer-Verlag, Berlin; New York.

Dunne, J.A., Williams, R.J., and Martinez, N.D., 2002a. Food-web structure and network theory: The role of connectance and size. Proc. Natl. Acad. Sci. USA, **99**:12917-12922.

Dunne, J.A., Williams, R.J., and Martinez, N.D., 2002b. Network topology and biodiversity loss in food webs: robustness increases with connectance. Ecol. Lett., **5**:558-567.

Dunne, J.A., Williams, R.J., and Martinez, N.D., 2004. Network structure and robustness of marine food webs. Mar. Ecol. Progr. Ser., **273**:291-302.

Erdös, P., and Rényi, A., 1959. On random graphs. Publicationes mathematicae, **6**:290-297.

Erdös, P., and Rényi, A., 1960. On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Sciences, **5**:17-61.

Farkas, I., Derényi, I., Barabási, A.-L., and Vicsek, T., 2001. Spectra of "real-world" graphs: Beyond the semicircle law. Physical Review E, **64**:026704.

Fath, B.D., 1998. Network analysis: foundations, extensions, and applications of a systems theory of the environment. Ph.D. Dissertation. University of Georgia, Athens, Georgia.

Fath, B.D., 2004. Network analysis applied to large-scale cyber-ecosystems. Ecol. Model., **171**:329-337.

Fath, B.D., and Patten, B.C., 1998. Network synergism: emergence of positive relations in ecological systems. Ecol. Model., **107**:127-143.

Fath, B.D., and Patten, B.C., 1999. Quantifying resource homogenization using network flow analysis. Ecol. Model., **107**:193-205.

Gardner, M.R., and Ashby, W.R., 1970. Connectance of large dynamical (cybernetic) systems: Critical values for stability. Nature, **228**:784.

Godsil, C.D., 1993. Algebraic combinatorics. Chapman & Hall, New York.

Goh, K.I., Kahng, B., and Kim, D., 2001. Spectra and eigenvectors of scale-free networks. Physical Review E, **6405**.

Goldwasser, L., and Roughgarden, J., 1993. Construction and analysis of a large Caribbean food web. Ecology, **74**:1216-1233.

Hall, S.J., and Raffaelli, D.G., 1991. Food web patterns: lessons from a species rich web. Journal of Animal Ecology, **60**:823-841.

Havens, K., 1992. Scale and structure in natural food webs. Science, **257**:1107-1109.

Higashi, M., and Patten, B.C., 1986. Further aspects of the analysis of indirect effects in ecosystems. Ecol. Model., **31**:69-77.

Higashi, M., and Patten, B.C., 1989. Dominance of indirect causality in ecosystems. Am. Nat., **133**:288-302.

Hill, J.V., unpublished ms. On the largest real root of a polynomial.

Huxham, M., Beaney, S., and Raffaelli, D., 1996. Do parasites reduce the chances of triangulation in a real food web? Oikos, **76**:284-300.

Jain, S., and Krishna, S., 2003. Graph theory and the evolution of autocatalytic networks. In: S. Bornholdt and H.G. Schuster (Editors), Handbook of graphs and networks: from the genome to the internet, Wiley-VCH, Berlin, pp 355-395.

Jeong, H., Tomnor, B., Oltvai, Z.N., and Barabási, A.-L., 2000. The large-scale organization of metabolic networks. Nature, **407**:651-654.

Jones, C.G., Lawton, J.H., and Shachak, M., 1997. Positive and negative effects of organisms as physical ecosystem engineers. Ecology, **78**:1946-1957.

Jørgensen, S.E., Patten, B.C., and Straškraba, M., 1992. Ecosystems emerging: toward an ecology of complex systems in a complex future. Ecol. Model., **62**:1-27.

Karp, R.M., 1990. The transitive closure of a random digraph. Random Structures & Algorithms, **1**:73-93.

Kenny, D., and Loehle, C., 1991. Are food webs randomly connected. Ecology, **72**:1794-1799.

Krause, A.E., Frank, K.A., Mason, D.M., Ulanowicz, R.E., and Taylor, W.W., 2003. Compartments revealed in food-web structure. Nature, **426**:282-285.

Laland, K.N., Odling-Smee, F.J., and Feldman, M.W., 1999. Evolutionary consequences of niche construction and their implications for ecology. Proc. Natl. Acad. Sci. USA, **96**:10242-10247.

Lawler, L.R., 1978. A comment on randomly constructed model ecosystems. Am. Nat., **112**:445-447.

Link, J., 2002. Does food web theory work for marine ecosystems? Mar. Ecol.-Prog. Ser., **230**:1-9.

Luczak, T., 1990. The phase transition in the evolution of random digraphs. Journal of Graph Theory, **14**:217-223.

Martinez, N.D., 1991. Artifacts or attributes - effects of resolution on the Little Rock Lake food web. Ecol. Monogr., **61**:367-392.

Martinez, N.D., Hawkins, B.A., Dawah, H.A., and Feifarek, B.P., 1999. Effects of sampling effort on characterization of food-web structure. Ecology, **80**:1044-1055.

May, R., 1972. Will a large complex system be stable. Nature, **238**:413-414.

Maynard Smith, J., and Szathmâry, E., 1995. The major transitions in evolution. W.H. Freeman Spektrum, Oxford ; New York.

Memmott, J., Martinez, N.D., and Cohen, J.E., 2000. Predators, parasitoids and pathogens: species richness, trophic generality and body sizes in a natural food web. Journal of Animal Ecology, **69**:1-15.

Montoya, J.M., and Sole, R.V., 2002. Small world patterns in food webs. J. Theor. Biol., **214**:405-412.

Müller, F., and Leupelt, M., 1998. Eco targets, goal functions, and orientors. Springer, Berlin; New York.

Newman, M.E.J., 2001a. Scientific collaboration networks. I. Network construction and fundamental results - art. no. 016131. Physical Review E, **6401**:6131-+.

Newman, M.E.J., 2001b. The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. USA, **98**:404-409.

Newman, M.E.J., 2002. Random graphs as models of networks. In: S. Bornholdt and H.G. Schuster (Editors), Handbook of graphs and networks, Wiley-VCH, Berlin

Newman, M.E.J., 2003. The structure and function of complex networks. Siam Review, **45**:167-256.

Odling-Smee, F.J., Laland, K.N., and Feldman, M.W., 2003. Niche construction: the neglected process in evolution. Princeton University Press, Princeton, N.J.

Opitz, S., 1996. Trophic interactions in Caribbean coral reefs. Tech Report 43, ICLARM.

Patten, B.C., 1983. On the quantitative dominance of indirect effects in ecosystems. In: W.K. Lauenroth, G.V. Skogerboe, and M. Flug (Editors), Analysis of ecological systems: state-of-the-art in ecological modelling, Elsevier, Amsterdam, pp 27-37.

Patten, B.C., 1985a. Energy cycling in the ecosystem. Ecol. Model., **28**:1-71.

Patten, B.C., 1985b. Energy cycling, length of food chains, and direct versus indirect effects in ecosystems. Can. Bull. Fish. Aqu. Sci., **213**:119-138.

Patten, B.C., Higashi, M., and Burns, T.P., 1990. Trophic dynamics in ecosystem networks: significance of cycles and storage. Ecol. Model., **51**:1-28.

Patten, B.C., and Odum, E.P., 1981. The cybernetic nature of ecosystems. Am. Nat., **118**:886-895.

Patten, B.C., Richardson, T.H., and Barber, M.C., 1982. Path analysis of a reservoir ecosystem model. Can. Water Resourc. J., **7**:252-282.

Pimm, S.L., 1979. The structure of food webs. Theoretical Population Biology, **16**:144-158.

Pimm, S.L., 1982. Food webs. Chapman and Hall, London; New York.

Pimm, S.L., and Lawton, J.H., 1977. On feeding on more than one trophic level. Nature, **275**:542-544.

Pimm, S.L., and Lawton, J.H., 1980. Are food webs divided into compartments? Journal of Animal Ecology, **49**:879-898.

Pimm, S.L., Lawton, J.H., and Cohen, J.E., 1991. Food web patterns and their consequences. Nature, **350**:669-674.

Polis, G.A., 1991. Complex trophic interactions in desserts: an empirical critique of food-web theory. Am. Nat., **138**:123-155.

Pollis, G.A., and Strong, D.R., 1996. Food web complexity and community dynamics. Am. Nat., **157**:813-846.

Ponstein, J., 1966. Matrices in graph and network theory. Van Gorcum & Comp., Assen,.

Post, D.M., 2002. The long and short of food-chain length. Trends Ecol. Evol., **17**:269-277.

Post, D.M., Pace, M.L., and Hairston, N.G., 2000. Ecosystem size determines food-chain length in lakes. Nature, **405**:1047-1049.

Root, R.B., 1967. The niche exploitation pattern of the blue-gray gnatcatcher. Ecol. Monogr., **37**:317-350.

Seneta, E., 1973. Non-negative matrices; an introduction to theory and applications. Wiley, New York,.

Townsend, C.R., Thompson, R.M., McIntosh, A.R., Kilroy, C., Edwards, E., and Scarsbrook, M.R., 1998. Disturbance, resource supply, and food-web architecture in streams. Ecol. Lett., **1**:200-209.

Ulanowicz, R.E., 1986. Growth and development: ecosystems phenomenology. Springer-Verlag, New York.

Ulanowicz, R.E., 1997. Ecology, the ascendent perspective. Columbia University Press, New York.

Ulanowicz, R.E., and Wolff, W.F., 1991. Ecosystem flow networks - loaded dice? Math. Biosci., **103**:45-68.

Waide, R.B., and Reagan, W.B., 1996. The food web of a tropical rainforest. University of Chicago Press, Chicago, IL.

Warren, P.H., 1989. Spatial and temporal variation in the structure of a fresh-water food web. Oikos, **55**:299-311.

Watts, D.J., 1999. Small worlds: the dynamics of networks between order and randomness. Princeton University Press, Princeton, N.J.

Watts, D.J., and Strogatz, S.H., 1998. Collective dynamics of "small-world" networks. Nature, **393**:440-442.

Whipple, S.J., 1998. Path-based network unfolding: A solution for the problem of mixed trophic and non-trophic processes in trophic dynamic analysis. J. Theor. Biol., **190**:263-276.

Williams, R.J., Berlow, E.L., Dunne, J.A., Barabasi, A.-L., and Martinez, N.D., 2002. Two degrees of separation in complex food webs. Proc. Natl. Acad. Sci. USA, **99**:12913-12916.

Williams, R.J., and Martinez, N.D., 2000. Simple rules yield complex food webs. Nature, **404**:180-183.

Yodzis, P., 1982. The compartmentation of real and assembled ecosystems. Am. Nat., **120**:551-570.

Yodzis, P., 1998. Local trophodynamics and the interaction of marine mammals and fisheries in the Benguela ecosystem. Journal Of Animal Ecology, **67**:635-658.

Yodzis, P., and Winemiller, K.O., 1999. In search of operational trophospecies in a tropical aquatic food web. Oikos, **87**:327-340.

3.7 TABLES

Table 3.1: Topological properties of 17 empirical food-webs

Habitat	**Food Web**	**Original Reference**	**Taxa**†	**n**†	**C**†	**%B**†	**%I**†	**%T**†	**L/n**†	**λ_1(A)**	**d**	**Pr(d)**	**#K**	**%K**
Terrestrial	Coachella Valley	Polis 1991	30	29	0.31	10	90	0	9.03	6.35	2.7	0.001 *	2	59
	St. Martin Island	Goldwasser & Roughgarden 1993	44	42	0.12	14	69	17	4.88	0.00	4.9	0.001 *	0	0
	El Verde Rainforest	Waide & Reagen 1996	156	155	0.06	18	69	13	9.74	10.25	0.5	0.001 *	1	45
	UK Grassland	Martinez et al. 1999	75	61	0.03	18	69	13	1.59	0.00	1.6	0.001 *	0	0
	Scotch Broom	Memmott et al. 2000	154	85	0.03	1	40	59	2.62	1.00	1.6	0.001 *	0	0
Lake/Pond	Skipworth Pond	Warren 1989	35	25	0.32	4	92	4	7.88	3.00	4.9	0.001 *	2	20
	Bridge Brook Lake	Havens 1992	75	25	0.17	32	68	0	4.28	2.00	2.3	0.001 *	1	8
	Little Rock Lake	Martinez 1991	181	92	0.12	13	86	1	10.84	6.20	4.6	0.001 *	2	26
Stream	Canton Creek	Townsend et al. 1998	108	102	0.07	53	22	25	6.83	1.00	5.8	0.001 *	0	0
	Stony Stream	Townsend et al. 1998	112	109	0.07	56	27	17	7.61	1.00	6.6	0.001 *	0	0
Estuary	Chesapeake Bay	Baird & Ulanowicz 1989	33	31	0.07	16	52	32	2.19	1.00	1.2	0.001 *	0	0
	St. Marks Estuary	Christian & Luczkovich 1999	48	48	0.10	10	80	10	4.60	1.00	3.6	0.001 *	0	0
	Ythan Estuary, 1991	Hall & Raffaelli 1991	92	83	0.06	9	54	37	4.76	1.62	3.1	0.001 *	1	2
	Ythan Estuary, 1996	Huxham et al. 1996	134	124	0.04	4	56	40	4.67	1.62	3.1	0.001 *	1	2
Marine	Benguela	Yodzis 1998	29	29	0.24	7	93	0	7.00	3.00	4.0	0.001 *	2	21
	Carribean Reef, small	Opitz 1996	50	50	0.22	6	94	0	11.12	8.63	2.5	0.001 *	2	60
	NE US Shelf	Link 2002	81	79	0.22	3	94	4	17.76	4.87	12.9	0.001 *	2	39

Taxa refers to the original number of species; n is the number of nodes or trophospecies; C = L/n^2 is connectance; %B, %I and % T are the proportions of basal (indegree = 0), intermediate (indegree & outdegree > 0), and top (outdegree = 0) trophospecies; L/n is link density; $\lambda_1(A)$ is the dominant eigenvalue of the entire foodweb; $d = |\ \lambda_1(A) - L/n|$; Pr(d) is the fraction of an ensemble of 1001 random digraphs in which d is greater than or equal to that observed in K_α; * indicates statistically significant Pr(d) at $\alpha = 0.05$; #K is the number of non-trivial strongly connected components; %K is the percent of species in a non-trivial strongly connected component. † marks topological properties previously reported for these food-webs (Williams and Martinez, 2000; Dunne et al., 2002a; Dunne et al., 2004).

Table 3.2: Topological properties of strongly connected components in 17 empirical food webs

Model	**K_a**	**n**	**L**	**C**	**L/n**	**$\lambda_1(K_a)$**	**mult(λ_1)**	**r**	**d**	**Pr(d)**
Coachella Valley	1	11	71	0.59	6.45	6.35	1	3.40	0.1	0.001 *
	2	6	22	0.61	3.67	3.56	1	3.56	0.1	0.001 *
El Verde Rainforest	1	70	633	0.13	9.04	10.25	1	2.43	1.2	0.001 *
Skipworth Pond	1	3	9	1.00	3.00	3.00	1	--	0.0	1.001
	2	2	4	1.00	2.00	2.00	1	--	0.0	1.001
Bridge Brook Lake	1	2	4	1.00	2.00	2.00	1	--	0.0	1.001
Little Rock Lake	1	21	167	0.38	7.95	6.20	1	2.42	1.7	0.001 *
	2	3	9	1.00	3.00	3.00	1	--	0.0	1.001
Ythan Estuary, 1991	1	2	3	0.75	1.50	1.62	1	2.62	0.1	0.212
Ythan Estuary, 1996	1	2	3	0.75	1.50	1.62	1	2.62	0.1	0.226
Benguela	1	3	9	1.00	3.00	3.00	1	--	0.0	1.001
	2	3	7	0.78	2.33	2.25	1	4.05	0.1	0.007 *
Carribean Reef, small	1	2	3	0.75	1.50	1.62	1	2.62	0.1	0.217
	2	28	244	0.31	8.71	8.63	1	4.11	0.1	0.001 *
NE US Shelf	1	4	11	0.69	2.75	2.88	1	4.41	0.1	0.003 *
	2	27	243	0.33	9.00	4.87	1	1.69	4.1	0.001 *

K_α indicates the non-trivial strongly connected component number; n is the number of nodes (trophospecies), L is the number of links; $C = L/n^2$ is connectance, L/n is the link density, $\lambda_1(K_\alpha)$ is the dominant eigenvalue of K_α, mult(λ_1) is the multiplicity of the dominant eigenvalue, $\rho = \frac{\lambda_1(K_\alpha)}{|\lambda_2(K_\alpha)|}$ is the damping ratio (-- indicates ρ is undefined because $|\lambda_2(K_\alpha)| = 0$), $d = |\lambda_1(A) - L/n|$ is the absolute difference between the dominant eigenvalue and link density, Pr(d) is the fraction of an ensemble of 1001 random digraphs in which d is greater than or equal to that observed in K_α, and * indicates statistically significant Pr(d) at $\alpha = 0.05$.

Table 3.3: Trophospecies in the two non-trivial strongly connected components of Coachella Valley food-web

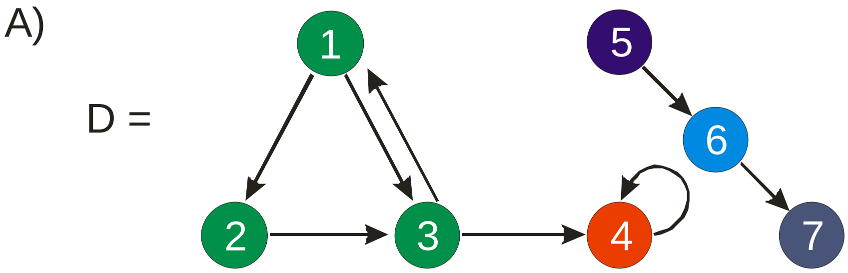
	K_1		K_2
n	**Trophospecies**	**n**	**Trophospecies**
21	primarily herbivorous mammals and birds	11	small arthropod predators
22	small omnivorous mammals and birds	12	medium arthropod predators
26	primarily carnivorous lizards	13	large arthropod preadators
27	primarily carnivorous snakes	14	facultative arthropod predators
28	large primarily predacious birds	15	life-history arthropod omnivore
29	large primarily predacious mammals	16	spider parasitoids
		17	primary parasitoids
		18	hyperparisitoids
		23	predacious mammals and birds
		24	arthropodivorous snakes
		25	primarily arthropodivorous lizards

3.8 FIGURE LEGENDS

Figure 3.1: A) Example digraph D with strongly connected components $K_1 = \{1,2,3\}$, $K_2 = \{4\}$, $K_3 = \{5\}$, $K_4 = \{6\}$, $K_5 = \{7\}$, B) adjacency matrix associated with the digraph D (dotted lines denote strongly connected components), and C) plot demonstrating pathway proliferation from node 3 to 1, as the number of pathways $a_{3,1}^{(m)}$ increases as pathway length m increases. Transient effects created by smaller eigenvalues are visible in the first ten pathway lengths, but the rate of pathway proliferation has nearly converged to $\lambda_1(K_1) = 1.32$ by a pathway length of 16.

Figure 3.2: Dominant eigenvalue in random digraphs. Points show the relationship between dominant eigenvalues $\lambda_1(A)$ and link density L/n in 99,000 uniform random digraphs where $a_{ij} = 1$ with probability p and $a_{ij} = 0$ with probability (1–p) (50 replicates of each combination of n = {2,3,…,100} and p = {0.05, 0.10, …, 1}). The line indicates the expected $\lambda_1(A) = L/n$ relationship. The inset graph shows the distribution of residuals indicating that as L/n increases it becomes a better predictor of $\lambda_1(A)$.

Figure 3.1



B)

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

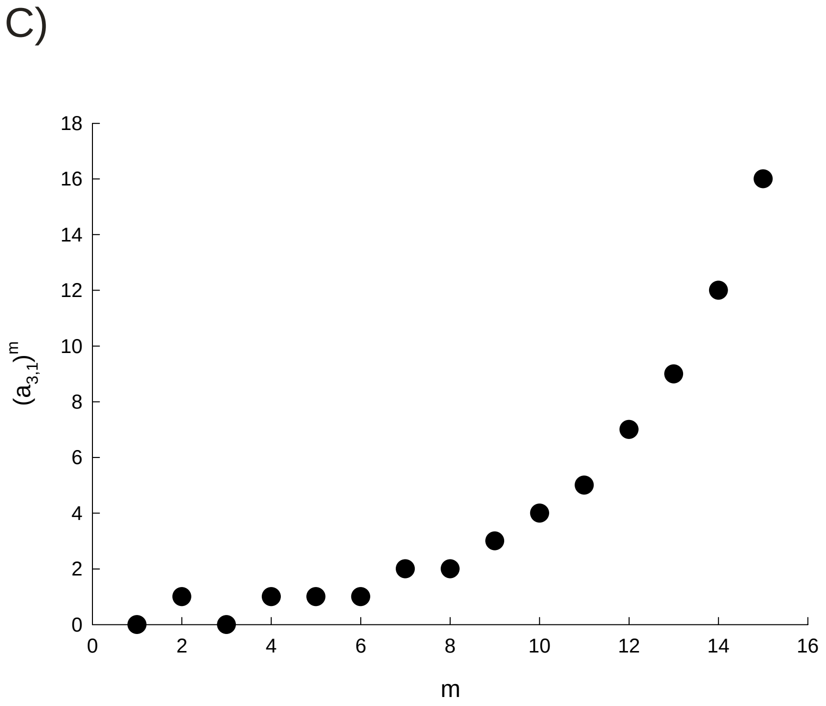
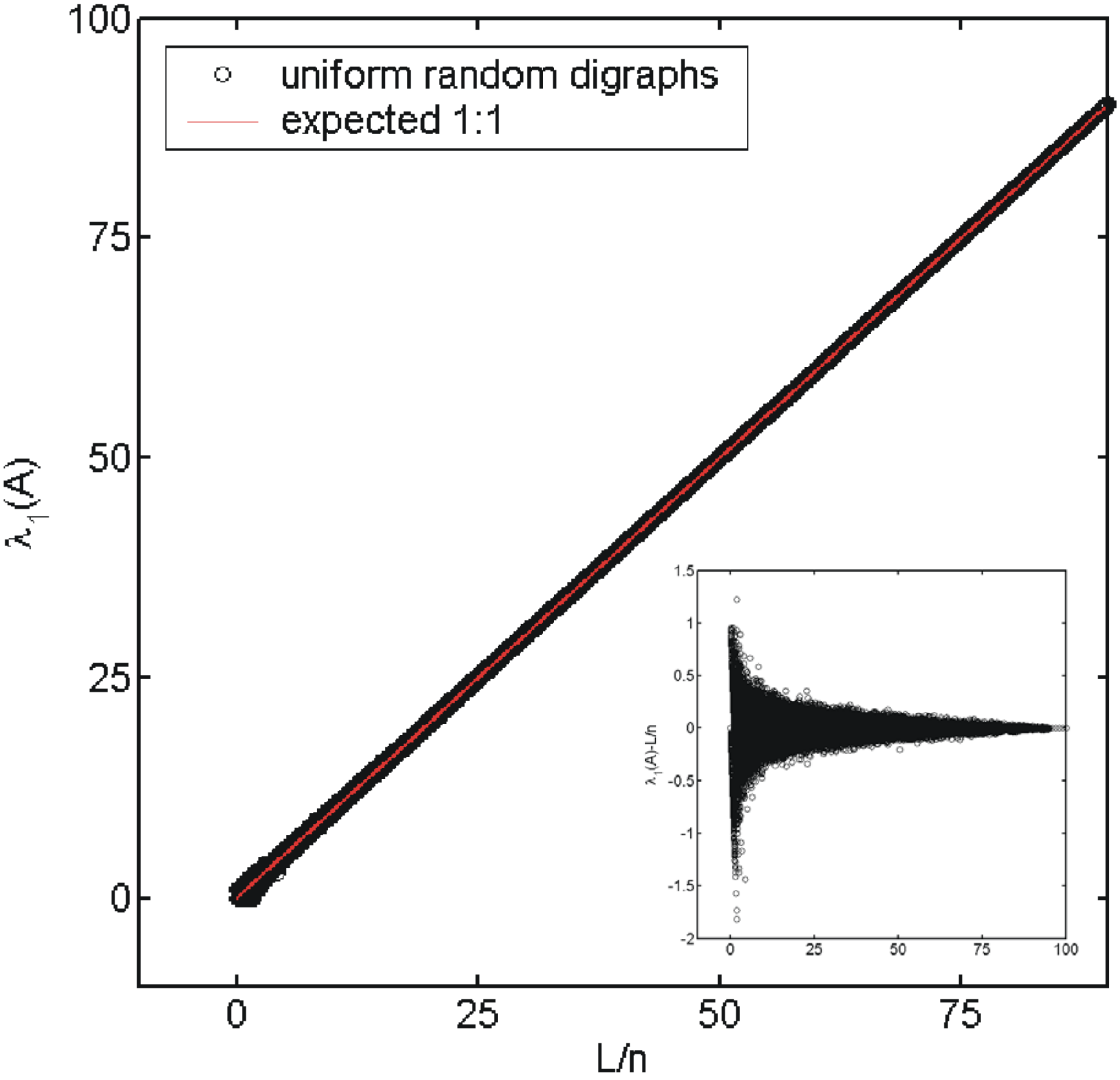


Figure 3.2: Dominant Eigenvalue in Random Digraphs



CHAPTER 4

TEMPORAL VARIATION OF INDIRECT EFFECTS IN A SEVEN-COMPARTMENT MODEL OF NITROGEN FLOW IN THE NEUSE RIVER ESTUARY, USA: TIME SERIES ANALYSIS[1]

---

[1] Borrett, S.R., S.J. Whipple, B.C. Patten, R.R. Christian, Accepted by *Ecological Modelling*. Reprinted here with permission of publisher.

ABSTRACT

Indirect effects, the ability of one element in a system to impact another without direct interaction, are critical components of complex adaptive systems. Network Environ Analysis (NEA) is an environmental extension of economic Input-Output Analysis that characterizes and quantifies indirect effects transmitted by a specific currency (e.g., carbon, nitrogen) within a steady state system. Previous NEA results suggest that indirect effects tend to dominate direct effects in ecosystems. Although steady state models may be useful, ecologists are often interested in system dynamics. In this article, we use NEA to investigate the temporal dynamics of indirect effects in sixteen steady state models of nitrogen cycling in the Neuse River Estuary, USA. These models were originally constructed by Christian and Thomas (2003. Network analysis of nitrogen inputs and cycling in the Neuse River Estuary, North Carolina, USA. Estuaries **26**:815-828.). Our work had two primary objectives: 1) describe and quantify the temporal dynamics of indirect flows and 2) identify the importance of model factors in determining the magnitude of indirect flows. Our results indicate that indirect flows transmitted by the network of nitrogen transactions, especially microbial transformations, dominated direct flows in the Neuse River Estuary between spring 1985 and winter 1989. When we examined the indirect flow index (ratio of indirect to total system throughflow), we were surprised by the relatively small amount of temporal variation within the Neuse River Estuary. In all seasons, indirect flow was more than 80% of TST, and there was no significant interannual variation. Our analysis also revealed 1) no clear relationship between boundary or direct flow and indirect flow, but 2) a strong association between cycled and indirect flow. We conclude that while an increase in boundary, direct, and cycled flows can influence indirect flows, in the Neuse River Estuary models indirect flows are predominantly due to cycling. If nitrogen loading had been reduced

during our study period, this system likely would have remained eutrophic for an extended period of time. If we assume the system remained similar in 1997, we would expect the legislated reduction of nitrogen loading to have little immediate effect, although this action will be important for long–term change. Dominance of indirect flow due to cycling makes this system state relatively difficult to change rapidly by external forcing. In this sense the indirect effects stabilize the system, making it more resistant to external perturbations, though it may be more sensitive in the long–term.

“We are far less powerful than we think, but have much more impact than we perceive.”

Mitchel Thomashow (2002)

4.1 INTRODUCTION

Indirect effects, the ability of one species or system element to influence another without directly interacting with it, appear to have a significant role in the organization and transformation of biological systems. They have long been recognized by biologists (Darwin, 1959), and the challenge they pose for environmental management helped spark the environmental movement (Carson, 1962). Though challenging to investigate, ecologists are increasingly recognizing the significance of indirect effects as components of ecological interactions (Patten, 1983; Andrewartha and Birch, 1984; Patten, 1984; Miller and Kerfoot, 1987; Strauss, 1991; Wootton, 1994, 2002) and evolution (Miller and Travis, 1996; Laland et al., 1999; Odling-Smee et al., 2003). Wootton (2002) recently argued that indirect effects are “…a fundamental cause of ecosystem complexity.” In this article, we characterize the temporal variation of indirect effects in sixteen seasonal models of nitrogen cycling for the Neuse River Estuary, USA.

Part of the challenge of analyzing indirect effects is that there appear to be multiple types that ecologists have struggled to classify and empirically demonstrate (Miller and Kerfoot, 1987; Strauss, 1991; Wootton, 1994; Abrams, 1995). Here, we adopt the categorization scheme of Wootton (1993; 1994; 2002) that begins with two categories: interaction chains and interaction modifications. Interaction chains occur when one species impacts another by affecting a third. This occurs by linking two or more direct interactions together. For example, in the chain A → B → C, A indirectly influences C by directly influencing B. As ecologists are often interested in the abundance or density of species, these have also been labeled ‘density mediated interactions’

(Abrams et al., 1996). Several commonly studied ecological interactions arise from this type of indirect interaction including keystone predation, trophic cascades, apparent competition, indirect mutualisms, and exploitative competition (Wootton, 2002). Interaction modifications, on the other hand, occur when a species modifies the interaction between two species. At least two types of interaction modifications have been identified. 'Trait–mediated indirect effects' (Abrams, 1995; Abrams et al., 1996; Peacor and Werner, 1997; Bolker et al., 2003; Luttbeg et al., 2003; Werner and Peacor, 2003) occur when one species changes the traits or behavior of a second that alters how it interacts with a third. For example, Pacific killifish (*Fundulus parvipinnis*) parasitized by the trematode *Euhaporchis californiensis* tend to exhibit conspicuous behavior making them more susceptible to predation by avian predators than unparasitized killifish (Lafferty and Morris, 1996). 'Environment–mediated indirection modifications' occur when one species changes the environmental context in which two species interact. Ecosystem engineers, such as beavers that construct ponds providing habitat for aquatic organisms, are one example (Jones et al., 1997).

In ecosystems, organisms and elements of their abiotic environments are coupled together through an intricate network of energy–matter exchanges (Patten et al., 1976; Ulanowicz, 1986; Higashi and Burns, 1991). While observation and analysis of these transaction networks cannot identify specific mechanisms or depict all types of indirect interactions (Loehle, 1990; Wootton, 1994), they do capture many types of indirect interactions that are reflected in trophic dynamics and biogeochemistry (Patten, 1990; Higashi and Burns, 1991). These include some indirect interactions from each of the three broad categories identified earlier. Therefore, analysis of these flow–storage networks may illuminate the consequences of particular system organizations, especially the indirect effects mediated by this transaction network.

One purpose for developing input–output analysis (IOA) in economics was to trace indirect effects through the network of economic transactions (Leontief, 1966). Network environ analysis (NEA) is an environmental application and extension of economic IOA. Through NEA, Patten and colleagues (Patten, 1984, 1985; Higashi and Patten, 1986, 1989; Patten, 1991) characterized indirect flows transmitted through the network of energy–matter exchanges in an ecosystem. At the whole-system level, the relative significance of indirect flows is indicated by the ratio of indirect to direct flow (Indirect/Direct). This ratio indicates the significance of indirect flows within a system's *environs*, which are within system, compartment-specific, input and output oriented environments (Patten, 1978, 1981, 1982, 1992). Previous results suggest that the Indirect/Direct flow ratio tends to be greater than unity in model ecosystems, implying that indirect effects are dominant (Patten, 1984; Higashi and Patten, 1986, 1989; Patten, 1991; Fath, 2004). This has led Patten (in prep.) to hypothesize that the natural world is unified through indirect effects.

Applications of NEA have been limited to steady state models (e.g., Dame and Patten, 1981; Patten and Matis, 1982; Flebbe, 1983; Gattie et al., 2005). Though these models can be useful, ecologists are often interested in system dynamics. In this paper we use NEA to investigate the temporal dynamics of indirect flows in the Neuse River Estuary. We take a small step toward a true dynamic NEA by investigating both the seasonal and interannual variation in indirect effects in a discrete–time sequence of steady state models. Further, while past work indicates that aspects of both system structure and function influence the degree of indirect effects (Patten et al., 1990; Patten, 1991; Fath, 2004), the relative importance of various factors is unknown. This knowledge is crucial to our understanding of how indirect effects integrate ecological systems and alter their responses to environmental impacts. The Neuse River Estuary data set provides an opportunity to begin to address these questions. Understanding the causes and consequences of indirect effects is necessary for sustainable management of ecosystems and the services they provide, a critical mission for ecological sciences (Palmer et al., 2004; Palmer et al., 2005).

The Neuse is a large river in North Carolina draining a 16,000 km^2 watershed. Water leaving the river enters the Neuse River Estuary and then flows into Pamlico Sound. This estuary has received a great deal of political and scientific attention in the last two decades because it has become highly eutrophic and at times hypoxic (Christian et al., 1986; Baird et al., 2004). In 1997, the State of North Carolina legislated a thirty percent reduction in nitrogen loading to the estuary. In addition, the US EPA required the state to develop a total maximum daily load (TMDL) of nitrogen entering the estuary by the summer of 1997. The sixteen seasonal nitrogen models were originally constructed by Christian and Thomas (2000; 2003) as part of the larger Neuse River Modeling and Monitoring (ModMon) program to study the estuary's response to new environmental management (Reckhow and Gray, 2000). Their analysis using a complementary type of ecological network analysis indicated that nitrogen dynamics in the estuary are dominated by internal recycling (Christian and Thomas, 2003).

Our first objective was to characterize the temporal dynamics of indirect flows transmitted through the nitrogen transaction networks. We used NEA to describe and quantify the indirect flows because it decomposes observed throughflows into their boundary, direct, and indirect portions. Although ecosystems are open thermodynamic systems that eventually dissipate imported energy–matter, indirect flows tend to dominant direct because of the large number of indirect pathways available in cyclic systems (Patten et al., 1982; Patten, 1985; Borrett and Patten, 2003, see Chapters 2 and 3). As the Neuse River Estuary nitrogen models are cyclic and well connected, we expected indirect flows to be dominant. We also anticipated the proportion of indirect flow to vary seasonally, contributing more in the spring and summer when biological activity increases, and less in the winter when biological activity slows and physical processes are more pronounced. River discharge and nutrient loading tend to be high and flushing times short during colder months (Christian et al., 1991). Alternatively, more biological activity may decrease the proportion of indirect flow.

Our second objective was to determine the relative importance of factors determining indirect flow in the Neuse models. Higashi and Patten (Higashi and Patten, 1986; Patten et al., 1990; Patten, 1991) showed algebraically that the ratio of indirect–to–direct effects should increase as model size (number of nodes, n), connectivity (proportion of possible links connected, $C = L/n^2$), strength of direct flows (Direct), and magnitude of cycling (Cycled) increase. This algebra, however, does not determine the relative importance of these factors or indicate whether this is a complete set of factors. Other attributes such as network topology (Borrett and Patten, 2003, see Chapter 2 and 3) or boundary flow also might be influential. In the model ecosystems we analyzed, network structure (n, C, and topology) was constant for the 16 seasons, but functional attributes (Direct, Cycled, and Boundary flow) varied. Given Higashi's algebra, we expected Direct and Cycled flow to significantly influence the magnitude of indirect effects. Further, we expected cycling to be the most significant factor as it establishes the feedbacks, amplifying direct flows into indirect components.

4.2 MATERIALS AND METHODS

4.2.1 NEUSE RIVER ESTUARY NITROGEN CYCLING MODELS

For our investigation, we used sixteen seasonal network models of nitrogen cycling in the Neuse River estuary from Spring 1985 to Winter 1989 originally constructed by Christian and

Thomas (2000; 2003). Each model has the same basic structure (Figure 4.1), with seven compartments (n = 7), 22 within–system observed flows, and boundary loading to and losses from each compartment. Compartments, or network nodes, represent nitrogen storage (mmol N m^{-2}) in phytoplankton (PN-Phyto), heterotrophs (PN-Hetero), detritus (PN-Aerobic), sediments, dissolved organic nitrogen (DON), nitrate and nitrite (NO_x), and ammonium (NH_4). Nitrogen flows (mmol N m^{-2} $season^{-1}$) are represented by directed arcs in the network. Each compartment has nitrogen imports from upstream and the watershed as well as boundary losses due to denitrification and exports downstream into Pamlico Sound. The sediment compartment has an additional boundary loss due to sediment burial. Baird et al. (1991) suggested this structural consistency is critical for the type of model comparisons we make in this paper.

As reported in Christian and Thomas (2003), nitrogen flow and storage data for the models were drawn largely from a four year study of the Neuse River Estuary (Christian et al., 1991; Christian et al., 1992; Rizzo et al., 1992; Boyer et al., 1993; Boyer et al., 1994; Rizzo and Christian, 1996). The least well known data, including sediment–water column interactions, burial, and denitrification, were estimated based on related studies; sediment burial rates were adjusted to balance the model to a steady state[2] (i.e., inputs equal outputs for each node). See Christian and Thomas (2000; 2003) and Christian et al. (1992) for additional model details.

---

[2] The model for Summer 1987 was not exactly at steady state. However, the difference between input and output throughflow in NO_x, the only compartment not at steady state, was only 0.9 mmol N m^{-2} $season^{-1}$. Analysis of an adjusted model (data not shown) showed this had no significant impact on our results.

4.2.2 NETWORK ENVIRON ANALYSIS AND INDIRECT EFFECTS

4.2.2.1 ENVIRON INDIRECT EFFECTS

Empirically measured or observed energy–matter flows from compartment j to i ($F_{n\times n} = (f_{ij})$, i, j = 1,…, n) in natural systems are comprised of both direct and indirect flows (Patten et al., 1976; Whipple and Patten, 1993; Whipple, 1999; Gattie et al., 2005). NEA (Patten et al., 1976; Patten, in prep.) is a family of input–output methods descended from economics (Leontief, 1965, 1966) that analytically decomposes observed flows to identify their origins or fates within the system of interest (see Fath and Patten, 1999 for review). NEA includes structural analysis for path enumeration as well as functional analyses to investigate flow, storage, utility, and control. In this work, we focused on the output oriented throughflow-specific flow analysis. NEA methods are extensively described in the literature (Patten et al., 1976; Matis and Patten, 1981; Fath and Patten, 1999; Patten, in prep.) and were recently collected into a single MATLAB® function to facilitate its application (Fath and Borrett, 2005, see Appendix A). We will briefly review the NEA elements we used, and define the response variables for this study.

There are several basic steps to output oriented throughflow-specific flow analysis in NEA. Given the observed intercompartmental flows ($F_{n\times n}$) and boundary inputs ($z_{n\times 1}$) and outputs ($y_{1\times n}$) of an n node system, we first define the total amount of energy–matter flowing into and out of each node as $T_k^{(in)} \equiv \sum_{j(\neq k)=1}^{n} f_{kj} + z_k$ and $T_k^{(out)} \equiv \sum_{i(\neq k)=1}^{n} f_{ik} + y_k$ respectively, where $T_k^{(in)} = T_k^{(out)} = T_k$ $(k = 1,\ldots,n)$ at steady state. Total system throughflow ($TST = \sum T_k$) is a system-level measure of activity and is often used to characterize ecosystem models (Finn, 1976; Ulanowicz, 1986).

In the second step, observed flows (F) are normalized by the donor compartment throughflow. The resultant matrix, $G_{n\times n}$, represents the donor-specific direct flow intensities from j to i ($G = (g_{ij}) = \left(\frac{f_{ij}}{T_j}\right)$). Elements of G are interpreted as the probability that boundary material entering j will flow to i over a direct path (i.e., path length, m = 1). Next, flow from j to i over indirect pathways (m > 1) is determined by raising G to the m^{th} power, G^m. The elements of G^m are interpreted as the fractional transfer coefficient for boundary material entering j flowing over all pathways of length m between any j and i. Finally, total node throughflow is recovered by post multiplying the sum of the infinite power series of flow intensities by the boundary inflows:

$$T = \left(\underbrace{I}_{\text{Boundary}} + \underbrace{G^1}_{\text{Direct}} + \underbrace{G^2 + \ldots G^m + \ldots}_{\text{Indirect}} \right) * z, \tag{4.1}$$

where $I = G^0$ is the matrix multiplicative identity. In open, dissipative, thermodynamic systems like ecosystems, $0 \leq g_{ij} < 1$ $(\forall i, j)$ and at least one column (or row) sum is less than unity ensuring that the dominant eigenvalue of G is less than one (Berman and Plemmons, 1979). This criterion ensures that the power series converges to the transitive closure matrix $N = (n_{ij}) = (I - G)^{-1}$, where (n_{ij}) represents the integral (boundary + direct + indirect) throughflow from j to i generated by an input to j. Thus, $T = Nz$. The analysis used here only incorporates flow over pathways that Higashi and Patten (1989) termed procedurally indirect. Pathways with adjacent repeated nodes (e.g., i → i → i → i → j) that create only temporal indirectness are included as direct flows.

We can rewrite the throughflow decomposition in (4.1) by distributing z across the flow intensity partition, and summing the elements of each n×1 resultant vector to derive the following restatement in terms of TST:

$$\mathrm{TST} = \underbrace{\sum \mathrm{Iz}}_{\text{Boundary}} + \underbrace{\sum \mathrm{Gz}}_{\text{Direct}} + \underbrace{\sum (\mathrm{N}-\mathrm{I}-\mathrm{G})\mathrm{z}}_{\text{Indirect}} . \tag{4.2}$$

This equation identifies the dimensionalized flow variables for the whole system scaled by the model inputs that we used as response variables: Boundary ($\sum \mathrm{Iz}$), Direct ($\sum \mathrm{Gz}$), and Indirect ($\sum (\mathrm{N}-\mathrm{I}-\mathrm{G})\mathrm{z}$). Dividing both sides of (4.2) by TST generates:

$$1 = \frac{\sum \mathrm{z}}{\mathrm{TST}} + \frac{\sum \mathrm{Gz}}{\mathrm{TST}} + \frac{\sum (\mathrm{N}-\mathrm{I}-\mathrm{G})\mathrm{z}}{\mathrm{TST}} . \tag{4.3}$$

Elements of the right-hand side of (4.3) represent the portion of TST derived from each of the three categories of flow. From (4.2) the ratio of indirect–to–direct flows (Indirect/Direct) can be constructed as

$$\text{Indirect}\big/\text{Direct} = \frac{\sum (\mathrm{N}-\mathrm{I}-\mathrm{G})\mathrm{z}}{\sum \mathrm{Gz}} . \tag{4.4}$$

Indirect/Direct is a system–level indicator of the relative significance of indirect flow in TST. To address our first objective, we calculated this ratio and the TST partition in (4.3) for the Neuse River Estuary models using a modification of NEA.m (Fath and Borrett, 2005, see Appendix A). Further, we compared the mean indirect flow index ($\mathrm{IFI} \equiv \frac{\text{Indirect}}{\mathrm{TST}} = \frac{\sum (\mathrm{N}-\mathrm{I}-\mathrm{G})\mathrm{z}}{\mathrm{TST}}$, $0 \le \mathrm{IFI} < 1$), averaging first across seasons, and second across four years. Year classes were constructed by grouping spring, summer, and fall of a year with the following winter (e.g., year class 1985, yc1985 = {Spring 1985, Summer 1985, Fall 1985, and Winter 1986}). Following a Kolmogorov-Smirnov test to insure IFI was normally distributed, we analyzed statistical significance of season and year class using a two-way ANOVA (Dalgaard, 2002). When necessary, multiple pairwise comparison t-tests were evaluated with Bonferroni adjusted p-values. We calculated all statistics with R (version 1.7.3, R Development Core Team, 2004).

Finn (1976) proposed an alternative partition of TST into portions from acyclic throughflow and cycled flow. Cycled flow, calculated as Cycled $\equiv \sum_{i=1}^{n}((n_{ii}-1)z_i)$, was an additional response variable in this study. The Finn cycling index (FCI), the ratio of cycled flow to TST, is another common metric used to characterize ecosystem organization. Allesina and Ulanowicz (2004) show that FCI does not account for all recycling, and introduce a new measure called the comprehensive cycling index (CCI). In their analysis of 23 ecosystem models, however, they found CCI to be linearly related to FCI by a factor of 1.14. As CCI is computationally intensive to determine and appears to be a linear multiple of FCI, we focus on FCI.

4.2.2.2 DETERMINANTS OF INDIRECT EFFECTS

To evaluate the relative influence of system attributes in determining the portion of indirect flow, we constructed a series of ordinary least squares regression models. While we could not evaluate the role of factors such as n and C because structure was constant in these networks, flow variability allowed us to consider the role of direct and cycled flow identified by Higashi and Patten's algebra as important determinants of indirect flows. We also considered boundary flow as a potential factor as it is a component of TST (2). We first examined the independent relationship between Indirect and Boundary, Direct and Cycled. We further considered additional pairwise combinations and multiple regression where appropriate.

