



Original Articles

Artificial intelligence for biodiversity: Exploring the potential of recurrent neural networks in forecasting arthropod dynamics based on time series

Sébastien Lhoumeau^{a,*}, João Pinelo^b, Paulo A.V. Borges^{a,c,d}

^a cE3c-Centre for Ecology, Evolution and Environmental Changes, Azorean Biodiversity Group, CHANGE—Global Change and Sustainability Institute, School of Agricultural and Environmental Sciences, University of the Azores, Rua Capitão João d'Ávila, Pico da Urze, 9700-042 Angra do Heroísmo, Azores, Portugal

^b Atlantic International Research Centre, 9700-702 Angra do Heroísmo, Azores, Portugal

^c IUCN SSC Atlantic Islands Invertebrate Specialist Group, 9700-042 Angra do Heroísmo, Azores, Portugal

^d IUCN SSC Species Monitoring Specialist Group, 9700-042 Angra do Heroísmo, Azores, Portugal

ARTICLE INFO

Keywords:

Conservation scenario
Ecological modelling
Azores arthropods
Machine learning
Long-term ecological data
Model comparison

ABSTRACT

In the current biodiversity crisis, the increasing demand for effective conservation tools aligns with significant advancements in artificial intelligence (AI). There is the need for the development of more robust and accurate forecasting methods, ultimately enhancing our understanding of ecological dynamics and supporting the formulation of effective conservation strategies. This research conducted a comparative analysis of Local Polynomial Regression (LOESS), Seasonal Autoregressive Integrated Moving Average (SARIMA), and Recurrent Neural Network (RNN) models for time-series prediction. Using a unique Long-Term Monitoring Program for island forest arthropods (2012–2023), wherein we selected the 39 most prevalent species collected using SLAM (Sea Land Air Malaise) traps within a native forest fragment on Terceira Island in the Azores archipelago. The results indicate that RNN outperformed LOESS in terms of both goodness of fit and overall accuracy. Although RNN did not surpass classical SARIMA in data prediction, it demonstrated superior goodness-of-fit on the training dataset. Furthermore, we investigated extinction and invasion scenarios within the Terceira arthropod assemblage, providing insight into broader implications and avenues for future research. This study discusses the utility and limitations of RNN models in biodiversity conservation through various scenarios. It contributes to the ongoing discourse at the convergence of conservation, ecology, and artificial intelligence (AI), highlighting advancements and innovative solutions crucial for the effective implementation of conservation strategies.

1. Introduction

The global biodiversity crisis has reached a critical stage (Driscoll et al., 2018), needing urgent action to address the multifaceted challenges. Across all biota and ecosystems, we are witnessing an alarming rate of species extinction (Pievani and Belardinelli, 2023; Urban, 2015), driven primarily by human activities (Borges et al., 2019; Cardoso et al., 2020a; Chichorro et al., 2022; Rønsted et al., 2022; Strona and Bradshaw, 2022). The dramatic decline of biodiversity threatens the multitude of species that are disappearing but also the stability of global ecosystems and the critical services they provide (Bellard et al., 2012; Eisenhauer et al., 2023; Van Genuchten, 2023). Concerted global efforts are needed to reverse this crisis, through active conservation measures,

sustainable practices, and policy changes aligned to the preservation of biodiversity (Shivanna, 2020).

Predicting biodiversity patterns and trends is a highly complex task that fundamentally depends on understanding its temporal variability (Dornelas et al., 2013). Long-term biodiversity monitoring programs are indispensable tools for mapping changes over an extended period (Loh et al., 2005). Time series analysis is of immense value to biodiversity conservation, tracking and monitoring changes in biodiversity over time (Hamilton, 1994; Wauchope et al., 2021). They can reveal underlying ecological dynamics, such as population cycles or responses to climate change (See for example Basset et al.(2023), Fischer et al. (2023) and Lhoumeau & Borges, (2023)). Analysis may provide critical insights into trends and patterns in species populations, enabling scientists to predict

* Corresponding author at: cE3c-Centre for Ecology, Evolution and Environmental Changes, Azorean Biodiversity Group, CHANGE—Global Change and Sustainability Institute, School of Agricultural and Environmental Sciences, University of the Azores, Rua Capitão João d'Ávila, Pico da Urze, 9700-042 Angra do Heroísmo, Azores, Portugal.

E-mail address: seb.lhoumeau@gmail.com (S. Lhoumeau).

<https://doi.org/10.1016/j.ecolind.2025.113119>

Received 29 July 2024; Received in revised form 13 January 2025; Accepted 15 January 2025

Available online 1 February 2025

1470-160X/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

future shifts and potential threats (Dornelas et al., 2014). In addition, the information collected by these programmes can be used to make future predictions and help to design effective conservation strategies (García-Barón et al., 2021) and correctly develop ecological indicators that take into account how ecosystems change over time (see for example Tsafack et al., 2023).

Regression models such as Seasonal Autoregressive Integrated Moving Average (SARIMA) and Local Polynomial Regression (LOESS) have been widely used to analyse and predict ecological dynamics (Fox et al., 2022; Ledolter, 2008; Shyu et al., 1992; Sun et al., 2002; Yap and Musa, 2023). These modelling methods are built on pre-existing assumptions and mathematical formulas, which can potentially lead to biases and limitations. This is especially the case when these models are selected without meticulous consideration or in situations where there is insufficient previous knowledge about the studied system (Dornelas et al., 2013). Prior information can be included in a theory-driven approach like Bayesian analysis. Bayesian methods have been shown to be powerful and flexible tools for ecological modelling and forecasting (Ellison, 2004; Hefley et al., 2017). However, their implementation often requires specific statistical skills and can be time-consuming, which may render them less suitable in situations where resources or expertise are limited. For those trained in Bayesian methods, their complexity and specificity can be seen as strengths, offering robust frameworks for modelling complex ecological systems.

The integration of algorithms that operate without a priori assumptions has emerged as a transformative approach in scientific research, enabling a more expansive exploration of potential scenarios. By setting aside predefined constraints, these algorithms allow data to reveal patterns and relationships that might otherwise remain unexplored. It can broaden the scope of data analysis and serves as a foundation for refining hypotheses and testing other models. The insights obtained from such unconstrained analyses provide a robust framework for guiding further investigation, ensuring that subsequent models are informed by a comprehensive understanding of the underlying data landscape. This paradigm shift underscores the value of exploration as a precursor to hypothesis-driven science, fostering innovation and discovery across diverse disciplines.

Being exclusively data-driven, deep learning algorithms operate in a more 'blind' manner, employing iterative methods to identify features and discern patterns, making them less dependent on prior assumptions (Wu and Feng, 2018). However, while they minimise the reliance on explicit assumptions imposed by the modeler, they remain susceptible to biases intrinsic to the dataset itself. This differentiates them from traditional model-led approaches, where assumptions are more transparent but may also limit model flexibility. Neural networks are a promising modelling tool for investigating complex temporal data because they can approximate any function due to the Universal Approximation Theorem (Hornik et al., 1989). Artificial Intelligence (AI) has recently been recognized as a transformative tool in ecology (Christin et al., 2019), particularly in conservation (Branco et al., 2023), where advancements in computer vision and bioacoustics have proven effective for species identification and ecological interaction analysis (Cardoso et al., 2020; Fujisawa et al., 2023; Hirn et al., 2022; Maglietta et al., 2020; Pichler et al., 2020; Wäldchen and Mäder, 2018). These applications largely rely on extensive datasets, which can be automated through Internet of Things (IoT) technologies (Besson et al., 2022; Kays and Wikelski, 2023; Wang and Bu, 2022). Additionally, AI has been applied in time series analysis and forecasting (Lim and Zohren, 2021; Masini et al., 2023), with studies exploring the strengths and limitations of neural networks in handling temporal data. However, the lack of comparative studies with other methodologies and the scarcity of temporal data—often requiring significant resources for collection (Caughlan and Oakley, 2001)—remain key challenges in leveraging AI for predictive purposes in ecology.

Predicting biodiversity patterns and trends is a highly complex task that fundamentally depends on understanding its temporal variability

(Dornelas et al., 2013). Long-term biodiversity monitoring programs are indispensable tools for mapping changes over an extended period (Loh et al., 2005). The Azores archipelago, located in the North Atlantic Ocean, is a unique island system with distinct conservation challenges (Myers et al., 2000; Neff, 2001). The islands are characterized by high levels of endemism and vulnerability to habitat degradation and biological invasions (Borges et al., 2022; Ferreira et al., 2016; Florencio et al., 2021).

A unique biomonitoring program (SLAM Project – Long Term Ecological Study of the Impacts of Climate Change in the Natural Forest of Azores) has been collecting biodiversity data in the Azores archipelago since 2012. This program conducts seasonal repeated arthropod sampling. It already achieved a comprehensive understanding of biodiversity trends of arthropod assemblages within the native forest of Terceira (Borges et al., 2020, 2017; Lhoumeau and Borges, 2023; Matthews et al., 2019) and inspired the development of new indicators for forest arthropod monitoring and conservation (Tsafack et al., 2023). However, accurate predictions of biodiversity changes are still needed for developing effective conservation strategies in this region.

The aim of this study is twofold. Firstly, it will compare how well Recurrent Neural Networks (RNNs) work with traditional modelling approaches. This will help to improve the field of ecological forecasting. Secondly, it will provide practical advice for conservation in the face of the global biodiversity crisis. The specific aims of the study are as follows: i) to test the power of RNNs in predicting arthropod assemblage dynamics through time using the SLAM Project dataset; and ii) to explore extinction and invasion scenarios to demonstrate how RNN models can inform effective conservation strategies.

2. Materials and methods

2.1. Study area Description

The Azores, an archipelago of nine volcanic islands in the North Atlantic Ocean, lie between 36° 45'N–39° 43'N and 24° 45'W–31° 17'W. With a temperate maritime climate, mild temperatures range from 14 °C in winter to 22 °C in summer, accompanied by high humidity and year-round rainfall that sustain lush landscapes (Cropper and Hanna, 2014). The islands' volcanic soil supports diverse flora and fauna, including unique evergreen forests Fig. 1 (Elias et al., 2016; Keith et al., 2022) and numerous endemic species (Borges et al., 2022), particularly arthropods closely tied to native flora (Rego et al., 2019; Tsafack et al., 2022). This distinctive ecosystem faces significant conservation challenges due to high endemism, habitat degradation, and biological invasions (Fernández-Palacios et al., 2021; Ferreira et al., 2016; Florencio et al., 2021).

2.2. Data collection

The long-term monitoring of arthropods in the Azores has been ongoing since July 2012. This initiative is linked to the EU-NETBIOME-ISLANDBIODIV project (Borges et al., 2018). It is well established in Terceira's native forest fragments, which are the best preserved in the archipelago, but covers only 5.8 % of the island's total area (Triantis et al., 2010) and are now limited to areas above 500 m due to historical human deforestation (Borges et al., 2017).

Ten permanent sampling sites measuring 50 m x 50 m have been set up within this forest (Fig. 2). Each site is equipped with a SLAM (Sea, Land, and Air Malaise) trap, a passive flight-intercept trap measuring 110 × 110 × 110 cm. These traps are designed to capture arthropods from all directions (Skvarla et al., 2021).

