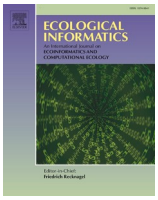
Ecological Informatics 84 (2024) 102865

Contents lists available at ScienceDirect

Ecological Informatics

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



Jump around: Selecting Markov Chain Monte Carlo parameters and diagnostics for improved food web model quality and ecosystem representation

Gemma Gerber [*,1], Ursula M. Scharler

School of Life Sciences, University of KwaZulu-Natal, Westville Campus, South Africa

ARTICLE INFO

Keywords:
Data variability
Food web
Markov Chain Monte Carlo
Model quality
Linear inverse modelling
Ecological network analysis

ABSTRACT

Capturing ecological data variability in food web models is an important step for improving model representation of empirical systems. One approach is to use linear inverse modelling and Markov Chain Monte Carlo (LIM-MCMC) techniques to set up an inverse LIM problem using empirical data constraints, and then sample multiple plausible food webs from the inverse problem using an MCMC algorithm. We describe the set of plausible food webs as an 'ensemble' of solutions to the inverse problem sampled with the LIM-MCMC algorithm. The extent of data variability eventually integrated into an ensemble depends on how well the LIM-MCMC algorithm samples the solution space. Algorithm quality can be adjusted via user-defined parameters describing starting points, jump sizes, and number of iterations or food webs produced. However, little information exists on how each LIM-MCMC algorithm parameter affects the degree of empirical data variability introduced into the ensemble. Further, post hoc algorithm quality diagnostics with commonly used trace plots and the coefficient of variation (CoV) rarely address critical aspects of algorithm quality, such as (1) if the returned ensemble successfully targeted the solution space distribution (stationarity), (2) correlation between ensemble solutions (mixing), and (3) if the ensemble contains enough solutions to adequately capture input data variability (sampling efficiency). Therefore, we used several established MCMC convergence diagnostics to (1) quantify how algorithm parameters affect ensemble flow values and if these differences propagate to ecological indicators and (2) evaluate algorithm quality and compare to current evaluation and ecosystem modelling methods. We applied 30 LIM-MCMC algorithm combinations of varying starting points, jump sizes, and number of iterations to solve food web ensembles from a single food web model. We analysed ensembles with Ecological Network Analysis (ENA) to calculate indicators describing system function. Results show that LIM-MCMC algorithm parameters, in particular the jump size, affect ensemble flow values, which propagate to ecological indicators describing different ecosystem function of the same model. Thereafter, comparisons of post hoc diagnostics show that MCMC convergence diagnostics provided more robust estimates of algorithm quality than trace plots and CoV. Together, these findings underpin several novel recommendations to enhance LIM-MCMC algorithm parameter selection and quality assessments applicable to any ecological ensemble network study.

1. Introduction

Ecosystems are incredibly variable due to dynamic environmental parameters and ecological interactions (Scharler and Borrett, 2021). As such, recent efforts have focussed on incorporating ecological data variability into food web models, thereby estimating ecosystem flow and function uncertainty (Hines et al., 2018; Kones et al., 2009a). Food web models mathematically describe spatial and temporal 'snapshots' of exchanges of energy or material (as 'flows' or 'links') between ecosystem components (Scharler and Borrett, 2021). Empirical data variability is commonly incorporated into model outputs by calculating multiple plausible food webs via linear inverse modelling and Markov Chain Monte Carlo (LIM-MCMC) techniques (Niquil et al., 2011; van Oevelen et al., 2010).

This approach uses ecological input data packed into linear equations defining the boundaries of a multi-dimensional Euclidean solution

* Corresponding author.
E-mail address: gerber@iiasa.ac.at (G. Gerber).
[1] Current address: International Institute for Applied Systems Analysis, Biodiversity, Ecology and Conservation Research Group, Laxenburg, Austria.

https://doi.org/10.1016/j.ecoinf.2024.102865
Received 18 March 2024; Received in revised form 20 October 2024; Accepted 21 October 2024
Available online 24 October 2024

space, within which exist infinite plausible food webs (Niquil et al., 2011). The popular 'mirror' MCMC algorithm (van den Meersche et al., 2009) samples the solution space by first solving a parsimonious solution ('starting solution'). From there, a second solution is drawn from a normal distribution centred on the first point with a fixed standard deviation called the jump length (for details see Niquil et al., 2011; van den Meersche et al., 2009). This process is repeated for *n* number of iterations, solving a chain of solutions (or an 'ensemble' of solutions) through the solution space. Within the ensemble, each solution is one unique plausible food web with one solved value per flow, all valid within the input data constraints (Soetaert and van Oevelen, 2009; van Oevelen et al., 2010). The ensemble is subsequently analysed with Ecological Network Analysis (ENA) (e.g., Borrett and Lau, 2014; Butts, 2008; Kones et al., 2009b) to calculate ecological indicators describing ecosystem function. Taking the uncertainty arising from an ensemble into account as opposed to focussing only on a single solution improves interpretations of empirical ecosystem function (Hines et al., 2018) and allows statistical comparison between models (e.g., Nogues et al., 2021; Tecchio et al., 2016; Zhang et al., 2022).

Ensemble quality, i.e. how well it represents the empirical system, is subject to some uncertainty depending on several factors, including the accuracy of the ecological data assumptions (see Scharler and Borrett, 2021 for details) and the sampling algorithm quality. The latter can be defined as how 'well' the algorithm samples the solution space, i.e., to return an ensemble that approximates the range of the underlying ecological input data. A 'good quality' algorithm adequately samples the solution space, returning an ensemble reflective of the flow variability as specified in the input data (i.e., a good quality ensemble). In contrast, a 'poor quality' algorithm samples a fraction of the solution space, thus returning an ensemble that only partially reflects the input data flow range (i.e., a poor-quality ensemble). Network analysis of a good quality ensemble results in ecological indicators describing a more complete range of potential ecosystem status, which is partially lost with poor quality ensembles. Given the calls for incorporating ecosystem models and ecological indicators into ecosystem management and policy (de Jonge and Schückel, 2021; Fath et al., 2019), ensuring algorithm quality that adequately captures flow variability in empirical ecosystems is important for more confident inferences of ecosystem status.

In this manuscript, we focus on two aspects of algorithm quality. The first is: what makes a good quality algorithm? Algorithm quality depends on user-defined parameters of starting points, jump size, and number of iterations (Niquil et al., 2011; van den Meersche et al., 2009; van Oevelen et al., 2010). Current broad parameter recommendations include (1) solving the starting solution in a 'central' region of the solution space (Soetaert et al., 2009), (2) choosing jump sizes based on the input flow range magnitudes (van den Meersche et al., 2009), and (3) returning numerous samples, where 3000 iterations are considered adequate (Soetaert et al., 2009), but commonly $\geq$10,000 iterations are used (e.g., Bentley et al., 2019; Olli et al., 2019; Zhang et al., 2022). However, as each food web has a unique model structure and parameters, case study specific algorithm parameters may be required for adequate solution space sampling. While the general goal is for the algorithm to adequately sample the solution space (e.g., Meddeb et al., 2019; Saint-Béat et al., 2013; Tecchio et al., 2016), there is little information on how algorithm parameter combinations of starting solution, jump size and number of iterations influences the resulting food web ensemble quality.

The second question we address is how to confidently assess algorithm quality. There are no existing a priori methods to determine which algorithm parameters result in a good quality algorithm. In practice, modellers apply post hoc diagnostics on the ensemble flows. Such assessments include trace plots, which visualise flow values over a number of iterations to assess MCMC stationarity and mixing (Raj et al., 2016; van den Meersche et al., 2009; van der Heijden et al., 2020). Another approach using the coefficient of variation (CoV) determines the 'convergence' of each flow mean and standard deviation (Bell et al., 2017; van Oevelen et al., 2011; Zhang et al., 2022), and is interpreted as a degree of 'stationarity'towards a stable value (e.g., Saint-Béat et al., 2013, 2020). However, trace plots and CoV can neglect critical aspects of algorithm quality from an MCMC convergence perspective.

MCMC convergence is the convergence of the posterior probability density function (PDF) of the solution chain *P(x)* to the 'target' PDF of the solution space *Q(x)* (Roy, 2020; van Ravenzwaaij et al., 2018). The solution space is considered well-sampled, indicating good algorithm quality, where *P(x)* approaches *Q(x)*. While true convergence cannot be achieved, it can be inferred via several MCMC convergence diagnostics (Hogg and Foreman-Mackey, 2018; Plummer et al., 2020; Roy, 2020) applied to each ensemble flow. Together, the diagnostics assess three main MCMC convergence aspects: stationarity (how well *P(x)* approaches *Q(x)*), mixing (correlation between samples), and sampling efficiency whether enough solutions have been returned to adequately describe *Q(x)*) (Roy, 2020). While trace plots can infer stationarity and mixing, they can be subjectively interpreted. Additionally, the 'convergence' described by the CoV indicates the residual uncertainty in the flow solutions, where large uncertainties have larger CoV (e.g., van Oevelen et al., 2011), which is different to the definition of MCMC convergence. Therefore, situations may arise where trace plots and CoV diagnostics report a 'good' quality algorithm, whereas a more rigorous assessment with multiple MCMC convergence diagnostics may lead to different interpretations (Du et al., 2022; Mengersen et al., 1999).

In summary, the aims of this study were three-fold. First, we examined how different LIM-MCMC 'mirror' algorithm parameters (starting points, jump size, number of iterations) affect the sampling of the solution space of a single case-study food web using 30 LIM-MCMC 'mirror' algorithm combinations of the three parameters. Secondly, we investigated whether existing algorithm quality assessments in ecological models using trace plots and CoV can be improved with MCMC convergence diagnostics by comparing the proportion of flows that 'pass' or 'fail' each diagnostic. Thirdly, we investigated whether the differences in flow values from each LIM-MCMC scenario propagates to ecological indicators describing ecosystem status.

2. Materials and methods

2.1. Model construction

As a case study, we constructed an empirical food web model of uMdloti Estuary (June 2015) following network construction guidelines (see Fath et al., 2007; Scharler and Borrett, 2021). uMdloti Estuary is a large temporarily closed estuary on the east coast of South Africa (van Niekerk et al., 2020) (Fig. 1 A). The rationale for using this case study site is (1) applying this methodology on an empirical system, rather than a 'toy' or theoretical model, provides more insight for future empirical research applications and implications, and (2) we had an available dataset from a recent study of invertebrate responses to prolonged mouth closure (Scharler et al., 2020), providing a unique opportunity to construct a detailed invertebrate food web model.

As this manuscript is not focussed on the case study site itself, but rather *if* and *how* flow and ecological indicator uncertainty is introduced by varying LIM-MCMC algorithm parameters, we do not give a full description of the food web model construction here. Rather, we provide a brief description, and detail the model construction process, data sources, and parameterisation process in Appendix A.

We used empirical datasets and published literature (Tables A1–5) to define the a priori food web model, i.e., the compartments, directed flows between them, and flows across the system boundary (boundary gains/losses). Next, we parameterised compartment biomasses and flow inequalities using empirical data (Scharler et al., 2020), taxon-specific estimates (e.g., Nozais et al., 2005; Ortega-Cisneros et al., 2016; Tagliarolo et al., 2019), allometry (e.g., Brey, 2010, 2012; Huntley and Lopez, 1992), and open-source databases (e.g., Parr et al., 2014). We standardised all parameters to a common model currency, expressing

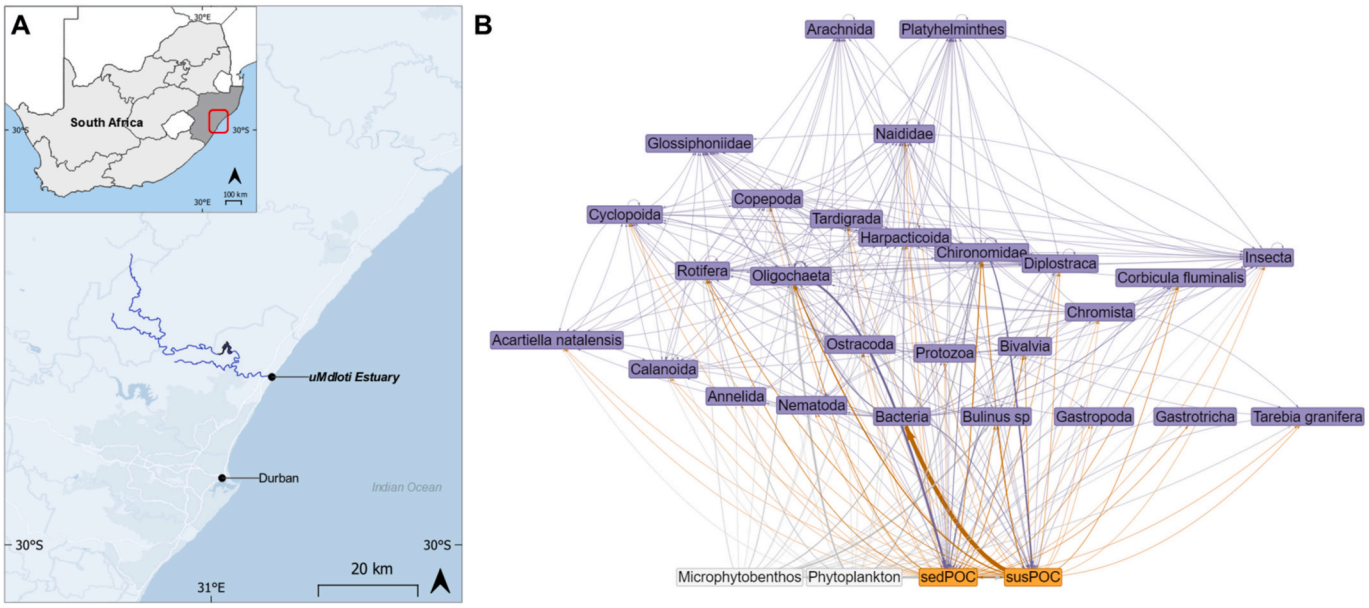


Fig. 1. uMdloti Estuary (29°39′2.1348" S, 31°7.'44.9328″ E) and associated rivers on the east coast of South Africa near Durban (eThekwini Municipality), KwaZulu-Natal (A) and uMdloti Estuary June 2015 food web, showing internal compartments and the flows between them (B). Boundary flows of gross primary production, respiration, imports, and exports are omitted for clarity.

biomass as storage dimensions (mgC m^{-2}) and flows as rates of carbon transfer per day (mgC m^{-2} d^{-1}) (e.g., Le Guen et al., 2019; Meddeb et al., 2019).

