

1 MicroxanoX: an R package for simulating an aquatic
2 *MIC*Robial ecosystem that can occupy *OX*ic or
3 *ANOX*ic states.

4 Rainer M Krug^a, Owen L. Petchey^a

5 ^a*Department of Evolutionary Biology and Environmental Studies, Winterthurerstrasse 190,*
6 *8057 Zurich*

7 **Abstract**

Ca. 100 words.

8 *Keywords:* reproducibility; regime shift;

9 **1. TODO**

- 10 • **update links**
11 • **update cross references**
12 • **check keywords**
13 • **add more references**
14 • **expand on Impact and Conclusion section**

*Corresponding Author

**Equal contribution

Email addresses: `Rainer.Krug@uzh.ch`; `Rainer@krugs.de` (Rainer M Krug),
`Owen.Petchey@ieu.uzh.ch` (Owen L. Petchey)

15 **2. Required Metadata**

16 *2.1. Current code version*

17 Ancillary data table required for subversion of the codebase.

Nr.	Code metadata description	Please fill in this column
C1	Current code version	v0.9.0
C2	Permanent link to code/repository used for this code version	https://github.com/UZH-PEG/microxanox
C3	Code Ocean compute capsule	
C4	Legal Code License	CC BY 4.0
C5	Code versioning system used	git
C6	Software code languages, tools, and services used	R
C7	Compilation requirements, operating environments	R ($\geq 4.1.0$) magrittr tibble ggplot2 patchwork grDevices stats mgcv deSolve dplyr tidyr stringr multidplyr
C8	If available Link to developer documentation/manual	TO BE ADDED
C9	Support email for questions	Rainer.Krug@uzh.ch; Rainer@krugs.de

18 **3. Motivation and significance**

19 Mathematical models play a key role in the development of understanding
 20 about how ecosystems work and how they respond to environmental changes
 21 (Binzer et al., 2016; Chaparro-Pedraza, 2021; Vasseur and McCann, 2005).

22 They are also important for developing hypotheses that can then be tested in
23 empirical studies. One area of ecology in which simple models have played
24 a influential role is the area of how ecosystems respond to gradual change in
25 an environmental driver (Scheffer et al., 2001). An environmental driver is a
26 environmental condition that affects an ecosystem, but is assumed to not be
27 affected by the ecosystem, such as the rate of nutrient input into a lake.

28 It is conceivable that an ecosystem state, such as the total biomass of a
29 particular type of bacteria, may remain unchanged when an environmental driver
30 changes. It is also possible that the ecosystem state changes gradually. It is
31 also possible that the ecosystem state changes abruptly to a new state that
32 is difficult to recover from (Scheffer et al., 2001). This possibility for abrupt,
33 perhaps catastrophic changes that are difficult to reverse causes considerable
34 concern (Collins et al., 2021; Northrop et al., 2021; Vandermeer and Perfecto,
35 2019).

36 An example where a gradual change of an environmental variable causes an
37 abrupt change of the system is the switch from an aerobic (oxygen is available
38 for metabolism) to anaerobic (oxygen generally unavailable) state in a microbial
39 ecosystem. Three types of microbes occur in the model: cyanobacteria (CB)
40 dominating the oxic state, and two types of sulfur bacteria that dominate the
41 anoxic state (sulfate reducing bacteria (SB) and phototrophic sulfur bacteria
42 (PB)). This system has been investigated by Bush et al. (2017) in a simulation
43 study of a mathematical model. They showed that gradual change in the rate at
44 which oxygen could diffuse into the ecosystem (termed the oxygen diffusivity)
45 could cause catastrophic changes in the ecosystem state that would be difficult
46 to reverse.

47 One feature of the study by Bush et al. (2017) was limited biodiversity.
48 Specifically, there was no biodiversity within each of the three types of bacteria.
49 This leaves open the question of if and how biodiversity within these types
50 (i.e. functional groups) of bacteria, in their model, affects the ecosystem response
51 to environmental change. This limitation is not specific to the study of Bush et
52 al. (2017), there are few if any studies of the effects of biodiversity on abrupt
53 transitions between ecosystem states.

