Biodiversity increases the forecastability of species abundances in changing environments

Romana Limberger[1§*], Uriah Daugaard[1§*], Yves Choffat[1], Anubhav Gupta[1], Martina Jelić[1], Sabina Jyrkinen[1], Rainer M. Krug[1], Seraina Nohl[1], Frank Pennekamp[1], Sofia J. van Moorsel[1,2], Xue Zheng[1], Debra Zuppinger-Dingley[1], Owen L. Petchey[1]

§ These authors contributed equally: Romana Limberger, Uriah Daugaard

[1]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

[2]Department of Geography, University of Zurich, Zurich, Switzerland

***Corresponding authors:**

Romana Limberger

Department of Evolutionary Biology and Environmental Studies

University of Zurich

Winterthurerstrasse 190

CH-8057 Zurich, Switzerland

e-mail: romana.limberger@ieu.uzh.ch

Uriah Daugaard

Department of Evolutionary Biology and Environmental Studies

University of Zurich

Winterthurerstrasse 190

CH-8057 Zurich, Switzerland

e-mail: uriah.daugaard@ieu.uzh.ch

Abstract

Accurate forecasts of ecological dynamics are critical for ecosystem management and conservation, yet the drivers of forecastability are poorly understood. Here we show that experiments are a powerful (but underutilized) tool to explore the limits of ecological forecasting. We conducted a long-term microcosm experiment with aquatic protists and manipulated two potential drivers of forecastability: biodiversity and environmental change (light). We applied data-driven forecasting methods to finely-resolved time series of species abundances and ecosystem properties. The experimental treatments had interactive effects on abundance forecasts. Specifically, biodiversity increased forecastability in changing environments but reduced forecastability in constant environments. However, forecasts of ecosystem properties were only weakly affected by the experimental treatments. Our results indicate that the combination of biodiversity loss and environmental change could make populations particularly unpredictable. The implications for ecological conservation are worrying: diversity loss might impair decision-taking by reducing the forecastability of ecological dynamics in changing environments.

Key words: biodiversity, complexity, ecological forecasting, empirical dynamic modeling (EDM), environmental change, light, machine learning, microcosm experiment, microorganisms, predictability, species richness

Introduction

Forecasting the future state of populations, communities, and ecosystems is challenging but necessary to take well-informed decisions in ecosystem conservation and management[1–4]. Ecological dynamics are difficult to forecast because organisms are embedded in complex networks of species interactions, respond simultaneously to many different drivers, often in a nonlinear way, and can display stochastic or chaotic dynamics[5,6]. However, global change and biodiversity loss have made ecological forecasting all the more relevant, resulting in rising numbers of published forecasts[7]. Typically, forecasts are based on observational data of organisms that are relevant in conservation and management (e.g. endangered species, harmful algae)[7]. Likewise, attempts to gain a more general understanding about the constraints of forecastability have leveraged observational data to compare forecast skill across different species and ecosystem properties[7,8]. Only few studies, however, investigated potential drivers of forecastability with experimental approaches[9–12]. Forecasts of ecological dynamics require long and finely-resolved time series[13], which can be challenging to accomplish with experiments. Yet, controlled experiments have been a fundamental tool in many fields of ecology[14–17] and could be useful to advance our understanding about the potential and the limits of ecological forecasting.

The complexity of communities is at the heart of the notion that ecological systems are difficult to forecast. However, this idea has never been tested with an experimental manipulation of community complexity (e.g. biodiversity). Communities with higher species richness have the potential for more direct and indirect pathways among species[18,19], making the outcome of species interactions less predictable[20,21]. Moreover, higher biodiversity often increases the temporal variability of species abundances[22,23]. Whether higher variability leads to lower forecastability (i.e. to greater discrepancy of predicted and observed values) depends on the

nature of the variation, with more stochastic or chaotic variability resulting in reduced forecastability[24].

The effect of biodiversity on forecastability could also depend on whether we forecast species abundances or aggregate ecosystem properties (e.g. community biomass). Biodiversity can have contrasting effects on the temporal variation of population abundances and community biomass[25,26], which might translate into contrasting effects on forecastability. In general, aggregate properties are expected to be easier to forecast because they might integrate over environmental variability[27].

When environmental conditions change, ecological forecasting is particularly relevant but could also be particularly difficult. Data-driven forecasting methods (e.g. machine learning) perform better when past time points hold more information about the future[24]. Changes in abiotic conditions might alter population dynamics[28,29], and could therefore reduce forecastability due to greater dissimilarity of past and future dynamics. In addition, abiotic change often alters species interactions[30,31], which could further constrain forecastability[11]. On the other hand, environmental change can reduce stochasticity[32] and increase autocorrelation of population time series[33,34]. Consequently, future states of such systems depend more on past states, and should thus be more predictable[24]. Despite the importance of ecological forecasting in changing environments, we lack empirical studies that test how environmental change influences forecastability.

Here we present a microcosm experiment about the effect of biodiversity and environmental change on the forecastability of ecological dynamics. Our experiment is a unique combination of high numbers of time points (123), experimental units (30), and experimental taxa (18). We assembled communities of aquatic protists with three different levels of species

richness (7, 10, 14 species) and exposed the communities to constant or gradually declining light levels. We chose light as environmental driver because aquatic light regimes are changing, primarily due to the browning of lakes resulting from human-induced changes in land-use and precipitation patterns[35–37]. More generally, however, our manipulation of light is exemplary for other anthropogenic pressures that are changing gradually and directionally[38]. We sampled our experimental communities three times per week for 41 weeks, and measured species abundances and two aggregate ecosystem properties (community biomass and oxygen). We then forecasted these variables using data-driven forecasting models, and quantified forecast error, i.e. the deviation of observed and predicted values. With this dataset we tested the hypotheses that (1) higher biodiversity leads to higher forecast error of species abundances, (2) directional environmental change increases forecast error because forecasts are made for less-often observed conditions, and (3) forecast error of aggregate ecosystem properties is lower than forecast error of species abundances and decreases with biodiversity.

Results

Effects of species richness and environmental change on forecast error

Diversity and environmental change had interactive effects on the forecast error of species abundances (Figure 1a, Table 1). Forecast error increased with species richness in constant light but decreased with richness in declining light (Figure 1a, Table 1). Specifically, when median richness increased by one species, forecast error increased by 5.6% (95% CI: -0.8% – 12.1%) in constant light and decreased by 4.6% (95% CI: -11.0% – 1.0%) in declining light relative to the baseline forecast error of 1 (i.e. the expected forecast error when the average value of a time series is used as the prediction).

