

# Revealing perturbation responses with limited observations of biological communities

Nikolaos Alexandridis<sup>a,\*</sup>, Cédric Bacher<sup>b</sup>, Fred Jean<sup>c</sup>, Jeffrey M. Dambacher<sup>d</sup>

<sup>a</sup> Lund University, Centre for Environmental and Climate Science (CEC), Sölvegatan 37, 22362 Lund, Sweden

<sup>b</sup> Ifremer, Centre de Bretagne, DYNECO-LEBCO, CS 10070, 29280 Plouzané, France

<sup>c</sup> Université de Brest, UBO/CNRS/IRD/Ifremer, LEMAR, Institut Universitaire Européen de la Mer, 29280 Plouzané, France

<sup>d</sup> CSIRO, GPO Box 1538, Hobart, Tasmania 7001, Australia

## ARTICLE INFO

### Keywords:

Biological traits  
Biotic interactions  
Community structure  
Ecological perturbation  
Environmental variability  
Qualitative modelling

## ABSTRACT

Restrictions in empirical research of biological communities have limited our understanding of the combined influence of environmental variability and system structure on community composition. Spatial patterns of community composition in less accessible systems, such as marine benthos, can often not be explained by many factors beyond the direct impact of the environment on community members. We present a method that combines commonly collected data of community composition with analyses of qualitative mathematical models, to assess not only direct impacts of environmental variability, but also the propagation of impacts through complex interaction networks. Transformed spatial data of community composition describe the community members' observed similarity of response to an external input. The output of qualitative mathematical models describes the community members' predicted similarity of response to input entering the system through any of its variables. A statistically significant agreement between the observed and any of the predicted response similarities indicates the respective system variable as a likely gateway for environmental variability into the system. The method is applied to benthic macroinvertebrate communities in the Rance estuary (Brittany, France). Organisms identified as likely gateways have traits that agree with their predicted response to documented spatially and temporally structured environmental variability. We suggest use of this novel framework for more comprehensive identification of environmental drivers of community change, including gateway community members and cascades of environmentally driven change through community structure.

## 1. Introduction

Fluctuating abiotic conditions, also known as environmental variability, and interactions among community members, outlined as system structure, are central to the study of biological communities. Understanding the combined effect of these two elements on community composition remains a difficult challenge for ecological research (Sutherland et al., 2013). Logistical restrictions often limit controlled community experiments that jointly investigate environment and structure to small subsets of natural systems. Such experiments have provided valuable theoretical insight, but their inference potential regarding the behaviour of complex real-world systems remains limited (Wernberg et al., 2012). Observational studies can produce more comprehensive community descriptions, but these are typically based on snapshots of a system that are restricted in space or time. Statistical

analysis of such sample data can reveal environmental drivers that shape community composition. Nevertheless, most multivariate statistics detect patterns shaped by organisms' requirement niche, ignoring the influence of biotic interactions, or impact niche components, sensu Leibold (1995). Consequently, indication of environmental variability based on its direct effects on a community is not robust against confounding spatial and temporal dynamics controlled by community interactions (James and McCulloch, 1990; Gotelli et al., 2009). Structural equation models are increasingly used to investigate networks of biotic interactions (Grace et al., 2010). However, representation of feedback cycles and the resulting dynamics is still challenging (Hayduk, 2009). As a result, this crucial property of biological communities is often disregarded in favor of acyclic representations (Fan et al., 2016).

The study of benthic macroinvertebrates in the Rance estuary (Brittany, France) exhibits some typical challenges of linking environmental

\* Corresponding author at: The Ecology Building, Sölvegatan 37, 22362 Lund, Sweden.  
E-mail address: [nikos.alexandridis@cec.lu.se](mailto:nikos.alexandridis@cec.lu.se) (N. Alexandridis).

<https://doi.org/10.1016/j.ecolind.2021.107840>

Received 19 October 2020; Received in revised form 22 May 2021; Accepted 24 May 2021

Available online 27 May 2021

1470-160X/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

variability and community structure. The estuary's macrobenthos was comprehensively sampled twice, in 1976 and 1995 (Kirby and Retière, 2009). Spatial patterns of community composition in each year are explained as the direct effect of the environment independently on each species (Desroy, 1998). Differences in population abundances between these observations, separated by almost twenty years, are linked to short- or long-term environmental variability (Desroy and Retière, 2004). A portion of the observed variation in community composition is explained by direct environmental impacts, but a large portion is also left unresolved. Undetermined variation in community composition could result from the propagation of impacts of external drivers through complex interaction networks. The signal of such drivers may be confounded by internal community dynamics, but likely not in a random manner. We, instead, expect this signal to be shaped deterministically by system structure. Control observations that can isolate this signal from other sources of variation, or enough knowledge and data to develop dynamic models of benthic communities, are lacking. Consequently, there is currently no means to test hypotheses regarding structure-modulated impacts of environmental variability, or predict perturbation responses of benthic communities in the Rance and other similar systems worldwide (Alexandridis, 2017).

Here, we present a new approach to identifying environmentally driven community change, by detecting system-wide perturbation effects and organisms that act as their points of entry into a community. We argue that comprehensive assessment of perturbation impacts, not only directly on community members, but also through cascading community responses, increases the robustness of environmental variability indication. Our approach combines widely available snapshot community observations and relatively easy-to-build qualitative mathematical models that represent both requirement and impact niches of community members (Levins, 1998). Qualitative predictions of population responses are compared with transformed observations of species abundances. When predicted community responses to perturbation affecting the system through a specific variable agree with observations of community composition under perturbation, this variable is identified as a likely point of entry, or gateway, for environmental variability into the system. We provide an example application of this framework based on recent functional grouping and qualitative modelling of benthic macroinvertebrate communities in the Rance estuary (Alexandridis et al., 2017a, 2017b), and discuss results in relation to the environmental history of the system. We identify research directions that can employ this framework to improve understanding of the combined role of environmental variability and community structure.

## 2. Theory

Qualitative mathematical modelling allows incorporation of community structure, especially complex feedback cycles, in assessment of the impacts of environmental variability (Levins, 1998). In this context, sustained environmental variability can be considered as a press perturbation, or external input into a system through any of its variables. Given a specific input, a qualitative mathematical model generates predictions about the direction of change of each system variable. Puccia and Levins (1985) suggest using these predictions to identify the source of an external input, by examining observed correlation patterns among species populations in a system that is suspected to have received an input. Populations that change in the same direction in response to this input will show positive correlations over a range of the input's magnitude, while those that change in opposite directions will show negative correlations. No correlation will be observed between two populations if any of them have a zero response to this input. Similarly, qualitative predictions for a modelled system regarding responses to each potential input can be translated into positive, negative or no correlations between the model variables. Predicted correlation signs that agree with those observed should, in theory, provide evidence for the source or cause of the actual input into the system. While this

approach has the potential to identify unknown perturbation sources, it requires a definitive set of model predictions. In complex systems, however, these predictions can be difficult to obtain, as an input can propagate both positive and negative effects to a response variable, rendering ambiguous qualitative predictions (Dambacher et al., 2002). Furthermore, natural variability and sampling error can easily limit the agreement between predicted and observed correlation patterns (Puccia and Levins, 1985).