54 We decided to fill this research gap by making a simulation study of how
55 within functional group biodiversity affects ecosystem responses to environmental
56 change Limberger (2022), and to base our work on the work and model of Bush
57 et al. (2017). It was with this goal in mind that we developed the *microxanox*
58 package. The first stage of development was to write code from scratch (as there
59 was no available code to start from) and to confirm that this new implementation
60 would reproduce the previously published results. The resulting reproduction is
61 available as one of the package vignettes: [vignette Partial reproduction of Bush](#)
62 [et al.](#)

63 The second stage was to add functionality that would be necessary to answer
64 our research question. Most importantly, we made it possible to have multiple
65 species of bacteria within each of the three functional groups, for the multiple
66 species to differ in their characteristics, and to vary the number of species and
67 amount of variability among them. We also added functionality that allowed:

temporally varying environmental conditions, addition of random noise to state variables, and immigration. In addition to the model itself, the package includes some functions to analyse the results as well as visualize the results to provide a starting point for customized visualizations based on own requirements. The basic and additional functionality is described in the package [User Guide](#).

4. Software description

{»Describe the software in as much as is necessary to establish a vocabulary needed to explain its impact.«}

The *microxanoX* package is for simulating a three functional group system (*CB*: cyanobacteria, *PB*: phototrophic sulfur bacteria, and *SB*: sulfate-reducing bacteria) with four chemical substrates (*P*: phosphorus, *O*: oxygen, *SR*: reduced sulfur, *SO*: oxidized sulfur). It includes feedback between organisms and biogeochemical processes and is based on [Bush et al. \(2017\)](#) (See [Bush et al. \(2017\)](#) for a detailed discussion of the model). At the core of the simulations is a set of ordinary differential equations (specified in the function `bushplus_dynamic_model()`, though this function need not be directly called). There are functions for running individual simulations and for running a set of simulations across, for example, a range of environmental conditions.

To make the simulation run with multiple species, we expressed different species characteristics in the elements of vectors and matrices. We also coded the ordinary differential equations to include the vectors and matrices, and to use matrix mathematics. In this way, we made it possible to run simulations with different numbers of species without having to change the underlying code.

The package functions and code have modular structure, so that new functionality can be easily added. E.g. temporally defined events of any type could be specified. Further, all parameter values required to run a simulation are stored in one object. Lastly, the general structure of the code should make it straightforward to adapt the model to other similar systems (described in more detail in the Impact section).

4.1. Software architecture

{»Give a short overview of the overall software architecture; provide a pictorial component overview or similar (if possible). If necessary provide implementation details.«}

The framework used when writing this package aimed to maximise simplicity for the user, and to make it easy to reproduce results (see the supplement to [Limberger \(2022\)](#) for an example of how this is used). As such, all the parameters needed to run a simulation or find a stable state (i.e. the final state of the ecosystem) are contained in a single object (which can easily be created using included functions). This parameter object is given to a function that runs the simulations and returns the results. The returned results object is identical to the parameter object but with an additional slot named `results`,

109 which contains the results of the run. Thus the returned results object contains
110 the simulation conditions (parameters) as well as the results, and can be used to
111 run the simulation again. This promotes reproducibility and makes incremental
112 changes of individual parameters and re-running the simulations straightforward.

113 In the following sections we describe how to use the package to run one
114 simulation and to find steady states across an environmental gradient.

115 4.2. Running one simulation

116 A typical simulation would look as shown in @ref(fig:runsim_example).

117 A simulation is run using the `run_simulation()` function. In this function,
118 the ODEs are solved using the function `ode()` in the package *deSolve* package
119 (Soetaert et al., 2010). The `run_simulation()` function needs only one argu-
120 ment - an object as created by the function `new_runsim_parameter()`. The
121 parameter object returned by `new_runsim_parameter()` contains among other
122 things the `strain_parameter` object, which can be created by the function
123 `new_strain_parameter()`. For a detailed description of the parameter objects,
124 their meaning and how they are created and have values set and changed please
125 see the *User Guide* which accompanies the package or is available at [User Guide](#)

126 After the parameter object has been defined, it can be used in the `run_simulation()`
127 function. The function returns an object which is identical to the parameter
128 object, except of an additional slot containing the results. This design produces
129 a fully reproducible object as it can be used as a parameter object to be fed
130 back into the `run_simulation()` function to run the simulation again from the
131 parameter used to generate the results before.

132 4.3. Finding Stable States

133 The general approach used to find the stable state of the system with a
134 specific parameter set is to run the simulation for a long time and record the
135 final state. When one does this across a range of environmental conditions, one
136 discovers how the steady state of the system responds to the environmental
137 conditions. The package contains functionality for finding steady states that
138 correspond to values of one environmental driver, namely the value of oxygen
139 diffusivity.

