Can neural networks predict steady annual cycles of marine ecosystems?

Thomas Slawig1 · Markus Pfeil1

Accepted: 10 July 2022 / Published online: 8 October 2022
© The Author(s) 2022

Abstract
We used artificial neural networks to replace the complete spin-up procedure that computes a steady annual cycle of a marine ecosystem driven by ocean transport. The networks took only the few biogeochemical model parameters and attempted to predict the spatially distributed concentrations of the ecosystem, in this case only nutrients, for one time point of the annual cycle. The ocean circulation was fixed for all parameters. Different network topologies, sparse networks, and hyperparameter optimization using a genetic algorithm were used. This showed that all studied networks can produce a distribution that is point-wise close to the original spin-up result. However, these predictions were far from being annually periodic, such that a subsequent spin-up was necessary. In this way, the overall runtime of the spin-up could be reduced by 13% on average. It is debatable whether this procedure is useful for the generation of initial values, or whether simpler methods can achieve faster convergence.

Können neuronale Netzwerke stetige Jahreszyklen mariner Ökosysteme vorhersagen?

Zusammenfassung
Wir haben künstliche neuronale Netze verwendet, um den kompletten Spin-up zu ersetzen, mit dem ein stetiger Jahreszyklus eines marinen, durch den Ozeantransport angetriebenen Ökosystems berechnet wird. Die Netze nahmen nur die wenigen biogeochemischen Modellparameter und versuchten, die räumlich verteilten Konzentrationen des Ökosystems, hier nur Nährstoffe, für einen Zeitpunkt des Jahreszyklus vorherzusagen. Die Ozeanzirkulation war für alle Parameter fest. Es wurden verschiedene Netzwerktopologien, „sparse networks“ und ein Hyperparametertuning durch einen genetischen Algorithmus verwendet. Alle Netze konnten eine Verteilung erzeugen, die dem ursprünglichen Spin-up-Ergebnis punktweise ähnlich war. Allerdings waren die Vorhersagen weit davon entfernt, jahresperiodisch zu sein, weshalb ein nachträglicher Spin-up nötig war. So konnte die Gesamtlauzeit des Spin-ups im Durchschnitt um 13 % reduziert werden. Es bleibt fraglich, ob dieses Verfahren sinnvoll ist, um Anfangswerte zu generieren, oder ob einfachere Methoden eine schnellere Konvergenz erreichen können.

Marine ecosystems: expert knowledge

The marine ecosystem forms an essential part of the global carbon cycle. The ocean absorbs CO₂ via its surface, which is converted by photosynthesis in algae, i.e., light and available nutrients are important. Animal plankton uses the phytoplankton as food, and dead parts sink into deeper layers, see Fig. 1 for a schematic. Simulations are used to study the behavior of the marine ecosystem. These consist of the ocean circulation and the biogeochemical processes of the ecosystem itself. The modeling of the circulation is described by the equations of fluid mechanics, but there is a variety of biogeochemical models of different complexity [1]. They represent growing and dying of the species, their interdependency as food for each other, etc. Thus, the models’ design depends heavily on human expert knowledge, reflected in their structure. Moreover, they contain parameters that cannot be determined experimentally. Our example here was the N-model from [1] with phosphate as the only substance and five parameters (attenuation coeffi-
Optimization: fusion of measurement and simulation data

Two questions arise in this context: Which model represents reality (i.e., measured data) best, and what are optimal model parameters for the individual models [2]? For this purpose, simulation data are compared with measurement data. Climatological data, averaged over several years but spatially distributed, are used to form a somehow “ideal” year. However, the available data are spatially sparse, see Fig. 2. Thus, a simulation output does not need to be perfect everywhere, it suffices to catch the system behavior mostly at the coasts and in the surface layers.

