

Forecasting in the face of ecological complexity: number and strength of species interactions determines forecast skill in ecological communities

Uriah Dugaard¹, Stephan B. Munch², David Inauen¹, Frank Pennekamp¹, and Owen L. Petchey¹

¹Department of Evolutionary Biology and Environmental Studies, University of Zurich,
Winterthurerstrasse 190, 8057 Zurich, Switzerland

²Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA
95064, USA

Running title: Forecasting in the face of complexity

Corresponding author: Uriah Dugaard. Email address: uriah.daugaard@ieu.uzh.ch

ORCID and emails of authors:

- Uriah Dugaard: 0000-0003-4092-717X, uriah.daugaard@ieu.uzh.ch
- Stephan B. Munch: 0000-0002-9997-4934, smunch@ucsc.edu
- David Inauen: 0000-0002-8686-8008, david.inauen@uzh.ch
- Frank Pennekamp: 0000-0003-0679-1045, frank.pennekamp@ieu.uzh.ch
- Owen Petchey: 0000-0002-7724-1633, owen.petchey@ieu.uzh.ch

Statement of authorship: DI, FP and OLP conceived the experiment; UD, SBM and OLP conceived the research questions and designed the methodology; DI collected the data; UD analysed the data and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data accessibility statement: Data available from the Zenodo repository <https://doi.org/10.5281/zenodo.6323239>

Article type: Letter.

Number of words: 150 in abstract, 4989 in main text, 0 in boxes, 426 in figure legends

Number of figures: 4

Number of tables: 0

Number of boxes: 0

Number of references: 70

Keywords: Ecological forecasting, prediction, interaction strength, complexity, temperature, community ecology, food webs, empirical dynamic modelling, experiment, microbial community

30 **Abstract**

31 The potential for forecasting the dynamics of ecological systems is currently unclear, with contrasting opinions
32 regarding its feasibility due to ecological complexity. To investigate forecast skill within and across system
33 complexity, we monitored a microbial system exposed to either constant or fluctuating temperatures in a five
34 months long laboratory experiment. We tested how forecasting of species abundances depends on number and
35 strength of interactions and on model size (number of predictors). We also tested how greater system complexity
36 (i.e. the fluctuating temperatures) impacted these relations. We found that the more a species interacted, the
37 weaker these interactions were and the better its abundance was predicted. Forecast skill increased with model size.
38 Greater system complexity decreased forecast skill for three out of eight species. These insights into how abundance
39 prediction depends on the embedding of the species within the system and on overall system complexity could
40 improve species forecasting and monitoring.

41 **Introduction**

42 Over the last decades, it has become increasingly important to proficiently predict the consequences of climate
43 change and biodiversity loss (e.g. Godfray and May, 2014; Dietze, 2017). Ecological forecasting, formally defined as
44 the prediction of natural capital and ecosystems states and services, has advanced to be an imperative scientific and
45 applied discipline (Clark et al., 2001; Houlahan et al., 2017; Dietze et al., 2018). Examples of its applications include
46 predicting ecotoxicological effects on community responses (e.g. Clements and Rohr, 2009), forecasting the
47 successes of species invasions (e.g. Romanuk et al., 2009) and predicting how communities respond to climate
48 change (e.g. Hattab et al., 2016; Gaüzère et al., 2018; McCarthy et al., 2018). However, in the context of the
49 complexity of real-world systems, skilful ecological forecasting remains a major challenge to the point that its
50 feasibility has been questioned (Hayes and Barry, 2008; Beckage et al., 2011; Planque, 2016).