In total, the food web model consisted of two non-living detrital compartments (suspended (susPOC) and sedimented (sedPOC) particulate organic matter), two primary-producer compartments (phytoplankton and microphytobenthos), and 19 consumer compartments (Fig. 1 B). A total of 215 flows consisted of 167 internal flows (diet, egestion, mortality) and 48 boundary flows described imports, exports, gross primary production, and respiration flows for living compartments. Twenty-three mass-balance equalities and 127 inequalities constrained the flows. Lastly, we coded the model into the required 'LIM declaration file' input format (Appendix A) using function autoGen from R package **autoLIMR v.3.0.1** (Gerber et al., 2023).

2.2. Flow & ecological indicator uncertainty analysis with LIM-MCMC algorithm scenarios

2.2.1. Algorithm scenario development

We developed 30 LIM-MCMC mirror algorithm scenarios, consisting of different starting points, jump sizes, and number of iterations (Fig. 2). By default, the mirror algorithm calculates the starting point for underdetermined LIM problems using the LSEI algorithm (Haskell and Hanson, 1981; van den Meersche et al., 2009). The starting solution can be changed to the 'Central' solution, for which the mean of each flow minimisation and maximisation values are calculated by LIM::Xranges (van Oevelen et al., 2010). The difference between the LSEI and the Central starting solutions is that they solve the starting solution in different regions of the solution space. We selected jump sizes based on the recommendation that they should be within the same order of magnitude as the input data flow inequalities (van den Meersche et al., 2009). As the case study food web model flows varied by orders of magnitude, we selected jump sizes based on the approximate orders of flow magnitudes, ranging from 0.01 mgC m^{-2} d^{-1} to 100 mgC m^{-2} d^{-1}. We selected the number of iterations based on the commonly used 10,000 iterations (e.g., Bentley et al., 2019; Olli et al., 2019; Zhang et al., 2022), along with half (5000) and double (20,000) the number of iterations.

2.2.2. Calculating multiple plausible food webs

Ensembles of multiple plausible food webs, consisting of 215 flows per food web, were calculated for each algorithm scenario (Fig. 2) using function multi_net in R package **autoLIMR v3.0.1** (Gerber et al., 2023). To ensure reproducibility, we applied the base R function set.seed before

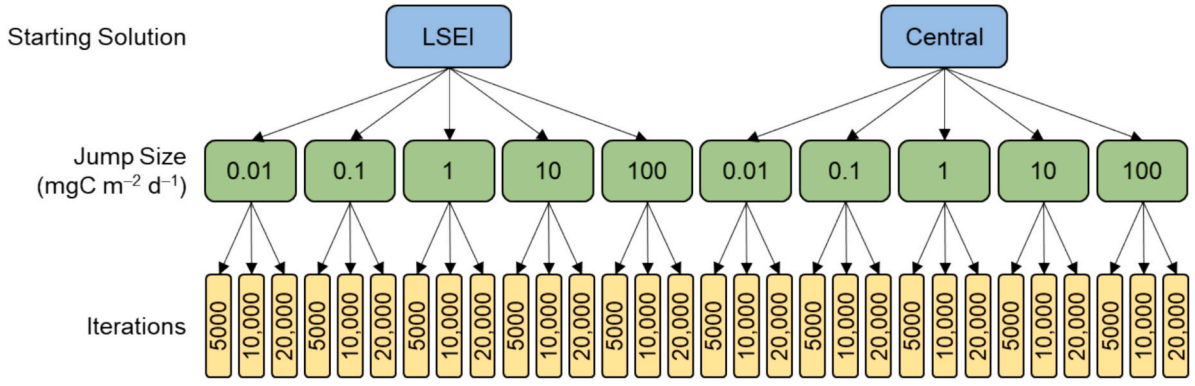


Fig. 2. Nested design of the LIM-MCMC mirror algorithm scenarios, consisting of two different starting solutions, five jump sizes per starting point, and three different numbers of iterations per jump size.

each simulation to initialise a pseudorandom number generator, ensuring that for each algorithm scenario, the random processes start and end in the same point and return the same results when the function is rerun with the same algorithm scenario.

2.2.3. Ecological network analysis

For each ensemble returned by the 30 LIM-MCMC algorithm scenarios, we applied Ecological Network Analysis (ENA) on all ensemble flows (215) with R package **enaR v3.0.0** (Lau et al., 2017). We calculated six flow-based ecological indicators (Table 1). We first calculated three ecological indicators that are often strongly linked to trophic level I (via primary production or detritivory) to describe ecosystem function. Total System Throughflow (TST_{flow}) is the sum of all compartmental inputs or outputs and is interpreted as a measure of system size or activity (Patten, 1995). More productive systems have a larger TST_{flow}, indicating positive 'health' (Patten, 1995), or could result from system stress or eutrophication (Luong et al., 2014). Finn Cycling Index (FCI) is the sum of all nodal cycling as a fraction of the TST_{flow} (Finn, 1976, 1980), quantifying how much energy is reused by the system (de la Vega et al., 2018; Safi et al., 2019). A larger FCI indicates increased recycling, interpreted as an indicator of a stressed system (Pezy et al., 2018; Scharler and Baird, 2005; Tecchio et al., 2015) or resilience to perturbations (Saint-Béat et al., 2015). The Detritivory: Herbivory ratio (D:H) is the ratio of total detritivory (consumption of detritus) to total herbivory (consumption of primary producers) by trophic level II consumers (Kay et al., 1989). In this study, D:H was calculated as the total detritus consumption divided by the total microalgae consumption. Ecosystems with a high D:H indicate that detritus is important for medium cycling, such as carbon cycling (Chrystal and Scharler, 2014; de Jonge and Schückel, 2019), and can be interpreted as more resilient to external perturbations (Lassalle et al., 2011).

Additionally, we calculated flow-based information indices to describe system organisation and development. Average Mutual Information (AMI) represents the degree of flow specialisation or efficiency by estimating the constraint exerted on a unit of energy to flow from one compartment to the next (Ulanowicz, 1986). A higher AMI indicates greater efficiency or more constrained organisation of the flows, whereas excessive efficiency indicates a 'brittle' system prone to disturbance (Ulanowicz et al., 2009). A lower AMI indicates less flow efficiency (i.e., more redundant flows), but too little efficiency can mean stagnation (Ulanowicz et al., 2009). Flow diversity (H) applies the Shannon Diversity equation (MacArthur, 1955) to individual flows to describe the diversity of interactions between ecosystem components (Ulanowicz, 1986). A higher system H indicates a high degree of flow complexity and diversity (Ulanowicz, 1986), whereas a low H indicates greater flow specialisation. AMI and H are related, such that H is the upper boundary of AMI (Ulanowicz, 1986). A single metric, called robustness, incorporates both AMI and H, capturing the opposing nature of efficiency (AMI) and diversity (H) (Ulanowicz et al., 2009). Robustness can measure system sustainability, where a system with a proportion of AMI/H just below 0.5 is considered optimal for system resilience (Ulanowicz et al., 2009).

Table 1
Summary of ecological indicator formulas and units.

Indicator	Formula	Units	Reference
Total System Throughflow (TST_{flow})	$TST_{flow} = \sum_{i=1}^{n} T_i$, where $T_i^{in} = z_i + \sum_{j=1}^{n} f_{ji}$; $T_i^{out} = y_i + \sum_{j=1}^{n} f_{ij}$ At steady state, $T_i^{in} = T_i^{out} = T_i$	mgC m^{-2} d^{-1}	(Patten, 1995)
Finn Cycling Index (FCI)	$FCI = \frac{\sum TST_{Ci}}{TST_{flow}}$ Where the cycled flow of node i is: $TST_{Ci} = \left(\frac{n_{ii}-1}{n_{ii}}\right) T_i$	None	(Finn, 1976, 1980)
Detritivory: Herbivory ratio (D:H)	$DH = \frac{\sum Detritivory}{\sum Herbivory}$	None	(Kay et al., 1989)
Average Mutual Information (AMI)	$AMI = K\sum_{ij}\left(\frac{T_{ij}}{T_{..}}\right) log\left(\frac{T_{ij}T_{..}}{T_{i.}T_{.j}}\right)$	Bits	(Ulanowicz, 1986)
Flow diversity (H)	$H = -k\sum_{ij}\left(\frac{T_{ij}}{T_{..}}\right) log\left(\frac{T_{ij}}{T_{..}}\right)$	Bits	(Ulanowicz, 1986)
Robustness	Robustness $= -a \bullet ln(a)$, where $a = AMI/H$	None	(Ulanowicz et al., 2009)

2.2.4. Statistical analyses

We compared flows, and ecological indicators, between algorithm scenarios using Cliff's Delta (δ) (Cliff, 1993; Macbeth et al., 2011), a non-parametric effect size measure often used to statistically compare multiple plausible food webs (e.g. Niquil et al., 2020) with function cliff.delta from R package **effsize v0.8.1** (Torchiano, 2020). Cliff's δ is interpreted as a measure of dominance or overlap between two group distributions (Macbeth et al., 2011). It is not associated with a probability of null hypothesis rejection but instead uses threshold values to indicate the magnitude of difference between two groups. We considered groups as significantly different where Cliff's δ magnitudes were 'medium' ($|\delta| < 0.474$), and 'large' ($|\delta| \geq 0.474$), but not for 'small' ($|\delta| < 0.33$) or 'negligible' ($|\delta| < 0.147$) magnitudes (e.g., Nogues et al., 2021; Tecchio et al., 2016). As Cliff's δ does not compare mean and standard deviations, we display results with box plots to aid interpretation.

2.3. Algorithm quality assessments

We tested algorithm quality by analysing ensemble flows from each of the 30 algorithm scenarios using (1) the two established diagnostics commonly used in ecological modelling (trace plots, CoV), and (2) a further six Markov Chain Monte Carlo (MCMC) convergence diagnostics. For this assessment, we chose to apply the diagnostics only to bounded flows (i.e., flows with defined minima and maxima). Regarding the network used here, all but 15 of the total 215 defined flows were bounded, i.e. were assigned maxima according to the function Xranges, that takes the initial ecological input information into account. For the 15 unbounded flows, the maxima were automatically set to $1e^{+30}$ (Soetaert et al., 2009; van Oevelen et al., 2010). These maxima far exceed any constraint in the input data, so we omitted these flows for reporting.

For each of the eight convergence diagnostics, we counted the number of bounded flows (200) that 'pass' or 'fail' the respective diagnostic test (Table 2). First, we assessed algorithm quality with commonly used diagnostics in ecological modelling: trace plots and CoV. We visually assessed flows with trace plots and considered that a flow was 'well-sampled', thus indicating good algorithm quality, if the trace plot looked like a 'hairy caterpillar' (Roy, 2020). After that, we assessed the algorithm quality by calculating the CoV ($CoV = SD/Mean$) on each bounded flow. Flows were considered adequately representative of the ecological input data where there was little residual uncertainty (CoV < 1) (e.g., de Jonge et al., 2020) (Table 2).

Next, we assessed algorithm quality using MCMC convergence diagnostics. We calculated such on each bounded flow using function mcmc_diags from **autoLIMR** (Gerber et al., 2023), which is a convenience function that calculates six MCMC convergence diagnostics at once, two describing stationarity, two describing mixing, and two describing sampling efficiency. For details of the diagnostic calculations for each assessment with the mcmc_diags function, we strongly

Table 2
Diagnostic tests to assess Markov Chain Monte Carlo convergence aspects: 1) stationarity, 2) mixing, and 3) sampling efficiency.

Convergence Criteria	Diagnostic	Algorithm quality Test
Stationarity	Coefficient of variation (CoV)	CoV < 1
	Geweke *Z*-score	$-1.96 > Z < 1.96$ ($p = 0.05$)
	Heidelberger-Welch diagnostic	Stationarity test = "Pass" Halfwidth test = "Pass"
Mixing	Lag-*k* autocorrelation	Lags between the 1st and 5th samples are less correlated than the 1st and 50th samples.
	Raftery-Lewis dependent factor (*I*)	$I < 5$
Sampling Efficiency	Effective sample size (*ESS*)	*ESS* > 10 % user-defined number of iterations
	Raftery-Lewis required number of samples (*N*)	*N* < user-defined number of iterations

encourage the reader to consult the relevant cited literature below. Stationarity was assessed through Geweke *Z*-scores (Geweke, 1992) and Heidelberger-Welch (HW) diagnostics (Heidelberger and Welch, 1983). The Geweke Z-scores test for equality of a Markov chain's first (10 %) and last (50 %) part, by calculating the mean of each and comparing them via a simple two-sample test of means. If the resulting Z-score is between −1.96 and 1.96 (corresponding to $p = 0.05$) (Du et al., 2022), the means of the two parts of the chain are not significantly different from one another, and thus it can be inferred that the flow has reached stationarity.