140 Two methods for finding steady states are implemented. The first runs a
141 separate simulation for each combination of starting conditions and oxygen
142 diffusivity (we term this the *Replication method*). This is the method used in
143 the Bush et al. (2017) study. The second runs two simulations, with step-wise
144 and slowly temporally increasing oxygen diffusivity, and the other with step-wise
145 and slowly decreasing or decreasing oxygen diffusivity. During this temporal
146 environmental change, the state of the system was recorded just before change to
147 a new oxygen diffusivity (we term this the *Temporal method*). We implemented
148 two methods since there is no definitive best method, and in order to check if
149 results were sensitive to choice of method.

150 The replication method is implemented in the function `run_replication_ssfind()`
151 which takes a parameter object as returned by the function `new_replication_ssfind_parameter()`

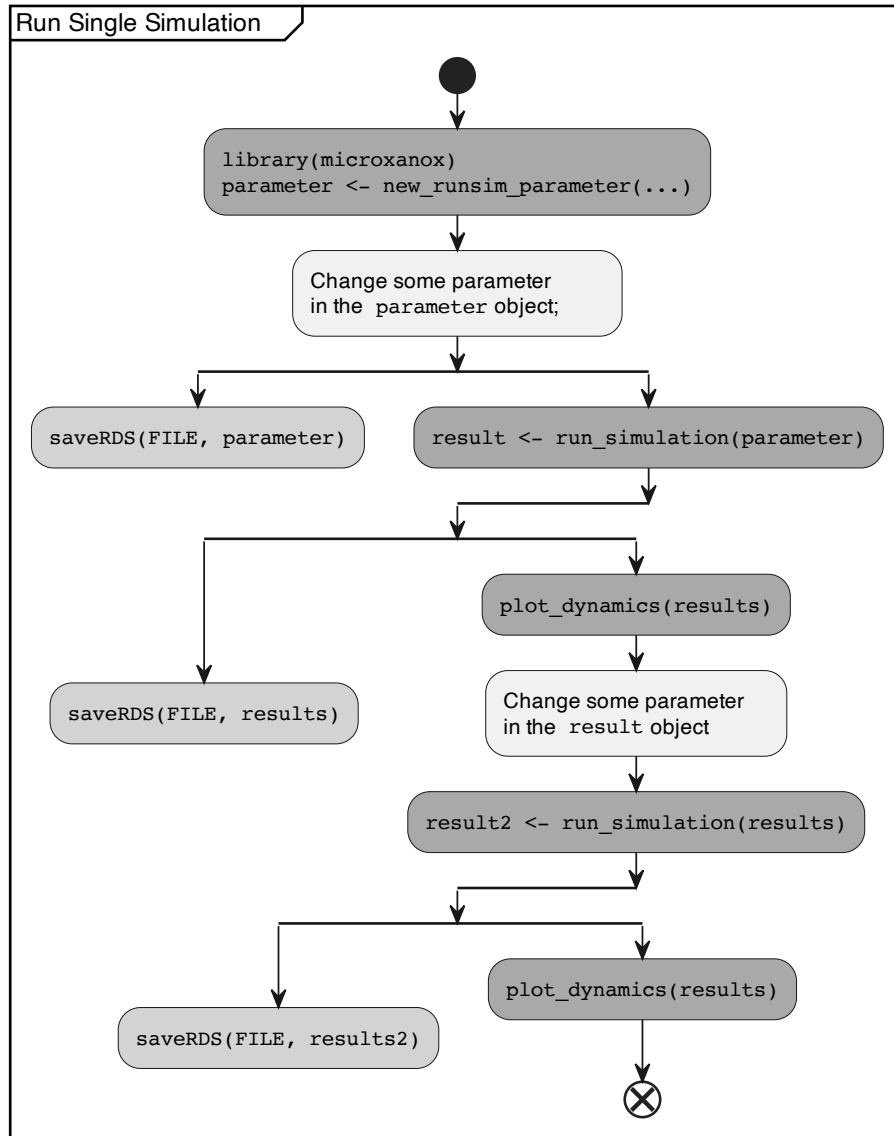
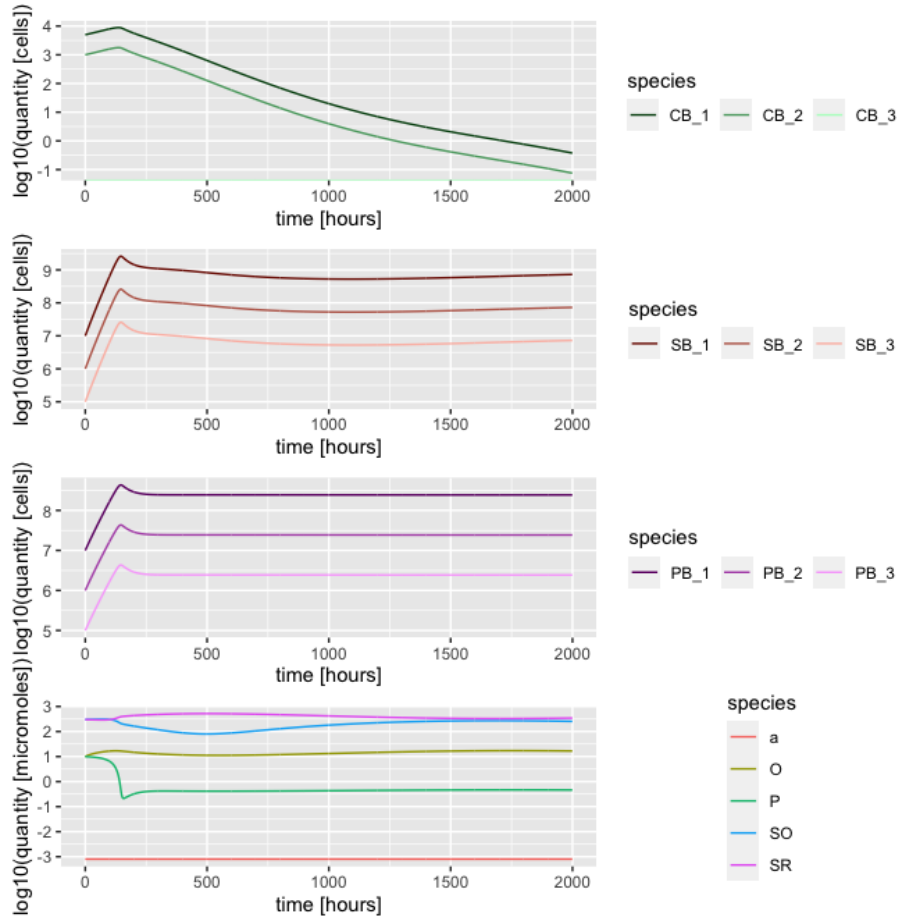