51 Generally, an ecological network or system is more complex the more variables (e.g. species) are part of it (Mitchell,
52 2009; Bradbury and Vehrencamp, 2014). The number of possible indirect interaction pathways between variables

53 rapidly increases with increasing network size (Borrett and Patten, 2003), and this is believed to hinder skilful
54 predictions (Yodzis, 1988; Wootton, 2002). In fact, some studies have found that prediction skill deteriorated with
55 increasing system complexity (e.g. Doak et al., 2008; Novak et al., 2011; Jonsson et al., 2018) and that species
56 interactions can reduce community predictability (Thompson et al., 2021), resulting in view that ecology is
57 unpredictable due to its complexity (Beckage et al., 2011). Yet, some recent results suggest the opposite: complexity
58 can increase rather than decrease prediction skill (Iles and Novak, 2016; Mougi, 2017). For instance, it was found
59 that the total abundance (the sum of all species abundances) was more predictable when the system consisted of
60 more species (Dornelas et al., 2011), while another study showed that the prediction of interaction strengths
61 improved with increasing food web size (Berlow et al., 2009). With evidence pointing in both directions, it remains
62 unclear whether there is a general relation between system complexity and forecast skill or whether each result is
63 specific to the system and to the quantity forecasted.

64 Within a system of a given complexity, commonly only a few strong species interactions are present with most
65 interactions being weak (e.g. Paine, 1992; Berlow et al., 2004; Wootton and Emmerson, 2005; Bascompte et al.,
66 2005), though still being important for system stability (e.g. McCann et al., 1998; O’Gorman and Emmerson, 2009;
67 Kadoya et al., 2018). Moreover, there is some evidence indicating that generalists (i.e. species with many possible
68 interacting partners) mostly have weak interactions, while specialist species (i.e. fewer possible interacting partners)
69 show the stronger interactions (Wootton and Stouffer, 2016). As in the case of increasing system complexity, the
70 more interactions a species has in a network the more indirect pathways exist that can influence its abundance.
71 Hence if prediction skill generally decreases with complexity, we might hypothesize that prediction skill for a given
72 species will also decrease the more interactions that species has within a network. However, we might alternatively
73 hypothesize that the sum of many weak interactions will have lower variance through time, making the focal species
74 easier to predict.

75 Whether the prediction of species abundances depends on how many interactions the species have has remained
76 untested so far, to our knowledge. In this study, we investigated this relation as its confirmation could help explain
77 why some species can be forecast better than others (Harris, 1994). In this context, we also tested whether the
78 forecast skill of species abundances depends on how much of the system is used in the forecast model.

79 Using a laboratory-based aquatic microbial community as our study system, we carried out a five months long
80 experiment. The community consisted of algae, bacteria, ciliates, flagellates and rotifer species. These species are
81 characterized by short generation times (e.g. Altermatt et al., 2015), which renders them convenient study organisms
82 for our experiment and questions. We exposed replicates of the community to either a constant temperature or a
83 fluctuating temperature setting. The fluctuating temperatures added a layer of complexity to the system, by
84 potentially affecting the species and their interactions in both direct and indirect ways.

85 We forecasted species abundances and estimated the number and the strength of species interactions using the
86 nonparametric time-series analysis framework empirical dynamic modelling (EDM). We build iterative forecasts that
87 included increasingly more variables as predictors. We hypothesized that the more a species is isolated (i.e. fewer
88 and weaker interactions), the better its abundance can be predicted as it is less dependent on the system state. We
89 expected that the fluctuating temperatures would decrease forecast skill, unless they are a strong enough driver of
90 system dynamics to outweigh the effect of increased system complexity. Further, we hypothesized that, in general,
91 forecasting improves when more system variables are included in the prediction, but that fewer variables are
92 necessary to achieve the highest or close to the highest forecast skill for more isolated species. Lastly, we also tested
93 whether a variable that interacts strongly with a focal species is also a good predictor variable of the abundance of
94 said species.

95 **Material and methods**

96 **Experiment: design, setup and sampling**

97 We carried out a laboratory-based experiment to record the dynamics of microbial communities (i.e. microcosms)
98 at constant (17.3 °C) and at fluctuating temperatures over a period of 154 days. We used three different fluctuating
99 temperature time series. One was identical to temperature time series of a local small stream (Furtbach ZH,
100 Switzerland) while we constructed the other two via spectral mimicry to resemble the recorded time series and its
101 characteristics (equal mean temperature with similar variance and autocorrelation, see Cohen et al., 1999; Petchey,

102 2000). For further information regarding the temperature time series see Fig. S1 and Section S2.1 in the
103 supplementary information.