The monitoring process involved a periodic seasonal collection with a target frequency of 90 days, allowing each sample to cover one season of the year. Arthropods were sorted into morphospecies by parataxonomists. One of the authors then identified most of them to the species level and classified based on their biogeographical origin

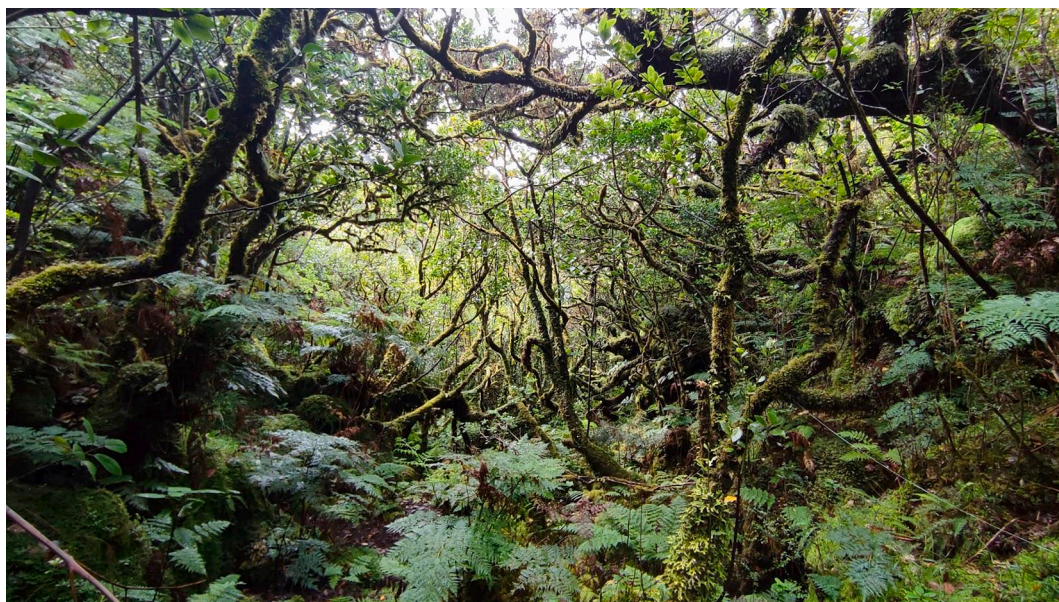


Fig. 1. Typical view within the native forest of the Azores archipelago (Terra-Brava fragment on the island of Terceira).

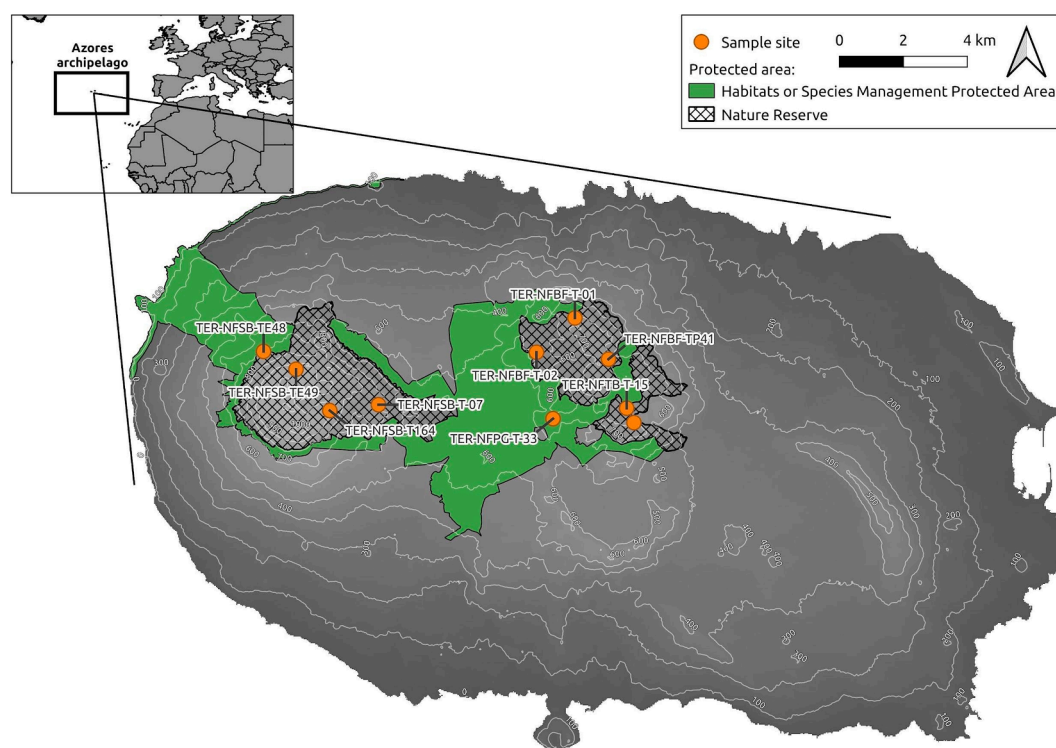


Fig. 2. Map showing the location of the ten sampling sites involved in the long-term monitoring project on Terceira Island.

following the most recent checklist of Azorean arthropods (Borges et al., 2022). Identification was focused on specific groups: Diplopoda (Chordeumatida, Julida), Chilopoda (Geophilomorpha, Lithobiomorpha, Scolopendromorpha), Arachnida (Araneae, Opiliones, Pseudoscorpiones), and Insecta (Archaeognatha, Blattaria, Coleoptera, Hemiptera, Neuroptera, Psocodea, Thysanoptera, Trichoptera, Hymenoptera: Formicidae). The total abundance of each morphospecies was taken, including the distinction between adults and juveniles. Only the abundance of adult arthropods was considered in the following analysis. For more detailed information about the sampling protocol and methodology, refer to the previous data (Costa and Borges, 2021; Lhoumeau et al.,

2022a,b).

2.3. Data Pre-processing

The data was sourced from our previously published open-access databases (Costa and Borges, 2021; Lhoumeau et al., 2022a,b). The data underwent a meticulous process of cleaning and standardisation before being used in this study.

First, we selected a part of the dataset that represented a continuous time series since the start of the long-term monitoring. We considered all seasons from the summer 2013 (sampled in September 2013) to the

spring 2023 (sampled in June 2023). This represented a total of 40 seasons. Sampling events less than 70 days or more than 110 days were excluded to ensure a sampling consistency ($90 \text{ days} \pm 20 \text{ days}$) and arthropod abundances were then proportionally standardised to 90 days. The deviations from the original temporal sampling design are primarily attributable to the adverse weather conditions that prevented the samples from being collected at the scheduled time.

After selecting sampling events, not all ten sites were available for each season, with a minimum of three sites being sampled in the autumn of 2020. To achieve a balanced sampling design, we used a bootstrap method previously developed by Lhoumeau & Borges, 2023. It consisted of an iterative process of seasonally selecting three sites and averaging the standardised abundances. Morphospecies with a total abundance of less than 100 individuals sampled over the 10 consecutive years were removed from the dataset to minimise the possible noise that rarely detected species could introduce in the analysis.

A final step in the data preparation process included a visual assessment of the time series, per species. Species with erratic variations over time were removed from the dataset, which resulted in discarding six morphospecies from the final dataset.

2.4. Data Description

The dataset used in this study represented a continuous time series of 40 seasons covering a time frame from spring 2013 (sampled in June 2013) to spring 2023 (sampled in June 2023). It consisted of adult abundance data for 39 arthropod morphospecies. Detailed information on each of these morphospecies can be found in the supplementary material (Table S1).

75 % of the data was used to train the algorithms (i.e. data from summer 2013 to autumn 2020, $n = 30$ seasons), while the remaining 25 % (from winter 2021 to spring 2023, $n = 10$ seasons, hereafter called test

window) of the data was reserved for testing and validation purposes. This standard split ensured that the model is trained on most of the data, while still having a significant amount available to evaluate its predictive performance.

2.5. Model Descriptions

Three distinctive models were employed: Recurrent Neural Networks (RNN), Seasonal Autoregressive Integrated Moving Average (SARIMA), and Local Polynomial Regression (LOESS). The following Table 1 provides details about each model’s structure and parameters.

RNNs are a type of neural network designed for sequential data, using recurrent connections to capture complex temporal dependencies and patterns, making them effective for non-linear time series. SARIMA, a statistical model, combines autoregressive and moving average components with differencing to address non-stationarity, explicitly modelling seasonal trends and periodic behaviours with high interpretability. LOESS, a non-parametric regression method, fits local polynomial models to subsets of the data, allowing it to adapt flexibly to smooth, non-linear trends without assuming a global functional form.

2.5.1. RNN training and models fitting

The model training was conducted on a 64-bit Ubuntu 22.04.2 LTS platform with an Intel® Core™ i5-5250U CPU at 1.60 GHz, 4.0 GB of memory, and a Mesa Intel® HD Graphics 6000 GPU.

In the RNN training, the input data was a vector of arthropod abundance for a given time step T . The output was a similar vector at the next time step, $T + 1$. The model was trained over 1,000,000 epochs using a gradient descent optimizer.

For SARIMA fitting, we looped over each species abundance and transformed the data into a seasonal time series with a fixed seasonal parameter of 4. For each species, we selected the best SARIMA model

Table 1
Comprehensive overview of the models, highlighting their specifications.

	RNN	SARIMA	LOESS
Model type	Neural networks Non-parametric training Parametric model	Parametric	Non-parametric regression
Mathematical formulation	$y = f(x)$ Succession of 4 computational units: Gated Recurrent Unit (Cho et al., 2014): 39 input, 100 outputs Gated Recurrent Unit: 100 inputs, 100 outputs Fully connected layer: 100 inputs, 100 outputs) Fully connected layer: 100 inputs, 39 outputs Rectified Linear Unit (ReLU) functions are used as activation function within the fully connected layers	$SARIMA(p, d, q)(P, D, Q, s)$ Where: p: The order of the non-seasonal autoregressive component d: The order of non-seasonal differencing, q: The order of the non-seasonal moving average component P: The order of the seasonal autoregressive component D: The order of seasonal differencing Q: The order of the seasonal moving average component s: The number of time steps in each season or the seasonal period	$y_t = f(x_t) + \epsilon_t$ Where: f: The smooth function ϵ_t : The model residuals
Design principle	Process and model sequential data by capturing temporal dependencies using feedback loops.	Model time series data with seasonality and trends using a combination of autoregressive, differencing and moving average components.	Use of weighted local polynomial approximations to estimate the relationship between the independent variable and the dependent variable.
Modelling approach	Data-driven and assemblage based	Data-driven and species based	Data-driven and species based
Underlying assumption	None	Stationarity Linearity Independence Seasonality	Locality Independence No multicollinearity
Allow scenarios explorations	Yes	No	No
Programming language (implemented on)	Julia version 1.9.3 (Bezanson et al., 2017)	R version 4.3.1 (R Core Team, 2023)	R version 4.3.1 (R Core Team, 2023)
Package used	Flux.jl version 0.14.6 (Innes, 2018)	forecast version 8.21.1 (Hyndman and Khandakar, 2008)	stats version 4.3.1 (R Core Team, 2023)
Examples of use	(Aslan et al., 2022; Rew et al., 2019; Salinas et al., 2020)	(Kumar and Rao, 2023; Yap and Musa, 2023)	(Regonda et al., 2005; Yang et al., 2023)

using maximum likelihood estimation, conducting a search over all possible models. We prioritised the selection of non-stationary models. Any negative values returned were set to zero, as negative abundance has no ecological meaning.

Finally, for LOESS fitting, we looped over each species. For each species, we fit a polynomial surface using local fitting by the least-square method. Any negative values returned were set to zero, as they have no ecological meaning.

2.5.2. Models testing

The predictive ability assessment of the three models involves a stepwise process. The assessment of each species is done on an individual basis, following the subsequent steps. The overall accuracy of a model is calculated as the mean of the accuracies for each species.

The first step is to use a specified model (RNN, SARIMA or LOESS) to predict the last 10 seasons for each species. This process results in a 10-dimensional abundance vector for each species.

In the second step, we compare the actual 10-dimensional abundance vector from the test dataset with the predicted abundance vector calculated in step 1, using the following formula:

$$\text{Species accuracy} = \cos\left(\frac{\overrightarrow{\text{Real data}} \cdot \overrightarrow{\text{Model data}}}{\|\overrightarrow{\text{Real data}}\| \cdot \|\overrightarrow{\text{Model data}}\|}\right) \times \left(1 - \frac{\left| \|\overrightarrow{\text{Model data}}\| - \|\overrightarrow{\text{Real data}}\| \right|}{\|\overrightarrow{\text{Model data}}\| + \|\overrightarrow{\text{Real data}}\|}\right)$$

The comparison of the two vectors with this formula produced an output between 0 and 1 representing the degree of similarity between the two vectors of positive values. The criteria used in this formula are based on the collinearity of the vectors and their norm.