## 3. Material and methods

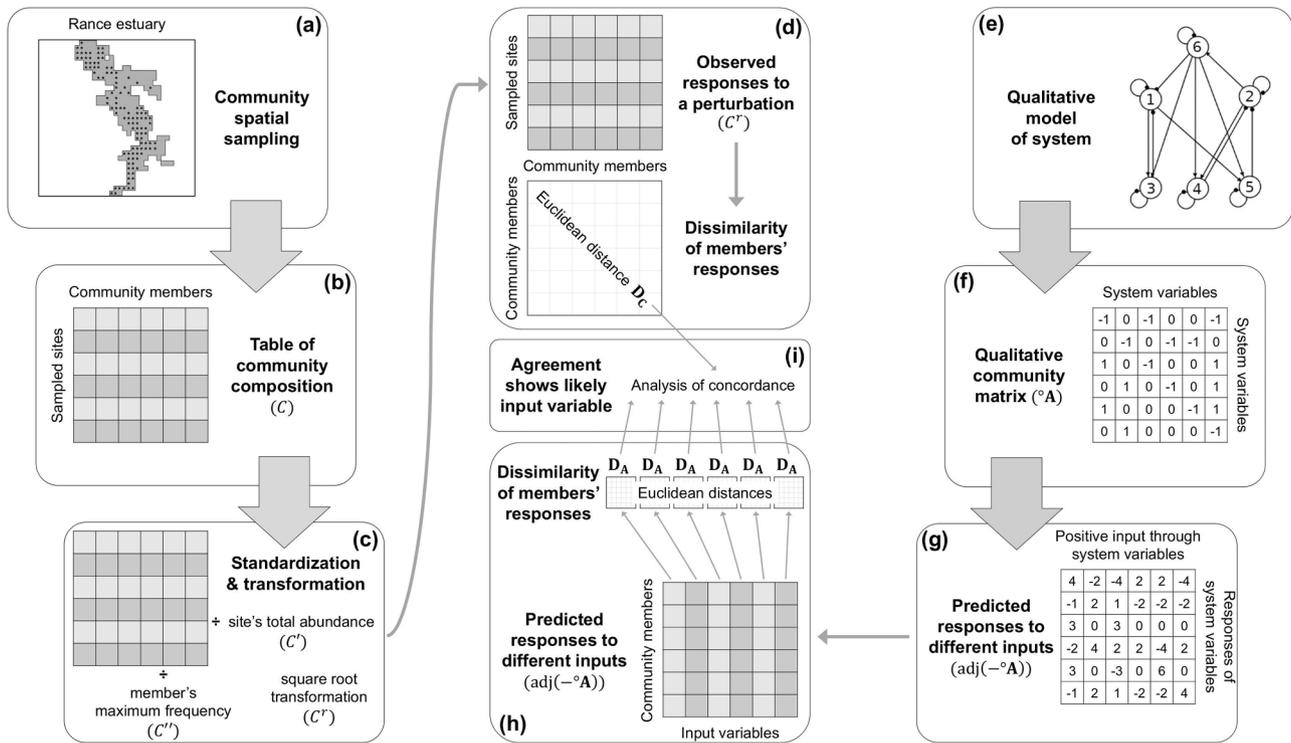
### 3.1. General assumptions

1) A system that consists of a biological community and its basic resources is assumed to be at, or near, a state of equilibrium under the influence of environmental variability in the form of press perturbation (i.e., sustained input into the system), whose sign and magnitude may vary in space and time. The equilibrium assumption allows prediction of perturbation responses through qualitative mathematical modelling, and is typically supported by a system's ability to persist in the face of disturbance (Puccia and Levins, 1985). 2) A press perturbation is expected to affect the system disproportionately through one of its variables, but the approach can also be extended to include multiple gateway variables (see section 6). 3) Detection of system-wide perturbation impacts over independently generated, pre-existing spatial abundance patterns requires that unobserved pre-perturbation abundances of each community member at different sites show random variability within a smaller range than any change caused by the press perturbation. 4) Internal community-shaping processes are assumed to act independently at each sampled site. Community samples can then be considered as independent realizations of a community model.

According to these assumptions, a perturbation forces community members with varying strength across sites away from their pre-perturbation abundances. These unobserved abundances vary, inter alia, at each site according to the members' level of numerical dominance in the community, and across sites according to the sites' total abundance, or site productivity (Borcard et al., 2011). Since perturbation impacts are assumed to exceed this variability in space, and we are interested in the community members' perturbation responses, such variation can, for our purposes, be considered as noise. Our goal is to reduce this noise, by standardizing abundances for site productivity and community dominance. Site-paired community member distances regarding standardized abundances should also eliminate the influence of site-dependent perturbation magnitude, quantifying only community members' dissimilarity in terms of their perturbation responses. This dissimilarity can then be compared with the same measure calculated from predictions of mathematical community models for each potential input variable. Agreement between observations and predictions should indicate the respective system variable as a likely gateway for environmental variability into the system.

### 3.2. Community observations

Samples of community abundance (Fig. 1a) are arranged into a data table  $C = [c_{ij}]$  of size  $(s \times m)$  with sites (rows)  $i = \{1 \dots s\}$  and community members (columns)  $j = \{1 \dots m\}$  (Fig. 1b); row sums are noted as  $c_i$  and column maxima as  $c_{\sqrt{j}}$ . Abundances of community members at each site are first standardized for site productivity through division by the site's total community abundance. The resulting table  $C' = [c'_{ij}]$ , where  $c'_{ij} = \frac{c_{ij}}{c_i}$ , describes the frequencies of community members at each site. Site frequencies of each community member are then standardized for community dominance through division by the community member's maximum frequency over all sites. The resulting table  $C'' = [c''_{ij}]$ , where  $c''_{ij} = \frac{c'_{ij}}{c'_{\sqrt{j}}}$ , describes each community member's relative frequencies



**Fig. 1.** Schematic representation of the analytical framework. Left-hand panels illustrate the treatment of community observations. (a) A spatial sampling of a community that meets the general assumptions of the methodology results in (b) a table describing community composition at the sampled sites (table C). (c) Abundances at each site are divided by the site's total abundance to standardize for site productivity (table C\*). The resulting frequencies of each community member are divided by the member's maximum frequency to standardize for community dominance (table C\*\*). The output is square root transformed to reduce the importance of extremely high values (table C\*). (d) This table describes the community members' responses to an assumed perturbation. The Euclidean distance of community members with respect to site-paired values of table C\* represents their perturbation response dissimilarity (matrix D<sub>C</sub>). Right-hand panels illustrate the derivation of qualitative model predictions. (e) A signed digraph model of the system that includes the sampled community is translated into (f) a qualitatively specified community matrix (°A). (g) Standard analysis of °A (Dambacher et al., 2002) generates the adj(-°A) matrix, each column of which describes the system variables' responses to positive input through a specific variable. (h) Euclidean distances, or dissimilarities, of community members with respect to their predicted responses to each input variable (matrices D<sub>A</sub>) are in turn (i) compared with matrix D<sub>C</sub> through concordance analysis. A statistically significant level of agreement between observations and predictions indicates the respective system variable as a likely gateway for external input into the system.

across sites (Fig. 1c). Double standardization makes these values comparable between sites and community members, regardless of systematic differences in site productivity and community dominance. Finally, square root transformation reduces the importance of observations with extremely high values, resulting in table  $C^* = [c^*_{ij}]$ , where  $c^*_{ij} = \sqrt{c^{**}_{ij}}$  (Fig. 1c). For each pair  $\{x, y\} \in \binom{\{1 \dots m\}}{2}$  of the  $m$  community members, the Euclidean distance  $D(x, y) = \sqrt{\sum_{i=1}^s (c^*_{ix} - c^*_{iy})^2}$ , calculated among values of the respective  $s$ -length columns of table  $C^*$ , quantifies the community members' dissimilarity in terms of their relative frequencies across all sites. Calculations of  $D(x, y)$  for all elements of  $\{x, y\}$  are combined into the distance matrix  $D_C$  (Fig. 1d).