Corresponding to the climatological data, the marine ecosystem models are also simulated as long as they converge to an annually periodic state, i.e., a stable annual cycle. Such a simulation run, called a “spin-up”, extends over a long period of 3000–10,000 model years until the converged, stable annual cycle is reached. It starts from default, constant values for the substances. Since the simulation has to be performed for the entire ocean, this results in a very high computational effort. The whole process, starting from the setting of initial state and model parameters and ending with the converged steady annual cycle, is depicted in Fig. 3, a snapshot of an exemplary resulting steady distribution in Fig. 4. The 3D output is given at 52,749 grid points, in a still coarse 2.8-degree horizontal resolution. In addition to spatial parallelization and high-performance computing, the high computational cost can be further reduced: The feedback of biogeochemistry on ocean transport is small and might be neglected, i.e., the transport may be precomputed and held fixed for all parameters. The transport matrix method [5] stores the effect of an ocean transport step (without biogeochemistry) in time-averaged matrices. Nevertheless, it is necessary to further reduce computational time, especially for optimization runs that require a large number of model evaluations.

Artificial neural networks as a replacement for the spin-up process

We first trained artificial neural networks (ANNs) to directly predict the final result of the spin-up from the biogeochemical model parameters alone (see the red arrows in Fig. 3). This approach is very ambitious: The input parameters do not contain any spatial information about the geometry, but we try the network to produce a spatial distribution of the biogeochemical substances. On the other hand, the final output has some geometric structure (as can been in Fig. 4), which is retained for a wide range of model parameters, at least to some extent. We started with a fully connected network (FCN) with three hidden layers, see Fig. 5.

We had 1100 data sets, i.e., pairs of model parameter sets (obtained by Latin Hypercube sampling) and corresponding 3D spin-up results. These were used partially for training, validation, and testing. Fig. 6 shows that the differences between the ANN’s prediction and the spin-up result are still visible, here for one exemplary parameter set. However, an error of $10^{-3}$ does not seem too great, knowing that the average range of the nutrient values is $O(1)$.

Thus, we checked how “far” the prediction by the ANN is from the converged spin-up result, measured in necessary iterations in the spin-up. To this end, we run the spin-up, starting with the ANN’s prediction, until convergence (i.e., until the difference between two successive years in the model was below a given threshold). We reached a far better result, see Fig. 7. Nevertheless, the maximal error is in the
Fig. 3 Flow chart of a spin-up. Orange arrows mark the in- and output used in the training of the artificial neural network, red ones its action as a replacement of the whole spin-up, magenta ones the use of the prediction in a subsequent spin-up.

Fig. 4 Distribution of nutrients (phosphate, in mmmol P m$^{-3}$) at the ocean surface as a result of a converged spin-up.

same range of $10^{-3}$, but only at some barely visible points. The overall accuracy is much better.

We performed these tests for 100 test parameter sets. Fig. 8 shows how long the spin-ups took when starting from the prediction rather than the constant default values. Even though the differences between the prediction and the standard spin-up result did not seem that great (see again Fig. 6), it still took quite a high number of iterations to finally reach a converged annual cycle.

These were results for the fully connected network depicted above. Next, we compared different network topologies. We tested convolutional networks and networks adapted by the Sparse Evolutionary Training (SET) algorithm [8]. This adaptively modifies the network topology by deleting edges with small weights and randomly adding new ones in each training epoch. Additionally, we performed hyperparameter tuning by a Genetic Al-

Fig. 5 Simple topology of the neural network first used.

Fig. 6 Difference between the prediction of the neural network and the result of the spin-up at the ocean surface, in mmmol P m$^{-3}$.

Fig. 7 As in Fig. 6, but now between result of the standard spin-up and the one starting from the ANN's prediction.
The parameters we optimized were the number of layers, their respective number of neurons, the activation function, learning rate and optimization algorithm as well as two parameters of the SET algorithm. None of these techniques significantly improved the results. Networks designed and trained using the SET algorithm showed the same quality of results as the standard ones, which is a justification for the algorithm’s use since it is not designed to produce better predictions, but only to lower training cost. Hyperparameter tuning by the GA did not lead to significant improvements, see [7] for details. Even with sparse networks, the number of network parameters to be optimized in the training process remains too large for the available number of training data. However, obtaining more data sets is quite difficult and computationally expensive in this context.