4.3 RESULTS

The sixteen seasonal models of nitrogen cycling in the Neuse River Estuary demonstrate an expected seasonality in TST (Figure 4.2). TST varies from a minimum of 5,732 mmol N m^{-2} $season^{-1}$ in fall 1988 to a maximum of 20,182 mmol N m^{-2} $season^{-1}$ in summer 1985. Mean TST was 10,373 mmol N m^{-2} $season^{-1}$ (± 4,207 SD). There appears to be a seasonal trend to the TST data with highs in summers and lows in winters. These results match our expectation that biogeochemical activity increases in spring as temperature and light levels increase, reaching a maximum in summer, and then declining through the fall to a minimum in winter. One exception to this general pattern is the summer of 1986. It is lower than the previous spring and is surprisingly lower than that found in the other four years. The cause of this deviation is unknown. The Finn Cycling Index (FCI) was high in these models with a mean of 88% (± 9% SD). Though we used a different set of algorithms, our TST and FCI values appear similar to those reported by Christian and Thomas (2003), demonstrating analytic consistency. Further, they showed that the values of nitrogen cycling found in the Neuse River Estuary are some of the highest reported in the literature. While there appears to be some seasonality to FCI, its regression with TST does not explain much of the variation ($R^2 = 0.31$).

In summary, nitrogen fluxes in the sixteen seasons of the Neuse River Estuary analyzed here were large and varied in a seasonal manner consistent with ecological expectations. More surprising was the magnitude of nitrogen recycling in the estuary. TST and the cycling index provide the starting point for our analysis of indirect flows in this system.

4.3.1 INDIRECT EFFECTS IN THE NEUSE RIVER ESTUARY MODELS

Nitrogen flux in the Neuse River Estuary was dominated by indirect flows in all sixteen seasons (Figure 4.3). Indirect/Direct consistently exceeded unity, ranging from 9.9 in winter

1987 to 174.4 in summer 1987 (Figure 4.3A). Indirect/Direct is linearly associated with TST ($R^2 = 0.72$), indicating that it is influenced by variation in TST. The fractional decomposition of TST using equation (4.3) again shows the dominance of indirect flows (Figure 4.3B). In all sixteen seasons indirect flow was more than 80% of TST. This partition also hints at a strong similarity in the amount of nitrogen loading or boundary inputs and the direct flows in the system.

There were small but not significant seasonal and interannual variations in the indirect flow index (IFI), where indirect flows were normalized by TST (Figure 4.4). Seasonal means ranged from 0.98 (± 0.01 SD) in summer to 0.91 (± 0.06 SD) in winter (Figure 4.4A). Year class means ranged from 0.97 (± 0.02 SD) in yc1985 and yc1987 to 0.94 (± 0.07 SD) in yc1986 (Figure 4.4B). Winter and yc1986 were the most variable groups, as they contained winter 1987 which had the lowest IFI. This may be because winter 1987 had the highest loading of nitrogen, resulting from extensive rainfall that season, and shortest flushing time within the estuary (Christian et al., 1991, Christian and Thomas, 2000). A two-way ANOVA suggests that seasonal variation is significant at $\alpha = 0.1$ but not $\alpha = 0.05$, and there is no significant difference between year classes (Table 4.1). We used a multiple pairwise t-test with Bonferroni corrected p-values to examine the potential seasonal differences suggested by the ANOVA. As the seasonal variances were not equal, violating the assumptions of the traditional t-test, we used a modified test in which standard deviations are not pooled (Dalgaard, 2002). These results suggest that there are no significant differences between the paired seasons (Table 4.2). Thus, the marginal statistical significant difference among seasons that appears in the two-way ANOVA probably results from unequal variances.

We were surprised at the relatively small amount of temporal variation in indirect flows when normalized by TST within the Neuse River Estuary. We expected more pronounced seasonal variation, increasing with biological activity in the spring, climaxing in summer and then declining to a minimum in the winter. This pattern occurs but the difference between seasons was small; all were over 90%.

4.3.2 DETERMINANTS OF INDIRECT FLOW

The first step of this analysis was to investigate the pairwise relationship between Indirect and 1) Boundary, 2) Direct, and 3) Cycled. An initial scatter plot of Boundary and Direct versus Indirect revealed no clear relationship with Indirect (Figure 4.5A). However, it did reveal a strong linear relationship between Boundary and Direct. When the regression was forced through the origin (by definition direct is zero if there is no boundary flow), it was Direct = 0.96*Boundary ($R^2 = 0.99$, $p<0.001$; Figure 4.5B). Thus, boundary and direct flow were nearly identical in this model. This occurs because each model compartment has a high transfer efficiency. Little nitrogen is lost from the system at each node, allowing most of the boundary inputs to pass into the model over paths of length one, which is how Direct was defined. There was also a strong linear relationship between Indirect and Cycled (Figure 4.6). The regression equation was Indirect = 906 + 0.98*Cycled ($R^2 = 0.99$, $p<0.001$), where both coefficients were statistically significant. The large intercept suggests that a portion of indirect flow occurs without cycling. This is understandable as interaction chains can create indirect flows without cycling. In these models, cycled flow by itself explains nearly all of the variation in indirect flow; a multiple factor investigation was unnecessary.

4.4 DISCUSSION

We make two primary contributions with this work. First, we used a sequence of seasonal network models to investigate the discrete-time dynamics of indirect flows. In its present form, the mathematics of NEA requires input data from a steady-state system. Due to this analytical constraint, NEA characterizes the organizational complexity of ecosystems, but not their dynamic complexity that is frequently of ecological interest. However, the seasonal networks we analyzed provide a temporal sequence of snapshots that begins to reveal system dynamics, much like constructing a motion picture film from a series of still frames. Contrary to our expectations, our results revealed no significant seasonal or interannual variation in indirect flow as a fraction of TST; it was constantly larger than 80%. While there are temporal changes in nitrogen loading and total nitrogen flux, the macro organization of the estuary ecosystem with respect to nitrogen flux was remarkably constant.

The consistency of IFI results from the overwhelming dominance of indirect and recycling flows in determining TST. There are several features of the Neuse River Estuary that contribute to this. First, boundary inputs are minor components of TST. Christian and Thomas (2003) calculated that TN loading averaged 2% of TST, ranging from <1% to the unusually high value of 8% during the winter of 1987. Second, the importance of indirect flows and cycling of nitrogen is directly related to differences in the time scales of physical and biological processes. The flushing time of freshwater is of the order of weeks to months with an average of about 7 weeks (Christian et al., 1991). The turnover of dissolved nitrogen due to biological uptake and release is of the order of hours to days (Boyer et al., 1994). Phytoplankton turnover occurs in days (Boyer et al., 1993). Thus, there is considerable opportunity for imported nitrogen to be cycled through biological components numerous times from entry to exit from the system

(Christian and Thomas, 2003). Third, the biological activity is largely microbial and potentially rapid, resulting from planktonic uptake and release of nitrogen and benthic biogeochemical processes. Cycles with quantitatively important, and biologically mediated, indirect flows thus control the overall dynamics of nitrogen. Biological processing is highly variable among seasons and is such a dominant part of TST that the resultant indices of FCI and IFI show little variation.

Second, we evaluated the relative contribution of several system attributes to the development of indirect flow. Higashi and Patten (Higashi and Patten, 1986; Patten et al., 1990; Patten, 1991) demonstrated algebraically that indirect flows could be influenced by n, C, cycling, and direct flows, but this work did not evaluate the relative contribution of these variables or determine if it was a complete set. We used linear regression models to evaluate the factors most significant in determining indirect flows. This statistical analysis supported and extended our findings for an explanation of consistency in IFI and FCI. We discovered that in the sixteen Neuse River Estuary models cycled flow was indeed the most important factor in explaining variation in indirect flow. Boundary and direct flows were only important in that they are necessary for cycling to occur. Though indirect flows are composed of both cyclic and acyclic flows such that IFI > CCI > FCI, cyclic flows appear to be dominant in these models. This occurs because the model compartments have a high transfer efficiency with respect to nitrogen use; only a small fraction of nitrogen imports are immediately lost to the system boundary. In this study, we were unable to appraise the relative significance of the model structure, including n, C, and network topology, as structure was constant. Unfortunately we cannot generalize from these observations – they are particular to our study system. However, they are the first results to begin to address this question.

Ultimately, the Neuse River Estuary was very retentive of nitrogen between spring 1985 and winter 1989. The system was dominated by internal, biological processes. If nitrogen loading had been significantly reduced during this period, this ecosystem likely would have remained eutrophic for an extended period of time. If we assume that the system organization remained similar in 1997, we would expect the legislated reduction of nitrogen loading to have little immediate effect, though this action may be important for long-term change. This conclusion is consistent with that of Christian and Thomas (2003).

It is difficult to compare our results with previous reports of indirect flows determined by NEA for two reasons. First, previous publications usually do not report the same statistics we use in this paper. Higashi and Patten (1989) reported a measure of total indirect-to-direct flows for several ecosystem models taken from the literature. A five node model of energy flux (kcal m^{-2} y^{-1}) in a cold spring and a five node network of nitrogen flux (g N m^{-2} y^{-1}) in a Puerto Rico mountain rainforest had total indirect flow factors of 2.2 and 12.8 respectively (models reported in (Patten et al., 1976)). They also reported a total indirect flow measure for a six compartment oyster reef model of energy flux (kcal m^{-2} day^{-1}) as 3.3 (model from Dame and Patten (1981)). Unfortunately, they do not provide a clear description for how these total indirect flows were calculated. However, we suspect the calculation 1) is based on an alternate formulation for the indirect-to-direct flow ratio presented in Fath and Patten (1999) – $\text{Indirect}/\text{Direct} = \frac{\sum\sum(n_{ij} - i_{ij} - g_{ij})}{\sum\sum(g_{ij})}$, where i_{ij} are the elements of the identity matrix I, and 2) includes indirect flows from both the input and output orientations. Using Fath and Patten's alternative formulation the indirect-to-direct ratios in the three models are 0.91, 6.18 and 1.53 respectively; when calculated as in equation (4.4) we found them to be 1.02, 6.14, and 1.58 respectively.

Regardless of how calculated, the Indirect/Direct ratios for the Neuse River Estuary are substantially larger. Fath (2004) reported an indirect-to-direct ratio for 250 large–scale cyber–ecosystem models of trophic dynamics with 30, 60, 120, 300, or 600 nodes, with varying connectivity, cycling, and TST. His results, again using the alternative formulation, show that Indirect/Direct increases in these models with the number of nodes from approximately 8-10 for models with 30 nodes to 80-95 in models with 600 nodes. It also tends to increase with the Finn cycling index (FCI), though this had a restricted range (0.14 – 0.26) in these models. Given these results, the magnitudes of the Indirect/Direct ratio in the Neuse River Estuary models are quite surprising. Some of these very small models display much larger indirect flows than even the largest cyber–ecosystem models. We suspect that this difference is a reflection of the differences in compartment efficiencies and cycling, which result from distinctions between models of trophic processes and biogeochemistry. Christian et al. (1996) discussed the ramifications of the differences between trophic and biogeochemical networks. Biogeochemical networks, such as analyzed here, may be minimally dissipative, highly aggregated in biological nodes and disaggregated in chemical nodes, focused on microbial processes, and inclusive of inorganic components instead of treating them as external to the system (e.g., CO_2 in foodweb networks). These attributes generally foster cycling and resultant indirect flow.

The second difficulty however is perhaps more important. Our results show that the Indirect/Direct ratio increases linearly with TST. Thus, when we compare Indirect/Direct ratios between systems, we cannot be sure if we are identifying differences in indirect flows or variation in TST. However, if we normalize the magnitude of indirect flow by TST as in the indirect flow index (IFI) we should remove the effects of TST variation. This creates an indicator of indirect flows that is more comparable across systems, similar in form to FCI or

CCI. This should facilitate comparative ecosystem analysis of indirect flows in the spirit of Baird et al. (1991) and Krivtsov (2004).

Strong indirect influences in the Neuse River Estuary with respect to nitrogen, made possible by strong cycling, appear to make this system initially more resistant to external manipulations to manage the desired change of state (less eutrophic). In this sense, indirect flows seem to be a stabilizing force, making the system more resistant to external perturbations, an idea consistent with previous theoretical and empirical ecological investigations of nutrient cycling (DeAngelis et al., 1989 and references therein). However, this may be a short-term or transient dynamic. Systems tightly integrated or strongly connected by network indirect flows may ultimately be more vulnerable due to their interconnectedness (Barabási, 2002; Gunderson and Holling, 2002). The system transformation may take a long time to be fully expressed, but we suspect that systems like the Neuse River Estuary that are highly integrated will be more sensitive to both external and internal perturbations including internal changes such as alterations in a lower hierarchical level like an individual node becoming more or less dissipative (efficient) or disappearing completely. This is because system integration through indirect flows creates what is known as the multiplier effect in economics (Samuelson, 1948). The multiplier effect is indicated by the number of times an input is expected to be used or travel through the system before it exits. Ecologists have termed this average path length (Finn, 1976) or more recently network aggradation (Jørgensen et al., 2000). We expect perturbations to be multiplied more in more integrated systems, making them more sensitive or less stable. This increasing sensitivity with increasing connectedness has been described as a loss of ecological resilience in the general model of system transformations or adaptive cycle at the heart of Panarchy Theory (Holling, 1973, 2001; Gunderson and Holling, 2002). This potential loss of stability is one argument for a modular or hierarchical arrangement of system components, allowing the potential benefits of increasing system integration while localizing instabilities (Simon, 1962; Krause et al., 2003). This potential increased sensitivity to perturbations suggests that targeted attacks on the most well connected nodes (as in Albert et al., 2000; Dunne et al., 2002) or nodes with the largest throughflow might rapidly generate system wide changes. If true this suggests that management actions to increase the nitrogen dissipation from Sediment (most well connected) or PN-Hetero (most throughflow) might initiate a more rapid system transformation. Control analysis (Schramski et al., 2005) of this system, however, suggests that NOx may be a better management target. Further investigation is required to understand these conflicting possibilities.

In conclusion, nitrogen flux in the Neuse River Estuary between spring 1985 and winter 1989 was dominated by indirect flows. The indirect flow index – total indirect nitrogen flow normalized by total system throughflow – showed little seasonal and no significant interannual variation. Nitrogen cycling was the overriding factor determining the magnitude of indirect flows. The Neuse River Estuary was highly nitrogen retentive, ensuring that the same atom of nitrogen revisited individual compartments multiple times before exiting the system. Our analysis suggests that because this system is dominated by internal processes management efforts focused on decreasing nitrogen loading would not have rapidly impacted the eutrophic state of the estuary. In this sense indirect flows seem to be a stabilizing force, making it more resistant to external perturbations. This may be a transient effect, however, as indirect flows may make the system more sensitive in the long-term, as perturbations are propagated and potentially magnified. This prediction may currently be in operation as the ecosystem commonly experiences periods of anoxia and hypoxia with subsequent fish kills linked to the substantial microbial processing of energy (Baird et al., 2004).

4.5 ACKNOWLEDGEMENTS

This work benefited from conversations with D.K. Gattie, J.R Schramski, B.D. Fath, and S. Bata. We would also like to thank M.B. Beck, D.K. Gattie, H.R. Pulliam, P.G. Verity, and two anonymous scientists for manuscript reviews. SRB and SJW were supported in part by the National Science Foundation biocomplexity grant (OPP-00-83381) to the Skidaway Institute of Oceanography.

4.6 REFERENCES

Abrams, P.A., 1995. Implications of dynamically variable traits for identifying, classifying, and measuring direct and indirect effects in ecological communities. Am. Nat., **146**:112-134.

Abrams, P.A., Menge, B.A., Mittlebach, G.G., Spiller, D., and Yodzis, P., 1996. The role of indirect effects in food webs. In: G. Polis and K.O. Winemiller (Editors), Food webs: dynamics and structure, Chapman and Hall, New York, pp 371-395.

Albert, R., Jeong, H., and Barabasi, A.L., 2000. Error and attack tolerance of complex networks. Nature, **406**:378-382.

Allesina, S., and Ulanowicz, R.E., 2004. Cycling in ecological networks: Finn's index revisited. Computational Biology and Chemistry, **28**:227-233.

Andrewartha, H.G., and Birch, L.C., 1984. The ecological web: more on the distribution and abundance of animals. University of Chicago Press, Chicago.

Baird, D., Christian, R.R., Peterson, C.H., and Johnson, G.A., 2004. Consequences of hypoxia on estuarine ecosystem function: energy diversion from consumers to microbes. Ecological Applications, **14**:805-822.

Baird, D., McGlade, J.M., and Ulanowicz, R.E., 1991. The comparative ecology of six marine ecosystems. Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci., **333**:15-29.

Barabási, A.-L., 2002. Linked: the new science of networks. Perseus, Cambridge, Mass.

Berman, A., and Plemmons, R.J., 1979. Nonnegative matrices in the mathematical sciences. Academic Press, New York.

Bolker, B., Holyoak, M., Krivan, V., Rowe, L., and Schmitz, O., 2003. Connecting theoretical and empirical studies of trait-mediated interactions. Ecology, **84**:1101-1114.

Borrett, S.R., and Patten, B.C., 2003. Structure of pathways in ecological networks: relationships between length and number. Ecol. Model., **170**:173-184.

Boyer, J.N., Christian, R.R., and Stanley, D.W., 1993. Patterns of phytoplankton primary productivity in the Neuse River Estuary, North Carolina, USA. Mar. Ecol. Progr. Ser., **97**:287-297.

Boyer, J.N., Stanley, D.W., and Christian, R.R., 1994. Dynamics of NH4+ and NO3- uptake in the water column of the Neuse River Estuary, North Carolina. Estuaries, **17**:361-371.

Carson, R., 1962. Silent spring. Houghton Mifflin, Boston.

Christian, R.R., Boyer, J.N., and Stanley, D.W., 1991. Multiyear distribution patterns of nutrients within the Neuse River Estuary, North Carolina. Mar. Ecol. Progr. Ser., **71**:259-274.

Christian, R.R., Boyer, J.N., Stanley, D.W., and Rizzo, W.M., 1992. Network analysis of nitrogen cycling in an estuary. In: C.J. Hurst (Editor), Modeling the Metabolic and Physiologic Activities of Microorganisms, John Wiley & Sons, New York, pp 217-247.

Christian, R.R., Bryant Jr., W.L., and Stanley, D.W., 1986. The relationship between river flow and *Microcystis aeruginosa* blooms in the Neuse River, North Carolina. Report 223, Water Resources Research Institute of the University of North Carolina, Raleigh, NC.

Christian, R.R., and Thomas, C.R., 2000. Neuse River Estuary modeling and monitoring project stage 1: network analysis for evaluating the consequences of nitrogen loading. Water Resources Research Institute 325-F, East Carolina University, Greenville, NC.

Christian, R.R., and Thomas, C.R., 2003. Network analysis of nitrogen inputs and cycling in the Neuse River Estuary, North Carolina, USA. Estuaries, **26**:815-828.

Dalgaard, P., 2002. Introductory statistics with R. Springer, New York.

Dame, R.F., and Patten, B.C., 1981. Analysis of energy flows in an intertidal oyster reef. Mar. Ecol. Progr. Ser., **5**:115-124.

Darwin, C., 1959. On the origin of species by means of natural selection. John Murray, London.

DeAngelis, D.L., Mulholland, P.J., Palumbo, A.V., Steinman, A.D., Huston, M.A., and Elwood, J.W., 1989. Nutrient dynamics and food-web stability. Ann. Rev. Ecol. Syst., **20**:71-95.

Dunne, J.A., Williams, R.J., and Martinez, N.D., 2002. Network topology and biodiversity loss in food webs: robustness increases with connectance. Ecol. Lett., **5**:558-567.

Fath, B.D., 2004. Network analysis applied to large-scale cyber-ecosystems. Ecol. Model., **171**:329-337.

Fath, B.D., and Borrett, S.R., 2005. A Matlab function for Network Environ Analysis. Environmental Modelling & Software, **XX**:XXX-XXX.

Fath, B.D., and Patten, B.C., 1999. Review of the foundations of network environ analysis. Ecosystems, **2**:167-179.

Finn, J.T., 1976. Measures of ecosystem structure and function derived from analysis of flows. J. Theor. Biol., **56**:363-380.

Flebbe, P.A., 1983. Aquatic carbon cycle in Okefenokee Swamp habitats: environ analysis. In: W.K. Lauenroth, G.V. Skogerboe, and M. Flug (Editors), Analysis of Ecological Systems: State-of-the-Art in Ecological Modelling, Elsevier, Amsterdam, pp 751-760.

Gattie, D.K., Schramski, J.R., Borrett, S.R., Patten, B.C., and Whipple, S.J., 2005. Indirect effects and distributed control in ecosystems 2. Network environ analysis of a seven-compartment model of nitrogen flow in the Neuse River Estuary, USA: steady-state analysis. Ecol. Model., **XXX**:XXX-XXX.

Gunderson, L.H., and Holling, C.S., 2002. Panarchy: understanding transformations in human and natural systems. Island Press, Washington, DC.

Higashi, M., and Burns, T.P., 1991. Theoretical studies of ecosystems: the network perspective. Cambridge University Press, Cambridge.

Higashi, M., and Patten, B.C., 1986. Further aspects of the analysis of indirect effects in ecosystems. Ecol. Model., **31**:69-77.

Higashi, M., and Patten, B.C., 1989. Dominance of indirect causality in ecosystems. Am. Nat., **133**:288-302.

Holling, C.S., 1973. Resilience and stability of ecological systems. Ann. Rev. Ecol. Syst., **4**:1-23.

Holling, C.S., 2001. Understanding the complexity of economic, ecological, and social systems. Ecosystems, **4**:390-405.

Jones, C.G., Lawton, J.H., and Shachak, M., 1997. Positive and negative effects of organisms as physical ecosystem engineers. Ecology, **78**:1946-1957.

Jørgensen, S.E., Patten, B.C., and Straškraba, M., 2000. Ecosystems emerging: 4. Growth. Ecol. Model., **126**:249-284.

Krause, A.E., Frank, K.A., Mason, D.M., Ulanowicz, R.E., and Taylor, W.W., 2003. Compartments revealed in food-web structure. Nature, **426**:282-285.

Krivtsov, V., 2004. Investigations of indirect relationships in ecology and environmental sciences: a review and the implications for comparative theoretical ecosystem analysis. Ecol. Model., **174**:37-54.

Lafferty, K.D., and Morris, A.K., 1996. Altered behavior of parasitized killifish increases susceptibility to predation by bird final hosts. Ecology, **77**:1390-1397.

Laland, K.N., Odling-Smee, F.J., and Feldman, M.W., 1999. Evolutionary consequences of niche construction and their implications for ecology. Proc. Natl. Acad. Sci. USA, **96**:10242-10247.

Leontief, W.W., 1965. The structure of the American economy. Sci. Am., **212**:25-35.

Leontief, W.W., 1966. Input-output economics. Oxford University Press, New York.

Loehle, C., 1990. Indirect effects - a critique and alternate methods. Ecology, **71**:2382-2386.

Luttbeg, B., Rowe, L., and Mangel, M., 2003. Prey state and experimental design affect relative size of trait- and density-mediated indirect effects. Ecology, **84**:1140-1150.

Matis, J.H., and Patten, B.C., 1981. Environ analysis of linear compartmental systems: the static, time invariant case. Bull. Int. Stat. Inst., **48**:527-565.

Miller, T.E., and Kerfoot, W.C., 1987. Redefining indirect effects. In: W.C. Kerfoot and A. Sih (Editors), Predation: direct and indirect impacts on aquatic communities, University Press of New England, London, pp 33-37.

Miller, T.E., and Travis, J., 1996. The evolutionary role of indirect effects in communities. Ecology, **77**:1329-1335.

Odling-Smee, F.J., Laland, K.N., and Feldman, M.W., 2003. Niche construction: the neglected process in evolution. Princeton University Press, Princeton, N.J.

Palmer, M.A., Bernhardt, E.S., Chornesky, E.A., Collins, S.L., Dobson, A.P., Duke, C.S., Gold, B.D., Jacobson, R., Kingsland, S., Kranz, R., Mappin, M.J., Martinez, M.L., Micheli, F., Morse, J.L., Pace, M.L., Pascual, M., Palumbi, S., Reichman, O.J., Townsend, A., and Turner, M.G., 2004. Ecological science and sustainability for a crowded planet: 21st century vision and action plan for the Ecological Society of America. Ecological Society of America.

Palmer, M.A., Bernhardt, E.S., Chornesky, E.A., Collins, S.L., Dobson, A.P., Duke, C.S., Gold, B.D., Jacobson, R.B., Kingsland, S.E., Kranz, R.H., Mappin, M.J., Martinez, M.L., Micheli, F., Morse, J.L., Pace, M.L., Pascual, M., Palumbi, S.S., Reichman, O., Townsend, A.R., and Turner, M.G., 2005. Ecological science and sustainability for the 21st century. Front. Ecol. Environ., **3**:4-11.

Patten, B.C., 1978. Systems approach to the concept of environment. Ohio J. of Sci., **78**:206-222.

Patten, B.C., 1981. Environs: the superniches of ecosystems. Am. Zool., **21**:845-852.

Patten, B.C., 1982. Environs: relativistic elementary particles for ecology. Am. Nat., **119**:179-219.

Patten, B.C., 1983. On the quantitative dominance of indirect effects in ecosystems. In: W.K. Lauenroth, G.V. Skogerboe, and M. Flug (Editors), Analysis of ecological systems: state-of-the-art in ecological modelling, Elsevier, Amsterdam, pp 27-37.

Patten, B.C., 1984. Further developments toward a theory of the quantitative importance of indirect effects in ecosystems. Verh. Gesellschaft für Ökologie, **13**:271-284.

Patten, B.C., 1985. Energy cycling in the ecosystem. Ecol. Model., **28**:1-71.

Patten, B.C., 1990. Environ theory and indirect effects: a reply to Loehle. Ecology, **71**:2386-2393.

Patten, B.C., 1991. Network ecology: indirect determination of the life-environment relationship in ecosystems. In: M. Higashi and T. Burns (Editors), Theoretical Studies of Ecosystems: the network perspective, Cambridge University Press, New York, New York, pp 288-351.

Patten, B.C., 1992. Energy, emergy and environs. Ecol. Model., **62**:29-69.

Patten, B.C., in prep. Holoecology: the unification of nature by network indirect effects. Columbia University Press, New York.

Patten, B.C., Bosserman, R.W., Finn, J.T., and Cale, W.G., 1976. Propagation of cause in ecosystems. In: B.C. Patten (Editor), Systems Analysis and Simulation in Ecology, Vol. IV, Academic Press, New York, pp 457-579.

Patten, B.C., Higashi, M., and Burns, T.P., 1990. Trophic dynamics in ecosystem networks: significance of cycles and storage. Ecol. Model., **51**:1-28.

Patten, B.C., and Matis, J.H., 1982. The water environs of the Okefenokee swamp: an application of static linear environ analysis. Ecol. Model., **16**:1-50.

Patten, B.C., Richardson, T.H., and Barber, M.C., 1982. Path analysis of a reservoir ecosystem model. Can. Water Resourc. J., **7**:252-282.

Peacor, S.D., and Werner, E.E., 1997. Trait-mediated indirect interactions in a simple aquatic food web. Ecology, **78**:1146-1156.

R Development Core Team, 2004. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reckhow, K.H., and Gray, J., 2000. Neuse River Estuary Modeling and Monitoring Project Stage 1: Stage 1 Executive Summary and Long-term Modeling Recommendations. Report 325-A, Water Resources Research Institute of the University of North Carolina, Raleigh, NC.

Rizzo, W.M., and Christian, R.R., 1996. Significance of subtidal sediments to heterotrophically-mediated oxygen and nutrient dynamics in a temperate estuary. Estuaries, **19**:475-487.

Rizzo, W.M., Lackey, G.J., and Christian, R.R., 1992. Significance of euphotic, subtidal sediments to oxygen and nutrient cycling in a temperate estuary. Mar. Ecol.-Prog. Ser., **86**:51-61.

Samuelson, P.A., 1948. Economics, an introductory analysis, 1st edition. McGraw-Hill Book Co., New York,.

Schramski, J.R., Gattie, D.K., Patten, B.C., Borrett, S.R., Fath, B.D., Thomas, C.R., and Whipple, S.J., 2005. Indirect effects and distributed control in ecosystems 5. Distributed control in the environ networks of a seven-compartment model of nitrogen flow in the Neuse River Estuary, USA: steady-state analysis. Ecol. Model., **XXX**:XXX-XXX.

Simon, H.A., 1962. The architecture of complexity. Proc. Am. Phil. Soc., **106**:467-482.

Strauss, S.Y., 1991. Indirect effects in community ecology - their definition, study and importance. Trends Ecol. Evol., **6**:206-210.

Thomashow, M., 2002. Bringing the biosphere home: learning to perceive global environmental change. MIT Press, Cambridge, Mass.

Ulanowicz, R.E., 1986. Growth and development: ecosystems phenomenology. Springer-Verlag, New York.

Werner, E.E., and Peacor, S.D., 2003. A review of trait-mediated indirect interactions in ecological communities. Ecology, **84**:1083-1100.

Whipple, S.J., 1999. Analysis of ecosystem structure and function: extended path and flow analysis of a steady-state oyster reef model. Ecol. Model., **114**:251-274.

Whipple, S.J., and Patten, B.C., 1993. The problem of nontrophic processes in trophic ecology: Toward a network unfolding solution. J. Theor. Biol., **163**:393-411.

Wootton, J.T., 1993. Indirect effects and habitat use in an intertidal community: interaction chains and interaction modifications. Am. Nat., **141**:71-89.

Wootton, J.T., 1994. The nature and consequences of indirect effects in ecological communities. Ann. Rev. Ecol. Syst., **25**:443-466.

Wootton, J.T., 2002. Indirect effects in complex ecosystems: recent progress and future challenges. Journal of Sea Research, **48**:157-172.

4.7 TABLES

Table 4.1: Results of two-way ANOVA indicating temporal variation (seasonal and interannual) of indirect flow index (IFI)

source of variation	df	SSE	MSE	F value	Pr(>F)
Season	3	0.012119	0.00404	3.67	0.05637 •
Year Class	3	0.002369	0.00079	0.7174	0.56616
Residuals	9	0.009906	0.001101		

· indicates significance at α = 0.1 level.

Table 4.2: Results of a pairwise comparison (t-test with non-pooled SD) of seasonal mean indirect flow index (IFI) indicating no significant differences. Values in the table are Bonferroni corrected p-values.

	Fall	Spring	Summer
Spring	1	-	-
Summer	0.89	0.52	-
Winter	0.81	1	0.55

4.8 FIGURE LEGENDS

Figure 4.1: Nitrogen cycling model of the Neuse River Estuary (modified from Christian and Thomas (2000)).

Figure 4.2: Seasonal total system throughflow (TST, mmol N m^{-2} $season^{-1}$, left axis) and proportion of total system throughflow due to cycling (Finn Cycling Index, FCI, right axis) in the Neuse Estuary from spring 1985 to winter 1989. The mean cycling index was 0.88 (±0.09 SD) and the mean total system throughflow was 10,373 mmol N m^{-2} $season^{-1}$ (± 4,207 SD).

Figure 4.3: Indirect effects in 16 seasonal models of nitrogen cycling of the Neuse River Estuary (Spring 1985 to Winter 1989). A) Ratio of indirect–to–direct flow partitions of TST. In all seasons, the Indirect/Direct ratio is substantially greater than one (marked by thin horizontal line), indicating the dominance of indirect effects. B) Seasonal partition of TST into boundary, direct, and indirect flow components.

Figure 4.4: Temporal variation in indirect flow index (IFI = Indirect/TST). A) Seasonal. B) Interannual, based on year classes grouping 4 sequential seasons to form a year class (i.e., yc1985 = {spring 1985, summer 1985, fall 1985, and winter 1986}). There are no statistically significant differences between seasons or years classes.

Figure 4.5: Relationship between flow types in 16 seasonal nitrogen cycling models of the Neuse River Estuary (1985-1989). A) Indirect flow shows no clear relationship with either boundary or direct flow; however, B) there is a strong linear relationship between boundary and direct flows (Direct = 0.96*Boundary, $R^2 = 0.99$, $p<0.001$).

Figure 4.6: Relationship between indirect and cycled flows in the 16 seasonal nitrogen cycling models of the Neuse River Estuary (Indirect = 906 + 0.98*Cycled, $R^2 = 0.99$, $p<0.001$).

Figure 4.1

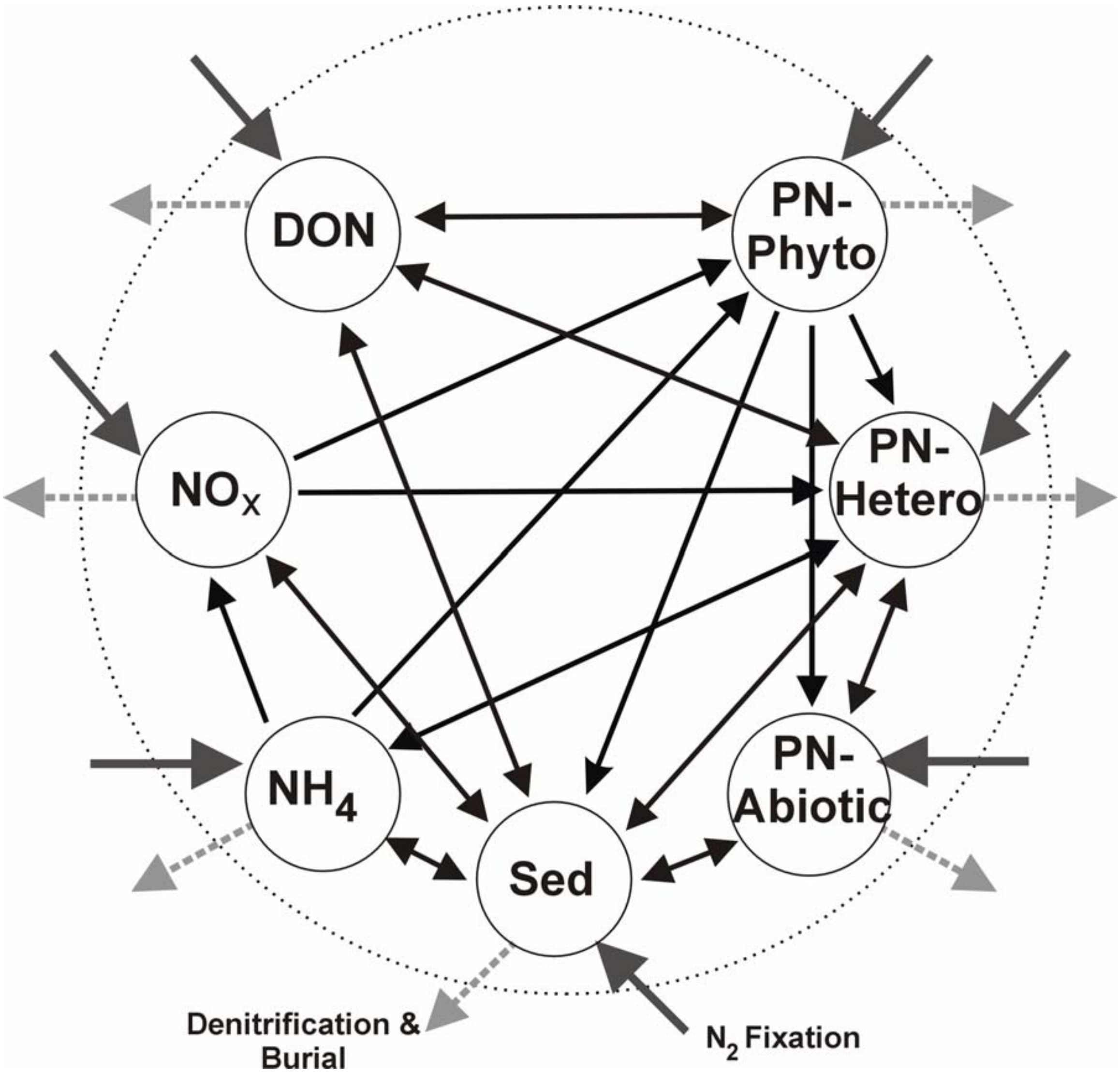
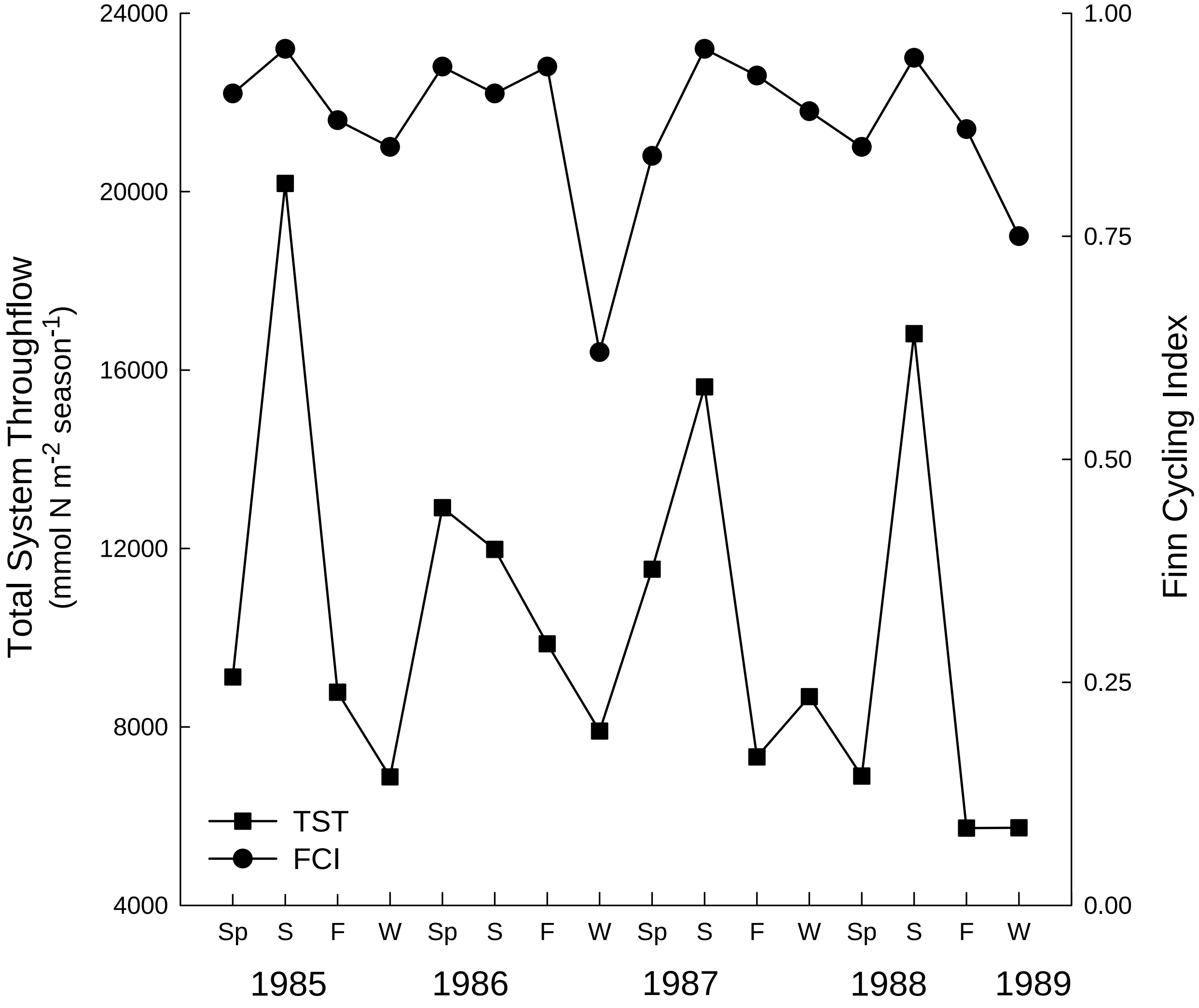

DON
PN-Phyto
NO$_X$
PN-Hetero
NH$_4$
PN-Abiotic
Sed
Denitrification & Burial
N$_2$ Fixation

Figure 4.2



24000
20000
16000
12000
8000
4000
Total System Throughflow
(mmol N m^{-2} season^{-1})
1.00
0.75
0.50
0.25
0.00
Finn Cycling Index
TST
FCI
Sp S F W Sp S F W Sp S F W Sp S F W
1985
1986
1987
1988
1989

Figure 4.3

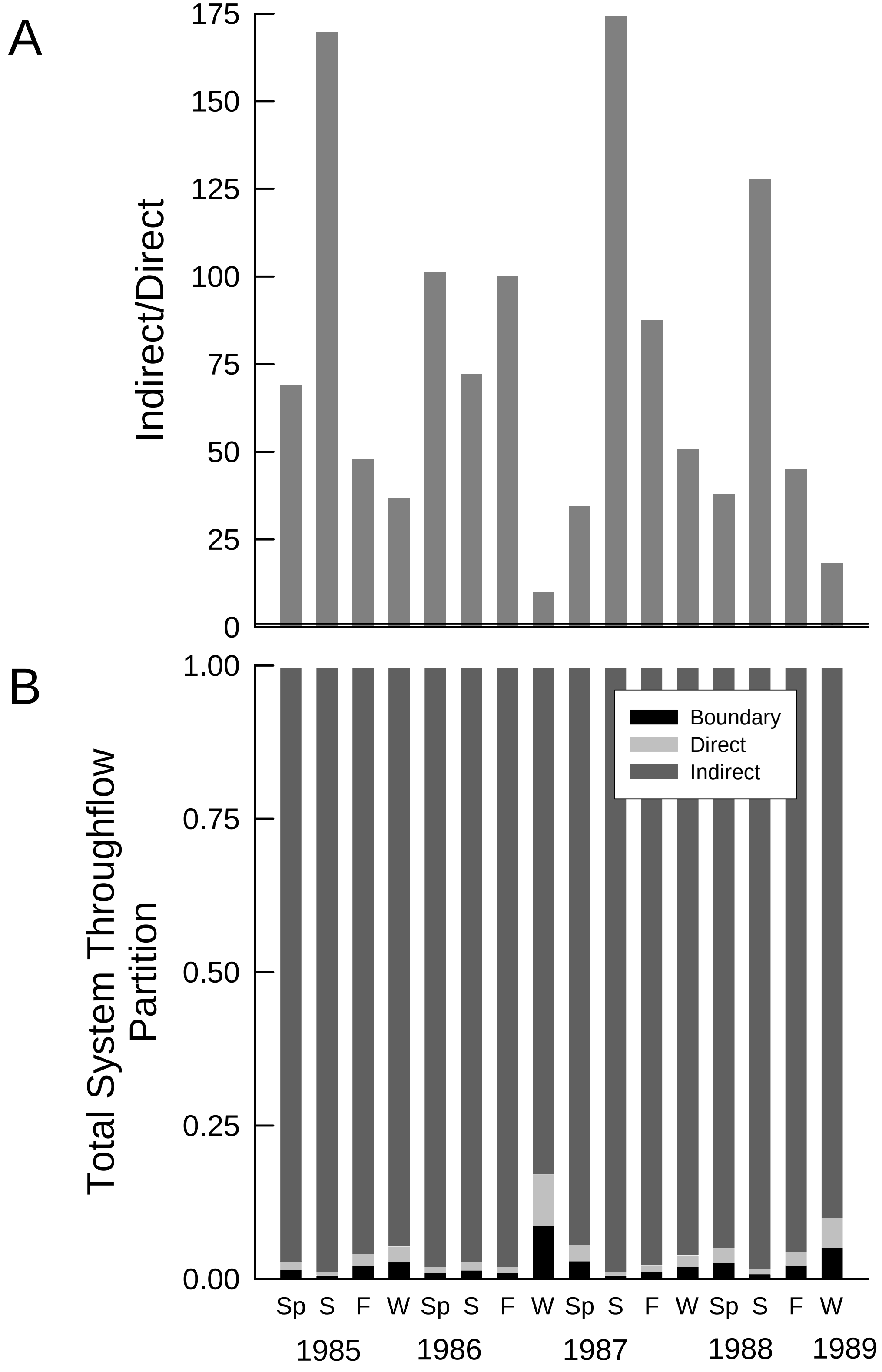

A
175
150
125
100
75
50
25
0
Indirect/Direct
B
1.00
0.75
0.50
0.25
0.00
Total System Throughflow Partition
Boundary
Direct
Indirect
Sp
S
F
W
Sp
S
F
W
Sp
S
F
W
Sp
S
F
W
1985
1986
1987
1988
1989