Assuming that the system is affected by a press perturbation, responses of community members across sampled sites are 1) described by comparable values along the respective columns of table  $C^*$  and 2) dictated by the input's magnitude at each site and the community members' position in the system structure. If the system position of two community members  $\{x, y\}$  forces them to respond to this input in a correlated manner, site-paired values along the respective columns of table  $C^*$  will be on average more similar than expected by chance, since they will only depend on the input's magnitude at each site. As a result,  $D(x, y)$  will be minimized. The same measure will be maximized between community members whose position in the system structure results in opposite responses to that input. Under these conditions, distance matrix  $D_C$  describes the community members' dissimilarity of response to an assumed perturbation across sampled sites.

### 3.3. Model predictions

In describing a theoretical community's dynamics, the community matrix consists of the partial derivatives of each community member's population growth rate with respect to each community member's abundance (Berlow et al., 2004; Novak et al., 2016). Given a community matrix  $A$ , Eq. (1) predicts the difference between the community member  $i$ 's equilibrium abundance ( $N_i^*$ ) before and after sustained change in parameter  $h$  ( $p_h$ ) of community member  $j$ 's growth rate function ( $f_j$ ). The sign of this difference is defined by the first term of the expression,  $\text{adj}(-A)$ , as the second term is a positive scalar in stable systems and the third is by convention positive (for positive input) or negative (for negative input).

$$dN_i^* = N_i^* \text{new} - N_i^* \text{old} = \text{adj}(-A) \times \frac{1}{\det(-A)} \times \frac{\partial f_j}{\partial p_h} dp_h \quad (1)$$

The community matrix can be specified qualitatively (i.e., with entries of +1, -1 and 0), in which case it is denoted as °A (Fig. 1e, f). For a system with  $n$  variables, the  $\text{adj}(-°A) = (a_{kl})$  is a  $(n \times n)$  matrix with rows  $k = \{1 \dots n\}$  and columns  $l = \{1 \dots n\}$ . Assuming that the modelled system is stable, each column vector  $\mathbf{a} = (a_{k0})$  ( $0 = \text{constant}$ ) of the  $\text{adj}(-°A)$  predicts the direction of the equilibrium response of all system variables to a press perturbation that affects the system through a single variable (Dambacher et al., 2002) (Fig. 1g). These values represent the net number of positive and negative effects that contribute to the direction of responses, and have been found to at least moderately correspond with quantitatively predicted response strengths (Dambacher et al.,

2003). Therefore, they describe not only the direction but, to some degree, also the magnitude of the system's response to external input entering through a single variable. For each pair  $\{w, z\} \in \binom{\{1 \dots n\}}{2}$  of the  $n$  system variables, the Euclidean distance  $D(w, z) = \sqrt{(a_{w0} - a_{z0})^2}$ , calculated between the respective values of the column vector  $\mathbf{a}$ , quantifies the predicted dissimilarity of response of two system variables to such an external input. Calculations of  $D(w, z)$  for all elements of  $\{w, z\}$  are combined into the distance matrix  $\mathbf{D}_A$ , and the process is repeated for all columns of the  $\text{adj}(-^\circ\mathbf{A})$  (Fig. 1h).

### 3.4. Observations vs. Predictions

Impacts of environmental variability would ideally be detected by sampling biological communities before and after a perturbation. The first set of samples would act as controls, minimizing the effects of drivers other than the perturbation in the second set of samples. However, environmental variability can operate over temporal scales that are too large to allow observation of non-perturbed communities. Furthermore, the use of destructive sampling methods can preclude this approach, while the need for sampling often arises only after a perturbation has occurred. As a result, pre-perturbation (pre-input) abundances of community members ( $N_i^*$  old in Eq. (1)) are often unobserved, so we have to generate hypotheses about external input into the system based only on observed post-perturbation (post-input) abundances of community members ( $N_i^*$  new in Eq. (1)).

Qualitative community models predict the signal of external input as the perturbation is propagated through the system's structure. Identifying a single system variable as a perturbation gateway depends on separating this signal in  $N_i^*$  new from systematic differences in site productivity, community dominance and other random sources of abundance variation. Transformation of community members' abundances is aimed at reducing this background noise. The resulting distance matrix  $\mathbf{D}_C$  is here intended to describe the community members' observed response dissimilarities to an assumed perturbation. Each distance matrix  $\mathbf{D}_A$  describes the community members' predicted response dissimilarities to perturbation in the form of external input through a single variable. The problem of identifying likely input variables then becomes one of testing the agreement between observed ( $\mathbf{D}_C$ ) and predicted dissimilarities of response to input through each system variable ( $\mathbf{D}_A$ ). To this end, matrix  $\mathbf{D}_C$  is compared in turn with the array of  $\mathbf{D}_A$  matrices, each calculated from a column of the  $\text{adj}(-^\circ\mathbf{A})$  (Fig. 1i). Agreement between each pair of matrices can be assessed by a test of congruence among distance matrices (CADM), which measures Kendall's  $W$  coefficient of concordance, and quantifies its significance level through a permutation procedure (Legendre and Lapointe, 2004). The null hypothesis is monotonic independence of matrices. The alternative hypothesis is that matrices are congruent, i.e., they have similar rankings of community members' distances.

## 4. Calculation

### 4.1. System characteristics

The Rance estuary is situated on the northern coast of Brittany, France. A tidal power plant spans its mouth, which opens to the English Channel. The operating constraints of the power plant have reduced the tidal range in the estuary compared to the open sea. Maximum water depth is 17 m at low tide, but the main part of the estuary is 5–6 m deep. From downstream to upstream of the estuary, pebbles and coarse sand are replaced by medium and fine sand, muddy sand and finally mud. A similar sequence is observable from the central channel to the banks. Since operation of the power plant started in 1966, natural silting has increased due to long periods of slack water, while strong sluice and

turbine currents have eroded parts of the riverbed (Retière, 1994). Macrobenthic community composition approximately stabilized 10 years after start of the power plant's operation (Desroy, 1998). Infaunal organisms dominate macrobenthos, but proliferation of the slipper limpet (*Crepidula fornicata*) since the 1970's has enhanced epifaunal species assemblages. Fluctuations of environmental conditions are rather regular, and macrobenthos is protected from the disruptive effects of swells or storms. Extremely low temperatures can however reduce benthic abundances (Desroy, 1998). Harmful algal blooms can occur, but have been declining in magnitude since the mid-90's, probably due to decreasing phosphate input to the Rance drainage basin (Le Bec et al., 2016). In the Rance estuary's relatively stable environment, benthic abundances are considered to fluctuate around equilibrium levels in response to abiotic drivers, while biotic interactions are expected to have stronger influence on benthic community composition than in the open sea (Desroy and Retière, 2004).