The Heidelberger-Welch (HW) diagnostic assesses stationarity in two parts: first, it tests the stationarity of the sampled chain (stationarity test), and then tests if there are enough iterations to estimate the mean of the flow values with an acceptable degree of precision (halfwidth test) (Heidelberger and Welch, 1983). The stationarity test uses the Cramer-von Mises statistic to test the null hypothesis that sampled values are derived from a normal distribution (Plummer et al., 2020). The test is first applied to the whole sample chain, then again after incrementally discarding the first 10 %, 20 %, … of the chain until either the null hypothesis is accepted (i.e., that the sampled values come from a normal distribution), or 50 % of the chain has been discarded. If the stationarity test is passed, the half-width test is applied. It calculates a 95 % confidence interval (CI) for the mean, using the part of the sample chain that passed the stationarity test. Then, half the width of the 95 % CI interval is compared with the mean of the chain. If the ratio between the half-width and the mean is lower than a specified target value (the default is 0.1), then the halfwidth test is passed. The target value can be manually changed, but the default is sufficient for most applications. For our study, we determined that flow values (i.e., sample chains) 'passed' the HW diagnostic if the flow passed both the stationarity and the halfwidth tests.

We assessed mixing with lag-*k* autocorrelation (Roy, 2020) and the Raftery-Lewis dependence factor (*I*) (Raftery and Lewis, 1992). Autocorrelation is the degree to which sampled flows in the chain are correlated, where a low lag-k autocorrelation indicates good mixing and a better-quality sampling algorithm (Plummer et al., 2020; Roy, 2020). The lag-*k* autocorrelation assessment calculates the correlation between every sample and the sample *k* steps before (see Plummer et al., 2020; Roy, 2020 for calculation details). For a well-mixed chain of samples (indicative of good mixing), autocorrelation should decrease with an increase in *k*. Theoretically, one could compare samples across any two relevant (i.e., sufficiently different) lag-*k* periods, but this study considered that flow values 'passed' the autocorrelation diagnostic if the 1st and 50th samples (i.e., lag 50, or 50 iterations apart) were less correlated than the 1st and 5th samples (i.e., lag 5, or 5 iterations apart), indicating that correlation between samples decreased with an increase in number of iterations.

The Raftery-Lewis dependence factor (*I*) assesses autocorrelation between samples, where $I > 5$ indicates strong autocorrelation, potentially due to poor choices of starting solutions, high posterior correlations, or 'stickiness' of the sampled chain (Plummer et al., 2020). For this study, flow values 'passed' the Raftery-Lewis dependence factor diagnostic if $I < 5$ (Raftery and Lewis, 1992).

We assessed sampling efficiency with the Raftery-Lewis required number of iterations (*N*) test (Raftery and Lewis, 1992) and the effective sample size (*ESS*) statistic (Robert and Casella, 2004). The *N* diagnostic is the required number of samples to meet convergence criteria, where an adequate number of samples is achieved where $N <$ number of user-defined iterations. Flow values 'passed' the *N* test if the required number of iterations (*N*) was less than the defined number of iterations in the algorithm, which indicates a good quality algorithm (Raftery and Lewis, 1992). The *ESS* is the estimated number of independent observations that the ensemble is equivalent to (given the degree of autocorrelation between samples within the ensemble). Flow values 'passed' the *ESS* test if *ESS* > 10 % of the total number of iterations, indicating that an adequate number of iterations has been returned (Gabry et al., 2019; Plummer et al., 2020).

2.3.1. Statistical analyses

For each separate diagnostic (Table 2), Fisher's Exact Test for Count Data (stats::fisher.test, R Core Team, 2022) was used to determine significant differences in the number of flows that passed and failed each diagnostic across all 30 algorithm scenarios. After that, we applied post hoc analysis with function pairwise_fisher_test (**rstatix v.0.7.1**, Kassambara, 2022) to determine if differences between the number of flows that passed and failed the diagnostic were due to starting solution, jump size, or number of iterations. We set statistical significance at a 5 % threshold ($\alpha = 0.05$), and all reported *p* values are two-tailed. While we conducted statistical analyses on count data, we report the proportion of bounded flows that passed each diagnostic (%) to aid results interpretation.

3. Results

3.1. Flow uncertainty analysis

3.1.1. Flow magnitude representation

It is impractical to report the flow values for all 215 food web flows derived from each algorithm scenario, as there are simply too many flows to derive any succinct results. Instead, we chose to report how each algorithm scenario affects select so-called bounded flows (200 flows with defined minima and maxima) of different range magnitudes ('small', 'medium', and 'large') based on the smallest, median, and largest flow ranges (Table 3). The underlying methodology for selecting bounded flows of different range magnitudes is presented in **Appendix B**.

3.1.2. Uncertainty of select flows solved with various LIM-MCMC scenarios

All 30 LIM-MCMC algorithm scenarios adequately sampled 'small' flows but showed decreasing sampling adequacy for 'medium' and 'large' flows. For 'small' flows (Platyhelminthes, Arachnida_R), flow values were not significantly different between all 30 algorithm scenarios, and their full range of potential flow values were sampled by each LIM-MCMC algorithm scenario (Fig. 3; Fig. 4).

In contrast, 'medium' flows (Gastropoda_EX, susPOC_Q_Nematoda) were well sampled by all algorithm scenarios except those of the smallest jump size (0.01 mgC m^{-2} d^{-1}), which only sampled a smaller proportion of possible flow values (Fig. 3, Fig. 4). There were no significant differences between 'medium' flow values for algorithm scenarios with jump sizes 0.1–100 mgC m^{-2} d^{-1}, nor were flow values affected by an increase in number of iterations within each jump size. One notable result is how starting solution affected 'medium' flow values at the smallest jump sizes (0.01 and 0.1 mgC m^{-2} d^{-1}). With

Table 3
Select flows representing flow magnitudes of 'large', 'medium' and 'small' in the June 2015 uMdloti food web network, and their theoretical minimum, maximum, and ranges (mgC m^{-2} d^{-1}) calculated with function Xranges (van Oevelen et al., 2010).

Flow Name	Flow Description	Minimum	Maximum	Range	Magnitude
Bacteria_R	Bacteria respiration to external CO_2 sink	7.02	1357.39	1350.37	Large
susPOC_Q_Oligochaeta	Oligochaeta consumption of susPOC	0.00	818.71	818.71	
Gastropoda_EX	Gastropoda Export across the system boundary	0.00	5.05	5.05	Medium
susPOC_Q_Nematoda	Nematoda consumption of susPOC	0.00	4.04	4.04	
Platyhelminthes_R	Platyhelminthes respiration to external CO_2 sink	0.08	0.18	0.11	Small
Arachnida_R	Arachnida respiration to external CO_2 sink	0.11	0.17	0.06	

algorithm scenarios of larger jump sizes (1–100 mgC m^{-2} d^{-1}), the starting solution bares no effect on sampled flow values (Fig. 4). In contrast, the starting solution significantly impacts the sampled flow value for algorithm scenarios with the smallest jump size (0.01 mgC m^{-2} d^{-1}). For example, at the smallest jump size, Gastropoda_EX values returned by the algorithm scenarios with the Central starting solution were significantly smaller than those of the LSEI starting solution (Fig. 4). For algorithm scenarios with jump sizes of 0.1 mgC m^{-2} d^{-1}, the choice of starting solution affected sample flow values only for scenarios with 5000 iterations, but not for 10,000 or 20,000 iterations. Thus, 'medium' flows may be adequately sampled by algorithms of jump size $\geq$1 mgC m^{-2} d^{-1}, with no differences in sampling efficiency introduced by differences in starting solutions or number of iterations.

The two 'large' flows (Bacteria_R, susPOC_Q_Oligochaeta) were better sampled by algorithm scenarios of the largest jump sizes (10 and 100 mgC m^{-2} d^{-1}) (Fig. 3, Fig. 4). The large jump size allowed the algorithm to sample the full range of potential flow values, and to negate effects of starting points and number of iterations on sampling efficiency. The largest jump size (100 mgC m^{-2} d^{-1}) allowed the algorithm to sample a greater range of potential flow values than any other jump size. Generally, for algorithm scenarios with jump sizes <100 mgC m^{-2} d^{-1}, the choice of starting solution and number of iterations significantly affected the sampled flow values. For example, Bacteria_R flow values calculated from algorithm scenarios with the Central starting solution were significantly larger than those with the LSEI starting solution (Fig. 4). Within algorithm scenarios of different starting solutions, Bacteria_R flow values were often significantly different with an increase in iterations (5000 to 20,000) (Fig. 3). Similarly, susPOC_Q_Oligochaeta flow values calculated from algorithm scenarios with the Central starting solution were significantly different than those calculated from algorithm scenarios of the LSEI starting solution at smaller jump sizes, and showed differences in flow values with increasing number of iterations.

3.2. Algorithm quality assessments

3.2.1. Algorithm assessment with diagnostics commonly used in ecological modelling

In ecological modelling, often only larger flows are visually assessed with trace plots (e.g., Nogues et al., 2021). For conciseness, we report trace plots of one large flow, 'Bacteria_R' here, and trace plots for all six selected flows (Table 3) in Figs. B4–6. Bacteria_R trace plots show that at the smallest jump size (0.01 mgC m^{-2} d^{-1}), flow values derived from the Central solution algorithm scenarios are larger than those derived from the LSEI solution algorithm scenarios (Fig. 5). At jump size 0.1 mgC m^{-2} d^{-1}, flow values derived from both starting solutions begin to converge after approximately 200 iterations. For jump sizes 1–100 mgC m^{-2} d^{-1}, flow values derived from both starting points almost immediately converge to a similar value.

For algorithm scenarios of jump sizes 0.01–10 mgC m^{-2} d^{-1}, Bacteria_R trace plots show little mixing ('stickiness'), i.e., little difference between values of each iteration due to a limited degree of polytope exploration (Fig. 5). At these jump sizes, an increase in number of iterations did not improve mixing of the flow but did result in an increased range of sampled flow values. In contrast, flow values derived from algorithm scenarios with the largest jump size (100 mgC m^{-2} d^{-1}) showed much variability indicative of good mixing (could be described as 'hairy caterpillars'). When compared to the potential range of flow values (7.02–1357.39 mgC m^{-2} d^{-1}, Table 3), algorithms of the smallest jump size (0.01 mgC m^{-2} d^{-1}) sampled only a small fraction of potential flow values (7.02–10.45 mgC m^{-2} d^{-1}, starting solutions combined). The mirror algorithm sampled larger proportions of the potential flow values with an increase in jump size. Algorithm scenarios of jump sizes 10 and 100 mgC m^{-2} d^{-1} returned similarly large flow ranges (7.02–1198.50 and 7.02–1357.39 mgC m^{-2} d^{-1}, respectively), with jump size 100 mgC m^{-2} d^{-1} fully sampling all potential flow values (7.02–1357.39 mgC m^{-2} d^{-1}) for both starting solutions and all number of iterations.

For the CoV diagnostic, there was a significant difference between the number of flows that passed and failed the diagnostic across jump sizes, number of iterations, and starting solutions (Fisher's Exact $p <$ 0.001; Table C1). Post hoc comparisons show that for algorithm scenarios with a jump size of 1 mgC m^{-2} d^{-1} or greater, 100 % of the flows passed the CoV diagnostic (Fig. 6). For Central solution algorithm scenarios, there was no statistical difference in the number of flows that passed the CoV diagnostic ($p > 0.05$). For LSEI solution algorithm scenarios, the number of flows that passed the CoV diagnostic was significantly different only for one algorithm scenario (0.01 mgC m^{-2} d^{-1} jump size, 10,000 iterations; $p < 0.05$).

3.2.2. Algorithm assessment with MCMC convergence diagnostics

For each MCMC convergence diagnostic, there was a statistically significant difference between the number of flows that passed and failed the diagnostic test across jump sizes, number of iterations, and starting solutions (Fisher's Exact, $p < 0.001$; Table C1). The number of flows that passed diagnostic tests of stationarity (Geweke and HW diagnostics), mixing (lag-*k* autocorrelation and *I*), and sampling efficiency (*ESS* and *N*) were not significantly different between algorithm scenarios of the Central and LSEI starting solutions ($p > 0.05$). The proportion of flows that passed MCMC convergence diagnostics of stationarity, mixing, and sampling efficiency significantly increased with increased jump size ($p < 0.05$; Fig. 6).

Within each jump size, the proportion of flows that passed MCMC convergence diagnostics were generally not significantly different with an increase in number of iterations ($p > 0.05$), except in a few key cases. At the smallest jump size (0.01 mgC m^{-2} d^{-1}), there was a significant decrease in the proportion of flows that passed the *I* diagnostic with an increase in number of iterations ($p < 0.05$). Conversely, at the smallest jump size there was a significant increase in the proportion of flows the passed the *N* diagnostic ($p < 0.05$).