Figure 4.4

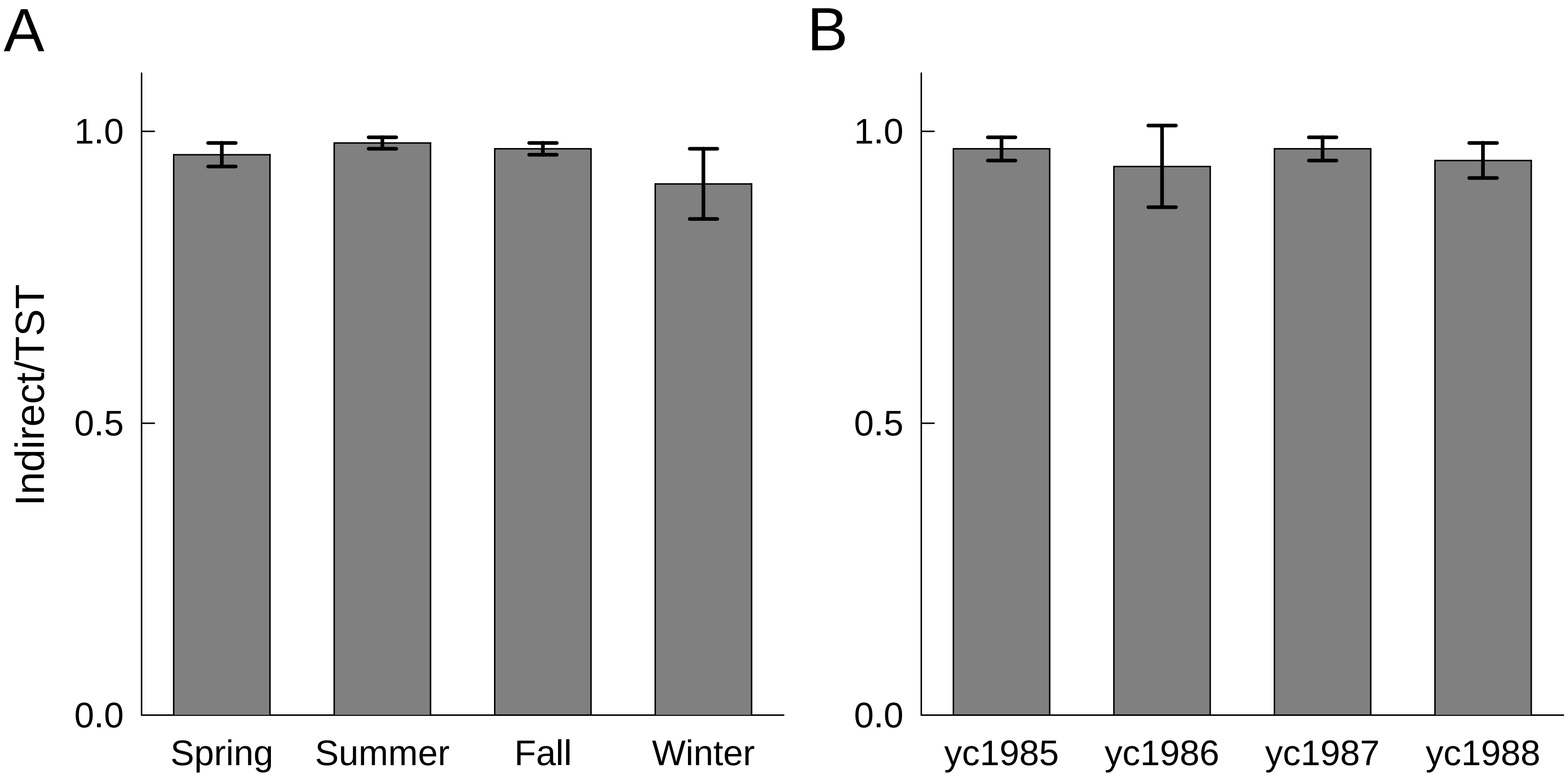

A
B
Indirect/TST
0.0
0.5
1.0
Spring
Summer
Fall
Winter
0.0
0.5
1.0
yc1985
yc1986
yc1987
yc1988

Figure 4.5

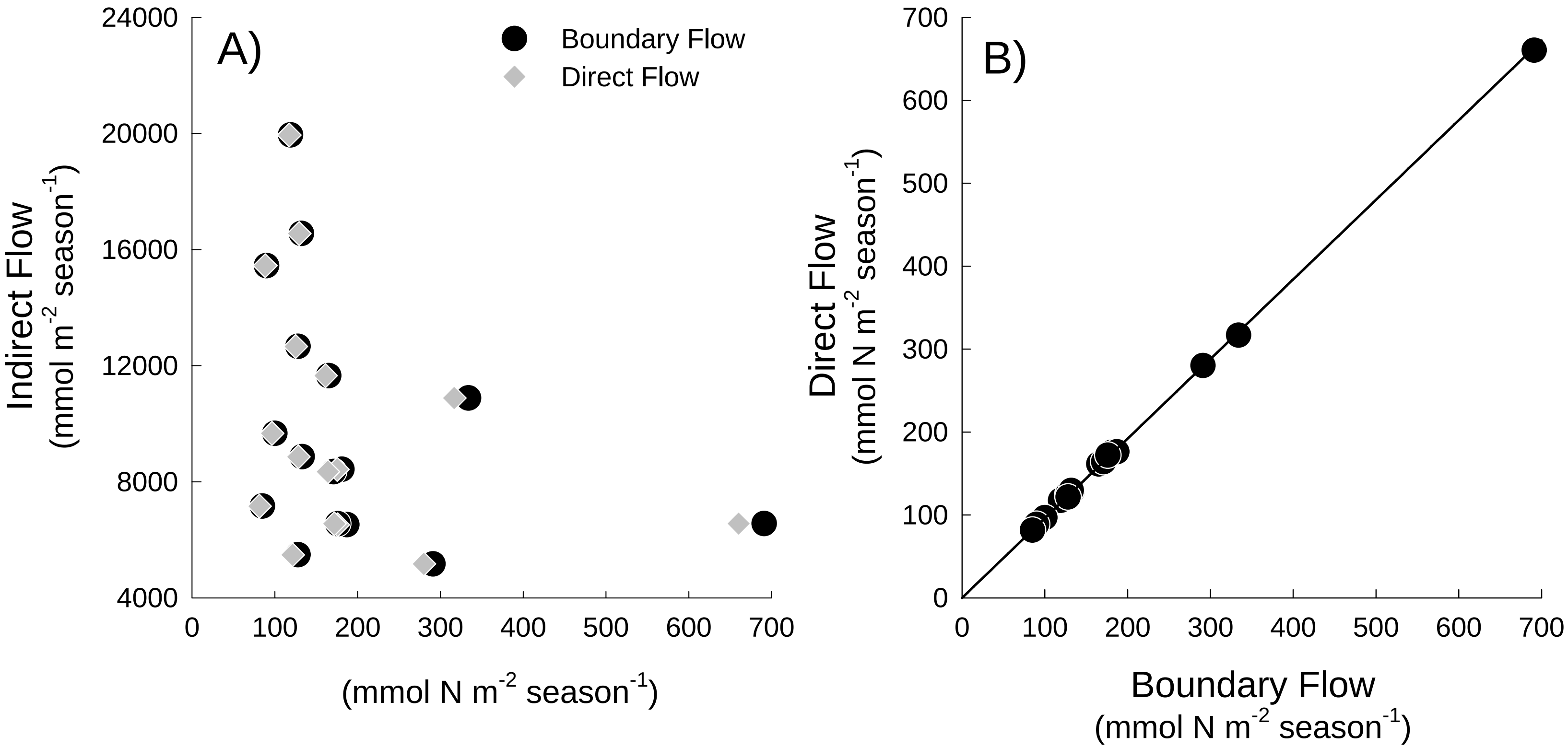
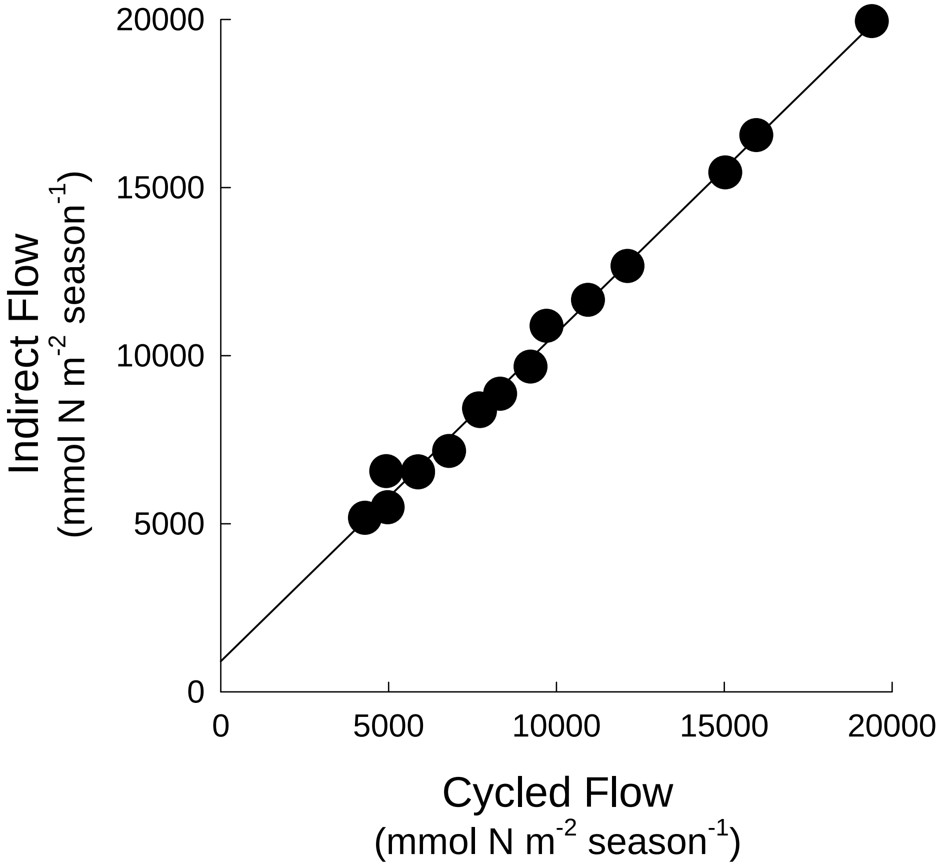

A)
Boundary Flow
Direct Flow
Indirect Flow
(mmol m-2 season-1)
24000
20000
16000
12000
8000
4000
0
100
200
300
400
500
600
700
(mmol N m-2 season-1)
B)
Direct Flow
(mmol N m-2 season-1)
700
600
500
400
300
200
100
0
Boundary Flow
(mmol N m-2 season-1)

Figure 4.6


20000
15000
10000
5000
0
0
5000
10000
15000
20000
Indirect Flow
(mmol N m^{-2} season^{-1})
Cycled Flow
(mmol N m^{-2} season^{-1})

CHAPTER 5

ENVIRON INDICATOR SENSITIVITY TO FLUX UNCERTAINTY IN A PHOSPHORUS MODEL OF LAKE SIDNEY LANIER, USA[1]

---

[1] Borrett, S.R., O.O. Osidele, M.B. Beck. To be submitted to *Ecological Modelling*

ABSTRACT

Effective environmental impact assessment and management requires improved understanding of the organization and transformation of environmental systems, especially ecosystems. Ecosystems are complex adaptive systems in which many independent agents are linked through an intricate network of energy, matter, and informational interactions. While advances have been made in the last three decades, we still lack a complete theoretical understanding of the processes that create and maintain ecosystems. Network Environ Analysis (NEA) is one type of ecosystem network analysis distinguished by its explicit environment focus. It provides one approach for building novel ecosystem insights. NEA, however, is applied to mathematical models of natural systems, and is thus model dependent. As ecological modeling is an imprecise art, often complicated by incomplete empirical data for model calibration and evaluation, the utility of NEA may be limited by model uncertainty. In this work, we investigated the sensitivity of nine whole-system indicators of ecosystem growth and development to flow and storage uncertainty in a phosphorus model of Lake Sidney Lanier, USA. The indicators included total system throughflow (TST), total system storage (TSS), total boundary input (Boundary), Finn cycling index (FCI), ratio of indirect-to-direct flows (Indirect/Direct), indirect flow index (IFI), network aggradation (AGG), network homogenization (HMG), and network amplification (AMP). The results of our study make two primary contributions to ecosystem network analysis. First, they demonstrate that five of the ecosystem network analysis indicators – FCI, Indirect/Direct, IFI, AGG and HMG – were relatively robust to the flow and storage uncertainty in the Lake Lanier phosphorus model. This allows us to circumvent at least part of the modeling problem to draw stronger conclusions about the organization of the Lake Lanier ecosystem. Phosphorus flux in the lake is heavily influenced

if not dominated by internal processes. Second, the results show that the majority of the nine ecosystem indicators investigated co-varied and that their common variation could largely be mapped into two latent factors. We have tentatively interpreted these two factors as 1) system integration and 2) boundary influences.

"The *desire* for simplification is justified, but the same does not always apply to simplification itself, which is a working hypothesis, useful as long as it is recognized as such and not mistaken for reality. The greater part of historical and natural phenomena are [sic] not simple, or not simple in the way that we would like." (Levi, 1988)

5.1 INTRODUCTION

Ecosystems are open non-equilibrium thermodynamic systems knit together by an intricate network of energy, matter, and information exchanges among biological organisms and their environments (Ulanowicz, 1986; Higashi and Burns, 1991; Capra, 1996; Patten, 1998; Jørgensen, 2002). Like other kinds of complex adaptive hierarchical systems, they appear to self-organize in response to thermodynamic gradients (Schneider and Kay, 1994; Müller, 1996; Levin, 1998; Patten et al., 2002). A clear formal understanding of ecosystem organization and transformation has yet to emerge, however, despite several empirical and theoretical attempts (e.g., Odum, 1969; Ulanowicz, 1986; Gunderson and Holling, 2002; Jørgensen, 2002; Fath et al., 2004). This understanding is critical as ecosystems provide the natural capital and services that support human endeavors (Costanza et al., 1997; Daily, 1997); it is essential for effective ecosystem assessment and management (Christensen, 1996; Reichman and Pulliam, 1996).

Ecosystem science has a long history of characterizing patterns of organization and development (e.g., Lindeman, 1942; Odum, 1962; Teal, 1962; Margalef, 1963; Odum, 1969). Numerous measures of ecosystem organization have been proposed including gross production-to-respiration ratio, gross production-to-biomass ratio, species diversity, flow diversity, cycling (Odum, 1969), emergy (Odum, 1983), and ascendency (Ulanowicz, 1986, 1997). As systems develop, several of these measures appear to consistently increase or decrease, suggesting possible orienting forces or constraints on self-organization (Odum, 1969; Ulanowicz, 1986; Schneider and Kay, 1994; Müller and Leupelt, 1998; Fath et al., 2001).

Jørgensen (2002), in his efforts to weave together multiple ecosystem theories into a coherent whole, acknowledged that a plurality of approaches remains necessary. Holoecology (Patten, in prep.) offers one distinct approach. The overarching goals of the Holoecology Research Program are 1) to develop a deep understanding of the lawful processes that create, constrain, and sustain ecological systems and 2) to create a formal theory of environment. Network Environ Analysis (NEA) is the primary methodology of holoecology (Patten, 1978; Matis and Patten, 1981; Patten, 1982; Fath and Borrett, 2005; Patten, in prep.). Fundamentally, NEA is an environmental application and extension of economic Input–Output Analysis (Leontief, 1966). Though it has elements in common with other forms of ecosystem network analysis such as Input-Output Analysis as implemented in Ecopath (Christensen and Pauly, 1992; Christensen and Walters, 2004) or Acendency Theory (Ulanowicz, 1986, 1997; Allesina and Bondavalli, 2004), NEA is distinguished by its explicit environment focus. It is used in the holistic study of ecological networks to describe and quantify component-level bounded environments within systems, termed environs (Patten, 1978, 1981, 1982).

A particular strength of NEA is that it includes several whole-system indicators that summarize system organization including total system storage (TSS), total boundary input or output (Boundary), total system throughflow (TST), Finn cycling index (FCI), average path length or network aggradation (AGG), ratio of indirect-to-direct flows (Indirect/Direct), indirect flow index (IFI), network homogenization (HMG), and network amplification (AMP). TSS, Boundary, TST, FCI, and AGG are commonly used in other forms of ecosystem network analysis; Indirect/Direct, IFI, HMG, and AMP are unique to NEA (Patten, 1998; Fath and Patten, 1999b). Collectively, these indicators provide insight into the flow and storage organization of ecosystem models. Based in part on the application of these indicators to ecosystem models, Patten (1998; in prep.) argues that ecological systems are integrated by indirect effects.

The Achilles heel of these indicators, NEA, and ecosystem network analyses in general is that they are based on models of real ecosystems. Ecological modeling remains a challenging endeavor with no consensus on the best approach or evaluation procedures (Levins, 1966; Oreskes et al., 1994; Oreskes, 1998; Ginzburg and Jensen, 2003). In truth, an infinite number of models fit a given data set, and choosing among models remains challenging. As the opening quote from Levi (1988) states, natural phenomena are not always simple in the way we would like. Previous research revealed that ecosystem indicators can be highly dependent upon the model form, especially the degree of model aggregation (lumping vs. splitting) (Cale and Odell, 1979; Gardner et al., 1982; Cale and O'Neill, 1988; Pahl-Wostl, 1992; Sugihara et al., 1997; Yodzis and Winemiller, 1999; Thompson and Townsend, 2000; Abarca-Arenas and Ulanowicz, 2002). Less research has explored the sensitivity of these indicators to uncertainty in the magnitude of flows and storages (but see Bosserman, 1983; Fath, 2004). This type of uncertainty can arise from incomplete or unknown data for model calibration and evaluation or empirical measurement error. In the work reported here, we investigated the sensitivity and interrelations of the indicators of ecosystem organization listed above to flux and storage uncertainty in a phosphorus based ecosystem model of Lake Sidney Lanier, USA. We assumed a particular model structure (compartments and connections) and held it constant. We then used Monte Carlo simulations to generate a population of plausible parameterizations given a limited set of empirical data which lead to different combinations of phosphorus flows and storages.

We investigated two specific hypotheses. First, we hypothesized that six of our focal system indicators (FCI, Indirect/Direct, IFI, AGG, AMP, and HMG) would be robust under model flux uncertainty. Quantitatively we expected each of these indicators to be relatively less variable than TST and TSS; qualitatively we expected their interpretations to be consistent. Support for this hypothesis will allow us to draw more robust conclusions about Lake Lanier despite model uncertainty, effectively circumventing part of the modeling problem. Second, we hypothesized that these six indicators measure slightly different aspects of the same unmeasured underlying or latent factor, which we tentatively term system integration by indirect effects. This hypothesis implies that the indicators will be highly correlated. We also expected they would largely map into one principle latent factor. Support for this hypothesis would potentially allow us to reduce the number of NEA indicators required to characterize the degree of system integration by indirect effects.

5.2 MATERIALS AND METHODS

5.2.1 STUDY SYSTEM

Lake Sydney Lanier is a large reservoir in the headwaters of the Chattahoochee River in Georgia, USA. It drains a 2704 km^2 watershed, and at the conservation pool elevation of 326.4 m covers 150 km^2 with 869 km of shoreline (Guan, 1993; Fath and Beck, 2005). Constructed in the 1950's by the U.S. Army Corp of Engineers, the reservoir was initially to provide flood control, hydroelectric power, and downstream navigation regulation (US Army Corps of Engineers, 2005). The reservoir lies just northeast of downtown Atlanta, one of the fastest growing metropolitan areas in the United States for the past two decades (Metro Atlanta Chamber of Commerce, 2000). It has become an increasingly important source of drinking water, wastewater dilution, and recreation for the region. Lake Lanier is a key resource supporting the expansion of Atlanta and the rapid urbanization of its own watershed (Fath and Beck, 2005)

5.2.2 MODEL CONSTRUCTION

For this work, we started with a modified version of the thirteen compartment Lanier ecosystem model introduced by Osidele and Beck (2004). The original model was converted to a consistent currency – phosphorus – and non-linear mechanistic functions were replaced with phenomenological donor controlled functions so that the model would readily achieve steady-state. These alterations were essential to meet the assumptions of NEA. In our model, the donor controlled functions are represented by first-order differential equations of the generic form:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = -\lambda_X \cdot X \tag{5.0}$$

where X is a generic compartmental state variable, λ_X is a rate constant, and t is continuous time. Because the function is donor controlled, the negative sign indicates a loss or transfer from the donor compartment. Our modified model has eleven state variables or storage compartments (mg P m^{-2}), twenty six within–system flows (mg P m^{-2} d^{-1}), five boundary inputs, and 4 boundary losses (Figure 5.1). The compartments represent 1) epilimnion phosphorus (EPI-P), 2) hypolimnion phosphorus (HYPO-P), 3) phytoplankton (PHYTO), 4) microbes (bacteria, etc.; MIC-B), 5) microzooplankton (MIC-Z), 6) macrozooplankton (MAC-Z), 7) macroinvertebrates (insect larvae, etc.; MAC-I), 8) larval/juvenile fish (FISH), 9) suspended detritus (SSD), 10) sediment detritus (SEDD), and 11) pore space phosphorus (PORE). Phosphorus enters the model from the Chattahoochee River, Chestatee River, and the watershed via fluxes into EPI-P, HYP-P, FISH, SSD, and PORE. Phosphorus is lost from the system from EPI-P, PHYTO, MAC-I, and FISH.

In this investigation model structure (network topology) was fixed, while flow and storage magnitudes were variable. We used Monte Carlo simulations (Manly, 1991) and regionalized sensitivity analysis (Osidele and Beck, 2001) to identify parameter values that generated model behavior that matched the observed summer behavior in the Lake. Behavior was defined as a range of values for phosphorus, phytoplankton and fish in order to account for uncertainty and spatial variability in the available observational data. Thus, model behavior matched empirical behavior (i.e., successful) when the model ouputs fell within the prescribed range of values. Data for phosphorus and phytoplankton were obtained from studies conducted for the Georgia Department of Natural Resources (DNR) under the U.S. Environmental Protection Agency's Clean Lakes Program (Hatcher et al., 1994). Data on fish were obtained from fisheries management reports of the DNR Wildlife Resources Division (Weaver, 2000). These data records indicate that average annual phosphorus concentration in the photic zone ranges from 2 to 4 μg L^{-1}, with no distinct spatial variation across the reservoir. Phytoplankton chlorophyll concentrations range from 7 μg L^{-1} at the tributary inlets to 2 μg L^{-1} at the dam site, and larval-juvenile fish average 20 to 50 kg.ha^{-1} lakewide. These observations were converted to the model currency and units for prescribing the following summer behavior definition:

- steady state epilimnion SRP concentration should be between 58 – 70 mg P m^{-2};
- steady state phytoplankton biomass should be between 13 – 52 mg P m^{-2};
- steady state larval-juvenile fish biomass should be between 100 – 250 mg P m^{-2}.

Thirty parameters were considered in the regionalized sensitivity analysis. They consisted of the 30 mass transfer rate constants λ_X in (5.0), representing 26 within-system flows and 4 boundary losses. Each parameter was sampled from a uniform distribution over the interval [0.0., 0.5]. To insure our sample of parameter values was representative of the larger population of possibilities, we compared mean and variance of the ecosystem indicators in successful samples from trial runs of 500, 1000, and 5000 realizations. We used a Welsh t-test to compare the means as it does not assume identical variance (Dalgaard, 2002).

5.2.3 NETWORK ENVIRON ANALYSIS – THROUGHFLOW ANALYSIS

NEA (Patten et al., 1976; Patten, in prep.) is a family of input–output methods descended from economics (Leontief, 1965, 1966) that analytically decomposes observed system flows and storages to identify their origins or fates within a system of interest (see Fath and Patten, 1999b for review). NEA methods are extensively described in the literature (Patten et al., 1976; Matis and Patten, 1981; Fath and Patten, 1999b; Patten, in prep.) and include input and output analyses of structure, throughflow, storage, utility, and control within systems. In this paper, we concentrate on the output oriented throughflow analyses. We briefly review the NEA methods used in this paper and formally define the system indicators investigated (Table 5.1).

Let $F_{n\times n} = (f_{ij})$, i,j = 1,…, n, and $f_{ii} = 0$ be the empirically observed flows from compartment j to i, z_i be the inputs to i, y_i be the outputs from i, and x_i be the storage values of an n compartment system. The total currency (e.g., energy, carbon, nitrogen, phosphorus) flowing into or out of a compartment, called throughflow, is defined as $T_k^{(in)} \equiv \sum_{j(\neq k)=1}^{n} f_{kj} + z_k$ and $T_k^{(out)} \equiv \sum_{i(\neq k)=1}^{n} f_{ik} + z_k$, respectively. At steady state, energy–matter continues to flow through the system, but storage is constant through time, whereupon $T_k^{(in)} = T_k^{(out)} = T_k \quad k = 1,\ldots,n$. Total system throughflow ($TST = \sum T_k$) is a system-level measure of activity often used to characterize ecosystem models (Finn, 1976; Ulanowicz, 1986)[2]. Fath et al. (2001) further suggested that, at least in models of energy flux, TST was a network measure of system power in the sense of Lotka (1922). Total system storage ($\mathrm{TSS} = \sum_{i=1}^{n} x_i$) is the total amount of currency stored in the system compartments. Jørgensen and Mejer (1979) suggested that TSS would tend to be maximized during ecosystem growth and development. Total boundary input ($\sum_{i=1}^{n} z_i$) indicates the amount of energy-matter initially available to the system, while total boundary output ($\sum_{i=1}^{n} y_i$) is the amount lost to the environment. Again, at steady-state these are equal such that total boundary flow (Boundary) is $\mathrm{Boundary} = \sum_{i=1}^{n} z_i = \sum_{i=1}^{n} y_i$.

At least four variations for partitioning throughflow among different pathway types have been reported previously, three of which build off the original Leontief analysis (Gattie et al., in prep.). Many of the indicators we investigate here are derived from two of these: Leontief and Finn.

5.2.3.1 LEONTIEF

The first type of throughflow partition, the Leontief model, is based on Leontief's original economic Input-Output Analysis (Leontief, 1966). To determine the integral (boundary

---

[2] Total system throughflow as defined here is consistent with most applications of NEA, but different from what Ulanowicz (1986) calls total system throughput (TSTP). TSTP is the summation of all within system flows, boundary inputs and boundary outputs, such that $\mathrm{TSTP} = \sum\sum f_{ij} + \sum z_i + \sum y_i$, where as $\mathrm{TST} = \sum\sum f_{ij} + \sum z_i = \sum\sum f_{ij} + \sum y_i$.

+ direct + indirect) amount of material and service required to generate a given product, he developed a matrix to map outputs into throughflow. Augustinovics (1970) turned the problem around to find a matrix to map inputs into throughflows, which we will denote as N. The output oriented Leontief decomposition can initially be written as T = Nz. N can be derived by first normalizing the observed flows by the donor compartment throughflow ($G = (g_{ij}) = \left(\frac{f_{ij}}{T_j}\right)$). The resultant matrix, $G_{n\times n}$, represents the donor-specific direct flow intensities from j to i. The elements of G are interpreted as the probability that material entering j across the boundary will flow to i over a direct path (i.e., path length, m = 1). Flow from j to i over indirect pathways (m > 1) is then determined by raising G to the m^{th} power, G^m. The elements of G^m are typically interpreted as the probability of material flowing over all pathways of length m between any two compartments. Total node throughflow is recovered by post multiplying the sum of the infinite power series of flow intensities by the boundary inflows:

$$T = \left(\underbrace{I}_{\text{Boundary}} + \underbrace{G^1}_{\text{Direct}} + \underbrace{G^2 + \ldots G^m + \ldots}_{\text{Indirect}} \right) * z, \tag{5.1}$$

where $I = G^0$ is the matrix multiplicative identity. In open, dissipative, thermodynamic systems like ecosystems, the power series converges to the Leontief transitive closure matrix $N = (n_{ij}) = (I - G)^{-1}$, such that ($n_{ij}$) represents the integral (boundary + direct + indirect) throughflow from j to i generated by a boundary input to j. This power series decomposition reveals that, at steady-state the empirically observed flows are a composite of boundary, direct, and indirect flows. Patten (in prep.) terms this network enfolding, which is one of the cardinal hypotheses of Holoecology (see Chapter 2).

For a system level partition, we can rewrite the throughflow decomposition in (5.1) by distributing z across the flow intensity partition, and summing the elements of each n×1 resultant vector to derive the following restatement in terms of TST:

$$\mathrm{TST} = \underbrace{\sum \mathrm{Iz}}_{\text{Boundary}} + \underbrace{\sum \mathrm{Gz}}_{\text{Direct}} + \underbrace{\sum (\mathrm{N}-\mathrm{I}-\mathrm{G})\mathrm{z}}_{\text{Indirect}}. \tag{5.2}$$

Dividing both sides of (5.2) by TST generates:

$$1 = \frac{\sum \mathrm{z}}{\mathrm{TST}} + \frac{\sum \mathrm{Gz}}{\mathrm{TST}} + \frac{\sum (\mathrm{N}-\mathrm{I}-\mathrm{G})\mathrm{z}}{\mathrm{TST}}. \tag{5.3}$$

The boundary flow index ($\frac{\sum \mathrm{z}}{\mathrm{TST}}$), direct flow index ($\frac{\sum \mathrm{Gz}}{\mathrm{TST}}$), and the indirect flow index ($\mathrm{IFI} = \frac{\sum (\mathrm{N}-\mathrm{I}-\mathrm{G})\mathrm{z}}{\mathrm{TST}}$) are dimensionless numbers between 0 and 1 that indicate the proportion of TST derived from the three categories of flow in the Leontief type decomposition. IFI was first used to compare the seasonality of indirect effects in the Neuse River Estuary (Borrett et al., in press, Chapter 4). Notice that in (5.2) and (5.3), direct effects are defined as only the first transfer of the boundary inputs within the system; indirect effects encompass all other transfers.

Three NEA indicators have been developed from the Leontief partition to characterize the flow organization in model systems. The first is the ratio of indirect-to-direct flow (Higashi and Patten, 1989; Fath and Patten, 1999b; Borrett et al., in press). This is calculated using the elements of the Leontief decomposition in (5.2) as follows:

$$\mathrm{Indirect}\big/\mathrm{Direct} = \frac{\sum (\mathrm{N}-\mathrm{I}-\mathrm{G})\mathrm{z}}{\sum \mathrm{Gz}}. \tag{5.4}$$

When greater than unity, this ratio indicates that indirect flows dominate direct. The second indicator is network homogenization (HMG), which is said to occur when the integral flow transfer coefficients (n_{ij}) are more evenly distributed than the direct (g_{ij}) (Patten et al., 1990). It is calculated as the ratio of the coefficient of variation of G, CV(G), and the coefficient of variation of N, CV(N) (Fath and Patten, 1999a), such that

$$\mathrm{HMG} = {}^{\mathrm{CV(G)}}\!\big/_{\mathrm{CV(N)}}, \tag{5.5}$$

where the coefficient of variation is the standard deviation of the matrix elements divided by their mean. Like Indirect/Direct, when HMG is greater than unity the network operates to more evenly distribute the system resources. Thus, each node is receiving a more equal amount of flow from the other nodes over the integral (direct + indirect) pathways. The third measure is network amplification (AMP), which indicates "…that the summed total amount of flow through a compartment can be greater that the total amount of input into the network" (Fath and Patten, 1999b, p. 175). This has been quantified by counting the number of off diagonal elements of the N matrix that are greater than unity ($\# n_{ij} > 1, i \neq j$) (Patten et al., 1990). To facilitate future cross system comparisons, we report this number as the fraction of possible amplified positions, such that $\mathrm{AMP} = \dfrac{(\# n_{ij} > 1)\ (i \neq j)}{n(n-1)}$.

5.2.3.2 FINN CYCLING

Finn (1976) proposed an alternative partition of TST into portions from acyclic throughflow and cycled flow. Cycled flow is calculated using the Leontief transitive closure matrix where $\mathrm{Cycled} \equiv \sum_{i=1}^{n} ((n_{ii} - 1) z_i)$, and Acyclic = TST − Cycled. Based on this flow partition, he proposed an index of cycled flow, known today as the Finn cycling index (FCI). It is the ratio of cycled flow to TST (Finn, 1980), and is a common metric used to characterize ecosystem organization. Allesina and Ulanowicz (2004) recently found that FCI slightly underestimates the magnitude of cycling; however, they report a linear association between FCI and their estimate of true cycling (1.142 * FCI) in 23 ecosystem models. Here we report the uncorrected FCI because it is not necessarily a bad estimate of cycling, it is more commonly reported in the literature, and it is simple to convert if desired.

Finn (1976) also introduced another network indicator he called average path length, calculated as the ratio of TST to total boundary input. It specifies the average number of times an average input passes through the system before exiting. Han (1997) suggested that this should be called the flow multiplying ability of the system. Patten (Jørgensen et al., 2000) re-introduced the same measure, but with a thermodynamic interpretation as an indicator of system growth and organization. He termed it network aggradation, which is the name we adopt. Thus, network aggradation (AGG) is

$$\mathrm{AGG} = \mathrm{TST} \big/ \sum z_i \, . \tag{5.6}$$

Observe that this is the inverse of the boundary flow index in (5.3). If TST is a measure of the system power, then AGG is the power generated by an average boundary input.

5.2.4 DATA ANALYSIS

We first used a modified version of NEA.m (Fath and Borrett, 2005, see Appendix A) to calculate the indicators for each of the plausible parameterizations. To address our first hypothesis that the ecosystem indicators would be robust to flow and storage uncertainty, we began by characterizing the indicator distributions. However, comparing their variability was challenging as these indicators are not measured on similar scales. For example, TST has units of mg P m^{-2}, Indirect/Direct is a dimensionless ratio where $0 \le \mathrm{Indirect/Direct} \le \infty$, and AMP is a dimensionless ratio where $0 \le \mathrm{AMP} \le 1$. To dodge this issue, we compared their relative variability by rank ordering the indicator coefficients of variability (CV). Variability of TST, TSS, and Boundary provided a baseline indication of the flow and storage variability in the plausible model parameterizations.

To evaluate our second hypothesis that the ecosystem indicators we are investigating measure different aspects of the same underlying factor, we initially used ordinary least squares regression to evaluate the correlation strength of the indicators. We then used a principle components factor analysis to identify the existence of any latent or underlying variables (Grimm and Yarnold, 1995; Johnson, 1998). We used the Kaiser criterion to select the number of relevant factors.

Calculations and statistics were accomplished using a combination of MATLAB® (version 6.5, The Mathworks, Inc.) and R (version 2.0.1, R Development Core Team, 2004).

5.3 RESULTS

Our first task was to demonstrate that our sample of plausible model parameterizations was representative of the larger population of possibilities. We ran three Monte Carlo simulations with three different samples: 500, 1000, and 5000. These produced 14, 18, and 90 plausible parameterizations that generated steady-state models with behavior that sufficiently matched the empirically known summer behavior of epilimnion soluble reactive phosphorus, phytoplankton, and fish in the lake (Figure 5.3). We then compared the mean and variability of the nine NEA indicators in the three samples (Figure 5.4). There were no statistically significant differences of the mean and variance between the three samples for TST, TSS, FCI, Indirect/Direct, IFI, AGG, and AMP. Sample variance of the total boundary inputs Boundary were significantly different, increasing with sample size. Network homogenization mean and variance was statistically different between the 500 and 5000 trials. While a larger sample size might have stabilized results for total boundary flow and homogenization, we concluded that for our purposes, the 90 plausible parameterizations would be sufficiently representative and provide a large enough sample size to address our hypotheses. All subsequent data presented is from the 90 plausible parameterizations.

While our work focuses on the whole system indicators derived from NEA, we present the average (± SD) flows and storage values for the 90 plausible parameterizations for completeness (Figure 5.2). Phosphorus stored in the epilimnion (EPI-P), phytoplankton (PHYTO), and fish (FISH) were the least variable, as expected as these were model behaviors matched to the empirical data. In contrast, the storage in the other compartments was quite variable. For example, phosphorus in the sediment pore space (PORE) had a mean of 341.3 mg P m^{-2} and a standard deviation of 588.7. Boundary flows were relatively less variable with the largest standard deviation being 21.5 associated with the largest mean boundary loss 53 mg P m^{-2} d^{-1} from FISH. Internal system flows ranged from a minimum of 4.3 mg P m^{-2} d^{-1} (± 2.7) from phytoplankton to epilimnion P to a maximum of 55.8 mg P m^{-2} d^{-1} (± 47.2) from the hypolimnion to the microbial compartment.

5.3.1 INDICATOR VARIABILITY

Our results show a range of variability among the nine network statistics evaluated (Table 5.2). For the 90 parameterizations that matched the known behavior of Lake Lanier, the mean total system throughflow (TST) was 749 (± 303) mg P m^{-2} d^{-1}, and the mean total system storage (TSS) was 1634 (± 985) mg P m^{-2}. Total phosphorus loading or boundary inputs (Boundary) was 96 (± 20) mg P m^{-2} d^{-1}. As these models were necessarily at steady-state, this is also the mean and standard deviation of the total boundary losses form the lake. Mean Finn cycling (FCI) was 0.39 (± 0.11), indicating that approximately 39% of phosphorus TST was derived from recycling in Lake Lanier. This degree of recycling is not surprising given that we are analyzing an ecosystem model of phosphorus cycling, although it does mean the majority of system activity ("power") is generated by boundary and straight chain flows. The indirect flow index (IFI) and ratio of indirect-to-direct flow (Indirect/Direct) were 0.75 (± 0.07) and 7.45 (± 2.83), respectively. These indicators suggest that a large fraction of TST was derived from indirect flows and that indirect flows were more than seven times greater than the direct flows. This is consistent with general findings from NEA theory and analyses (e.g., Higashi et al., 1989). The degrees of network aggradation (AGG), homogenization (HMG) and amplification (AMP) were 7.84 (± 2.69), 3.10 (± 0.31) and 0.21 (± 0.12), respectively. Theses indicators suggest that 1) an average input passes through an average of 7.84 nodes before exiting the system, 2) the off-diagonal integral flow probabilities were 3 times more evenly distributed than the direct flow probabilities, and 3) on average 22% of the integral flow probabilities exceeded unity. This degree of amplification indicates that quite a few nodes received more than face value of a given input, probably due to recycling.

Despite the quantitative variability derived from model uncertainty, the qualitative interpretations of the indicators are robust. In all 90 parameterizations we can conclude that recycling was a large if not dominant source of phosphorus flux. Indirect/Direct always surpassed unity implying that indirect flows were dominant, and network homogenization, amplification, and aggradation always occurred.

To compare the relative variability of the ecosystem indicators we calculated their coefficients of variation (CV; Table 5.1). The CV of TST (0.40) and TSS (0.60) reflect the whole system flow and storage variability due to model uncertainty. We used these values as benchmarks to compare the other seven indicators. IFI and HMG had the lowest CVs at 0.10, much lower than that of TST or TSS, while AMP was the most variable at 0.59. AMP was the only indicator to have a CV larger than TST, close to TSS. The CV of AGG and Indirect/Direct were similar at 0.34 and 0.38 respectively, which is quite close to that of TST. FCI was a little less variable at 0.28, and Boundary was 0.21. Given these quantitative results, the relative variability and hence the robustness of the indicators is variable. Indirect flow index and network homogenization were the most robust indicators in our sample.

We conclude that there is mixed support for our first hypothesis. The qualitative interpretations of the NEA system indicators are consistent, while their quantitative values are more variable.

5.3.2 INDICATOR INTERRELATIONS

Analysis of the relationships between the ecosystem indicators revealed several interesting patterns. For this analysis we utilized the natural log transformation of TST (ln(TST)), TSS (ln(TSS)), Indirect/Direct (ln(Indirect/Direct)), and AGG (ln(AGG)). This transformation normalized the distributions of TST and TSS, which was not true for the untransformed variables. It also straightened the relationship between ln(Indirect/Direct), ln(AGG) and the other indicators (Figure 5.5). Normal distributions and linear relations are important underlying assumptions of the statistics we used.

Pairwise scatter plots of the combinations of the nine indicators reveal several close associations (Figure 5.5). In almost all cases there is a statistically significant positive association between the variables (Figure 5.5 & Table 5.3). In contrast there are strong positive associations between FCI, ln(Indirect/Direct), IFI, ln(AGG), and AMP. Their correlation coefficients ranged from 0.89 to 0.98. Both the scatter plots and correlation coefficients suggest that several of these indicators may be influenced by the same underlying factor, lending support for our second hypothesis. In some cases, however, a linear model may not be the most appropriate (e.g., between ln(Indirect/Direct) and IFI). Boundary and homogenization are exceptions. Though Boundary does tend to increase with ln(TST) ($r = 0.46$) and ln(TSS) ($r = 0.23$), it is not significantly correlated with the other variables. Homogenization shows relatively weak correlations with the other variables, several of which are not statistically significant.

We used a principle component factor analysis to further evaluate these associations (Table 5.4). After evaluating several alternatives (not shown) we decided that two factors were appropriate for our data. This decision was supported by the Kaiser criterion; only two eigenvalues of the correlation matrix were larger than unity ($\lambda_1 = 5.94$ and $\lambda_2 = 1.46$). This suggests that only the first two factors account for more variability than any single variable. Further, the first factor captures 65% of the variance and the cumulative variance accounted for by the two factors was 80%; adding a third factor only increased this to 85%. Factor loadings (Table 5.4) indicate the correlation of each NEA indicator with the two factors. Factor 1 is highly associated with the variability in ln(TST) (0.84), FCI (0.95), ln(Indirect/Direct) (0.96), IFI (0.97), ln(AGG) (1.00) and AMP (0.91). HMG, ln(TSS), and Boundary have a role in the Factor 1, but much smaller than previous variables. Factor 2 is dominated by Boundary (0.99), with some contribution from ln(TST) (0.53) and a smaller piece from ln(TSS) (0.28). In contrast to principal components analysis (PCA), factor analysis only examines the common variation of the variables (Tabachnick and Fidell, 1983). Indicator uniqueness indicates the proportion of variation not in common with the other variables. HMG (0.85) was the most unique variable, followed by ln(TSS) (0.55), and then AMP (0.17). Other variables had a uniqueness factor below 0.10.

Plotting the factor loadings provides an alternative way of viewing the data (Figure 5.6). An indicator's distance from the origin indicates its strength of association with a factor represented by the axis. As ln(AGG) is the furthest distance from the origin of the y-axis, it is the most highly associated with the first factor (x-axis). This plot emphasizes the close relationship between the variability of ln(AGG), IFI, ln(Indirect/Direct), FCI, and AMP, and their distance from variation in total boundary flux (Boundary). Three factors do not fall close to an axis – HMG, ln(TSS), and ln(TST) – which could make interpretation of the factors difficult. The uniqueness vector, however, indicates that HMG and ln(TSS) variabilities are largely unique; ln(TST) is a special case we address further in the discussion.

These results generally support our second hypotheses, though they suggest it might be somewhat simplistic. Six of the nine NEA indicators – TST, FCI, Indirect/Direct, IFI, AGG, and AMP – are highly correlated and are strong elements of the underlying factor associated with most of the indicator variance. The second factor is dominated by boundary flows, while TSS and HMG have a high degree of uniqueness in these data.

5.4 DISCUSSION

We will discuss the primary results for each hypothesis in turn.

5.4.1 INDICATOR ROBUSTNESS

The NEA indicators of ecosystem organization were qualitatively consistent and quantitatively differentially robust, lending support to our first hypothesis. Qualitatively the index interpretations did not change. Indirect flows consistently dominated direct, indirect flows and cycled flows were consistently a large proportion of TST, and some degree of network aggradation, network homogenization and network amplification occurred. Quantitatively the network indicators were differentially robust, decreasing in variability as AMP > Indirect/Direct > AGG > FCI > HMG = IFI. We cannot determine how these indicators would change if model topology was altered, but given the Lake Lanier model structure and flow and storage uncertainty the magnitudes of network homogenization and indirect flow index were relatively consistent. This suggests it may be possible to circumvent part of the modeling problem, at least in this case, to draw more robust conclusions about the real ecosystem's condition.

The question remains, however, what do these indicators reveal about the Lake Lanier ecosystem? Here we will put the individual indicators into context. In the next section we will discuss their associations and their collective import for understanding ecosystem organization.

We can draw several important conclusions about the Lake Lanier ecosystem organization from this study if we assume the structure of our model is an adequate representation. First, phosphorus flows and storages are heavily influenced by internal ecological processes. This is evident in the consistently high indirect flow index, indirect–to–direct ratio, and degree of recycling. This suggests a relatively well developed ecosystem as the system is receiving high use of the phosphorus boundary inputs. Second, we conclude that although the phosphorus storages in each compartment are relatively heterogeneous, the phosphorus resources available in flows are relatively well mixed based on the fairly robust measure of network homogenization.

5.4.2 INDICATOR INTERRELATIONS

Our second hypothesis anticipated that the NEA indicators were different measures of the same underlying factor, which we hypothesized to be system integration – a form of functional connectivity. The results revealed many strong associations between indicators and that their common variance could be largely partitioned into two latent factors. The common variability of at least six of the nine measures (TST, FCI, Indirect/Direct, IFI, AGG, and AMP) could chiefly be attributed to one latent factor. This suggests that they indicate different aspects of a common unmeasured system variable. We hypothesize that this hidden variable is the degree of network integration by indirect effects as suggested by Patten (in prep.). Total boundary flows (Boundary) and a portion of TST could be attributed to a second factor. Together these latent factors accounted for about 80% of the total variance. Variation of HMG and TSS were fairly unique.

We are not the first to suggest that these indicators overlap. For example, Higashi (Higashi and Patten, 1986; Patten et al., 1990; Patten, 1991) showed algebraically that an increase in the number of nodes, connectance, storage, strength of direct flows, or recycling tends to increase the indirect-to-direct ratio. Therefore the positive association of Indirect/Direct with TSS, TST and FCI is not surprising. In a study of forty-one aquatic ecosystem models of various sizes and connectance but with a consistent currency (g wet weight m^{-2} y^{-1}), Christensen (1995) compared several proposed indicators of ecosystem maturity. He found a strong correlation between FCI and AGG, but they were not well correlated with TST. In his principle components analysis, FCI and AGG were closely associated with the first component, while TST was more closely associated with the second. This is not inconsistent with our results, although our results show a stronger relationship between TST and FCI and AGG. This difference may be driven by the differences in model currency. Fath et al. (2001) used a five mode pathway decomposition to reveal how simultaneous increases in TST, TSS, Boundary and FCI were reconcilable. In their discussion of AMP and HMG, Patten et al. (1990) illustrated how recycling pathways are the primary reason these conditions occur. We would then expect them to generally increase with recycling. When they introduced the quantitative measure, Fath and Patten (1999a) confirmed that HMG tended to increase with cycling in a twenty node model. Later, Fath (2004) used large ($n > 100$) cyber-ecosystem models to show that both

Indirect/Direct and HMG tended to increase with FCI. Network amplification, however, never occurred in the cyber-ecosystems (Fath, 2004). This may be because the cyber-ecosystem models had a very restricted range of cycling ($0.14 < \mathrm{FCI} < 0.26$). Borrett et al. (in press) also observed a strong association between Indirect/Direct and AGG in sixteen nitrogen network models of the Neuse River Estuary, though the reason for this association is not immediately transparent.

