

Recovering complex ecological dynamics from time series using state-space universal dynamic equations

Corresponding Author: Dr Jack Buckner

This file contains all editorial decision letters in order by version, followed by all author rebuttals in order by version.

Version 0:

Decision Letter:

**** Please ensure you delete the link to your author home page in this e-mail if you wish to forward it to your coauthors ****

Dear Dr Buckner,

Your manuscript titled "Recovering complex ecological dynamics from time series using state-space universal dynamic equations" has now been seen by 3 reviewers, whose comments are appended below. You will see that they find your work of some potential interest. However, they have raised quite substantial concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication, but would be interested in considering a revised version that fully addresses these serious concerns.

In revision, please address the following editorial thresholds:

- * Provide a novel framework integrating state-space models with universal dynamic equations for ecological dynamics, including clear explanations of the state-space framework and full comparisons with other models to illustrate features of the proposed framework.
- * Provide compelling applications of the proposed framework in complex ecosystems, including demonstrations of how the approach scales to large and complex models, raised by reviewer #2.
- * Clearly explain the validation and implementation of the model so that readers can easily capture the analysis.

We hope you will find the reviewers' comments useful as you decide how to proceed. Should additional work allow you to address these criticisms, we would be happy to look at a substantially revised manuscript. If you choose to take up this option, please either highlight all changes in the manuscript text file, or provide a list of the changes to the manuscript with your responses to the reviewers.

When resubmitting, please provide a point-by-point response to the reviewers' comments. Please submit your responses as a separate file, distinct from your cover letter where you can add responses to the Editors' comments that you do not want to be made available to the reviewers. Word files are preferred. We recommend that any figures, tables or graphs that are included in the response to reviewers are also included in the main article or Supplementary Information.

Please bear in mind that we will be reluctant to approach the reviewers again in the absence of substantial revisions.

If the revision process takes significantly longer than three months, we will be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the reviewers' comments with a list of your changes to the manuscript text (which should be in a separate document to any cover letter), a tracked-changes version of the manuscript (as a PDF file) and any completed checklist:

Link Redacted

**** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first ****

Please do not hesitate to contact us if you have any questions or would like to discuss the required revisions further. Thank you for the opportunity to review your work.

Best regards,

Lifen Jiang, PhD
Editorial Board Member
Communications Earth & Environment
orcid.org/0000-0002-1546-8189

Mengjie Wang
Associate Editor
Communications Earth & Environment
@CommsEarth

EDITORIAL POLICIES AND FORMAT

If you decide to resubmit your paper, please ensure that your manuscript complies with our editorial policies and complete and upload the checklist below as a Related Manuscript file type with the revised article:

Editorial Policy Policy requirements
(Download the link to your computer as a PDF.)

For Manuscripts that fall into the following fields:

- Behavioural and social science
- Ecological, evolutionary & environmental sciences
- Life sciences

An updated and completed version of our Reporting Summary must be uploaded with the revised manuscript

You can download the form here:

<https://www.nature.com/documents/nr-reporting-summary.zip>

For your information, you can find some guidance regarding format requirements summarized on the following checklist: (<https://www.nature.com/documents/commsj-phys-style-formatting-checklist-article.pdf>) and formatting guide (<https://www.nature.com/documents/commsj-phys-style-formatting-guide-accept.pdf>).

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

This paper presents an innovative approach by explicitly incorporating a state-space framework with UDEs to address common sources of uncertainty in ecological time series data, specifically handling both process and observation errors. While UDEs have previously been applied to ecological datasets, this work appears to pioneer the integration of state-space modeling with UDEs, offering a structured method to accommodate uncertainty in chaotic or noisy datasets. To strengthen this paper, it would be valuable to demonstrate how the proposed method scales to larger and more complex models beyond the current application on the JER dataset. Ecological systems typically involve high-dimensional data with numerous interacting species, spatial heterogeneity, and various environmental variables. By applying the state-space UDE approach to synthetic datasets that mimic these complexities, the authors could address the well-known scalability challenges in state-space models. Such an extension would not only validate the method's applicability to real-world ecological systems but also benefit researchers working with complex dynamic models. Furthermore, enhancing the comparative analysis could bolster the paper's impact. Evaluating the state-space UDE against traditional UDE models, probabilistic physics-informed neural networks (PINNs), or Bayesian state-space models would help establish its unique advantages and limitations. These comparisons would offer a more comprehensive understanding of the method's performance, especially in handling uncertainty and capturing complex dynamics inherent in ecological data.