The reported treatment effects on forecast error of species abundances persisted in several robustness analyses. To exclude that the result depended on the forecasting method, we compared the outcome of five different forecasting methods: Auto-Regressive Integrated Moving Average (ARIMA), Simplex Empirical Dynamic Modeling (EDM), Multiview EDM, Random Forests (RF), and Recurrent Neural Networks (RNN). Both the direction and the size of the treatment effects were similar among the five forecasting methods (Supplement Figure S18). We also tested if treatment effects depended on the richness variable. Some species did not establish or persist in our communities, and we therefore used median observed richness rather than planned richness as explanatory variable in our model. Treatment effects were similar when using other richness variables (including planned richness; Supplement Figure S18). Likewise, the results were robust to the method used to make time series stationary prior to forecasting (simple detrending versus segmented detrending; Supplement Section 7.2). Finally, the reported treatment effects were not driven by individual species (Supplement Figure S21).

Forecastability of aggregate ecosystem properties was largely unaffected by the experimental treatments. Light and richness had no effect on the forecast error of total community biomass (Table 1, Figure 1b) except when biomass was forecasted with RNN (Supplementary Table S7, Figure S16). In this case, treatments had similar effects as on forecast error of species abundances (i.e. light x richness interaction). Forecast error of oxygen was unaffected by light and richness (Supplementary Table S8, Figure S17). For most forecasting methods, forecast error of aggregate properties (community biomass, oxygen) did not differ from average forecast error of species abundances (Supplementary Tables S9-S11).

Explanations of the treatment effects on forecast error

To investigate possible reasons for differences in forecast error, we analyzed how the forecast error of species abundances was related with time series properties. Therefore, we calculated three metrics for each time series: autocorrelation, coefficient of variation (as a measure of temporal variability), and permutation entropy (as a measure of time series complexity[24]). We expected that greater autocorrelation and lower permutation entropy would be associated with lower forecast error, while greater variability (i.e. higher coefficient of variation) could be associated with lower or higher forecast error, depending on whether higher variability is associated with more or less determinism.

Forecast error decreased with increasing autocorrelation and coefficient of variation, and increased with increasing permutation entropy (Figure 2 a-c, Supplement Table S12). The experimental treatments affected the three time series metrics, albeit not in an interactive way: autocorrelation and coefficient of variation were lower in declining light than in constant light, whereas permutation entropy was unaffected by light but increased with richness (Figure 2 d-f, Supplement Table S13).

We used a structural equation model (SEM) to disentangle direct and indirect effects of richness and time series metrics on forecast error of species abundances (Figure 3, Supplementary Table S14). A multi-group model provided a good match between model-implied and observed covariance ($\chi^2 = 3.219$, DF = 4, p = 0.522, i.e., no evidence against a good model fit). The model explained 32.2% of the variation in forecast error in constant light, and 9.7% of the variation in forecast error in declining light (R^2, Table S15). In line with the bivariate correlation analyses, the SEM showed in both light conditions that forecast error declined with increasing autocorrelation, declined with an increasing coefficient of variation, and increased

with increasing permutation entropy (Figure 3; note that not all of these effects were statistically significant).

The SEM confirmed the interactive effect of light and richness on forecast error. Richness had a positive (albeit not significant) effect on forecast error in constant light but reduced forecast error in decreasing light (Figure 3, Supplementary Table S14). Furthermore, the effect of species richness on coefficient of variation and autocorrelation was reversed between the two light treatments (Figure 3). Although not significant, these reversed effects of richness on coefficient of variation and autocorrelation in the two light treatments might explain why increasing richness had opposite effects on forecast error in the two light treatments.

Discussion

Our experiment revealed that biodiversity increases forecastability (i.e. reduces forecast error) of species abundances in changing environments. In constant conditions, however, biodiversity tended to reduce forecastability. That is, only in constant environments did we find any support for the long-standing assumption[5] that biotic complexity makes ecological dynamics hard to forecast. Our hypothesis that environmental change reduces forecastability was supported only when diversity was low. Taken together, our results suggest that the combination of biodiversity loss and global change could make populations particularly unpredictable.

Environmental change increased the forecastability of species abundances at the upper end of the diversity gradient. This positive effect of environmental change on forecastability potentially resulted from higher autocorrelation in changing conditions (Figure 2d). Higher correlation between successive time points means that past states have higher information content, resulting in better performance of forecasting models[8]. Accordingly, higher autocorrelation was associated

with lower forecast error (Figure 2a, Figure 3). Higher autocorrelation in changing conditions was due to both direct and indirect effects of the environmental driver. Decreasing light put some species on deterministic trajectories of population decline, which indirectly led to abundance increases of inferior competitors and prey organisms (Supplement Figures S22, S23). Such indirect effects of environmental change via altered species interactions are widespread[39–41], and may be more frequent than direct effects[42]. As such, autocorrelation can spread through the community even if only few species are directly affected by environmental change. Taken together, our results illustrate that forecastability of species abundances can be high in changing environments if direct and indirect effects of environmental change result in higher autocorrelation of time series, and therefore in greater information content of past states.

When diversity was low, environmental change decreased the forecastability of species abundances. While low sample size (two compositions) at the very low end of the diversity gradient led to inconclusive statistical analyses, we can speculate about possible mechanisms. First, the difference in autocorrelation between constant and decreasing light was smaller when diversity was low (Figure 2d), maybe because lower species richness reduced the potential for high autocorrelation mediated by changes in species interactions. Second, in the two communities with lowest diversity, an entire functional group (i.e. consumers of algae) was represented by only one species, which declined or suddenly increased in response to the environmental driver, resulting in pronounced changes in prey abundances (Supplement Section S8.1). That is, previously unobserved dynamics towards the end of the experiment might have contributed to the high forecast error in these communities. In more diverse communities, however, the consumer functional group thrived across the entire light gradient and did not change suddenly as conditions changed. For example, the decline of the dominant consumer in

the decreasing light treatment was compensated by increases in other consumer species (Supplement Section S8.1). This process is known as the insurance effect of diversity[43,44], i.e. the loss of species due to environmental change is compensated by other species that perform similar functions but tolerate and potentially thrive in the new environmental conditions[45–47]. This mechanism might have contributed to the positive effect of diversity on forecastability in the changing light treatment by maintaining functional groups in the community, thereby preventing pronounced changes in population dynamics.