We conclude that many of the relationships between the NEA indicators were known or anticipated. Part of our contribution is to bring them together in one quantitative study. Further, in many cases the quantitative form of these relationships was unknown. We start to fill in this detail. In addition, our data support the hypothesis that many of these indicators are characterizing similar aspects of ecosystem organization.

A closer look at TST, FCI, Indirect/Direct, and IFI and their components may be able to explain at least part of their strong associations. In the context of NEA, indirect flows can be partitioned based on two different pathway types: chains and cycles. In the simple chain $i \rightarrow j \rightarrow k$, the influence of i on k is mediated by j. There are a finite number of chains possible in a network; they can be up to $m = n-1$ arcs long. Energy-matter flows in one direction along the chain before it exits the system. In contrast, cycles (pathways that begin and end at the same compartment, e.g., $i \rightarrow j \rightarrow i$), establish the possibility of recursive energy and material flow. While there are a finite number of simple cycles (cycles without repeated nodes) the total number of cycles is theoretically infinite (Chapter 2 & 3). Flows of a conserved currency over these cyclic pathways, however, are ultimately limited in an open thermodynamic system by boundary losses. It is possible for energy-matter in a cycle to pass through the same node multiple times before it is exported or dissipated from the system. Although not all indirect flows (Indirect) are cyclic, all cyclic flows (Cyclic) are necessarily indirect. Thus, we can partition Indirect into components from acyclic ($\text{Indirect}_{\text{AC}}$) and cyclic ($\text{Indirect}_{\text{C}}$) flows. Boundary flows and direct flows are acyclic by definition. Using this new notation, we can rewrite equation (5.2) as follows:

$$\text{TST} = \text{Boundary}_{\text{AC}} + \text{Direct}_{\text{AC}} + \text{Indirect}_{\text{AC}} + \text{Indirect}_{\text{C}}. \qquad (5.7)$$

This combines the Leontief and Finn throughflow decompositions. Equation (5.7) implies that IFI ≥ FCI. When cycled flow is a significant component of indirect flow, we expect these two indices to be closely associated. This should be true in more retentive networks (more efficient or less dissipative) such as models of biogeochemical currencies.

Caution is wise when interpreting resultant factors in factor analysis (Grimm and Yarnold, 1995). However, our correlation and factor analysis suggest there are two latent variables influencing most of the NEA indicator variables in this study. These factors can be linked to a conceptual model of ecosystem growth introduced by Jørgensen et. al. (2000), which was extended by Fath et. al. (2004) to include four forms. These are:

- Form 0 – Boundary growth. Open thermodynamic systems are sustained by a constant source of low-entropy material. Growth form 0 is the increase in boundary input. Fundamentally, boundary input limits the work that can be accomplished by the system.
- Form I – Structural growth. Increasing biomass caused by component growth in number, size, and types.
- Form II – Network growth. This growth form captures shifts in the internal organization of the system, which may include increasing connectivity and recycling of energy and matter.

- Form III – Informational growth. This is augmentation of the internal order of the informational content of the system. It includes behavioral changes from a more exploitative to a more conservative system as well as possible increases in genetic information of component organisms.

Again, Factor 1 is dominated by TST, FCI, Indirect/Direct, IFI, AGG, and AMP. Cycling and cyclic pathways appear to be the essential element linking these indicators. Energy–matter cycles are fundamental to system growth and development as they establish a route of cybernetic feedback (positive or negative) in the system (Patten and Odum, 1981; DeAngelis et al., 1986), and provide a means of integrating ecosystem flows and storage (Ulanowicz, 1983; Patten et al., 1990). These six indicators and therefore Factor 1 appear to indicate different aspects of Growth Form II, which Jørgensen et. al. (2000) termed "growth-to-throughflow". In addition, we tentatively suggest that this factor is Patten's hypothesized integration through indirect effects. Though the variability of HMG was largely unique, we would expect it to also be an indicator of Growth Form II.

Factor 2 captures the variability of boundary flows (Boundary) and a portion of the variability of TST. This factor represents the effects of environmental variability. TST is a component of this factor as well as the first because boundary flows and internal flows are combined in its calculation (5.2). We conclude that this factor relates to Growth Form 0.

Variation in TSS was largely unique such that it was only partially associated with Factor 1 and Factor 2. TSS is an indicator of Growth Form I.

5.4.3 CAVEATS AND CONCLUSIONS

We must be careful not to over generalize our results; this work is one case study. Our model and methodology allow us to investigate the effect of flow and storage uncertainty on the NEA indictors of ecosystem organization in this particular case. We are unable to assess the indicator variability due to model structural uncertainty; our demonstration that several of the indicators were relatively robust only allows us to evade a portion of the modeling problem. Our analysis of the relationships between indicators is also limited. The differential indicator robustness means that we did not encounter the full range of each of the indicators. The relationships we characterize here may not hold in all cases. The difference between our results and those of Christensen (1995) suggest that this may be the case.

Ultimately we would like to utilize these indicators to classify this ecosystem as healthy, mature, or to characterize its integrity as the system may be changing due to the rapid urbanization of its watershed. These classifications are relative and require meaningful comparisons, which at the moment do not yet exist for our model. The most pressing problem is that our model currency is phosphorus – which is not commonly used. Perhaps in the future enough case studies will exist to draw meaningful conclusions.

In conclusion, the results of our study make two primary contributions to ecosystem network analysis. First, they demonstrate that five of the ecosystem network analysis indicators – FCI, Indirect/Direct, IFI, AGG and HMG – were relatively robust to the flow and storage uncertainty in the Lake Lanier phosphorus model. Thus, if we assume the model structure is an adequate representation, we can conclude that phosphorus flux in the Lake Lanier ecosystem is heavily influenced by internal processes. Second, the results show that the majority of the nine ecosystem indicators investigated co-varied and that their common variation could largely be mapped into two latent factors. We have tentatively interpreted these two factors as 1) system integration (Growth Form II) and 2) boundary (Growth Form 0). This case study provides a window into the workings of network environ analysis and the Lake Lanier ecosystem. Continued research into the sensitivity of ecological network analysis to model uncertainty will lead to a better understanding of when our model simplifications are justified, when our simplifications need to be altered to better match natural phenomena, and when they are not justified.

5.5 ACKNOWLEDGEMENTS

We would like to thank B.C. Patten, D.K. Gattie, H.R. Pulliam, P.G. Verity and two anonymous reviewers for their examination of this manuscript. This work was supported by EPA/NSF Water and Watershed Program (#R825758).

5.6 REFERENCES

Abarca-Arenas, L.G., and Ulanowicz, R.E., 2002. The effects of taxonomic aggregation on network analysis. Ecol. Model., **149**:285-296.

Allesina, S., and Bondavalli, C., 2004. WAND: an ecological network analysis user-friendly tool. Environmental Modelling & Software, **19**:337-340.

Allesina, S., and Ulanowicz, R.E., 2004. Cycling in ecological networks: Finn's index revisited. Computational Biology and Chemistry, **28**:227-233.

Augustinovics, M., 1970. Methods of international and intertemporal comparison of structure. In: A.P. Carter and A. Brody (Editors), Contributions to input-output analysis, Amsterdam, North-Holland

Borrett, S.R., Whipple, S.J., Patten, B.C., and Christian, R.R., in press. Indirect effects and distributed control in ecosystems 3. Temporal variability of indirect effects in a seven-compartment model of nitrogen flow in the Neuse River Estuary (USA): time series analysis. Ecol. Model., **XXX**:XXX-XXX.

Bosserman, R.W., 1983. Flow-analysis sensitivities for models of energy or material flow. Bull. Math. Biol., **45**:807-826.

Cale, W.G., and Odell, P.L., 1979. Concerning aggregation in ecosystem models. In: E. Halfon (Editor), Theoretical Systems Ecology, Academic Press, New York, pp 55-77.

Cale, W.G., and O'Neill, R.V., 1988. Aggregation and consistency problems in theoretical models of exploitative resource competition. Ecol. Model., **40**:97-107.

Capra, F., 1996. The web of life: a new scientific understanding of living systems, 1st Anchor Books edition. Anchor Books, New York.

Christensen, N.L.B., Ann M.; Brown, James H.; Carpenter, Stephen; D'Antonio, Carla; Francis, Rober; Franklin, Jerry F.; MacMahon, James A.; Noss, Reed F.; Parsons, David J.; Peterson, Charles H.; Turner, Monica G.; Woodmansee, Robert G., 1996. The report of the Ecological Society of America Committee on the Scientific Basis for Ecosystem Management. Ecological Applications, **6**:665-691.

Christensen, V., 1995. Ecosystem maturity - towards quantification. Ecol. Model., **77**:3-32.

Christensen, V., and Pauly, D., 1992. Ecopath-II - a software for balancing steady-state ecosystem models and calculating network characteristics. Ecol. Model., **61**:169-185.

Christensen, V., and Walters, C.J., 2004. Ecopath with Ecosim: methods, capabilities and limitations. Ecol. Model., **172**:109-139.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R.V., Paruelo, J., Raskin, R.G., Sutton, P., and van den Belt, M., 1997. The value of the world's ecosystem services and natural capital. Nature, **387**:253-260.

Daily, G.C., Editor. 1997. Nature's Services. Island Press, Washington, D.C.

Dalgaard, P., 2002. Introductory statistics with R. Springer, New York.

DeAngelis, D.L., Post, W.M., and Travis, C.C., 1986. Positive feedback in natural systems. Springer-Verlag, Berlin; New York.

Fath, B.D., 2004. Network analysis applied to large-scale cyber-ecosystems. Ecol. Model., **171**:329-337.

Fath, B.D., and Beck, M.B., 2005. Elucidating public perceptions of environ mental behavior: a case study of Lake Lanier. Environmental Modelling & Software, **20**:485-498.

Fath, B.D., and Borrett, S.R., 2005. A Matlab function for Network Environ Analysis. Environmental Modelling & Software, **XX**:XXX-XXX.

Fath, B.D., Jorgensen, S.E., Patten, B.C., and Straškraba, M., 2004. Ecosystem growth and development. Biosystems, **77**:213-228.

Fath, B.D., and Patten, B.C., 1999a. Quantifying resource homogenization using network flow analysis. Ecol. Model., **107**:193-205.

Fath, B.D., and Patten, B.C., 1999b. Review of the foundations of network environ analysis. Ecosystems, **2**:167-179.

Fath, B.D., Patten, B.C., and Choi, J.S., 2001. Complementarity of ecological goal functions. J. Theor. Biol., **208**:493-506.

Finn, J.T., 1976. Measures of ecosystem structure and function derived from analysis of flows. J. Theor. Biol., **56**:363-380.

Finn, J.T., 1980. Flow analysis of models of the Hubbard Brook ecosystem. Ecology, **61**:562-571.

Gardner, R.H., Cale, W.G., and O'Neill, R.V., 1982. Robust analysis of aggregation error. Ecology, **63**:1771-1779.

Gattie, D.K., Borrett, S.R., and other?, in prep. Throughflow decompositions in network environ analysis. Ecological Complexity.

Ginzburg, L.R., and Jensen, C.X.J., 2003. Rules of thumb for judging ecological theories. Trends Ecol. Evol., **19**:121-126.

Grimm, L.G., and Yarnold, P.R., 1995. Reading and understanding multivariate statistics, 1st edition. American Psychological Association, Washington, D.C.

Guan, W., 1993. Integrating Water Quality Modeling with Geographic Information System -- Application to Lake Sidney Lanier. Doctoral Dissertation. University of Georgia, Athens, GA.

Gunderson, L.H., and Holling, C.S., 2002. Panarchy: understanding transformations in human and natural systems. Island Press, Washington, DC.

Han, B.P., 1997. On several measures concerning flow variables in ecosystems. Ecol. Model., **104**:289-302.

Hatcher, K.J., Callaham, M.A., Nearing, M.A., Pancorbo, O., Patten, B.C., Rogers, L.F., Sellers, J., and Van Den Avyle, M.J., 1994. Diagnostic/Feasibility Study of Lake Sidney Lanier, Georgia. Project Completion Report Prepared for Georgia Environmental Protection Division, Athens, Georgia.

Higashi, M., and Burns, T.P., 1991. Theoretical studies of ecosystems: the network perspective. Cambridge University Press, Cambridge.

Higashi, M., Burns, T.P., and Patten, B.C., 1989. Food network unfolding - an extension of trophic dynamics for application to natural ecosystems. J. Theor. Biol., **140**:243-261.

Higashi, M., and Patten, B.C., 1986. Further aspects of the analysis of indirect effects in ecosystems. Ecol. Model., **31**:69-77.

Higashi, M., and Patten, B.C., 1989. Dominance of indirect causality in ecosystems. Am. Nat., **133**:288-302.

Johnson, D.E., 1998. Applied multivariate methods for data analysts. Duxbury Press, Pacific Grove, CA.

Jørgensen, S.E., 2002. Integration of ecosystem theories: A pattern, 3rd edition. Kluwer Academic Publishers, Dordrecht; Boston.

Jørgensen, S.E., and Mejer, H.F., 1979. A holistic approach to ecological modelling. Ecol. Model., **7**:169-189.

Jørgensen, S.E., Patten, B.C., and Straškraba, M., 2000. Ecosystems emerging: 4. Growth. Ecol. Model., **126**:249-284.

Leontief, W.W., 1965. The structure of the American economy. Sci. Am., **212**:25-35.

Leontief, W.W., 1966. Input-output economics. Oxford University Press, New York.

Levi, P., 1988. The drowned and the saved. Summit Books, New York.

Levin, S.A., 1998. Ecosystems and the biosphere as complex adaptive systems. Ecosystems, **1**:431-436.

Levins, R., 1966. The strategy of model building in population biology. Am. Scientist, **54**:421-431.

Lindeman, R.L., 1942. The trophic-dynamic aspect of ecology. Ecology, **23**:399-418.

Lotka, A.J., 1922. Contribution to the energetics of evolution. Proc. Natl. Acad. Sci. USA, **8**:147-151.

Manly, B.F.J., 1991. Randomization and Monte Carlo methods in biology, 1st edition. Chapman and Hall, London ; New York.

Margalef, R., 1963. Certain unifying principles in ecology. Am. Nat., **97**:357-374.

Matis, J.H., and Patten, B.C., 1981. Environ analysis of linear compartmental systems: the static, time invariant case. Bull. Int. Stat. Inst., **48**:527-565.

Metro Atlanta Chamber of Commerce, 2000. Final report of the Clean Water Initiative. Technical Report, Metro Atlanta Chapter of Commerce.

Müller, F., 1996. Emergent properties of ecosystems - consequences of self-organizing processes? Senckenbergiana maritima, **27**:151-168.

Müller, F., and Leupelt, M., 1998. Eco targets, goal functions, and orientors. Springer, Berlin; New York.

Odum, E.P., 1969. The strategy of ecosystem development. Science, **164**:262-270.

Odum, H.T., 1962. Quantitative classification and limits of ecosystems. Am. Zool., **2**:434-434.

Odum, H.T., 1983. Systems ecology: an introduction. Wiley, New York.

Oreskes, N., 1998. Evaluation (not validation) of quantitative models. Environmental Health Perspectives, **106**:1453-1460.

Oreskes, N., Shraderfrechette, K., and Belitz, K., 1994. Verification, Validation, and Confirmation of Numerical-Models in the Earth-Sciences. Science, **263**:641-646.

Osidele, O.O., and Beck, M.B., 2001. Identification of model structure for aquatic ecosystems using regionalized sensitivity analysis. Water Sci. Technol., **43**:271-278.

Osidele, O.O., and Beck, M.B., 2004. Food web modelling for investigating ecosystem behaviour in large reservoirs of the south-eastern United States: lessons from Lake Lanier, Georgia. Ecol. Model., **173**:129-158.

Pahl-Wostl, C., 1992. The possible effects of aggregation on the quantitative interpretation of flow patterns in ecological networks. Math. Biosci., **112**:177-183.

Patten, B.C., 1978. Systems approach to the concept of environment. Ohio J. of Sci., **78**:206-222.

Patten, B.C., 1981. Environs: the superniches of ecosystems. Am. Zool., **21**:845-852.

Patten, B.C., 1982. Environs: relativistic elementary particles for ecology. Am. Nat., **119**:179-219.

Patten, B.C., 1991. Network ecology: indirect determination of the life-environment relationship in ecosystems. In: M. Higashi and T. Burns (Editors), Theoretical Studies of Ecosystems: the network perspective, Cambridge University Press, New York, New York, pp 288-351.

Patten, B.C., 1998. Network orientors: Steps toward a cosmography of ecosystems: Orientors for directional development, self-organization, and autoevolution. In: F. Müler and M. Leupelt (Editors), Eco Targets, Goal Functions, and Orientors, Springer, Berlin, pp 137-160.

Patten, B.C., in prep. Holoecology: the unification of nature by network indirect effects. Columbia University Press, New York.

Patten, B.C., Bosserman, R.W., Finn, J.T., and Cale, W.G., 1976. Propagation of cause in ecosystems. In: B.C. Patten (Editor), Systems Analysis and Simulation in Ecology, Vol. IV, Academic Press, New York, pp 457-579.

Patten, B.C., Fath, B.D., Choi, J.S., Bastianoni, S., Borrett, S.R., Brandt-Williams, S., Debeljak, M., Fonseca, J., Grant, W.E., Karnawati, D., Marques, J.C., Moser, A., Müller, F., Pahl-Wostl, C., Seppelt, R., Seinborn, W.H., and Svirezhev, Y.M., 2002. Complex adaptive hierarchical systems. In: R. Costanza and S.E. Jørgensen (Editors), Understanding and solving environmental problems in the 21st century: Toward a new, integrated hard problem science, Elsevier Science, Ltd, Oxford, pp 41-87.

Patten, B.C., Higashi, M., and Burns, T.P., 1990. Trophic dynamics in ecosystem networks: significance of cycles and storage. Ecol. Model., **51**:1-28.

Patten, B.C., and Odum, E.P., 1981. The cybernetic nature of ecosystems. Am. Nat., **118**:886-895.

R Development Core Team, 2004. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reichman, O.J., and Pulliam, H.R., 1996. The scientific basis for ecosystem management. Ecological Applications, **6**:694-696.

Schneider, E.D., and Kay, J.J., 1994. Life as a manifestation of the second law of thermodynamics. Mathematical and Computer Modelling, **19**:25-48.

Sugihara, G., Bersier, L.F., and Schoenly, K., 1997. Effects of taxonomic and trophic aggregation on food web properties. Oecologia, **112**:272-284.

Tabachnick, B.G., and Fidell, L.S., 1983. Using multivariate statistics. Harper & Row, New York.

Teal, J.M., 1962. Energy-flow in salt-marsh ecosystem of Georgia. Ecology, **43**:614-624.

Thompson, R.M., and Townsend, C.R., 2000. Is resolution the solution?: the effect of taxonomic resolution on the calculated properties of three stream food webs. Freshw. Biol., **44**:413-422.

Ulanowicz, R.E., 1983. Identifying the structure of cycling in ecosystems. Math. Biosci., **65**:219-237.

Ulanowicz, R.E., 1986. Growth and development: ecosystems phenomenology. Springer-Verlag, New York.

Ulanowicz, R.E., 1997. Ecology, the ascendent perspective. Columbia University Press, New York.

US Army Corps of Engineers. 2005. Lake Sidney Lanier. *in*. US Army Corps of Engineers, Mobile District.

Weaver, O.R., 2000. Lake Lanier Annual Report, 1999. Technical Report, Georgia Department of Natural Resources, Fisheries Section of the Wildlife Resources Division, Gainsville, Georgia.

Yodzis, P., and Winemiller, K.O., 1999. In search of operational trophospecies in a tropical aquatic food web. Oikos, **87**:327-340.

5.7 TABLES

Table 5.1: Network Environ Analysis indicators of whole system organization

Indicator	**Label**	**Description**	**Formula**
Total System Throughflow	TST	Sum of total flow into or out of nodes	$\mathrm{TST} = \sum_{k=1}^{n} T_k = \sum_{k=1}^{n}\sum_{j=1}^{n}(f_{kj} + z_k) = \sum_{i=1}^{n}\sum_{k=1}^{n}(f_{ik} + y_k)$
Total System Storage	TSS	Total amount of model currency stored in nodes	$\mathrm{TSS} = \sum_{k=1}^{n} x_k$
Total Boundary Flow	Boundary	Total amount of boundary input or output	$\mathrm{Boundary} = \sum_{i=1}^{n} z_i = \sum_{j=1}^{n} y_j$
Finn Cycling Index	FCI	Proportion of TST from recycling	$\mathrm{FCI} = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{n_{ij}}{n_{ii}}(n_{ii} - 1)z\right)_j \Big/ \mathrm{TST}$
Indirect/Direct	Indirect/Direct	Ratio of indirect and direct flows	$\mathrm{Indirect}/\mathrm{Direct} = \frac{\sum(N - I - G)z}{\sum Gz}$
Indirect Flow Index	IFI	Proportion of TST derived from indirect flows	$\mathrm{IFI} = \frac{\sum(N - I - G)z}{\mathrm{TST}}$
Homogenization	HMG	tendency to uniformly distribute causality across the network	$\mathrm{HMG} = \frac{\mathrm{CV(G)}}{\mathrm{CV(N)}} = \frac{\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(\bar{g} - g_{ij})}{n-1}}}{\bar{g}} \Big/ \frac{\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(\bar{n} - n_{ij})}{n-1}}}{\bar{n}}$
Amplification	AMP	Proportion of flows obtaining more than face value	$\mathrm{AMP} = \frac{\# n_{ij} > 1 \ (i \neq j)}{n(n-1)}$
Aggradation (a.k.a. average path length, flow multiplying ability, multiplier effect)	AGG	Average number of times an average input passes through the system	$\mathrm{AGG} = \mathrm{TST}/\mathrm{Boundary}$

Table 5.2: Mean, standard deviation (SD), and coefficient of variation (CV) of Network Environ Analysis ecosystem indicators to flow and storage uncertainty in a phosphorus model of Lake Lanier. They are ordered by decreasing relative variability shown by the coefficient of variation.

Indicator	Mean	SD	CV
TSS	1634	985	0.60
AMP	0.21	0.12	0.59
TST	749	303	0.40
Indirect/Direct	7.45	2.83	0.38
AGG	7.84	2.69	0.34
FCI	0.39	0.11	0.28
Boundary	96	20	0.21
HMG	3.10	0.31	0.10
IFI	0.75	0.07	0.10

Table 5.3: Pairwise linear ordinary least squares regression coefficients (Y=b + mX)

	Y	X	b		m		Pr(>F)	
1	ln(TST)	ln(TSS)	2.94	*	0.50	*	0.00	*
2	ln(TST)	Boundary	5.78	*	0.01	*	0.00	*
3	ln(TST)	FCI	5.53	*	2.63	*	0.00	*
4	ln(TST)	ln(Indirect/Direct)	5.01	*	0.80	*	0.00	*
5	ln(TST)	IFI	3.62	*	3.94	*	0.00	*
6	ln(TST)	ln(AGG)	4.75	*	0.90	*	0.00	*
7	ln(TST)	HMG	5.88	*	0.23		0.07	
8	ln(TST)	AMP	6.09	*	2.22	*	0.00	*
9	ln(TSS)	Boundary	6.75	*	0.01	*	0.03	*
10	ln(TSS)	FCI	6.41	*	2.20	*	0.00	*
11	ln(TSS)	ln(Indirect/Direct)	5.87	*	0.72	*	0.00	*
12	ln(TSS)	IFI	4.44	*	3.78	*	0.00	*
13	ln(TSS)	ln(AGG)	5.48	*	0.89	*	0.00	*
14	ln(TSS)	HMG	5.64	*	0.54	*	0.00	*
15	ln(TSS)	AMP	6.92	*	1.66	*	0.00	*
16	Boundary	FCI	105.76	*	-22.67		0.26	
17	Boundary	ln(Indirect/Direct)	108.03	*	-5.75		0.33	
18	Boundary	IFI	132.44	*	-47.43		0.11	
19	Boundary	ln(AGG)	116.34	*	-9.66		0.14	
20	Boundary	HMG	140.14	*	-14.09	*	0.04	*
21	Boundary	AMP	99.95	*	-14.55		0.40	
22	FCI	ln(Indirect/Direct)	-0.16	*	0.28	*	0.00	*
23	FCI	IFI	-0.68	*	1.43	*	0.00	*
24	FCI	ln(AGG)	-0.24	*	0.31	*	0.00	*
25	FCI	HMG	0.17		0.07	*	0.05	
26	FCI	AMP	0.22	*	0.81	*	0.00	*
27	ln(Indirect/Direct)	IFI	-1.74	*	4.92	*	0.00	*
28	ln(Indirect/Direct)	ln(AGG)	-0.21	*	1.07	*	0.00	*
29	ln(Indirect/Direct)	HMG	0.64		0.43	*	0.00	*
30	ln(Indirect/Direct)	AMP	1.36	*	2.73	*	0.00	*
31	IFI	ln(AGG)	0.31	*	0.22	*	0.00	*
32	IFI	HMG	0.47	*	0.09	*	0.00	*
33	IFI	AMP	0.64	*	0.52	*	0.00	*
34	ln(AGG)	HMG	0.91	*	0.36	*	0.00	*
35	ln(AGG)	AMP	1.51	*	2.37	*	0.00	*
36	HMG	AMP	2.98	*	0.45		0.08	

* indicates significant at α = 0.05

Table 5.4: Principle components factor analysis

Variable	Factor 1	Factor 2	Uniqueness
ln(TST)	0.84	0.53	0.01
ln(TSS)	0.61	0.28	0.55
Boundary	-0.09	0.99	0.01
FCI	0.95	-0.04	0.09
ln(Indirect/Direct)	0.96	-0.02	0.07
IFI	0.98	-0.09	0.04
ln(AGG)	1.00	-0.07	0.01
HMG	0.34	-0.18	0.85
AMP	0.91	-0.01	0.17
Loading Sums of Squares	5.81	1.40	
Proportion of Variance	0.65	0.16	
Cumulative Variance	0.65	0.80	

5.8 FIGURE LEGENDS

Figure 5.1: Eleven compartment model of phosphorus flux in Lake Sidney Lanier, Georgia, USA.

Figure 5.2: Average (± SD) inter-compartment flows (F) oriented from column to row, inputs (z), outputs (y) in mg P m^{-2} d^{-1}, and storage (x, mg P m^{-2}) for 90 plausible models of phosphorus flow and storage in Lake Lanier. Compartment labels are in the vector Names.

Figure 5.3: Example of eleven successful model outputs for A) epilimnion phosphorus, B) phytoplankton, and C) larval-juvenile fish. Constraints for the behavior definitions are represented by the dashed lines and square markers.

Figure 5.4: Sample distributions of nine Network Environ Analysis indicators in plausible parameterizations of the Lake Lanier phosphorus model from simulations with 500, 1000, and 5000 trials. The trials generated 14, 18, and 90 plausible model parameterizations respectfully. The indicators are A) Total System Throughflow (TST), B) Total System Storage (TSS), C) Total Boundary Input (Boundary), D) Finn cycling index (FCI), E) ratio of indirect-to-direct flows (Indirect/Direct), F) indirect flow index (FCI), G) network aggradation (AGG), H) network homogenization (HMG), and I) network amplification (AMP).

Figure 5.5: Pairwise scatter plots of nine Network Environ Analysis indicators of ecosystem organization in 90 plausible parameterization of the Lake Lanier phosphorus model. Scatter plots and ordinary least squares regression lines are plotted above the principle diagonal and corresponding Pearson correlation coefficients are positioned below.

Notice that TST, TSS, Indirect/Direct, and AGG have been transformed by the natural logarithm.

Figure 5.5: Factor analysis loadings plot.

Figure 5.1

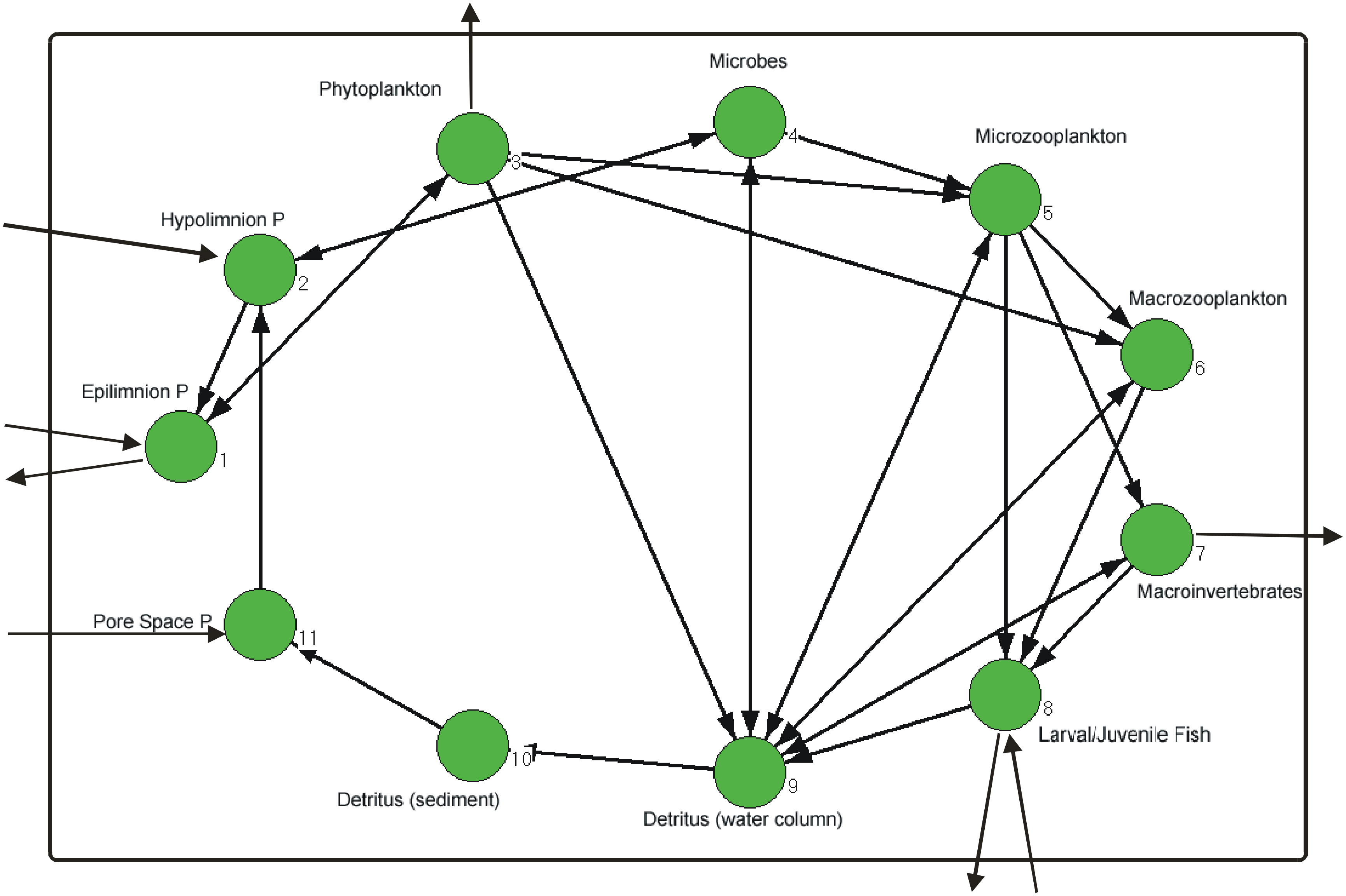

Phytoplankton
Microbes
Microzooplankton
Hypolimnion P
Macrozooplankton
Epilimnion P
Macroinvertebrates
Pore Space P
Larval/Juvenile Fish
Detritus (sediment)
Detritus (water column)
1
2
3
4
5
6
7
8
9
10
11

Figure 5.2

$$F = \left|\begin{array}{rrrrrrrrrrr}
0 & 33.6\ (10.1) & 4.3\ (2.7) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 28.3\ (29.6) & 0 & 0 & 0 & 0 & 0 & 0 & 39.9\ (25.8) \\
25.3\ (5.3) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 55.8\ (47.2) & 0 & 0 & 0 & 0 & 0 & 0 & 33.9\ (28.4) & 0 & 0 \\
0 & 0 & 5.2\ (3.1) & 29.6\ (25.2) & 0 & 0 & 0 & 0 & 38.2\ (32.4) & 0 & 0 \\
0 & 0 & 5.1\ (3.3) & 0 & 16.8\ (15.1) & 0 & 0 & 0 & 35.7\ (24.0) & 0 & 0 \\
0 & 0 & 0 & 0 & 18.8\ (16.7) & 0 & 0 & 0 & 38.3\ (25.4) & 0 & 0 \\
0 & 0 & 0 & 0 & 18.1\ (13.9) & 28.4\ (17.2) & 18.8(16.0) & 0 & 0 & 0 & 0 \\
0 & 0 & 5.4\ (3.2) & 31.7\ (29.1) & 19.3\ (18.1) & 29.1\ (25.6) & 19.2\ (16.9) & 39.8\ (25.7) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 26.2\ (23.3) & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 26.2\ (23.3) & 0
\end{array}\right|$$

$$\text{Names} = \left|\begin{array}{c}
\text{Epilimnion P} \\
\text{Hypolimnion P} \\
\text{Phytoplankton} \\
\text{Microbes} \\
\text{Microzooplankton} \\
\text{Macrozooplankton} \\
\text{Macroinverts} \\
\text{Larval/Juvenile fish} \\
\text{Suspended detritus} \\
\text{Sediment detritus} \\
\text{Pore Space P}
\end{array}\right| \quad z = \left|\begin{array}{r}
6.9\ (5.3) \\
21.1\ (9.9) \\
0 \\
0 \\
0 \\
0 \\
0 \\
27.5\ (12.6) \\
27.6\ (11.2) \\
0 \\
13.8\ (8.6)
\end{array}\right| \quad y^T = \left|\begin{array}{r}
19.4\ (9.1) \\
0.0 \\
5.3\ (2.9) \\
0.0 \\
0.0 \\
0.0 \\
19.1\ (16.0) \\
53.0\ (21.5) \\
0.0 \\
0.0 \\
0.0
\end{array}\right| \quad x = \left|\begin{array}{r}
64.4\ (3.5) \\
228.5\ (208.7) \\
22.3\ (7.0) \\
129.9\ (106.7) \\
77.6\ (50.1) \\
154.0\ (223.0) \\
81.4\ (55.8) \\
171.1\ (42.5) \\
152.6\ (83.5) \\
240.0\ (398.4) \\
341.3\ (588.7)
\end{array}\right|$$

Figure 5.3

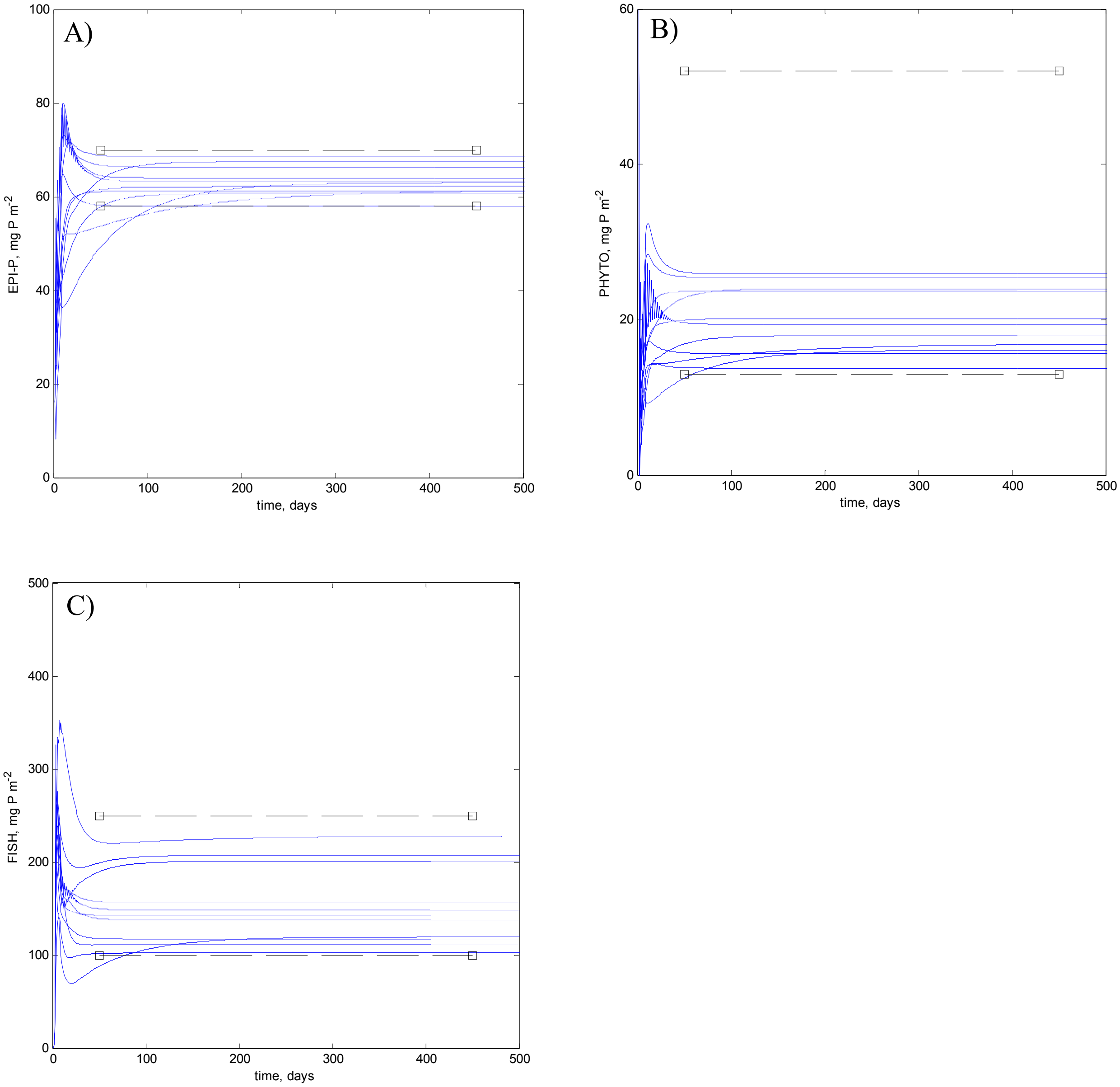

A)
100
80
60
40
20
0
EPI-P, mg P m^{-2}
0
100
200
300
400
500
time, days
B)
60
40
20
0
PHYTO, mg P m^{-2}
0
100
200
300
400
500
time, days
C)
500
400
300
200
100
0
FISH, mg P m^{-2}
0
100
200
300
400
500
time, days

Figure 5.4

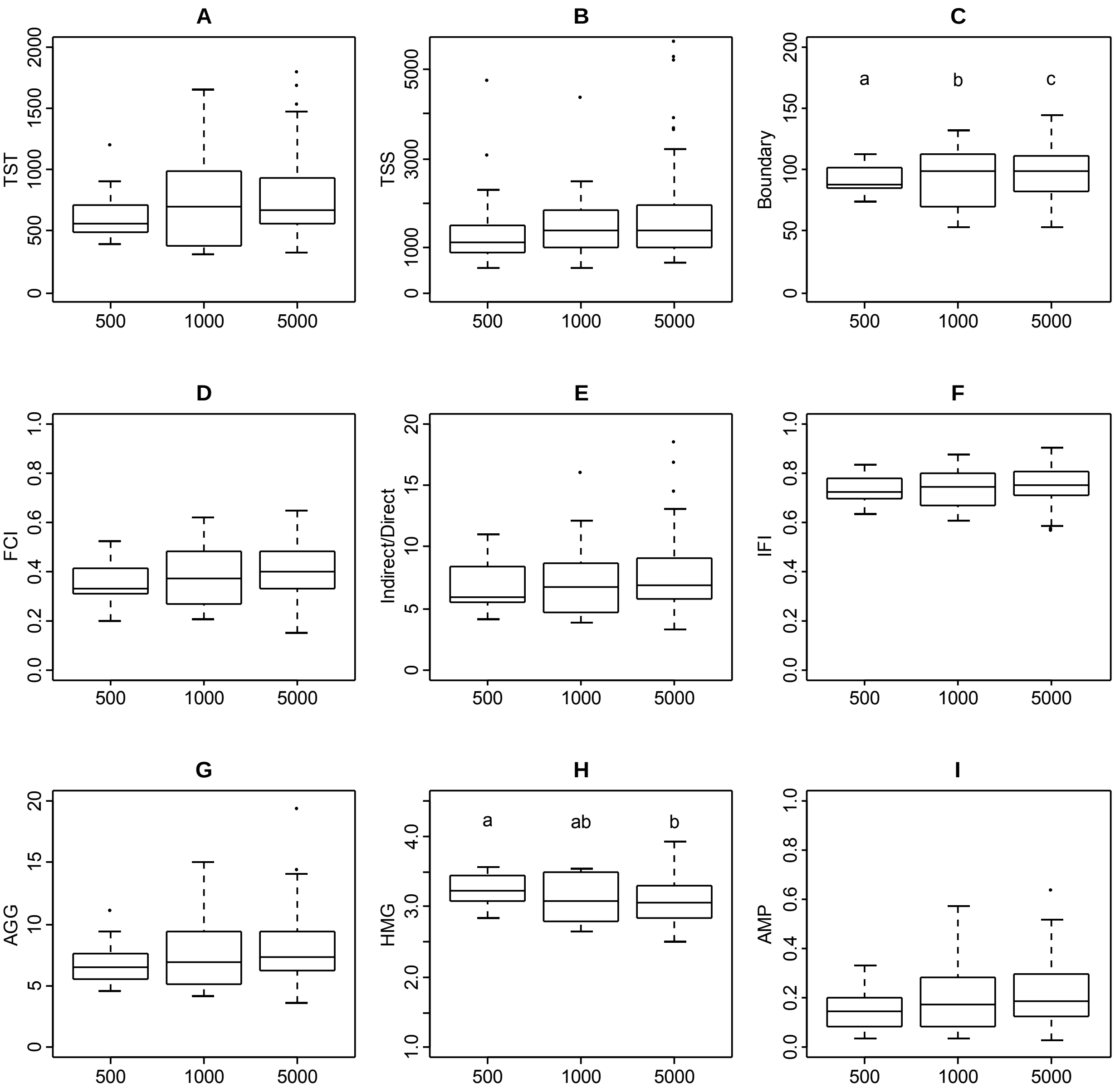
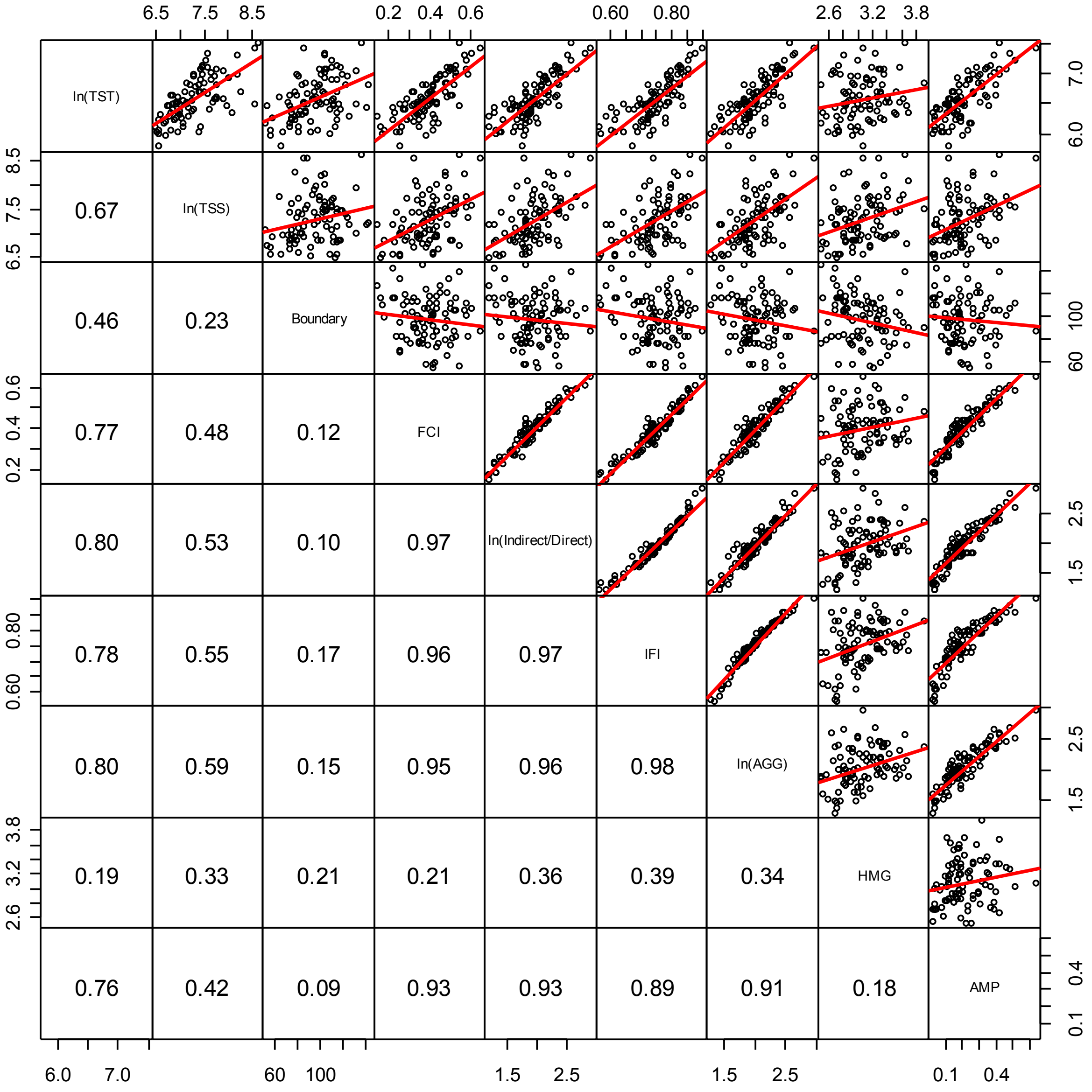
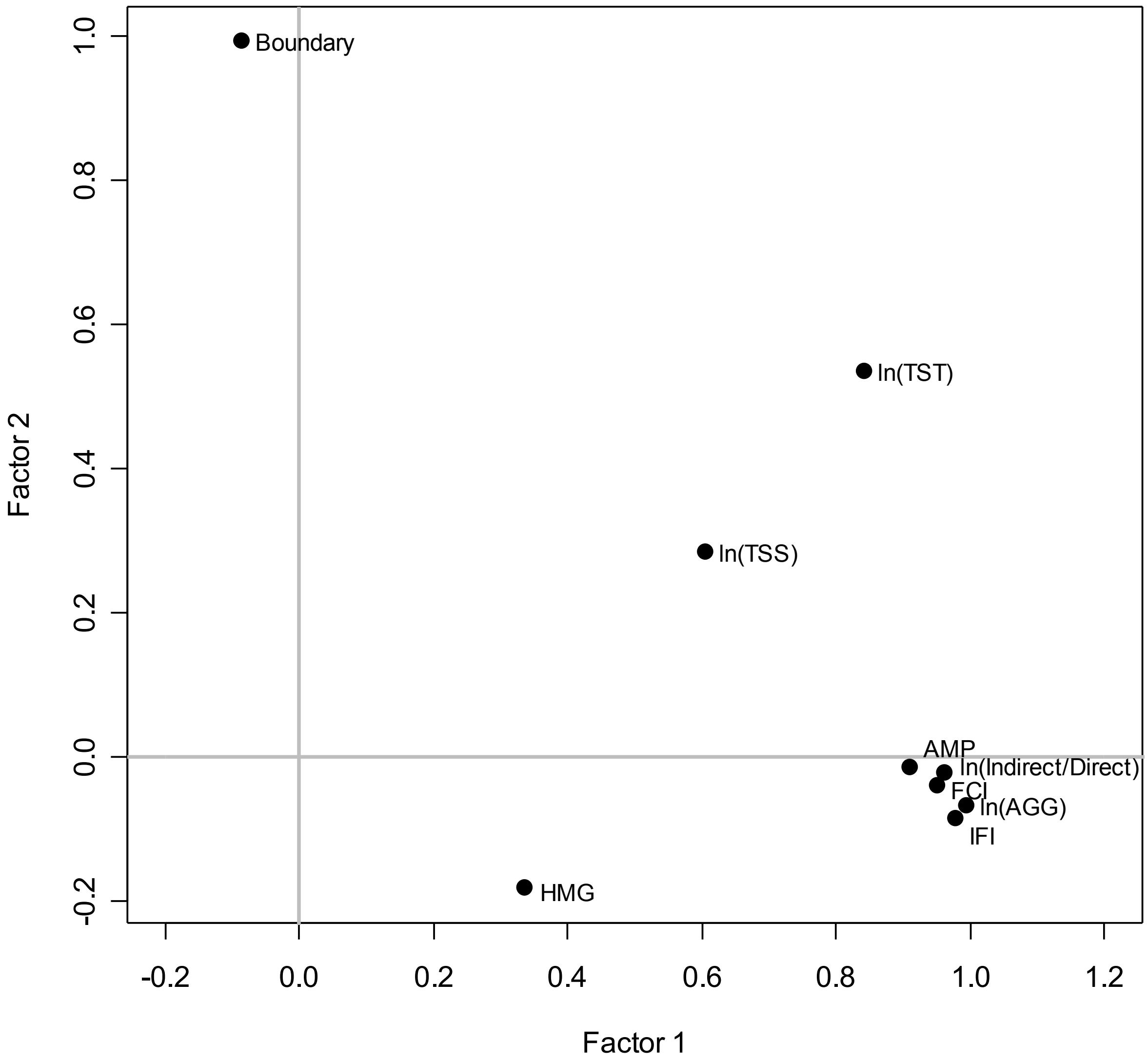