3.3. Ecological network analysis

All ecological indicators were significantly different between algorithm scenarios (Fig. 7). Total System Throughflow (TST_{flow}) and Finn Cycling Index (FCI) significantly increased with increasing jump sizes. Within each jump size, TST_{flow} significantly increased with an increase in number of iterations, and there was usually no significant difference in FCI at increasing number of iterations. At the smallest jump sizes (0.01 and 0.1 mgC m^{-2} d^{-1}), TST_{flow} and FCI were significantly larger for ensembles solved with algorithm scenarios with the Central starting solution than LSEI (Fig. 8). In comparison, D:H significantly decreased

Fig. 3. Select food web flows (mgC·m^{-2}·d^{-1}) of large (Bacteria_R, susPOC_Q_Oligochaeta), medium (Gastropoda_EX, susPOC_Q_Nematoda), and small (Platyhelminthes_R, Arachnida_R) range magnitudes of the June 2015 uMdloti Estuary food web network, calculated from thirty LIM-MCMC algorithm scenarios of varying starting points (LSEI and Central solution), jump sizes, and number of iterations. Boxplots show flow median (line inside box), interquartile ranges (IQR), error variance (± 1.5 * IQR), and outliers (filled circles). Grey indicates the potential flow ranges calculated from the input food web inequalities (van Oevelen et al., 2010). Lowercase letters indicate statistically significant differences between number of iterations within each jump size. Uppercase letters indicate statistically significant differences between jump sizes (Cliff's $|\delta| \geq 0.33$).

with increasing jump sizes and number of iterations (Fig. 7). At smaller jump sizes, D:H values resulting from algorithm scenarios of the Central starting solution were significantly smaller than with the LSEI starting solution (Fig. 8).

Average Mutual Information (AMI) significantly decreased with increased jump size for both starting solutions (Fig. 7). Within the larger jump sizes (10 and 100 mgC m^{-2} d^{-1}), there was no difference in AMI with increased number of iterations. At smaller jump sizes (0.01 and 0.1 mgC m^{-2} d^{-1}), AMI calculated from the Central solution algorithm scenarios were significantly larger than those calculated with LSEI. Flow diversity (H) significantly increased from jump size 0.01 to 1 mgC m^{-2} d^{-1}, but then significantly decreased with increasing jump size (10–100 mgC m^{-2} d^{-1}). Within jump sizes, H values significantly decreased with increased number of iterations. Flow diversity values calculated with the

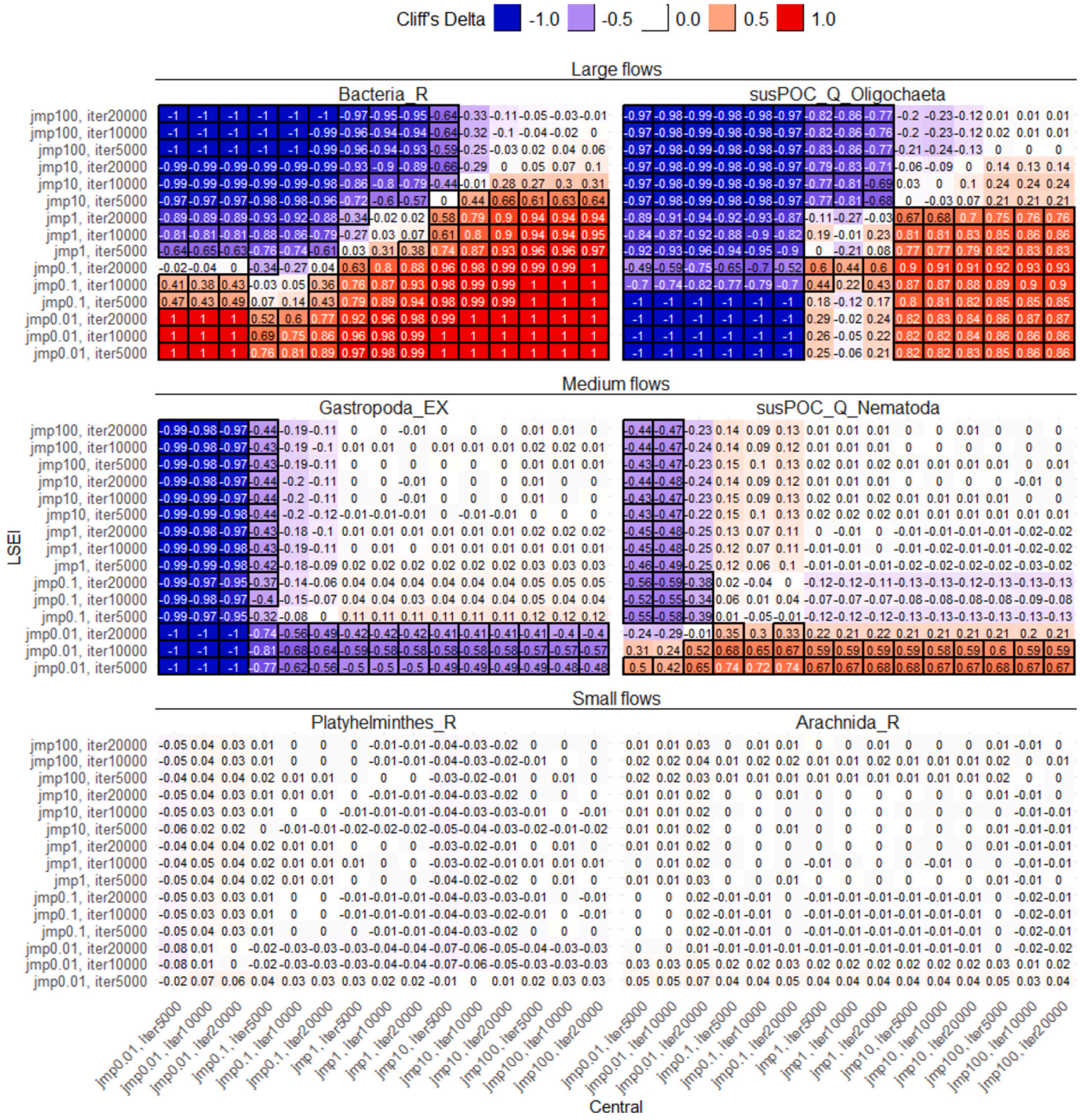


Fig. 4. of Cliff's δ pairwise comparisons of select uMdloti Estuary June 2015 food web flows (mgC·m^{-2}·d^{-1}) of large (Bacteria_R, susPOC_Q_Oligochaeta), medium (Gastropoda_EX, susPOC_Q_Nematoda), and small (Platyhelminthes_R, Arachnida_R) range magnitudes, calculated with different mirror algorithm scenarios of starting solutions central (x-axis) and LSEI (y-axis) at various jump sizes and number of iterations. Red tiles indicate a greater likelihood of larger flow values in group 1 (x-axis) than in group 2 (y-axis), and blue tiles indicate a greater likelihood of larger flow values in group 2 (y-axis) than in group 1 (x-axis). White tiles indicate a complete overlap of flow values between groups (δ = 0). Statistically significant differences ($|\delta| \geq 0.33$) are outlined in black.

Central solution algorithm scenarios were significantly smaller for jump sizes 0.01–10 than those calculated from the LSEI starting solution (Fig. 8).

Robustness calculated from Central solution algorithm scenarios showed no significant differences between the smallest (0.01 mgC m^{-2} d^{-1}) and largest (100 mgC m^{-2} d^{-1}) jump sizes (Fig. 7). However, robustness calculated from LSEI solution algorithm scenarios significantly differed between all jump sizes. For both starting point algorithm scenarios, at the largest jump size (100 mgC m^{-2} d^{-1}), there was no significant difference in robustness with an increase in number of iterations. However, at smaller jump sizes (0.01–10 mgC m^{-2} d^{-1}), flows derived from algorithm scenarios of the Central solution returned significantly larger robustness values than flows derived from algorithm scenarios of the LSEI solution (Fig. 8).

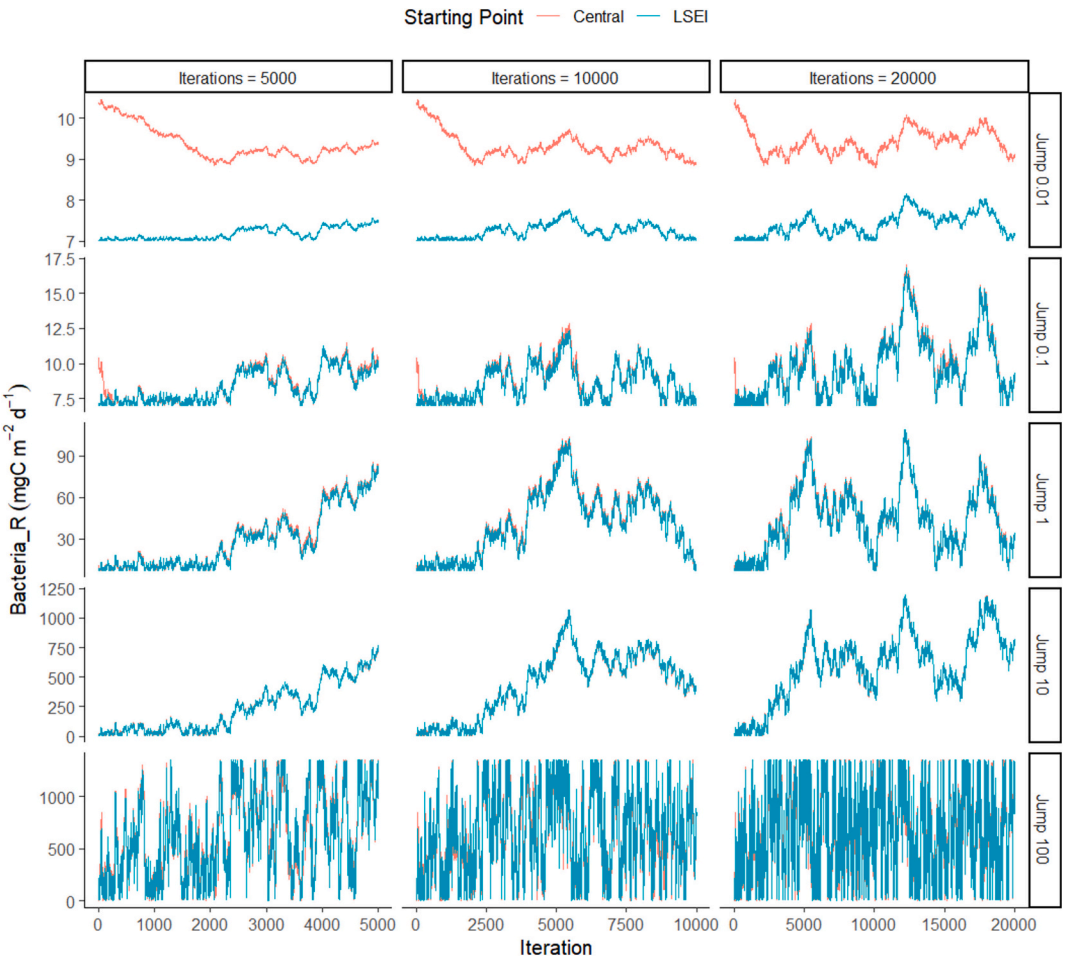


Fig. 5. Trace plots of 'Bacteria_R' flow values (mgC m^{-2} d^{-1}) at each iteration derived from thirty mirror algorithm scenarios. Each scenario consists of combinations of different starting solutions (Central and LSEI), jump sizes (0.01–100 mgC m^{-2} d^{-1}), and number of iterations (5000, 10,000, 20,000).

4. Discussion

4.1. Flow uncertainty analysis

Our results show that the mirror algorithm parameter selection (starting points, jump sizes, number of iterations) significantly affects LIM-MCMC sampled values of large and medium food web flows, but not smaller flows (Fig. 3). In particular, we found that *jump size* is a crucial parameter for algorithm quality, where larger jump sizes allowed for adequate solution space sampling to return food web ensembles that are representative of the ecological input data variability. Smaller jump sizes were too small to adequately sample the potential flow ranges. This can be conceived as the algorithm taking small 'steps' in the Euclidean space, exploring only a small area and thus sampling a small fraction of the potential flow ranges. In contrast, the small flows (Platyhelminthes_R, Arachnida_R) were well-sampled by all jump sizes because all jump sizes were large enough to fully sample the small range of potential flow values. These findings corroborate previous suggestions that high dimensional food web problems require larger jump sizes to adequately sample the solution space (Hogg and Foreman-Mackey, 2018; Nemeth and Fearnhead, 2021). Therefore, we propose an enhancement to the initial jump size recommendation (van den Meersche et al., 2009): in ecological models, an adequate jump size is the same as the magnitude of the *largest* flow inequalities.

We noted increased computational time with increased jump size (Fig. B9). Large jump sizes can lead to many proposal rejections and computational inefficiency (van den Meersche et al., 2009). A decrease in jump size allowed the ensembles to be solved quicker (Fig. B9) with the trade-off of a less adequately sampled solution space (Fig. 3, Fig. 4). As researchers are often limited by time and computational power, trade-offs between time and algorithm performance are warranted. However, even at smaller jump sizes, the computational efficiency can be improved by selecting an appropriate starting solution and number of iterations.