There were some exceptions. Given the expectation of seasonality in the abundance data, stationary SARIMA models were not evaluated. Similarly, if the RNN model displayed a constant abundance of zero, the accuracy for that species was not evaluated.

2.5.3. Model comparison

To compare the effectiveness of the three models, we combined our previous accuracy assessment with the computation of the Root-Mean-Square Error (RMSE) for each species in our dataset (Chai and Draxler, 2014). Lower RMSE indicated a higher goodness of fit of a given model. The result of the RMSE calculation does not belong to a finite interval, which complicates its interpretation when applied to a single model. For this reason, it was only used here to compare models.

For each species, we calculated the RMSE between the actual data and the corresponding data predicted by the model under consideration within the full dataset. We then calculated the mean RMSE for each model, which gave us a single representative value for the fitting and predictive performance of the model across all species.

The RMSE and accuracy of the three models were compared using the Kruskal-Wallis test and pairwise using the Wilcoxon test.

2.6. Scenario Designs

The RNN model allows scenarios to be explored as the algorithm uses any input data to predict subsequent data, thereby facilitating the simulation of these scenarios within a given ecological assemblage. The possibilities for scenario exploration are therefore almost unlimited if the scenarios design fit the framework in which the model was trained.

Here, we examined two distinct types of ecological scenario: Single Species Extinction (SSE) and Single Species Invasion (SSI). The abundance vector selected for scenario initialization corresponds to the final one in the dataset, reflecting the current situation in spring 2023. This vector is referred to as the “initial vector”.

In the SSE scenarios, we examined the effects of a single extinction

within the assemblage. For these scenarios, the abundance of the selected species was gradually reduced, removing from 0 % to 100 % (in 10 % increments) of the abundance in the initial vector. All species of the assemblage were tested.

In contrast, the SSI scenarios explored the impacts of an invasion by a single exotic species within the assemblage. In this scenario, the abundance of the chosen exotic species was increased by a factor of 10 in the initial vector.

For both scenario types, we predicted the outcomes for the following 40 seasons, equivalent to the next 10 years. To assess the implications of these scenarios on the species abundance distribution within the assemblage, we fit a GamBin model for each season, as proposed by Ugland et al. (2007) and Matthews et al. (2014). The Gambin alpha (α) parameter is a metric which describes the shape of a species abundance distribution (SAD). Low α values indicate distributions dominated by rare species (logseries-like), while high α values suggest a more even spread of abundances (lognormal-like). It serves as a flexible ecological indicator to compare community structures across different environments or disturbance gradients, providing insights into biodiversity and ecosystem changes with minimal computational complexity (Matthews et al., 2014).

Kruskal-Wallis tests were calculated to compare the mean value of each metric following an SSE or SSI. Pairwise Wilcoxon tests were also performed in the case of a significant result of a Kruskal-Wallis test in order to identify the species that responded differently.

3. Results

3.1. Modelling approach comparison

We observed that modelling approaches presented different RMSE values (Table 2, Kruskal-Wallis chi-squared = 17.301, df = 2, p-value < 0.05). We noticed a gradual decrease in goodness of fit from the RNN to the LOESS (Table 3). Our results indicated that the RNN model performed exceptionally well, both in terms of fitting accuracy and a low standard deviation around this value. This result emphasises that the fitting accuracy is almost identical for all species in this assemblage. The SARIMA approach is the second best in terms of fit accuracy but had a standard deviation twice as high as the mean RMSE, indicating significant differences in fitting accuracy between species. Finally, the LOESS approach showed the worst fitting accuracy and a significant standard deviation (almost twice the mean). This observation highlighted that this approach was unsuccessful in capturing the temporal variation in species abundance and showed significant inconsistencies in fitting accuracy across the dataset.

Similarly, the forecasting accuracy showed statistically difference among the different modelling approaches (Table 2, Kruskal-Wallis chi-squared = 15.499, df = 2, p < 0.05). RNN and SARIMA showed similar performances, where the LOESS had a statistically significant lower accuracy (Table 3). The neural network displayed a greater ability to predict more species, with a total of 36, as compared to the SARIMA, which could only predict 31 species. Despite the differences in prediction capabilities and mean accuracy values, all three modelling approaches exhibited the same standard deviation around the mean. This suggests that the forecasting accuracy of these models is relatively consistent among species considered in the assemblage, regardless of the approach used.

3.2. Scenarios explorations

3.2.1. Scenarios SSE

The results indicate that all SSE scenarios had a similar effect on the GamBin alpha parameter, maintaining equilibrium across seasons with no significant variation in value distribution (Fig. 3). Over the full predicted time series, none of the simulated extinction scenarios altered community structure (Table S2). However, some species, including

Table 2

Comparison of model RMSE and accuracy. RMSE evaluates the goodness of fit of a modelling approach across the entire dataset, and accuracy measures the forecast accuracy of a given model within the test windows. NA indicates that a particular modelling approach failed to fit or train with the species data.

Species name	RMSE LOESS	RMSE RNN	RMSE SARIMA	Accuracy LOESS	Accuracy RNN	Accuracy SARIMA
<i>Acorigone acoreensis</i> (Wunderlich, 1992)	14,43	4,42	8,52	78,14 %	90,36 %	87,05 %
<i>Anaspis proteus</i> Wollaston, 1854	11,37	4,89	9,67	28,28 %	52,01 %	67,31 %
<i>Atheta aeneicollis</i> (Sharp, 1869)	3,95	2,04	2,81	5,33 %	57,69 %	54,52 %
<i>Calacalles subcarinatus</i> (Israelson, 1984)	77,84	5,84	20,95	13,66 %	81,88 %	49,21 %
<i>Catops coracinus</i> Kellner, 1846	5,82	6,58	NA	0,00 %	28,03 %	NA
<i>Cinara juniperi</i> (De Geer, 1773)	14,54	2,98	NA	20,17 %	33,17 %	NA
<i>Cixius azoterceirae</i> Remane & Asche, 1979	334,01	43,01	229,40	65,87 %	93,90 %	79,29 %
<i>Cyphopterum adscendens</i> (Herrich-Schäffer, 1835)	70,07	8,02	26,31	17,99 %	81,15 %	59,50 %
<i>Drouetius borgesii borgesii</i> (Machado, 2009)	11,40	4,54	8,66	1,95 %	64,46 %	24,33 %
<i>Ectopsocus briggsi</i> McLachlan, 1899	6,06	NA	6,04	70,79 %	NA	62,73 %
<i>Elipsocus azoricus</i> Meinander, 1975	4,35	2,81	4,08	56,05 %	55,05 %	22,53 %
<i>Elipsocus brincki</i> Badonnel, 1963	56,03	12,64	75,55	30,94 %	69,64 %	28,15 %
<i>Eupteryx azorica</i> Ribaut, 1941	NA	NA	4,90	NA	NA	43,21 %
<i>Gibbaranea occidentalis</i> Wunderlich, 1989	23,25	5,11	6,46	31,40 %	74,49 %	73,98 %
<i>Hemerobius azoricus</i> Tjeder, 1948	12,57	8,33	NA	34,74 %	66,08 %	NA
<i>Hoplothrips corticis</i> (De Geer, 1773)	6,60	6,16	NA	22,55 %	29,64 %	NA
<i>Kleidocerys ericae</i> (Horváth, 1909)	3,57	3,31	3,70	25,07 %	16,33 %	34,12 %
<i>Lathys denticchelis</i> (Simon, 1883)	16,28	3,14	8,16	33,18 %	76,56 %	62,00 %
<i>Leiobunum blackwalli</i> Meade, 1861	78,22	20,29	NA	27,90 %	41,33 %	NA
<i>Leucognatha acoreensis</i> Wunderlich, 1992	4,31	3,53	5,09	47,39 %	35,91 %	52,86 %
<i>Lithobius pilicornis pilicornis</i> Newport, 1844	13,17	4,12	8,20	15,50 %	39,47 %	38,62 %
<i>Macarocoris cata</i> (Blackwall, 1867)	9,00	3,49	5,28	84,12 %	68,83 %	70,31 %
<i>Monalocoris filicis</i> (Linnaeus, 1758)	7,85	2,93	5,29	21,23 %	34,81 %	46,12 %
<i>Notothecta dryochares</i> (Israelson, 1985)	31,00	15,54	14,85	53,94 %	45,20 %	76,04 %
<i>Ommatoiulus moreleti</i> (Lucas, 1860)	3,15	2,10	2,96	6,40 %	43,96 %	11,44 %
<i>Pinalitus oromii</i> J. Ribes, 1992	52,55	10,06	21,53	58,71 %	96,97 %	89,42 %
<i>Pisaura acoreensis</i> Wunderlich, 1992	9,79	3,43	6,04	27,15 %	29,90 %	39,27 %
<i>Pseudophloeophagus tenax borgesii</i> Stüben, 2022	13,87	1,69	3,56	54,08 %	96,61 %	92,86 %
<i>Rugathodes acoreensis</i> Wunderlich, 1992	35,86	12,34	21,86	75,57 %	74,73 %	83,82 %
<i>Savigniorrhypis acoreensis</i> Wunderlich, 1992	19,25	7,03	19,29	50,23 %	68,29 %	56,21 %
<i>Strophingia harteni</i> Hodkinson, 1981	4,16	1,79	4,08	43,94 %	82,25 %	69,93 %
<i>Tachyporus nitidulus</i> (Fabricius, 1781)	NA	1,96	NA	NA	40,41 %	NA
<i>Tenuiphantes tenuis</i> (Blackwall, 1852)	5,98	1,86	3,84	40,85 %	85,39 %	74,85 %
<i>Trichopsocus clarus</i> (Banks, 1908)	7,24	2,13	NA	53,62 %	53,81 %	NA
<i>Trigoniphthalmus borgesii</i> Mendes, Gaju, Bach & Molero, 2000	98,55	23,23	NA	42,27 %	58,12 %	NA
<i>Trioza laurissilvae</i> Hodkinson, 1990	41,08	14,37	34,88	71,98 %	83,08 %	55,91 %
<i>Valenzuela flavidus</i> (Stephens, 1836)	24,92	7,04	15,65	61,76 %	74,92 %	77,57 %
<i>Walckenaeria grandis</i> (Wunderlich, 1992)	NA	NA	3,22	NA	NA	50,82 %
<i>Zetha simonyi</i> (Krauss, 1892)	12,21	3,99	9,15	26,32 %	43,34 %	46,68 %
Mean value	31,79	7,41	19,35	38,86 %	60,22 %	57,44 %
Standard deviation	57,45	8,05	41,48	22,93 %	22,25 %	21,00 %

Table 3

Outputs of Wilcoxon pairwise comparison tests for RMSE and accuracy.

	Model 1	Model 2	Data model 1	Data model 2	Statistic	p-value	Significance
Accuracy	LOESS	RNN	36	36	329	0.000243	***
	LOESS	SARIMA	36	31	311	0.002	**
	RNN	SARIMA	36	31	594	0.657	ns
RMSE	LOESS	RNN	39	39	1152	6.02e-05	***
	LOESS	SARIMA	39	39	904	0.154	ns
	RNN	SARIMA	39	39	468	0.003	**

Cixius azoterceirae, *Pseudophloeophagus tenax borgesii*, and *Anaspis proteus*, exhibited perturbations distinct from others in the assemblage (Fig. 4). These fluctuations, however, consistently returned to equilibrium regardless of the scenario.

3.2.2. Scenarios SSI

The SSI scenarios revealed that no single scenario had a greater impact than the others, with all producing similar effects on the assemblage (Fig. 5, Kruskal-Wallis test, $n = 160$, $\chi^2 = 2.86$, d.f. = 3, $p = 0.413$). The GamBin alpha value showed no clear trend, indicating stability in assemblage composition over time, even in the presence of invasive species. Notably, the simulated invasions of *Ommatoiulus moreleti* and *Tenuiphantes tenuis* caused deviations similar to those observed in the invasion scenario with *Ectopsocus briggsi*.