Figure 1: Typical flow of a simulation. Dark Grey boxes: commands necessary for simulation; Light Grey: Saving of parameter and results; Lightest Grey: Different non specified commands.

152 and the number of cores for multithreading the simulation. As the multithread-
 153 ing uses the package function `mclapply()` from the package `parallel` (R Core
 154 Team, 2022), the multithreading only works on Linux and Mac. It is planned to
 155 move to `parLapply()` (R Core Team, 2022) in a future release.
 156 This temporal method implemented in the function `run_temporal_ssfind()`,
 157 which takes a parameter object as created by the function `new_temporal_ssfind_parameter()`.
 158 It is planned for a later release, to run these two simulations in parallel.
 159 For a more detailed walk-through of these two approaches and explanation
 160 please see the [User Guide](#).

161 *4.4. Analysing and visualising results*

162 From the results returned, summary measures about how the ecosystem
 163 stable states respond to environmental change can be extracted. The function
 164 `get_stability_measures()` returns quantities such as the amount of environ-
 165 mental change required for the system to abruptly change to a different state.
 166 The function `plot_dynamics()` plots a single simulation run, as returned from
 167 the `run_simulation()` function. This function is only provided as a convenience
 168 function to provide a way to easily see the results of a simulation run. An
 169 example plot resulting from this function is shown in @ref(fig:plot-dynamics).
 170 `\begin{figure}`



{
 }
 \caption{Results of a simulation run shown using the function *plot_dynamics()*.
 In this case, there were three strains per functional group, though strains within
 functional groups had identical properties in this example. CB_1 =
 cyanobacteria strain 1; SB_1 = sulfur reducing bacteria stain 1; PB =
 phototrophic sulfur bacteria strain 1. Here we show the figure produce by
plot_dynamics(), which produces figures not with all features ideal for
 publication.} \end{figure}

5. Impact

{»This is the main section of the article and the reviewers weight the
 description here appropriately Indicate in what way new research
 questions can be pursued as a result of the software (if any). Indicate
 in what way, and to what extent, the pursuit of existing research
 questions is improved (if so). Indicate in what way the software

has changed the daily practice of its users (if so). Indicate how widespread the use of the software is within and outside the intended user group. Indicate in what way the software is used in commercial settings and/or how it led to the creation of spin-off companies (if so).«}

The open source implementation and extension of the model used in [Bush et al. \(2017\)](#) provides the means of reproducing the results published while at the same time provides the means of doing unique, innovative, and important investigations of how ecosystems respond to environmental change, and how biodiversity may modulate this response.