While the authors compare their model to MARSS and gpEDM methods, a deeper and rigorous discussion on the reasons behind the UDE's performance—whether superior or inferior—would be beneficial. Elaborating on the specific scenarios where the UDE excels or falls short would provide valuable insights and guide future research in this area.

Overall, while this paper makes strides in combining UDEs with state-space modeling, demonstrating its scalability and benchmarking against other state-space-based approaches would greatly strengthen its contributions to the field. In its current form, however, I do not feel it meets the novelty and rigor required for publication in this journal, and I recommend

rejection.

Reviewer #2 (Remarks to the Author):

In this paper, the authors developed state-space universal dynamic equations by combining universal differential equations with a state-space modeling framework, accounting for uncertainty.

The proposal system was tested on two simulated and two empirical case studies.

The authors claim that this method can recover nonlinear biological interactions that produce complex behaviors including chaos and regime shifts. While agreeing with majority part of the claim, I do have some questions.

A few technical details

1. Line 117, in the second sum, shouldn't the summation index be from $t=0$, to $t=T-1$?
2. In Figure 2A, the forecast results do not seem to be consistent to the testing data. Can the authors provide explanation?
3. In Figure 3E, the RMSE shown for all the cases seem to be fairly large. Can the authors provide explanation?
4. Can the authors explain why the Three-species food chain model and Kelp forest model are selected for the tests?

Reviewer #3 (Remarks to the Author):

This study introduces a promising framework integrating state-space models with universal dynamic equations (UDEs) to address uncertainties in ecological time series modeling. The authors rigorously validate the method through simulated (chaotic food chains, kelp-urchin systems) and empirical (fisheries, arid rangelands) case studies, demonstrating its ability to recover nonlinear dynamics and regime shifts. While the work is innovative and addresses a critical gap in ecological forecasting, the manuscript's complexity and inconsistent presentation hinder its accessibility to a broad audience. Key strengths include the integration of mechanistic knowledge with data-driven neural networks, while major weaknesses revolve around methodological clarity, reproducibility, and logical coherence across examples.

First, unclear Methodological Comparisons and Scope. The manuscript introduces multiple alternative models (e.g., MARSS, gpEDM, null models) across different case studies without a consistent rationale. For instance, the fisheries example compares UDEs to a null model and linear state-space models, while the food chain case adds Gaussian process models. This inconsistency creates confusion about the method's unique value. I would suggest focus comparisons on state-space UDEs, traditional UDEs (without state-space components) and a null model.

Second, insufficient explanation of state-space framework. The distinction between the process model (dynamic equations with neural networks) and data model (observation errors) is poorly defined. The manuscript mentions the data model $h(ut, yt)$ but never provides its explicit form. Is the data model just likelihood equation? This omission makes it unclear how observation errors are integrated into the framework. The introduction of state-space framework in the introduction section is too abstract, not suitable for broad readers.

Third, ambiguous cross-validation protocols. The manuscript inconsistently applies cross-validation methods. In the Method section, the authors describe two cross-validation methods. While, in the Result section, it does not clarify how the methods are applied.

Fourth, inconsistent model implementation across cases. In the food chain model, the process model introduces stochasticity via additive noise, and then the time series data were added observation errors. However, in the kelp-urchin model, auto regressive processes were used to simulation abiotic conditions but lack explicit definitions. It's unclear how process errors and observation errors were distinguished. In addition, it's unclear AR-1 is applied to X_t or y_t ?