4. Discussion

In this study, we show the effectiveness and limits of RNN in forecasting the assemblage of arthropods in the native forest of the Azores archipelago. We compare their performance with traditional modelling approaches (SARIMA and LOESS) and evaluate the accuracy and reliability in predicting biodiversity changes. Moreover, by exploring changes in arthropod assemblages within different scenarios, we emphasize that deep learning algorithms can serve as effective policy-making support tools.

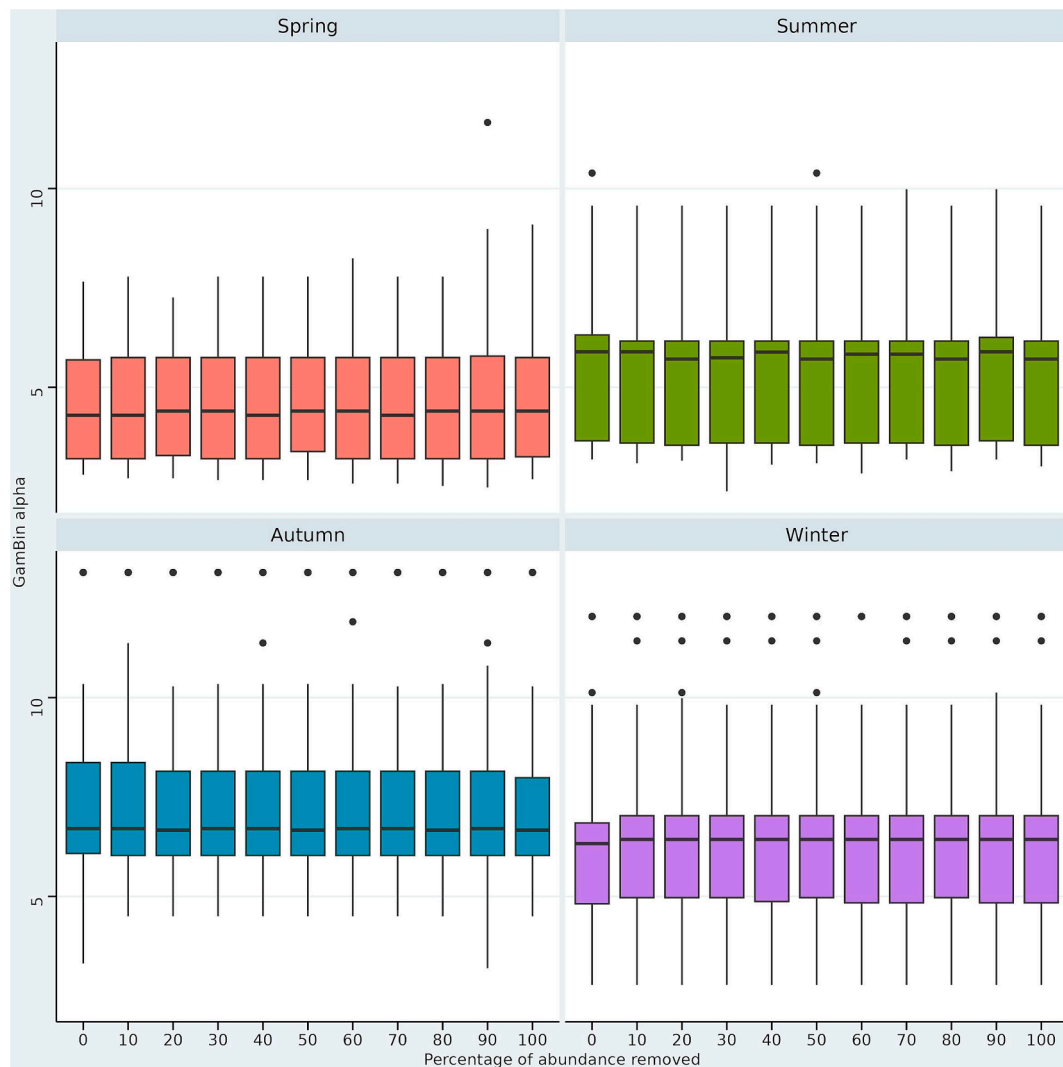


Fig. 3. Box plots depicting the distribution of GamBin alpha values across four seasons: Spring, Summer, Autumn, and Winter. The x-axis indicates the percentage of species abundance data removed (SSE scenarios), while the y-axis shows the GamBin alpha values. The boxes represent the interquartile range (IQR) between the first quartile (Q1) and the third quartile (Q3), with horizontal lines inside the boxes denoting the median (50th percentile) alpha values. Whiskers extend to 1.5 times the IQR from Q1 and Q3, indicating the range of the data excluding outliers, while individual points outside the whiskers represent outliers. Higher alpha values suggest a more log-normal shape of the species abundance distribution curve.

4.1. Recurrent Neural networks performed better in forecasting arthropod assemblage

Our observations indicate that the LOESS model oversimplifies the data, resulting in decreased goodness of fit and accuracy. This is attributed to both the model's definition and the characteristics of the data. Designed for smoothing scatter plots (Avery, 2013; Cleveland and Loader, 1996), LOESS iteratively fits simple polynomial models to local data subsets (Cleveland, 1979). However, its reliance on local data structure makes it less effective with insufficiently dense data sets, such as ours. Despite extensive data cleaning, LOESS remains susceptible to outliers, which can be misconstrued as significant seasonal variations due to the lack of temporal definition in our dataset (Cleveland, 1979). Given the importance of temporal resolution for understanding species dynamics, using only a single season is not feasible due to data limitations. Therefore, we suggest LOESS might be more suited for larger datasets, although obtaining such datasets can be challenging in ecology, particularly for arthropods where diversity and complexity are prominent (Gallé et al., 2022).

On the other hand, SARIMA is specialized in forecasting seasonal time series (Dubey et al., 2021; Yap and Musa, 2023). Models appeared

to fit our data better than LOESS and demonstrated good global mean accuracy. It was expected given the highly seasonal nature of arthropod abundance (Borges et al., 2017; Lhoumeau and Borges, 2023). SARIMA effectively captures both short-term and long-term dependencies in ecological data. One of SARIMA's main advantages lies in its capability to handle non-stationary data common in ecological time series. Through differencing, it removes trends, making the data stationary and enabling more accurate modelling and forecasting (Ghysels et al., 2006). SARIMA can handle many seasonal factors and appear useful for analysing complicated ecological systems with many interacting factors. However, even though SARIMA has been successfully used in the study of animals and plants, it still only looks at different types of animals and plants on their own. This means that it does not take into account how they might interact with each other. This limitation arises from our exclusion of abiotic factors like temperature or relative humidity in the dataset, relying only on long-term arthropod survey data to predict future assemblage. Therefore, although SARIMA models are efficient in fitting our long-term dataset and accurately predicting future species abundance, their species-based approach limits their flexibility in application, especially for scenarios that explore assemblage responses to induced changes in the dataset.

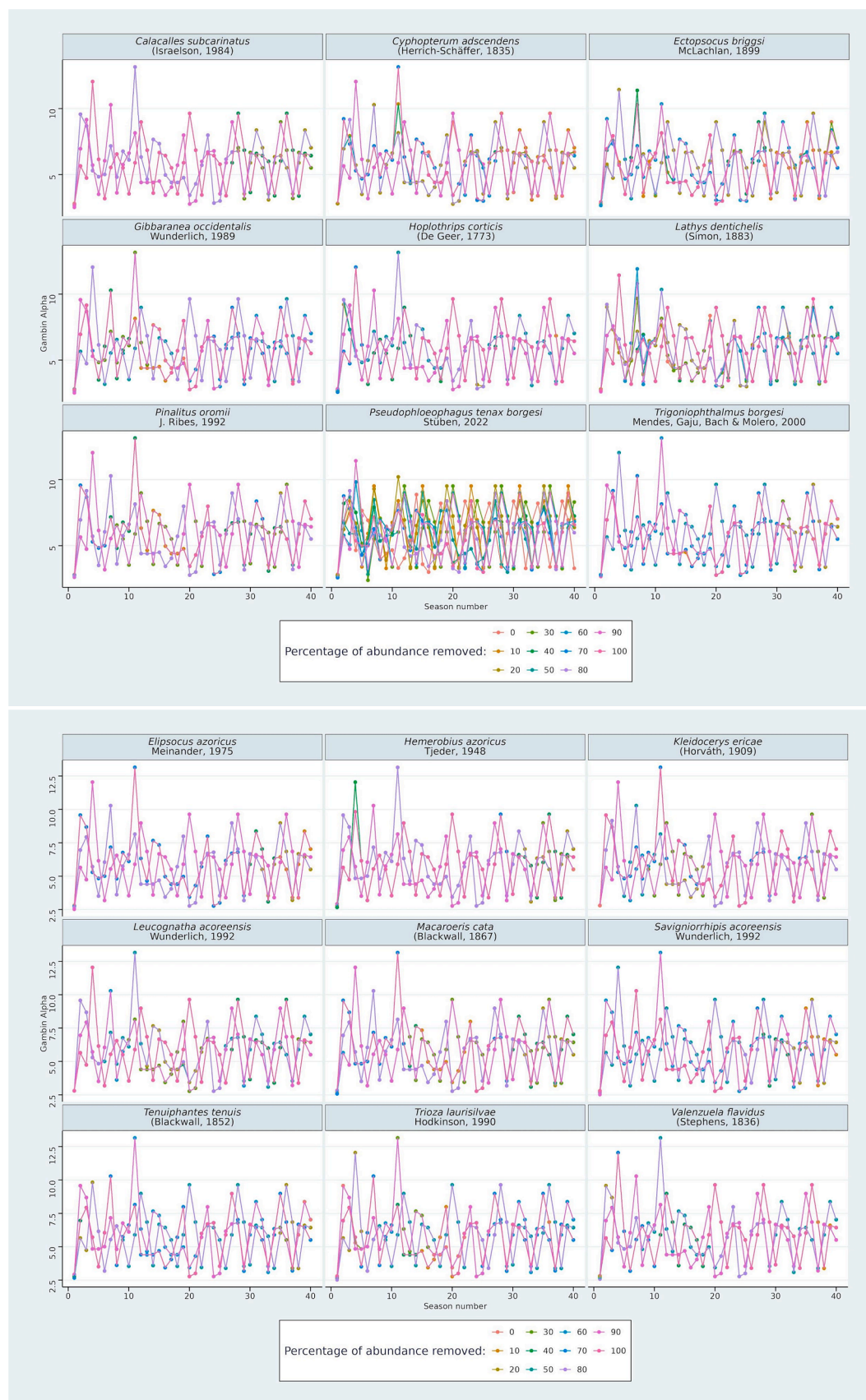


Fig. 4. Time series of the seasonal variation in GamBin alpha values following a given SSE. Each subplot corresponds to a different species, with the x-axis representing the season number and the y-axis showing the GamBin alpha values, which indicate the shape of the species abundance distribution curves. Higher alpha values suggest a more log-normal shape of the species abundance distribution curve.



Fig. 4. (continued).

