Environmental Modelling & Software 101 (2018) 117–127

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

# Uncertainty analyses for Ecological Network Analysis enable stronger inferences

David E. Hines [a, *], Santanu Ray [b], Stuart R. Borrett [a, c]

[a] Department of Biology and Marine Biology, University of North Carolina Wilmington, Wilmington, 28403 NC, USA
[b] Ecological Modelling Laboratory, Department of Zoology, Visva-Bharati University, Santiniketan 731 235, India
[c] Duke Network Analysis Center, Social Science Research Institute, Duke University, Durham, 27708 NC, USA

ARTICLE INFO

*Article history:*
Received 3 June 2017
Received in revised form
3 November 2017
Accepted 17 December 2017

*Keywords:*
Ecological Network Analysis
Linear Inverse Modeling
Uncertainty analysis
Network ecology
Estuaries
Food web

ABSTRACT

Uncertainty analyses show how variability inherent in model parameters affects model outcomes. While conducting uncertainty analyses is considered best practice, technical and conceptual challenges limit applications for network models. This work adapts Linear Inverse Modeling (LIM) techniques to conduct uncertainty analysis on ecosystem flow networks, which represent the movement of energy-matter through ecosystems. We present a new R function for the enaR package to perform the analysis and use two case studies of previously published networks to demonstrate the power of this approach. The first case study examines a system with available flow uncertainty data to show how LIM uncertainty analysis can support stronger statistical inference. The second case study examines a system without available uncertainty data to illustrate how these techniques can determine the relative strength of model conclusions, even without quantitative data. The tools presented here represent an important step in the maturation of Ecological Network Analysis.



**Software and data availability section**

Name of software enaR
Developer Systems Ecology and Ecoinformatics Laboratory
Contact Address Department of Biology and Marine Biology, University of North Carolina Wilmington, 601 S. College Rd., Wilmington, NC 28303
Telephone 1.910.962.2411
E-mail enaR.maintainer@gmail.com
Year first available 2012
Hardware required PC, Mac, Linux
Software required R
Availability Public, open-source, freely available from www.r-project.org
LanguageR

## 1. Introduction

Researchers, analysts, and managers use models to test hypotheses (Hilborn and Mangel, 1997; Jørgensen and Bendoricchio, 2001; Quinn and Bates, 2011; Jones and Lennon, 2014), inform recommendations (Costanza and Ruth, 1998; Miller, 2012), and draw inferences about ecological relationships from data (Johnson and Omland, 2004; Lafferty et al., 2015). While a variety of model types can be used to evaluate the state and function of systems (Weisberg, 2012), the application of network models to accomplish this goal has rapidly increased over the past several decades (Borrett et al., 2014). Ecosystem networks have been used to investigate a variety of topics ranging from the effects of predators in trophic cascades (Wallach et al., 2017), to identifying species interactions at multiple spatial scales (Ovaskainen et al., 2015), to evaluating the overall sustainability of ecosystems (Ulanowicz et al., 2009).

One common class of ecosystem network model, flow networks, characterizes budgets for the movement of energy-matter through ecosystems using nodes that represent resource pools and edges that represent the transfer of energy-matter between resource pools. These models can describe a diverse range of systems and interactions including food webs (Niquil et al., 1999; Dunne et al., 2013), biogeochemical cycles (Christian and Thomas, 2003; Borrett et al., 2016), and systems containing mostly non-living components such as urban metabolism networks (Samaniego and

* Corresponding author.
*E-mail address:* deh9951@gmail.com (D.E. Hines).

https://doi.org/10.1016/j.envsoft.2017.12.011
1364-8152/

Moses, 2008; Zhang, 2013; Zhang et al., 2016). For example, a food web network typically consists of nodes that are species or groups of species, and edges that represent the transfer of matter through physical consumption (Pascual and Dunne, 2005; Borrett et al., 2016). Ecosystem flow networks let researchers conduct Ecological Network Analysis (ENA), which quantifies and tracks the organization and movement of energy-matter in a system of interest (Hannon, 1973; Patten et al., 1976; Ulanowicz, 1986). These models and analyses can be used to evaluate specific components and interactions within a system, as well as the state of whole ecosystems.

ENA evaluation of network models can reveal hidden relationships that result from compound and indirect interactions (Bondavalli and Ulanowicz, 1999; Borrett et al., 2010; Jordán and Scheuring, 2002; Schückel et al., 2015). For example, Christian and Luczkovich (1999) used a trophic network of the relationships among species in St. Marks Bay, Florida to calculate the effective trophic levels of the ecosystem components, considering all of the interactions in the system. In a separate example, Schramski et al. (2006) used ENA of a nitrogen (N) cycling network model to quantify the role each N pool played in regulating the movement of N through an estuary. These types of analyses grant researchers insight into some of the complex interactions that occur within ecosystems and can facilitate monitoring of ecosystem indicators (Coll and Steenbeek, 2017), but their usefulness is dependent on the accuracy and precision of the parameters used to build the networks.