From an ecosystem management perspective, our results suggest that biodiversity conservation is fundamental for accurate population forecasts in changing environments. More-diverse communities might be less prone to ecological surprises as environmental conditions change because biodiversity can promote the resistance of functional groups[47,48], hinder the immigration of new species[49,50], and delay abrupt shifts to new ecosystem states[51]. The loss or rise of functional groups will have large impacts if it concerns species with strong ecosystem-effects. For example, the decline of keystone predators or consumers can have cascading effects on other groups[39,52], potentially leading to novel dynamics of many species in the community, and therefore to overall low forecastability. Taken together, biodiversity loss in changing environments might promote ecological surprises that reduce the forecastability of species abundances, thereby hampering decision-taking in ecosystem management and conservation.

In contrast to our hypothesis, the forecastability of aggregate properties (i.e. community biomass, oxygen) was largely unaffected by our experimental treatments. Trends of treatment effects went in the same rather than opposite direction as effects on forecastability of species abundances. We had expected that contrasting effects of biodiversity on the temporal variability of aggregate and non-aggregate properties[25,26] may translate into contrasting effects on

forecastability. However, in our experiment, species richness affected neither the temporal variability of population abundances (Figure 2e) nor that of community biomass or oxygen (Supplement Figure S24). In general, the experimental treatments had similar effects on time series properties irrespective of the level of aggregation of the response variable, possibly explaining the similarity in treatment effects on forecastability. Overall, our results suggest that biodiversity and environmental change are less influential in affecting the forecastability of aggregate ecosystem properties.

Also in contrast to our expectations, forecastability of community biomass did not differ from average forecastability of species abundances. Forecastability is expected to increase with the level of aggregation because aggregate properties might integrate over environmental variability[27]. Indeed, community biomass had lower temporal variability than the biomass of individual species (Supplement Figure S25). However, our results suggests that lower temporal variability does not necessarily translate into higher forecastability. Rather, we observed the opposite pattern: time series with higher variability were forecasted better, indicating that higher temporal variability did not reflect more stochastic but more deterministic variation (e.g. abundance changes in response to the light decline or to changes in competitor abundances). At the level of the community, compensatory dynamics among species might have led to reduced variability of community biomass without increasing its forecastability because biomass integrated over primarily deterministic (not stochastic) variation. In conclusion, our results do not support the prospect that forecastability increases by aggregating individual ecosystem components.

Our study illustrates that experiments with microbial organisms are a useful tool to investigate potential drivers of forecastability. Data-driven forecasting methods require long time

series[13], which can be collected within a comparatively short time frame for microbial organisms because of their short generation times. Moreover, the high level of control in laboratory microcosm experiments reduces confounding factors, which is more difficult in observational studies, in particular over the long timescales that are necessary for forecasting studies. However, such experiments also come with limitations. Specifically, the diversity of our experimental communities was lower than that of natural protist communities. Yet, our diversity gradient was in the range of previous studies which used experimental protist communities to draw important conclusions about the effect of diversity on the stability of communities[53,54].

In conclusion, our results suggest that biodiversity does not necessarily reduce the forecastability of ecological systems. Rather, when environmental conditions change, biodiversity might promote forecastability by buffering against ecological surprises that can result from a decline or rise of functional groups. In contrast, low diversity in combination with environmental change makes ecosystems more unpredictable. Our results therefore caution against assuming high predictability of ecosystems with low biotic complexity. Forecasting efforts for management or conservation purposes might require particularly finely-resolved time series in ecosystems of low diversity, so as not to miss unexpected changes in ecological dynamics as environmental conditions change. Considering anthropogenic pressures on biodiversity, reduced predictability of less-diverse systems may pose an additional challenge to conserve populations, communities, and ecosystems in changing environments.

Methods

Experimental design

In a nine-month experiment, we factorially manipulated light (2 levels) and species richness (3 levels). At each of the three richness levels, we assembled five different community compositions (see below for details), resulting in a total of 30 microcosms. Light intensity was either constantly high or declined gradually. The declining light treatment involved three phases: (1) three months of constantly high light, (2) three months of gradual light decline, and (3) three months at the final low light level (Supplementary Figure S1). To manipulate light, we used eight incubators (IPP260plus, Memmert), four per light level. Light was either 30% of the maximum incubator light intensity (~ 22,500 Lux) or declined in phase 2 from 30% to 1% (~ 800 Lux). Throughout the entire experiment, all incubators were set to a light:dark cycle of 16:8 hours and a temperature of 18°C.

We manipulated species diversity by assembling protist communities with different numbers of species (7, 10, and 14 species, respectively). The species were drawn from a pool of 18 species of algae and ciliates, comprising six functional groups: edible algae (two species), inedible algae (5), bacterivorous ciliates (4), omnivorous ciliates (4), mixotrophic ciliates (2), and a predatory ciliate (1) (Supplementary Table S1). The three diversity levels were generated by manipulating diversity within the functional groups of inedible algae, bacterivores, and omnivores (one, two, or three species per functional group), and in the mixotrophs (one, one, or two species) (Supplementary Table S2). To avoid confounding effects of species richness and species identity/composition[55,56], we assembled five different compositions at each diversity level. The compositions were as dissimilar as possible, and species occurred with similar

frequency both within and across diversity levels (see Supplement Section S1.3 for computation of community compositions).

Implementation of the experiment

The experiment started with a four-week inoculation phase. We filled 30 2L-bottles with 670 mL of medium, 20 wheat seeds for slow release of nutrients, and a magnetic stirrer for homogenization during samplings. Over a period of four weeks, we added a total of 330 mL of stock cultures growing at high densities, starting with prey (three species of bacteria and two edible algae) and inedible algae, followed by ciliate consumers, and finally the predator. When inoculation was complete, the composition of the medium in the bottles was 75% filtered protist pellet medium (0.55 g protist pellet in 1L Chalkley's medium) and 25% WC medium. Diversity was manipulated in a substitutive design, e.g. the volume added of inedible algae (80 mL) was divided among one to three species depending on the diversity level. During inoculation, we homogenized bacteria associated with algae and ciliate cultures so that communities did not differ in their initial bacterial composition. To this end, we added filtrates (< 1.2 µm) of all inedible algae and ciliates to each bottle. See Supplement Section S1.5 for further details on inoculation.

After the inoculation phase, we sampled the 30 bottles three times per week (Monday, Wednesday, Friday) for nine months (41 weeks), resulting in 123 sampling days. On sampling days, we removed 50 mL from each bottle after homogenization on a magnetic stirrer, and then added 50 mL of fresh medium to the bottle. Samples were taken under sterile conditions using 50-mL glass pipettes.