Finally, the trained RNN model showed significantly better goodness of fit performance compared to LOESS and SARIMA models, while having a similar overall mean accuracy to SARIMA. This high goodness

of fit was expected, as neural networks tend to overfit on training data (Vennerød et al., 2021). The low mean RMSE suggests that the RNN effectively captured complex temporal dynamics after training. RNNs

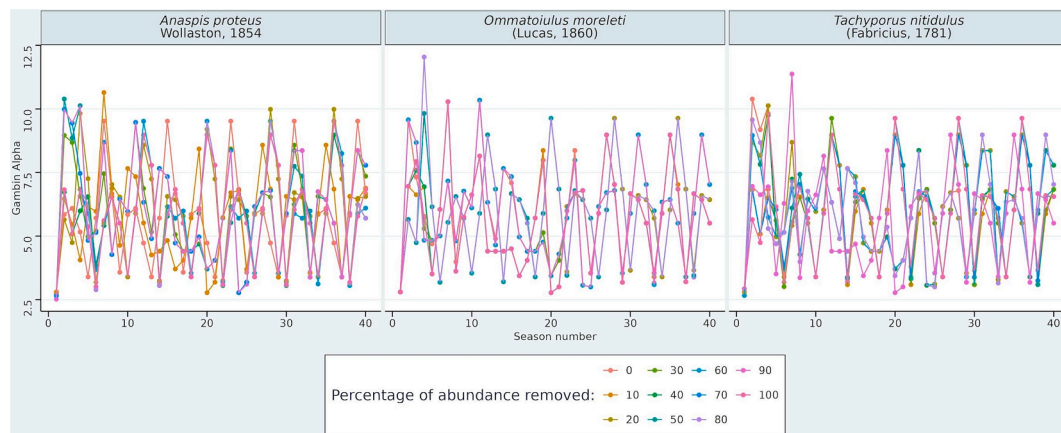


Fig. 4. (continued).

are good at capturing how things happen over time and the order in which they happen, which makes them useful for working out how ecological data is related to each other in complicated ways that change over time. The RNN model recognises patterns, dependencies, and relationships within the training dataset of time series of all arthropods simultaneously. This accounts for how animals of the same species and different species interact, which helps forecasting how assemblage of animals will behave. Recurrent connections in RNNs allow the model to retain a memory of past observations. Models capture both long-term trends and cyclical patterns in arthropod abundance. A single RNN model is used to forecasting arthropod assemblages at each time step. It uses all the information it has had in the past to make an informed prediction. This holistic approach is especially beneficial in ecological studies, where interactions between different species play a crucial role in shaping arthropod communities.

4.2. Quality and size of datasets showed crucial importance

The results highlight the importance of dataset size and quality for accurate ecological modelling, regardless of the approach chosen – SARIMA, LOESS or RNN. In all modelling approaches, data relevance, defined as ecologically meaningful data, is critical for model accuracy and interpretation. This focus on data quality is even more significant for exploratory approaches like RNN, which operate without a priori hypotheses. In such cases, the absence of predefined expectations makes it harder to identify whether the model outputs hold ecological relevance. A high-quality dataset helps mitigate this uncertainty, ensuring that the results are meaningful and grounded in ecological reality. Robust data serves as the foundation for reliable insights, particularly in hypothesis-free exploratory modelling.

The importance of dataset size is clear, as larger datasets offer more comprehensive training material for models, especially in capturing temporal patterns. Seasonal time series with visual consistency are correlated with better model performance. This suggests that the simple maths behind these series makes it easier to understand and predict the data. To achieve effective ecological forecasting, datasets with numerous time steps are required. Long-term biomonitoring programs that use consistent methods are crucial for creating adequate datasets. However, despite global collaborative efforts (Dornelas et al., 2018), data on this topic remain scarce.

Another consideration is the computational power required for developing these models, particularly artificial intelligence models (Christin et al., 2019). Nevertheless, recent advances in software and hardware have made these tools more accessible, expanding their potential for conservation applications (Parker, 2019; Roesch et al., 2023).

Finally, overfitting represents a common challenge when training machine learning models with limited datasets (Vennerød et al., 2021).

This limitation may be attributed to the relatively limited size of the training dataset, which constrains the model's capacity to generalise beyond the training data. The size of the datasets therefore represents a significant consideration. Indeed, preprocessing techniques, such as data cleaning and standardisation, often result in a reduction in the amount of data available. It is therefore recommended that a priori experiments be designed to generate a substantial amount of representative data in a relatively short period of time to mitigate the risk of overfitting.

4.3. Showcasing the usefulness and limits of RNN for the exploration of conservation scenarios

Efficient approaches for effective management are crucial considering the current global biodiversity crisis (Voskamp et al., 2023). Turney et al. (2020) emphasise the urgent need for user-friendly and time-efficient approaches. In this context, we aimed to illustrate the utility of RNN models for scientists, project managers and decision makers. The findings indicate that these models provide essential flexibility in simulating diverse ecological scenarios. We highlight the potential of these models as a systematic method for understanding the impact of SSE or SSI on the diversity of the assemblages.

In the SSE scenarios, the fact that there were no statistically significant impacts on the structure of the arthropod assemblage is promising. This indicates that it is resilient under the modelling conditions that have been considered. Although some species may exhibit distinct variations, this could be attributed to the model's ability to accurately predict species' temporal dynamics.

The extinction scenarios presented in this study are relatively basic. They assume a closed system and constant dynamics over time. It is important to note that the absence of statistically significant changes in the GamBin alpha metric used in this study does not necessarily imply the absence of other alterations. A more comprehensive analysis that includes metrics related to functional diversity, ecosystem service quality, or other interactions could provide further insights, but it is beyond the scope of this study.

However, the SSE scenarios involving the endemic Hemiptera species *Cixius azoterceirae* are particularly interesting. In contrast to scenarios where species' richness remains unchanged, the reduction in abundance of *C. azoterceirae* caused significant changes in the structure patterns of the assemblage. This means that the species is a key species in Terceira's native forest areas. Its loss may lead to potential restructuring of ecological dynamics.

The study of SSI scenarios within the native forest fragments of Terceira provided different insights. Specifically, the non-significant changes in the GamBin alpha metric indicate a stable distribution. This observation presents new research questions concerning the ecological resilience and invasion resistance of the Azorean forest. The

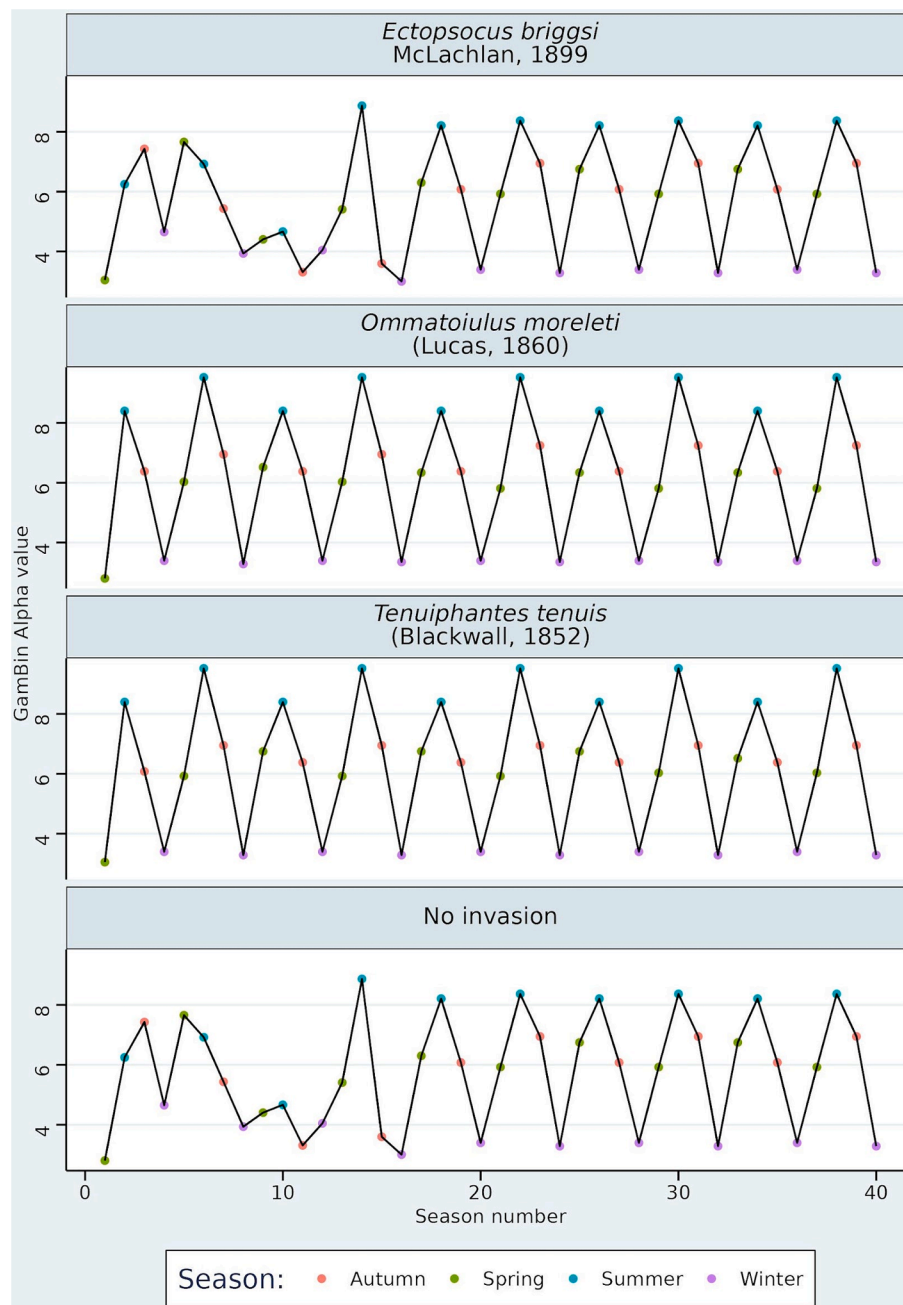


Fig. 5. Time series of the seasonal variation in GamBin alpha values for three species (*Ectopsocus briggsi*, *Ommatoiulus moreleti*, and *Tenuiphantes tenuis*) and a control scenario with no invasion. The x-axis represents the season number, while the y-axis shows the GamBin alpha values, which indicate the shape of the species abundance distribution curves. Points are colored by season: Autumn (red), Spring (green), Summer (blue), and Winter (purple). Higher alpha values suggest a more log-normal shape of the species abundance distribution curve. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

native forest's overall pristine condition and limited presence of exotic species may contribute to its resilience against significant changes in diversity metrics. The phenomenon described aligns with the biotic resistance hypothesis, which suggests that ecosystems with a high number of native species are more resistant to the establishment and proliferation of exotic species (Beaury et al., 2020). It can be argued that native forests are in a "lagged period" of invasion, as demonstrated by Liu et al. (2024). In their study, the authors concluded that native ecosystems tend to experience a longer time between the initial invasion by non-native species and a significant change in the ecosystem.

However, previous literature documents increasing pressure from exotic species in this habitat (Borges et al., 2020). This suggests that

external factors beyond community composition alone need to be considered to fully understand the nuances of invasion dynamics. The presence of exotic species may not have a significant impact on native biodiversity based solely on community composition. The complexity of these interactions is highlighted by the lack of significant changes in diversity metrics despite an increase in exotic species abundance. In our scenarios, exotic species may have light negative impacts, as the native ecosystem may possess resistance characteristics such as limited human disturbance, minimal introduction of exotic species, or specific biotic and abiotic features preventing alien species establishment. Alternatively, similar species abundance distribution can arise from replacement of indigenous species by exotic species which maintain overall

species rank and relative abundances. This can have significant impacts on native forest arthropod communities and should not be neglected (Borges et al., 2020; Lhoumeau and Borges, 2023).

This aspect encourages a more comprehensive study of how species can spread by including outside factors in ecological models. To develop effective conservation and management strategies, it is needed to understand how outside factors influence their population. These data, when available, can be seamlessly integrated into RNN models.

One notable advantage of the use of RNN is their quick application, which provide almost instantaneous results once neural networks are trained and validated. This speed is crucial for making conservation decisions. In addition, the results are not only rapid but also directly interpretable, making it easier to take direct conservation measures. The models are versatile and users can configure the output type based on specific study objectives. This tool allows exploration of a lots of different scenarios, like the simultaneous decline of different species, focusing on functional groups and combining invasions and extinctions. However, it is essential to base simulation choices on relevant ecological arguments for the studied system.