The design of the package code and functionality is with reproducibility in mind: the combination of all parameters being in a single parameter object as well as the return of the simulation as a result object which inherits from the parameter object provides a relatively easy to use framework to implement reproducible experiments.

Here we evidence the impact of the *microxanax* package by describing three use cases and then by describing how the package can be a starting point for models of other ecosystems. The first two use cases are described in detail (including the code for reproducing them) in the *User Guide* and the *Partial Reproduction* vignettes. The third is taken from [REF NEEDED \(2222\)](#).

5.1. Use case 1: Regime shifts during temporal environmental change

The study of [Bush et al. \(2017\)](#) includes simulations of the effect of oxygen diffusivity (an environmental driver, in the sense that it affects the ecosystem but is not affected by it) on the ecosystem state (oxic or anoxic). The *microxanax* package contains functionality to make a specific temporal pattern of change in the oxygen diffusivity. As well as allowing individual simulations during which oxygen diffusivity varies, this functionality forms the basis of the temporal method for finding stable states.

An example of this functionality is given in the *Partial Reproduction* vignette, which we briefly show here (Figure @ref(fig:uc1)). The example is composed of a single simulation, at the beginning of which the system is in the oxic state with high abundance of cyanobacteria. Oxygen diffusivity is then slowly decreased and eventually, around hour 30'000 the system switches to the anoxic state, with high abundance of both sulfur bacteria types. The oxygen diffusivity is then increased and at around hour 38'000 the system abruptly switches back to the oxic state.

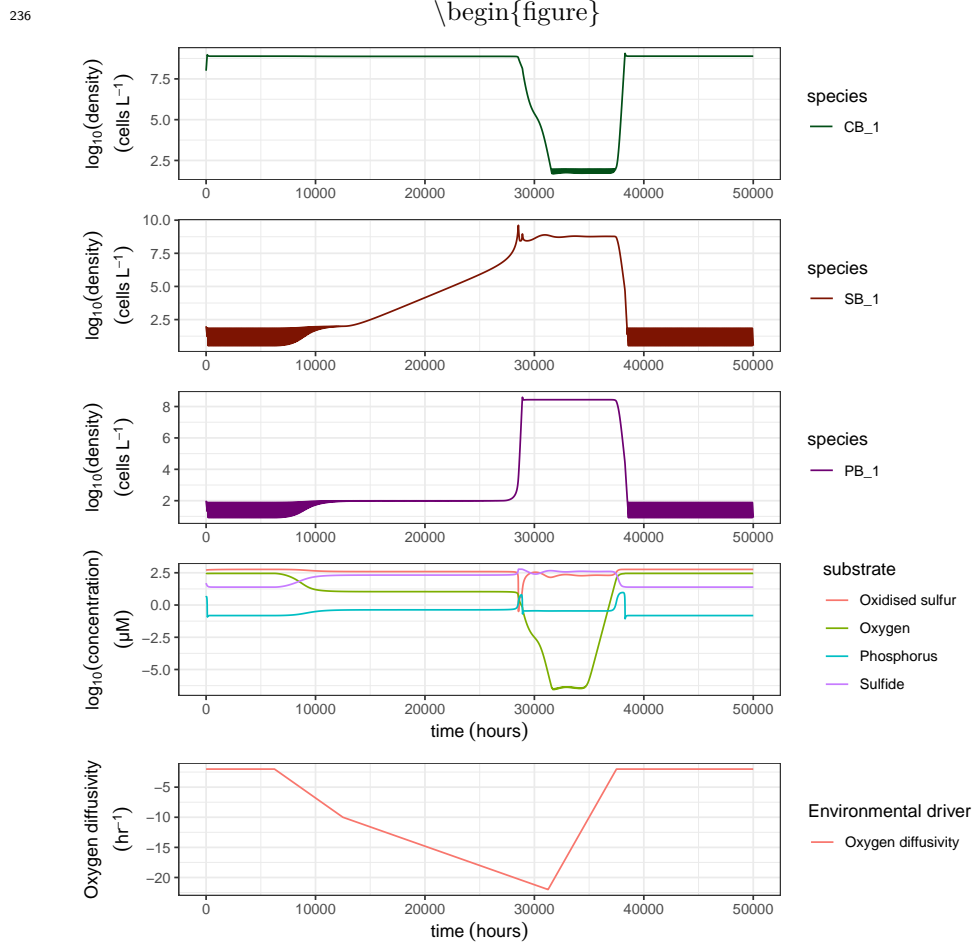
Also visible in the results are thick lines showing abundances of bacteria when abundances are low. This is due to the implementation of a function that at regular intervals, increases the abundance to a preset level. This prevents abundances reducing to very small numbers. The function that implements this increase abundance can also be made to add a certain abundance to each strain at regular intervals, thus simulating immigration in to the system.