104 The tri-trophic microcosms were semi-naturalistic with respect to the potential co-occurrences of the species and the
105 functional groups present (Table S5, Fig. S2). The first trophic level (i.e., the bottom level) of the community consisted
106 of three bacteria (*Serratia fonticola*, *Brevibacillus brevis* and *Bacillus subtilis*), an autotroph alga (*Chlamydomonas*
107 *reinhardtii*), a mixotroph alga (*Euglena gracilis*) and a mixotroph ciliate (*Euplotes daidaleos*). *E. gracilis* and *E.*
108 *daidaleos* are mixotrophic species and their trophic level is between the first and the second level (Ward and Follows,
109 2016). The second level contained three bacterivore ciliate species (*Colpidium striatum*, *Dexiostoma campylum* and
110 *Spirostomum* sp.), one omnivore ciliate species (*Paramecium caudatum*) and one omnivore rotifer species (*Rotifer*
111 sp.), while one ciliate predator species (*Didinium nasutum*) made up the top level. Further, small non-identified
112 flagellate species present in the stock cultures of the used species were also part of the microbial communities and
113 we classified them into the three groups "*small and white flagellates*", "*green and white flagellates*" and "*big and*
114 *white flagellates*". Prior to the experiment, we kept the ciliate and algae species in stock culture jars at 20 °C
115 containing organic protozoan pellet medium (Carolina Biological Supply Company, Burlington NC; concentration of
116 0.55 gL⁻¹, Altermatt et al., 2015). For the heterotrophic and mixotrophic species we bacterized the medium with the
117 above listed bacteria species. We fed *D. nasutum* with *P. caudatum ad libitum* and freshly established all stock
118 cultures two weeks prior to the experiment.

119 We set up the microcosms in 2L screw-capped glass bottles filled with 250 ml of the non-bacterized medium
120 containing *C. reinhardtii* at 50 cells/ml, 750 ml of the bacterized medium, a magnetic stirrer and 20 wheat seeds
121 for slow and continuous nutrient release. We added the remaining species (except *D. nasutum* and
122 *Spirostomum*) at a density of 0.1 cells/ml. As *Spirostomum* sp. only reached low abundances in the stock cultures
123 we inoculated it at a density of 0.005 cells/ml. We added the predator *D. nasutum* at a density of
124 0.02 cells/ml nine days after the start of the experiment. We reintroduced all species (except bacteria and *C.*
125 *reinhardtii*) at very low densities (<0.01 cells/ml) once a week. This rate was high enough so that extinct species
126 could potentially re-establish in the long-term and low enough to not influence population dynamics in the short-
127 term.

128 We kept the experimental bottles in temperature-controlled incubators with a 14/10h light-dark cycle. We had 18
129 replicates distributed across six incubators. We set three of the incubators to the constant temperature and assigned
130 one incubator to each of the three fluctuating temperature time series. Thus, nine replicates were in the constant
131 and nine in the fluctuating temperature environment.

132 We sampled the microcosms three times per week (Mondays, Wednesdays and Fridays) for 22 weeks (66 data points
133 per microcosm). We measured dissolved oxygen concentration using a non-invasive oxygen recorder (Precision
134 Sensing GmbH, Germany) with oxygen sensing optodes attached to the inside of the bottles. Before sampling, we
135 homogenized the microcosms on a magnetic plate. We sampled 65 ml from each replicate and added the same
136 amount of bacterized medium to them afterwards. We measured the abundances of the grouped bacteria species
137 and of the small, intermediate and large species by respectively using flow cytometry, FlowCAM imaging, video
138 microscopy and manual counts (supplementary Table S5). Video microscopy involved the R-package *bemovi*
139 (Pennekamp et al., 2015). For the video- and the FlowCAM data, we used automated species classification methods.
140 For more details regarding measurements and classifications see supplementary Section S2.2.

141 **Processing of recorded time series**

142 Preceding analyses, we processed the recorded time as follows (based on e.g. Benincà et al., 2008): we first
143 interpolated the time series using a cubic hermite spline to obtain equally distanced time points (time step of 2.3
144 days). To flatten sharp changes in abundances we carried out a fourth-root power transformation. We then
145 regressed the time series against time and henceforth used the resulting residuals, which are trendless, after we
146 standardized them.