A
TST
0
500
1000
1500
2000
500
1000
5000
B
TSS
0
1000
3000
5000
500
1000
5000
C
Boundary
0
50
100
150
200
a
b
c
500
1000
5000
D
FCI
0.0
0.2
0.4
0.6
0.8
1.0
500
1000
5000
E
Indirect/Direct
0
5
10
15
20
500
1000
5000
F
IFI
0.0
0.2
0.4
0.6
0.8
1.0
500
1000
5000
G
AGG
0
5
10
15
20
500
1000
5000
H
HMG
1.0
2.0
3.0
4.0
a
ab
b
500
1000
5000
I
AMP
0.0
0.2
0.4
0.6
0.8
1.0
500
1000
5000

Figure 5.5:



ln(TST)
ln(TSS)
Boundary
FCI
ln(Indirect/Direct)
IFI
ln(AGG)
HMG
AMP
0.67
0.46
0.23
0.77
0.48
0.12
0.80
0.53
0.10
0.97
0.78
0.55
0.17
0.96
0.97
0.80
0.59
0.15
0.95
0.96
0.98
0.19
0.33
0.21
0.21
0.36
0.39
0.34
0.76
0.42
0.09
0.93
0.93
0.89
0.91
0.18
6.5
7.5
8.5
0.2
0.4
0.6
0.60
0.80
2.6
3.2
3.8
6.0
7.0
60
100
1.5
2.5
0.1
0.4

Figure 5.6


Boundary
ln(TST)
ln(TSS)
AMP
ln(Indirect/Direct)
FCI
ln(AGG)
IFI
HMG
Factor 1
Factor 2
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
-0.2
0.0
0.2
0.4
0.6
0.8
1.0

CHAPTER 6

DECOUPLING ECOSYSTEM AGGRADATION AND INDIRECT FLOW: COMPARATIVE NETWORK ENVIRON ANALYSIS OF TWENTY ENERGY MODELS[1]

---

[1] Borrett, S.R. and B.C. Patten, to be submitted to *Biosystems*

ABSTRACT

Ecosystems are open thermodynamic systems composed of an intricate network of energy, matter, and information flows connecting species and their environments. They receive flows of energy and matter at their boundaries, perform work, and produce heat and degraded by products which are dissipated to the environment. In this study we evaluate the relationship between two measures of ecosystem organization: the ratio of total system throughflow and total boundary inputs (termed aggradation, AGG) and the ratio of indirect-to-direct flows (Indirect/Direct). A strong association between Indirect/Direct and AGG is compelling because AGG has a strong thermodynamic interpretation, while Indirect/Direct is wholly derived from ecosystem network analysis. Linking these two measures offers the possibility of bringing together these two approaches. However, our comparative Network Environ Analysis (NEA) of 20 energy flux ecosystem models drawn from the literature partially decouples these measures. We show they necessarily start at different values, but that they appear to converge as compartment transfer efficiency increases. In addition, we evaluate alternate methods of calculating Indirect/Direct and argue that it is more appropriate to make this calculation using input vectors that reflect the observed distribution of inputs because indirect flows are generated by the internal system organization and the boundary flows that activate them. Finally, our work identifies six ecosystem models in which Indirect/Direct is less than unity, but more generally provides additional evidence supporting the hypothesis that indirect flows tend to dominate direct flows in model ecosystems.

6.1 INTRODUCTION

Living systems appear to violate the second law of thermodynamics (Schrödinger, 1946). The second law states that all natural processes are irreversible and entropy generating or degrading, driving the natural world toward thermodynamic equilibrium (Zemansky and Dittman, 1997). Instead, living systems construct order and organization: aggrading rather than degrading. The crux of the problem is to recognize that living systems are open, far-from-equilibrium systems (Schrödinger, 1946; Prigogine and Stengers, 1984). They acquire relatively low-entropy (high-exergy) energy–matter from their surroundings and use it to perform work, degrading the energy and rearranging the matter. Local organization is gained by exporting relatively high-entropy (low-exergy) residues.

Ecosystems are open, dissipative, far-from-equilibrium thermodynamic systems composed of biotic and abiotic elements transferring and transforming energy, matter, and information. They receive high-exergy flows of energy and matter at their boundaries, perform organizational work of aggradation (departure from equilibrium), and produce heat and degraded byproducts which are dissipated to the environment. Jørgensen et al. (2000) define ecosystem aggradation as a growth process. Like other kinds of complex adaptive hierarchical systems, ecosystems appear to self-organize in response to thermodynamic gradients (Schneider and Kay, 1994; Capra, 1996; Müller, 1996; Levin, 1998; Müller, 1998; Patten, 1998b; Patten et al., 2002).

A plurality of theoretical tactics has led to the development of multiple measures of ecosystem organization (e.g., gross-production-to-respiration ratio (Odum, 1969), cycling (Finn, 1976; Ulanowicz, 1983), emergy (Odum, 1983), indirect-to-direct flow ratio (Patten, 1983; Higashi and Patten, 1986, 1989), and acendency (Ulanowicz, 1986, 1997)). Given that the fundamental subject matter is the same, however, it would be highly surprising if information embedded in some of the measures did not partially overlap. As the methodologies mature and are compared, relationships between indices are emerging. For example, Christensen (1995) analyzed forty-one ecosystem models of biomass flow and storage created with Ecopath software (Christensen and Pauly, 1992) to determine the maturity of their organization. Using a statistical approach, he found strong associations between many of the indicators, and was able to select a reduced set of orthogonal indices to rank order the ecosystem maturities, largely based on the ideas of Odum (1969). Building on the work of Patten (1995), Fath et al. (2001) demonstrated that several measures of ecosystem growth and development (proposed as system orientors or goal functions) were complementary by translating them into a common network analysis framework.

In Chapters 4 and 5 we used a statistical approach to uncover hidden relationships between ecosystem indicators derived from Network Environ Analysis (NEA). One relationship was the strong association between *network aggradation*, which is the ratio of total system throughflow to total boundary inputs, AGG = TST/Boundary, and *environ indirect effects*, which are characterized by the ratio of indirect-to-direct flows (Indirect/Direct). AGG was initially introduced by Finn (1976) as *average path length*, but was rediscovered and proposed as a thermodynamically grounded measure of system growth by Patten (Jørgensen et al., 2000). Indirect/Direct is a measure of system organization developed in NEA (Patten, 1982b, 1984a, b, 1985b; Higashi and Patten, 1986, 1989; Patten, in prep.). In Chapter 4 we found a strong linear association (AGG = 0.99 *Indirect/Direct + 0.75, $R^2=0.99$, $p<0.001$) between AGG and Indirect/Direct in the sixteen seasonal models of nitrogen flux in the Neuse River Estuary. The relation in this case suggested the two indices started at different points, but increased at the same rate. Their relationship weakened slightly in the ninety plausible parameterization of the

Lake Lanier phosphorus ecosystem model in Chapter 5 (AGG = 0.90 * Indirect/Direct + 0.75, $R^2 = 0.92$, $p<0.001$). This difference may be due to the lower cycling in the Lanier models, but could also be influenced by differences in the model currency (nitrogen and phosphorus, respectively) or underlying model structure (number of nodes, number of links, pattern of link connection).

In the work presented here we use comparative Network Environ Analysis of 20 energy based ecosystem models to further characterize this relationship. We evaluate two primary hypotheses. The first is the strong association between system aggradation (AGG) and environ indirect effects (Indirect/Direct). Here, we provide a more robust test of the relationship by investigating energy models with different structure. We expect energy models to have lower transfer efficiencies and therefore have less cycling than the nutrient based models. We anticipate this will decrease the magnitude of both AGG and Indirect/Direct although not necessarily in the same way. The number of model compartments or nodes (n) is also suspected to influence AGG (Finn, 1976), and Higashi (Patten et al., 1990; Patten, 1991) showed algebraically that Indirect/Direct tends to increase as system size (n), connectivity (C) or proportion of possible direct connections completed, system feedback, strength of direct interactions, storage, and cycling increase.

Second, this study provides a statistical test of the hypothesis that indirect flows tend to dominate direct flows in ecosystems, one of the cardinal hypotheses of Holoecology (Patten, in prep., see also Chapter 2). Previous results from NEA suggest that Indirect/Direct is typically greater than unity in model ecosystems, implying that indirect flows are dominant (Patten, 1983, 1984a; Higashi and Patten, 1986, 1989; Patten, 1991, in prep.). Indirect flows are only one type of indirect effects (Miller and Kerfoot, 1987; Wootton, 1994, 2002), though they may phenomenologically reflect several (but not all) types (see Chapter 4). Indirect effects act as an integrating factor binding a system together as a functional and perhaps co-evolutionary unit. As systems become more complex, organized and grow further from thermodynamic equilibrium we expect Indirect/Direct to increase. Conversely, we anticipate reduced transfer efficiencies in energy models to constrain indirect flow magnitude and therefore Indirect/Direct, but we are unsure if this ratio will drop below unity at which point indirect flows do not dominate direct.

In the process of considering these issues we used structural NEA developed in Chapter 3 to determine the presence of strongly connected components in the model structures and characterized their rates of pathway proliferation. Presence of strongly connected components and large rates of pathway proliferation are hypothesized to increase the likelihood that indirect flows dominate direct flows. We also evaluated two methods of calculating Indirect/Direct.

6.2 MATERIALS AND METHODS

For this work we adopted an empirical approach. We used comparative Network Environ Analysis (NEA, Fath and Patten, 1999; Patten, in prep.) to characterize the organization of 20 ecosystem models of energy or carbon flux collected from the literature (Table 6.1). Using a modified version of NEA.m (Fath and Borrett, 2005, see Appendix A), we conducted an output oriented structural and throughflow analysis. In this section we will first introduce the models selected for this study and then describe the elements of network environ analysis used.

6.2.1 ECOSYSTEM MODELS

While all 20 steady-state models selected for this study trace the flux of energy or carbon through an ecosystem, they span a range of aquatic ecosystem types and vary in their structural and functional representations of the systems (Tables 6.1 & 6.2). The ecosystem types include an oyster reef community (Dame and Patten, 1981), tidal marsh (Ulanowicz, 1986), estuaries (Baird and Milne, 1981; Rybarczyk, 2003), seagrass beds (Baird et al., 1998), open bays (Baird and Ulanowicz, 1989; Monaco and Ulanowicz, 1997; Ulanowicz et al., 1998), freshwater and tidal marshes (Ulanowicz et al., 1997; Ulanowicz et al., 2000), and mangroves (Ulanowicz et al., 1999). The models also range in size (n) and connectivity (C, Table 6.2). The oyster reef community model (Dame and Patten, 1981) is the smallest model with six compartments but it has the highest connectivity at 33%. As this model has been well studied, it provides a benchmark for comparisons (Dame and Patten, 1981; Patten, 1985a; Higashi and Burns, 1991). In contrast, the wet and dry season models of Florida Bay are the largest models with 125 compartments (Ulanowicz et al., 1998). Their connectivity is near the lowest with 13% and 12%, respectively. The lowest connectivity (5%) occurs in the fifty-one compartment model for the seagrass beds near Sprague Island off the coast of the St. Marks National Wildlife Refuge in Florida. Notice that the models do not have the same flow units; the units include kcal m^{-2} d^{-1}, mg C m^{-2} $summer^{-1}$, mg C m^{-2} y^{-1}, and mg C m^{-2} d^{-1}. Nor were the models constructed for similar spatial extents. As we are primarily interested in the relationship between non-dimensional indicators of ecosystem organization these differences will be less important. However, the specific ecosystem observations are likely to be significantly influenced by the disparity in temporal and spatial scale, as well as the modeling decisions of the original authors. This prevents us from drawing robust comparisons of the actual ecosystems. Further, Baird et al. (1991) argued that it is inappropriate to compare network analysis results from models with different structures (but see Christensen, 1995).

A further complication of this data set is that the models are not strictly independent. Several of the models ostensibly represent the same system at different times, in different places, or under alternate conditions. For example, the Crystal River Marsh ecosystem is modeled twice. One represents a section of the marsh impacted by thermal pollution; the other represents an unimpacted section (Ulanowicz, 1986). There are six seagrass ecosystem models for different points in space and time off the coast of the St. Marks National Wildlife Refuge (Baird et al., 1998). The seagrass system near Live Oak Island and the one near Wakulla Beach are modeled in both January and February. Two additional seagrass sites, Sprague Point and Lighthouse Point, were modeled in January and February, respectively. The Florida Bay (Ulanowicz et al., 1998), gramminoid marsh (Ulanowicz et al., 2000), and mangrove models (Ulanowicz et al., 1999) developed as part of the Across Trophic Levels System Simulation (ATLSS; http://atlss.org/) project were modeled in both a wet and dry season. The structure, flow, and storage relationships of the models varied despite the physical relationship among several of them. Therefore, we will treat the 20 models as independent for our analysis, although we acknowledge that strictly speaking we have only eleven independent samples.

6.2.2 NETWORK ENVIRON ANALYSIS

Network Environ Analysis (NEA, Patten, 1978; Matis and Patten, 1981; Patten, 1982a; Fath and Patten, 1999; Patten, in prep.) is an environmental application and extension of economic Input-Output Analysis (Leontief, 1966) used to investigate ecological systems holistically. It is the primary methodology of the Holoecology Research Program (HRP), which is striving to develop a deep understanding of the processes that create, constrain, and sustain ecological systems, as well as construct a formal theory of environment. NEA operates like a macroscope to characterize whole-system organization by describing, quantifying, and analyzing component-level environments or *environs* that comprise a system (Patten, 1978, 1981, 1982a, 1992). Fath and Patten (1999) recently reviewed the foundations of NEA. The methodology includes analyses of model structure, throughflow, utility, and control. Our focus in this paper is on aggradation (AGG) and the ratio of indirect-to-direct flows (Indirect/Direct) which can be derived from the throughflow analysis. Therefore, here we limit our description of NEA methods to the structure and throughflow analyses used. Additional details of the methodology are described in the literature (Patten et al., 1976; Matis and Patten, 1981; Patten, 1985a, in prep.). NEA also includes calculations for both input and output orientations, but we will again limit our analysis to the output case. We expect the input case to be qualitatively similar.

6.2.2.1 STRUCTURE

Employing the network perspective (Higashi and Burns, 1991), ecosystem structure has three primary elements: the number of compartments or nodes, n; the number of flows or directed links L connecting node j to node i, and the pattern of the connections. The system structure can be represented as an unweighted directed graph (digraph) or through its isomorphic associated adjacency matrix $A_{n\times n} = (a_{ij})$, where $a_{ij} = 1$ if there is a link from j to i or $a_{ij} = 0$ if there is not. Two measures – link density (L/n) and connectance ($C = L/n^2$) – are often used to characterize aspects of the network structure.

For this study, we focus on two large-scale patterns possible in network structure: strongly connected components and pathway proliferation. As described in Chapter 3, networks can be decomposed into a unique set of *strongly connected components* (K). An induced subdigraph includes a subset of nodes and all links that both start and terminate on the node subset. The maximally induced subdigraph is the largest induced subdigraph that remains strongly connected, and a strongly connected component is a maximally induced subdigraph in which it is possible to reach each node from every other node over a pathway of some length. For each model, we report the number of non-trivial (more than one node) strongly connected components (#K) and the proportion of model nodes included in a strongly connected component (%K). We also report the largest rate of pathway proliferation – the rate at which the number of pathways increases as pathway length increases – of the strongly connected components. Strongly connected components and pathway proliferation have been hypothesized to influence the development of indirect effects (Borrett and Patten, 2003, see also Chapters 2 and 3).

6.2.2.2 THROUGHFLOW

Throughflow analysis in NEA begins with information about three types of energy–matter flows: observed internal system transfers $F_{n\times n} = (f_{ij})$, boundary inputs z_i, and boundary losses y_j (i,j = 1,…,n). With this information from a steady-state model of conservative energy–matter transfers, we can determine TST/Boundary and Indirect/Direct.

The total currency flowing into our out of a node, termed throughflow, is defined as $T_k^{(in)} \equiv \sum_{j(\neq k)=1}^{n} f_{kj} + z_k$ and $T_k^{(out)} \equiv \sum_{i(\neq k)=1}^{n} f_{ik} + y_k$ respectively. At steady state, an assumption of NEA, $T_k^{(in)} = T_k^{(out)} = T_k \ (k = 1,\ldots,n)$. Total system throughflow ($TST = \sum T_k$) is a system-level measure of activity (Finn, 1976). Finn (1976) partitioned TST into two components: 1) TST from straight-chain flow (Chain), and 2) TST from recycled flow (Cycled). He used this partition to construct a cycling index, FCI = Cycled/TST that could be used as another measure of system organization (Finn, 1980).

Again, network aggradation (AGG) is TST divided by the total boundary input $\sum_{i=1}^{n} z_i$ or output $\sum_{j=1}^{n} y_j$, which at steady state are identical such that $\text{Boundary} = \sum_{i=1}^{n} z_i = \sum_{j=1}^{n} y_j$. Aggradation is then calculated as

$$\text{AGG} = \frac{\text{TST}}{\text{Boundary}} = \frac{\sum\sum f_{ij} + \text{Boundary}}{\text{Boundary}} = \frac{\sum\sum f_{ij}}{\text{Boundary}} + 1. \quad (6.1)$$

Normalizing TST by boundary flows creates a measure of system organization. AGG relates to the internal organization generated by energy input from the system environment. In his early work with flow analysis, Finn (1976) proposed AGG as a measure of *average path length*. He found that it indicates the average path length traveled in the system by an average boundary inflow, and noted that like TST this measure was sensitive to n. Han (1997) re-examined Finn's average path length, suggesting that it should be called the *flow multiplying ability* of the system because it indicates the reutilization of energy–matter inputs into the system. Jørgensen et al. (2000) used AGG to distinguish the organization of two hypothetical steady-state systems with four compartments each and identical total boundary flows. One system was a straight chain; the other had a number of additional connections which allowed recycling and changed the system into one strongly connected component. As expected, the system with recycling had a larger AGG value.

The ratio of indirect-to-direct flows (Indirect/Direct) is derived from an alternative decomposition of throughflows into boundary, direct, and indirect flows based on economic Input-Output Analysis (Leontief, 1966). This analysis starts by finding a matrix N that maps inputs into throughflows such that $T = Nz$. T and z are known so we can simply solve for N. However, an alternative method of calculating N is informative. The first step is to divide the observed flows by the donor compartment throughflow ($G = (g_{ij}) = \left(\frac{f_{ij}}{T_j}\right)$). The elements of $G_{n\times n}$ are interpreted as the fractional transfer coefficients, or the probability that energy–matter entering j will flow to i over a direct pathway (i.e., pathway of length m = 1). The elements g_{ij} are non-dimensional and range between 0 and 1. Fractional transfer coefficients for flow from j to i over indirect pathways (m > 1) are found by raising G to the power m, G^m. Total node throughflow is recovered by post multiplying the convergent infinite series by the boundary inputs

$$T = \left(\underbrace{I}_{\text{Boundary}} + \underbrace{G^1}_{\text{Direct}} + \underbrace{G^2 + \ldots G^m + \ldots}_{\text{Indirect}} \right) * z, \tag{6.2}$$

where $I = G^0$ is the identity matrix. The power series for thermodynamically conserved currencies converges to the transitive closure matrix, $N = (n_{ij}) = (I - G)^{-1}$, because ecosystems are open and dissipative. Therefore, the (n_{ij}) elements are interpreted as the integral (boundary + direct + indirect) fractional transfer coefficients of energy–matter from j to i over all pathways of all lengths generated by a boundary input at j.

By distributing z and summing the resultant vectors, we create a whole system level partition of TST as follows

$$\mathrm{TST} = \underbrace{\sum Iz}_{\text{Boundary}} + \underbrace{\sum Gz}_{\text{Direct}} + \underbrace{\sum (N - I - G)z}_{\text{Indirect}}. \tag{6.3}$$

Dividing both sides of (6.2) by TST generates:

$$1 = \frac{\sum z}{\mathrm{TST}} + \frac{\sum Gz}{\mathrm{TST}} + \frac{\sum (N - I - G)z}{\mathrm{TST}}, \tag{6.4}$$

where $\frac{\sum z}{\mathrm{TST}}$ is the boundary flow index (BFI), $\frac{\sum Gz}{\mathrm{TST}}$ is the direct flow index (DFI), and $\frac{\sum (N - I - G)z}{\mathrm{TST}}$ is the Indirect Flow Index (IFI). These are dimensionless numbers between 0 and 1 that indicate the proportion of boundary, direct, and indirect flow in the system. These indices are similar in concept to Finn's cycling index. In Chapter 5 we observed that all cyclic flow was necessarily indirect such that IFI ≥ FCI. Here, we introduce the ratio of FCI to IFI (FCI/IFI) to evaluate the significance of cycled flow in indirect flow.

As observed in Chapter 4, there are at least two ways of calculating Indirect/Direct. Fath and Patten (1999) calculate it by summing the appropriate fractional transfer coefficients,

$$\text{Indirect/Direct}\,(1) = \frac{\sum\sum(n_{ij} - \delta_{ij} - g_{ij})}{\sum\sum(g_{ij})}, \tag{6.5}$$

where $\delta_{ij} = 1$ if i = j, otherwise it is zero. This approach tacitly assumes there is a unit input into every node. Alternatively, Indirect/Direct can be calculated as

$$\text{Indirect/Direct}\,(2) = \frac{\sum(\mathrm{N} - \mathrm{I} - \mathrm{G})\mathrm{z}}{\sum \mathrm{Gz}}. \tag{6.6}$$

This calculation incorporates the magnitude and distribution of boundary inputs from the original model. We use this comparative NEA to further evaluate the differences between these two calculations.

6.3 RESULTS

The models included in this study demonstrate a range of structure and throughflow organization.

6.3.1 STRUCTURE

While the 20 ecosystem energy flow models vary in size (6 ≤ n ≤ 125), connectance (0.05 ≤ C ≤ 0.33), and link density (2 ≤ L/n ≤ 15.75), all but one has one large non-trivial strongly connected component (#K; Table 6.2). The Chesapeake Bay model (n = 36) has two strongly connected components: one with n = 6 and another with n =16. The percentage of model components participating in the non-trivial strongly connected components ranges from a minimum of 41% in the January model of the Sprague Point seagrass to a maximum of 100% in the Ythan estuary model (Figure 6.1A). The Ythan estuary model was the only energy model analyzed to be fully contained in one strongly connected component. On average, 75% (± 0.18 SD) of the nodes were involved in a non-trivial strongly connected component. The maximum rate of pathway proliferation ($\lambda_1(A)$) ranged from 2.15 in the oyster reef model to 14.17 in the dry season of the mangrove model (Table 6.2). These rates appear to be similar to link density (L/n; Figure 6.1B), which is the expected value for random digraphs (see Chapter 3). However, when we compared the absolute difference between $\lambda_1(A)$ and L/n ($d = | \lambda_1(A)-L/n|$) in each model to an ensemble of 1001 random digraphs with the same number of nodes and probability of connectance, we found that 11 of the 20 models were significantly different at $\alpha = 0.05$ (Pr(d); Table 6.2). This indicates that these 11 models have a structure unlikely to be created by random processes.

Collectively, these structural measures indicate that while the networks as a whole were fairly sparsely connected in terms of connectance, many of the model nodes were connected by at least one cycle. By definition a strongly connected component has at least one cycle that links all nodes in the component. The maximal rates of pathway proliferation in the strongly connected components suggest the presence of multiple simple cycles (cycles without repeated nodes except the starting and ending point). Given these structural results, we conclude that each of these models has the potential to have relatively large indirect flows.

6.3.2 THROUGHFLOW

Results of the throughflow analysis are summarized in Table 6.2. As the model flows had different units, the magnitudes of TST and Boundary were highly variable, ranging from a minimum of 84 and 41 kcal m^{-2} d^{-1}, respectively, in the oyster reef model to a maximum TST of

3,917,246 mg C^{-2} y^{-1} in the Ythan estuary and a maximum boundary flow of 888,791 mg C^{-2} y^{-1} in the Narragansett Bay model. These values are an important starting point for our analysis, but comparing them is uninformative. The remaining throughflow measures are dimensionless ratios that are more comparable. The proportion of TST derived from recycling (FCI) ranges from 0.01 in the Sprague Point seagrass model to 0.51 in the Narragansett Bay model (Table 6.2). It is perhaps surprising that the Ythan estuary model which was relatively well connected and completely encapsulated in a strongly connected component did not have a greater FCI. It has the second largest FCI, but it is less than half of the Narragansett Bay model. This reinforces the idea that system structure is a necessary, but not sufficient, element for ecosystem analysis. Flow and storage magnitudes and distributions in the network can be critical.

Indirect/Direct was quite variable in these models, and the two methods of calculating it (6.5 and 6.6) generated different magnitudes (Figure 6.2). More importantly the qualitative interpretation changed in two cases. Indirect/Direct (1) (using equation 6.5) ranged from 0.35 to 8.13. In fourteen models this ratio was greater than unity, implying that indirect flows were a dominant fraction of the internal model flows. However, in six models (Somme Estuary, Crystal River Marsh – control and thermally impacted, Sprague Point and Lighthouse Point seagrass beds, and the wet season graminoid marsh) Indirect/Direct (1) was less than unity. Indirect/Direct (2) (using equation 6.6) also suggests that indirect flows do not dominate in six of the models and has a minimum of 0.12 and maximum of 6.93 (Figure 6.3A). With this measure, however, indirect flows dominate direct in the wet season graminoid marsh model, but not in the wet season cypress wetland model. A scatter plot of Indirect/Direct (1) and (2) shows that in most cases the different measures generate similar results; the points are very close to the expected 1:1 line (Figure 6.2A). These results support the hypothesis that indirect flows tend to dominate direct flows in model ecosystems, as this was the case in 70% of the models.

Network aggradation ranged from 1.62 in the Sprague Point seagrass model to 5.65 in the Narragansett Bay model. This implies that the average input into the Sprague Point model did not make it to a second compartment, while in the Narragansett Bay model the average input traveled to five nodes before exiting the system. Mean AGG for these 20 models was 2.59 (± 0.93 SD). It is interesting to note that the boundary flow and therefore the AGG values we report for the six models from the St. Marks National Wildlife Refuge do not match those reported by Baird et al. (1998). Our boundary flow values are consistently higher, which causes our AGG numbers to be lower. We are not sure what accounts for this difference.

We chose to evaluate the relationship between Indirect/Direct and AGG using the second method of calculation (6.6), as this is consistent with Chapters 4 and 5 (Figure 6.2B). AGG and Indirect/Direct (2) again appear to be linearly associated (AGG = 0.60 * Indirect/Direct (2) + 1.65, $R^2 = 0.80$, $p<0.0001$) and both parameters are statistically different from zero ($p < 0.001$). However, this association is weaker than reported in the Neuse River Estuary and Lanier models. The Narragansett Bay model has an Indirect/Direct (2) that is much larger than the other models and it is the only point to fall below the 1:1 line. We conclude that while there remains some association between Indirect/Direct and AGG, it has changed from that seen in the Lake Lanier and Neuse River Estuary models and is not well represented by our simple linear regression.

The partition of TST into boundary, direct, and indirect flows (6.4; Figure 6.3B) reveals that boundary flows were 50% or more of TST in five of these models, but never exceeded 62%. In the Narragansett Bay model boundary flows were only 18% of TST. The Indirect Flow Index (IFI) ranged from 0.04 in the Sprague Point seagrass model to 0.72 in the Narragansett Bay model. It averaged 0.33 (± 0.15 SD). The FCI/IFI ratio reveals that the proportion of indirect flows derived from cycling is fairly variable, ranging from 0.71 in the Narragansett Bay model to 0.06 and 0.12 in the two wet and dry seasons of the graminoid marsh model (Figure 6.4).

Our final analysis evaluated the relationship of Indirect/Direct (2) and the number of model nodes (n), connectance (C), maximal rate of pathway proliferation ($\lambda_1(A)$), and the Finn cycling index (FCI; Figure 6.4). Higashi's algebra suggests that Indirect/Direct will tend to increase with n and C (Patten et al., 1990; Patten, 1991). In Chapters 2 and 3 we hypothesized that Indirect/Direct would also tend to increase as the pathway proliferation increased. However, we found no significant relationship between Indirect/Direct (2) and n, C, or $\lambda_1(A)$ in these models. Indirect/Direct (2) does tend to increase with FCI ($R^2 = 0.86$), although the Narragansett Bay model could be driving the relationship. Without the outlying point, the association falls to ($R^2 = 0.49$).

6.4 DISCUSSION

Results from this study raise three issues we address in this section. First, we presented two methods of calculating the Indirect/Direct ratio and found they did not necessarily generate the same quantitative results and, while there was generally good qualitative agreement, in two cases the qualitative results changed. Is more appropriate than another? Second, our Indirect/Direct ratios – regardless of how calculated – show that indirect flows do tend to dominate direct flows in 70% of the models we investigated. We take this as weak support for the second cardinal hypothesis of Holoecology – that indirect effects are dominant in ecosystems (Patten, in prep., see also Chapter 2). These models require closer examinations to determine why. As these models were not constructed with a fixed modeling criterion set for the purpose of comparative analysis, these results may be driven by modeling decisions, rather than actual differences in ecosystem organization. Finally, the relationship between Indirect/Direct and network aggradation needs further attention. While the statistical relationship between the two variables seems to weaken as cycling decreases, our results are not clear.

6.4.1 CALCULATING INDIRECT/DIRECT

In this paper we contrast two methods of calculating Indirect/Direct (6.5 and 6.6). The first is reported in Fath and Patten (1999) in their review of the foundations of Network Environ Analysis (NEA). The second method was introduced in Chapter 4. Results of this study show that while in some cases the two methods generate similar values, in other cases the results are quite different.

The first method tacitly assumes a vector of unit inputs into each node. Therefore, the magnitude of total boundary inputs in this calculation is n, which is evenly distributed across all nodes. When comparing Indirect/Direct from systems of different sizes, we might be concerned about the differences in assumed boundary input magnitudes. The ratio construction – where this assumption is applied to the numerator and denominator – should effectively negate this issue. This is why it is possible to compare Indirect/Direct between systems with different boundary inputs. The more serious issue is the assumed even distribution of inputs, as it equally weights the network flows when this may not be the case in the observed input flow vector. For example, the Narragansett Bay model only has inputs into three of its thirty-two compartments (Benthic Algae, Phytoplankton, and Detritus). Even among these three compartments the total boundary input is unevenly distributed – 21%, 79%, and 0.6%. This uneven distribution of inputs, propagated throughout the network, is responsible for the observed differences between the two methods of calculating Indirect/Direct.

One approach to use a “unit” vector of inputs and correct both the input magnitude and match their possibly heterogeneous distribution is to divide the elements of the input vector by the total boundary inputs ($\bar{z}_i = \frac{z_i}{\sum_{i=i}^{n} z_i}$). In all cases the resultant vector magnitude is unity, and the observed input flow distribution is retained. When we did this for the models in our study and recalculated Indirect/Direct as in (6.5), substituting the implicit vector of ones with $\bar{z}$, we recovered the exact values of Indirect/Direct (2). We conclude that Indirect/Direct (2) is a more appropriate measure of the ratio of indirect-to-direct flows because it more accurately captures the empirically observed input distribution. This input distribution is significant because it differentially weights flow pathways in the model, affecting the ultimate flow partitions. Direct and indirect flows are determined by both the internal system transfer efficiencies and the external environment that activates the system, driving it away from thermodynamic equilibrium.

6.4.2 DOMINANCE OF DIRECT FLOW

In 14 of the 20 models we investigated, indirect flows dominated direct. This was apparent both in the ratio of indirect-to-direct flows (Table 6.2) and complete partition of throughflows into the boundary, direct, and indirect flow indices (Figure 6.1). Patten (Patten, 1983, 1985a; Higashi and Patten, 1989; Patten, 1998a, in prep.) argues this occurs because of the large number of indirect pathways which, while they may individually only transmit a small amount of indirect flow, are cumulatively quite powerful. However, in six models direct flows were larger than indirect flows. For this discussion, we will focus on the six models identified by Indirect/Direct (2) (e.g., Somme Estuary, Crystal River Marsh (control and thermally impacted), Sprague Point and Lighthouse Point seagrass beds, the wet season cypress wetlands, and the wet season graminoid marshes). We wonder why this occurs. In all six systems the large numbers of indirect pathways were present; they all had a large non-trivial strongly connected component that encompassed a sizable fraction of the model compartments (>40%). In addition, these strongly connected components experienced pathway proliferation, as the proliferation rates $\lambda_1(A)$ were all larger than unity ($2.15 \leq \lambda_1(A) \leq 14.17$). As discussed in Chapter 3, this indicates that the model structure provided the opportunity for indirect flows to be significant components of TST. This is true despite several of these models having relatively low connectivity. However, the potential in the structure was not realized. It appears that in these models the cyclic indirect pathways were not a significant component of TST. For example, by comparison to the other five seagrass models explicitly constructed by Barid et al. (1998) for network analysis comparison, the Sprague Point seagrass model appears to have been operating not as a well integrated reticulate network, but as a leaky pipe. The system seems to have shut off. The Indirect/Direct (2) ratio is the lowest of the models examined at 0.12, and the IFI is only 0.01. In contrast, boundary flow accounts for 62% of TST. The vast majority of the energy flowing into the system is immediately lost from the system; the internal transfer coefficients must be low. This point is echoed by the network aggradation of 1.62, indicating that the average input does not even make it to a second node before exiting the system. However, there is not an obvious consistent pattern that demarcates the six systems other than Indirect/Direct (2). The network aggradation measures of the six ecosystems are low, but there are systems with equally low values of AGG that still have Indirect/Direct > 1 (e.g., the oyster reef model). The same can be said of the FCI. Understanding why these six systems have so little indirect flow will require closer investigation of the model details than is available in this study.

6.4.3 INDIRECT/DIRECT VS. AGGRADATION

Finally, we turn our attention to the relationship between Indirect/Direct and AGG. The relationship between Indirect/Direct and AGG was not clarified by the statistical approach used in this paper, although the strength of their association appears to decline as the degree of cycling falls. Here, we consider the theoretical minimum and maximum values of Indirect/Direct and AGG to clarify their relationship.

TST achieves a minimum when the nodes of a network are 100% dissipative (internal transfer coefficients are zero), which is equivalent to a structurally disconnected system. In this case, all boundary inputs are immediately lost to the external environment in a steady state system. Using the definition, TST can be partitioned into boundary and internal flows as

$$\mathrm{TST} = \underbrace{\sum_{i=1}^{n} z_i}_{\text{Boundary}} + \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}}_{\text{Internal}} . \tag{6.7}$$

If the internal flows are zero ($\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij} = 0$) then TST = Boundary. In this case, network aggradation reduces to unity because the internal flows are zero and we are left with Boundary/Boundary. When there is at least one connection between compartments, what Patten et al. (1976) identify as a causal bond essential to system formation, TST must then exceed Boundary and network aggradation exceeds unity.

Indirect/Direct is a ratio of a within-system flow decomposition based on pathway types. To show this more clearly, we can set the right hand side of the throughflow partition in 6.7 equal to the right hand side of 6.3, such that

$$\underbrace{\sum_{i=1}^{n} z_i}_{\text{Boundary}} + \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}}_{\text{Internal}} = \underbrace{\sum \mathrm{I}z}_{\text{Boundary}} + \underbrace{\sum \mathrm{G}z}_{\text{Direct}} + \underbrace{\sum (\mathrm{N}-\mathrm{I}-\mathrm{G})z}_{\text{Indirect}} . \tag{6.8}$$

As I is the identity matrix, $\sum_{i=1}^{n} z_i = \underbrace{\sum Iz}_{\text{Boundary}}$. These terms then cancel and we are left with

$$\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}}_{\text{Internal}} = \underbrace{\sum Gz}_{\text{Direct}} + \underbrace{\sum (N - I - G)z}_{\text{Indirect}}. \tag{6.9}$$

Reconsidering our hypothetical system in which $f_{ij} = 0$ ($\forall i, j$), it is now clear that Direct = Indirect = 0, and Indirect/Direct is mathematically indeterminate but physically does not exist. Adding only one flow from node j to i will cause $f_{ij} > 0$ and Direct > 0, but Indirect remains 0. Thus, in this hypothetical system, Indirect/Direct = 0 while AGG > 1. These exercises demonstrate that while Indirect/Direct might be driven by similar elements of ecosystem organization, the measures start at different points. This difference is reflected to some degree in the intercept values of the statistical models.

Alternatively, we consider the maximum values of Indirect/Direct and AGG. This occurs in systems in which indirect flows are a large proportion of TST, and the boundary and direct flow indices are necessarily small (6.4). In this case, the indirect flows will closely approximate TST such that Indirect ≈ TST. A large Indirect Flow Index can only occur in systems where the internal flow transfer efficiencies are very large. When this happens, the vast majority of energy–matter entering the system will be passed into the system via the direct flow pathways, and therefore Direct ≈ Boundary. These relations account for why TST/Boundary ≈ Indirect/Direct in the Neuse River Estuary (Chapter 4), which was a highly retentive system. More generally, a large Indirect Flow Index implies that energy–matter entering the system as boundary flows remain in the system for an extended period of time. We would expect this to be reflected in the aggradation ratio, which is what happens.

We conclude that Indirect/Direct and the aggradation ratio necessarily start at different points (min(Indirect/Direct) = 0, min(AGG) = 1), but converge as the system transfer efficiencies increase (Indirect/Direct ≈ AGG).

6.5 SUMMARY

The idea that Indirect/Direct and AGG are strongly associated, as suggested by the results in Chapters 4 and 5, is compelling, and it is reasonable that alternate indicators of ecosystem organization might measure different aspects of the same underlying system elements (Christensen, 1995; Fath et al., 2001, see also Chapter 5). The relationship between network aggradation and the ratio of indirect-to-direct effects is exciting, however, because one measure has a strong thermodynamic interpretation, while the other is wholly derived from network analysis. Linking these measures offers the possibility of bringing together these two approaches. Our analysis partially decouples these measures of ecosystem growth and development. We argue that they must start at different values, but necessarily converge as the internal fractional transfer coefficients increase. In this work, we also evaluate alternate methods of calculating Indirect/Direct and argue that it is more appropriate to make this calculation using input vectors that reflect the observed distribution of inputs. Finally, we identify six ecosystem models in which Indirect/Direct is less than unity, but generally provide additional evidence supporting the hypothesis that indirect flows tend to dominate direct flows in model ecosystems.

6.6 ACKNOWLEDGEMENTS

This work benefited from discussion with the Systems Ecology and Engineering faculty and students at the University of Georgia. We would also like to thank M.B. Beck, D.K. Gattie, H.R. Pulliam, and P.G. Verity for comments on early drafts of this manuscript. SRB was supported in part by the National Science Foundation biocomplexity grant (OPP-00-83381) to the Skidaway Institute of Oceanography.

6.7 REFERENCES

Baird, D., Luczkovich, J., and Christian, R.R., 1998. Assessment of spatial and temporal variability in ecosystem attributes of the St Marks National Wildlife Refuge, Apalachee Bay, Florida. Estuar. Coast. Shelf Sci., **47**:329-349.

Baird, D., McGlade, J.M., and Ulanowicz, R.E., 1991. The comparative ecology of six marine ecosystems. Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci., **333**:15-29.

Baird, D., and Milne, H., 1981. Energy flow in the Ythan Estuary, Aberdeenshire, Scotland. Estuar. Coast. Shelf Sci., **13**:455-472.

Baird, D., and Ulanowicz, R.E., 1989. The seasonal dynamics of the Chesapeake Bay ecosystem. Ecol. Monogr., **59**:329-364.

Borrett, S.R., and Patten, B.C., 2003. Structure of pathways in ecological networks: relationships between length and number. Ecol. Model., **170**:173-184.

Capra, F., 1996. The web of life: a new scientific understanding of living systems, 1st Anchor Books edition. Anchor Books, New York.

Christensen, V., 1995. Ecosystem maturity - towards quantification. Ecol. Model., **77**:3-32.

Christensen, V., and Pauly, D., 1992. Ecopath-II - a software for balancing steady-state ecosystem models and calculating network characteristics. Ecol. Model., **61**:169-185.

Dame, R.F., and Patten, B.C., 1981. Analysis of energy flows in an intertidal oyster reef. Mar. Ecol. Progr. Ser., **5**:115-124.

Fath, B.D., and Borrett, S.R., 2005. A Matlab function for Network Environ Analysis. Environmental Modelling & Software, **XX**:XXX-XXX.

Fath, B.D., and Patten, B.C., 1999. Review of the foundations of network environ analysis. Ecosystems, **2**:167-179.

Fath, B.D., Patten, B.C., and Choi, J.S., 2001. Complementarity of ecological goal functions. J. Theor. Biol., **208**:493-506.

Finn, J.T., 1976. Measures of ecosystem structure and function derived from analysis of flows. J. Theor. Biol., **56**:363-380.

Finn, J.T., 1980. Flow analysis of models of the Hubbard Brook ecosystem. Ecology, **61**:562-571.

Han, B.P., 1997. On several measures concerning flow variables in ecosystems. Ecol. Model., **104**:289-302.

Higashi, M., and Burns, T.P., 1991. Theoretical studies of ecosystems: the network perspective. Cambridge University Press, Cambridge.

Higashi, M., and Patten, B.C., 1986. Further aspects of the analysis of indirect effects in ecosystems. Ecol. Model., **31**:69-77.

Higashi, M., and Patten, B.C., 1989. Dominance of indirect causality in ecosystems. Am. Nat., **133**:288-302.

Jørgensen, S.E., Patten, B.C., and Straškraba, M., 2000. Ecosystems emerging: 4. Growth. Ecol. Model., **126**:249-284.

Leontief, W.W., 1966. Input-output economics. Oxford University Press, New York.

Levin, S.A., 1998. Ecosystems and the biosphere as complex adaptive systems. Ecosystems, **1**:431-436.