#>

```

229 #> Attaching package: 'dplyr'
230 #> The following objects are masked from 'package:stats':
231 #>
232 #>     filter, lag
233 #> The following objects are masked from 'package:base':
234 #>
235 #>     intersect, setdiff, setequal, union

```



```

237 {
238     }
239     \caption{The temporal dynamics of the ecosystem model when an
240     environmental condition (here parameter  $a$ , the oxygen diffusivity) changes.
241     Plot of the stable states of the simulation runs under different oxygen diffusivity.
242     In this simulation there is only one strain in each functional group. CB_1 =
243     cyanobacteria strain 1; SB_1 = sulfur reducing bacteria stain 1; PB =
244     phototrophic sulfur bacteria strain 1. Here we show a figure adapted from the
245     output of the plot_dynamics() function.} \end{figure}

```

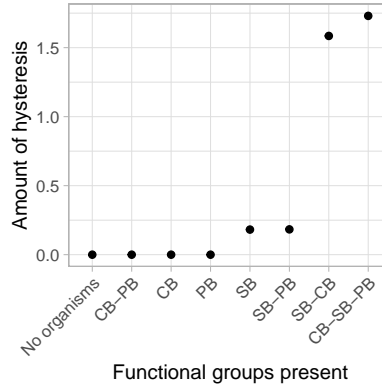


Figure 2: The amount of hysteresis depends on the combination of types of organisms present. The model is entirely deterministic, hence there are no error bars.

5.2. Use case 2: The extent of hysteresis depends on community composition

The package contains a function to extract summary features of ecosystem responses to environmental change, such as the amount of hysteresis displayed by the ecosystem. Hysteresis is a key feature of ecosystem responses to environmental change, because it is related to how difficult it can be to reverse the effects of environmental change (Scheffer et al., 2001). The amount of hysteresis is measured as the extent of the environmental condition (here oxygen diffusivity) for which there were two stable states. I.e. it is the extent of the environmental conditions for which historical conditions play an important role in determining the current system state (a definition of hysteresis).

Using the package to calculate the extent of hysteresis involves setting ecosystem and simulation parameters, including parameters for the finding of stable states across an environmental gradient, running the stable state finding function, and analysing the results with the function that calculates extent of hysteresis. The code for this is provided in the *User Guide* vignette.

The results show that the amount of hysteresis depends greatly on the combinations of organisms present (Figure @ref(fig:uc2)). For example, with only the CB (cyanobacteria) present, there was no hysteresis. In contrast, the presence of both CB and SB (sulfate reducing bacteria) led to a large amount of hysteresis. (These results are also given in the *Partial Reproduction* vignette.)

5.3. Use case 3: Effects of functional diversity on regime shifts

As discussed in the Introduction section, the package was motivated by the question of how biodiversity influences ecosystem responses to environmental change. Extensive results concerning this question are given in a separate publication Limberger (2022). Here we describe one of the results, which is that having biodiversity in a functional group can allow state changes to occur that otherwise would not have. I.e. biodiversity can qualitatively change the state of the ecosystem.