### 4.2. Community observations

Samples of the upper sediment layers were collected at 113 sites across the bed of the Rance estuary in April 1995, prior to spring recruitment. Invertebrate organisms larger than 1 mm were counted and determined at species level. Species abundances were extrapolated to a standard sediment volume across samples (Desroy, 1998). A set of 14 biological traits (Table 1) described the role of each of the 240 collected species in an array of ecological processes, including environmental filtering, consumption of algae or detritus, predation, use of space, biogenic habitat modification and trade-offs in utilization and allocation of resources. These traits allowed application of the emergent group hypothesis (Héroult, 2007) for the classification of species into 20 functional groups (Table 1). The groups were given trait values and were found to suffice for the representation of functional diversity in the Rance estuary (Alexandridis et al., 2017a). Abundances of the 20 functional groups at the 113 sampled sites (see Supplemental Information of Alexandridis et al. (2018)) were split between tables  $C_1$  and  $C_2$ , separating the 12 epifaunal from the 9 infaunal groups, respectively (one group was included in both matrices, because of the high mobility and variable substrate preference of its member species). These tables were standardized and transformed into tables  $C_1^r$  and  $C_2^r$ , from which Euclidean distance matrices  $\mathbf{D}_{C1}$  and  $\mathbf{D}_{C2}$  were calculated, respectively.

Mantel correlograms quantified spatial correlation in the multivariate domain of functional group abundances, in order to investigate the spatial structure of community observations, and assess violations of general assumptions. Correlograms are based on calculation of the normalized Mantel statistic between pairs of site dissimilarity matrices. One matrix in each pair quantifies differences in multivariate community composition, and the other is derived by attributing the value 0 to pairs of sites that belong to the same distance class and the value 1 to all other pairs of sites. The process is repeated for each distance class and values of the Mantel statistic are tested by permutations. Mantel correlograms were derived from the square root of Wisconsin-transformed tables  $C_1$  and  $C_2$ . This is the same procedure producing tables  $C_1^r$  and  $C_2^r$ , except that division by row maxima precedes division by column totals, as sites are compared instead of community members. The dataset was detrended, and the permutation test results were Holm-corrected for multiple testing. The number of distance classes was calculated based on Sturge's rule, and the correlograms were restricted to distances including all sites (Borcard et al., 2011).

### 4.3. Model predictions

General rules of biotic interaction were formulated with the help of ecological theory and expert knowledge (Alexandridis et al., 2017b), towards constructing two signed-directed graph, or signed digraph, models (Fig. 2). Model variables included the previously built functional

**Table 1**  
Functional groups of species with their assigned representative species and biological trait values. Groups starting with 'H' and 'C' correspond to algae or detritus feeders and predators/scavengers, respectively. Unitless quantities do not represent absolute, but only relative values; see Alexandridis et al. (2017a) for details.

Groups	Representative species	Temperature	Development	Dispersal	Fecundity	Tide/salinity	Substrate	Size (cm)	Area	Position	Mobility	Growth rate	Lifespan (yr)	Epibiosis	Engineering
H1	<i>Morchellium argus</i>	eurythermal	brooded	short	low	stenohaline	rock	3.3	0.1	epifauna	sessile	2.6	1.7	epibiont	neutral
H2	<i>Lepidochitona cinerea</i>	stenothermal	planktonic	short	high	stenohaline	rock	10.8	4.1	epifauna	mobile	0.9	11.6	epibiont	neutral
H3	<i>Balanus crenatus</i>	eurythermal	planktonic	long	high	euryhaline	rock	2.0	0.8	epifauna	sessile	2.5	2.0	epibiont	neutral
H4	<i>Crepidula fornicata</i>	stenothermal	planktonic	long	high	stenohaline	rock	7.6	0.0	epifauna	sessile	1.9	11.2	basibiont	neutral
H5	<i>Oligochaeta</i>	stenothermal	laid	short	low	emersed	muddy sand	4.5	5.0	infauna	mobile	3.4	2.0	neutral	destabilizer
H6	<i>Thyasira flexuosa</i>	eurythermal	planktonic	short	low	stenohaline	mud	3.6	0.8	infauna	mobile	1.0	10.0	neutral	stabilizer
H7	<i>Melinna palmata</i>	stenothermal	brooded	short	low	stenohaline	mud	7.5	0.3	interface	sessile	2.6	3.6	neutral	stabilizer
H8	<i>Notomastus latericeus</i>	stenothermal	brooded	short	low	stenohaline	muddy sand	6.0	2.9	interface	mobile	2.6	1.9	neutral	destabilizer
H9	<i>Hedise diversicolor</i>	eurythermal	laid	short	high	emersed	muddy sand	12.8	0.2	interface	mobile	2.1	3.4	neutral	destabilizer
H10	<i>Malacoeris fuliginosus</i>	eurythermal	planktonic	long	high	euryhaline	mud	8.5	1.9	interface	mobile	2.5	2.7	neutral	destabilizer
H11	<i>Galathea oculata</i>	eurythermal	planktonic	long	high	euryhaline	mud	11.1	0.0	interface	sessile	2.7	4.4	neutral	stabilizer
H12	<i>Glycymeris glycymeris</i>	stenothermal	planktonic	short	high	stenohaline	muddy gravel	8.0	1.4	infauna	mobile	0.8	15.0	neutral	stabilizer
H13	<i>Anapagurus hynchmanni</i>	stenothermal	planktonic	long	high	stenohaline	gravel	10.0	0.1	epifauna	mobile	0.6	10.0	neutral	neutral
H14	<i>Cerastoderma edule</i>	stenothermal	planktonic	long	high	emersed	muddy sand	8.6	0.5	interface	mobile	0.7	8.9	neutral	stabilizer
C1	<i>Sphaerosyllis bulbosa</i>	stenothermal	brooded	short	low	stenohaline	gravel	1.3	0.3	epifauna	mobile	4.7	1.9	neutral	neutral
C2	<i>Marphysa bellii</i>	stenothermal	planktonic	short	high	stenohaline	muddy sand	23.3	0.5	interface	mobile	1.1	4.7	neutral	neutral
C3	<i>Nephtys hombergii</i>	stenothermal	planktonic	long	high	stenohaline	gravel	10.5	0.3	interface	mobile	2.2	7.3	neutral	neutral
C4	<i>Myrtilanida edwardsi</i>	stenothermal	planktonic	long	low	stenohaline	mud	1.4	3.1	interface	mobile	5.8	1.9	neutral	neutral
C5	<i>Urticina felina</i>	eurythermal	planktonic	short	high	euryhaline	rock	16.7	10.3	epifauna	sessile	1.1	14.0	epibiont	neutral
C6	<i>Syllis cornuta</i>	stenothermal	planktonic	long	low	stenohaline	rock	7.4	5.2	epifauna	mobile	2.3	2.3	epibiont	neutral

groups and two resource variables representing food and space. Interactions between variables (positive or negative direct effects) were defined based on trait values of the functional groups (Alexandridis et al., 2017b). Model 1 was built for epifaunal organisms that prefer rock or gravel substrates, and model 2 for infaunal or interface-positioned organisms preferring finer sediment substrates. Matrices  ${}^{\circ}A_1$  and  ${}^{\circ}A_2$  are an equivalent representation of models 1 and 2, respectively (Dambacher et al., 2002). Model analysis revealed a moderate to high potential for stability (Alexandridis et al., 2017b). Each column of the  $adj(-{}^{\circ}A_1)$  and  $adj(-{}^{\circ}A_2)$  describes the predicted responses of functional groups to input through a single model variable. As there are no field observations of available food and space resources, predicted responses of these variables were omitted. Should these variables respond strongly to perturbation, their omission could reduce the detecting power of the analysis. The omission, however, of variables with different units also removes the need for additional standardization of system observations. Euclidean distances were calculated within each column of the  $adj(-{}^{\circ}A_1)$  and  $adj(-{}^{\circ}A_2)$ , producing two sets of distance matrices  $D_{A1}$  and  $D_{A2}$ , respectively.