ANN as a generator of better initial states for the spin-up?

We might look again at Fig. 8 and observe that, taking the ANN’s prediction as an initial value, we can reduce the necessary steps in the spin-up to some extent (here, by 13% on average). This could lead to the idea that an ANN can be used to generate initial values that accelerate the spin-up. In climate research, it is assumed that the stable annual cycle of the ecosystem is independent of the initial state. A constant initial state is usually chosen in this application. In [6], it was shown that the resulting stable annual cycle for the model hierarchy presented in [1] is in fact only weakly dependent on the initial state. Fig. 9 shows that different initial values, randomly chosen using different probability distributions or concentrating all substances in one box, produce very similar results compared to those with the default constant ones. However, it is conceivable that a different initial state might shorten the runtime of the spin-up, a fact that we just saw when using the ANN’s prediction. However, it has to be further investigated whether this is really a reasonable procedure. In [6], we were mainly interested to see if the converged spin-up results differ depending on the initial states, but we also found out that the speed of the spin-up’s convergence differs. A similar or even higher acceleration of the spin-up might be achievable by simpler methods, e.g., taking the average or another combination of available converged spin-up data.

Summary and outlook

We investigated whether ANNs of different structure can completely replace the computationally expensive spin-up process in marine ecosystem simulations. The networks took only the model parameters of the biogeochemistry as inputs and tried to predict the spatially resolved, steady annual cycle of the substances in the marine. The ocean circulation was pre-computed and fixed for all parameters. It turned out that the ANNs could predict the distribution of the steady annual cycles obtained by the standard spin-up with some accuracy. However, these predictions were still far from being annually periodic. It still took a subsequent spin-up to reach the converged steady cycle. In the average over 100 test parameter sets, the runtime for the spin-up was reduced by about 13% in this way. Different network topologies, sparse networks and hyperparameter tuning by a genetic algorithm were used, with no big influence on the quality of the results. A future improvement could be the use of different network types (for example Conditional Generative Adversarial Nets [9]), including results of the first steps of a spin-up as input for the networks and reducing the output data dimension by methods such as principal component analysis or similar.
Acknowledgements  We would like to thank the reviewers who gave us helpful ideas to improve the manuscript and also drew our attention to promising further options and directions for the research in this application.

Funding  Open Access funding enabled and organized by Projekt DEAL.

Open Access  This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Kriest I, Khatiwala S, Oschlies A (2010) Towards an assessment of simple global marine biogeochemical models of different complexity. Prog Oceanogr 86(3-4):337–360
2. Schartau M, Wallhead P, Hemmings J, Loeptin U, Kriest I, Krishna S, Ward BA, Slawig T, Oschlies A (2017) Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling. Biogeoosciences 14(6):1647–1701
3. Boyer TP, Antonov JJ, Baranova OK, Coleman C, Garcia HE, Grotch A, Johnson DR, Locarnini RA, Mishonov AV, O’Brien TD, Paver CR, Reagan JF, Seidov D, Smolyar IV, Zweng MM (2013) World ocean database 2013. In: Levitus S, Mishonov A (eds) Technical Ed.; NOAA Atlas NESDIS 72
4. Reimer J (2019) Statistical analysis of the phosphate data of the world ocean database 2013. arXiv:1912.07384
5. Khatiwala S (2007) A computational framework for simulation of biogeochemical tracers in the ocean. Global Biogeochem Cycles 21:3
6. Pfeil M, Slawig T (2021) Unique steady annual cycle in marine ecosystem model simulations. arXiv:2111.15424
7. Pfeil M, Slawig T (2022) Approximation of a marine ecosystem model by artificial neural networks. Electron Trans Numer Anal 56:138–156
8. Mocanu DL, Mocanu E, Stone P, Nguyen PH, Gibescu M, Liotta A (2018) Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science. Nat Commun 9:2383–2394
9. Mirza M, Osindero S (2014) Conditional generative adversarial nets. arXiv:1411.1784

Publisher’s Note  Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.