To allow species to re-colonize bottles if they had gone extinct, we added the species from the stock cultures every three weeks, mimicking a very low rate of immigration. In protist

microcosms, species often go extinct through competitive or trophic interactions and would have no chance of re-colonizing the microcosm when conditions become favourable. At these immigration events, we added a very low number (10 individuals) of each species that was part of the community of a given bottle. Soon after the start of the experiment, two species went extinct in the stock cultures (*Stylonychia mytilus* and *Didinium nasutum*) and could therefore not be re-immigrated. Stock cultures were cultivated at high light conditions (i.e. 30% light).

Quantification of species abundances and ecosystem properties

We measured abiotic ecosystem properties (dissolved oxygen, carbon, and nitrogen), species abundances of ciliates and algae, and total abundance of bacteria. To quantify dissolved oxygen, we used a hand-held oxygen meter (PreSens Precision Sensing GmbH, Germany) and two optical oxygen sensors attached to the inner wall of each bottle. We measured dissolved organic carbon (DOC) and dissolved nitrogen with a TOC/TN analyzer (Formacs HT). We quantified bacterial abundance with a flowcytometer (BD Accuri C6 Plus; BD Biosciences), algae abundances and morphological traits with a FlowCAM (VS1; Yokogawa Fluid Imaging Technologies), and ciliate abundances and movement and morphological traits with video microscopy (R-package *bemovi*[57]). The predatory ciliate *Didinium* was counted manually under a dissecting microscope when present.

We identified the recorded algae and ciliate individuals by using support vector machine (SVM) classifiers (R-package e1071[58]). We trained these classifiers with trait data collected from monocultures cultivated at different light levels. We further improved classifier performance by extending the training data with manually annotated data from the experiment, if necessary, and by including additional classes that occurred in the experimental communities but not in the

monocultures (e.g. debris, *Chlamydomonas* clumps). The classifiers were composition-specific, i.e. the training data set for a given classifier only contained taxa that could potentially occur in the respective composition. See Supplement Section S2 for further details on measurements and classifications.

Lastly, we computed community biomass by summing the biomass of all algae, ciliates, and bacteria in a sample. To this end, we estimated the volume of each particle classified as ciliate or algae by using their extracted morphological traits and their three-dimensional shape, determined based on Hillebrand et al.[59] and on personal observations. To calculate the volume of bacteria, we first quantified the relationship between forward scatter and particle size by measuring beads of known size on our flowcytometer. We used this calibration curve to estimate the size of the bacteria measured in our samples, and then computed their volume assuming an ellipsoidal shape. We converted volume to biomass assuming that the organisms had the density of water. Further details are given in Supplement Section S4.3.

Processing of recorded time series

After data collection and classification, we carried out several quality controls and made corrections if needed (for a complete description see Supplement Section S3). Then, we averaged over technical replicates made during sampling with flowcytometry, video microscopy, and oxygen meter.

The recorded data included some missing data (flowCAM: 0.81%, flowcytometer: 4.84%, videos: 0.84%, oxygen meter: 1.67%, TOC/TN analyzer: 0.92%, Supplementary Table S5). We used cubic hermite splines to impute the missing data and to interpolate the time series such that all time points were equidistant (time step of 2.3 days). To make the time series stationary, we

detrended and standardized them, as is commonly done in time series analyses[9] by taking the standardized residuals of the regressions of the recorded values against time. See Supplementary Figures S5-S12 for untransformed and transformed time series.

Forecasting

We forecasted species abundances and two aggregate ecosystem properties (total community biomass and oxygen concentration). The forecasted species (i.e. forecast targets) included the ciliate and algae species, bacteria (considered one group), and additional morphotypes ("*Chlamydomonas* clumps", "*Desmodesmus* clumps", "Dividing *Chlamydomonas*" and "Small cells", i.e. small unidentified cells), 14 targets in total. However, we did not forecast the abundance of a target if it went extinct during the experiment (Supplement Section S3.2). Across all bottles, we forecasted 253 taxa time series.

We used different forecast methods to ensure that the results were independent of the chosen methodology. We used one linear model (Auto-Regressive Integrated Moving Average, i.e. ARIMA), two nonlinear approaches based on Empirical Dynamic Modeling (Simplex EDM and Multiview EDM[60,61]), and two machine learning methods (Random Forest RF and Recurrent Neural Networks RNN). Simplex and Multiview EDM differ insofar as the former allows better parameter optimization, while the latter is designed for high-dimensional systems. See Supplement Section S5 for a more detailed description of the forecast models and their parametrization.

We used the first 111 time points of the time series to train the models and then forecasted the remaining 12 time points (i.e. ~10% of the data, corresponding to the last 4 weeks of sampling) with one-step-ahead forecasts. Two of the forecast methods (ARIMA, Simplex EDM)

included only the forecast target itself as predictor, whereas the other three methods included additional predictors. As predictors we used all forecast targets of a given bottle as well as the dissolved carbon and nitrogen concentrations. We quantified the forecast error (i.e. the disagreement between forecasted and recorded values) as the root mean square error (RMSE). Because we forecasted standardized time series, an RMSE value below one indicates a better performing model than the average value of the time series (which has an expected RMSE of one).

Forecast error analyses

We computed several linear mixed models (R-package *lme4*[62]) to analyze the effects of the experimental treatments on forecast error of either species abundances, community biomass, or oxygen. As explanatory variables, we used light condition (binary), species richness (continuous), and their interaction. Observed species richness deviated from planned richness because not all species persisted throughout the experiment (Supplement Section S3.1). Furthermore, in one bottle the alga *Desmodesmus* was introduced by cross-contamination. We therefore used realized richness (i.e. the median of the observed richness values) rather than planned richness in our models. Realized species richness was unaffected by the light conditions and was positively correlated with planned species richness (for more information and alternative realized species richness metrics see Supplement Section S4.1). In all models, we included random intercepts following the experimental design and depending on the model fitted (e.g. "Taxon" was used for the abundance forecasts mixed model, but not in the oxygen model), as specified in the model tables in Supplement Section S6. We used *F*-tests to assess the importance of covariates and interactions.

We also investigated if treatment effects depended on the level of aggregation of the response variable. In this case, the response variable included the mean forecast error of species abundances and the forecast error of community biomass. As explanatory variables, we used realized richness, light condition, a binary variable that specified the level of aggregation (species abundances vs. community biomass), and all possible pairwise interaction terms. Bottle and incubator were included as random effects. Finally, we repeated this model for the forecast error of oxygen instead of community biomass.