4.4. Observations vs. Predictions

Distance matrix  $D_{C1}$  ( $D_{C2}$ ) was in turn compared with each of the  $D_{A1}$  ( $D_{A2}$ ) distance matrices, corresponding to each potential input variable. Agreement between each pair of distance matrices was quantified by Kendall's *W* coefficient of concordance. The null hypothesis of incongruence between distance matrices was tested by calculating *W*'s significance level through a permutation procedure with 9999 permutations and Holm correction for multiple testing (Campbell et al., 2009).

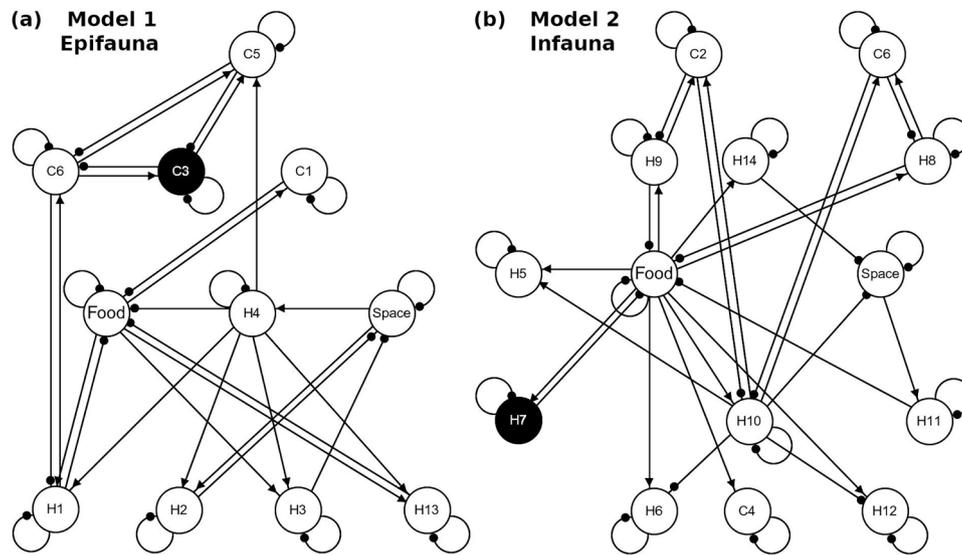
5. Results

Mantel correlograms of epifaunal (Fig. 3a) and infaunal (Fig. 3b) community observations show community samples that are more similar than expected by chance at approximately 1 km and more dissimilar at approximately 3 km of distance. Patterns of compositional similarity and dissimilarity also appear at larger distances among the sampled sites, but these are weaker and not statistically significant.

Permutation tests of Kendall's *W* coefficient of concordance (Table 2) show almost all pairs of distance matrices to be statistically independent (i.e., to convey distinct information about observed and predicted response dissimilarities of functional groups to external input through the respective variables). These system variables were therefore considered as less likely to be gateways for external input into the system. Only two comparisons resulted in statistically significant *W* values, one each for the epifaunal and infaunal communities. In the epifaunal community, predator/scavenger group C3, represented by the errant polychaete *Nephtys hombergii* (Table 1), was thus identified as a likely gateway for external input. In the infaunal community, input was shown to likely enter through algae- or detritus-feeding group H7, represented by the sedentary polychaete *Melinna palmata* (Table 1). Input through groups C3 and H7 corresponds to the highest values of *W* for each benthic community and overall (Table 2). *W* values vary between 0.76 (calculated for group C3) and 0.31, both observed in the epifaunal community. *W* values in the infaunal community vary between 0.67 (calculated for group H7) and 0.43. Greater range of *W* values in the epifaunal community shows clearer distinction between stronger and weaker agreements of observed perturbation response dissimilarities with predicted perturbation response dissimilarities. This is an indication of a stronger perturbation signal through one epifaunal system variable, compared to the infaunal community.

6. Discussion

Comparison of model predictions with community observations



**Fig. 2.** Signed digraph models of processes that shape (a) the epifaunal and (b) the infaunal benthic communities. Nodes represent functional groups or the basic resources of food and space. Links ending in arrows and filled circles represent positive and negative direct effects, respectively; see Alexandridis et al. (2017b) for details. Black nodes indicate variables identified in Section 5 as likely gateways for external input into the system.

detected only two statistically significant levels of agreement regarding functional groups' response dissimilarities, one for each benthic community (Table 2). This finding supports the expectation of input entering each system through primarily one of its variables. However, results should only be considered as an indicator of the relative likelihood of each variable to be a gateway for external input, as the sensitivity of the analysis and its robustness to violations of assumptions have not been assessed. Before the framework can be thus validated, we suggest careful evaluation of the general assumptions. Our conservative choice of assumptions means that they may often not be perfectly valid. Therefore, the objective should not simply be to assess whether assumptions are violated, but rather whether detected signals are the result of such assumption violations, instead of environmental variability.

Environmental variability, viewed as external input into benthic communities, can take many forms in the Rance estuary. Substrate characteristics primarily differentiate the spatial distribution of epifaunal versus infaunal macrobenthos (Retière, 1994). The effect of substrate on macrobenthic organisms has been previously detected with the use of multivariate statistical techniques (Desroy, 1998); we largely removed this effect on epifaunal and infaunal relative frequencies, by separately comparing observations and predictions for the two communities (Alexandridis et al., 2017b). Tidal emersion, along with associated levels of water salinity, is a further potential source of environmental variability (Desroy and Retière, 2004). Exploratory models of community dynamics have identified tidal zonation as the driver of spatial dependence in community composition in the Rance estuary (Alexandridis et al., 2018), resulting in spatial correlations similar to the small-scale patterns detected here in correlograms of epifaunal and infaunal communities.

In addition to spatial drivers, environmental variability can operate in time, but these effects are much more difficult to detect in snapshot observations of a system. The extremely cold winter of 1993-'94 has been hypothesized to affect macrobenthos in the Rance estuary (Desroy and Retière, 2004). The impact of this winter has been recorded in similar systems, including on the representative species of group C3, the polychaete *Nephtys hombergii*, in the Dutch Wadden Sea (Beukema et al., 2000). This long-term study detects perturbation impacts in spite of background random variation, as well as their cascading through a much simpler trophic chain of 3 species. Cold winters are shown to intervene on average once every 3 or 4 years, while the system appears to recover rapidly. These observations support the assumptions of adequately small

random pre-perturbation abundance variability and persistence at, or near, a state of equilibrium under the influence of press perturbation (Donohue et al., 2016) in a system that is similar to the Rance estuary.