Although RNN models raised as powerful conservation tool, they also have limitations. The accuracy of the model is dependent on the quality and representativeness of the training data. Researchers must carefully evaluate the quality and relevance of the data to ensure the reliability of the model's predictions. In addition, it is crucial to understand the underlying modelling framework to identify assumptions, limitations, and potential biases. This is particularly important due to the 'black box' nature of RNN models, which can pose challenges in terms of interpretability.

Incorporating additional data or factors into the model requires fine-tuning the model. Researchers must weigh the impact of added complexity on model performance. Training RNN models, especially with large datasets, is time-consuming and demands computational resources. Additionally, RNN models are deterministic, meaning that a given situation always produces the same result. Although this approach may be appropriate for straightforward ecological scenarios with limited variables, it raises concerns about the possibility of overlooking confounding factors that have not been considered.

5. Conclusion

Predicting the future in ecological systems remains challenging due to the numerous factors involved. While Recurrent Neural Network (RNN) models may not outperform classical SARIMA in data forecasting, they are distinguished by their flexibility in scenario explorations. Using data of arthropod assemblage within Terceira's native forests, it showed a comprehensive modelling and indirect consideration of potential interactions. This holistic approach positions RNN as an effective decision-making tool for species and spatial conservation. The model's rapid results come with limitations: accuracy, data quality, awareness of the modelling framework and the challenges of interpretability.

As a call to action, our study demonstrates the practical use of RNN in a specific context and encourages researchers to explore its potential in diverse ecosystems. It can advance biodiversity conservation research and decision-making.

CRedit authorship contribution statement

Sébastien Lhoumeau: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **João Pinelo:** Writing – review & editing. **Paulo A.V. Borges:** Writing – original draft, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition, Data curation, 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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ecolind.2025.113119>.

Data availability

Data are available in already published Data papers (cited in the manuscript). Code is available at <https://github.com/SebEyes/SLAM-RNN>.

References

- Aslan, S., Zennaro, F., Furlan, E., Critto, A., 2022. Recurrent neural networks for water quality assessment in complex coastal lagoon environments: A case study on the Venice Lagoon. *Environ. Model. Softw.* 154, 105403. <https://doi.org/10.1016/j.envsoft.2022.105403>.
- Avery, M., 2013. Literature review for local polynomial regression. Unpublished manuscript.
- Basset, Y., Butterill, P.T., Donoso, D.A., Lamarre, P.A., G., Souto-Vilarós, D., Perez, F., Bobadilla, R., Lopez, Y., Alejandro Ramírez Silva, J., Barrios, H., 2023. Abundance, occurrence and time series: long-term monitoring of social insects in a tropical rainforest. *Ecol. Ind.* 150, 110243. <https://doi.org/10.1016/j.ecolind.2023.110243>.
- Beaury, E.M., Finn, J.T., Corbin, J.D., Barr, V., Bradley, B.A., 2020. Biotic resistance to invasion is ubiquitous across ecosystems of the United States. *Ecol. Lett.* 23, 476–482. <https://doi.org/10.1111/ele.13446>.
- Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., Courchamp, F., 2012. Impacts of climate change on the future of biodiversity. *Ecol. Lett.* 15, 365–377. <https://doi.org/10.1111/j.1461-0248.2011.01736.x>.
- Besson, M., Alison, J., Bjerger, K., Gorochowski, T.E., Høye, T.T., Jucker, T., Mann, H.M., R., Clements, C.F., 2022. Towards the fully automated monitoring of ecological communities. *Ecol. Lett.* 25, 2753–2775. <https://doi.org/10.1111/ele.14123>.
- Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B., 2017. Julia: A fresh approach to numerical computing. *SIAM Rev.* 59, 65–98. <https://doi.org/10.1137/141000671>.
- Borges, P., Lamelas-Lopez, L., Andrade, R., Lhoumeau, S., Vieira, V., Soares, A., Borges, I., Boeiro, M., Cardoso, P., Crespo, L.C., Karsholt, O., Schülke, M., Serrano, A., Quartau, J.A., Assing, V., 2022. An updated checklist of Azorean arthropods (Arthropoda). *Biodivers. Data J.* 10, e97682. <https://doi.org/10.3897/BDJ.10.e97682>.
- Borges, P.A.V., Cardoso, P., Fattorini, S., Rigal, F., Matthews, T.J., Di Biase, L., Amorim, I.R., Florencio, M., Borda-de-Agua, L., Rego, C., Pereira, F., Nunes, R., Carvalho, R., Ferreira, M.T., López, H., Pérez Delgado, A.J., Otto, R., Fernández Lugo, S., De Nascimento, L., Caujapé-Castells, J., Casquet, J., Danflous, S., Fournel, J., Sadeyen, A.-M., Elias, R.B., Fernández-Palacios, J.M., Oromí, P., Thébaud, C., Strasberg, D., Emerson, B.C., 2018. Community structure of woody plants on islands along a bioclimatic gradient. *Frontiers of Biogeography* 10, e40295. <https://doi.org/10.21425/F5FBG40295>.
- Borges, P.A.V., Gabriel, R., Fattorini, S., 2019. Biodiversity Erosion: Causes and Consequences, in: Leal Filho, W., Azul, A.M., Brandli, L., Özuyar, P.G., Wall, T. (Eds.), *Life on Land, Encyclopedia of the UN Sustainable Development Goals*. Springer International Publishing, Cham, pp. 1–10. https://doi.org/10.1007/978-3-319-71065-5_78-1.
- Borges, P.A.V., Pimentel, R., Carvalho, R., Nunes, R., Wallon, S., Prieto, A., 2017. Seasonal dynamics of arthropods in the humid native forests of Terceira Island (Azores). *Arquipélago Life and Marine Sciences* 34, 105–122.
- Borges, P.A.V., Rigal, F., Ros-Prieto, A., Cardoso, P., 2020. Increase of insular exotic arthropod diversity is a fundamental dimension of the current biodiversity crisis. *Insect Conserv. Divers.* 13, 508–518. <https://doi.org/10.1111/icad.12431>.
- Branco, V.V., Correia, L., Cardoso, P., 2023. The use of machine learning in species threats and conservation analysis. *Biol. Conserv.* 283, 110091. <https://doi.org/10.1016/j.biocon.2023.110091>.
- Cardoso, P., Barton, P.S., Birkhofer, K., Chichorro, F., Deacon, C., Fartmann, T., Fukushima, C.S., Gaigher, R., Habel, J.C., Hallmann, C.A., Hill, M.J., Hochkirch, A., Kwak, M.L., Mammola, S., Ari Noriega, J., Orfinger, A.B., Pedraza, F., Pryke, J.S., Roque, F.O., Settele, J., Simaika, J.P., Stork, N.E., Suhling, F., Vorster, C., Samways, M.J., 2020a. Scientists' warning to humanity on insect extinctions. *Biol. Conserv.* 242, 108426. <https://doi.org/10.1016/j.biocon.2020.108426>.
- Cardoso, P., Branco, V.V., Borges, P.A.V., Carvalho, J.C., Rigal, F., Gabriel, R., Mammola, S., Cascalho, J., Correia, L., 2020b. Automated Discovery of Relationships, Models, and Principles in Ecology. *Front. Ecol. Evol.* 8. <https://doi.org/10.3389/fevo.2020.530135>.
- Caughlan, L., Oakley, K.L., 2001. Cost considerations for long-term ecological monitoring. *Ecol. Ind.* 1, 123–134. [https://doi.org/10.1016/S1470-160X\(01\)00015-2](https://doi.org/10.1016/S1470-160X(01)00015-2).

- Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* 7, 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>.
- Chichorro, F., Correia, L., Cardoso, P., 2022. Biological traits interact with human threats to drive extinctions: A modelling study. *Eco. Inform.* 69, 101604. <https://doi.org/10.1016/j.ecoinf.2022.101604>.
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation.
- Christin, S., Hervet, É., Lecomte, N., 2019. Applications for deep learning in ecology. *Methods Ecol. Evol.* 10, 1632–1644. <https://doi.org/10.1111/2041-210X.13256>.
- Cleveland, W.S., 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *J. Am. Stat. Assoc.* 74, 829–836. <https://doi.org/10.1080/01621459.1979.10481038>.
- Cleveland, W.S., Loader, C., 1996. *Smoothing by Local Regression: Principles and Methods*. In: Härdle, W., Schimek, M.G. (Eds.), *Statistical Theory and Computational Aspects of Smoothing*. Physica-Verlag HD, Heidelberg, pp. 10–49.
- Costa, R., Borges, P., 2021. SLAM Project - Long Term Ecological Study of the Impacts of Climate Change in the natural forest of Azores: I - the spiders from native forests of Terceira and Pico Islands (2012-2019). *BDJ* 9, e69924. <https://doi.org/10.3897/BDJ.9.e69924>.
- Cropper, T.E., Hanna, E., 2014. An analysis of the climate of Macaronesia, 1865–2012. *Int. J. Climatol.* 34, 604–622. <https://doi.org/10.1002/joc.3710>.
- Dornelas, M., Antão, L.H., Moyes, F., Bates, A.E., Magurran, A.E., Adam, D., Akhmetzhanova, A.A., Appeltans, V., Arcos, J.M., Arnold, H., Ayvayppan, N., Badihi, G., Baird, A.H., Barbosa, M., Barreto, T.E., Bässler, C., Bellgrove, A., Belmaker, J., Benedetti-Cecchi, L., Bett, B.J., Bjorkman, A.D., Blažewicz, M., Blowes, S.A., Bloch, C.P., Bonebrake, T.C., Boyd, S., Bradford, M., Brooks, A.J., Brown, J.H., Brulheide, H., Budy, P., Carvalho, F., Castañeda-Moya, E., Chen, C.A., Chumbley, J.F., Chase, T.J., Siegwart Collier, L., Collinge, S.K., Condit, R., Cooper, E., Cornelissen, J.H.C., Cotano, U., Kyle Crow, S., Damasceno, G., Davies, C.H., Davis, R.A., Day, F.P., Degraer, S., Doherty, T.S., Dunn, T.E., Durigan, G., Duffy, J.E., Edelist, D., Edgar, G.J., Elahi, R., Elmendorf, S.C., Enemar, A., Ernest, S.K.M., Escoribano, R., Estiarte, M., Evans, B.S., Fan, T.-Y., Turini Farah, F., Loureiro Fernandes, L., Farneda, F.Z., Fidelis, A., Fitt, R., Fosaa, A.M., Franco, D.C., G.A., Frank, G.E., Fraser, W.R., García, H., Cazzolla Gatti, R., Givan, O., Gorgone-Barbosa, E., Gould, W.A., Gries, C., Grossman, G.D., Gutierrez, J.R., Hale, S., Harmon, M.E., Harte, J., Haskins, G., Henshaw, D.L., Hermanutz, L., Hidalgo, P., Higuchi, P., Hoey, A., Van Hoey, G., Hofgaard, A., Holec, K., Hollister, R.D., Holmes, R., Hoogenboom, M., Hsieh, C., Hubbell, S.P., Huettmann, F., Hufferd, C.L., Hurlbert, A.H., Macedo Ivanaukas, N., Janík, D., Jandt, U., Jazdzewska, A., Johannessen, T., Johnstone, J., Jones, J., Jones, F.A.M., Kang, J., Kartawijaya, T., Keeley, E.C., Kelt, D.A., Kinnear, R., Klanderud, K., Knutsen, H., Koenig, C.C., Kortz, A.R., Král, K., Kuhn, L.A., Kuo, C.-Y., Kushner, D.J., Laguionie-Marchais, C., Lancaster, L.T., Min Lee, C., Lefcheck, J. S., Lévesque, E., Lightfoot, D., Lloret, F., Lloyd, J.D., López-Baucells, A., Louzao, M., Madin, J.S., Magnússon, B., Malamud, S., Matthews, I., McFarland, K.P., McGill, B., McKnight, D., McLarney, W.O., Meador, J., Meserve, P.L., Metcalfe, D.J., Meyer, C.F. J., Michelsen, A., Milchakova, N., Moens, T., Moland, E., Moore, J., Mathias Moreira, C., Müller, J., Murphy, G., Myers-Smith, I.H., Myster, R.W., Naumov, A., Neat, F., Nelson, J.A., Paul Nelson, M., Newton, S.F., Norden, N., Oliver, J.C., Olsen, E.M., Onipchenko, V.G., Pabis, K., Pabst, R.J., Paquette, A., Paredes, S., Paterson, D.M., Péliissier, R., Peñuelas, J., Pérez-Matus, A., Pizarro, O., Pomati, F., Post, E., Prins, H.T., Priscu, J.C., Provoost, P., Prudic, K.L., Pulliainen, E., Ramesh, B.R., Mendivil Ramos, O., Rassweiler, A., Rebelo, J.E., Reed, D.C., Reich, P.B., Remillard, S.M., Richardson, A.J., Richardson, A.J., van Rijn, L., Rocha, R., Rivera-Monroy, V.H., Rixen, C., Robinson, K.P., Ribeiro Rodrigues, R., de Cerqueira Rossa-Feres, D., Rudstam, L., Ruhl, H., Ruz, C.S., Sampaio, E.M., Rybicki, N., Rypel, A., Sal, S., Salgado, B., Santos, F.A.M., Savassi-Coutinho, A.P., Scanga, S., Schmidt, J., Schooley, R., Setiawan, F., Shao, K.-T., Shaver, G.R., Sherman, S., Sherry, T.W., Sicinski, J., Sievers, C., da Silva, A.C., Rodrigues da Silva, F., Silveira, F.L., Slingsby, J., Smart, T., Snell, S.J., Soudzilovskaia, N.A., Souza, G.B.G., Maluf Souza, F., Castro Souza, V., Stallings, C.D., Stanforth, R., Stanley, E.H., Mauro Sterza, J., Stevens, M., Stuart-Smith, R., Rondon Suarez, Y., Supp, S., Yoshio Tamashiro, J., Tarigan, S., Thiede, G. P., Thorn, S., Tolvanen, A., Teresa Zugliani Toniato, M., Totland, Ø., Twilley, R.R., Vaitkus, G., Valdivia, N., Vallejo, M.I., Valone, T.J., Van Colen, C., Vanaverbeke, J., Venturoli, F., Verhey, H.M., Vianna, M., Vieira, R.P., Vrška, T., Quang Vu, C., Van Vu, L., Waide, R.B., Waldock, C., Watts, D., Webb, S., Wesolowski, T., White, E.P., Widdicombe, C.E., Wilgers, D., Williams, R., Williams, S.B., Williamson, M., Willig, M.R., Willis, T.J., Wipf, S., Woods, K.D., Woehler, E.J., Zawada, K., Zettler, M.L., 2018. BioTIME: A database of biodiversity time series for the Anthropocene. *Glob. Ecol. Biogeogr.* 27, 760–786. <https://doi.org/10.1111/geb.12729>.
- Dornelas, M., Gotelli, N.J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C., Magurran, A. E., 2014. Assemblage Time Series Reveal Biodiversity Change but Not Systematic Loss. *Science* 344, 296–299. <https://doi.org/10.1126/science.1248484>.
- Dornelas, M., Magurran, A.E., Buckland, S.T., Chao, A., Chazdon, R.L., Colwell, R.K., Curtis, T., Gaston, K.J., Gotelli, N.J., Kosnik, M.A., McGill, B., McCune, J.L., Morlon, H., Mumby, P.J., Øvreås, L., Studeny, A., Vellend, M., 2013. Quantifying temporal change in biodiversity: challenges and opportunities. *Proc. R. Soc. B Biol. Sci.* 280, 20121931. <https://doi.org/10.1098/rspb.2012.1931>.
- Driscoll, D.A., Bland, L.M., Bryan, B.A., Newsome, T.M., Nicholson, E., Ritchie, E.G., Doherty, T.S., 2018. A biodiversity-crisis hierarchy to evaluate and refine conservation indicators. *Nat. Ecol. Evol.* 2, 775–781. <https://doi.org/10.1038/s41559-018-0504-8>.
- Eisenhauer, N., Ochoa-Hueso, R., Huang, Y., Barry, K.E., Gebler, A., Guerra, C.A., Hines, J., Jochum, M., Andrzejczak, K., Bucher, S.F., Buscot, F., Ciobanu, M., Chen, H., Junker, R., Lange, M., Lehmann, A., Rillig, M., Römermann, C., Ulrich, J., Weigelt, A., Schmidt, A., Türke, M., 2023. Ecosystem consequences of invertebrate decline. *Curr. Biol.* S0960982223012289. <https://doi.org/10.1016/j.cub.2023.09.012>.
- Elias, R.B., Gil, A., Silva, L., Fernández-Palacios, J.M., Azevedo, E.B., Reis, F., 2016. Natural Zonal Vegetation of the Azores Islands: Characterization and Potential Distribution. *phyto* 46, 107–123. <https://doi.org/10.1127/phyto/2016/0132>.
- Ellison, A.M., 2004. Bayesian inference in ecology. *Ecol. Lett.* 7, 509–520. <https://doi.org/10.1111/j.1461-0248.2004.00603.x>.
- Fernández-Palacios, J.M., Kreft, H., Irl, S.D.H., Norder, S., Ah-Peng, C., Borges, P.A.V., Burns, K.C., de Nascimento, L., Meyer, J.-Y., Montes, E., Drake, D.R., 2021. Scientists' warning – The outstanding biodiversity of islands is in peril. *Global Ecol. Conserv.* 31, e01847. <https://doi.org/10.1016/j.gecco.2021.e01847>.
- Ferreira, M.T., Cardoso, P., Borges, P.A.V., Gabriel, R., de Azevedo, E.B., Reis, F., Araújo, M.B., Elias, R.B., 2016. Effects of climate change on the distribution of indigenous species in oceanic islands (Azores). *Clim. Change* 138, 603–615. <https://doi.org/10.1007/s10584-016-1754-6>.
- Fischer, F.M., Chytrý, K., Chytrý, H., Chytrý, M., Těšitel, J., 2023. Seasonal beta-diversity of dry grassland vegetation: Divergent peaks of above-ground biomass and species richness. *J. Vegetation Science* 34, e13182. <https://doi.org/10.1111/jvs.13182>.
- Florencio, M., Patiño, J., Nogué, S., Traveset, A., Borges, P.A.V., Schaefer, H., Amorim, I. R., Arnedo, M., Ávila, S.P., Cardoso, P., de Nascimento, L., Fernández-Palacios, J.M., Gabriel, S.I., Gil, A., Gonçalves, V., Haroun, R., Illera, J.C., López-Darias, M., Martínez, A., Martins, G.M., Neto, A.L., Nogales, M., Oromí, P., Rando, J.C., Raposeiro, P.M., Rigal, F., Romeiras, M.M., Silva, L., Valido, A., Vanderpoorten, A., Vasconcelos, R., Santos, A.M.C., 2021. Macaronesia as a Fruitful Arena for Ecology, Evolution, and Conservation Biology. *Front. Ecol. Evol.* 9, 718169. <https://doi.org/10.3389/fevo.2021.718169>.
- Fox, H.K., Molina, A.C., Swearingen, T.C., 2022. An interrupted time series approach to assess Marine Protected Area impacts on recreational fishing licence sales. *Aquat. Conserv.* 32, 1970–1982. <https://doi.org/10.1002/aqc.3889>.
- Fujisawa, T., Nogueiras, V., Meramveliotakis, E., Papadopolou, A., Vogler, A.P., 2023. Image-based taxonomic classification of bulk insect biodiversity samples using deep learning and domain adaptation. *Syst. Entomol.* 48, 387–401. <https://doi.org/10.1111/syen.12583>.
- Gallé, R., Tölgyesi, C., Császár, P., Bátor, Z., Gallé-Szpisjak, N., Kaur, H., Maák, I., Torma, A., Batáry, P., 2022. Landscape structure is a major driver of plant and arthropod diversity in natural European forest fragments. *Ecosphere* 13, e3905.
- García-Barón, I., Giakoumi, S., Santos, M.B., Granado, I., Louzao, M., 2021. The value of time-series data for conservation planning. *J. Appl. Ecol.* 58, 608–619. <https://doi.org/10.1111/1365-2664.13790>.
- Ghysels, E., Osborn, D.R., Rodrigues, P.M.M., 2006. Forecasting Seasonal Time Series, in: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), *Handbook of Economic Forecasting*. Elsevier, pp. 659–711. [https://doi.org/10.1016/S1574-0706\(05\)01013-X](https://doi.org/10.1016/S1574-0706(05)01013-X).
- Hamilton, J.D., 1994. *Time series analysis*. Princeton University Press, Princeton, N.J.
- Hefley, T.J., Hooten, M.B., Russell, R.E., Walsh, D.P., Powell, J.A., 2017. When mechanism matters: Bayesian forecasting using models of ecological diffusion. *Ecol. Lett.* 20, 640–650. <https://doi.org/10.1111/ele.12763>.
- Hirn, J., García, J.E., Montesinos-Navarro, A., Sánchez-Martín, R., Sanz, V., Verdú, M., 2022. A deep Generative Artificial Intelligence system to predict species coexistence patterns. *Methods Ecol. Evol.* 13, 1052–1061. <https://doi.org/10.1111/2041-210X.13827>.
- Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. *Neural Netw.* 2, 359–366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- Hyndman, R.J., Khandakar, Y., 2008. Automatic time series forecasting: the forecast package for R. *J. Stat. Softw.* 26, 1–22. <https://doi.org/10.18637/jss.v027.i03>.
- Innes, M., 2018. Flux: Elegant Machine Learning with Julia. *Journal of Open Source Software*. <https://doi.org/10.21105/joss.00602>.
- Kays, R., Wikelski, M., 2023. The Internet of Animals: what it is, what it could be. *Trends Ecol. Evol.* <https://doi.org/10.1016/j.tree.2023.04.007>.
- Keith, D.A., Ferrer-Paris, J.R., Nicholson, E., Bishop, M.J., Polidoro, B.A., Ramirez-Llodra, E., Tozer, M.G., Nel, J.L., Mac Nally, R., Gregr, E.J., Watermeyer, K.E., Essl, F., Faber-Langendoen, D., Franklin, J., Lehmann, C.E.R., Etter, A., Roux, D.J., Stark, J.S., Rowland, J.A., Brummitt, N.A., Fernandez-Arcaya, U.C., Suthers, I.M., Wiser, S.K., Donohue, I., Jackson, L.J., Pennington, R.T., Iliffe, T.M., Gerovasileiou, V., Giller, P., Robson, B.J., Petterelli, N., Andrade, A., Lindgaard, A., Tahvanainen, T., Terauds, A., Chadwick, M.A., Murray, N.J., Moat, J., Plissock, P., Zager, I., Kingsford, R.T., 2022. A function-based typology for Earth's ecosystems. *Nature* 1–6. <https://doi.org/10.1038/s41586-022-05318-4>.
- Dubey, A.K., Kumar, A., García-Díaz, V., Sharma, A.K., Kanhaiya, K., 2021. Study and analysis of SARIMA and LSTM in forecasting time series data. *Sustainable Energy Technol. Assess.* 47, 101474. <https://doi.org/10.1016/j.seta.2021.101474>.
- Kumar, M.T., Rao, M.C., 2023. Studies on predicting soil moisture levels at Andhra Loyola College, India, using SARIMA and LSTM models. *Environ. Monit. Assess.* 195, 1426. <https://doi.org/10.1007/s10661-023-12080-1>.
- Ledolter, J., 2008. Smoothing Time Series with Local Polynomial Regression on Time. *Commun. Stat.- Theory Methods* 37, 959–971. <https://doi.org/10.1080/03610920701693843>.
- Lhoumeau, S., Borges, P.A.V., 2023. Assessing the Impact of Insect Decline in Islands: Exploring the Diversity and Community Patterns of Indigenous and Non-Indigenous Arthropods in the Azores Native Forest over 10 Years. *Diversity* 15, 753. <https://doi.org/10.3390/d15060753>.
- Lhoumeau, S., Cardoso, P., Boieiro, M., Ros-Prieto, A., Costa, R., Lamelas-Lopez, L., Leite, A., Amorim do Rosário, I., Gabriel, R., Malumbres-Olarte, J., Rigal, F., Santos, A., Tsafack, N., Ferreira, M., Borges, P., 2022a. SLAM Project - Long Term