Time series properties

To explore possible mechanisms behind observed treatment effects on forecast error, we investigated how forecast error was related with time series properties. For each time series, we calculated three metrics: the coefficient of variation, autocorrelation, and permutation entropy. The coefficient of variation is a measure of temporal variability and was calculated as the ratio of the standard deviation of the detrended time series and the mean of the original time series. Autocorrelation describes the correlation of current and past values and was calculated for a lag of 5 (i.e. correlation with values 5 time points ago); the choice of lag did not affect the results. Permutation entropy is a measure of time series complexity and is inversely related with the intrinsic predictability of a time series[24]. We calculated autocorrelation and permutation entropy based on the transformed time series. See Supplement Section S4.2 for further details.

We used two approaches to investigate the relationship of the time series metrics with the experimental treatments and with forecast error of species abundances. First, we computed separate linear mixed models to test (i) how the experimental treatments (realized richness, light, and their interaction) affected each of the three time series metrics, and (ii) how each of the three

metrics, light condition, and their interaction affected forecast error. Again, we included appropriate random factors in the models (see model tables in Supplement Section S6) and used *F*-tests to assess the importance of model covariates. As a second approach, we computed a structural equation model (SEM) to further clarify the relations found with the first approach. For this, we used the R-packages *lavaan*[63] and *lavaan.survey*[64]. *Lavaan.survey* allowed us to include random effects according to our experimental design. Furthermore, we set up the model to estimate the relations between the variables in both light conditions by using a multiple group analysis. We tested the significance of SEM path coefficients with two-sided t-tests. All data analyses and forecasts were done in R (version 4.3, package information in Supplement Section S9).

Data availability statement

All data and code used in this study will be made **publicly** available on Zenodo upon publication (DOI: https://doi.org/10.5281/zenodo.11067497). At this stage, file access is restricted. Access will be given to editors and reviewers upon request.

References

1. Clark, J. S. *et al.* Ecological forecasts: an emerging imperative. *Science* **293**, 657–660 (2001).
2. Dietze, M. C. *et al.* Iterative near-term ecological forecasting: needs, opportunities, and challenges. *Proc. Natl. Acad. Sci.* **115**, 1424–1432 (2018).
3. Henden, J.-A. *et al.* End-user involvement to improve predictions and management of populations with complex dynamics and multiple drivers. *Ecol. Appl.* **30**, e02120 (2020).

4. Deyle, E. R. *et al.* A hybrid empirical and parametric approach for managing ecosystem complexity: water quality in Lake Geneva under nonstationary futures. *Proc. Natl. Acad. Sci.* **119**, e2102466119 (2022).
5. Beckage, B., Gross, L. J. & Kauffman, S. The limits to prediction in ecological systems. *Ecosphere* **2**, art125 (2011).
6. Doak, D. F. *et al.* Understanding and predicting ecological dynamics: are major surprises inevitable? *Ecology* **89**, 952–961 (2008).
7. Lewis, A. S. L. *et al.* Increased adoption of best practices in ecological forecasting enables comparisons of forecastability. *Ecol. Appl.* **32**, e2500 (2022).
8. Ward, E. J., Holmes, E. E., Thorson, J. T. & Collen, B. Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. *Oikos* **123**, 652–661 (2014).
9. Benincà, E. *et al.* Chaos in a long-term experiment with a plankton community. *Nature* **451**, 822–825 (2008).
10. Daugaard, U., Munch, S. B., Inauen, D., Pennekamp, F. & Petchey, O. L. Forecasting in the face of ecological complexity: number and strength of species interactions determine forecast skill in ecological communities. *Ecol. Lett.* **25**, 1974–1985 (2022).
11. Dumandan, P. K. T., Simonis, J. L., Yenni, G. M., Ernest, S. K. M. & White, E. P. Transferability of ecological forecasting models to novel biotic conditions in a long-term experimental study. *bioRxiv* (2023) doi:https://doi.org/10.1101/2023.11.01.565145.
12. Fujita, H. *et al.* Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics. *Microbiome* **11**, 63 (2023).
13. Dietze, M. C. *Ecological Forecasting*. (Princeton University Press, Princeton, 2017).

14. Jessup, C. M. *et al.* Big questions, small worlds: microbial model systems in ecology. *Trends Ecol. Evol.* **19**, 189–197 (2004).

15. Jochum, M. *et al.* The results of biodiversity–ecosystem functioning experiments are realistic. *Nat. Ecol. Evol.* **4**, 1485–1494 (2020).

16. Petchey, O. L., McPhearson, P. T., Casey, T. M. & Morin, P. J. Environmental warming alters food-web structure and ecosystem function. *Nature* **402**, 69–72 (1999).

17. Schindler, D. E., Carpenter, S. R., Cole, J. J., Kitchell, J. F. & Pace, M. L. Influence of food web structure on carbon exchange between lakes and the atmosphere. *Science* **277**, 248–251 (1997).

18. Polis, G. A. & Strong, D. R. Food web complexity and community dynamics. *Am. Nat.* **147**, 813–846 (1996).

19. Yodzis, P. Diffuse effects in food webs. *Ecology* **81**, 261–266 (2000).

20. Montoya, J. M., Woodward, G., Emmerson, M. C. & Solé, R. V. Press perturbations and indirect effects in real food webs. *Ecology* **90**, 2426–2433 (2009).

21. Novak, M. *et al.* Predicting community responses to perturbations in the face of imperfect knowledge and network complexity. *Ecology* **92**, 836–846 (2011).

22. Hector, A. *et al.* General stabilizing effects of plant diversity on grassland productivity through population asynchrony and overyielding. *Ecology* **91**, 2213–2220 (2010).

23. Hu, J., Amor, D. R., Barbier, M., Bunin, G. & Gore, J. Emergent phases of ecological diversity and dynamics mapped in microcosms. *Science* **378**, 85–89 (2022).

24. Pennekamp, F. *et al.* The intrinsic predictability of ecological time series and its potential to guide forecasting. *Ecol. Monogr.* **89**, e01359 (2019).

25. Tilman, D. Biodiversity: population versus ecosystem stability. *Ecology* **77**, 350–363 (1996).

26. Xu, Q. *et al.* Consistently positive effect of species diversity on ecosystem, but not population, temporal stability. *Ecol. Lett.* **24**, 2256–2266 (2021).

27. Lewis, A. S. L. *et al.* The power of forecasts to advance ecological theory. *Methods Ecol. Evol.* **14**, 746–756 (2023).