Characteristics of the two functional groups that correspond to the highest agreement levels support their role as gateways for suspected environmental variability into the system. Both groups have been assigned stenothermal and stenohaline trait values, meaning that they cannot tolerate single-digit temperatures or salinities that differ greatly from those of the open sea (Alexandridis et al., 2017a). The concomitant occurrence of low temperatures and tidal exposure in space and time probably contributed to these groups' identification as likely gateways for environmental variability into their communities. However, an apparently stronger signal of external input in the epifaunal community, where spatial correlation patterns are weaker, suggests a stronger contribution of the previous winter. Other groups may have the same trait values, but the two identified as likely gateways for press perturbation are among the most abundant (C3 is the 6th and H7 the 2nd most abundant group overall), making their impact on community dynamics stronger compared to similarly affected organisms (Grime, 1998). Both groups are also represented by polychaetes, which have been previously identified as efficient indicators of adverse environmental conditions (Giangrande et al., 2005).

Our framework employs relatively easy-to-build qualitative mathematical models of communities to increase the mechanistic understanding that can be derived from analysis of commonly collected, snapshot data of community composition. Eq. (1) only requires that the strength of any individual input is sufficiently strong to produce an observable impact on the system. As observed distances of community members are calculated from site-paired values, no distinction is made based on the sign of the input, which can be the same or opposite across sites. The expectation for a single such input can be relaxed to consider whether one or more perturbations are acting on the system through multiple variables. The community's response to combined input through more variables can be explored by adding or subtracting the respective columns of the adjoint of the negative community matrix (Dambacher et al., 2002). If potential perturbations are known to have a direct impact of the same or opposite direction on several variables, applying such an adjustment to these variables entails no additional risk of spurious agreement between predictions and observations. In all other cases, this risk should also be considered (Puccia and Levins, 1985). For instance, Coll et al. (2019) could not exclude the explanatory

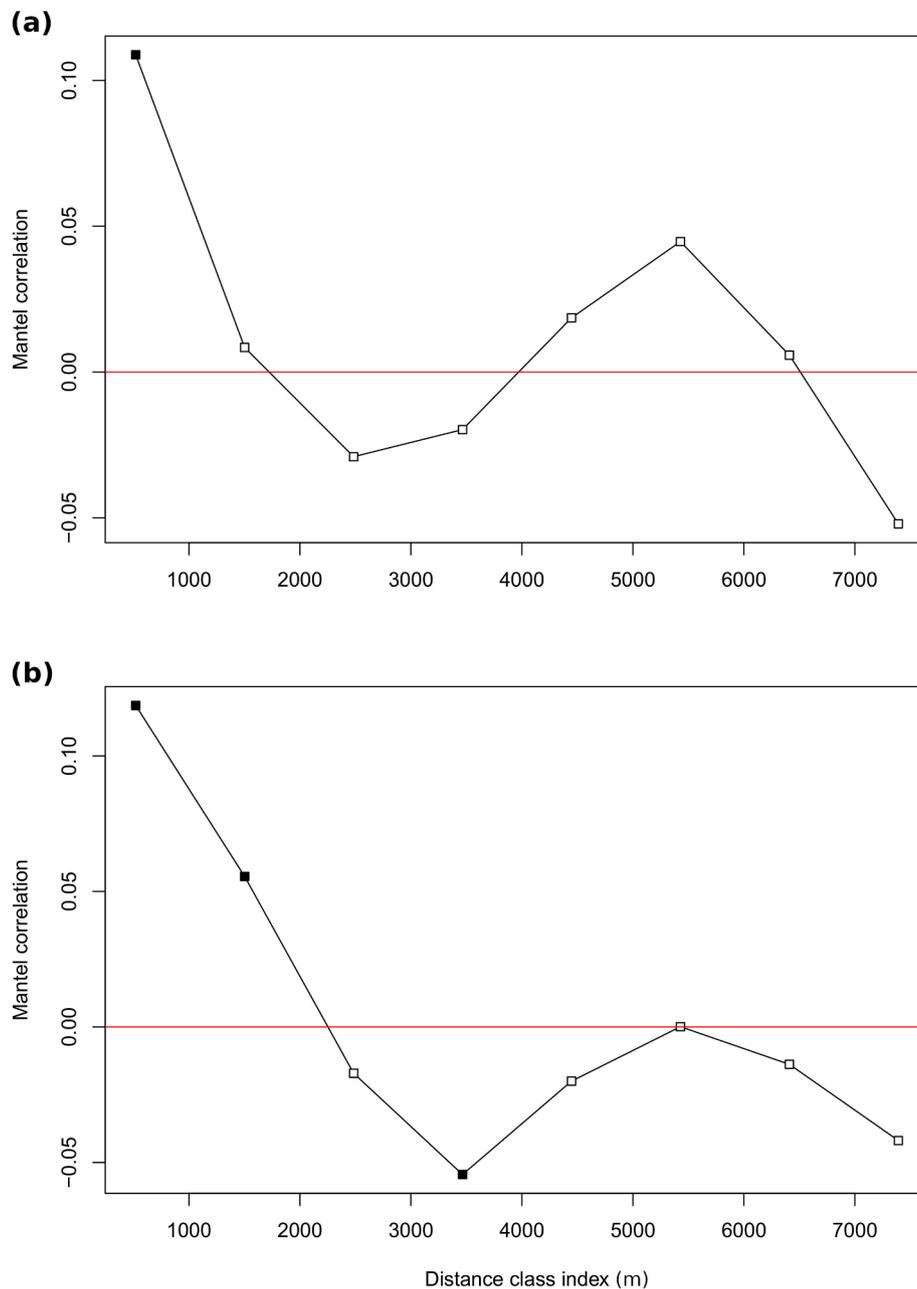


Fig. 3. Correlograms of the Mantel statistic at different distance classes for (a) the epifaunal and (b) the infaunal benthic community samples. Filled squares indicate statistically significant values at 0.05 level.

potential of alternative combinations of input variables when comparing observed system changes with the predictions of qualitative mathematical models. They employed their knowledge of the system and the principle of parsimony to identify the most likely source of input.

Assessing the framework’s assumptions in the Rance estuary is a challenging, but necessary task in lack of formal sensitivity and robustness analysis. We expect such analysis to allow targeted relaxation of assumptions, as we were rather conservative in their definition. Observations of relatively stable composition of benthic communities in the Rance estuary support the assumption of a system near equilibrium (Desroy, 1998). Assumptions of a single gateway variable and adequately small, random pre-perturbation abundance variability are more difficult to assess. Pre-perturbation abundance variability can exceed acceptable levels because of internal system dynamics, such as alternative stable states, or external input forcing organisms in different directions. No such dynamics are predicted by our qualitative models, or

have been observed in the Rance estuary (Alexandridis et al., 2017b). The possibility of external sources of input to act in opposite directions cannot be excluded, but no such known effect appears to be strong enough to mask expected perturbation signals in our results. The assumption of independent community functioning at the sampled sites can be violated by spatial autocorrelation, due to the processes of local dispersal and biogenic habitat modification, but such patterns appear at site distances of around 7 km (Alexandridis et al., 2018), further apart than any significant patterns observed here. Spatial segregation of epifaunal and infaunal communities is the major component of such large-scale patterns. Consequently, our separate analysis of the two communities has probably reduced the respective processes’ signal below levels that could affect the results.