Other comments:

1. Standardize the presentation of simulated data generation (e.g., always include plots of raw vs. noisy data).
2. Clarify how covariates (e.g., $t > 1992$ in fisheries) are encoded and integrated into neural networks.
3. Quantify how much state-space UDEs improve predictions over traditional UDEs (e.g., % reduction in RMSE).
4. In simulated cases with known ground-truth errors (e.g., food chain), evaluate whether the model accurately partitions process vs. observation errors.
5. Specify how the nine simulated datasets in Figure 3f differ.
6. It's better to have both the simulated data and empirical data in same cases.

** Visit Nature Research's author and referees' website at www.nature.com/authors for information about policies, services and author benefits**

Communications Earth & Environment is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' create and link their Open Researcher

and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the Manuscript Tracking System by clicking on 'Modify my Springer Nature account' and following the instructions in the link below. Please also inform all co-authors that they can add their ORCIDs to their accounts and that they must do so prior to acceptance.

<https://www.springernature.com/gp/researchers/orcid/orcid-for-nature-research>

For more information please visit <http://www.springernature.com/orcid>

If you experience problems in linking your ORCID, please contact the [Platform Support Helpdesk](http://platformsupport.nature.com/).

Version 1:

Decision Letter:

**** Please ensure you delete the link to your author home page in this e-mail if you wish to forward it to your coauthors ****

Dear Dr Buckner,

Your manuscript titled "Recovering complex ecological dynamics from time series using state-space universal dynamic equations" has now been seen by our reviewers, whose comments appear below. In light of their advice, we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment, provided that you can add more ecological insights based on real data applications, as raised by reviewer #1.

We therefore invite you to revise your paper one last time to address the remaining concerns of our reviewers. At the same time we ask that you edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table".

*****Please take care to match our formatting and policy requirements. We will check revised manuscript and return manuscripts that do not comply. Such requests will lead to delays. *****

Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files as a Related Manuscript file.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at <https://www.nature.com/documents/commsj-file-checklist.pdf>.

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication. For further information about article processing charges, open access funding, and advice and support from Nature Portfolio, please visit <https://www.nature.com/commsenv/open-access>

At acceptance, you will be provided with instructions for completing the open access licence agreement on behalf of all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:

Link Redacted

**** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first ****

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Mengjie Wang

Associate Editor, Communications Earth & Environment
Consulting Editor, Communications Sustainability
Bluesky: @commsearth.nature.com; @commssustain.nature.com

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

This revision includes substantial methodological, structural, and interpretive updates. It generalizes the concept of universal differential equations to encompass both continuous-time and discrete-time formulations, allowing the framework to model a broader range of ecological systems, including those with seasonal or discrete data structures. The revised version introduces two statistical approaches—joint-likelihood and marginal-likelihood methods—for estimating hidden states and parameters within a state-space framework. The marginal-likelihood formulation employs the Unscented Kalman Filter to account for uncertainty in latent states, thereby improving inference accuracy.

The new version explicitly models process and observation errors and discusses appropriate strategies for fixing or estimating these terms. The updated manuscript also provides a more comprehensive comparison of state-space UDEs against alternative approaches, including Gaussian process empirical dynamic models (gpEDMs), multivariate autoregressive state-space (MARSS) models, and UDEs trained with alternative optimization schemes, evaluated across multiple forecast horizons. Figures and analyses have been thoroughly revised to include standardized tables and consistent performance metrics, enhancing clarity and reproducibility.

The framework builds directly on prior work in universal differential equations (Rackauckas et al., 2021), neural ODEs (Rubanova et al., 2019), and state-space UDE models (Ouala et al., Learning Latent Dynamics for Partially-Observed Chaotic Systems, 2020). The manuscript positions itself as a unification of these ideas, but the combination is largely straightforward and does not represent a substantial conceptual leap. However, the application of a UDE framework to ecological modeling remains relatively novel. The approach itself is well designed and could be applied to a wide range of complex dynamical systems beyond ecology.