Biodiversity was added to the functional groups using the `new_strain_parameter()` function to create a parameter set with multiple species per functional group (albeit all with identical features) and then to add variability among the species by calling the `add_strain_var()` function. This function takes an already existing parameter set and adds the specified amount of variation. The new parameter object is then used as before. Figure @ref(fig:uc3) shows a simulation with two species (strains) in each of the three functional groups. The ecosystem starts in the oxic state, though with relatively high abundance of each functional group. The strain of SB that is more tolerant to oxygen (SB_1) initially decreases in abundance, but then increases, and the other (SB_2) strain then becomes abundant and SB_1 declines. Furthermore, the cyanobacteria crash in abundance, and the system switches to the anoxic state. In contrast, if there are two identical strains with tolerance half way between those in Figure @ref(fig:uc3) the ecosystem remains in the oxic state.

5.4. Adapting for other ecosystems and organisms

We anticipate that the package can be a useful starting point for investigating models of other types of ecosystem and how biodiversity in them affects responses to environmental change. The overall framework of the package, the purpose of each function, and the objects used for storing parameters and results could be retained. E.g., all such models would have parameters that differ among species and need to be described in an object, and studies will often need to run simulations and sets of simulations across environmental conditions. Researchers wanting to model a new ecosystem do not, therefore, have to start from scratch. This will relieve researchers from needing to make software design decisions, and rather focus on appropriately representing their ecosystem, and finding the results that interest them. Nevertheless, adaptation of the code in the package will require a person / persons that can take a conceptual model of an ecosystem and then represent that in terms of parameters and rate equations, and that is relatively proficient in R programming.

6. Conclusions

{»Set out the conclusion of this original software publication.«}

The *microxanox* R package allows the simulation, visualisation, and analysis of a model of a microbial ecosystem while allowing variation in the amount of diversity contained in each of the functional groups of organisms present. It has been used for the research described in another paper that describes one of the first investigations of the effects of diversity on ecosystem resilience [Limberger \(2022\)](#). In that paper, we show that diversity can have large and important effects of ecosystem responses, highlighting the need for models such as ours, with which one can easily manipulate the amount of biodiversity. The *microxanox* package has also been used to reproduce the results of the paper that inspired the package development [Bush et al. \(2017\)](#).

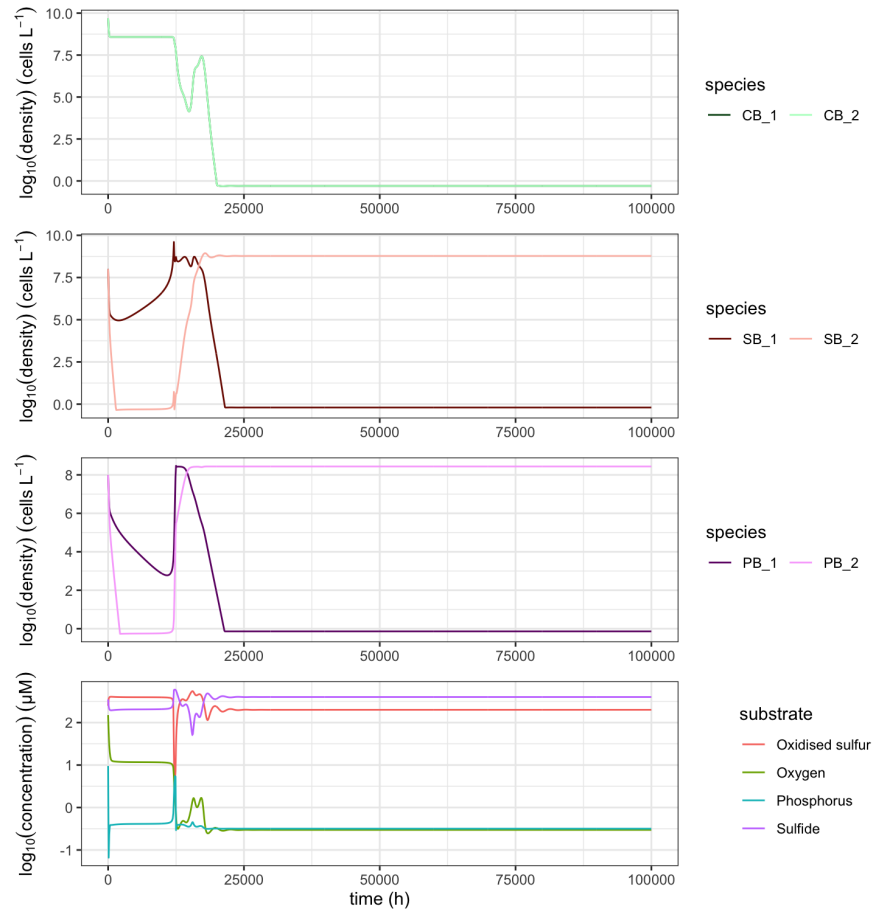


Figure 3: The dynamics of the ecosystem when there are two species in each functional group, and some variation (diversity) in species parameters.