Consequences of violated assumptions can be assessed through validation of the analytical framework, combined with sensitivity and robustness analysis. This task requires application of the framework on

**Table 2**

Kendall's *W* coefficient of concordance between distance matrix  $D_{C1}$  ( $D_{C2}$ ) and each of the  $D_{A1}$  ( $D_{A2}$ ) distance matrices, corresponding to input entering the system through the respective variable of the epifauna (infauna) community model (see Fig. 2). Statistically significant values at 0.05 level are indicated with an asterisk.

Epifauna input variable	Kendall's <i>W</i>	Infauna input variable	Kendall's <i>W</i>
Food	0.41	Food	0.49
Space	0.49	Space	0.49
H1	0.37	H5	0.57
H2	0.46	H6	0.48
H3	0.57	H7	0.67 *
H4	0.56	H8	0.56
H13	0.31	H9	0.64
C1	0.38	H10	0.58
C3	0.76 *	H11	0.49
C5	0.65	H12	0.44
C6	0.40	H14	0.45
–	–	C2	0.51
–	–	C4	0.44
–	–	C6	0.43

biological communities affected by known environmental drivers that vary in space and time. The ability of the framework to detect these drivers can then be assessed, along with the framework's sensitivity to diverse environmental drivers and robustness to different violations of assumptions. These analyses can be applied to systems simulated using mechanistic community models (e.g., Alexandridis et al., 2018), along with a systematic assessment of different community structures (similar to Dambacher et al., 2003; Hosack et al., 2008; Melbourne-Thomas et al., 2012). Alternatively, the framework can be applied to real-world systems with diverse characteristics that have been surveyed repeatedly (e.g., Dauvin, 2000; Boyé et al., 2019), offering both detailed data and good understanding of community functioning.

Ultimately, evaluation of this framework and exploitation of its full potential can be achieved through its integration with other techniques aimed at the analysis of spatial and temporal community observations. For instance, variation partitioning using canonical ordination techniques (Borcard et al., 1992) separates spatial variation in community composition into independent components: pure spatial, pure environmental, spatial component of environmental influence and undetermined. Explicit assignment of observed variation to direct environmental impacts will facilitate assessment of their influence on general assumptions and separation of such impacts from the signal of environmental variability propagating through system structure. Comparison of qualitative predictions with community observations, as illustrated in this study, can thus be the first step towards incorporating complex networks of biotic interactions into ecological variation partitioning. The resulting combined investigation of community structure and environmental variability will increase the ability of researchers across ecological systems to explain typically large undetermined variation in limited community observations. Enriching assessments of environmental impacts with complex feedbacks shaped by diverse biotic interactions, will allow for more comprehensive and, therefore, robust identification and prediction of community responses to perturbation.

#### CRediT authorship contribution statement

**N.A.:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Visualization, Writing - original draft, Writing - review & editing. **C.B.:** Funding acquisition, Supervision, Writing - review & editing. **F.J.:** Funding acquisition, Supervision, Writing - review & editing. **J.M.D.:** Conceptualization, Methodology, Writing - review & editing.

#### 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

Thanks to Scott Foster and Olivier Gauthier for help with the statistical analyses. This work was supported by the LabexMer French project (ANR-10-LABX-19-01).

#### References

- Alexandridis, N., 2017. Models of general community assembly mechanisms simulating the spatial and temporal dynamics of benthic biodiversity. Université de Bretagne Occidentale, Brest, France.
- Alexandridis, N., Bacher, C., Desroy, N., Jean, F., 2018. Individual-based simulation of the spatial and temporal dynamics of macroinvertebrate functional groups provides insights into benthic community assembly mechanisms. PeerJ 6, e5038. <https://doi.org/10.7717/peerj.5038>.
- Alexandridis, N., Bacher, C., Desroy, N., Jean, F., 2017a. Building functional groups of marine benthic macroinvertebrates on the basis of general community assembly mechanisms. J. Sea Res. 121, 59–70. <https://doi.org/10.1016/j.seares.2017.01.007>.
- Alexandridis, N., Dambacher, J.M., Jean, F., Desroy, N., Bacher, C., 2017b. Qualitative modelling of functional relationships in marine benthic communities. Ecol. Model. 360, 300–312. <https://doi.org/10.1016/j.ecolmodel.2017.07.021>.
- Berlow, E.L., Neutel, A.-M., Cohen, J.E., de Ruiter, P.C., Ebenman, B.o., Emmerson, M., Fox, J.W., Jansen, V.A.A., Iwan Jones, J., Kokkoris, G.D., Logofet, D.O., McKane, A. J., Montoya, J.M., Petchey, O., 2004. Interaction strengths in food webs: issues and opportunities. J. Anim. Ecol. 73 (3), 585–598. <https://doi.org/10.1111/jae.2004.73.issue-310.1111/j.0021-8790.2004.00833.x>.
- Beukema, J.J., Essink, K., Dekker, R., 2000. Long-term observations on the dynamics of three species of polychaetes living on tidal flats of the Wadden Sea: the role of weather and predator–prey interactions. J. Anim. Ecol. 69 (1), 31–44. <https://doi.org/10.1046/j.1365-2656.2000.00368.x>.
- Borcard, D., Gillet, F., Legendre, P., 2011. In: Numerical Ecology with R. Springer New York, New York, NY, pp. 227–292. [https://doi.org/10.1007/978-1-4419-7976-6\\_7](https://doi.org/10.1007/978-1-4419-7976-6_7).
- Borcard, D., Legendre, P., Drapeau, P., 1992. Partialling out the Spatial Component of Ecological Variation. Ecology 73, 1045–1055. <https://doi.org/10.2307/1940179>.
- Boyé, A., Thiébaud, É., Grall, J., Legendre, P., Broudin, C., Houbin, C., Le Garrec, V., Maguer, M., Droual, G., Gauthier, O., 2019. Trait-based approach to monitoring marine benthic data along 500 km of coastline. Divers. Distrib. 25 (12), 1879–1896. <https://doi.org/10.1111/ddi.v25.1210.1111/ddi.12987>.
- Campbell, V., Legendre, P., Lapointe, F.-J., 2009. Assessing Congruence Among Ultrametric Distance Matrices. J. Classif. 26 (1), 103–117. <https://doi.org/10.1007/s00357-009-9028-x>.
- Coll, M., Albo-Puigserver, M., Navarro, J., Palomera, I., Dambacher, J.M., 2019. Who is to blame? Plausible pressures on small pelagic fish population changes in the northwestern Mediterranean Sea. Mar. Ecol. Prog. Ser. 617–618, 277–294. <https://doi.org/10.3354/meps12591>.
- Dambacher, J.M., Li, H.W., Rossignol, P.A., 2003. Qualitative predictions in model ecosystems. Ecol. Model. 161 (1–2), 79–93. [https://doi.org/10.1016/S0304-3800\(02\)00295-8](https://doi.org/10.1016/S0304-3800(02)00295-8).
- Dambacher, J.M., Li, H.W., Rossignol, P.A., 2002. Relevance of Community Structure in Assessing Indeterminacy of Ecological Predictions. Ecology 83, 1372–1385. [https://doi.org/10.1890/0012-9658\(2002\)083\[1372:ROCSIA\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[1372:ROCSIA]2.0.CO;2).
- Dauvin, J.-C., 2000. The Muddy Fine Sand Abra alba–Melinna palmata Community of the Bay of Morlaix Twenty Years After the Amoco Cadiz Oil Spill. Mar. Pollut. Bull. 40 (6), 528–536. [https://doi.org/10.1016/S0025-326X\(99\)00242-8](https://doi.org/10.1016/S0025-326X(99)00242-8).
- N. Desroy Les peuplements benthiques de substrats meubles du bassin maritime de la Rance : évolution de la biodiversité et effets de l'activité prédatrice de Nephtys hombergii (annelide polychète) sur le recrutement (thesis) 1998 Rennes 1.
- Desroy, N., Retière, C., 2004. Using benthos as a tool for coastal management: The impact of the tidal power station on benthic communities of the Rance basin. Aquat. Ecosyst. Health Manage. 7 (1), 59–72. <https://doi.org/10.1080/14634980490281263>.
- Donohue, I., Hillebrand, H., Montoya, J.M., Petchey, O.L., Pimm, S.L., Fowler, M.S., Healy, K., Jackson, A.L., Lurgi, M., McClean, D., O'Connor, N.E., O'Gorman, E.J., Yang, Q., Adler, F., 2016. Navigating the complexity of ecological stability. Ecol. Lett. 19 (9), 1172–1185. <https://doi.org/10.1111/ele.12648>.
- Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S.R., Park, H., Shao, C., 2016. Applications of structural equation modeling (SEM) in ecological studies: an updated review. Ecol. Process 5, 19. <https://doi.org/10.1186/s13717-016-0063-3>.
- Giangrande, A., Licciano, M., Musco, L., 2005. Polychaetes as environmental indicators revisited. Mar. Pollut. Bull. 50 (11), 1153–1162. <https://doi.org/10.1016/j.marpolbul.2005.08.003>.
- Gotelli, N.J., Anderson, M.J., Arita, H.T., Chao, A., Colwell, R.K., Connolly, S.R., Currie, D.J., Dunn, R.R., Graves, G.R., Green, J.L., Grytnes, J.-A., Jiang, Y.-H., Jetz, W., Lyons, S.K., McCain, C.M., Magurran, A.E., Rahbek, C., Rangel, T.F.L.V.B., Soberón, J., Webb, C.O., Willig, M.R., 2009. Patterns and causes of species richness:

- a general simulation model for macroecology. *Ecol. Lett.* 12, 873–886. <https://doi.org/10.1111/j.1461-0248.2009.01353.x>.
- Grace, J.B., Anderson, T.M., Olff, H., Scheiner, S.M., 2010. On the specification of structural equation models for ecological systems. *Ecol. Monogr.* 80 (1), 67–87. <https://doi.org/10.1890/09-0464.1>.
- Grime, J.P., 1998. Benefits of plant diversity to ecosystems: immediate, filter and founder effects. *J. Ecol.* 86 (6), 902–910. <https://doi.org/10.1046/j.1365-2745.1998.00306.x>.
- Hayduk, L.A., 2009. Finite Feedback Cycling in Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal* 16, 658–675. <https://doi.org/10.1080/10705510903206030>.
- Héroult, B., 2007. Reconciling niche and neutrality through the Emergent Group approach. *Perspect. Plant Ecol. Evolut. Systemat.* 9 (2), 71–78.
- Hosack, G.R., Hayes, K.R., Dambacher, J.M., 2008. Assessing Model Structure Uncertainty Through an Analysis of System Feedback and Bayesian Networks. *Ecol. Appl.* 18 (4), 1070–1082. <https://doi.org/10.1890/07-0482.1>.
- James, F.C., McCulloch, C.E., 1990. Multivariate Analysis in Ecology and Systematics: Panacea or Pandora's Box? *Annu. Rev. Ecol. Syst.* 21 (1), 129–166. <https://doi.org/10.1146/annurev.es.21.110190.001021>.
- Kirby, R., Retière, C., 2009. Comparing environmental effects of Rance and Severn barrages. *Proceedings of the Institution of Civil Engineers - Maritime Engineering* 162, 11–26. <https://doi.org/10.1680/maen.2009.162.1.11>.
- Le Bec, C., Legendre, A., Messiaen, G., 2016. Changes in the annual harmful algal blooms of *Alexandrium minutum*: effects of environmental conditions and drainage basin inputs in the Rance estuary (Brittany, France). *Aquat. Living Resour.* 29 (1), 104. <https://doi.org/10.1051/alr/2016006>.
- Legendre, P., Lapointe, F.-J., 2004. Assessing Congruence among Distance Matrices: Single-Malt Scotch Whiskies Revisited. *Austral. New Zealand J. Statist.* 46 (4), 615–629. <https://doi.org/10.1111/anzs.2004.46.issue-410.1111/j.1467-842X.2004.00357.x>.
- Leibold, M.A., 1995. The Niche Concept Revisited: Mechanistic Models and Community Context. *Ecology* 76, 1371–1382. <https://doi.org/10.2307/1938141>.
- Levins, R., 1998. Qualitative mathematics for understanding, prediction, and intervention in complex ecosystems. *Ecosyst. Health* 178–204.
- Melbourne-Thomas, J., Wotherspoon, S., Raymond, B., Constable, A., 2012. Comprehensive evaluation of model uncertainty in qualitative network analyses. *Ecol. Monogr.* 82 (4), 505–519. <https://doi.org/10.1890/12-0207.1>.
- Novak, M., Yeakel, J.D., Noble, A.E., Doak, D.F., Emmerson, M., Estes, J.A., Jacob, U., Tinker, M.T., Wootton, J.T., 2016. Characterizing Species Interactions to Understand Press Perturbations: What Is the Community Matrix? *Annu. Rev. Ecol. Evol. Syst.* 47 (1), 409–432. <https://doi.org/10.1146/annurev-ecolsys-032416-010215>.
- Puccia, C.J., Levins, R., 1985. Qualitative modeling of complex systems: an introduction to loop analysis and time averaging. Harvard University Press, Cambridge, Massachusetts.
- Retière, C., 1994. Tidal power and the aquatic environment of La Rance. *Biol. J. Linn. Soc.* 51, 25–36. <https://doi.org/10.1111/j.1095-8312.1994.tb00941.x>.
- Sutherland, W.J., Freckleton, R.P., Godfray, H.C.J., Beissinger, S.R., Benton, T., Cameron, D.D., Carmel, Y., Coomes, D.A., Coulson, T., Emmerson, M.C., Hails, R.S., Hays, G.C., Hodgson, D.J., Hutchings, M.J., Johnson, D., Jones, J.P.G., Keeling, M.J., Kokko, H., Kumin, W.E., Lambin, X., Lewis, O.T., Malhi, Y., Mieszowska, N., Milner-Gulland, E.J., Norris, K., Phillimore, A.B., Purves, D.W., Reid, J.M., Reuman, D.C., Thompson, K., Travis, J.M.J., Turnbull, L.A., Wardle, D.A., Wiegand, T., Gibson, D., 2013. Identification of 100 fundamental ecological questions. *J. Ecol.* 101 (1), 58–67. <https://doi.org/10.1111/jec.2012.101.issue-110.1111/1365-2745.12025>.
- Wernberg, T., Smale, D.A., Thomsen, M.S., 2012. A decade of climate change experiments on marine organisms: procedures, patterns and problems. *Glob. Change Biol.* 18 (5), 1491–1498. <https://doi.org/10.1111/j.1365-2486.2012.02656.x>.