28. Fussmann, G. F., Ellner, S. P., Shertzer, K. W. & Jr, N. G. H. Crossing the Hopf bifurcation in a live predator-prey system. *Science* **290**, 1358–1360 (2000).

29. Fussmann, K. E., Schwarzmüller, F., Brose, U., Jousset, A. & Rall, B. C. Ecological stability in response to warming. *Nat. Clim. Change* **4**, 206–210 (2014).

30. Blois, J. L., Zarnetske, P. L., Fitzpatrick, M. C. & Finnegan, S. Climate change and the past, present, and future of biotic interactions. *Science* **341**, 499–504 (2013).

31. Gilman, S. E., Urban, M. C., Tewksbury, J., Gilchrist, G. W. & Holt, R. D. A framework for community interactions under climate change. *Trends Ecol. Evol.* **25**, 325–331 (2010).

32. Blasius, B., Rudolf, L., Weithoff, G., Gaedke, U. & Fussmann, G. F. Long-term cyclic persistence in an experimental predator–prey system. *Nature* **577**, 226–230 (2020).

33. Drake, J. M. & Griffen, B. D. Early warning signals of extinction in deteriorating environments. *Nature* **467**, 456–459 (2010).

34. Pace, M. L. *et al.* Reversal of a cyanobacterial bloom in response to early warnings. *Proc. Natl. Acad. Sci.* **114**, 352–357 (2017).

35. Leech, D. M., Pollard, A. I., Labou, S. G. & Hampton, S. E. Fewer blue lakes and more murky lakes across the continental U.S.: Implications for planktonic food webs. *Limnol. Oceanogr.* **63**, 2661–2680 (2018).

36. Rose, K. C., Greb, S. R., Diebel, M. & Turner, M. G. Annual precipitation regulates spatial and temporal drivers of lake water clarity. *Ecol. Appl.* **27**, 632–643 (2017).

37. Wollschläger, J., Neale, P. J., North, R. L., Striebel, M. & Zielinski, O. Editorial: Climate change and light in aquatic ecosystems: variability & ecological consequences. *Front. Mar. Sci.* **8**, (2021).

38. Tilman, D. *et al.* Forecasting agriculturally driven global environmental change. *Science* **292**, 281–284 (2001).

39. Amundrud, S. L. & Srivastava, D. S. Trophic interactions determine the effects of drought on an aquatic ecosystem. *Ecology* **97**, 1475–1483 (2016).

40. Bell, G. *et al.* Trophic structure modulates community rescue following acidification. *Proc. R. Soc. B Biol. Sci.* **286**, 20190856 (2019).

41. Martin, T. E. & Maron, J. L. Climate impacts on bird and plant communities from altered animal–plant interactions. *Nat. Clim. Change* **2**, 195–200 (2012).

42. Ockendon, N. *et al.* Mechanisms underpinning climatic impacts on natural populations: altered species interactions are more important than direct effects. *Glob. Change Biol.* **20**, 2221–2229 (2014).

43. Loreau, M. *et al.* Biodiversity as insurance: from concept to measurement and application. *Biol. Rev.* **96**, 2333–2354 (2021).

44. Yachi, S. & Loreau, M. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc. Natl. Acad. Sci.* **96**, 1463–1468 (1999).

45. Bestion, E., Barton, S., García, F. C., Warfield, R. & Yvon-Durocher, G. Abrupt declines in marine phytoplankton production driven by warming and biodiversity loss in a microcosm experiment. *Ecol. Lett.* **23**, 457–466 (2020).

46. Hong, P. *et al.* Biodiversity promotes ecosystem functioning despite environmental change. *Ecol. Lett.* **25**, 555–569 (2022).

47. Isbell, F. *et al.* Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature* **526**, 574–577 (2015).

48. Baert, J. M., Laender, F. D., Sabbe, K. & Janssen, C. R. Biodiversity increases functional and compositional resistance, but decreases resilience in phytoplankton communities. *Ecology* **97**, 3433–3440 (2016).

49. Beaury, E. M., Finn, J. T., Corbin, J. D., Barr, V. & Bradley, B. A. Biotic resistance to invasion is ubiquitous across ecosystems of the United States. *Ecol. Lett.* **23**, 476–482 (2020).

50. Shurin, J. B. Dispersal limitation, invasion resistance, and the structure of pond zooplankton communities. *Ecology* **81**, 3074–3086 (2000).

51. Limberger, R. *et al.* Functional diversity can facilitate the collapse of an undesirable ecosystem state. *Ecol. Lett.* **26**, 883–895 (2023).

52. Estes, J. A. *et al.* Trophic downgrading of planet Earth. *Science* **333**, 301–306 (2011).

53. Pennekamp, F. *et al.* Biodiversity increases and decreases ecosystem stability. *Nature* **563**, 109–112 (2018).

54. Steiner, C. F., Long, Z. T., Krumins, J. A. & Morin, P. J. Temporal stability of aquatic food webs: partitioning the effects of species diversity, species composition and enrichment. *Ecol. Lett.* **8**, 819–828 (2005).

55. Huston, M. A. Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. *Oecologia* **110**, 449–460 (1997).

56. Schmid, B., Baruffol, M., Wang, Z. & Niklaus, P. A. A guide to analyzing biodiversity experiments. *J. Plant Ecol.* **10**, 91–110 (2017).

57. Pennekamp, F., Schtickzelle, N. & Petchey, O. L. BEMOVI, software for extracting behavior and morphology from videos, illustrated with analyses of microbes. *Ecol. Evol.* **5**, 2584–2595 (2015).

58. Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. & Leisch, F. *E1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*. (2021).

59. Hillebrand, H., Dürselen, C.-D., Kirschtel, D., Pollingher, U. & Zohary, T. Biovolume calculation for pelagic and benthic microalgae. *J. Phycol.* **35**, 403–424 (1999).

60. Sugihara, G. & May, R. M. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* **344**, 734–741 (1990).

61. Ye, H. & Sugihara, G. Information leverage in interconnected ecosystems: overcoming the curse of dimensionality. *Science* **353**, 922–925 (2016).

62. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

63. Rosseel, Y. lavaan: an R package for structural equation modeling. *J. Stat. Softw.* **48**, 1–36 (2012).

64. Oberski, D. lavaan.survey: an R package for complex survey analysis of structural equation models. *J. Stat. Softw.* **57**, 1–27 (2014).

Acknowledgements

We thank Fabian Avesani for help in the lab, Bernhard Schmid for discussions on the experimental design, and the Predictive Ecology Group of the University of Zurich for feedback on the manuscript. We thank Frederik Hammes for input on flowcytometry settings and for

access to a backup flowcytometer, with help from Jürg Sigrist. This project was part of SNF Project 310030_188431.