The package greatly lowers the amount of work required in further investigations of the specific ecosystem modelled. There has, for example, been quite limited investigation of how biodiversity influences the short-term responses of the modelled ecosystem to environmental change. Likewise, the package could be used to power an investigation of the effects of biodiversity on the usefulness of early warning signals of abrupt ecosystem change (Scheffer et al., 2009). In addition this package could be used as a template for the implementation for developing models of other types of ecosystems and organism. By doing so, other models can profit from the overall framework used, and the reproducibility aspects as well as the flexibility implemented.

7. Conflict of Interest

The authors declare no known conflicting or competing interests associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

8. Acknowledgements

This project was part of SNF Project 310030_188431. The project was also supported by the University of Zurich Research Priority Programme in Global Change and Biodiversity.

References

- Binzer, A., Guill, C., Rall, B.C., Brose, U., 2016. Interactive effects of warming, eutrophication and size structure: Impacts on biodiversity and food-web structure. *Global Change Biology* 22, 220–227. doi:[10.1111/gcb.13086](https://doi.org/10.1111/gcb.13086)
- Bush, T., Diao, M., Allen, R.J., Sinnige, R., Muyzer, G., Huisman, J., 2017. Oxic-anoxic regime shifts mediated by feedbacks between biogeochemical processes and microbial community dynamics. *Nature Communications* 8, 789. doi:[10.1038/s41467-017-00912-x](https://doi.org/10.1038/s41467-017-00912-x)
- Chaparro-Pedraza, P.C., 2021. Fast environmental change and eco-evolutionary feedbacks can drive regime shifts in ecosystems before tipping points are crossed. *Proceedings of the Royal Society B: Biological Sciences* 288, 20211192. doi:[10.1098/rspb.2021.1192](https://doi.org/10.1098/rspb.2021.1192)
- Collins, S.L., Nippert, J.B., Blair, J.M., Briggs, J.M., Blackmore, P., Ratajczak, Z., 2021. Fire frequency, state change and hysteresis in tallgrass prairie. *Ecology Letters* 24, 636–647. doi:[10.1111/ele.13676](https://doi.org/10.1111/ele.13676)
- Limberger, R., 2022. ###TO BE UPDATED### - The role of functional diversity in ecosystem regime shifts.
- Northrop, A.C., Avalone, V., Ellison, A.M., Ballif, B.A., Gotelli, N.J., 2021. Clockwise and counterclockwise hysteresis characterize state changes in the same aquatic ecosystem. *Ecology Letters* 24, 94–101. doi:[10.1111/ele.13625](https://doi.org/10.1111/ele.13625)

354 R Core Team, 2022. R: A language and environment for statistical computing.
 355 R Foundation for Statistical Computing, Vienna, Austria.
 356 REF NEEDED, A., 2222. REFERNECE NEEDED. Journal of missing
 357 references.
 358 Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos,
 359 V., Held, H., van Nes, E.H., Rietkerk, M., Sugihara, G., 2009. Early-warning
 360 signals for critical transitions. *Nature* 461, 53–59. doi:[10.1038/nature08227](https://doi.org/10.1038/nature08227)
 361 Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., Walker, B., 2001.
 362 Catastrophic shifts in ecosystems. *Nature* 413, 591–596.
 363 doi:[10.1038/35098000](https://doi.org/10.1038/35098000)
 364 Soetaert, K., Petzoldt, T., Setzer, R.W., 2010. Solving differential equations in
 365 R: Package deSolve. *Journal of Statistical Software* 33, 1–25.
 366 doi:[10.18637/jss.v033.i09](https://doi.org/10.18637/jss.v033.i09)
 367 Vandermeer, J., Perfecto, I., 2019. Hysteresis and critical transitions in a coffee
 368 agroecosystem. *Proceedings of the National Academy of Sciences* 116,
 369 15074–15079. doi:[10.1073/pnas.1902773116](https://doi.org/10.1073/pnas.1902773116)
 370 Vasseur, D.A., McCann, K.S., 2005. A Mechanistic Approach for Modeling
 371 Temperature-Dependent Consumer-Resource Dynamics. *The American*
 372 *Naturalist* 166, 184–198. doi:[10.1086/431285](https://doi.org/10.1086/431285)