147 Throughout the experiment, *Spirostomum* sp. remained practically extinct (Fig. S3M). As this species was effectively
148 not part of the microbial community, we did not consider it in the subsequent analyses. Further, the predator *D.*
149 *nasutum* did not show stable abundance. We used this species only as a predictor and interactor variable.
150 Accordingly, the forecasted target species were *C. reinhardtii*, *E. gracilis*, *E. daidaleos*,
151 *C. striatum*, *D. campylum*, *P. caudatum*, *Rotifer* sp. and the three bacteria species considered as one group (note that
152 for simplicity henceforth we refer to the bacteria group as a species).

153 **Forecasting of species abundances**

154 We forecasted the abundances of species using multiview empirical dynamic modeling (EDM, Ye et al., 2015; Ye and
155 Sugihara, 2016), as species dynamics are often nonlinear (Blonder et al., 2017; Clark and Luis, 2020). In EDM,
156 forecasting is based on the assumption that similar system states will lead to subsequent system states that are
157 again similar. In this method, state variables are used as predictors in both a non-lagged and a lagged fashion,
158 following Takens' theorem that the time series of a variable contains information of interacting variables (Takens,
159 1981). The lagged and non-lagged time series re-construct the attractor manifold and the number of time series
160 used for this is the embedding dimension E . For more information regarding EDM, see Ye et al. (2015).

161 Multiview EDM (Ye and Sugihara, 2016) is an extension of this method in which for a fixed embedding dimension all
162 possible combinations (called "views") of the predictor time series are constructed, which are then ranked by in-
163 sample forecast skill and the best k views are used for an average out-of-sample forecast. We used an embedding
164 dimension of $E = 3$ and a maximum lag of $l = 3$ (i.e. we lagged the predictors by zero days, 2.3 days and 4.7 days,
165 with species generation time ranging from hours to days, see e.g. Leary and Petchey, 2009; Altermatt et al., 2015).
166 For more information see supplementary Section S3.

167 For each species, we repeatedly forecasted its abundance using increasingly more predictors (i.e. we increased
168 forecast model size). Excluding temperature, there were 13 possible non-lagged predictors for each target species
169 (the eight target species, *D. nasutum*, the three flagellate groups and the dissolved oxygen). As the number of
170 predictors, we used $n = \{1,2,3,4,6,8,10,13\}$. For each value of n we calculated the number of possible non-lagged
171 predictor combinations $\binom{13}{n}$. Out of these combinations, we randomly selected 200 if the number of combinations
172 exceeded this value. We used the function `Multiview()` (R-package `rEDM`, Park et al., 2021) which adds the lags to
173 the predictor variables. We used the first 44 time points of the time series as the in-sample data and the last 22
174 points as the time points to be predicted (one-step ahead forecasts). For each predictor combination we evaluated
175 up to 25 values for k logarithmically spaced between 1 and 100. We previously determined the upper bound of $k =$
176 100 to be high enough (Fig. S5). We then repeated all forecasts with the temperature added to the same
177 combinations of predictors and finally also for when temperature was the sole predictor. In total, we fitted more

178 than 7.4 million multiview EDM models (Table S6). As a measure of forecast error, we calculated the RMSE of each
179 fitted model, with lower RMSE indicating a better forecast skill. Because we standardized the time series, an RMSE
180 below one indicates that the used model predicts the abundance of a species better than its mean abundance does.

181 **Estimation of number and strength of interactions**

182 For each target species T we determined which and how many state variables influenced its abundance by employing
183 convergent cross-mapping (CCM, see Sugihara et al., 2012), following the recommendations of Deyle et al. (2016,
184 see supplementary Section S4). CCM is a test of causation that reveals whether there is a causal link between the
185 test variables. We defined the state variables that showed a significant effect on a target species in a replicate as its
186 interactors and their sum as the number of interactions N_T .