While the lack of methodological innovation is understandable given the journal's emphasis on applied ecological insight, this makes it even more important for the manuscript to clearly demonstrate the ecological understanding gained from the framework. Although the manuscript dedicates considerable space to methodological exposition and simulation results, its discussion remains largely mathematical and generic, offering limited ecological interpretation.

I find the manuscript generally well written and technically interesting, but several notational and methodological clarifications would improve its precision and readability.

In Equation 6, the function ϕ is not defined clearly.

The paper does not specify any prior information for the initial state values. It would be helpful to describe the assumed prior distribution or initialization procedure used by the UKF.

Line 163, the covariance matrix $\Sigma_{\{\epsilon, i, t\}}$ has t in its subscript; if it is time-invariant, then t can be dropped.

Line 174, the notation $p(u_{1:T}|y_{1:T}, \theta)$ is ambiguous. The standard UKF returns the filtered posterior conditioned on observations up to time t , i.e. $p(u_t|y_{1:t}, \theta)$. If the posterior conditioned on all observations up to T is intended, it should be referred to as the smoothed posterior.

Line 181. The statement "When training with the joint likelihood, these terms need to be fixed at pre-defined values" reflects a common practical choice, but it is not a strict requirement. These variance terms can, in principle, be treated as trainable parameters within the joint-likelihood framework.

Example 1 is labeled numerically, but the subsequent examples are not.

In Equation 29 and 30, for the stochastic differential equation, the notation dW is unclear. If it is the Wiener process, then it should be dW_t instead of dW .

Reviewer #3 (Remarks to the Author):

The authors have addressed my concerns and the manuscript has been improved a lot.

** Visit Nature Portfolio's author and referees' website at www.nature.com/authors for information about policies, services and author benefits**

Open Access This Peer Review File 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.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File 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 license 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 license, visit <https://creativecommons.org/licenses/by/4.0/>

Dear Reviewers,

Thank you for your consideration of our manuscript. Your comments helped us substantially improve the manuscript. Based on these comments and those of the editors we have made three significant changes to the manuscript: 1) we added a fifth case study training the models on a more complex dataset to illustrate how well the methods scale to higher-dimensional systems, 2) we included a more thorough discussion of the state-space modeling framework and training methods in the main text of the paper, and 3) we compared the forecasting skill of state-space UDE models to a consistent and more comprehensive set of alternatives across the four primary case studies. Finally, we have improved the basic methodology presented in the original draft by introducing an additional method for training the UDE models. We now present two training methods based on the state space modeling framework, one that maximizes the marginal likelihood and a second that maximizes the joint likelihood of the state space model. The marginal likelihood approach can be more accurate and allows the model to quantify the magnitude of process errors. In contrast, the joint likelihood has lower computational costs and scales more easily to data sets with many state variables. The trade-off is that the joint likelihood approach can (in theory) produce biased estimates of the model's dynamics. However, both approaches produced similar estimates of the system's dynamics in our case studies (e.g. bifurcation diagrams) and have quantitatively similar forecasting skill—detailed comments to each review are given below.

Reviewer comments:

Reviewer #1

This paper presents an innovative approach by explicitly incorporating a state-space framework with UDEs to address common sources of uncertainty in ecological time series data, specifically handling both process and observation errors. While UDEs have previously been applied to ecological datasets, this work appears to pioneer the integration of state-space modeling with UDEs, offering a structured method to accommodate uncertainty in chaotic or noisy datasets.

Comment 1: To strengthen this paper, it would be valuable to demonstrate how the proposed method scales to larger and more complex models beyond the current application on the JER dataset. Ecological systems typically involve high-dimensional data with numerous interacting species, spatial heterogeneity, and various environmental variables. By applying the state-space UDE approach to synthetic

datasets that mimic these complexities, the authors could address the well-known scalability challenges in state-space models. Such an extension would not only validate the method's applicability to real-world ecological systems but also benefit researchers working with complex dynamic models.