Although understanding how imprecisions and uncertainties in parameterization affect network models is essential for appropriately interpreting ENA results, procedures to directly address this question are underdeveloped and often overlooked in the assessment of ecosystem flow networks (Ulanowicz, 2004; Dame and Christian, 2006; Fath et al., 2007). Uncertainty analyses, which quantify how the combined error in all model parameters propagates through model calculations to generate uncertainty in outputs (Crosetto and Tarantola, 2001), can be useful tools for this task, but can be difficult to apply to network flow models. The scarcity of uncertainty analyses in network ecology literature may be a result of the phenomenological approach that is often applied to network model construction (Ulanowicz, 1992, 2012), as opposed to the more mechanistic approaches of other model types such as building ordinary differential equations. Ecosystem flow networks are often parameterized by synthesizing multiple experimentally observed flow and biomass estimations together, leading to ambiguity about the way that the errors inherent in these measurements interact to affect the final model outputs.

Despite these difficulties, forms of uncertainty analyses have been successfully applied to network flow models. For example, Dame and Christian (2006) advocate varying network inputs and structure to identify uncertainty in results, but point out that adequate methodologies for thorough uncertainty analyses are lacking. Borrett and Osidele (2007) used Monte Carlo simulations to evaluate the robustness of network properties in 122 plausible parameterizations of a phosphorous transport network for Lake Sidney Lanier, GA, USA, and Kaufman and Borrett (2010) analyzed these same 122 phosphorous transport networks to quantify the variability of 18 network analysis metrics given the model uncertainty. Others have applied uncertainty analyses at fixed levels of error to assess the robustness of ENA results (Borrett and Salas, 2010). For example, Salas and Borrett (2011) applied an uncertainty analysis that examined the range of ENA outputs after perturbing all network edges by ±5% and re-balancing the models by altering boundary output flows. Other approaches, such as the sensitivity-based analysis used by Ayers and Scharler (2011) to evaluate uncertainty in network models of the KwaZulu-Natal Bight, South Africa, or the perturbation approach used by Mukherjee et al. (2015) to evaluate network indicators under three different scenarios representing biomass changes, have also been applied to address uncertainty in ecosystem flow networks.

Linear inverse modeling (LIM) has emerged as a useful tool for evaluating uncertainty in network models and constructing plausible network parameterizations (Vézina and Platt, 1988; Vézina and Pace, 1994; Kones et al., 2009). For example, Taffi et al. (2015) used LIM to fill data gaps in a food web of the Adriatic Sea before conducting network analyses. Recently, Guesnet et al. (2015) released software for MATLAB® that uses LIM along with Latin Hypercube and Monte Carlo sampling techniques to generate uncertainty estimations for the ENA routines included in the Ecopath with Ecosim software (Christensen et al., 2005). Although these advances represent important steps towards incorporating uncertainty analyses into ENA, these techniques remain under-recognized and underutilized. Furthermore, there are several approaches to model construction for ENA including Ecopath with Ecosim (Christensen et al., 2005), LIM (Vézina and Platt, 1988), and phenomenological parameterization (Ulanowicz, 1986), and the best modeling approach for each problem may vary based on data availability and system type.

As the use of network flow models continues to expand, standardizing approaches for uncertainty analysis in these models is necessary to ensure that ENA results are interpreted consistently and appropriately. Further, recent work has called for increased applications of ENA to inform ecosystem assessment and management (de Jonge et al., 2012; Longo et al., 2015), and quantifying the uncertainty in network flow model results is essential if these applications are to be useful. In this work, we present a generalized methodology and software function to conduct uncertainty analysis on ecosystem flow networks and to make this analysis more accessible to researchers. We first adapt LIM techniques to introduce an uncertainty analysis step into the ENA workflow (Fig. 1) and provide a new R function to accomplish this task. We then demonstrate the power of these analyses using two case studies of previously published models that have different availability of uncertainty data to highlight the diverse applicability of our approach.

In the first case study, we introduce a methodology for uncertainty analysis that can be applied to models with specified data on the uncertainty of edge parameters. This data-guided uncertainty analysis enables hypothesis testing to detect statistically significant differences in the ENA metrics of models, given the uncertainty in network parameterization. In the second case study, we demonstrate how this uncertainty analysis can be modified to accommodate networks where uncertainty data are not readily available. We applied an increasing amount of uniform uncertainty across the network flows to quantify how much variability was required to eliminate observed differences between networks. These examples highlight the power of the LIM uncertainty analyses presented in this work to both enable stronger inferences when comparing network flow models and provide insight into the robustness of network modeling results.

## 2. Materials and methods

### 2.1. Workflow

We present a software function and modified workflow for ENA to make LIM uncertainty analysis for network models accessible to researchers (Fig. 1). This workflow introduces an uncertainty step between the initial construction of network flow models and the application of ENA. While Fath et al. (2007) outline guidelines for the construction of ecosystem flow networks, several approaches to this task are available. Each of these approaches requires substantial input data, and which method researchers use may depend on data availability. For example, if biomass measurements and