- Ecological Study of the Impacts of Climate Change in the natural forests of Azores: V - New records of terrestrial arthropods after ten years of SLAM sampling. *BDJ* 10, e97952. <https://doi.org/10.3897/BDJ.10.e97952>.
- Lhoumeau, S., Cardoso, P., Costa, R., Boieiro, M., Malumbres-Olarte, J., Amorim, I., Rigal, F., Santos, A., Gabriel, R., Borges, P., 2022b. SLAM Project - Long Term Ecological Study of the Impacts of Climate Change in the natural forest of Azores: IV - The spiders of Terceira and Pico Islands (2019-2021) and general diversity patterns after ten years of sampling. *Biodivers. Data J.* 10, e96442. <https://doi.org/10.3897/BDJ.10.e96442>.
- Lim, B., Zohren, S., 2021. Time-series forecasting with deep learning: a survey. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 379, 20200209. <https://doi.org/10.1098/rsta.2020.0209>.
- Liu, D., Essl, F., Lenzner, B., Moser, D., Semenchuk, P., Blackburn, T.M., Cassey, P., Biancolini, D., Capinha, C., Dawson, W., Dyer, E.E., Guénard, B., Economo, E.P., Kreft, H., Pergl, J., Pyšek, P., van Kleunen, M., Rondinini, C., Seebens, H., Weigelt, P., Winter, M., Purvis, A., Dullinger, S., 2024. Regional invasion history and land use shape the prevalence of non-native species in local assemblages. *Global Change Biol.* 30, e17426. <https://doi.org/10.1111/gcb.17426>.
- Loh, J., Green, R., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V., Randers, J., 2005. The Living Planet Index: Using species population time series to track trends in biodiversity. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences* 360, 289–295. <https://doi.org/10.1098/rstb.2004.1584>.
- Maglietta, R., Reno, V., Caccioppoli, R., Seller, E., Bellomo, S., Santacesaria, F.C., Colella, R., Cipriano, G., Stella, E., Hartman, K., Fanizza, C., Dimauro, G., Carlucci, R., 2020. Convolutional Neural Networks for Risso's Dolphins Identification. *IEEE Access* 8, 80195–80206. <https://doi.org/10.1109/ACCESS.2020.2990427>.
- Masini, R.P., Medeiros, M.C., Mendes, E.F., 2023. Machine learning advances for time series forecasting. *J. Econ. Surv.* 37, 76–111. <https://doi.org/10.1111/joes.12429>.
- Matthews, T.J., Borregaard, M.K., Ugland, K.I., Borges, P.A.V., Rigal, F., Cardoso, P., Whittaker, R.J., 2014. The gambin model provides a superior fit to species abundance distributions with a single free parameter: evidence, implementation and interpretation. *Ecography* 37, 1002–1011. <https://doi.org/10.1111/ecog.00861>.
- Matthews, T.J., Sadler, J., Carvalho, R., Nunes, R., Borges, P.A.V., 2019. Differential temporal beta-diversity patterns of native and non-native arthropod species in a fragmented native forest landscape. *Ecography* 42, 45–54. <https://doi.org/10.1111/ecog.03812>.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858. <https://doi.org/10.1038/35002501>.
- Neff, C., 2001. The Azores - a forgotten biodiversity hotspot. *Geoöko XXII* 189–201.
- Perkel, J.M., 2019. Julia: come for the syntax, stay for the speed. *Nature* 572, 141–142. <https://doi.org/10.1038/d41586-019-02310-3>.
- Pichler, M., Boreux, V., Klein, A., Schleuning, M., Hartig, F., 2020. Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. *Methods Ecol. Evol.* 11, 281–293. <https://doi.org/10.1111/2041-210X.13329>.
- T. Pievani S. Belardinelli M. Extinction . in: Wallenhorst, N., Wulf, C. Handbook of the Anthropocene 2023 Springer International Publishing Cham 263 269 10.1007/978-3-031-25910-4 41.
- R Core Team, 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rego, C., Boieiro, M., Rigal, F., Ribeiro, S.P., Cardoso, P., Borges, P.A.V., 2019. Taxonomic and functional diversity of insect herbivore assemblages associated with the canopy-dominant trees of the Azorean native forest. *PLoS One* 14, e0219493. <https://doi.org/10.1371/journal.pone.0219493>.
- Regonda, S., Rajagopalan, B., Lall, U., Clark, M., Moon, Y.-I., 2005. Local polynomial method for ensemble forecast of time series. *Nonlinear Processes Geophys.* 12, 397–406. <https://doi.org/10.5194/npg-12-397-2005>.
- Rew, J., Park, S., Cho, Y., Jung, S., Hwang, E., 2019. Animal Movement Prediction Based on Predictive Recurrent Neural Network. *Sensors* 19, 4411. <https://doi.org/10.3390/s19204411>.
- Roesch, E., Greener, J.G., MacLean, A.L., Nassar, H., Rackauckas, C., Holy, T.E., Stumpf, M.P.H., 2023. Julia for biologists. *Nat. Methods* 20, 655–664. <https://doi.org/10.1038/s41592-023-01832-z>.
- Rønsted, N., Walsh, S., Clark, M., Edmonds, M., Flynn, T., Heintzman, S., Loomis, A., Lorence, D., Nagendra, U., Nyberg, B., Opgenorth, M., Weisenberger, L., Williams, A., Wolkis, D., Wood, K., Keir, M., 2022. Extinction risk of endemic vascular flora of Kauai Island, Hawaii, based on IUCN assessments. *Conserv. Biol.* 36. <https://doi.org/10.1111/cobi.13896>.
- Salinas, D., Flunkert, V., Gasthaus, J., Januschowski, T., 2020. DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *Int. J. Forecast.* 36, 1181–1191. <https://doi.org/10.1016/j.ijforecast.2019.07.001>.
- Shivanna, K.R., 2020. The Sixth Mass Extinction Crisis and its Impact on Biodiversity and Human Welfare. *Reson* 25, 93–109. <https://doi.org/10.1007/s12045-019-0924-z>.
- Shyu, W.S.C., Eric Grosse, William M., 1992. Local Regression Models, in: *Statistical Models in S*. Routledge.
- Skvarla, M.J., Larson, J.L., Fisher, J.R., Dowling, A.P.G., 2021. A Review of Terrestrial and Canopy Malaise Traps. *Ann. Entomol. Soc. Am.* 114, 27–47. <https://doi.org/10.1093/aesa/saaa044>.
- Strona, G., Bradshaw, C.J.A., 2022. Coextinctions dominate future vertebrate losses from climate and land use change. *Sci. Adv.* 8, eabn4345. <https://doi.org/10.1126/sciadv.abn4345>.
- Sun, H., Liu, H.X., Xiao, H., Ran, B., 2002. Short Term Traffic Forecasting Using the Local Linear Regression Model. *Center for Traffic Simulation Studies, UC Irvine*.
- Triantis, K.A., Borges, P.A.V., Ladle, R.J., Hortal, J., Cardoso, P., Gaspar, C., Dinis, F., Mendonça, E., Silveira, L.M.A., Gabriel, R., Melo, C., Santos, A.M.C., Amorim, I.R., Ribeiro, S.P., Serrano, A.R.M., Quarta, J.A., Whittaker, R.J., 2010. Extinction debt on oceanic islands. *Ecography* 33, 285–294. <https://doi.org/10.1111/j.1600-0587.2010.06203.x>.
- Tsafack, N., Gabriel, R., Elias, R., Boieiro, M., Ferreira, M.T., Borges, P., 2022. Arthropods and other biota associated with the Azorean trees and shrubs: *Laurus azorica* (Seub) Franco (Magnoliophyta, Magnoliopsida, Laurales, Lauraceae). *BDJ* 10, e80088. <https://doi.org/10.3897/BDJ.10.e80088>.
- Tsafack, N., Lhoumeau, S., Ros-Prieto, A., Navarro, L., Kocsis, T., Manso, S., Figueiredo, T., Teresa Ferreira, M., Borges, P.A.V., 2023. Arthropod-based biotic integrity indices: A novel tool for evaluating the ecological condition of native forests in the Azores archipelago. *Ecol. Ind.* 154, 110592. <https://doi.org/10.1016/j.ecolind.2023.110592>.
- Turney, C., Ausseil, A.-G., Broadhurst, L., 2020. Urgent need for an integrated policy framework for biodiversity loss and climate change. *Nat. Ecol. Evol.* 4, 996. <https://doi.org/10.1038/s41559-020-1242-2>.
- Ugland, K.I., Lamshead, P.J.D., McGill, B., Gray, J.S., O'Dea, N., Ladle, R.J., Whittaker, R.J., 2007. Modelling dimensionality in species abundance distributions: description and evaluation of the Gambin model. *Evol. Ecol. Res.* 9, 313–324.
- M.C. Urban Accelerating extinction risk from climate change *Science* 348 2015 571 573 <https://doi.org/10.1126/science.1257573>
- Van Genuchten, E., 2023. How Biodiversity Impacts Our Health, in: *A Guide to a Healthier Planet*. Springer Nature Switzerland, Cham, pp. 107–111. https://doi.org/10.1007/978-3-031-34479-4_13.
- Vennerod, C.B., Kjær, A., Bugge, E.S., 2021. Long Short-Term Memory RNN. <https://doi.org/10.48550/arXiv.2105.06756>.
- Voskamp, A., Fritz, S.A., Köcke, V., Biber, M.F., Nogueira Brockmeyer, T., Bertzy, B., Forrest, M., Goldstein, A., Henderson, S., Hickler, T., Hof, C., Kastner, T., Lang, S., Manning, P., Mascia, M.B., McFadden, I.R., Niamir, A., Noon, M., O'Donnell, B., Opel, M., Schwede, G., West, P., Schenck, C., Böhning-Gaese, K., 2023. Utilizing multi-objective decision support tools for protected area selection. *One Earth* 6, 1143–1156. <https://doi.org/10.1016/j.oneear.2023.08.009>.
- Wäldchen, J., Mäder, P., 2018. Machine learning for image based species identification. *Methods Ecol. Evol.* 9, 2216–2225. <https://doi.org/10.1111/2041-210X.13075>.
- Wang, J., Bu, Y., 2022. Internet of Things-based smart insect monitoring system using a deep neural network. *IET Networks* 11, 245–256. <https://doi.org/10.1049/ntw2.12046>.
- Wauchope, H.S., Amano, T., Geldmann, J., Johnston, A., Simmons, B.I., Sutherland, W.J., Jones, J.P.G., 2021. Evaluating Impact Using Time-Series Data. *Trends Ecol. Evol.* 36, 196–205. <https://doi.org/10.1016/j.tree.2020.11.001>.
- Wu, Y., Feng, J., 2018. Development and Application of Artificial Neural Network. *Wires. Pers. Commun.* 102, 1645–1656. <https://doi.org/10.1007/s11277-017-5224-x>.
- Yang, X., Zhang, X., Zhang, P., Bidegain, G., Dong, J., Hu, C., Li, M., Zhang, Z., Guo, H., 2023. Ensemble habitat suitability modeling for predicting optimal sites for eelgrass (*Zostera marina*) in the tidal lagoon ecosystem: Implications for restoration and conservation. *J. Environ. Manage.* 330, 117108. <https://doi.org/10.1016/j.jenvman.2022.117108>.
- Yap, Z.N., Musa, S., 2023. Stream Flow Forecasting on Pahang River by Time Series Models, ARMA, ARIMA and SARIMA. *Rtcebe* 4, 331–341.