Response: We included a fifth case study based on a spatially-structured predator-prey model to illustrate how well our state-space modeling approach scales to datasets with a large number of state variables. We compared the results of our method to those of two other methods for training UDEs: gradient matching and shooting. We found that state-space models can scale to datasets with approximately 20 state variables, particularly when a continuous-time model formulation is used. However, the alternative methods scaled more efficiently, albeit at the cost of accuracy (Table 2).

Comment 2: Furthermore, enhancing the comparative analysis could bolster the paper's impact. Evaluating the state-space UDE against traditional UDE models, probabilistic physics-informed neural networks (PINNs), or Bayesian state-space models would help establish its unique advantages and limitations. These comparisons would offer a more comprehensive understanding of the method's performance, especially in handling uncertainty and capturing complex dynamics inherent in ecological data.

While the authors compare their model to MARSS and gpEDM methods, a deeper and rigorous discussion on the reasons behind the UDE's performance—whether superior or inferior—would be beneficial. Elaborating on the specific scenarios where the UDE excels or falls short would provide valuable insights and guide future research in this area.

Response: Thank you for these suggestions. We agree that a more rigorous benchmarking analysis and discussion of the conditions where our approach excels and why it is effective would improve the paper. We have addressed this in two ways. First, we used a standardized set of benchmarks across our four primary case studies, adding two additional methods for training UDE models along with the MARSS and gpEDM models. This allowed us to not only compare modeling techniques, but also to understand when and why the state-space training method improved the performance of UDEs in particular. Furthermore, we updated the Discussion section of our paper to provide a clear discussion of when the methods were most effective and why we obtained these results based on the benchmarking analysis (lines: 449-469 and 508-520).

Comment 3: Overall, while this paper makes strides in combining UDEs with state-space modeling, demonstrating its scalability and benchmarking against other state-space-based approaches would greatly strengthen its contributions to the field. In its current form, however, I do not feel it meets the novelty and rigor required for publication in this journal, and I recommend rejection.

Response: We have augmented our paper with additional methods to more rigorously benchmark UDEs versus other methods. In addition, we have developed a metacommunity model as a fifth case study to demonstrate the scalability of universal dynamic equations to larger, more complex ecological systems. We hope this revised version meets the reviewer's expectations for this journal.

Reviewer #2

In this paper, the authors developed state-space universal dynamic equations by combining universal differential equations with a state-space modeling framework, accounting for uncertainty.

The proposal system was tested on two simulated and two empirical case studies. The authors claim that this method can recover nonlinear biological interactions that produce complex behaviors including chaos and regime shifts. While agreeing with majority part of the claim, I do have some questions.

Response: Thank you for your review. Please see our detailed response to each comment below.

A few technical details

1. Line 117, in the second sum, shouldn't the summation index be from $t=0$, to $t=T-1$?

Response: Yes, you are correct, although it should be $t=1$ to $t=T-1$ because we need to have one data point before the first index in the summation.

2. In Figure 2A, the forecast results do not seem to be consistent to the testing data. Can the authors provide explanation?

3. In Figure 3E, the RMSE shown for all the cases seem to be fairly large. Can the authors provide explanation?

Response to 2 and 3: We have updated these panels as Figs. 2A and 3D, which now show more consistent forecasting results and lower RMSE, respectively. In addition, we

show model fit with mean absolute error in the main text and include the mean squared error in the Supplemental Information, because it was easier for us to quantify uncertainty in our mean estimates of model performance with these new metrics. The forecasting skill (MAE) looks reasonable to us given the levels of observational noise included in the simulated data sets. Part of the discrepancy might be due to the fact that in Fig. 2A, where we show a simulated dataset and a forecast, we selected a dataset in the low observational and process error case, which tended to have smaller forecasting errors than the high observational error case.

Comment 4: Can the authors explain why the Three-species food chain model and Kelp forest model are selected for the tests?