187 We then estimated the interaction strength time series of the species that were causally linked using Smap EDM
188 (Deyle et al., 2016). The estimated pairwise species interaction strength time series are $S_{T,I}(t) = \partial T(t + \tau) / \partial I(t)$,
189 where t is a time point and $\tau = 2.3$ is the smallest time step, ∂ indicates the partial derivative and $T(t)$ and $I(t)$ are
190 the transformed abundance time series of the target and the interactor state variable, respectively. In S-map EDM,
191 at each time point t the community matrix (i.e. the matrix with elements $(\partial T / \partial I)_{ij}$, where i and j are the different
192 target and interactor state variables) is calculated. This calculation is done by including information of when the
193 system was in a similar state at other times through the use of locally weighted multivariate linear regressions. The
194 parameter θ determines how nearby system states are weighted in the regression. We used an intermediate value
195 ($\theta = 5$) and carried out a sensitivity analysis for it (Section S6.1). The interactor variables were limited to those that
196 influenced the target in a given replicate (based on the CCM analysis described above). We used the same eight
197 target species as for the abundance forecasting.

198 **Forecast error analyses**

199 **Relation between number of interactions, mean interaction strength and forecast error**

200 We calculated the mean interaction strength μ_T of target species T with the N_T state variables it interacted with as:

201

$$\mu_T = \frac{1}{N_T} \frac{1}{L} \sum_{I \in I_T} \sum_{t=1}^L |S_{T,I}(t)|.$$

202

(1)

203 In equation 1, I_T is the set consisting of the interactors that affected the target T , $S_{T,I}$ is the interaction strength time
204 series between target T and interactor $I \in I_T$, $|S_{T,I}(t)|$ is its absolute value at time point t , N_T is the number of
205 interactions and L is the number of time points in the time series. We then computed the sum of interaction
206 strengths Σ_T by multiplying equation 1 with the number of interactions N_T : $\Sigma_T = N_T \mu_T$.

207 We investigated the relations between the three explanatory variables N_T , μ_T and Σ_T and the forecast error (RMSE)
208 of species abundances. The RMSE value was based on the forecast model in which all state variables were used as
209 predictors. We fitted three separate linear mixed models with RMSE as the response variable and one of the three
210 explanatory variables as the regressor. We fitted a fourth mixed model between μ_T and N_T . In all of the models we
211 included the temperature regime (i.e. constant or fluctuating) and its interaction with the other explanatory variable.

212 **Forecast error as a function of number of predictors and number of interactors**

213 We investigated the relation between median forecast error and the number of predictors and the temperature
214 regime (constant or fluctuating) with a linear mixed model conjointly for the eight different target species. We used
215 the median RMSE as the response variable, while the \log_{10} -transformed number of used predictors, the temperature
216 regime, a binary variable indicating whether temperature was used as a predictor and the target species were the
217 explanatory variables, alongside their pairwise interactions. We included bottle ID nested in target species as a
218 random intercept.

219 Using the same settings as before, we then forecasted species abundances again using as predictors only variables
220 that influenced the target species (based on the CCM analysis). Among all forecast models, we selected the ones
221 that predicted the target species the best. For this, from the models that yielded an RMSE within 1% of the lowest
222 achieved RMSE (for a given species and replicate) we selected the models with the least predictor variables (i.e. the
223 smallest models). We then used the number of predictors in the best forecast model as the response variable in a

224 mixed model that included the number of interactions, the temperature regime and their interaction as explanatory
225 variables.

226 **Interactor strength versus predictor importance**

227 We investigated whether stronger interactors are also better predictors. In each replicate we calculated the mean
228 interaction strength of each evaluated target-interactor species pair as the mean of the absolute values of their
229 interaction strength time series. We log₁₀-transformed this variable and used it in a mixed model in a three-way
230 interaction with the temperature regime and the target species. The response variable was the RMSE of the forecast
231 model in which the interacting species was the only predictor.

232 We fitted all linear mixed models using the function lmer (R-package lme4, Bates et al., 2015). We included bottle ID
233 as a random intercept in all models, if not specified otherwise.