In addition to jump size, our results show that the starting solution, or *where* the algorithm starts sampling in the solution space, may impact the algorithm quality at smaller jump sizes. For example, at the smallest jump size (0.01 mgC m^{-1} d^{-1}), medium-sized flow Gastropoda_EX values calculated from algorithm scenarios using the Central starting solution were significantly smaller than those of the LSEI starting solution (Fig. 3). The LSEI starting solution (Haskell and Hanson, 1981) is often solved at the intersection of the inequalities (van Oevelen et al., 2010), representing the extreme ecological input data values (minima, maxima) (Stukel et al., 2012; Stukel and Landry, 2010). The Central solution returns a starting point in a more central region of the solution space (Soetaert et al., 2009), allowing the mirror algorithm to begin sampling the solution space from less extreme ecological data values. In such cases where a small jump size is required due to computational limitations, starting the algorithm in a more representative region of the solution space may benefit modellers more than using the default LSEI solution.

The choice of the number of iterations may impact algorithm quality to a lesser extent than jump size and starting solution. Theoretically, increasing the number of iterations allows the algorithm to 'explore' more of the solution space by taking more 'steps' (Roy, 2020; van Ravenzwaaij et al., 2018). We found that at larger jump sizes, an

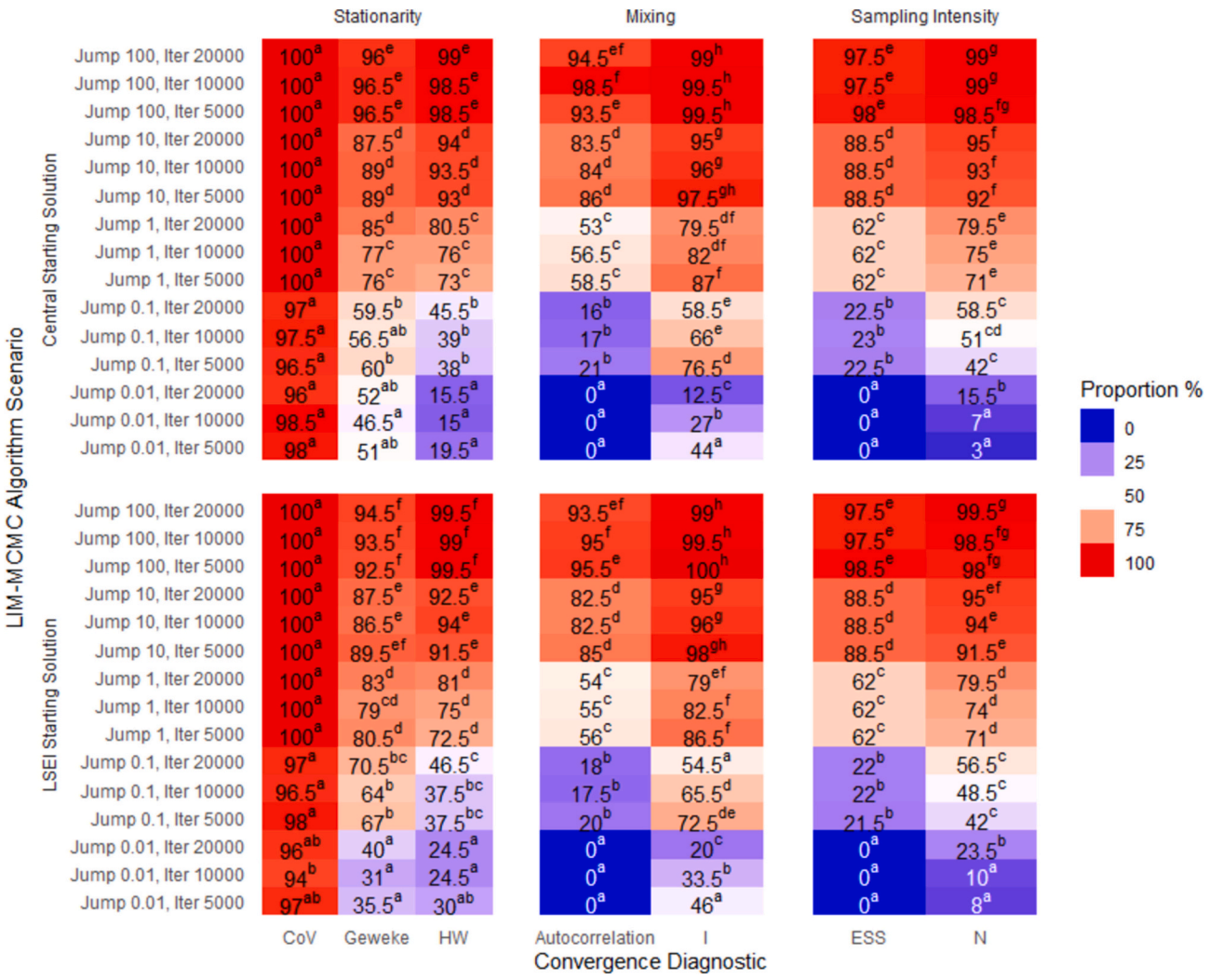


Fig. 6. Proportion (%) of bounded flows ($n = 200$) that passed the coefficient of variation (CoV) criteria (CoV <1) and various Markov Chain Monte Carlo (MCMC) convergence diagnostics. MCMC diagnostics assessed 1) stationarity with the Geweke and Heidelberger-Welch (HW) diagnostics, 2) mixing with autocorrelation and Raftery-Lewis independence factor (I), and 3) sampling intensity with effective sample size (ESS) and Raftery-Lewis required number of samples (N). Lowercase letters indicate Fisher's Exact post hoc significant differences between algorithm scenarios for each diagnostic.

increased number of iterations does not have a significant impact on flow values (Fig. 3).

With small jump sizes, we found that the degree of polytope exploration somewhat depends on the starting solution. For example, Bacteria_R flow values sampled by the algorithm with the smallest jump size (0.01 mgC m^{-1} d^{-1}) and LSEI starting solution were significantly different between 5000 and 20,000 iterations, with 20,000 iterations sampling a larger proportion of the potential flow values. In contrast, Bacteria_R flow values sampled by the same jump size, but different starting point (Central), showed no significant difference with an increase in number of iterations or range of potential flow values sampled. This may have implications for data management of large datasets, as it is more practical to handle, wrangle, and store 10,000 values per flow (e.g., Bentley et al., 2019; Tecchio et al., 2016; Zhang et al., 2022) than, for example, 200 million values per flow (e.g., Kelly et al., 2019).

This study found that the choice of starting solution and number of iterations may be insignificant if the selected jump size is large enough to sample areas further away from the initial solution within the solution space. For example, for algorithm scenarios of jump sizes 0.1–100 mgC m^{-1} d^{-1}, there were no significant differences in Gastropoda_EX flow values between the Central and LSEI starting solutions, nor between increasing number of iterations within each jump size.

Thus, in cases where smaller jump sizes are required due to computational limitations, we make three suggestions to enhance algorithm quality. First, modellers should assess a few different 'small' jump sizes to find a jump size that is computationally efficient, yet large enough to sample areas further away from the initial solution. Secondly, the Central solution should be used instead of the LSEI solution as the initial starting point to allow the algorithm to start sampling from a more 'central' region of the solution space (van den Meersche et al., 2009). This allows plausible food webs to be sampled from a more 'central' region, rather than at the intersections of the inequalities representing extreme ecological input data (e.g., Stukel et al., 2012; Stukel and Landry, 2010). Thirdly, we suggest increasing the number of iterations to improve the algorithm's solution space exploration (Roy, 2020; van Ravenzwaaij et al., 2018), thereby returning solutions more representative of the ecological input data. In all cases, we suggest applying appropriate diagnostics to assess whether the selected parameters were adequate to return ensembles of food webs representative of the ecological input data variability.

Progress towards improving computational efficiency to return ecologically representative food web ensembles includes parallel computation of multiple plausible food webs (Nemeth and Fearnhead, 2021). For example, de Jonge et al., 2020 employed a parallelisation approach to solve 100,000 plausible food webs through twenty-five parallel computations of 4000 iterations each. In the broader Markov

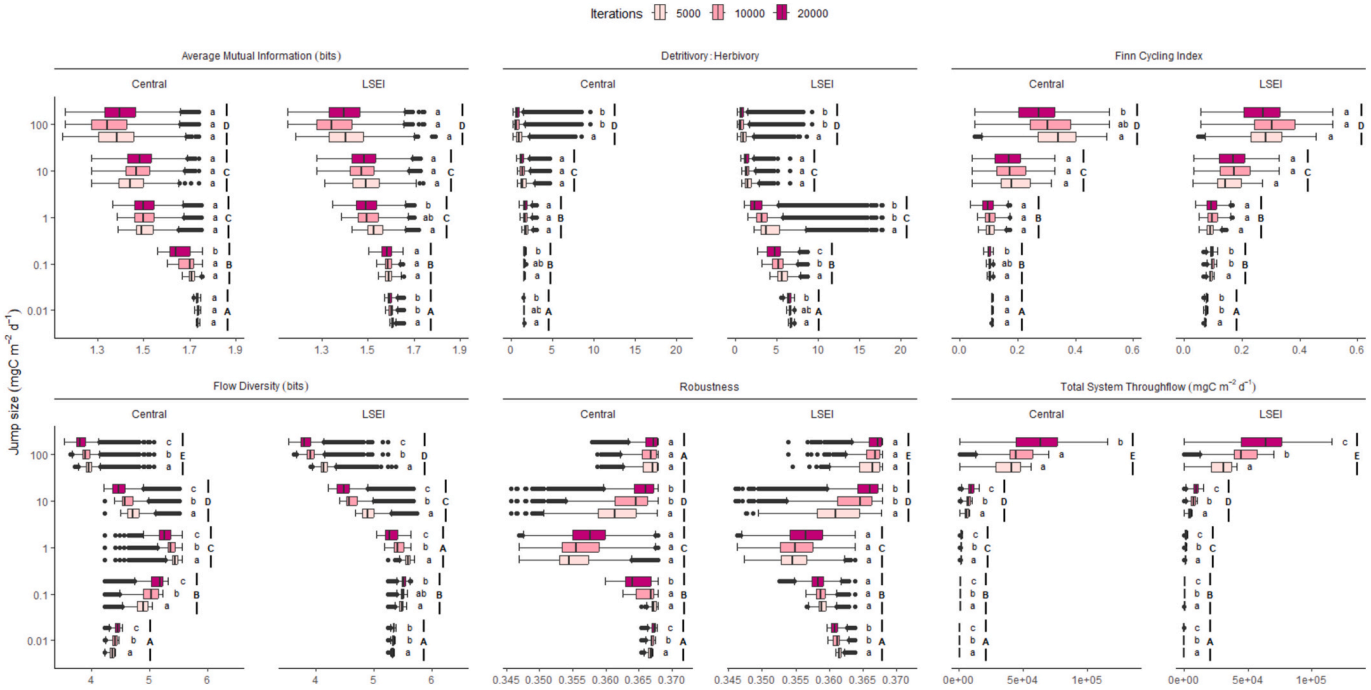


Fig. 7. Select Ecological Network Analysis (ENA) indicators describing the June 2015 uMdloti Estuary food web network derived from different mirror algorithm starting points (Central, LSEI) at increasing jump sizes and number of iterations. Lower and upper error lines represent ±1.5 * IQR (Interquartile Range). Boxplots show flow median (line inside box), interquartile ranges (IQR), error variance (± 1.5 * IQR), and outliers (filled circles). Lowercase letters indicate statistically significant differences between number of iterations within each jump size. Uppercase letters indicate statistically significant differences between jump sizes (Cliff's $|\delta| \geq 0.33$).

Chain Monte Carlo statistical method research, there are ongoing efforts towards optimising sampling efficiency, including progress towards finding more appropriate algorithm starting points (van Ravenzwaaij et al., 2018), subsampling optimisation and scaling (Hu and Wang, 2021; Nemeth and Fearnhead, 2021), and sequential proposal frameworks (Park and Atchadé, 2020). While these statistical methods have not yet made their way into food web modelling, we recommend that ecological modellers should keep in touch with these statistical techniques to potentially improve the incorporation of empirical data uncertainty in ecological modelling in future.

4.2. Algorithm quality assessments

While visual diagnostics such as trace plots are not precise estimates of convergence, they provide an interesting starting point to examine algorithm quality. In this study, we show that trace plots of a large flow (Bacteria_R) can subjectively indicate differences in algorithm quality (Fig. 5). Trace plots of Bacteria_R at the larger jump sizes show that the flow is integrating effectively over the posterior distribution and the algorithm samples the full range of potential flow values (Roy, 2020). Therefore, at the larger jump sizes, the selection of starting solution and number of iterations is irrelevant. Trace plots further show that the choice of starting solution is more important at smaller jump sizes than larger jump sizes. For example, at the smallest jump size (0.01 mgC m^{-1} d^{-1}), Bacteria_R flow values derived from algorithm scenarios of the Central starting solution are larger than those derived from algorithm scenarios of LSEI starting solution but rapidly converge at increasing jump sizes. Even though the algorithm may begin at different points in the solution space, the jump sizes are large enough to move away from the starting solution. These findings suggest that trace plots can be valuable tools for quick visual diagnostics of flow stationarity, mixing, and if the sampled flow is derived from the full input data range. Together, such findings can assist modellers with adequate algorithm parameter selection. The trace plot comparisons presented in this study (Fig. 5, Figs. B4–6) can be a helpful reference point for future ecological modelling studies aiming to use trace plots to diagnose algorithm quality.