Response: Yes, we chose the four examples in the original draft of the paper to include two examples that exhibited oscillating or chaotic dynamics (i.e., three-species food chain and fisheries), and two examples that exhibited regime shifts caused by a fold bifurcation (i.e., kelp forest and rangeland). We wanted one simulated and one empirical example of each. We selected the two simulated examples (i.e., three-species food chain and kelp forest models) based on these criteria and the authors' interests in their respective study systems. We have elaborated the reasons for choosing each of the case studies in the Introduction, discussing how each one helps address one of our primary research questions (lines 72-89).

Reviewer #3

Comment 1: This study introduces a promising framework integrating state-space models with universal dynamic equations (UDEs) to address uncertainties in ecological time series modeling. The authors rigorously validate the method through simulated (chaotic food chains, kelp-urchin systems) and empirical (fisheries, arid rangelands) case studies, demonstrating its ability to recover nonlinear dynamics and regime shifts. While the work is innovative and addresses a critical gap in ecological forecasting, the manuscript's complexity and inconsistent presentation hinder its accessibility to a broad audience. Key strengths include the integration of mechanistic knowledge with data-driven neural networks, while major weaknesses revolve around methodological clarity, reproducibility, and logical coherence across examples.

Response: Thank you for your time in reviewing our manuscript. We appreciate your concerns about the logical coherence across the examples used in our study, and have

significantly revised the methods to ensure they are consistent across examples and revised the text to describe our approach in greater detail.

Comment 2: First, unclear Methodological Comparisons and Scope. The manuscript introduces multiple alternative models (e.g., MARSS, gpEDM, null models) across different case studies without a consistent rationale. For instance, the fisheries example compares UDEs to a null model and linear state-space models, while the food chain case adds Gaussian process models. This inconsistency creates confusion about the method's unique value. I would suggest focusing comparisons on state-space UDEs, traditional UDEs (without state-space components) and a null model.

Response: We have selected a consistent set of models to compare in each of our examples, including two alternative methods for training UDEs that do not rely on the state-space framework (gradient matching, shooting). To guide the reader, we have noted which models are used in which examples in Table 1. We have chosen to keep the MARSS and gpEDM models in the benchmarking set because we found it helpful to compare UDEs to linear state-space methods (MARSS models) and nonlinear non-state-space methods (gpEDMs), which are useful for explaining why the methods work when they do. Using a more consistent set of models for benchmarking gave us a more comprehensive understanding of how and when our methods work well, which we believe has strengthened our Discussion section.

Comment 3: Second, insufficient explanation of state-space framework. The distinction between the process model (dynamic equations with neural networks) and data model (observation errors) is poorly defined. The manuscript mentions the data model $h(u_t, y_t)$ but never provides its explicit form. Is the data model just likelihood equation? This omission makes it unclear how observation errors are integrated into the framework. The introduction of state-space framework in the introduction section is too abstract, not suitable for broad readers.

Response: Thank you for bringing this concern to our attention. We have extended the description of the method in our paper to provide a more detailed overview of state-space models for a general readership and to clarify the details of our specific implementation (e.g., how observation errors enter the data model).

Comment 4: Third, ambiguous cross-validation protocols. The manuscript inconsistently applies cross-validation methods. In the Method section, the authors describe two cross-validation methods. While, in the Result section, it does not clarify how the methods are applied.

Response: We agree with this concern. We initially employed two cross-validation protocols: one method designed to assess the model's forecasting skill (leave-one-out cross-validation) and a second method aimed at evaluating the model's ability to interpolate missing data within the time period of the training dataset (k-fold cross-validation). However, in retrospect, we recognize that using both approaches could make our results hard to follow. To address these issues, we applied leave-future-out cross-validation to all the examples and omitted the second approach. We also added a section in the Methods detailing the implementation of the cross-validation routine (lines 607-618).

Comment 5: Fourth, inconsistent model implementation across cases. In the food chain model, the process model introduces stochasticity via additive noise, and then the time series data were added observation errors. However, in the kelp-urchin model, autoregressive processes were used to simulate abiotic conditions but lack explicit definitions. It's unclear how process errors and observation errors were distinguished. In addition, it's unclear AR-1 is applied to X_t or v_t ?