Contributions

Conceptualization: U.D., R.L., O.L.P.; **Data curation**: U.D., R.M.K., R.L., O.L.P.; **Formal analysis**: U.D., R.L.; **Funding acquisition**: O.L.P.; **Investigation**: Y.C., U.D., A.G., M.J., S.J., R.L., S.N., F.P., S.J.v.M, X.Z., D.Z-D.; **Methodology**: U.D., R.M.K., R.L., O.L.P; **Project administration**: R.L., O.L.P.; **Resources**: O.L.P.; **Software**: U.D., R.M.K.; **Supervision**: R.L., O.L.P.; **Validation**: U.D., R.M.K., R.L., F.P., O.L.P.; **Visualization**: U.D.; **Writing – original draft**: U.D., R.L.; **Writing – review & editing:** all authors.

Competing interests

The authors declare no competing interests.

Correspondence and requests for materials should be addressed to Romana Limberger or Uriah Daugaard.

Table 1: Results of two linear mixed models investigating the effects of species richness and light on the forecast error of taxa abundances and community biomass, respectively. We forecasted 253 taxa time series (across all bottles) and community biomass in 30 bottles. In both models, median centered richness and light treatment were fixed effects. The first model (response: forecast error of taxa abundances) included taxon, bottle, and incubator as random intercepts. The second model (response: forecast error of community biomass) included composition and incubator as random intercepts. Forecast error is based on forecasts computed with Simplex Empirical Dynamic Modeling (EDM). P-values < 0.05 in bold. See Supplement Table S6 for further details.

Response	Covariate	Estimate	95% CI	SS	DF	*F*-value	*p*-value
Forecast error of taxa abundances	Intercept	0.685	0.503 – 0.865				
	Richness	0.056	-0.008 – 0.121	0.008	1, 21.1	0.046	0.8328
	Light (decreasing)	-0.095	-0.265 – 0.081	0.225	1, 6.1	1.230	0.3092
	Richness:Light	-0.102	-0.193 – -0.014	0.922	1, 19.7	5.046	**0.0363**
Forecast error of community biomass	Intercept	0.668	0.456 – 0.879				
	Richness	0.038	-0.100 – 0.176	0.005	1, 19.5	0.060	0.8089
	Light (decreasing)	-0.034	-0.235 – 0.167	0.008	1, 13.1	0.108	0.7476
	Richness:Light	-0.046	-0.188 – 0.100	0.029	1, 13.4	0.387	0.5445

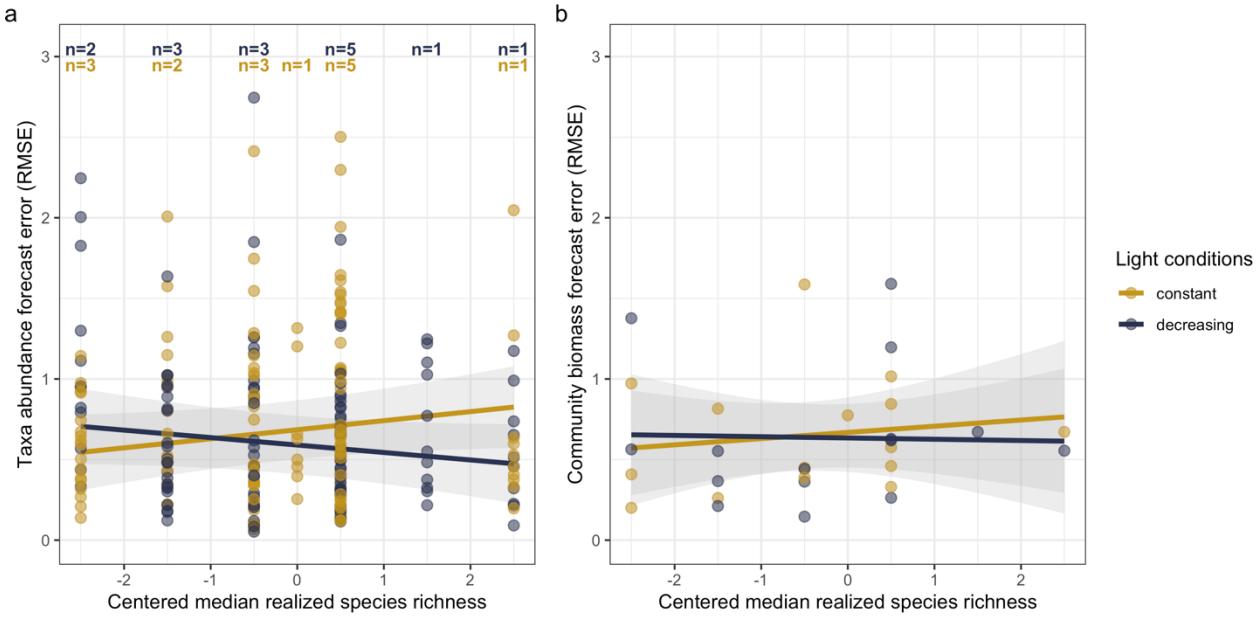

Figure 1: Effects of species richness and light on forecast error. The two treatments had interactive effects on forecast error of taxa abundances (a) but no effect on the forecast error of community biomass (b). Forecasts were computed with Simplex Empirical Dynamic Modeling (EDM), and forecast error was calculated as root mean square error (RMSE). RMSE is standardized (unitless). Observed richness deviated from planned richness because not all species persisted throughout the experiment. We therefore used realized richness (i.e. the median of the observed richness values) as explanatory variable. Median richness was centered; a centered median richness of zero corresponds to a median richness of 5.5 (minimum median richness: 3; maximum median richness: 8). Lines display the fit of the linear mixed model; shaded areas denote the 95% confidence intervals. In (a), we forecasted 253 taxa time series (across all bottles); each point represents a forecasted taxon in a bottle. Taxon, bottle, and incubator were included as random intercepts in the model. Inserts at the top indicate the number of bottles for the respective richness level (yellow: constant light, blue: decreasing light). In (b), we forecasted

community biomass in 30 bottles. Composition and incubator were included as random intercepts in the model.