Previously, visual assessments have been applied to only the larger flows to assess algorithm quality, assuming that if larger flows are well-sampled, the smaller flows will be too (e.g., Nogues et al., 2021). The findings of this study support this assumption. Therefore, we recommend that trace plots of the larger flows can be used as a quick indication of algorithm quality using base R functions or several R packages, including **autoLIMR** (Gerber et al., 2023), **coda** (Plummer et al., 2006, 2020), and **mcmcplots** (McKay Curtis, 2018).

In this study, algorithm quality assessment with CoV shows that almost all flows are well-sampled. The CoV has previously been interpreted as an indicator of algorithm quality by assessing residual uncertainty in the flow solutions (e.g., Bell et al., 2017; de Jonge et al., 2020; van Oevelen et al., 2011). However, little uncertainty around a flow value mean (CoV < 1) could also result from the algorithm sampling a small proportion of the solution space, neglecting other potential values. For example, flow Bacteria_R values derived from all 30 algorithm scenarios had a CoV < 1 (Fig. B10), indicating 'good quality'. Between algorithm scenarios, the CoV of Bacteria_R was smallest at algorithm scenarios of the smaller jump sizes, and larger at increased jump sizes. In this case, if a smaller CoV is indicative of 'good quality', then Bacteria_R flow values derived from algorithm scenarios of the smallest jump size would be of better quality than algorithm scenarios of larger jump sizes. In contrast, MCMC convergence diagnostics show that Bacteria_R was poorly sampled at smaller jump sizes and better sampled at the largest jump size (Fig. 6), suggesting that CoV may not be an appropriate diagnostic of algorithm quality. Instead, MCMC convergence diagnostics may provide a more robust assessment of algorithm quality.

Using MCMC convergence diagnostics, we show that most flows (92.5 %) sampled with algorithms scenarios of the largest jump size (100

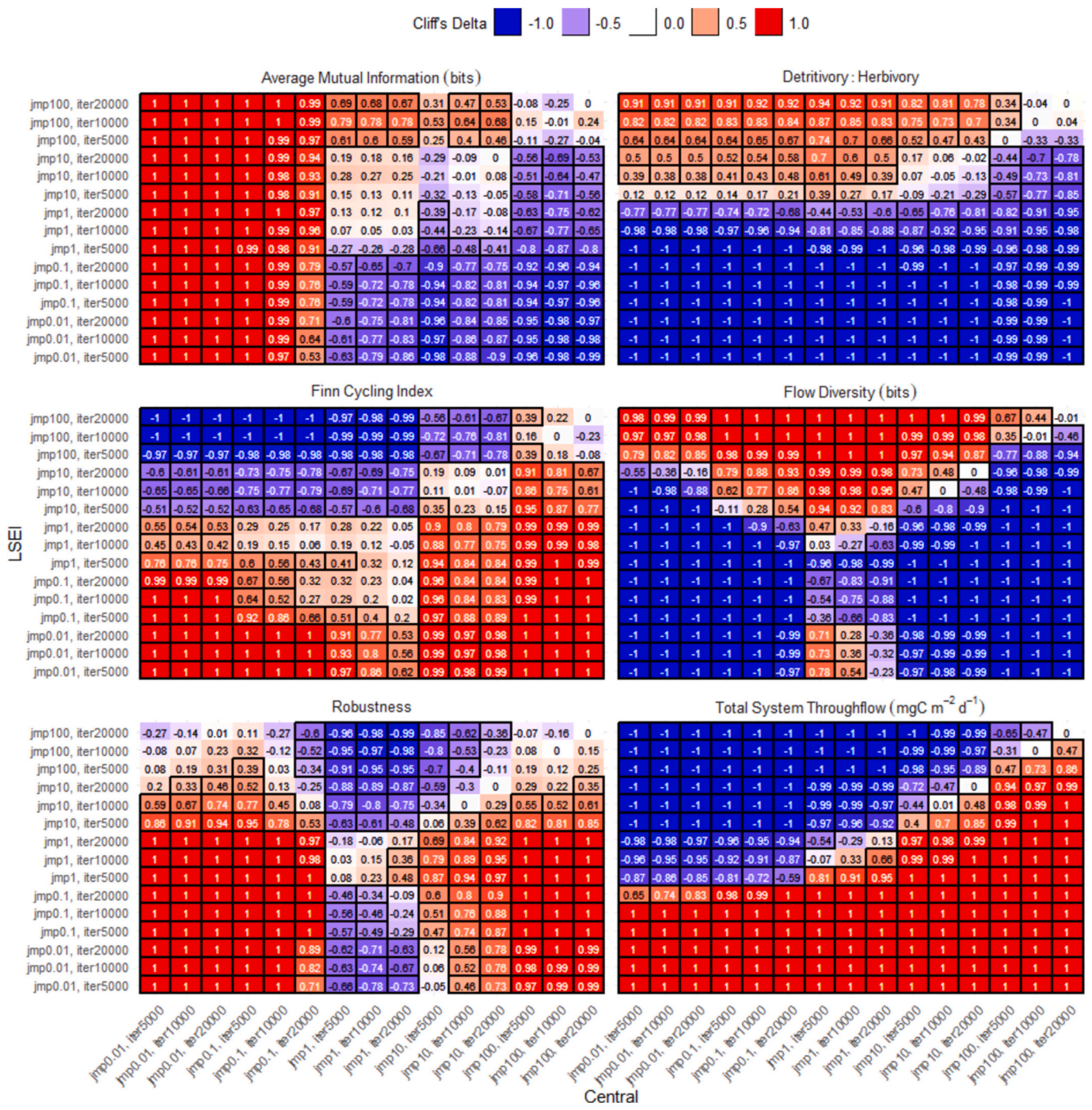


Fig. 8. Cliff's δ pairwise comparisons of select Ecological Network Analysis (ENA) indicators describing the June 2015 uMdloti Estuary food web network between different mirror algorithm starting points (Central, LSEI) at increasing jump sizes and number of iterations. Red tiles indicate a greater likelihood of larger flow values in group 1 (x-axis) than in group 2 (y-axis), and blue tiles indicate a greater likelihood of larger flow values in group 2 (y-axis) than in group 1 (x-axis). White tiles indicate a complete overlap of flow values between groups ($\delta = 0$). Statistically significant differences ($|\delta| \geq 0.33$) are outlined in black. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

mgC m^{-2} d^{-1}) passed all diagnostics of stationarity, mixing, and sampling efficiency (Fig. 6). This infers that the sampled food webs had reached stationarity (Geweke and HW diagnostics), were well-mixed (autocorrelation and *I*), and that enough samples had been returned to adequately describe the solution space distribution (*ESS* and *N*). Most flow values sampled with algorithm scenarios of the smallest jump size (0.01 mgC m^{-2} d^{-1}) did not pass MCMC convergence diagnostics, inferring poor convergence and poor algorithm quality. Poor convergence potentially results from, among others, poor choices of starting solutions, high posterior correlations, or the 'stickiness' of the sampled chain (Plummer et al., 2020).

Different proportions of flows 'passed' each separate diagnostic of stationarity, mixing, and sampling efficiency for the same algorithm scenarios, highlighting the methodological differences in how each diagnostic assesses convergence. For example, Geweke stationarity assessments of flows derived from the algorithm scenarios of the smallest

jump size showed that a greater proportion had 'passed' the diagnostic (35.5–52 %) in comparison to the HW diagnostic tests (15–30 %; Fig. 6). For the same algorithm scenarios, mixing assessments showed that none of the flows passed the lag-*k* autocorrelation diagnostic (0 %), whereas a larger proportion passed the *I* diagnostic (12.5–46 %). Sampling efficiency assessments show that none of the flows passed the *ESS* diagnostic (0 %), whereas a larger proportion passed the *N* diagnostic test (3–23.5 %). Such differences between diagnostic tests were expected as each diagnostic estimates convergence in a different method (Du et al., 2022; Hogg and Foreman-Mackey, 2018; Mengersen et al., 1999). Therefore, the findings of this study corroborate previous recommendations to use multiple MCMC convergence diagnostics to assess an aspect of stationarity, mixing, and sampling efficiency (Du et al., 2022; Hogg and Foreman-Mackey, 2018; Mengersen et al., 1999).

While true MCMC convergence is impossible to test (Hogg and Foreman-Mackey, 2018), to confidently state that a sampling algorithm is 'good' quality (i.e., fully samples the solution space), most flows of the sampled plausible food webs should meet convergence criteria (Roy, 2020). Therefore, we recommend applying multiple MCMC diagnostics on all flows to assess MCMC convergence and algorithm quality. This allows for greater confidence in the model outputs and their representation of the empirical system. When reported, these diagnostics can enhance model reporting and reproducibility, thus enabling improved ecosystem assessments and comparisons between studies. Statistical MCMC convergence diagnostics are available in several R packages, including **autoLIMR** (Gerber et al., 2023), **coda** (Plummer et al., 2006, 2020), and **boa** (Smith, 2007).

4.3. Ecological indicators and inferences of ecosystem status

Ecological indicators calculated on flow values sampled by the various algorithm scenarios significantly differed despite using the same a priori food web model and parameters. This highlights the importance of the mirror algorithm to adequately sample the solution space, as the difference in ecological indicators calculated from different algorithm scenarios can be large enough to make different inferences about ecosystem status.

For example, the ENA indicators derived from flow values calculated with a good quality algorithm (Central starting solution, 100 mgC m^{-2} d^{-1} jump size, and 20,000 iterations) suggest greater system activity, cycling (FCI = 0.27 ± 0.09), and reliance on primary producers to fuel the food web via herbivory (Fig. 7). More productive systems have higher activity and indicate more favourable conditions for primary production (Patten, 1995). The higher cycling indicates a decreased dependency on external energy inputs to maintain system function (de Jonge and Schückel, 2019). The low D:H indicates a high availability of primary producers as a food source, potentially from high primary production (e.g., Chrystal and Scharler, 2014).

In contrast, the ENA indicators derived from flow values calculated with a poor-quality algorithm (Central starting solution, 0.01 mgC m^{-2} d^{-1} jump size, and 20,000 iterations) suggest lower system activity, cycling, and reliance on primary producers to fuel the food web (Fig. 7). Lower system activity can be interpreted as decreased system productivity (Patten, 1995). Low cycling indicates a reliance on boundary inputs to sustain system function (de Jonge and Schückel, 2019). The large D:H ratio suggests increased detritivory, and together with low cycling, infers system dependency on detritus inputs from upstream to maintain system function (Heymans et al., 2002).

The only similarities between the two algorithm ecological indicator outputs are the Robustness indicators. Robustness incorporates AMI and H into one metric and, therefore, does not increase linearly but instead peaks on a 'robustness curve' (Ulanowicz et al., 2009). Robustness indicators from both above-mentioned algorithm scenarios are similar in value to one another and describes both systems as existent within the 'window of vitality' (Fig. B11), a theoretically narrow Robustness range where the system displays an optimal balance between flow efficiency and redundancy (de Jonge et al., 2019).

As there is increasing effort to include ecosystem models and ecological indicators in management and policy (de Jonge and Schückel, 2021; Fath et al., 2019; Schückel et al., 2022), it is important to highlight how such differences in described ecosystem status of the same model may influence future specific management actions. For example, outputs from the first scenario may lend towards management decisions of nutrient input management (to manage excessive primary production and potential eutrophication), whereas the second may suggest that freshwater input flow management is an important tool in maintaining adequate inputs of detritus into the system. However, while the linkages between ecological indicators of emergent ecosystem function and management are outside the scope of this study, we show that the methodological decisions of LIM-MCMC algorithm parameters can lead to different interpretations of ecosystem function from the same model. This underscores the importance of (1) selecting adequate LIM-MCMC parameters to ensure that the model outputs adequately represent the empirical system, and (2) the importance of adequate model quality diagnostics to provide more realistic inferences of ecosystem status, with the potential to better inform management decisions.

4.4. Limitations

One significant limitation to this study findings is that the LIM-MCMC algorithm parameters used here to adequately sample the solution space are not directly transferable to another case study. For this case study, LIM-MCMC algorithms of jump size 100 mgC m^{-1} d^{-1} were adequate to return ensembles of plausible food webs representative of the input ecological data, but this is case study dependent (Harrington et al., 2021; van Oevelen et al., 2010). For example, Nogues et al. (2021) used 500,000 iterations and a 0.5 gC m^{-2} jump size to sample the solution space of a 19-compartment, 144 flow food web model of the Courseulles-sur-Mer offshore wind farm, and Zhang et al. (2022) used 10,000 iterations with an unreported jump size to "sample the feasible region of energy flow" of five Baiyangdian Lake food web models, each consisting of 14 compartments and 69 flows. Thus, we cannot make any direct recommendations of adequate algorithm parameters. Instead, we suggest that future research uses MCMC convergence diagnostics to determine adequate algorithm parameters on a case by case basis, to determine if the user-defined algorithm parameters adequately sampled the solution space to return ensembles representative of the ecological input data variability.

Further, in empirical applications of this studies recommendations, using adequate LIM-MCMC algorithm sampling parameters may not be practically possible given computational and time constraints. In these cases, the calculated ecological indicators will not give a comprehensive overview of empirical ecosystem function. Until such time as more efficient LIM-MCMC algorithms are introduced into ecological modelling, we recommend that future studies use a combination of algorithm parameters that maximise algorithm performance, balanced by time and computation constraints. In future cases of complex food webs, where algorithm performance will not be comprehensive, it is important to robustly assess and report the model quality (e.g., with MCMC convergence diagnostics) along with the model outputs, so that interpretations of ecological function uncertainty can be reviewed within the limitations of model quality.