Response: We have substantially revised the methodology in the paper to compare the state-space UDE models with a consistent set of alternatives, including both alternative methods for training UDEs (gradient matching, shooting), an alternative nonlinear time series model (gpEDM), and an alternative state-space modeling framework (MARSS). We use a consistent set of simulation and cross-validation tests to compare forecasting skill of each model across examples.

Some of the differences in how noise was handled in the simulation models have been retained, because they help us understand how the models performed on datasets that exhibit different types of dynamics. Specifically, we chose to include autoregressive errors in the kelp-urchin model because autocorrelation in abiotic conditions is ubiquitous in marine environments, and because these longer-term deviations in abiotic conditions are more likely to cause the system to transition between the two alternative stable states. Adding autoregression to the errors for the three-species food chain model was less important for dynamics we wished to test with that model, so we chose the additive model structure.

Thank you for pointing out the lack of clear definitions for when the autoregressive errors were applied. We have updated the Methods section at line 655 to clarify this point.

Other comments:

1. Standardize the presentation of simulated data generation (e.g., always include plots of raw vs. noisy data).

Response: We really like this idea, but also think our current figures help communicate some key points about the analysis and the dataset. For example, Figure 2A shows both the simulated data and the forecast to help communicate how the simulation tests were performed. Similarly, in Figure 3C we used a longer simulated dataset than was used to train the models to help illustrate the flickering behavior exhibited by the system. However, we did update the figures to more clearly present the results of the cross validation and simulation testing results.

2. Clarify how covariates (e.g., $t > 1992$ in fisheries) are encoded and integrated into neural networks.

Response: Thanks for suggesting this. We have added a discussion of this point to the Methods section (see Eq. 22).

3. Quantify how much state-space UDEs improve predictions over traditional UDEs (e.g., % reduction in RMSE).

Response: This is a great suggestion. We show a quantitative comparison of state-space UDEs compared to other UDE training methods in main text Figures 2, 3, 4, and 5.

4. In simulated cases with known ground-truth errors (e.g., food chain), evaluate whether the model accurately partitions process vs. observation errors.

Response: We agree that this could be a valuable analysis to add to the paper; however, our primary goal in this manuscript is to characterize the state-space UDE framework's ability to recover nonlinear dynamics in the context of uncertainty. We center this manuscript on testing how well these methods work compared to alternative approaches and under different types of dynamics. Quantifying uncertainty for these methods is beyond the scope of this manuscript.

5. Specify how the nine simulated datasets in Figure 3f differ.

Response: In all cases, the different simulated datasets used different sequences of random numbers for both observation and process error terms. We have clarified this

point in the caption for Figure 3. We also added figures in the Supplemental Information that show all bifurcation diagrams (Figs. S2.1-S2.4)

6.It's better to have both the simulated data and empirical data in the same cases.

Response: This could be a valuable method for ensuring greater consistency between examples and increasing the rigor of our tests. We will strongly consider this comment in future work. However, using simulated and empirical examples from different systems allowed us to test the modeling technique in a broader range of conditions, which we also think is valuable.

My co-authors and I have made several substantial changes to the manuscript in response to your comments and requests. These revisions have primarily focused on updating the discussion section of the manuscript to draw out the key ecological insights from our case studies and describe the implications of our method for future studies in ecology. We have also addressed specific concerns about our mathematical notation and terminology raised by the reviewers. Detailed responses to each concern are provided below.

Major Comments on ecological insight and interpretation:

Editor: In light of their advice, we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment, provided that you can add more ecological insights based on real data applications, as raised by reviewer #1.

Reviewer 1: While the lack of methodological innovation is understandable given the journal's emphasis on applied ecological insight, this makes it even more important for the manuscript to clearly demonstrate the ecological understanding gained from the framework. Although the manuscript dedicates considerable space to methodological exposition and simulation results, its discussion remains largely mathematical and generic, offering limited ecological interpretation.