Matis, J.H., and Patten, B.C., 1981. Environ analysis of linear compartmental systems: the static, time invariant case. Bull. Int. Stat. Inst., **48**:527-565.

Miller, T.E., and Kerfoot, W.C., 1987. Redefining indirect effects. In: W.C. Kerfoot and A. Sih (Editors), Predation: direct and indirect impacts on aquatic communities, University Press of New England, London, pp 33-37.

Monaco, M.E., and Ulanowicz, R.E., 1997. Comparative ecosystem trophic structure of three US mid- Atlantic estuaries. Mar. Ecol.-Prog. Ser., **161**:239-254.

Müller, F., 1996. Emergent properties of ecosystems - consequences of self-organizing processes? Senckenbergiana maritima, **27**:151-168.

Müller, F., 1998. Gradients in ecological systems. Ecol. Model., **108**:3-21.

Odum, E.P., 1969. The strategy of ecosystem development. Science, **164**:262-270.

Odum, H.T., 1983. Systems ecology: an introduction. Wiley, New York.

Patten, B.C., 1978. Systems approach to the concept of environment. Ohio J. of Sci., **78**:206-222.

Patten, B.C., 1981. Environs: the superniches of ecosystems. Am. Zool., **21**:845-852.

Patten, B.C., 1982a. Environs: relativistic elementary particles for ecology. Am. Nat., **119**:179-219.

Patten, B.C., 1982b. Indirect causality in ecosystems: its significance for environmental protection. In: W.T. Mason and S. Iker (Editors), Research on Fish and Wildlife. Commemorative Monograph in Honor of the First Decade of the US Environmental Protection Agency, Office of Research and Development, US Environmental Protection Agency, Washington, D.C., pp 92-107.

Patten, B.C., 1983. On the quantitative dominance of indirect effects in ecosystems. In: W.K. Lauenroth, G.V. Skogerboe, and M. Flug (Editors), Analysis of ecological systems: state-of-the-art in ecological modelling, Elsevier, Amsterdam, pp 27-37.

Patten, B.C., 1984a. Further developments toward a theory of the quantitative importance of indirect effects in ecosystems. Verh. Gesellschaft für Ökologie, **13**:271-284.

Patten, B.C., 1984b. Toward a theory of the quantitative dominance of indirect effects in ecosystems. Verh. Gesellschaft für Ökologie, **13**:271-284.

Patten, B.C., 1985a. Energy cycling in the ecosystem. Ecol. Model., **28**:1-71.

Patten, B.C., 1985b. Energy cycling, length of food chains, and direct versus indirect effects in ecosystems. Can. Bull. Fish. Aqu. Sci., **213**:119-138.

Patten, B.C., 1991. Network ecology: indirect determination of the life-environment relationship in ecosystems. In: M. Higashi and T. Burns (Editors), Theoretical Studies of Ecosystems: the network perspective, Cambridge University Press, New York, New York, pp 288-351.

Patten, B.C., 1992. Energy, emergy and environs. Ecol. Model., **62**:29-69.

Patten, B.C., 1995. Network integration of ecological extremal principles: exergy, emergy, power, ascendency, and indirect effects. Ecol. Model., **79**:75-84.

Patten, B.C., 1998a. Network orientors: Steps toward a cosmography of ecosystems: Orientors for directional development, self-organization, and autoevolution. In: F. Müler and M. Leupelt (Editors), Eco Targets, Goal Functions, and Orientors, Springer, Berlin, pp 137-160.

Patten, B.C., 1998b. Steps toward a cosmography of ecosystems: 20 remarkable properties of life in environment. In: F. Mueller and M. Leupelt (Editors), Eco Targets, Goal Functions, and Orientors, Springer-Verlag, New York, New York

Patten, B.C., in prep. Holoecology: the unification of nature by network indirect effects. Columbia University Press, New York.

Patten, B.C., Bosserman, R.W., Finn, J.T., and Cale, W.G., 1976. Propagation of cause in ecosystems. In: B.C. Patten (Editor), Systems Analysis and Simulation in Ecology, Vol. IV, Academic Press, New York, pp 457-579.

Patten, B.C., Fath, B.D., Choi, J.S., Bastianoni, S., Borrett, S.R., Brandt-Williams, S., Debeljak, M., Fonseca, J., Grant, W.E., Karnawati, D., Marques, J.C., Moser, A., Müller, F., Pahl-Wostl, C., Seppelt, R., Seinborn, W.H., and Svirezhev, Y.M., 2002. Complex adaptive hierarchical systems. In: R. Costanza and S.E. Jørgensen (Editors), Understanding and solving environmental problems in the 21st century: Toward a new, integrated hard problem science, Elsevier Science, Ltd, Oxford, pp 41-87.

Patten, B.C., Higashi, M., and Burns, T.P., 1990. Trophic dynamics in ecosystem networks: significance of cycles and storage. Ecol. Model., **51**:1-28.

Prigogine, I., and Stengers, I., 1984. Order out of chaos: man's new dialogue with nature. Bantam Books, Toronto; New York.

Rybarczyk, H. 2003. Network model of Somme Estuary. *in* http://www.cbl.cees.edu/~ulan/ntwk/network.html. Ulanowicz, Robert E.

Schneider, E.D., and Kay, J.J., 1994. Life as a manifestation of the second law of thermodynamics. Mathematical and Computer Modelling, **19**:25-48.

Schrödinger, E., 1946. What is life? The physical aspect of the living cell. The Macmillan Company, Cambridge, England.

Ulanowicz, R.E., 1983. Identifying the structure of cycling in ecosystems. Math. Biosci., **65**:219-237.

Ulanowicz, R.E., 1986. Growth and development: ecosystems phenomenology. Springer-Verlag, New York.

Ulanowicz, R.E., 1997. Ecology, the ascendent perspective. Columbia University Press, New York.

Ulanowicz, R.E., Bondavalli, C., and Egnotovich, M.S., 1997. Network Analysis of Trophic Dynamics in South Florida Ecosystem, FY 96: The Cypress Wetland Ecosystem. Annual Report to the United States Geological Service Biological Resources Division Ref. No. [UMCES]CBL 97-075, Chesapeake Biological Laboratory, University of Maryland, Solomons, MD.

Ulanowicz, R.E., Bondavalli, C., and Egnotovich, M.S., 1998. Network Analysis of Trophic Dynamics in South Florida Ecosystem, FY 97: The Florida Bay Ecosystem. Annual Report to the United States Geological Service Biological Resources Division Ref. No. [UMCES]CBL 98-123, Chesapeake Biological Laboratory, University of Maryland, Solomons, MD.

Ulanowicz, R.E., Bondavalli, C., Heymans, J.J., and Egnotovich, M.S., 1999. Network Analysis of Trophic Dynamics in South Florida Ecosystem, FY 98: The Mangrove Ecosystem. Annual Report to the United States Geological Service Biological Resources Division Ref. No.[UMCES] CBL 99-0073; Technical Report Series No. TS-191-99, Chesapeake Biological Laboratory, University of Maryland, Solomons, MD.

Ulanowicz, R.E., Bondavalli, C., Heymans, J.J., and Egnotovich, M.S., 2000. Network Analysis of Trophic Dynamics in South Florida Ecosystem, FY 99: The Graminoid Ecosystem. Annual Report to the United States Geological Service Biological Resources Division Ref. No. [UMCES] CBL 00-0176, Chesapeake Biological Laboratory, University of Maryland, Solomons, MD.

Wootton, J.T., 1994. The nature and consequences of indirect effects in ecological communities. Ann. Rev. Ecol. Syst., **25**:443-466.

Wootton, J.T., 2002. Indirect effects in complex ecosystems: recent progress and future challenges. Journal of Sea Research, **48**:157-172.

Zemansky, M.W., and Dittman, R.H., 1997. Heat and thermodynamics; an intermediate textbook for students of physics, chemistry, and engineering, 7th edition. McGraw-Hill, New York.

6.8 TABLES

Table 6.1: Twenty ecosystem models of energy flux

Model #	Name	Ecosystem Type	Location	Flux Units	Spatial Extent	References
1	Oyster Reef Community	Oyster Reef	South Carolina	kcal m^{-2} d^{-1}	NR*	Dame and Patten, 1981
2	Cheasapeake Bay, Summer	Bay	mesohaline region of Chesapeake Bay, Maryland	mg C m^{-2} $summer^{-1}$	598,000 ha	Baird and Ulanowicz, 1989
3	Narragansett Bay	Bay	Narragansett Bay, Rhode Island	mg C m^{-2} y^{-1}	NR	Monaco and Ulanowicz, 1997
4	Ythan Estuary	Estuary	Aberdeenshire, Scotland	mg C m^{-2} y^{-1}	256 ha	Baird and Milne, 1989
5	Somme Estuary	Estuary	Northern France	mg C m^{-2} y^{-1}	NR	Rybarczyk, 2003
6	Crystal River Marsh, Control	Tidal Marsh	Crystal River, Florida	mg C m^{-2} d^{-1}	NR	Ulanowicz, 1986
7	Crystal River Marsh, Thermally Impacted	Tidal Marsh	Crystal River, Florida	mg C m^{-2} d^{-1}	NR	Ulanowicz, 1986
8	Seagrass, Live Oak Island, January	Seagrass	St. Marks National Wildlife Refuge, Apalachee Bay, Florida	mg C m^{-2} d^{-1}	4.5 ha	Baird et al., 1998
9	Seagrass, Live Oak Island, February	Seagrass	St. Marks National Wildlife Refuge, Apalachee Bay, Florida	mg C m^{-2} d^{-1}	4.5 ha	Baird et al., 1998
10	Seagrass, Wakulla Beach, January	Seagrass	St. Marks National Wildlife Refuge, Apalachee Bay, Florida	mg C m^{-2} d^{-1}	4.5 ha	Baird et al., 1998
11	Seagrass, Wakulla Beach, February	Seagrass	St. Marks National Wildlife Refuge, Apalachee Bay, Florida	mg C m^{-2} d^{-1}	4.5 ha	Baird et al., 1998
12	Seagrass, Sprague Point, January	Seagrass	St. Marks National Wildlife Refuge, Apalachee Bay, Florida	mg C m^{-2} d^{-1}	4.5 ha	Baird et al., 1998
13	Seagrass, Lighthouse Point, February	Seagrass	St. Marks National Wildlife Refuge, Apalachee Bay, Florida	mg C m^{-2} d^{-1}	4.5 ha	Baird et al., 1998
14	Florida Bay, Dry Season	Bay	South Florida	mg C m^{-2} y^{-1}	2,200,000 ha	Ulanowicz et al., 1998
15	Florida Bay, Wet Season	Bay	South Florida	mg C m^{-2} y^{-1}	2,200,000 ha	Ulanowicz et al., 1998
16	Cypress Wetland, Wet Season	Forested Wetland	Big Cypress Natural Preserve, Florida	mg C m^{-2} y^{-1}	295,000 ha	Ulanowicz et al., 1997
17	Graminoid Marshes and Sloughs, Dry Season	Freshwater Wetland	Everglades, Florida	mg C m^{-2} y^{-1}	214,000 ha	Ulanowicz et al., 2000
18	Graminoid Marshes and Sloughs, Wet Season	Freshwater Wetland	Everglades, Florida	mg C m^{-2} y^{-1}	214,000 ha	Ulanowicz et al., 2000
19	Mangrove, Dry Season	Mangrove	South Florida	mg C m^{-2} y^{-1}	171,000 ha	Ulanowicz et al., 1999
20	Mangrove, Wet Season	Mangrove	South Florida	mg C m^{-2} y^{-1}	171,000 ha	Ulanowicz et al., 1999

Model # is an identification number used on other tables and figures; NR indicates the spatial extent of the study was not reported.

Table 6.2: Network Environ Analysis of 20 energy/carbon ecosystem models

Model #	n	C	L/n	$\lambda_1(A)$	d	Pr(d)	TST	Boundary	BFI	DFI	FCI	IFI	AGG	Indirect/Direct (1)	Indirect/Direct (2)
1	6	0.33	2.00	2.15	0.15	0.00 *	84	41	0.50	0.20	0.11	0.31	2.02	1.53	1.58
2	36	0.09	3.39	2.85	0.54	0.00 *	3,227,453	888,791	0.28	0.19	0.19	0.53	3.63	3.07	2.76
3	32	0.15	4.94	5.99	1.05	0.33	3,917,246	693,846	0.18	0.10	0.51	0.72	5.65	8.13	6.93
4	13	0.23	3.00	3.23	0.23	0.01 *	4,182	1,259	0.30	0.24	0.24	0.46	3.32	2.15	1.89
5	9	0.30	2.67	2.53	0.14	0.00 *	2,035	876	0.43	0.33	0.14	0.24	2.32	0.67	0.74
6	21	0.19	3.90	4.70	0.79	0.57	15,063	7,358	0.49	0.31	0.07	0.20	2.05	0.62	0.67
7	21	0.14	2.90	3.87	0.97	0.20	12,032	6,018	0.50	0.29	0.09	0.21	2.00	0.69	0.71
8	51	0.08	3.86	3.68	0.18	0.07	1,316	515	0.39	0.23	0.13	0.38	2.56	2.09	1.70
9	51	0.08	4.27	3.83	0.44	0.68	1,591	602	0.38	0.23	0.11	0.39	2.64	1.73	1.66
10	51	0.07	3.55	3.73	0.18	0.08	1,383	603	0.44	0.25	0.09	0.31	2.29	1.40	1.24
11	51	0.08	3.94	3.67	0.27	0.00 *	1,921	801	0.42	0.25	0.08	0.33	2.40	1.37	1.30
12	51	0.05	2.49	3.42	0.93	0.22	12,651	7,809	0.62	0.34	0.01	0.04	1.62	0.35	0.12
13	51	0.08	3.98	3.72	0.26	0.00 *	2,865	1,433	0.50	0.29	0.04	0.21	2.00	0.74	0.71
14	125	0.13	15.75	11.01	4.74	0.00 *	1,779	548	0.31	0.30	0.08	0.39	3.25	1.45	1.29
15	125	0.12	15.50	10.97	4.54	0.00 *	2,722	739	0.27	0.27	0.14	0.46	3.69	1.91	1.73
16	68	0.12	8.01	6.85	1.17	0.00 *	2,572	1,419	0.55	0.28	0.04	0.17	1.81	1.71	0.62
17	66	0.18	12.02	11.06	0.96	0.00 *	7,520	3,473	0.46	0.22	0.04	0.31	2.17	1.00	1.41
18	66	0.18	12.02	11.06	0.96	0.00 *	13,677	6,272	0.46	0.25	0.02	0.30	2.18	0.81	1.20
19	94	0.15	14.24	14.17	0.08	0.38	3,272	1,531	0.47	0.20	0.10	0.34	2.14	1.74	1.70
20	94	0.15	14.26	14.16	0.10	0.29	3,266	1,532	0.47	0.20	0.10	0.33	2.13	1.69	1.69

Model # refers to the model identification number on Table 6.1; n is the number of nodes or compartments, $C = L/n^2$ is connectance where L is the number of observed flows; L/n is link density; $\lambda_1(A)$ is the maximum rate of pathway proliferation; $d = |\lambda_1(A) - L/n|$; Pr(d) is the fraction of an ensemble of 1001 random digraphs in which d is greater than or equal to that observed in the original model; #K is the number of non-trivial strongly connected components; %K is the proportion of model nodes involved in a strongly connected components; TST is total system throughflow (units as on Table 6.1); Boundary is the total boundary input (units as on Table 6.1); FCI is the Finn Cycling Index; IFI is the Indirect Flow Index; FCI/IFI is the proportion of indirect flow derived from cycling; AGG = TST/Boundary is network aggradation; Indirect/Direct (1) and Indirect/Direct (2) are the ratios of indirect-to-direct flows calculated using alternative methods (see text)

6.9 FIGURE LEGENDS

Figure 6.1: Network structure of 20 ecosystem models of energy/carbon flux. A) Proportion of model nodes involved in a strongly connected component (%K). B) Relationship between the maximal rate of pathway proliferation ($\lambda_1(A)$) and link density (L/n). Diagonal line represents the expected relationship in random digraphs.

Figure 6.2: Relationships between A) alternative methods of calculating the indirect-to-direct flow ratio (Indirect/Direct) and B) Indirect/Direct and network aggradation (AGG) in 20 ecosystem models of energy flux. Indirect/Direct (1) is calculated using the formula presented in Fath and Patten (1999) and Indirect/Direct (2) is calculated as in Chapter 4. Dark diagonal lines represent the 1:1 line where the compared indicators would be identical. Lighter gray lines represent the point at which Indirect/Direct or AGG equals unity. Above (or to the right of) this line indicates indirect flows dominate direct flows for Indirect/Direct. It identifies the theoretical minimum of AGG. The dashed line in B) is the ordinary least squares linear regression (AGG = 0.60 * Indirect/Direct (2) + 1.65, $R^2 = 0.80$, $p<0.0001$). Numbers in the figure correspond to Model # in Table 6.1.

Figure 6.3: Indirect flows in 20 ecosystem models of energy/carbon flux. A) Ratio of indirect–to–direct flow partitions of TST (method 2, see text). In 14/20 models, Indirect/Direct (2) is greater than one (marked by thin horizontal line), indicating the dominance of indirect effects. B) Partition of TST into boundary, direct, and indirect flow components. Model numbers correspond to Table 6.2. Underlined

models represent the same ecosystem at different points in time, space, or under different conditions. See Table 6.2 and text for details.

Figure 6.4: Proportion of indirect flow derived from recycling in 20 ecosystem models of energy/carbon flux. This is the ratio of the Finn Cycling Index (FCI) divided by the Indirect Flow Index (IFI).

Figure 6.5: Relationships between Indirect/Direct (2) and A) the number of nodes (n), B) connectance (C), C) maximum rate of pathway proliferation ($\lambda_1(A)$), and D) the Finn Cycling Index (FCI) in 20 ecosystem models of energy/carbon flux. The simple linear regression lines in A), B), and C) are not statistically significant; in D) $R^2 = 0.86$ including model 3 (Narragansett Bay) and $R^2 = 0.49$ with out it.

Figure 6.1

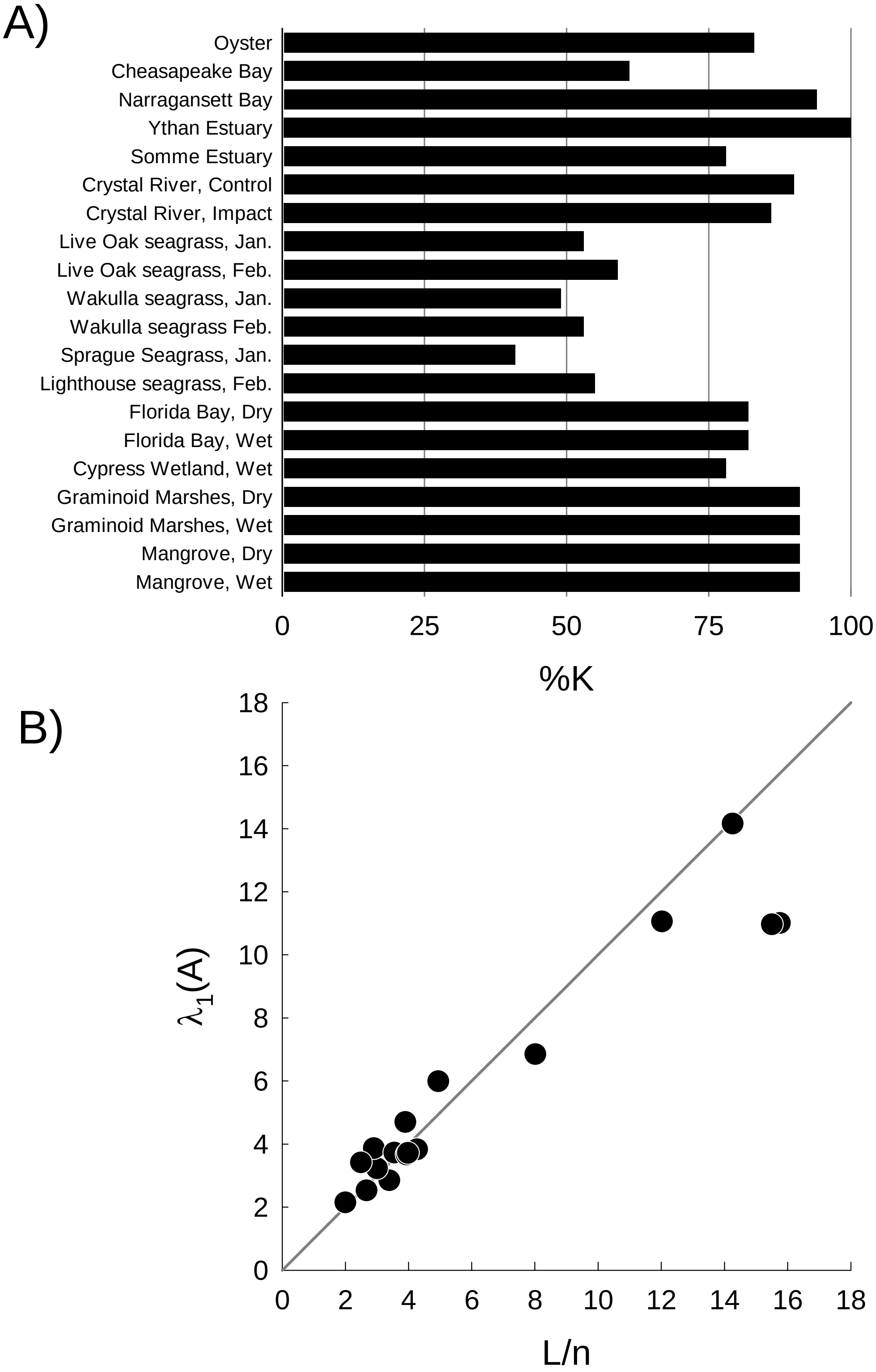

A)
Oyster
Cheasapeake Bay
Narragansett Bay
Ythan Estuary
Somme Estuary
Crystal River, Control
Crystal River, Impact
Live Oak seagrass, Jan.
Live Oak seagrass, Feb.
Wakulla seagrass, Jan.
Wakulla seagrass Feb.
Sprague Seagrass, Jan.
Lighthouse seagrass, Feb.
Florida Bay, Dry
Florida Bay, Wet
Cypress Wetland, Wet
Graminoid Marshes, Dry
Graminoid Marshes, Wet
Mangrove, Dry
Mangrove, Wet
0
25
50
75
100
%K
B)
18
16
14
12
10
8
6
4
2
0
$\lambda_1(A)$
0
2
4
6
8
10
12
14
16
18
L/n

Figure 6.2

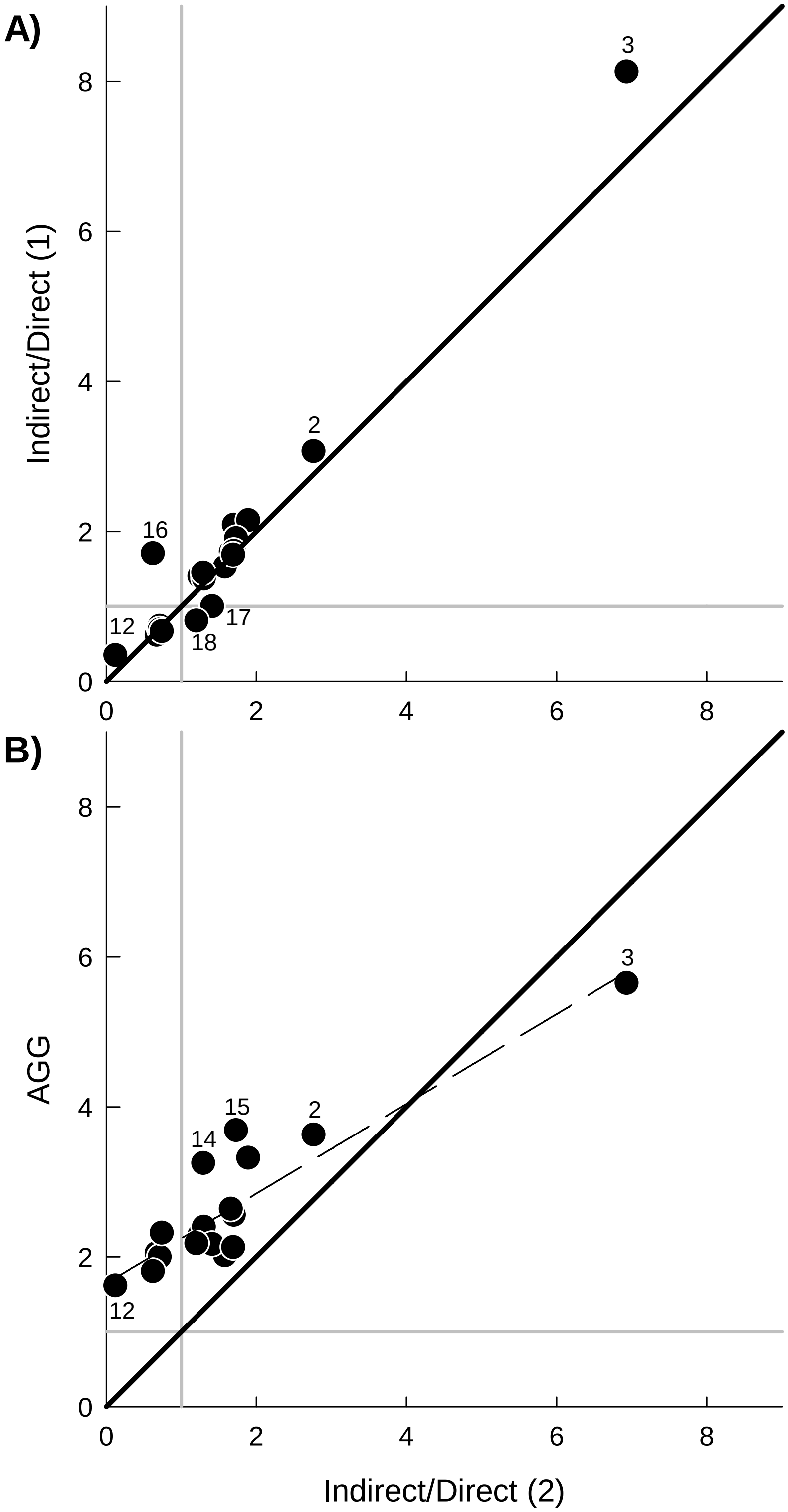

A)
Indirect/Direct (1)
0
2
4
6
8
3
2
16
12
17
18
0
2
4
6
8
B)
AGG
0
2
4
6
8
3
15
2
14
12
0
2
4
6
8
Indirect/Direct (2)

Figure 6.3

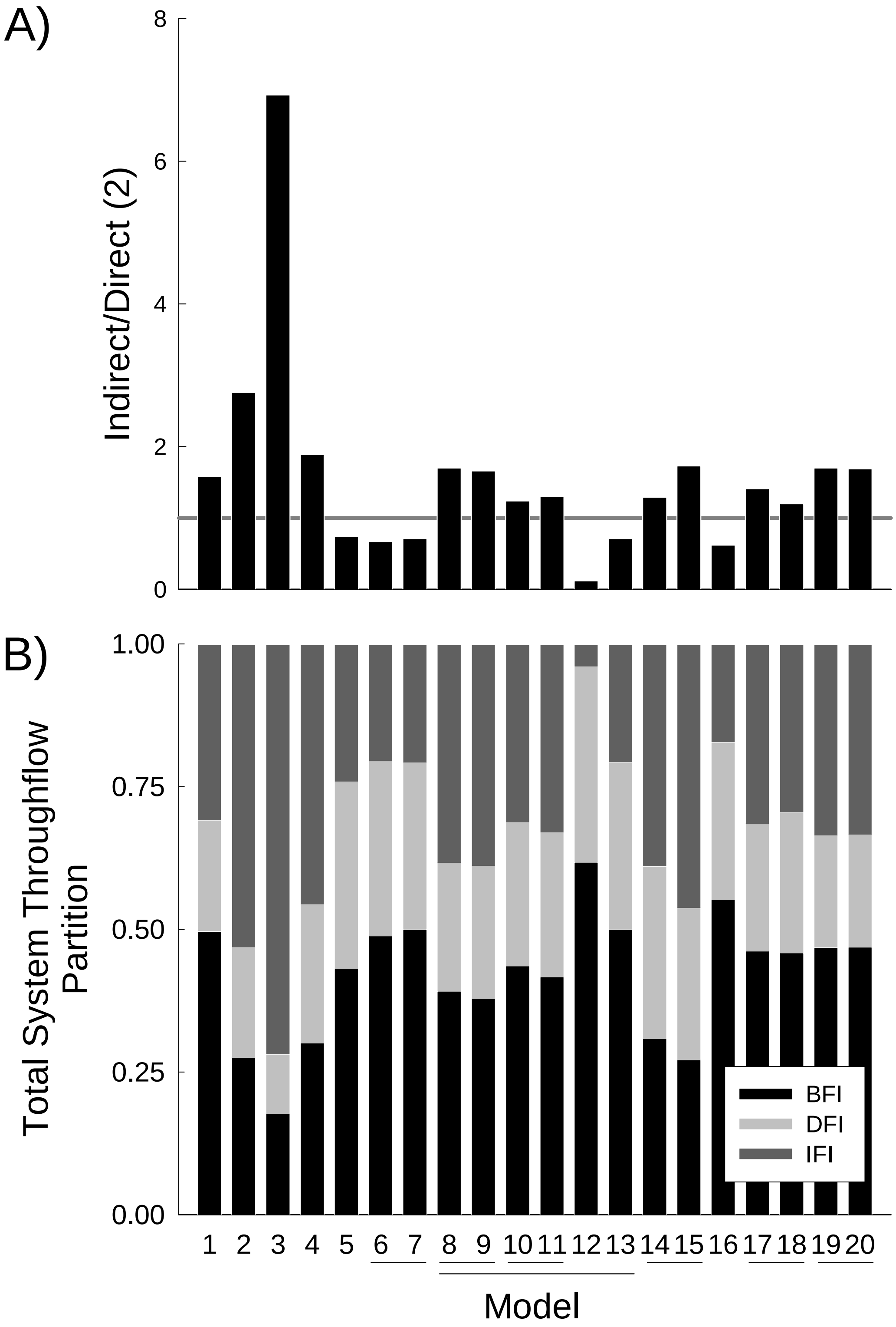

A)
Indirect/Direct (2)
0
2
4
6
8
B)
Total System Throughflow Partition
0.00
0.25
0.50
0.75
1.00
BFI
DFI
IFI
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Model

Figure 6.4

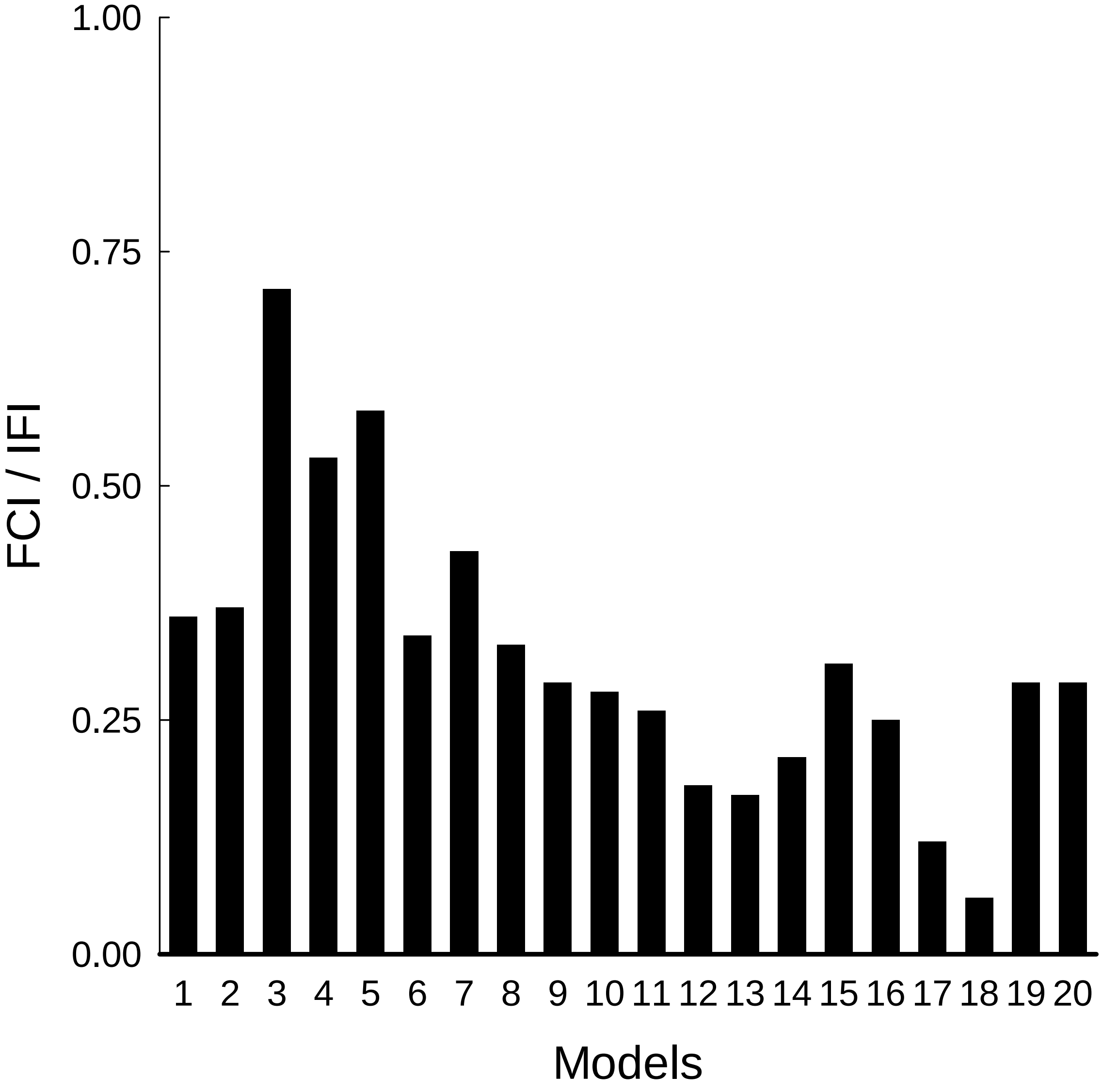
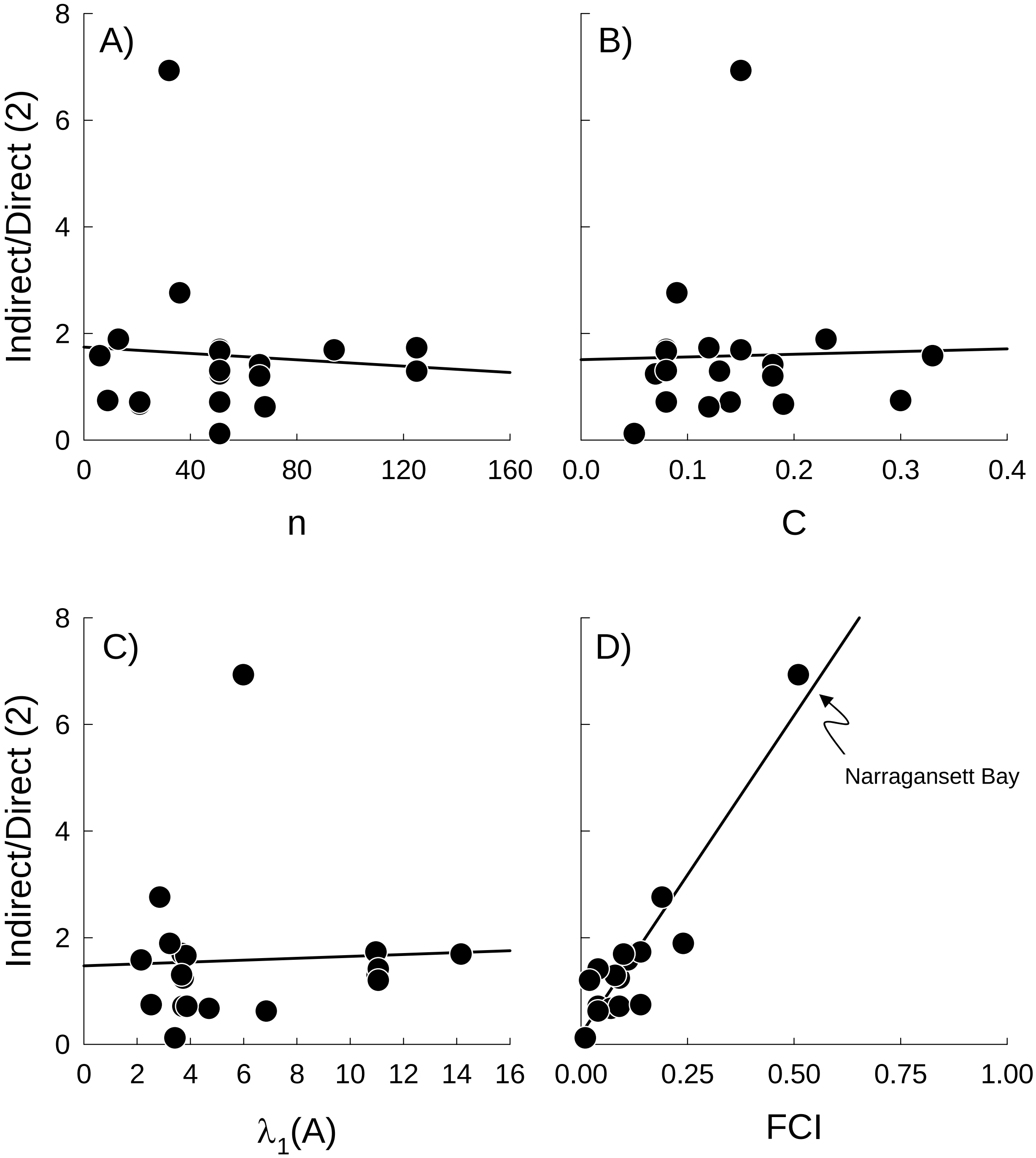

1.00
0.75
0.50
0.25
0.00
FCI / IFI
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Models

Figure 6.5


A)
B)
C)
D)
Indirect/Direct (2)
Indirect/Direct (2)
0
2
4
6
8
0
40
80
120
160
n
0.0
0.1
0.2
0.3
0.4
C
0
2
4
6
8
10
12
14
16
$\lambda_1(A)$
0.00
0.25
0.50
0.75
1.00
FCI
Narragansett Bay

CHAPTER 7

SUMMARY AND CONCLUSIONS

Here I summarize the key results of the individual chapters, discuss their significance, and develop some of their broader implications.

7.1 SUMMARY

The overarching goal of this work was to contribute to a deeper understanding of the processes that create, constrain, and sustain ecological systems. In this effort, we choose to focus on the role of ecosystem architecture (form and function) in the development of indirect effects because they appear to be key components of ecological interactions and evolution. Indirect effects have been distinguished into several types; our investigations focused on a category which results from energy–matter transactions between organisms and their environments, termed environ indirect effects.

This dissertation is comprised of two chapters that focus on ecosystem structure (2 and 3), two chapters that examine the effect of functional variability while structure is constant (4 and 5), and one chapter that investigates the relative role of structure and function in the development of environ indirect effects (6). In Chapters 2 and 3, my coauthors and I probed two structural patterns: *pathway proliferation* and *strongly connected components*. Pathway proliferation is the tendency for the number of pathways in an ecosystem to increase without bound as pathway length increases, and strongly connected components are a subset of system elements in which it is possible to get from any element to another. In Chapter 2, we introduced a measure of whole-system pathway proliferation and demonstrated it is sensitive to the number of model nodes, adjacent connections, and their arrangement. We examined the mathematical foundations for a node–node measure of the pathway proliferation rate in Chapter 3, and in the process revealed a form of system modularity (a hierarchical system subdivision into more or less interacting subsystems) based on strongly connected components. Further, our work found that the majority of 17 large empirically based food-web models contained one or more strongly connected components, each exhibiting pathway proliferation. These structures are ecologically relevant because they create channels for autocatalytic (positive) and cybernetic (negative) feedback and provide conditions we expect to increase the likelihood that indirect flows will dominate direct flows.

In Chapters 4 and 5, we delved into two case studies of Indirect/Direct in ecosystem models of the Neuse River Estuary and Lake Sidney Lanier, respectively. In both chapters we examined multiple models of the system that have a consistent structure, but differ in their flow and storage magnitudes. In Chapter 4, we examined a sixteen season time series of nitrogen flux in the Neuse River Estuary. Our results revealed surprisingly little seasonal and no significant interannual temporal variation in the proportion of total system nitrogen throughflow derived from indirect flows (indirect flow index); indirect flows were consistently large. This system appears to be tightly bound by environ indirect effects, largely generated by microbially mediated nitrogen recycling. In Chapter 5, we inspected the sensitivity of multiple measures of ecosystem growth and development – including Indirect/Direct – to uncertainty in the flow and storage magnitudes in a model of phosphorus flux in Lake Sidney Lanier. Like the Neuse River Estuary models, network structure remained constant while functional values varied. In this case the functional variability was generated in part by inadequate empirical data. Our analysis showed that in the population of plausible models (based on known empirical data), the ecosystem indicators were differentially robust to model uncertainty, and their common variation could be projected onto two latent factors. The indirect flow index and network homogenization were the most robust indicators, and we tentatively interpreted the two latent factors as 1) system integration, and 2) environmental (boundary) influences. As in the sixteen seasons of the Neuse

River Estuary, indirect flows were the dominant component of total system throughflow in each of the plausible parameterizations.

Building on the separate roles of structure (Chapters 2 and 3) and function (Chapters 4 and 5), in Chapter 6 we assessed their relative roles in the development of Indirect/Direct in 20 ecosystem models of energy flux drawn from the literature. We found that while certain structural elements of these models were variable (number of nodes and connectance), they all had at least one strongly connected component that encompassed a large fraction of the model compartments and exhibited some degree of pathway proliferation. Given these structural elements, we expected indirect flows to be consistently dominant. We learned, however, that structure is necessary, but not sufficient for Indirect/Direct to exceed unity. It creates the potential, but the functional information dictates how the structure is utilized. In the majority of the models indirect flows were dominant, but we found that in six models Indirect/Direct did not exceed unity. The number of model compartments and connectivity did not appear to influence Indirect/Direct. In this chapter, we also investigated the relationship between Indirect/Direct and a thermodynamic indicator proposed by Patten – dissipation specific total system throughflow – termed *network aggradation* (Jørgensen et al., 2000). Analysis of the Neuse and Lanier models showed a strong association between these two variables ($R^2 = 0.99$ and $R^2 = 0.96$, respectively). This statistical relationship remained in the energy model data set, but was weaker ($R^2 = 0.80$, $R^2 = 0.52$ when a potential outlier was excluded). Further consideration revealed that these measures necessarily start at different points, but converge as the within-system transfer efficiencies increase. Model structural differences appeared to have little impact on Indirect/Direct. Finally we evaluated alternative methods of calculating Indirect/Direct and concluded that it is important to account for the distribution of boundary inputs in the calculation. This is because indirect flows are generated by the integration of the internal system organization and the environmental boundary flows that activate the system.

7.2 CONCLUSIONS

From these results, conclusions can be drawn about system formation and the role of ecosystem architecture in the development of indirect effects.

Consideration of these results in a broader context reveals thier implications for our understanding of system formation in general, and ecosystems in particular. There appear to be at least two critical points in system formation. Patten et al. (1976) define a system generally "…as a collection of interacting or interdependent objects" (p. 465). They then argue that the creation of one causal bond between two objects is a critical point in system formation. This connection formally binds the two elements into a system, and generates the possibility of new emergent properties. Increased total system throughflow ("power") and growth away from thermodynamic equilibrium are examples. Our results suggest a second critical point occurs when a causal cycle forms, creating a strongly connected component. Again, at this second critical point new and novel system properties can emerge. Cycle formation creates a route for the development of cybernetic feedbacks to propagate throughout the strongly connected component. It is also possible for this strongly connected component to act as a positive feedback autocatalytic cycle – the subsystem catalyzes its own activity. Building on the work of Ulanowicz (1983; 1986; 1997), we argue in Chapter 3 that this is likely to occur in ecological networks. Ulanowicz (1986; 1997) further identifies the novel properties that can emerge in autocatalytic cycles, and thus in ecosystems. The significance of recycling is recognized in the ecological literature (e.g., Odum, 1953; Odum, 1969; Finn, 1977; Patten and Odum, 1981; Finn, 1983; Patten, 1985; Patten et al., 1990; Stone and Berman, 1993; Porter et al., 1996), but here I argue that it is perhaps more fundamental to system formation than perhaps previously realized.

The focus of this dissertation work is on the role of ecosystem architecture in the development of environ indirect effects. We found that while there are logical reasons for network forms, like strongly connected components and pathway proliferation, to be significant in the creation of indirect flows, they only enhanced the possibility. Realized indirect flows are finally determined by the heterogeneous magnitude and distribution of flow across the structure. In summary, ecosystem structure is necessary but not sufficient to determine the significance of indirect flows in the system. The distribution of internal and boundary flows is critical. Though Frank Lloyd Wright was discussing architecture, his statement that form and function are one may be equally correct when applied to ecosystems.

In closing, the research presented in this dissertation adds to our growing theoretical understanding of ecosystem organization and transformation. It characterizes new aspects of the role ecosystem architecture plays in the development of indirect effects, and extends the methodology of Network Environ Analysis. These fundamental developments should provide a firmer foundation for critical environmental management concepts including ecosystem health, integrity, and sustainability. However, there remains much work to fully understand the causes and consequences of ecosystem organization and transformations.

7.3 REFERENCES

Finn, J.T., 1977. Flow analysis: a method for tracing flows through ecosystem models. Ph.D. Dissertation. University of Georgia, Athens, GA.