5. Recommendations and conclusions

We provide a novel assessment of how user-defined LIM-MCMC mirror algorithm parameters (starting solution, jump size, number of iterations) capture ecological input data variability in food web ensembles of a case study food web. Algorithm parameter selection significantly influences how much ecological data variability is incorporated into the ensembles, and these differences propagate into ecological indicators describing ecosystem function, leading to different

ecological interpretations of the same a priori parameterised model. Therefore, selecting algorithm parameters that adequately capture ecological data variability is essential to quantify the uncertainty of the empirical flow values.

We show that jump size is the most critical algorithm parameter influencing algorithm quality, and therefore recommend selecting a jump size within the range of the larger food web flows to return ecologically representative ensembles. In situations where jump size is limited to smaller values due to computational inefficiencies, we recommend selecting a starting solution in a more 'central' region of the solution space, and increasing the number of iterations to improve algorithm quality to return ecologically representative food web models. In all cases, the output ensembles should be assessed post hoc to determine whether the selected algorithm parameters allow the LIM-MCMC algorithm to sufficiently capture the input data variability.

We found that currently established post hoc assessments of algorithm quality with trace plots and coefficient of variation (CoV) can be improved with additional MCMC convergence diagnostics of stationarity, mixing, and sampling efficiency. Statistical MCMC convergence diagnostics are not only more robust assessments of algorithm quality than trace plots and CoV, they also assess convergence in different ways. Therefore, we recommend using multiple MCMC convergence diagnostics to assess each aspect of stationarity, mixing, and sampling efficiency.

We trust that these novel recommendations of LIM-MCMC mirror algorithm parameter selection and robust quality diagnostics will facilitate improved representation of ecological input data into network models. Thereby, uncertainties in empirical flow values and of ecosystem function can be quantified with greater precision, enhancing confidence in interpretations of empirical ecosystem function.

CRediT authorship contribution statement

Gemma Gerber: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Ursula M. Scharler:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the National Research Foundation (NRF), South Africa, for funding (Grant No: 93554 and 118608, grant holder: UM Scharler, Grant No: 139404, grant holder: G Gerber).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoinf.2024.102865.

Data availability

All data used to construct the network are openly reported in the Supplementary Material. All network data and reproducible code used to simulate and statistically analyse results are publicly available in the authors GitHub repository https://github.com/gemmagerber/jump_around. Should any further information be required, please email the corresponding author or log an issue in via the GitHub repository https://github.com/gemmagerber/jump_around/issues.

References

Bell, J.B., Woulds, C., van Oevelen, D., 2017. Hydrothermal activity, functional diversity and chemoautotrophy are major drivers of seafloor carbon cycling. Sci. Rep. 7, 12025. https://doi.org/10.1038/s41598-017-12291-w.

Bentley, J.W., Hines, D., Borrett, S., Serpetti, N., Fox, C., Reid, D.G., Heymans, J.J., 2019. Diet uncertainty analysis strengthens model-derived indicators of food web structure and function. Ecol. Indic. 98, 239–250. https://doi.org/10.1016/J.ECOLIND.2018.11.008.

Borrett, S.R., Lau, M.K., 2014. enaR: an R package for ecosystem network analysis. Methods Ecol. Evol. 5, 1206–1213. https://doi.org/10.1111/2041-210X.12282.

Brey, T., 2010. An empirical model for estimating aquatic invertebrate respiration. Methods Ecol. Evol. 1, 92–101. https://doi.org/10.1111/j.2041-210x.2009.00008.x.

Brey, T., 2012. A multi-parameter artificial neural network model to estimate macrobenthic invertebrate productivity and production. Limnol. Oceanogr. Methods 10, 581–589. https://doi.org/10.4319/lom.2012.10.581.

Butts, C.T., 2008. Network: a package for managing relational data in R. J. Stat. Softw. 24. https://doi.org/10.18637/jss.v024.i02.

Chrystal, R.A., Scharler, U.M., 2014. Network analysis indices reflect extreme hydrodynamic conditions in a shallow estuarine Lake (lake St Lucia), South Africa. Ecol. Indic. 38, 130–140. https://doi.org/10.1016/j.ecolind.2013.10.025.

Cliff, N., 1993. Dominance statistics: ordinal analyses to answer ordinal questions. Psychol. Bull. 114, 494–509. https://doi.org/10.1037/0033-2909.114.3.494.

de Jonge, V.N., Schückel, U., 2019. Exploring effects of dredging and organic waste on the functioning and the quantitative biomass structure of the Ems estuary food web by applying input method balancing in ecological network analysis. Ocean Coast. Manag. 174, 38–55. https://doi.org/10.1016/j.ocecoaman.2019.03.013.

de Jonge, V.N., Schückel, U., 2021. A comprehensible short list of ecological network analysis indices to boost real ecosystem-based management and policy making. Ocean Coast. Manag. 208, 105582. https://doi.org/10.1016/j.ocecoaman.2021.105582.

de Jonge, V.N., Schückel, U., Baird, D., 2019. Subsets of food webs cannot be used as a substitute to assess the functioning of entire ecosystems. Mar. Ecol. Prog. Ser. 613, 49–66. https://doi.org/10.3354/meps12863.

de Jonge, D.S.W., Stratmann, T., Lins, L., Vanreusel, A., Purser, A., Marcon, Y., Rodrigues, C.F., Ravara, A., Esquete, P., Cunha, M.R., Simon-Lledó, E., van Breugel, P., Sweetman, A.K., Soetaert, K., van Oevelen, D., 2020. Abyssal food-web model indicates faunal carbon flow recovery and impaired microbial loop 26 years after a sediment disturbance experiment. Prog. Oceanogr. 189, 102446. https://doi.org/10.1016/J.POCEAN.2020.102446.

de la Vega, C., Schückel, U., Horn, S., Kröncke, I., Asmus, R., Asmus, H., 2018. How to include ecological network analysis results in management? A case study of three tidal basins of the Wadden Sea, South-Eastern North Sea. Ocean Coast. Manag. 163, 401–416. https://doi.org/10.1016/j.ocecoaman.2018.07.019.

Du, H., Ke, Z., Jiang, G., Huang, S., 2022. The performances of Gelman-Rubin and Geweke's convergence diagnostics of Monte Carlo Markov Chains in Bayesian analysis. J. Behav. Data Sci. 2, 47–72. https://doi.org/10.35566/jbds/v2n2/p3.

Fath, B.D., Scharler, U.M., Ulanowicz, R.E., Hannon, B., 2007. Ecological network analysis: network construction. Ecol. Model. 208, 49–55. https://doi.org/10.1016/J.ECOLMODEL.2007.04.029.

Fath, B.D., Asmus, H., Asmus, R., Baird, D., Borrett, S.R., de Jonge, V.N., Ludovisi, A., Niquil, N., Scharler, U.M., Schückel, U., Wolff, M., 2019. Ecological network analysis metrics: the need for an entire ecosystem approach in management and policy. Ocean Coast. Manag. 174, 1–14. https://doi.org/10.1016/J.OCECOAMAN.2019.03.007.

Finn, J.T., 1976. Measures of ecosystem structure and function derived from analysis of flows. J. Theor. Biol. 56, 363–380. https://doi.org/10.1016/S0022-5193(76)80080-X.

Finn, J.T., 1980. Flow analysis of models of the Hubbard brook ecosystem. Ecology 61, 562–571. https://doi.org/10.2307/1937422.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., Gelman, A., 2019. Visualization in Bayesian workflow. J. R. Stat. Soc. A. Stat. Soc. 182, 389–402. https://doi.org/10.1111/rssa.12378.

Gerber, G., Brooker, B., Scharler, U.M., 2023. Automated workflow for incorporation and evaluation of data uncertainty in ecological networks with autoLIMR. Eco. Inform. 78. https://doi.org/10.1016/j.ecoinf.2023.102375.

Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F. (Eds.), Bayesian Statistics. Clarendon, Oxford, UK, pp. 169–193.

Harrington, S.M., Wishingrad, V., Thomson, R.C., 2021. Properties of Markov Chain Monte Carlo performance across many empirical alignments. Mol. Biol. Evol. 38, 1627–1640. https://doi.org/10.1093/molbev/msaa295.

Haskell, K.H., Hanson, R.J., 1981. An algorithm for linear least squares problems with equality and nonnegativity constraints. Math. Program. 21, 98–118. https://doi.org/10.1007/BF01584232.

Heidelberger, P., Welch, P.D., 1983. Simulation run length control in the presence of an initial transient. Oper. Res. 31, 1109–1144. https://doi.org/10.1287/opre.31.6.1109.

Heymans, J.J., Ulanowicz, R.E., Bondavalli, C., 2002. Network analysis of the South Florida Everglades graminoid marshes and comparison with nearby cypress ecosystems. Ecol. Model. 149, 5–23. https://doi.org/10.1016/S0304-3800(01)00511-7.

Hines, D.E., Ray, S., Borrett, S.R., 2018. Uncertainty analyses for ecological network analysis enable stronger inferences. Environ. Model. Softw. 101, 117–127. https://doi.org/10.1016/j.envsoft.2017.12.011.