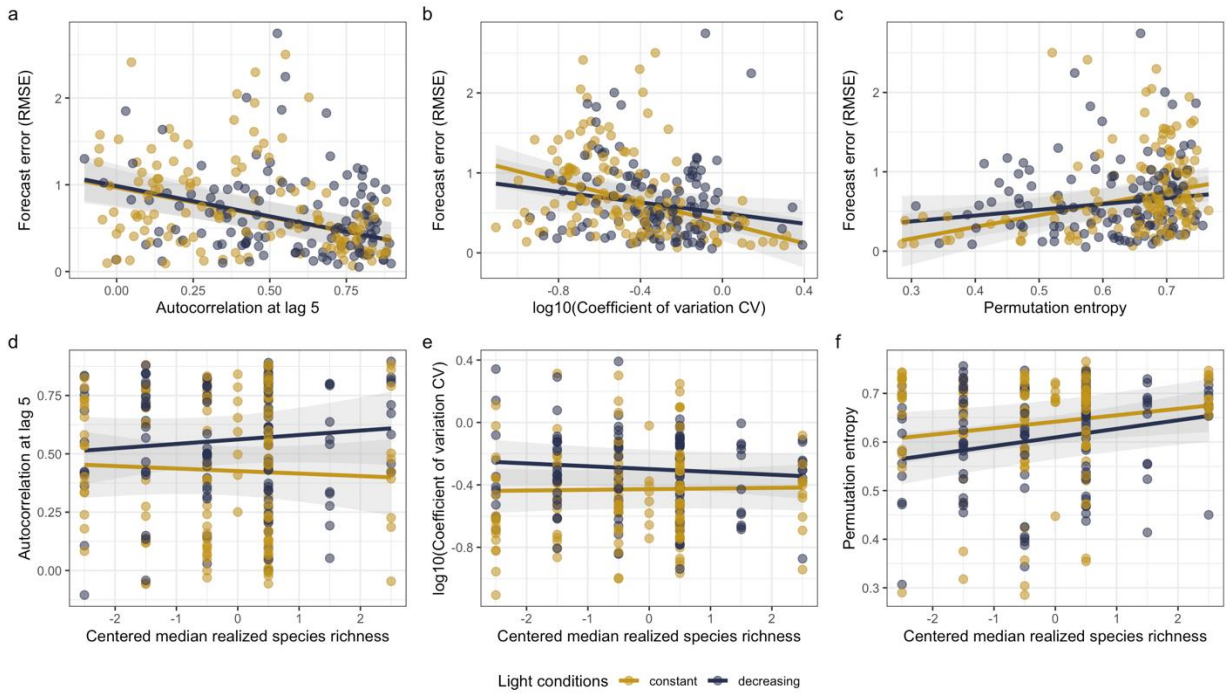

Figure 2: Relationship of time series metrics with forecast error (a-c), and with experimental treatments (d-f). Time series metrics are (a, d) autocorrelation at lag 5, (b, e) coefficient of variation as a measure of temporal variability, and (c, f) permutation entropy as a measure of time series complexity. All variables are unitless. Time series metrics were calculated for 253 taxa time series (across all bottles). Lines display the fit of the corresponding linear mixed model, shaded areas denote the 95% confidence intervals. Taxon, bottle, and incubator were included as random effects in the model. Forecast error is based on forecasts computed with Simplex EDM. Note that in (b) and (e) the coefficient of variation was $\log_{10}$-transformed so that linear model assumptions were met.

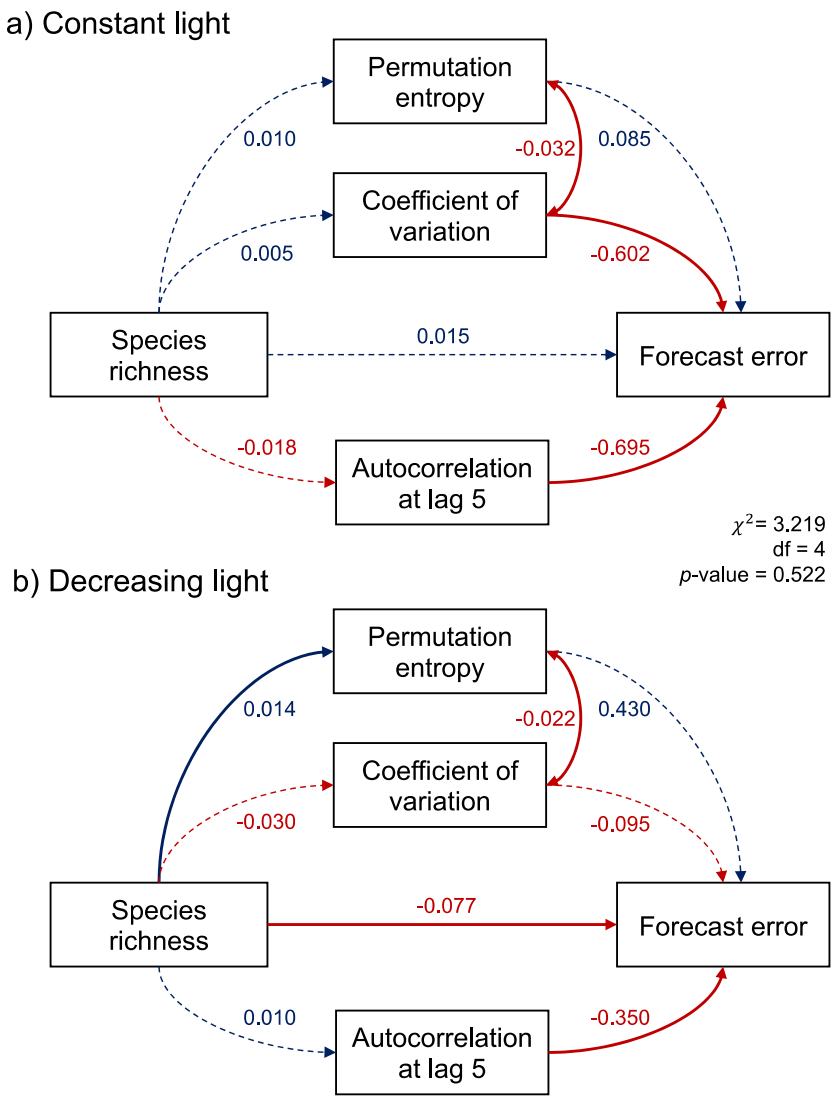

Figure 3: Structural equation model (SEM) showing the effects of species richness and time series metrics on forecast error of species abundances in (a) constant light, and (b) decreasing light. Solid and dashed arrows indicate significant and non-significant effects, respectively (see Supplementary Table S14). Values next to arrows are the path coefficients. The color matches the sign of the coefficients (blue=positive, red=negative). Bidirectional arrows are covariances between the error terms of the involved variables.