Finn, J.T., 1983. Ecology - Cycling Index. Nature, 304:397-397.

Jørgensen, S.E., Patten, B.C., and Straškraba, M., 2000. Ecosystems emerging: 4. Growth. Ecol. Model., 126:249-284.

Odum, E.P., 1953. Fundamentals of ecology. Saunders, Philadelphia,.

Odum, E.P., 1969. The strategy of ecosystem development. Science, 164:262-270.

Patten, B.C., 1985. Energy cycling in the ecosystem. Ecol. Model., 28:1-71.

Patten, B.C., Bosserman, R.W., Finn, J.T., and Cale, W.G., 1976. Propagation of cause in ecosystems. In: B.C. Patten (Editor), Systems Analysis and Simulation in Ecology, Vol. IV, Academic Press, New York, pp 457-579.

Patten, B.C., Higashi, M., and Burns, T.P., 1990. Trophic dynamics in ecosystem networks: significance of cycles and storage. Ecol. Model., 51:1-28.

Patten, B.C., and Odum, E.P., 1981. The cybernetic nature of ecosystems. Am. Nat., 118:886-895.

Porter, K.G., Saunders, P.A., Haberyan, K.A., Macubbin, A.E., Jacobsen, T.R., and Hodson, R.E., 1996. Annual cycle of autotrophic and heterotrophic production in a small, monomictic Piedmont lake (Lake Oglethorpe): Analog for the effects of climactic warming on dimictic lakes. Limnol. Oceanogr., 41:1041-1051.

Stone, L., and Berman, T., 1993. Positive feedback in aquatic ecosystems - the case of the microbial loop. Bull. Math. Biol., 55:919-936.

Ulanowicz, R.E., 1983. Identifying the structure of cycling in ecosystems. Math. Biosci., 65:219-237.

Ulanowicz, R.E., 1986. Growth and development: ecosystems phenomenology. Springer-Verlag, New York.

Ulanowicz, R.E., 1997. Ecology, the ascendent perspective. Columbia University Press, New York.

APPENDIX A

A MATLAB® FUNCTION FOR NETWORK ENVIRON ANALYSIS[1]

---

[1] Fath, B.D. and S.R. Borrett. In press. *Environmental Modelling & Software*. Reprinted here with permission of publisher.

ABSTRACT

Network Environ Analysis is a formal, quantitative methodology to describe an object's within system "environ"ment (Patten 1978a). It provides a perspective of the environment, based on general system theory and input-output analysis. This approach is one type of a more general conceptual approach called ecological network analysis. Application of Network Environ Analysis on ecosystem models has revealed several important and unexpected results (see e.g., Patten 1982, 1985, Higashi and Patten 1991, Fath and Patten 1999), which have been identified and summarized in the literature as network environ properties. To conduct the analysis one needs ecosystem data including the intercompartmental flows, compartmental storages, and boundary input and output flows. The software presented herein uses these data to perform the main network environ analyses and environ properties including unit environs, indirect effects ratio, network homogenization, network synergism, network mutualism, mode partitioning, and environ control. The software is available from The MathWorks MATLAB® Central File Exchange website (http://www.mathworks.com/matlabcentral/fileexchange/loadCategory.do).

A.1 INTRODUCTION

Ecological Network Analysis (ENA) is a mathematical methodology to study within system interactions for a given system structure (connectance pattern), function (flow regime), and boundary input. One could say that ENA is a formal realization of synecology, which is mostly concerned with interrelations of material, energy and information among system components as opposed to autecology, which focuses on the individual organisms and populations themselves. ENA starts with the assumption that a system can be represented as a network of nodes (vertices, compartments, components, storages, objects, etc.) and connections between them (arcs, links, flows, etc.). In ecological systems the connections are often based on the flow of conservative units such as energy, matter, or nutrients between the system compartments. If such a flow exists, then we say there is a direct transaction between the two connected compartments. These direct transactions give rise to both direct and indirect relations between all the objects in the system. Network analysis provides a system-oriented perspective because it is based on uncovering patterns and influences among all the objects in a system. Therefore, it gives a view on how components are tied to a larger web of interactions.

The intellectual lineage for Ecological Network Analysis comes from economics, which developed the ability to quantify indirect monetary flows in economic systems. Hannon (1973) first applied economic input-output analysis (Leontief 1951, 1966) to investigate flow distribution in ecosystems. His models were linked by the energy flow through the food web and he pursued this line of research primarily to determine interdependence of organisms in an ecosystem based on their direct and indirect energy flows. Several formulizations of ENA have arisen including Embodied Energy Analysis (Herendeen 1981, 1989), Ascendency Analysis

(Ulanowicz 1980, 1986, 1997) and Network Environ Analysis (Patten 1978a, 1981, 1982, 1985, in prep.).

A recent paper by Allesina and Bondavalli (2004) presented a user friendly, Windows-based version of Ascendency Analysis (Ulanowicz 1982, 1986, 1997). That paper contributes an important role to increase the visibility and usability of that specific network methodology. In a similar vein, we hope to increase the exposure and facilitate the use of Network Environ Analysis by presenting a MATLAB® function (m-file) to calculate the basic environ parameters and properties.

A.2 NETWORK ENVIRON ANALYSIS

There has been much confusion regarding environment as a concept and how one defines it. On the most basic level, environment is all that is external to an object. This has led to the standard object–environment duality where emphasis is placed on the direct flows that come into contact with the object. This has been carried so far as to consider all indirect flows as nonrelevant or historical, suggesting that including them would lead to an “infinite regress” (Mason and Langenheim 1957). Another interpretation of environment, based on an object embedded in a system, includes the summary contribution of the within-system flows that affect that object. In other words, the environment of an object within a specified system (and all objects are parts of a system) can be refined to recognize the special relationship an object has with the other objects within the system boundary. For distinction, Patten (1978a) termed this second, within-system environment the object’s “environ.” An object’s environ ends at the system boundary. Objects and connections in the external environment, beyond the system boundary, are not distinguished so exchanges between them are not material to the analysis. Exchanges across the boundary of the system with the external environment are deemed inputs or outputs. One important aspect of environ theory is an explicit representation of the two environs, input and output, for each object. Furthermore, it is possible to quantify the environs, and thus, the direct and indirect effects between any two objects in the system. In principle, environ analysis can be applied to a system as simple as a pair of interconnected objects. In practice it has been most readily applied to models of entire ecosystems.

Patten (1978a) introduced a systems theory of the environment and put forth three foundational tenets. First, each object has both an input environ, those flows introduced at the system boundary leading up to the object, and an output environ, those flows emanating from the object back to the other system objects before exiting at the system boundary. Second, the purpose of a system boundary is to provide a reference state for the system of study, without which environ analysis is impossible because the analysis collapses back down to the object's boundary. A system boundary is necessary to distinguish between the *system's environment* (the infinite regress) and its component *objects' environs* (within system processes). The third key realization is that the individual environs (and the flow carried in each one) are unique such that the system comprises the set union of all environs of each orientation (input or output), which in turn partition the system level of organization.

Network Environ Analysis has also been a fruitful way of investigating system level properties of ecosystems. In particular, a series of "network statistics" such as indirect effects ratio, homogenization, synergism, and mutualism have grown up around this analysis that express the role of each entity in a larger system. See Patten (in prep.), Fath and Patten (1999a), and Fath (2004) for further details regarding these properties and the history of Network Environ Analysis.

The purpose of this paper is to document software to calculate input and output environs and several of the other basic network properties. We first briefly describe the methodology and network properties. We then introduce the software, which is available from The MathWorks MATLAB® Central File Exchange website (http://www.mathworks.com/matlabcentral/fileexchange/loadCategory.do) and reproduced in Appendix A.A. Appendix A.B gives the program output for a well-studied example of an oyster reef model (Dame and Patten 1981). In Appendix A.C, we compile a glossary for the Network Environ Analysis notation and in Appendix A.D a glossary for common MATLAB notation.

A.3 METHODOLOGY

A.3.1 TERMS AND NOTATIONS

The software introduced in this paper summarizes algorithms for the well-documented methodology of Network Environ Analysis (Patten 1978a, 1982, 1985, in prep., Matis and Patten 1981, Patten et al. 1990, Fath and Patten 1999a). In Network Environ Analysis, systems are partitioned into compartments and a conservative substance such as energy, nitrogen, or phosphorus is propagated mathematically through the interconnected network. Let i, j = 1, ..., n represent n storage compartments (nodes) within an open physical system demarcated from its surroundings by a boundary across which conservative energy–matter is exchanged. The environment is traditionally indexed by 0, however, in the MATLAB® code environment is indexed as n+1 because the software does not allow a zero index. Within-system connections are expressed in an adjacency matrix, $\mathbf{A}=(a_{ij})$, corresponding to the model structure, where $a_{ij}=1$ if there is an observed flow from compartment j to compartment i, and $a_{ij}=0$ if there is no flow. Boundary transfers, z_{j0} (or just z_j) = input to j, y_{0i} (or y_i) = output from i, and internal exchanges between compartments, f_{ij} = flow directed from j to i, comprise a set of transactive flows, or transactions, meaning the transferred quantities are conserved. Let the ordered pairs (i, j), (j, 0) and (0, i) be flow arcs, j → i, 0 → j, and i → 0 carrying the corresponding flows f_{ij}, z_j, and y_i. Networks are a synthesis of such binary flows. Inputs (z_j) or flows (f_{ij}) retained in receiving compartments over time become storages (x_i), and storage at i, say, released as flow to j or the environment, respectively, becomes flow (f_{ji}) and output (y_i). The sum of flows into or out of the i'th compartment at any point in time is throughflow, $T_i^{(in)}$ and $T_i^{(out)}$, given by:

$$T_i^{(in)} = z_i + \sum_{j=1}^{n} f_{ij} \text{ and } T_i^{(out)} = \sum_{j=1}^{n} f_{ji} + y_i .$$

At steady state, compartmental inflows and outflows are equal such that $dx_i/dt = 0$, and therefore, incoming and outgoing throughflows are equal: $T_i^{(in)} = T_i^{(out)} \equiv T_i$. This notation is used to develop a structural analysis, and four functional analyses (flow, storage, utility, and control analysis); the later two functional analyses can be derived from either flows or storages.

A.3.2 STRUCTURAL ANALYSIS

Structural analysis provides important insight into the pattern and connectivity of a model. Path analysis is one type of structural analysis in Network Environ Analysis that enumerates pathways of a various lengths between components and the rate at which the number of pathways increases as path length increases. The analysis is performed on the adjacency matrix, $\mathbf{A}=(a_{ij})$, of the model. It is a property of matrix multiplication that $\mathbf{A}^m$ gives the number of paths of length m between two components, i and j in the model. In systems with feedback (i.e., all realistic ecological models), the number of pathways increases as m increases; therefore, $\sum_{m=0}^{\infty} A^m$ is a divergent series. This phenomenon is termed pathway proliferation in Network Environ Analysis. The proliferation rate is a significant system attribute because it describes the growth rate of indirect pathways, ultimately describing the number of pathways available for transactions and relations (Borrett and Patten 2003). Like the population growth rate in a Leslie Matrix (Caswell 2001), the pathway proliferation rate ($a_{ij}^{(m+1)}/a_{ij}^{(m)}$ as $m \to \infty$) is given by the largest eigenvalue of **A** (Fath 1998). The software returns the adjacency matrix, the pathway proliferation rate, the number of network nodes (n), the proportion of direct connections completed or the network connectance (L/n^2), and the link density (L/n), where $L = \sum_{i,j=1}^{n} a_{ij}$.

A.3.3 FUNCTIONAL ANALYSIS

Throughflow, Storage, Utility, and Control are four types of functional analysis used in Network Environ Analysis, each providing different insights into a system. Throughflow analysis is similar to input–output analyses performed by other ENAs, but storage, utility, and control analyses are unique to Network Environ Analysis. Here we describe storage analysis in detail, while the others follow as analogues.

A.3.3.1 STORAGE ANALYSIS

Nondimensional, storage-specific, output-oriented, intercompartmental flows are given by $p_{ij} = c_{ij}\Delta t$, for $i \neq j$, where, $c_{ij} = f_{ij}/x_j$; and for $i = j$, $p_{ii} = 1 + c_{ii}\Delta t$, where $c_{ii} = -T_i/x_i$ (Matis and Patten 1981). The dimensional quantities c_{ii} and c_{ij} are elements of a Jacobian "community" matrix, $\mathbf{C}=(c_{ij})$. This matrix type is typically employed in population–community ecology for stability or food-web analyses. By introducing small enough time steps Δt into the $c_{ii}\Delta t$ and $c_{ij}\Delta t$ values, dimensionless p_{ii} and p_{ij} quantities are obtained that lie in the range $0 \leq p_{ii}, p_{ij} \leq 1$, and thus are interpretable as probabilities (Barber 1978). From the matrix $\mathbf{P}=(p_{ij})$ of such probabilities, a dimensionless integral storage intensity matrix $\mathbf{Q}=(q_{ij})$ can be computed as the convergent power series:

$$\mathbf{Q} = \mathbf{P}^0 + \mathbf{P}^1 + \mathbf{P}^2 + \mathbf{P}^3 + \ldots + \mathbf{P}^m + \ldots = (\mathbf{I} - \mathbf{P})^{-1} \tag{A.1}$$

where $\mathbf{P}^0 = \mathbf{I}$ is the identity matrix. The m'th order terms, $m = 1, 2, \ldots$, account for interflows over all pathways in the system of lengths m, $\forall m$. In so doing the network has a graph theory property known as transitive closure (Patten et al. 1976). Input-oriented, storage-specific, intercompartmental flows are given by $p'_{ij} = c'_{ij}\Delta t$, where, for $i \neq j$, $c'_{ij} = f_{ij}/x_i$, and for $i = j$, $p'_{ii} = 1 + c'_{ii}\Delta t$, where $c'_{ii} = -T_i/x_i$ (MATLAB® uses **P'** to denote the transpose of **P**, therefore we use an additional "**P**" to designate "prime" such that **P'** becomes **PP**, and **Q'** becomes **QP** below, etc. in the code). Note the output-oriented values were normalized by the donating compartment storage, x_j, and the input-oriented values by the receiving compartmental storages, x_i. From the matrix $\mathbf{P'}=(p'_{ij})$, a dimensionless, integral, input-oriented, storage intensity matrix $\mathbf{Q'}=(q'_{ij})$ can be computed:

$$\mathbf{Q'} = (\mathbf{P'})^0 + (\mathbf{P'})^1 + (\mathbf{P'})^2 + (\mathbf{P'})^3 + \ldots + (\mathbf{P'})^m + \ldots = (\mathbf{I} - \mathbf{P'})^{-1}. \tag{A.2}$$

The nondimensional output-oriented integral storage matrix **Q** in (A.1) can be redimensionalized by multiplying by the input vector, **z** and the time step, Δt, such that $\mathbf{x}=\mathbf{Qz}\Delta t$; and the nondimensional input-oriented integral storage matrix **Q'** in (A.2) can be redimensionalized by pre-multiplying by the output vector, **y** and the time step, Δt, such that $\mathbf{x}=\mathbf{yQ'}\Delta t$ (order of multiplication for the scalar time step is irrelevant). In some cases, it is also useful to examine the dimensional form of the integral matrix, which we include as **S**=**Q**Δt. The software returns **C, C', P**, **P'**, **Q**, and **Q'**, **S**, and **S'**.

A.3.3.2 THROUGHFLOW ANALYSIS

Similar to the case for storage analysis, nondimensional, output-oriented, intercompartmental *flows* are given by, $g_{ij} = f_{ij}/T_j$; and, input-oriented, intercompartmental *flows* are given by, $g'_{ij} = f_{ij}/T_i$. The dimensionless g_{ij} and g'_{ij} quantities lie in the range $0 \leq g_{ij}, g'_{ij} \leq 1$, and thus are interpretable as probabilities. From the matrices **G**=(g_{ij}) and **G'**=(g'_{ij}), dimensionless integral output and input flow intensity matrices **N**=(n_{ij}) and **N'**=(n'_{ij}) can be computed similar to Equation (A.1) from the convergent power series:

$$\mathbf{N} = \mathbf{G}^0 + \mathbf{G}^1 + \mathbf{G}^2 + \mathbf{G}^3 + \ldots + \mathbf{G}^m + \ldots = (\mathbf{I} - \mathbf{G})^{-1}, \text{ and}$$

$$\mathbf{N'} = (\mathbf{G'})^0 + (\mathbf{G'})^1 + (\mathbf{G'})^2 + (\mathbf{G'})^3 + \ldots + (\mathbf{G'})^m + \ldots = (\mathbf{I} - \mathbf{G'})^{-1}, \text{ and} \qquad \text{(A.4)}$$

where $\mathbf{G}^0 = \mathbf{I}$ is again the identity matrix, and the m'th order terms, m = 1, 2, ..., account for interflows over all pathways in the system of lengths m. The nondimensional output-oriented integral flow matrix can be redimensionalized by multiplying by the input vector, **z**, such that **T**=**Nz** and the nondimensional input-oriented integral flow matrix can be redimensionalized by pre-multiplying by the output vector, **y**, such that **T**=**yN'**. The software returns **G**, **G'**, **N**, and **N'** as well as several system-level summary flow parameters.

A.3.3.3 UTILITY ANALYSIS

Intercompartmental *flow utilities* are given by, $d_{ij} = (f_{ij}–f_{ji})/T_i$. The dimensionless d_{ii} quantities lie in the range $–1 \leq d_{ii} \leq 1$, and thus are *not* interpretable as probabilities. They do, however, conform to the requirements of a convergent series so long as the magnitude of the largest eigenvalue is less than one. Therefore, a test must be executed before the flow utility power series is taken. Networks found not to fulfill the convergence property are excluded from the utility analysis and flagged in the software output as –9999. From the matrix $\mathbf{D}=(d_{ij})$, a dimensionless integral utility intensity matrix $\mathbf{U}=(u_{ij})$ can be computed:

$$\mathbf{U} = \mathbf{D}^0 + \mathbf{D}^1 + \mathbf{D}^2 + \mathbf{D}^3 + \ldots + \mathbf{D}^m + \ldots = (\mathbf{I} - \mathbf{D})^{-1}. \tag{A.5}$$

Note for completeness direct *storage utilities* could be derived from $ds_{ij} = (f_{ij} - f_{ji})/x_i$ and integral storage utilities given by the following power series (convergence restrictions apply):

$$\mathbf{US} = \mathbf{DS}^0 + \mathbf{DS}^1 + \mathbf{DS}^2 + \mathbf{DS}^3 + \ldots + \mathbf{DS}^m + \ldots = (\mathbf{I} - \mathbf{DS})^{-1}, \tag{A.6}$$

however, this parameter has not been thoroughly investigated or presented in the NEA literature. The nondimensional integral flow and storage utility matrices can be redimensionalized by multiplying by the diagonalized throughflow vector, $\check{\mathbf{T}}$, such that $\mathbf{Y}=\mathbf{U}\check{\mathbf{T}}$ and $\mathbf{YS}=\mathbf{US}\check{\mathbf{T}}$. The software returns **D**, **DS**, **U**, **US**, **Y**, and **YS**.

A.3.3.4 CONTROL ANALYSIS

Patten (1978b) introduced a Network Environ Analysis based measure of control or dominance (see also Patten and Auble (1981), Patten 1982, and Fath 2004). The measure, expressed in a matrix $\mathbf{CN} = (cn_{ij})$, is based on the ratio of integral flow from compartment j to i to the integral flow from i to j. Compartment j is said to dominate i if its output environ effect on i is larger than i's input environ effect on j ($cn_{ij} = n_{ij}/n_{ji}' > 1$). This control relationship was further modified such that when $n_{ij}/n_{ji}' < 1$, $cn_{ij} = 1 - n_{ij}/n_{ji}'$ otherwise $cn_{ij} = 0$. The storage version of this measure is **CQ**, where $cq_{ij}=q_{ij}/q'_{ji}$. The storage–based dominance relationships are always identical to the throughflow-based measures because intensive throughflows and storages are related by the compartment turnover rates – which cancel out in the ratio measure. NEA.m calculates both **CN** and **CQ**.

A.4 NETWORK AND ENVIRON PROPERTIES

In this section, we introduce the specific properties the software returns. A listing and brief description of each of these is given in Table 1. The formulas can be found in the code in Appendix A.B.

A.4.1 UNIT ENVIRON ANALYSIS

The first property is the quantitative environ, both in the input and output orientation. Since each compartment has two distinct environs there are 2*n* environs in total in an *n*-compartment system. The output environ $\mathbf{E_k} = (e_{ijk})$ ($i = 1, 2, \ldots, n+1$, $j = 1, 2, \ldots, n+1$, $k = 1, 2, \ldots, n$) for the kth compartment is calculated by multiplying **G** times the diagonalized matrix of the kth column of **N** minus the diagonal of the kth column of **N**, such that $e_{ijk} = g_{ij}$*diag(n_{ik})-diag(n_{ik}) for $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$. In MATLAB code: E(1:n,1:n,k) = G*diag(N(:,k))-diag(N(:,k)); When constructed in the manner, elements on the principle diagonal of $\mathbf{E_k}$ are the negative throughflows of each compartment generated by the unit input. The column sums of $\mathbf{E_k}$ ($i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$) give the negative output vector which is multiplied by −1 and incorporated into $\mathbf{E_k}$ as row $n + 1$, and row sums the negative input vector which is multiplied by −1 and incorporated into $\mathbf{E_k}$ as column $n + 1$. When assembled, the result is the output oriented flow from each compartment to each other compartment in the system and across the system boundary. In a similar fashion, input environs are calculated by multiplying the diagonalized matrix of the kth column of **N'** by **G'** minus the diagonal of the kth column of **N'** (in code: EP(1:n,1:n,i)=diag(NP(k, :))*GP -diag(NP(k, :))). Input (**SE'**) and output (**SE**) oriented storage environs are calculated in analogous fashion using **P**, **Q**, **P'**, and **Q'** (Matis and Patten 1981). These results comprise the foundation of Network Environ Analysis since they allow for the quantification of all within system interactions, both direct and indirect, on a compartment-by-compartment basis. In the software, **E**, **E'**, **SE**, and **SE'** are returned as $n+1 \times n+1 \times n$ arrays. The n+1 row is the unit output vector and the n+1 column is the unit input vector for each n environs.

A.4.2 INDIRECT TO DIRECT EFFECTS MEASURE

This property is one of the most important since it compares the strength of indirect (non-touching) flow in the compartment's environ to the direct flow. Indirect effects are calculated as the integral contributions minus the direct and initial boundary input. The indirect to direct effects ratio is a measure of the relative strength of these two factors. Mathematically, this is simply the following ratio for the output oriented throughflow case: $I/D = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n}(n_{ij} - i_{ij} - g_{ij})}{\sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}}$.

When the ratio is greater than one, indirect flows are greater than direct flows. This ratio can be calculated for the input and output oriented throughflow and storage analyses (respectively these are denoted as I/D(T,in), I/D(T,out), I/D(S,in), I/D(S,out)). Analysis of many models has shown that these ratios are often greater than one, indicating the non-intuitive result that indirect effects have greater contribution than direct effects (Higashi and Patten 1989). This is core evidence for the rationale behind systems modeling and systems perspective because it states that indirect effects are greater, therefore exerting greater dominance, than direct effects in a network. This clearly has implications for understanding feedback and direct versus indirect control in networks.

A.4.3 NETWORK HOMOGENIZATION

The homogenization property yields a comparison of resource distribution between the direct and integral flow intensity matrices. Due to the contribution of indirect pathways, flow in the integral matrix tends to be more evenly distributed than that in the direct matrix. A statistical comparison of the distributions can be made by calculating the coefficient of variation of the direct and integral matrices (Fath and Patten 1999b). For example, the coefficient of variation of the direct flow intensity matrix **G** is given by:

$$\mathrm{CV}(\mathbf{G}) = \frac{\sqrt{\dfrac{\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\bar{g}_{ij} - g_{ij}\right)^2}{(n-1)}}}{\bar{g}_{ij}} \qquad \text{(A.7)}$$

where $\bar{g}_{ij}$ is the mean of the elements of **G**.

Network homogenization occurs in the output oriented throughflow case when the coefficient of variation of **N** is less than the coefficient of variation of **G** because this indicates that the network flow is more evenly distributed in the integral matrix. The test statistic employed here looks at whether or not the ratio CV(**G**)/CV(**N**) exceeds one. In a similar fashion, this measure can be applied to the input throughflow and the input and output oriented storage cases. These measures are denoted as Homog(T,in), Homog(T,out), Homog(S,in), Homog(S,out).

A.4.4 NETWORK AMPLIFICATION

The amplification property deals explicitly with the values in the integral flow or storage matrices. Diagonal elements of **N** (**Q**) are almost always great than one. Off-diagonal elements, representing the integral flow from j to i, rarely are, but can exceed one when cycling drives more than the equivalent of one unit of input flow over the pathways linking those two compartments. If any off-diagonal element of **N** (**Q**) is greater than one, then amplification is said to occur because a virtual input or output of one is implied by the nondimensional analysis. This measure can be applied to both the input and output orientations, and is notated as Amp(T,in), Amp(T,out), Amp(S,in), Amp(S,out).

A.4.5 NETWORK SYNERGISM

Synergism implies that positive utility exceeds negative utility in the system. Utility is the throughflow or storage scaled value of net transactions between entity pairs. To determine whether or not this occurs a comparison is made between positive and negative utilities of the dimensionalized integral utility matrix, $\mathbf{Y}=\mathbf{U}\check{\mathbf{T}}$, which quantifies the magnitude of the positive and negative utilities. Synergism is said to occur when the magnitude of positive utility exceeds the magnitude of negative utility, which is the same as saying the ratio of the positive to negative utility exceeds one (Patten 1991, Fath and Patten 1998). For completeness, we include the application of this measure to the storage case, though this parameter has not been thoroughly investigated nor previously presented in the literature. These ratios are denoted as Synergism(T) and Synergism(S).

A.4.6 NETWORK MUTUALISM

In addition to quantifying the direct and indirect relations the utility matrix can be used to determine qualitative relations between any two components in the network such as predation, mutualism, or competition. Entries in the direct utility matrix, **D**, or integral utility matrix, **U,** can be positive or negative ($-1 \leq d_{ij}, u_{ij} \leq 1$). The elements of **D** represent the direct relation between that (i,j) pairing and the elements of **U** the integral relations, respectively (Patten 1991, Fath and Patten 1998). The direct matrix **D**, being zero-sum between complementary pairs d_{ij} and d_{ji}, always has the same number of positive and negative signs. Signs in the integral matrix, **U**, are determined by the entire web of system interactions. If there are more positive signs than negative signs in the integral utility matrix, then network mutualism is said to occur. Network mutualism reveals the preponderance of positive mutualistic relations in the system. Again for completeness, we include the application of this measure to the storage case, though it has not been thoroughly investigated nor previously presented in the NEA literature. These ratios are denoted as Mutualism(T) and Mutualism(S).

A.4.7 MODE PARTITIONING

Flow (and flow derived storage) into and out of a specific compartment can be partitioned into five categories or modes: (0) boundary input, (1) first passage, (2) cycled, (3) dissipative, and (4) boundary output, depending on its position relative to the focal compartment. Because this is a partition the modes are mutually exclusive and exhaustive (Fath et al. 2001). Boundary input is flow that starts from the environment and crosses the system boundary into a compartment within the system. It is calculated as **Iz**, where **I** is the identity matrix. First passage flow, or mode 1, is flow from any compartment that reaches another (focal) compartment for the very first time. Note, since this is compartment specific, flow cycled between other compartments en route to the focal compartment for the first time is still considered first passage. For example, first–passage flow to compartment k from i could have traveled the following path: i→j→i→j→k. Cycled flow, or mode 2, is calculated using a derivation of the Finn (1976) cycling index and represents the amount of flow that has exited a compartment but will return again to that same compartment before being lost from the system such that the compartment in question is both the originating and terminating node for that pathway. Dissipative, mode 3 flow has left the focal compartment never to return again, although it passes through other compartments before crossing the system boundary. Boundary output, mode 4, is flow that exits the system boundary directly from the focal compartment in question. Mode partitioning is described more fully by Higashi et al. (1993) and Fath et al. (2001).

A.5 SOFTWARE

`NEA.m` is a MATLAB® function created to rapidly perform NEA on flow-storage network models (available from http://www.mathworks.com/matlabcentral/fileexchange/loadCategory.do and in Appendix A.A). The function implements algorithms for all analyses described in Sections 3 and 4. It requires one input variable, $\Delta_{n+1 \times n+2}$, that summarizes the flow and storage information for the system of interest. The function returns a vector of the system-level environ properties ($\mathrm{ep}_{30\times 1}$) to the workspace (Table 1), displays the comprehensive list of analytical results in the command window, and saves all results as a MATLAB® data file called `NEA_output.mat`.

A.5.1 INPUT DATA

The input argument $\Delta_{n+1\times n+2}$ is a $(n+1)\times(n+2)$ composite matrix that summarizes the system information of the flow–storage network to be analyzed, where n is the number of compartments or nodes of the network. The composite input matrix is

$$\Delta = \begin{bmatrix} \mathrm{F}_{\mathrm{n}\times\mathrm{n}} & \mathrm{z}_{\mathrm{n}\times 1} & \mathrm{x}_{\mathrm{n}\times 1} \\ \mathrm{y}_{1\times\mathrm{n}} & 0 & 0 \end{bmatrix}_{\mathrm{n}+1\times\mathrm{n}+2},$$

where $\mathbf{F}_{n\times n}$ is the steady-state intercompartmental flow matrix, $z_{n\times 1}$ are the steady-state boundary inputs, $y_{1\times n}$ are the steady-state boundary outputs, and $\mathrm{x}_{\mathrm{n}\times 1}$ are storage values. Δ is oriented such that flows are from columns to rows.

In its present form, Network Environ Analysis makes two critical assumptions about the input data. First, flows and storages must be measured in a consistent conservative energy–matter unit. For example, all fluxes could have units of g C m^{-2} y^{-1} or mg P cm^{-3} d^{-1}, while the corresponding storages would be g C m^{-2} or mg P cm^{-3}. Second, the data must represent a static, steady-state system ($T^{(in)} = T^{(out)}$). When implemented, `NEA.m` first checks to ensure the system meets a steady-state requirement. If the model does not meet this requirement, then a warning is given and the analysis does not proceed. It would be necessary to first balance the network flows; algorithms for this are available (Savenkoff et al. 2001, Allesina and Bondavalli 2003). The static, steady-state assumption is a limitation of the methodology because few ecological systems exist in this condition. Despite this limitation, important insights emerge that appear to challenge conventional ecological theory – such as energy cycling (Patten 1985) and the dominance of indirect effects (Higashi and Patten 1989, Patten in prep.). Nonetheless, further work, like that initiated by Hippe (1983), to develop a dynamic Network Environ Analysis, is needed.

The input variable can be coded as a MATLAB® function so that system data do not need to be reentered multiple times. An example data function for an oyster reef model (Dame and Patten, 1981) is included with `NEA.m`.

A.5.2 IMPLEMENTING NEA.M & DATA OUTPUT

Once `NEA.m` is installed in the MATLAB® operating directory, the function can be implemented on Δ by typing "`ep=NEA(Δ);`" in the command window. Resultant matrices from structural, throughflow, utility, unit environ analyses as well as a table of the system-level environ indices will be displayed in the command window. The system-level indices are also returned to the workspace as the vector 'ep' (Table 1). All results are stored in the MATLAB® data file '`NEA_output.mat`'. This file can be loaded into the workspace, giving the user access to all resultant matrices for additional investigation and manipulation. Typing "`ep=NEA(Δ,0);″` will prevent the results from displaying in the command window.

While all results displayed in the command window can by cut and pasted into word processing or spreadsheet programs, there is a simple way to capture the results in an ASCII text file using the diary function in MATLAB®. To use this function, type `diary('file_name.txt')` in the command window, where 'file_name' is the name of the output file you are creating. Then, run `NEA.m` by typing "`NEA(Δ);`" or "`ep=NEA(Δ);`". When the computations are complete, type "`diary off`" to turn off the diary function. This file can then be opened in any text editor. Example NEA_output.mat and diary files for the Oyster Reef Model (Dame and Patten 1981) are included with the software and in Appendix A.B.

A.6 CONCLUSION

Network Environ Analysis is one branch of Ecological Network Analysis. It is a powerful tool for investigating the within–system transactions and relations in ecological systems. The software presented herein can be used to calculate the primary parameters and properties of Network Environ Analysis. The analysis itself is not computationally challenging, but does require some familiarity with matrix algebra and graph theory concepts. The software compiles the algorithms and should facilitate use of the methodology. Network Environ Analysis is an active area of research such that not all of its facets could be included here. Also, while the software provides the quantitative results, the challenging task of interpretation is left to the user. Previous applications of Network Environ Analysis (Matis and Patten 1981, Patten and Matis 1982, Flebbe 1983, Patten 1983, in prep.) may be useful guides for interpretation. It is our intention that dissemination of this software will encourage others to look more closely at the environ methods and be aided in applying them in their own research.

ACKNOWLDEGEMENTS

The computational methods and algorithms for the mathematical system theory of environment that we summarize and code here were developed through collaboration between several people over many years. We wish to acknowledge Bernard C. Patten who has served as the chief orchestrator and the following collaborators: M. Craig Barber, Robert W. Bosserman, Thomas P. Burns, John T. Finn, Masahiko Higashi, James Hill, IV, James H. Matis, and Stuart J. Whipple. This manuscript benefited from comments by Bernard C. Patten and anonymous reviewers. SRB was supported in part by a grant from the National Science Foundation (OPP-00-83381).

A.7 REFERENCES

Allesina, S., and Bondavalli, C., 2003. Steady state of ecosystem flow networks: a comparison between balancing procedures. Ecological Modelling 165, 221-229.

Allesina, S., and Bondavalli, C., 2004. Wand: an ecological network analysis user-friendly tool. Environmental Modelling & Software 19 (4), 337-340.

Barber, M.C., 1978. A markovian model for ecosystem flow analysis. Ecological Modelling 5, 193–206.

Borrett, S.R., Patten, B.C., 2003. Structure of pathways in ecological networks: relationships between length and number. Ecological Modelling 170 (2-3), 173-184.

Caswell, H., 2001. Matrix population models: Construction, analysis, and interpretation. Sinauer Associates: Sunderland, Mass.

Dame, R.F., Patten, B.C., 1981. Analysis of energy flows in an intertidal oyster reef. Marine Ecol. Progr. Series 5, 115–124.

Fath, B. D., 1998. Network analysis: foundations, extensions, and applications of a systems theory of the environment. (Ph.D. Thesis). University of Georgia, Athens, Georgia.

Fath, B.D., 2004. Network analysis in perspective: comments on WAND: an ecological network analysis user friendly tool. Environmental Modelling and Software 19, 341–343.

Fath, B.D., 2004. Distributed control in ecological networks. Ecological Modelling 179, 235-246.

Fath, B.D., Patten, B.C., 1998. Network synergism: emergence of positive relations in ecological systems. Ecological Modelling 107, 127–143.

Fath, B.D., Patten, B.C., 1999a. Review of the foundations of network environ analysis. Ecosystems 2, 167–179.

Fath B.D., Patten B.C., 1999b. Quantification of resource homogenization using network flow analysis. Ecological Modelling 123, 193–205.

Fath, B. D., Patten, B.C., Choi, J.S., 2001. Complementarity of ecological goal functions. Journal of Theoretical Biology 208, 493-506.

Finn, J.T., 1976. Measures of ecosystem structure and function derived from analysis of flows. Journal of Theoretical Biology 56, 363–380.

Flebbe, P. A., 1983. Aquatic carbon cycle in Okefenokee swamp habitats: environ analysis. In: Lauenroth, W. K., Skogerboe, G. V., Flug, M., (Eds.), Analysis of ecological systems: state-of-the-art in ecological modelling. Elsevier: Amsterdam.

Hannon, B., 1973. The structure of ecosystems. Journal of Theoretical Biology 41, 535–546.

Herendeen, R.A., 1981. Energy intensity in ecological and economic systems. Journal of Theoretical Biology 91, 607–620.

Herendeen, R.A., 1989. Energy intensity, residence time, exergy, and ascendency in dynamic ecosystems. Ecological Modelling 48, 19–44.

Higashi M. Burns, T.P., (Eds.). 1991. Theoretical Studies of Ecosystems: the network perspective. Cambridge University Press: New York.

Higashi, M., Patten, B. C., and Burns, T. P., 1993. Network trophic dynamics – the modes of energy-utilization in ecosystems. Ecological Modelling 66 (1-2) 1-42.

Higashi, M., Patten, B.C., 1989. Dominance of indirect causality in ecosystems. American Naturalist 133, 288–302.

Jørgensen, S. E., Patten, B. C., Straškraba, M., 2000. Ecosystems emerging: 4. growth. Ecological Modelling 126 (2-3), 249-284.

Leontief, W.W., 1951. The structure of American economy, 1919–1939; an empirical application of equilibrium analysis. Oxford University Press: New York.

Leontief, W.W., 1966. Input-output economics. Oxford University Press: New York.

Mason, H.L., Langenheim, J.H., 1957. Language and the concept of environment. Ecology 38, 325–340.

Matis J.H., Patten, B.C., 1981. Environ analysis of linear compartmental systems: the static, time invariant case. Proceedings 42nd Session International Statistics Institute, Manila, Philippines, December 4–14, 1979.

Odum, E.P., 1953. Fundamentals of ecology. W.B. Saunders. Philadelphia, Pennsylvania.

Patten, B.C., 1978a. Systems approach to the concept of environment. Ohio Journal of Science 78, 206–222.

Patten, B.C., 1978b. Energy environments in ecosystems. In: Fazzolare, R.A., Smith, C.B., (Eds.), Energy Use Management, Vol. IV. Pergamon Press: New York.

Patten, B.C., 1981. Environs: the superniches of ecosystems. American Zoologist 21, 845–852.

Patten, B.C., 1982. Environs: relativistic elementary particles or ecology. American Naturalist 119, 179–219.

Patten, B. C., 1983. On the quantitative dominance of indirect effects in ecosystems. In: Lauenroth, W. K., Skogerboe, G. V., Flug, M., (Eds.), Analysis of ecological systems: state-of-the-art in ecological modelling. Elsevier: Amsterdam.

Patten, B. C., 1985. Energy cycling in the ecosystem. Ecological Modelling 28, 1-71.

Patten, B.C., 1991. Network ecology: indirect determination of the life–environment relationship in ecosystems. In: Higashi, M. Burns, T. P., (Eds.), Theoretical Ecosystem Ecology: The Network Perspective. Cambridge University Press: London.

Patten, B.C., in prep. Holoecology: the unification of nature by network indirect effects. Columbia University Press: New York.

Patten, B.C., Bosserman, R.W., Finn, J.T., Cale, W.G., 1976. Propagation of cause in ecosystems. In: Patten, B.C., (Ed.), Systems analysis and simulation in ecology, Vol. IV. Academic Press: New York.

Patten, B. C., Matis, J. H., 1982. The water environs of the Okefenokee Swamp: an application of static linear environ analysis. Ecological Modelling 16 1-50.

Savenkoff, C., Vezina, A.F., Bundy, A., 2001. Inverse analysis of the structure and dynamics of the whole Newfoundland—Labrador Shelf ecosystem. Can. Tech. Rep. Fish. Aquat. Sci. 2354, 56.

Ulanowicz, R.E., 1980. An hypothesis on the development of natural communities. Journal of Theoretical Biology 85, 223–245.

Ulanowicz, R.E., 1982. NETWRK 4.2b: A Package of Computer Algorithms to Analyze Ecological Flow Networks. Copyright 1982, 1987, 1998, 1999, 2002. UMCEES Ref. No. 82–7 CBL. http://www.cbl.cees.edu/~ulan/ntwk/netwrk.txt

Ulanowicz, R.E., 1986. Growth and Development: ecosystem phenomenology. Springer-Verlag: New York.

Ulanowicz, R.E., 1997. Ecology, The Ascendent Perspective. Columbia University Press: New York.

Whipple, S.J., Patten, B.C., 1993. The problem of nontrophic processes in trophic ecology: towards a network unfolding solution. Journal of Theoretical Biology 163, 393–411.