Hogg, D.W., Foreman-Mackey, D., 2018. Data analysis recipes: using Markov chain Monte Carlo. Astrophys. J. Suppl. Ser. 236, 11. https://doi.org/10.3847/1538-4365/aab76e.
Hu, G., Wang, H., 2021. Most likely optimal subsampled Markov chain Monte Carlo. J. Syst. Sci. Complex. 34, 1121–1134. https://doi.org/10.1007/s11424-020-9335-1.
Huntley, M.E., Lopez, M.D.G., 1992. Temperature-dependent production of marine copepods: a global synthesis. Am. Nat. 140, 201–242. https://doi.org/10.1086/285410.
Kassambara, A., 2022. rstatix: pipe-friendly framework for basic statistical tests [WWW Document]. URL https://cran.r-project.org/package=rstatix.
Kay, J.J., Graham, L.A., Ulanowicz, R.E., 1989. A detailed guide to network analysis. In: Network Analysis in Marine Ecology. Springer, Berlin Heidelberg, pp. 15–61. https://doi.org/10.1007/978-3-642-75017-5_2.
Kelly, T.B., Davison, P.C., Goericke, R., Landry, M.R., Ohman, M.D., Stukel, M.R., 2019. The importance of mesozooplankton diel vertical migration for sustaining a mesopelagic food web. Front. Mar. Sci. 6, 508. https://doi.org/10.3389/fmars.2019.00508.
Kones, J.K., Soetaert, K., van Oevelen, D., Owino, J.O., 2009a. Are network indices robust indicators of food web functioning? A Monte Carlo approach. Ecol. Model. 220, 370–382. https://doi.org/10.1016/j.ecolmodel.2008.10.012.
Kones, J.K., Soetaert, K., van Oevelen, D., Owino, J.O., 2009b. Package "NetIndices.". Ecol. Model. 220, 370–382.
Lassalle, G., Lobry, J., Le Loc'h, F., Bustamante, P., Certain, G., Delmas, D., Dupuy, C., Hily, C., Labry, C., Le Pape, O., Marquis, E., Petitgas, P., Pusineri, C., Ridoux, V., Spitz, J., Niquil, N., 2011. Lower trophic levels and detrital biomass control the Bay of Biscay continental shelf food web: implications for ecosystem management. Prog. Oceanogr. 91, 561–575. https://doi.org/10.1016/j.pocean.2011.09.002.
Lau, M.K., Borrett, S.R., Hines, D.E., Singh, P., 2017. enaR: tools for Ecological Network Analysis [WWW Document]. URL https://cran.r-project.org/package=enaR.
Le Guen, C., Tecchio, S., Dauvin, J.-C., De Roton, G., Lobry, J., Lepage, M., Morin, J., Lassalle, G., Raoux, A., Niquil, N., 2019. Assessing the ecological status of an estuarine ecosystem: linking biodiversity and food-web indicators. Estuar. Coast. Shelf Sci. 228, 106339. https://doi.org/10.1016/J.ECSS.2019.106339.
Luong, A.D., De Laender, F., Olsen, Y., Vadstein, O., Dewulf, J., Janssen, C.R., 2014. Inferring time-variable effects of nutrient enrichment on marine ecosystems using inverse modelling and ecological network analysis. Sci. Total Environ. 493, 708–718. https://doi.org/10.1016/j.scitotenv.2014.06.027.
MacArthur, R., 1955. Fluctuations of animal populations and a measure of community stability. Ecology 36, 533. https://doi.org/10.2307/1929601.
Macbeth, G., Razumiejczyk, E., Ledesma, R.D., 2011. Cliff's Delta calculator: a non-parametric effect size program for two groups of observations. Univ. Psychol. 10, 545–555.
McKay Curtis, S., 2018. mcmcplots: Create plots from MCMC output [WWW Document]. URL https://cran.r-project.org/package=mcmcplots.
Meddeb, M., Niquil, N., Grami, B., Mejri, K., Haraldsson, M., Chaalali, A., Pringault, O., Hlaili, A.S., 2019. A new type of plankton food web functioning in coastal waters revealed by coupling Monte Carlo Markov chain linear inverse method and ecological network analysis. Ecol. Indic. 104, 67–85. https://doi.org/10.1016/j.ecolind.2019.04.077.
Mengersen, K., Knight, S., Robert, C.P., 1999. MCMC: How Do we Know when to Stop?.
Nemeth, C., Fearnhead, P., 2021. Stochastic gradient Markov chain Monte Carlo. J. Am. Stat. Assoc. 116, 433–450. https://doi.org/10.1080/01621459.2020.1847120.
Niquil, N., Saint-Béat, B., Johnson, G.A., Soetaert, K., van Oevelen, D., Bacher, C., Vézina, A.F., 2011. 9.07 - inverse modeling in modern ecology and application to coastal ecosystems. In: Wolanski, E., McLusky, D.B.T. (Eds.), Treatise on Estuarine and Coastal Science. Academic Press, Waltham, pp. 115–133. https://doi.org/10.1016/B978-0-12-374711-2.00906-2.
Niquil, N., Haraldsson, M., Sime-Ngando, T., Huneman, P., Borrett, S.R., 2020. Shifting levels of ecological network's analysis reveals different system properties. Philos. Trans. R. Soc. B 375, 20190326. https://doi.org/10.1098/rstb.2019.0326.
Nogues, Q., Raoux, A., Araignous, E., Chaalali, A., Hattab, T., Leroy, B., Ben Rais Lasram, F., David, V., Le Loc'h, F., Dauvin, J.C., Niquil, N., 2021. Cumulative effects of marine renewable energy and climate change on ecosystem properties: sensitivity of ecological network analysis. Ecol. Indic. 121. https://doi.org/10.1016/j.ecolind.2020.107128.
Nozais, C., Perissinotto, R., Tita, G., 2005. Seasonal dynamics of meiofauna in a South African temporarily open/closed estuary (Mdloti Estuary, Indian Ocean). Estuar. Coast. Shelf Sci. 62, 325–338. https://doi.org/10.1016/J.ECSS.2004.09.020.
Olli, K., Halvorsen, E., Vernet, M., Lavrentyev, P.J., Franzè, G., Sanz-Martin, M., Paulsen, M.L., Reigstad, M., 2019. Food web functions and interactions during spring and summer in the Arctic water inflow region: investigated through inverse modeling. Front. Mar. Sci. 6, 244. https://doi.org/10.3389/fmars.2019.00244.
Ortega-Cisneros, K., Scharler, U.M., Whitfield, A.K., 2016. Carbon and nitrogen system dynamics in three small South African estuaries, with particular emphasis on the influence of seasons, river flow and mouth state. Mar. Ecol. Prog. Ser. 557, 17–30. https://doi.org/10.3354/meps11870.
Park, J., Atchadé, Y., 2020. Markov chain Monte Carlo algorithms with sequential proposals. Stat. Comput. 30, 1325–1345. https://doi.org/10.1007/s11222-020-09948-4.
Parr, C.S., Wilson, N., Leary, P., Schulz, K.S., Lans, K., Walley, L., Hammock, J.A., Goddard, A., Rice, J., Studer, M., Holmes, J.T.G., Corrigan, R.J.J., 2014. The encyclopedia of life v2: providing global access to knowledge about life on earth. Biodiv. Data J., e1079 https://doi.org/10.3897/BDJ.2.e1079.
Patten, B.C., 1995. Network integration of ecological extremal principles: exergy, emergy, power, ascendency, and indirect effects. Ecol. Model. 79, 75–84. https://doi.org/10.1016/0304-3800(94)00037-I.
Pezy, J.-P., Raoux, A., Marmin, S., Balay, P., Dauvin, J.-C., 2018. What are the most suitable indices to detect the structural and functional changes of benthic community after a local and short-term disturbance? Ecol. Indic. 91, 232–240 doi: 0.1016/j.ecolind.2018.04.009.
Plummer, M., Best, N., Cowles, K., Vines, K., 2006. coda: convergence diagnosis and output analysis for MCMC [WWW Document]. R News. URL https://cran.r-project.org/package=coda.
Plummer, M., Best, N., Cowles, K., Vines, K., Sarkar, D., Bates, D., Almod, R., Magnusson, A., 2020. Package coda: output analysis and diagnostics for MCMC [WWW document]. URL https://cran.r-project.org/package=coda.
R Core Team, 2022. R: A Language and Environment for Statistical Computing [WWW Document]. URL https://www.r-project.org/.
Raftery, A.E., Lewis, S.M., 1992. How many iterations in the Gibbs sampler? In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics. Clarendon, Oxford, UK, pp. 763–773.
Raj, R., Hamm, N.A.S., Tol, C.V.D., Stein, A., 2016. Uncertainty analysis of gross primary production partitioned from net ecosystem exchange measurements. Biogeosciences 13, 1409–1422. https://doi.org/10.5194/bg-13-1409-2016.
Robert, C.P., Casella, G., 2004. Monte Carlo Statistical Methods. Springer Texts in Statistics, Springer, New York, NY.
Roy, V., 2020. Convergence diagnostics for Markov chain Monte Carlo. Ann. Rev. Stat. Appl. 7, 387–412. https://doi.org/10.1146/annurev-statistics-031219-041300.
Safi, G., Giebels, D., Arroyo, N.L., Heymans, J.J., Preciado, I., Raoux, A., Schückel, U., Tecchio, S., de Jonge, V.N., Niquil, N., 2019. Vitamine ENA: a framework for the development of ecosystem-based indicators for decision makers. Ocean Coast. Manag. 174, 116–130. https://doi.org/10.1016/j.ocecoaman.2019.03.005.
Saint-Béat, B., Vézina, A.F., Asmus, R., Asmus, H., Niquil, N., 2013. The mean function provides robustness to linear inverse modelling flow estimation in food webs: a comparison of functions derived from statistics and ecological theories. Ecol. Model. 258, 53–64. https://doi.org/10.1016/j.ecolmodel.2013.01.023.
Saint-Béat, B., Baird, D., Asmus, H., Asmus, R., Bacher, C., Pacella, S.R., Johnson, G.A., David, V., Vézina, A.F., Niquil, N., 2015. Trophic networks: how do theories link ecosystem structure and functioning to stability properties? A review. Ecol. Indic. 52, 458–471. https://doi.org/10.1016/J.ECOLIND.2014.12.017.
Saint-Béat, B., Fath, B.D., Aubry, C., Colombet, J., Dinasquet, J., Fortier, L., Galindo, V., Grondin, P.-L., Joux, F., Lalande, C., LeBlanc, M., Raimbault, P., Sime-Ngando, T., Tremblay, J.-E., Vaulot, D., Maps, F., Babin, M., 2020. Contrasting pelagic ecosystem functioning in eastern and western Baffin Bay revealed by trophic network modeling. Elementa: Sci. Anthropocene 8. https://doi.org/10.1525/elementa.397.
Scharler, U.M., Baird, D., 2005. A comparison of selected ecosystem attributes of three south African estuaries with different freshwater inflow regimes, using network analysis. J. Mar. Syst. 56, 283–308. https://doi.org/10.1016/j.jmarsys.2004.12.003.
Scharler, U.M., Borrett, S.R., 2021. Network construction, evaluation and documentation: a guideline. Environ. Model Softw. 140, 105020. https://doi.org/10.1016/j.envsoft.2021.105020.
Scharler, U.M., Lechman, K., Radebe, T., Jerling, H.L., 2020. Effects of prolonged mouth closure in a temporarily open/closed estuary: a summary of the responses of invertebrate communities in the uMdloti Estuary, South Africa. Afr. J. Aquat. Sci. 45, 121–130. https://doi.org/10.2989/16085914.2019.1689911.
Schückel, U., Nogues, Q., Brito, J., Niquil, N., Blomqvist, M., Sköld, M., Hansen, J., Jakobsen, H., Morato, T., 2022. Pilot Assessment of Ecological Network Analysis Indices. In: OSPAR, 2023: The 2023 Quality Status Report for the North-East Atlantic. OSPAR Commission, London. Available at: https://oap.ospar.org/en/ospar-assessments/quality-status-reports/qsr-2023/indicator-assessments/pilotassessment-ecological-network-analysis-indices/.
Smith, B.J., 2007. boa: an R package for MCMC output convergence assessment and posterior inference. J. Stat. Softw. 21, 1–37. https://doi.org/10.18637/jss.v021.i11.
Soetaert, K., van Oevelen, D., 2009. Modelling food web interactions in benthic deep-sea ecosystems. Oceanography 22, 128–143. https://doi.org/10.5670/oceanog.2009.13.
Soetaert, K., van den Meersche, K., van Oevelen, D., 2009. limSolve: Solving Linear Inverse Models [WWW Document]. URL https://cran.r-project.org/package=limSolve.
Stukel, M.R., Landry, M.R., 2010. Contribution of picophytoplankton to carbon export in the equatorial Pacific: a reassessment of food web flux inferences from inverse models. Limnol. Oceanogr. 55, 2669–2685. https://doi.org/10.4319/lo.2010.55.6.2669.
Stukel, M.R., Landry, M.R., Ohman, M.D., Goericke, R., Samo, T., Benitez-Nelson, C.R., 2012. Do inverse ecosystem models accurately reconstruct plankton trophic flows? Comparing two solution methods using field data from the California current. J. Mar. Syst. 91, 20–33. https://doi.org/10.1016/j.jmarsys.2011.09.004.
Tagliarolo, M., Porri, F., Garvie, C.D., Lechman, K., Scharler, U.M., 2019. Zooplankton metabolism in South African estuaries: does habitat type influence ecological strategies? J. Plankton Res. https://doi.org/10.1093/plankt/fbz035.
Tecchio, S., Rius, A.T., Dauvin, J.C., Lobry, J., Lassalle, G., Morin, J., Bacq, N., Cachera, M., Chaalali, A., Villanueva, M.C., Niquil, N., 2015. The mosaic of habitats of the Seine Estuary: insights from food-web modelling and network analysis. Ecol. Model. 312, 91–101. https://doi.org/10.1016/j.ecolmodel.2015.05.026.
Tecchio, S., Chaalali, A., Raoux, A., Tous Rius, A., Lequesne, J., Girardin, V., Lassalle, G., Cachera, M., Riou, P., Lobry, J., Dauvin, J.C., Niquil, N., 2016. Evaluating ecosystem-level anthropogenic impacts in a stressed transitional environment: the case of the Seine Estuary. Ecol. Indic. 61, 833–845. https://doi.org/10.1016/j.ecolind.2015.10.036.
Torchiano, M., 2020. effsize: efficient effect size computation [WWW Document]. URL https://cran.r-project.org/package=effsize.

Ulanowicz, R.E., 1986. A phenomenological perspective of ecological development. In: Poston, T.M., Purdy, R. (Eds.), Aquatic Toxicology and Environmental Fate: Ninth Volume. ASTM International, West Conshohocken, PA, pp. 73–81. https://doi.org/10.1520/STP29016S.

Ulanowicz, R.E., Goerner, S.J., Lietaer, B., Gomez, R., 2009. Quantifying sustainability: resilience, efficiency and the return of information theory. Ecol. Complex. 6, 27–36. https://doi.org/10.1016/j.ecocom.2008.10.005.

van den Meersche, K., Soetaert, K., van Oevelen, D., 2009. xsample() : An R function for sampling linear inverse problems. J. Stat. Softw. 30. https://doi.org/10.18637/jss.v030.c01.

van der Heijden, L.H., Niquil, N., Haraldsson, M., Asmus, R.M., Pacella, S.R., Graeve, M., Rzeznik-Orignac, J., Asmus, H., Saint-Béat, B., Lebreton, B., 2020. Quantitative food web modeling unravels the importance of the microphytobenthos-meiofauna pathway for a high trophic transfer by meiofauna in soft-bottom intertidal food webs. Ecol. Model. 430. https://doi.org/10.1016/j.ecolmodel.2020.109129.

van Niekerk, L., Adams, J.B., James, N.C., Lamberth, S.J., MacKay, C.F., Turpie, J.K., Rajkaran, A., Weerts, S.P., Whitfield, A.K., 2020. An estuary ecosystem classification that encompasses biogeography and a high diversity of types in support of protection and management. Afr. J. Aquat. Sci. 45, 199–216. https://doi.org/10.2989/16085914.2019.1685934.

van Oevelen, D., van den Meersche, K., Meysman, F.J.R., Soetaert, K., Middelburg, J.J., Vézina, A.F., 2010. Quantifying food web flows using linear inverse models. Ecosystems 13, 32–45. https://doi.org/10.1007/s10021-009-9297-6.

van Oevelen, D., Soetaert, K., Garcia, R., de Stigter, H.C., Cunha, M.R., Pusceddu, A., Danovaro, R., 2011. Canyon conditions impact carbon flows in food webs of three sections of the Nazaré canyon. Deep-Sea Res. II Top. Stud. Oceanogr. 58, 2461–2476. https://doi.org/10.1016/j.dsr2.2011.04.009.

van Ravenzwaaij, D., Cassey, P., Brown, S.D., 2018. A simple introduction to Markov Chain Monte–Carlo sampling. Psychon. Bull. Rev. 25, 143–154. https://doi.org/10.3758/s13423-016-1015-8.

Zhang, X., Yi, Y., Yang, Z., 2022. The long-term changes in food web structure and ecosystem functioning of a shallow lake: implications for the lake management. J. Environ. Manag. 301, 113804. https://doi.org/10.1016/j.jenvman.2021.113804.