Response: Key applications of our method in ecology are leveraging a neural network to help quantify and explain nonlinear functional responses and species-species and species-environment interactions. In this way, our method enables deeper ecological understanding. These applications are tested with simulated and empirical data in examples two, three, and four, with results presented in Figures 3-5. Our findings highlight that UDEs can be used to identify these relationships, which have implications for experimental design, determining what data and model parameters are required to estimate the relationships, and deepening theoretical understanding of the fundamental ecological processes governing ecosystem dynamics.

We also show how estimating nonlinear interaction terms can help inform our understanding of the stability and resilience of different ecosystem states. In examples two and four, we explicitly derive bifurcation diagrams that show the stability and basins of attraction around alternative ecosystem states. This is a new approach to quantifying bifurcation diagrams for identifying ecosystem regime shifts. In example three, we calculated the feedback between a fish population and fishing effort. In this context, tighter feedback loops lead to higher resilience.

To improve the manuscript in light of your comments, we have expanded two paragraphs (lines 396 - 410) in the discussion section that describe the methods' potential to capture nonlinear species interaction terms and quantify the stability of ecosystem states. We have also added a paragraph discussing other possible applications of the technique to ecological problems (lines 412-423). We have also added a section to the introduction highlighting the components of our analysis focused on issues specific to ecology (lines 90-97).

Line-by-line comments:

Comment: In Equation 6, the function ϕ is not defined clearly.

Response: ϕ is the density function of a multivariate normal distribution. Now clearly defined in the text on line 155: “where Σ_ϵ is a diagonal matrix with $\Sigma_{\epsilon,t,t} = \sigma_t^2$, ϕ is the probability density function for a multivariate normal distribution, and the function UDE stands in for the right-hand sides of Eq. 4 or Eq. 5, depending on the model’s formulation.”

Comment: The paper does not specify any prior information for the initial state values. It would be helpful to describe the assumed prior distribution or initialization procedure used by the UKF.

Response: The UKF is initialized at the first time point with a multivariate equal to the first observation and variance equal to the observation errors. We updated line 608 to reflect this: “The algorithm is initialized at the first time point in the data set using the observed value as the initial mean ($\hat{x}_0 = y_0$) and the observation errors as the initial covariance ($\Sigma_{u,0} = \Sigma_\epsilon$).”

Comment: Line 163, the covariance matrix $\Sigma_{\epsilon,i,t}$ has t in its subscript; if it is time-invariant, then t can be dropped.

Response: Good catch, we dropped the subscript. Thank you.

Comment: Line 174, the notation $p(u_{1:T}|y_{1:T},\theta)$ is ambiguous. The standard UKF returns the filtered posterior conditioned on observations up to time t , i.e. $p(u_t|y_{1:t},\theta)$. If the posterior conditioned on all observations up to T is intended, it should be referred to as the smoothed posterior.

Response: We intended the filtered posterior and have adopted the notation you suggested. “Evaluating Eq. 7 requires an approximation of the distribution of the states given the observations and the parameter estimates $p(u_t|y_{1:t},\theta)$.”

Comment: Line 181. The statement “When training with the joint likelihood, these terms need to be fixed at pre-defined values” reflects a common practical choice, but it is not a strict requirement. These variance terms can, in principle, be treated as trainable parameters within the joint-likelihood framework.

Response: Good catch, we have updated our text to reflect that this is a modeling choice, not a mathematical necessity: “When training with the joint-likelihood, we chose these terms to be fixed at pre-defined values.”

Comment: Example 1 is labeled numerically, but the subsequent examples are not.

Response: We have added labels to the other example sections in the methods.

Comment: In Equations 29 and 30, for the stochastic differential equation, the notation dW is unclear. If it is the Wiener process, then it should be dW_t instead of dW .

Response: We also corrected this notation in equations 16, 17, and 18. Thank you.