APPENDIX A.A: A FUNCTION FOR NETWORK ENVIRON ANALYSIS EXPRESSED IN MATLAB NOTATION

```
function [ep]=NEA(data,varargin);
% y=NEA(data) performs network environ analysis on model "data".
%
% y=NEA(data,0) performs network environ analysis on model "data", but
% does not show the results in the workspace (they are still saved as
% 'NEA_ouput.mat')
%
% VERSION REV 1.0.0
% Brian D. Fath & Stuart R. Borrett 2004
% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
% I. INTRODUCTION
% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
% This program is a compilation of algorithms for network environ
% analysis.  This is a self-contained program; all functions required are
% included in this file.
%
% REQUIRED DATA INPUT
% The input variable 'data' is an (n+1 x n+2) matrix composed of an nxn
% flow matrix (F), an nx1 input vector (z), an nx1 storage vector (x),
% a 1xn output vector (y), and a 1x2 vector of zeros.  Data should reflect
% a system at steady-state (though some analyses remain valid
% for non-steadystate (i.e., structural analysis)).
%
% DATA OUTPUT
% 'ep' is a vector of environ properties and network statistics.  To
% return additional variables to the Matlab workspace, place the variable
% name into the output definition.  For example, [A,G,ep]=NEA(data) will
% return A, G and ep to the workspace. All variables are stored in
% NEA_output.mat.
%
% PROGRAM OUTLINE
% I.Introduction
% II.Initialize Parameters
% III.Main Program
%       a.  Verify Steady-State Assumption
%       b.Network Environ Analysis
%           i.Structural Analysis
%           ii.Throughflow Analysis
%           iii. Storage Analysis
%           iv.Utility Analysis
%           v.Unit Environs
%           vi.Control Analysis
%       c.Summary of Environ Properties and Network Statistics
% IV.Subfunctions
%       a.NEA_structure
%       b.NEA_throughflow
%       c.NEA_storage
%       d.NEA_utility
%       e.NEA_u_environs
%       f.  NEA_control
% V.Auxiliary Programs
```

```
%       a.              Unpack
%       b.              Environ Error Tolerance
%       c.              Bcratio
%       d.  Mode
% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
tic     % Starts program timer

% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
% II.  Initialize Parameters -=-=-=-==-=-=-=-=-=-=-=-=-=-==-=-=-=-=-=-=-=-=
% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

% control variable number of inputs (sets disp_ctrl to default)
if length(varargin)==0, disp_ctrl=1;
elseif length(varargin)==1, disp_ctrl=varargin{1};
end

global n I
[F, y, z, x]=unpack(data);  % Unpacks the data matrix into component parts
n=length(F);        % length of F gives the dimensions of the flow matrix
T=sum(F')'+z;       % total throughflow at each compartment including input
FD=F-diag(T);       % flow matrix with negative throughflows on diagonal
I=eye(n);           % nxn identity matrix

% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
% III.  Main Program  =-=-=-=-=-==-=-=-=-=-=-=-=-=-=-=-==-=-=-=-=-=-=-=-=-=-=
% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

% Check Steady-State Assumption
Tin=sum(F')+z';     % inputs
Tout=sum(F)+y;      % outputs
pd=abs((Tin-Tout))./Tin; % proportional difference in node throughflow
pd_count=length(find(pd>=0.0005));  % find number of proportional throughflow
                                    % differences that are greater than
                                    % 0.0005 or 0.05%.
if pd_count==0
   disp('Steady-state assumption met')
else
   disp('Proportoinal Difference between Tin and Tout = '), pd
   error('Model does not meet steady-state requirement');
   return         % terminates program
end

% Network Environ Analysis
% ---------------------------------------
% Structural Analysis (SUBFUNCTION_1)
[A,A1,structure_ep]=NEA_structure(F);
% Throughflow Analysis (SUBFUNCTION_2)
[G,GP,N,NP,flow_ep]=NEA_throughflow(F,y,z,T,FD);
% Storage Analysis (SUBFUNCTION_3)
[C,CP,S,SP,P,Q,PP,QP,dt,stor_ep]=NEA_storage(F,T,x,FD);
% Utility Analysis (SUBFUNCTION_4)
[D,DS,U,Y,US,YS,utility_ep]=NEA_utility(FD,T,x);
% Unit Environ Analysis (SUBFUNCTION_5)
[E,EP,SE,SEP,environ_error_tol]=NEA_u_environs(G,N,GP,NP,P,Q,PP,QP);
% Control Analysis (SUBFUNCTION_6)
[CN,CQ]=NEA_control(N,NP,Q,QP);
```

```
% Summary of Network and Environ Properties
ep=[structure_ep,flow_ep,stor_ep,utility_ep]';
ep_labels1={'# nodes, n','# links, L','connectance, L/n^2', ...
    'link density, L/n','path proliferation',...
    'TST','Cycling Index (T)',...
    'MODE_0 boundary','MODE_1, 1st pass',...
    'MODE_2, cycled','MODE_3, dissipative','MODE_4, boundary'...
    'Amp (T,output)','Amp (T,input)',...
    'I/D (T,output)','I/D (T,input)','Homog (T,output)',...
    'Homog (T,input)', 'Aggradation',...
    'Cycling Index (S)','Amp (S,output)',...
    'Amp (S,input)','I/D (S,output)','I/D (S,input)',...
    'Homog (S,output)','Homog (S,input)',...
    'Synergism (T)','Mutualism (T)',...
    'Synergism (S)','Mutualism (S)'}';   % ep labels

% Table of Environ Properties
indx=1:length(ep_labels1); eee=num2cell(indx)'; ep_labels=[eee ep_labels1];
eeee=num2cell(ep); ep_table=[eee ep_labels1 eeee];

contents={'F' 'z' 'y' 'x' 'T' 'A' 'A1' 'G',...
        'GP' 'N' 'NP' 'CN' 'C' 'P' 'Q' 'S' 'CP' 'PP' 'QP' 'SP' 'CQ',...
        'D' 'DS' 'U' 'Y' 'US' 'YS' 'E' 'EP' 'SE' 'SEP',...
        'ep' 'ep_table' 'contents'};

save NEA_output F z y x T A A1 G GP N NP CN C CP S SP P Q dt PP QP CQ D DS U Y US YS E EP SE SEP ep
ep_table contents

% ----------------------------------------------------------------------
%                       DISPLAY CONTROL
% ----------------------------------------------------------------------
% This section allows you to turn on and off the display of various results
% in the workspace.

switch disp_ctrl
    case 1          % insert the parameter you want to see into this list
        disp('Original System Data')
        F,z,y,x
        disp('Structural Analysis')
        A,A1
        disp('Throughflow Analysis')
        T,G,GP,N,NP
        disp('Storage Analysis')
        C,CP,S,SP,P,Q,dt,PP,QP
        disp('Control Analysis')
        disp('Throughflow'),CN
        disp('Storage'),CQ
        disp('Utility Analysis')
        D,DS,U,Y,US,YS
        disp('Unit Environs')
        disp('numerical error tolerance'),environ_error_tol
        disp('Unit output flow environs'),E
        disp('Unit input flow environs'),EP
        disp('Unit output storage environs'),SE
        disp('Unit input storage environs'),SEP
```

```
        disp('Environ Properties')
        ep_table
    case 0
end
prog_time=toc               % time elapsed during program

% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
% IV.   SUBFUNCTIONS  =-=-=-=-=-==-=-=-=-=-=-=-=-=-=-==-=-=-=-=-=-=-=-=-=-=
% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

% ---------------------------------------------------------------------
% SUBFUNCTION_1: Structural Analysis (PRIMARY)
% ---------------------------------------------------------------------

function [A,A1,structure_ep]=NEA_structure(F)
% y=netstructure(data,struct_plots)
% This subfunction calculates several statistics that describe the network
% structure of the system.
% ***************************************************************
global n I
A=sign(F);          % nxn adjacency matrix
A1=A+I;     % nxn adjacency walk matrix
L=nnz(A);           % number of links or arcs in the network
C=L/(n^2);          % network connectance
Ln=L/n;             % link density
max_eig=max(abs(eig(A))); % dominant eigenvalue of A = rate of
       % pathway proliferation.  This can serve as a complexity index
structure_ep=[n,L,C,Ln,max_eig];   % return variable

% ---------------------------------------------------------------------
% SUBFUNCTION_2: Throughflow Analysis
% ---------------------------------------------------------------------
function [G,GP,N,NP,flow_ep]=NEA_throughflow(F,y,z,T,FD);
% [G,GP,N,NP,flow_ep]=NEA_throughflow(F,y,z,T,FD)
% This subfunction perfoms the input and output oriented throughflow
% normalized environ analysis
% ***************************************************************
global n I
% Direct throughflow
G=I+FD*inv(diag(T));    % fij/Tj for i,j=1:n -- output matrix
GP=I+inv(diag(T))*FD;% fij/Ti for i,j=1:n -- input matrix
% Integral throughflow
N=inv(I-G);    % integral output flow matrix -- I+G+G^2+G^3+...
NP=inv(I-GP);     % integral input flow matrix -- I+GP+GP^2+GP^3+...
dN=diag(N);

[MODE_0,MODE_1,MODE_2,MODE_3,MODE_4]=mode(N,z); % mode analysis
```

```
% Throughflow environ properties
% *******************************************************************
p=ones(n,1);                                    % ones vector
TSTc=sum(((dN-p)./dN).*T);                      % cycled (mode 2) throughflow
TST=sum(T);                                     % total system throughflow
CIF=TSTc/TST;                                   % cycling index (modified from Finn 1976)
Z=sum(z);                                       % total input

% Amplification parameter
NAF=length(find((N-diag(dN))>1));               % output
NAFP=length(find((NP-diag(diag(NP)))>1));       % input

% Indirect effects parameter
IDF=sum(sum(N-I-G))/sum(G(:));                  % indirect to direct ratio (output)
IDFP=sum(sum(NP-I-GP))/sum(GP(:));              % indirect to direct ratio (input)

% Homogenization parameter
CVG=std(G(:))/mean(G(:));                       % coefficient of variation for G
CVN=std(N(:))/mean(N(:));                       % coefficient of variation for N
HF=[CVG/CVN];                                   % homogenization parameter (output)

CVGP=std(GP(:))/mean(GP(:));                    % coefficient of variation for G
CVNP=std(NP(:))/mean(NP(:));                    % coefficient of variation for N
HFP=[CVG/CVN];                                  % homogenization parameter (input)

% Network Aggradation or Average Path Length
AGG=TST/Z;      % Jorgensen, Patten and Straskraba (2000)
                % Original formulation of average path length (Finn 1976)
                % This parameter is expected to increase as systems develop.

flow_ep=[TST,CIF, MODE_0,MODE_1,MODE_2,MODE_3,MODE_4,...
    NAF,NAFP,IDF,IDFP,HF,HFP,AGG];

% ---------------------------------------------------------------------
% SUBFUNCTION_3: Storage Analysis
% ---------------------------------------------------------------------
function [C,CP,S,SP,P,Q,PP,QP,dt,stor_ep]=NEA_storage(F,T,x,FD) % Storage Analysis
% [P,Q,PP,QP,dt,stor_ep]=NEA_storage(F,T,x,FD)
% This subfunction perfoms the input and output oriented storage
% normalized environ analysis
% *******************************************************************
global I n
% Direct storage matrices
C=FD*inv(diag(x));                              % fij/xj for i,j=1:n -- output matrix
CP=inv(diag(x))*FD;                             % fij/xi for i,j=1:n -- input matrix
dt=-1/floor(min(diag(C)));                      % smallest whole number to make diag(C) nonnegative
P=I+C*dt;                                        % non-dimensional direct output storage matrix
PP=I+CP*dt;                                     % non-dimensional direct input storage matrix
% Integral storage matrices
S=-inv(C);                                      % dimensionalized integral output community matrix
SP=-inv(CP);                                    % dimensionalized integral input community matrix
Q=inv(I-P);                                     % integral output storage matrix -- I+P+P^2+P^3+...
QP=inv(I-PP);                                   % integral input storage matrix -- I+PP+PP^2+PP^3+...
dQ=diag(Q);                                     % diag of integral output storage matrix (=diag(QP))

% Storage environ properties
```

```
% *****************************************************************
p=ones(n,1);                                  % ones vector
TSTcs=sum(((dQ-p)./dQ).*T);                   % cycled (mode 2) throughflow
TSTs=sum(T);                                  % total system throughflow
CIS=TSTcs/TSTs;                               % cycling index (storage)

% Amplification parameter
NAS=length(find((Q-diag(diag(Q)))>1));
NASP=length(find((QP-diag(diag(QP)))>1));

% Indirect effects parameter
IDS=sum(sum(Q-I-P))/sum(P(:));                % indirect to direct ratio (output matrix)
IDSP=sum(sum(QP-I-PP))/sum(PP(:));            % indirect to direct ratio (input matrix)

% Homogenization parameter
CVP=std(P(:))/mean(P(:));                     % Coefficient of variation for G
CVQ=std(Q(:))/mean(Q(:));                     % Coefficient of variation for N
HS=[CVP/CVQ];                                 % homogenization parameter (output storage)

CVPP=std(PP(:))/mean(PP(:));                  % Coefficient of variation for GP
CVQP=std(QP(:))/mean(QP(:));                  % Coefficient of variation for NP
HSP=[CVPP/CVQP];                              % homogenization parameter (input storage)
stor_ep=[CIS,NAS,NASP,IDS,IDSP,HS,HSP];

% -----------------------------------------------------------------------
% SUBFUNCTION_4: Utility Analysis
% -----------------------------------------------------------------------
function [D,DS,U,Y,US,YS,utility_ep]=NEA_utility(FD,T,x)
% ****************************************************************
global I n
% Direct Utility, Throughflow ---------------------------------------
D=inv(diag(T))*(FD-FD');                      % (fij-fji)/Ti for i,j=1:n, (GP-G') -- utility matrix
e=eig(D);                                     % convergence test
if abs(max(e))>=1                             % check for convergence
   disp('WARNING: Throughflow Utility matrix does not converge');
   U=-9999;                                   %  flag if no convergence
   Y=-9999;                                   %  flag if no convergence
   NSF=-9999; PNF=-9999;                      %  flag if no convergence
else
   % Integral Utility, Throughflow
                     U=inv(I-D);              % Nondimensional integral flow utility
                     Y=diag(T)*U;             % Dimensional integral flow utility

   % Throughflow Utility Indices
   NSF=bcratio(Y);                            % flow benefit cost ratio (calls other function) (Synergism)
                     B=[1 1;1 -1];            % coefficient matrix
                     Z=[n^2;sum(sum(sign(U)))];   % vector with total n and addition of all entries
                     X=B\Z;                   % solve for number of positive and negative signs
                     PNF=X(1,1)/X(2,1);       % ratio of positive to negative signs (mutualism)
end
```

```
% Direct Utility, Storage -------------------------------------------
DS=inv(diag(x))*(FD-FD');                       % (fij-fji)/xi for i,j=1:n, (CP-C') -- utility matrix
e=eig(DS);
if abs(max(e))>=1                               % check for convergence
   disp('WARNING: Storage Utility matrix does not converge');
   % Integral Utility, Storage
                  US=-9999;                        % flag if no convergence
                  YS=-9999;                        % flag if no convergence
   NSS=-9999; PNS=-9999;% flag if no convergence
else
   % Integral Utility, Storage
                  US=inv(I-DS);                 % Nondimensional integral storage utility
                  YS=diag(T)*US;                % Dimensional integral storage utility

   % Storage Utility Indices
   NSS=bcratio(YS);                             % storage benefit cost ratio (calls other function)
   B=[1 1;1 -1];                                % coefficient matrix
                  Z=[n^2;sum(sum(sign(US)))]; % vector with total n and addition of all entries
                  X=B\Z;                        % solve for number of positive and negative signs
                  PNS=X(1,1)/X(2,1);            % storage ratio of positive to negative signs
end

utility_ep=[NSF,PNF,NSS,PNS];

% -----------------------------------------------------------------------
% SUBFUNCTION_5: Unit Environs
% -----------------------------------------------------------------------
function [E,EP,SE,SEP,environ_error_tol]=NEA_u_environs(G,N,GP,NP,P,Q,PP,QP)
% This subfunction calculates the unit environs (input, output,
% throughflow, and storage) for the given system.  Noticeable numerical
% error is usually apparent in the resultant matricies. Here, I use the
% subfunction "environ_error" removes an arbitrary amount of error by
% setting values less than "environ_error_tol" to 0.  A more appropriate way
% might be to round the values to a particular decimal place.

% IMPORTANT:  Check the error tolerance level to make sure it is
% appropriate
% ***********************************************************************

global I n
E=zeros(n+1,n+1,n); EP=zeros(n+1,n+1,n);
SE=zeros(n+1,n+1,n); SEP=zeros(n+1,n+1,n);
                    % these statements dimensionalize E, EP, SE, and SEP as 3-D
                    % variables.

environ_error_tol=1e-10;
% The value of this is arbitrary. Other ways to set this variable are possible.

% Throughflow unit environs ----------------
for i=1:n
                  E(1:n,1:n,i)=G*diag(N(:,i));
                  E(1:n,1:n,i)=E(1:n,1:n,i)-diag(N(:,i));
                  E(n+1,1:n,i)=sum(-E(1:n,1:n,i));
   E(1:n,n+1,i)=sum(-E(1:n,1:n,i)')';           % unit output flow environs
   E=environ_error(E,environ_error_tol); % AUX1
end
```

```
for i=1:n
                    EP(1:n,1:n,i)=diag(NP(i,:))*GP;
                    EP(1:n,1:n,i)=EP(1:n,1:n,i)-diag(NP(i,:));
                    EP(1:n,n+1,i)=sum(-EP(1:n,1:n,i)')';
   EP(n+1,1:n,i)=sum(-EP(1:n,1:n,i));                    % unit input flow environs
   EP=environ_error(EP,environ_error_tol); % AUX1
end

% Storage unit environs                                  -------------------
for i=1:n
                    SE(1:n,1:n,i)=P*diag(Q(:,i));
                    SE(1:n,1:n,i)=SE(1:n,1:n,i)-diag(Q(:,i));
                    SE(n+1,1:n,i)=sum(-SE(1:n,1:n,i));
   SE(1:n,n+1,i)=sum(-SE(1:n,1:n,i)')';                  % unit output storage environs
   SE=environ_error(SE,environ_error_tol); % AUX1
end
for i=1:n
                    SEP(1:n,1:n,i)=diag(QP(i,:))*PP;
                    SEP(1:n,1:n,i)=SEP(1:n,1:n,i)-diag(QP(i,:));
                    SEP(1:n,n+1,i)=sum(-SEP(1:n,1:n,i)')';
                    SEP(n+1,1:n,i)=sum(-SEP(1:n,1:n,i));                          % unit input storage environs
   SEP=environ_error(SEP,environ_error_tol); % AUX1
end

% ------------------------------------------------
% SUBFUNCTION_6: Control Analysis
% ------------------------------------------------
function [CN,CQ,CN_diff, CQ_diff]=NEA_control(N,NP,Q,QP)
% This subfunciton calculates the ratio control or dominance matrix.
%***********************************
global I n
warning off MATLAB:divideByZero %temporarily turn off divide by 0 warning
% Throughflow
CN_temp=N./NP';
j=find(CN_temp<1 & CN_temp>=0);
CN=zeros(n);
CN(j)=1-CN_temp(j);

% Storage
CQ_temp=Q./QP';
i=find(CQ_temp<1 & CQ_temp>=0);
CQ=zeros(n);
CQ(j)=1-CQ_temp(j);

warning on MATLAB:divideByZero %turn on divide by 0 warning
```

```
% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
% =-=-=-=-=-=-=            AUXILIARY PROGRAMS         =-=-=-=-=-=-=-=-=
% =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

% ----------------------------------------------------------------------
% AUX1: Unpack (AUXILIARY PROGRAM)
% ----------------------------------------------------------------------
function [F, y, z, x]=unpack(DATA)
% function [F, y, z, x]=unpack2(DATA), where data= n+1 x n+2 matrix when
% n=number of network nodes function that unpacks data from a condensed
% format.  data matrix contains F,z,x,y
% data unpack
% Stuart R. Borrett | 2002
% ************************

n=length(DATA)-2;
F=DATA(1:n,1:n);
y=DATA(n+1,1:n);
z=DATA(1:n,n+1);
x=DATA(1:n,n+2);

% ----------------------------------------------------------------------
% AUX2: Environ Error (AUXILIARY PROGRAM)
% ----------------------------------------------------------------------
function ret=environ_error(E,tolerance)
% ret=environ_error2(E), where E is a 3-D environ matrix
% This program removes some numerical error by replacing very small values
% (under the error tolerance) with a 0.  The suggested error level is 1e-10,
% although there is no formal reason for choosing this level.  Further
% analysis is needed to determine the most appropriate level.
% ************************************************

et=tolerance;          % error tolerance level
[m,n,o]=size(E);
L=m*n*o;
for i=1:L
   if E(i)>0&E(i)<et
      E(i)=0;
   end
   if E(i)<0&E(i)>(-1*et)
      E(i)=0;
   end
end
ret=E;

% ----------------------------------------------------------------------
% AUX3: Bcratio (AUXILIARY PROGRAM)
% ----------------------------------------------------------------------
function r=bcratio(Y);
%  This calculates the ratio of sum of positive to sum of
%  negative interactions in the system.
%  The B matrix sets up a pair of linear equations where
%  pos+neg=sum(sum(abs(Y))) and pos-neg=sum(sum(Y))
%  This set of equations is solved, X, and a ratio is taken.
%  The next line zeros out the diagonal elements
%  Y=Y-diag(diag(Y));
```

```
% ********************************************************

plus=sum(sum(abs(Y)));
minus=sum(sum(Y));
B=[1 1;1 -1];
Z=[plus;minus];
X=B\Z;
r=X(1,1)/X(2,1);

% ------------------------------------------------------------------------
% AUX4: MODE (AUXILIARY PROGRAM)
% ------------------------------------------------------------------------
function [T0,T1,T2,T3,T4]=mode(N,z);
% This function partitions flow into five different modes.  Mode 0 is
% the boundary input -- flow that reaches a compartment from across the
% system boundary.  Mode 1 is internal first passage flow -- total internal
% flow from compartment j to compartment i for the first time along all
% available pathways (including cycles that do not touch i).  Mode 2 is
% cycled flow -- total contribution that returns to a compartment after its
% initial visit.  Modes 3 and 4 are dissipative equivalents to Modes 1
% and 0, respectively.
% ***************************************************
global I n
mode0=diag(I*z);
mode1=inv(diag(diag(N)))*N*diag(z)-diag(I*z);
mode2=(diag(diag(N))-I)*inv(diag(diag(N)))*N*diag(z);
TSC = sum((diag(diag(N))-I)*inv(diag(diag(N)))*N*diag(z));
T0=sum(sum(mode0));
T1=sum(sum(mode1));
T2=sum(sum(mode2));
T3=T1;
T4=T0;
T=T0+T1+T2;
```

APPENDIX A.B: RESULTS FROM NEA.M ANALYSIS OF OYSTER REEF MODEL

```
Steady-state assumption met.
```

ORIGINAL SYSTEM DATA

```
F =
         0         0         0         0         0         0
   15.7915         0         0    4.2403    1.9076    0.3262
         0    8.1721         0         0         0         0
         0    7.2745    1.2060         0         0         0
         0    0.6431    1.2060    0.6609         0         0
    0.5135         0         0         0    0.1721         0
z =
   41.4697
         0
         0
         0
         0
         0
y =
   25.1646    6.1759    5.7600    3.5794    0.4303    0.3594
x =
  1.0e+003 *
    2.0000
    1.0000
    0.0024
    0.0241
    0.0163
    0.0692
```

STRUCTURAL ANALYSIS

```
A =
     0     0     0     0     0     0
     1     0     0     1     1     1
     0     1     0     0     0     0
     0     1     1     0     0     0
     0     1     1     1     0     0
     1     0     0     0     1     0
A1 =
     1     0     0     0     0     0
     1     1     0     1     1     1
     0     1     1     0     0     0
     0     1     1     1     0     0
     0     1     1     1     1     0
     1     0     0     0     1     1
```

THROUGHFLOW ANALYSIS

```
T =
   41.4697
   22.2656
    8.1721
    8.4805
    2.5100
    0.6856
G =
         0         0         0         0         0         0
    0.3808         0         0    0.5000    0.7600    0.4758
         0    0.3670         0         0         0         0
         0    0.3267    0.1476         0         0         0
         0    0.0289    0.1476    0.0779         0         0
    0.0124         0         0         0    0.0686         0
GP =
         0         0         0         0         0         0
    0.7092         0         0    0.1904    0.0857    0.0147
         0    1.0000         0         0         0         0
         0    0.8578    0.1422         0         0         0
         0    0.2562    0.4805    0.2633         0         0
    0.7490         0         0         0    0.2510         0
N =
    1.0000         0         0         0         0         0
    0.5369    1.3885    0.2775    0.7800    1.1006    0.6606
    0.1971    0.5096    1.1019    0.2863    0.4039    0.2425
    0.2045    0.5288    0.2533    1.2971    0.4192    0.2516
    0.0605    0.1565    0.1904    0.1659    1.1241    0.0745
    0.0165    0.0107    0.0131    0.0114    0.0771    1.0051
NP =
    1.0000         0         0         0         0         0
    1.0000    1.3885    0.1019    0.2971    0.1241    0.0203
    1.0000    1.3885    1.1019    0.2971    0.1241    0.0203
    1.0000    1.3885    0.2441    1.2971    0.1241    0.0203
    1.0000    1.3885    0.6198    0.5604    1.1241    0.0203
    1.0000    0.3485    0.1556    0.1407    0.2822    1.0051
```

STORAGE ANALYSIS

```
C =
   -0.0207         0         0         0         0         0
    0.0079   -0.0223         0    0.1758    0.1172    0.0047
         0    0.0082   -3.3880         0         0         0
         0    0.0073    0.5000   -0.3516         0         0
         0    0.0006    0.5000    0.0274   -0.1542         0
    0.0003         0         0         0    0.0106   -0.0099
CP =
   -0.0207         0         0         0         0         0
    0.0158   -0.0223         0    0.0042    0.0019    0.0003
         0    3.3880   -3.3880         0         0         0
         0    0.3016    0.0500   -0.3516         0         0
         0    0.0395    0.0741    0.0406   -0.1542         0
    0.0074         0         0         0    0.0025   -0.0099
S =
   48.2280         0         0         0         0         0
   24.1140   62.3604   12.4643   35.0326   49.4283   29.6703
    0.0582    0.1504    0.3252    0.0845    0.1192    0.0716
    0.5817    1.5042    0.7204    3.6893    1.1923    0.7157
    0.3924    1.0149    1.2342    1.0754    7.2881    0.4829
    1.6696    1.0838    1.3181    1.1485    7.7833  101.5031
SP =
   48.2280         0         0         0         0         0
   48.2280   62.3604    0.0301    0.8450    0.8044    2.0543
   48.2280   62.3604    0.3252    0.8450    0.8044    2.0543
   48.2280   62.3604    0.0720    3.6893    0.8044    2.0543
   48.2280   62.3604    0.1829    1.5939    7.2881    2.0543
   48.2280   15.6538    0.0459    0.4001    1.8295  101.5031
P =
    0.9948         0         0         0         0         0
    0.0020    0.9944         0    0.0439    0.0293    0.0012
         0    0.0020    0.1530         0         0         0
         0    0.0018    0.1250    0.9121         0         0
         0    0.0002    0.1250    0.0068    0.9614         0
    0.0001         0         0         0    0.0026    0.9975
Q =
  192.9119         0         0         0         0         0
   96.4560  249.4416   49.8573  140.1304  197.7131  118.6812
    0.2327    0.6017    1.3009    0.3380    0.4769    0.2863
    2.3266    6.0168    2.8816   14.7572    4.7690    2.8627
    1.5697    4.0594    4.9370    4.3016   29.1522    1.9314
    6.6783    4.3353    5.2725    4.5939   31.1334  406.0125
dt =
    0.2500
PP =
    0.9948         0         0         0         0         0
    0.0039    0.9944         0    0.0011    0.0005    0.0001
         0    0.8470    0.1530         0         0         0
         0    0.0754    0.0125    0.9121         0         0
         0    0.0099    0.0185    0.0102    0.9614         0
    0.0019         0         0         0    0.0006    0.9975
```

```
QP =
  192.9119         0         0         0         0         0
  192.9119  249.4416    0.1203    3.3801    3.2176    8.2171
  192.9119  249.4416    1.3009    3.3801    3.2176    8.2171
  192.9119  249.4416    0.2882   14.7572    3.2176    8.2171
  192.9119  249.4416    0.7317    6.3758   29.1522    8.2171
  192.9119   62.6151    0.1837    1.6005    7.3178  406.0125
```

CONTROL ANALYSIS

THROUGHFLOW

CN =

```
         0    1.0000    1.0000    1.0000    1.0000    1.0000
         0    0.0000    0.8001    0.4382    0.2074         0
         0         0         0         0    0.3483         0
         0         0    0.1475         0    0.2520         0
         0         0         0         0         0    0.7361
         0    0.4724    0.3584    0.4409         0         0
```

STORAGE

```
CQ =
         0    1.0000    1.0000    1.0000    1.0000    1.0000
         0   -0.0000    0.8001    0.4382    0.2074         0
         0         0         0         0    0.3483         0
         0         0    0.1475         0    0.2520         0
         0         0         0         0         0    0.7361
         0    0.4724    0.3584    0.4409         0         0
```

UTILITY ANALYSIS

```
D =
         0   -0.3808         0         0         0   -0.0124
    0.7092         0   -0.3670   -0.1363    0.0568    0.0147
         0    1.0000         0   -0.1476   -0.1476         0
         0    0.3578    0.1422         0   -0.0779         0
         0   -0.5038    0.4805    0.2633         0   -0.0686
    0.7490   -0.4758         0         0    0.2510         0
DS =
         0   -0.0079         0         0         0   -0.0003
    0.0158         0   -0.0082   -0.0030    0.0013    0.0003
         0    3.3880         0   -0.5000   -0.5000         0
         0    0.1258    0.0500         0   -0.0274         0
         0   -0.0777    0.0741    0.0406         0   -0.0106
    0.0074   -0.0047         0         0    0.0025         0
U =
    0.8332   -0.2228    0.0706    0.0128   -0.0270   -0.0117
    0.4244    0.5994   -0.1938   -0.0359    0.0652   -0.0009
    0.3936    0.5472    0.7414   -0.2000   -0.0609    0.0073
    0.2077    0.2871    0.0021    0.9457   -0.0563    0.0055
    0.0010    0.0664    0.4369    0.1663    0.9109   -0.0615
    0.4224   -0.4354    0.2548    0.0684    0.1774    0.9762
Y =
   34.5510   -9.2415    2.9291    0.5320   -1.1207   -0.4864
    9.4493   13.3455   -4.3144   -0.7997    1.4517   -0.0210
    3.2164    4.4718    6.0592   -1.6347   -0.4978    0.0598
    1.7617    2.4347    0.0175    8.0204   -0.4777    0.0466
    0.0026    0.1667    1.0967    0.4174    2.2863   -0.1544
    0.2896   -0.2985    0.1747    0.0469    0.1216    0.6693
US =

    0.9999   -0.0077    0.0001   -0.0000   -0.0000   -0.0003
    0.0154    0.9741   -0.0075    0.0010    0.0050    0.0003
    0.0487    3.0821    0.9178   -0.4862   -0.4416    0.0057
    0.0043    0.2721    0.0430    0.9757   -0.0479    0.0006
    0.0025    0.1638    0.0703    0.0035    0.9649   -0.0102
    0.0073   -0.0042    0.0002    0.0000    0.0024    1.0000
YS =
   41.4646   -0.3189    0.0025   -0.0003   -0.0017   -0.0107
    0.3425   21.6883   -0.1679    0.0226    0.1108    0.0058
    0.3981   25.1876    7.5002   -3.9729   -3.6091    0.0463
    0.0365    2.3079    0.3648    8.2746   -0.4062    0.0050
    0.0063    0.4112    0.1766    0.0088    2.4219   -0.0255
    0.0050   -0.0029    0.0001    0.0000    0.0016    0.6856
```

UNIT ENVIRONS

```
numerical error tolerance
environ_error_tol =
  1.0000e-010
```

UNIT OUTPUT FLOW ENVIRONS

```
E(:,:,1) =
   -1.0000         0         0         0         0         0    1.0000
    0.3808   -0.5369         0    0.1023    0.0460    0.0079         0
         0    0.1971   -0.1971         0         0         0         0
         0    0.1754    0.0291   -0.2045         0         0         0
         0    0.0155    0.0291    0.0159   -0.0605         0         0
    0.0124         0         0         0    0.0042   -0.0165         0
    0.6068    0.1489    0.1389    0.0863    0.0104    0.0087         0
E(:,:,2) =
         0         0         0         0         0         0         0
         0   -1.3885         0    0.2644    0.1190    0.0051    1.0000
         0    0.5096   -0.5096         0         0         0         0
         0    0.4536    0.0752   -0.5288         0         0         0
         0    0.0401    0.0752    0.0412   -0.1565         0         0
         0         0         0         0    0.0107   -0.0107         0
         0    0.3851    0.3592    0.2232    0.0268    0.0056         0
E(:,:,3) =
         0         0         0         0         0         0         0
         0   -0.2775         0    0.1266    0.1447    0.0062         0
         0    0.1019   -1.1019         0         0         0    1.0000
         0    0.0907    0.1626   -0.2533         0         0         0
         0    0.0080    0.1626    0.0197   -0.1904         0         0
         0         0         0         0    0.0131   -0.0131         0
         0    0.0770    0.7766    0.1069    0.0326    0.0068         0
E(:,:,4) =
         0         0         0         0         0         0         0
         0   -0.7800         0    0.6486    0.1261    0.0054         0
         0    0.2863   -0.2863         0         0         0         0
         0    0.2548    0.0422   -1.2971         0         0    1.0000
         0    0.0225    0.0422    0.1011   -0.1659         0         0
         0         0         0         0    0.0114   -0.0114         0
         0    0.2164    0.2018    0.5475    0.0284    0.0060         0
E(:,:,5) =
         0         0         0         0         0         0         0
         0   -1.1006         0    0.2096    0.8543    0.0367         0
         0    0.4039   -0.4039         0         0         0         0
         0    0.3596    0.0596   -0.4192         0         0         0
         0    0.0318    0.0596    0.0327   -1.1241         0    1.0000
         0         0         0         0    0.0771   -0.0771         0
         0    0.3053    0.2847    0.1769    0.1927    0.0404         0



E(:,:,6) =
         0         0         0         0         0         0         0
         0   -0.6606         0    0.1258    0.0566    0.4782         0
         0    0.2425   -0.2425         0         0         0         0
         0    0.2158    0.0358   -0.2516         0         0         0
```

```
         0    0.0191    0.0358    0.0196   -0.0745         0         0
         0         0         0         0    0.0051   -1.0051    1.0000
         0    0.1832    0.1709    0.1062    0.0128    0.5269         0
```

UNIT INPUT FLOW ENVIRONS

```
EP(:,:,1) =
    -1     0     0     0     0     0     1
     0     0     0     0     0     0     0
     0     0     0     0     0     0     0
     0     0     0     0     0     0     0
     0     0     0     0     0     0     0
     0     0     0     0     0     0     0
     1     0     0     0     0     0     0
EP(:,:,2) =
   -1.0000         0         0         0         0         0    1.0000
    0.9848   -1.3885         0    0.2644    0.1190    0.0203         0
         0    0.1019   -0.1019         0         0         0         0
         0    0.2548    0.0422   -0.2971         0         0         0
         0    0.0318    0.0596    0.0327   -0.1241         0         0
    0.0152         0         0         0    0.0051   -0.0203         0
         0    1.0000         0         0         0         0         0
EP(:,:,3) =
   -1.0000         0         0         0         0         0    1.0000
    0.9848   -1.3885         0    0.2644    0.1190    0.0203         0
         0    1.1019   -1.1019         0         0         0         0
         0    0.2548    0.0422   -0.2971         0         0         0
         0    0.0318    0.0596    0.0327   -0.1241         0         0
    0.0152         0         0         0    0.0051   -0.0203         0
         0         0    1.0000         0         0         0         0
EP(:,:,4) =
   -1.0000         0         0         0         0         0    1.0000
    0.9848   -1.3885         0    0.2644    0.1190    0.0203         0
         0    0.2441   -0.2441         0         0         0         0
         0    1.1126    0.1845   -1.2971         0         0         0
         0    0.0318    0.0596    0.0327   -0.1241         0         0
    0.0152         0         0         0    0.0051   -0.0203         0
         0         0         0    1.0000         0         0         0
EP(:,:,5) =
   -1.0000         0         0         0         0         0    1.0000
    0.9848   -1.3885         0    0.2644    0.1190    0.0203         0
         0    0.6198   -0.6198         0         0         0         0
         0    0.4807    0.0797   -0.5604         0         0         0
         0    0.2880    0.5401    0.2960   -1.1241         0         0
    0.0152         0         0         0    0.0051   -0.0203         0
         0         0         0         0    1.0000         0         0
EP(:,:,6) =
   -1.0000         0         0         0         0         0    1.0000
    0.2472   -0.3485         0    0.0664    0.0299    0.0051         0
         0    0.1556   -0.1556         0         0         0         0
         0    0.1207    0.0200   -0.1407         0         0         0
         0    0.0723    0.1356    0.0743   -0.2822         0         0
    0.7528         0         0         0    0.2523   -1.0051         0
         0         0         0         0         0    1.0000         0
```

UNIT OUTPUT STORAGE ENVIRONS

```
SE(:,:,1) =
```

```
   -1.0000         0         0         0         0         0    1.0000
    0.3808   -0.5369         0    0.1023    0.0460    0.0079         0
         0    0.1971   -0.1971         0         0         0         0
         0    0.1754    0.0291   -0.2045         0         0         0
         0    0.0155    0.0291    0.0159   -0.0605         0         0
    0.0124         0         0         0    0.0042   -0.0165         0
    0.6068    0.1489    0.1389    0.0863    0.0104    0.0087         0
SE(:,:,2) =
         0         0         0         0         0         0         0
         0   -1.3885         0    0.2644    0.1190    0.0051    1.0000
         0    0.5096   -0.5096         0         0         0         0
         0    0.4536    0.0752   -0.5288         0         0         0
         0    0.0401    0.0752    0.0412   -0.1565         0         0
         0         0         0         0    0.0107   -0.0107         0
         0    0.3851    0.3592    0.2232    0.0268    0.0056         0
SE(:,:,3) =
         0         0         0         0         0         0         0
         0   -0.2775         0    0.1266    0.1447    0.0062         0
         0    0.1019   -1.1019         0         0         0    1.0000
         0    0.0907    0.1626   -0.2533         0         0         0
         0    0.0080    0.1626    0.0197   -0.1904         0         0
         0         0         0         0    0.0131   -0.0131         0
         0    0.0770    0.7766    0.1069    0.0326    0.0068         0
SE(:,:,4) =
         0         0         0         0         0         0         0
         0   -0.7800         0    0.6486    0.1261    0.0054         0
         0    0.2863   -0.2863         0         0         0         0
         0    0.2548    0.0422   -1.2971         0         0    1.0000
         0    0.0225    0.0422    0.1011   -0.1659         0         0
         0         0         0         0    0.0114   -0.0114         0
         0    0.2164    0.2018    0.5475    0.0284    0.0060         0
SE(:,:,5) =
         0         0         0         0         0         0         0
         0   -1.1006         0    0.2096    0.8543    0.0367         0
         0    0.4039   -0.4039         0         0         0         0
         0    0.3596    0.0596   -0.4192         0         0         0
         0    0.0318    0.0596    0.0327   -1.1241         0    1.0000
         0         0         0         0    0.0771   -0.0771         0
         0    0.3053    0.2847    0.1769    0.1927    0.0404         0
SE(:,:,6) =
         0         0         0         0         0         0         0
         0   -0.6606         0    0.1258    0.0566    0.4782         0
         0    0.2425   -0.2425         0         0         0         0
         0    0.2158    0.0358   -0.2516         0         0         0
         0    0.0191    0.0358    0.0196   -0.0745         0         0
         0         0         0         0    0.0051   -1.0051    1.0000
         0    0.1832    0.1709    0.1062    0.0128    0.5269         0
```

UNIT INPUT STORAGE ENVIRONS

```
SEP(:,:,1) =
    -1     0     0     0     0     0     1
     0     0     0     0     0     0     0
     0     0     0     0     0     0     0
     0     0     0     0     0     0     0
     0     0     0     0     0     0     0
     0     0     0     0     0     0     0
     1     0     0     0     0     0     0
SEP(:,:,2) =
   -1.0000         0         0         0         0         0    1.0000
    0.9848   -1.3885         0    0.2644    0.1190    0.0203         0
         0    0.1019   -0.1019         0         0         0         0
         0    0.2548    0.0422   -0.2971         0         0         0
         0    0.0318    0.0596    0.0327   -0.1241         0         0
    0.0152         0         0         0    0.0051   -0.0203         0
         0    1.0000         0         0         0         0         0
SEP(:,:,3) =
   -1.0000         0         0         0         0         0    1.0000
    0.9848   -1.3885         0    0.2644    0.1190    0.0203         0
         0    1.1019   -1.1019         0         0         0         0
         0    0.2548    0.0422   -0.2971         0         0         0
         0    0.0318    0.0596    0.0327   -0.1241         0         0
    0.0152         0         0         0    0.0051   -0.0203         0
         0         0    1.0000         0         0         0         0
SEP(:,:,4) =
   -1.0000         0         0         0         0         0    1.0000
    0.9848   -1.3885         0    0.2644    0.1190    0.0203         0
         0    0.2441   -0.2441         0         0         0         0
         0    1.1126    0.1845   -1.2971         0         0         0
         0    0.0318    0.0596    0.0327   -0.1241         0         0

    0.0152         0         0         0    0.0051   -0.0203         0
         0         0         0    1.0000         0         0         0
SEP(:,:,5) =
   -1.0000         0         0         0         0         0    1.0000
    0.9848   -1.3885         0    0.2644    0.1190    0.0203         0
         0    0.6198   -0.6198         0         0         0         0
         0    0.4807    0.0797   -0.5604         0         0         0
         0    0.2880    0.5401    0.2960   -1.1241         0         0
    0.0152         0         0         0    0.0051   -0.0203         0
         0         0         0         0    1.0000         0         0
SEP(:,:,6) =
   -1.0000         0         0         0         0         0    1.0000
    0.2472   -0.3485         0    0.0664    0.0299    0.0051         0
         0    0.1556   -0.1556         0         0         0         0
         0    0.1207    0.0200   -0.1407         0         0         0
         0    0.0723    0.1356    0.0743   -0.2822         0         0
    0.7528         0         0         0    0.2523   -1.0051         0
         0         0         0         0         0    1.0000         0
```

ENVIRON PROPERTIES

```
ep_table =
    [ 1]    '# nodes, n'              [       6]
    [ 2]    '# links, L'              [      12]
    [ 3]    'connectance, L/n^2'      [  0.3333]
    [ 4]    'link density, L/n'       [       2]
    [ 5]    'path proliferation'      [  2.1479]
    [ 6]    'TST'                     [ 83.5835]
    [ 7]    'Cycling Index (T)'       [  0.1102]
    [ 8]    'MODE_0 boundary'         [ 41.4697]
    [ 9]    'MODE_1, 1st pass'        [ 32.9056]
    [10]    'MODE_2, cycled'          [  9.2082]
    [11]    'MODE_3, dissipative'     [ 32.9056]
    [12]    'MODE_4, boundary'        [ 41.4697]
    [13]    'Amp (T,output)'          [       1]
    [14]    'Amp (T,input)'           [       4]
    [15]    'I/D (T,output)'          [  1.5341]
    [16]    'I/D (T,input)'           [  1.7166]
    [17]    'Homog (T,output)'        [  1.8916]
    [18]    'Homog (T,input)'         [  1.8916]
    [19]    'Aggradation'             [  2.0155]
    [20]    'Cycling Index (S)'       [  0.9133]
    [21]    'Amp (S,output)'          [      20]
    [22]    'Amp (S,input)'           [      21]
    [23]    'I/D (S,output)'          [294.1484]
    [24]    'I/D (S,input)'           [454.2225]
    [25]    'Homog (S,output)'        [  1.1160]
    [26]    'Homog (S,input)'         [  1.4645]
    [27]    'Synergism (T)'           [  4.9152]
    [28]    'Mutualism (T)'           [  2.2727]
    [29]    'Synergism (S)'           [ 13.0899]
    [30]    'Mutualism (S)'           [  2.6000]
```

APPENDIX A.C: GLOSSARY OF NETWORK ENVIRON ANALYSIS NOTATION

GENERAL

n = number of compartments or nodes
I = $n \times n$ identity matrix
A = $n \times n$ adjacency matrix oriented from columns to rows
L = number of links or arcs in **A** (number of direct connections)
F = $n \times n$ steady-state flow matrix oriented from columns to rows (M $L^{-2 \text{ or } -3}$ T^{-1})
z = $n \times 1$ steady-state input vector (M $L^{-2 \text{ or } -3}$ T^{-1})
y = $1 \times n$ steady-state output vector (M $L^{-2 \text{ or } -3}$ T^{-1})
x = $n \times 1$ steady-state input vector (M $L^{-2 \text{ or } -3}$)

$$\Delta = \begin{bmatrix} F_{n\times n} & z_{n\times 1} & x_{n\times 1} \\ y_{1\times n} & 0 & 0 \end{bmatrix}_{n+1\times n+2}$$ = composite variable of system data

T = $\sum_{i}^{n} f_{ij} + z_i = \sum_{j}^{n} f_{ij} + y_j$ = $n \times 1$ vector of steady-state node throughflow (M $L^{-2 \text{ or } -3}$ T^{-1})

THROUGHFLOW

G = $n \times n$ donor-throughflow normalized output-oriented direct flow intensity matrix
G' = **GP** = $n \times n$ recipient-throughflow normalized input-oriented direct flow intensity matrix
N = $n \times n$ output-oriented integral flow intensity matrix
N' = **NP** = $n \times n$ input-oriented integral flow intensity matrix

STORAGE

C = $n \times n$ donor-storage normalized output-oriented direct flow intensity matrix (T^{-1})
C' = **CP** = $n \times n$ recipient-storage normalized input-oriented direct flow intensity matrix (T^{-1})
dt = discrete time interval
P = $n \times n$ non-dimensional storage-normalized output-oriented direct flow matrix
P' = **PP**= $n \times n$ non-dimensional storage-normalized input-oriented direct flow matrix
Q = $n \times n$ output-oriented integral flow intensity matrix
Q' = **QP** = $n \times n$ input-oriented integral flow intensity matrix
S = Q*dt = $-(\mathrm{C})^{-1}$ = $n \times n$ dimensional donor-storage normalized output-oriented integral flow intensity matrix (T^{-1})
S' = **SP** = Q'*dt = $-(\mathrm{C}')^{-1}$ = $n \times n$ dimensional recipient-storage normalized output-oriented integral flow intensity matrix (T^{-1})

UTILITY

D = $n \times n$ throughflow-normalized utility matrix
DS = $n \times n$ storage-normalized utility matrix
U = $n \times n$ non-dimensional integral flow utility
US = $n \times n$ non-dimensional integral storage utility
Y = $n \times n$ integral flow utility scaled by original throughflow (M $L^{-2 \text{ or } -3}$ T^{-1})
YS = $n \times n$ integral storage utility scaled by original throughflow (M $L^{-2 \text{ or } -3}$ T^{-1})

UNIT ENVIRONS

E = $n+1 \times n+1 \times n$ unit output throughflow environs

EP = E' = $n+1 \times n+1 \times n$ unit input throughflow environs

SE = $n+1 \times n+1 \times n$ unit output storage environs

SEP = SE' = $n+1 \times n+1 \times n$ unit input storage environs

Table A.1 Network parameters and environ properties returned by MATLAB® function in ep31×1 vector.

#	abbreviation	short description
[1]	'# nodes, n'	number of nodes or compartments
[2]	'# links, L'	number of direct flows or arcs
[3]	'connectance, L/n^2'	connectance
[4]	'link density, L/n'	link density
[5]	'path proliferation'	$\lambda_1(A)$ = rate of pathway proliferation (dominant eigenvalue of A)
[6]	'TST'	total system throughflow
[7]	'Cycling Index (T)'	cycling index for throughflow = TST_c/TST
[8]	'MODE_0 boundary'	boundary input
[9]	'MODE_1, 1st pass'	first–passage flow
[10]	'MODE_2, cycled'	cycled flow
[11]	'MODE_3, dissipative'	last passage dissipative flow
[12]	'MODE_4, boundary'	boundary output
[13]	'Amp (T, output)'	network amplification (throughflow, output)
[14]	'Amp (T, input)'	network amplification (throughflow, input)
[15]	'I/D (T, output)'	indirect-to-direct effects ratio or network nonlocality (throughflow, output)
[16]	'I/D (T, input)'	indirect-to-direct effects ratio or network nonlocality (throughflow, input)
[17]	'Homog (T,output)'	network homogenization (throughflow, output)
[18]	'Homog (T,input)'	network homogenization (throughflow, input)
[19]	'Aggradation'	network aggradation = TST/∑z = average path length
[20]	'Cycling Index (S)'	cycling index calculated for storage analysis
[21]	'Amp (S, output)'	network amplification (storage, output)
[22]	'Amp (S, input)'	network amplification (storage, input)
[23]	'I/D (S, output)'	indirect–to–direct ratio or network nonlocality (storage, output)
[24]	'I/D (S, input)'	indirect–to–direct ratio or network nonlocality (storage, input)
[25]	'Homog (S, output)'	network homogenization (storage, output)
[26]	'Homog (S, input)'	network homogenization (storage, input)
[27]	'Synergism (T)'	benefit–cost ratio or network synergism (throughflow)
[28]	'Mutualism (T)'	positive to negative interaction ratio or network mutualism (throughflow)
[29]	'Synergism (S)'	benefit–cost ratio or network synergism (storage)
[30]	'Mutualism (S)'	positive to negative interaction ratio or network mutualism (storage)

APPENDIX A.D: A GLOSSARY OF PRIMARY MATLAB NOTAION USED IN APPENDIX A.A

% – starts a comment line.

1:6 – creates the sequence [1,2,3,4,5,6], while 1:2:6 creates the sequence [1,3,5]

X=[a b] – creates a 1×2 vector with elements a and b; X=[a;b] – creates a 2×1 vector with elements a and b.

X' – transposes matrix X

sum(A) – returns the column sums of A; sum(A(1:6,:)) sums the first six rows of all the columns; sum(A(:)) adds all the elements of A together.

A^2 – squares the matrix A; A.^2 – squares the elements of A

length(A) – returns the largest dimension of A

diag(A) – returns a vector of the principle diagonal of A if A is a square matrix: If A is a 1×n vector, it places the elements on the principle diagonal of a square matrix where all entries other than the diagonal are zero.

mean(A); std(A); max(A) – these commands return the mean, standard deviation, and maximum value of the columns of A.

find(A) – returns the array address of the non-zero elements of A

eig(A) – returns a vector of the eigenvalues of A

inv(A) – returns the inverse matrix of A when it exists

ones(m,n) – returns an m×n vector filled with ones.

zeros(m,n) – returns an m×n vector filled with zeros.

eye(n) – creates an n×n matrix where elements on the principle diagonal are 1 and all others are 0.