234 **Results**

235 **Relation between number of interactions, mean interaction strength and forecast error**

236 Both the number of interactions and the mean interaction of a target species had a significant effect on the forecast
237 error of species abundances (Fig. 1, Table S1). Forecast error decreased (i.e., forecast skill increased) the more
238 interactions a species had (Fig. 1A). Quantitatively, with every unit increase in number of interactions the forecast
239 error decreased by 5.2% (constant temperature) and 3.9% (fluctuating temperature) with respect to its biggest value
240 (t -value=-6.04, df =140, p -value<0.001). Meanwhile, the stronger a species interacted on average, the worse it was
241 predicted (Fig. 1B): with respect to its biggest value the forecast error increased by 6.5% (constant) and 4.8%
242 (fluctuating) for every 0.1 increase in mean interaction strength (t -value=6.45, df =127, p -value<0.001). The number
243 of interactions and the mean interaction strength of a species were negatively correlated, with the mean interaction
244 strength decreasing by 0.053 (constant) and 0.055 (fluctuating) for every unit increase in number of interactions (t -
245 value=-9.66, df =140, p -value<0.001, Fig. 1C), indicating that the more interactions a species had the weaker these
246 were. This resulted in the sum of interaction strength (the product of these two quantities) to be unrelated with

247 forecast error (t -value=1.40, df =140, p -value=0.163, Fig. 1D). Neither the temperature to which the microcosms
248 were exposed (constant or fluctuating) nor its interaction with the other considered explanatory variables had
249 significant effects on any of these results (Table S1). Each target species had a comparable number of interactions
250 across replicates (Fig. S8).

251 The described patterns persisted in a sensitivity analysis for the parameter θ (Fig. S10) and across several robustness
252 analyses reported in detail in the supplementary Section S6 (e.g. using a different measure of forecast proficiency
253 and estimating interactions strengths with a multivariate auto-regressive system state model, see Holmes et al.,
254 2012, and Fig. S15).

255 **Forecast error as a function of number of predictors and temperature**

256 In general, as would be expected, median forecast error of species abundances decreased the more state variables
257 were used as predictors (F -value=1847.81, p -value<0.001, Fig. 2, Table S2): the respective slopes ranged from -0.579
258 to -0.055 across target species (F -value=175.86, p -value<0.001), with 0.022 subtracted to these slopes in the case
259 of fluctuating temperatures (F -value=5.76, p -value=0.016). Overall, fluctuating temperatures increased forecast
260 errors (F -value=25.09, p -value<0.001), but this was not the case for the forecasting of all target species as the
261 difference in forecast error ranged from -0.015 to 0.392 across them (F -value=4.79, p -value<0.001). Specifically, the
262 fluctuating temperatures increased the forecasting error of *C. reinhardtii*, and there was moderate evidence for the
263 same being the case for the species *Rotifer* sp. and *P. caudatum* (respective differences in forecast errors of 0.392,
264 0.215 and 0.202, see Fig. 2 and Table S2). Further, the inclusion of temperature as a predictor decreased the forecast
265 error (F -value=26.62, p -value<0.001), with the change ranging from -0.079 to -0.005 across targets (F -value=7.53, p -
266 value<0.001, with no significant difference between the two temperature regimes (F -value=0.31, p -value=0.580).
267 The value of using temperature as a predictor decreased the more other predictors were used in the forecasting,
268 with temperature reducing forecast errors by 0.048 less for every ten predictors added to the model (F -value=27.53,
269 p -value<0.001).

270 The number of predictors used in the best forecast model of the abundance of a species was independent of the
271 number of interactions of said species (t -value=0.47, df =130, p -value=0.637, Fig. 3A, Table S3) regardless of the
272 temperature regime (t -value=-0.26, df =137, p -value=0.793). Across replicates, when the temperature was constant
273 38.9% of the best models had three predictors, 27.8% had two predictors and 22.2% had four predictors, and
274 similarly when the temperature varied in most cases the best models had three (43.1%), two (30.6%) and four
275 (15.3%) predictors (Fig. 3B). The highest forecast skill (smallest RMSE values) as a function of number of predictors
276 confirmed this result (Fig. S6).

277 **Interactor strength versus predictor importance**

278 Overall, we found no relation between the interaction strength of a state variable with a target species and the
279 forecast error of the abundance of the target species with the state variable as the sole predictor (F -value=2.30, p -
280 value=0.129, Fig. 4, Table S4), regardless of temperature regime (F -value=0.05, p value=0.822). The estimated slopes
281 varied across target species from -0.523 to 0.165 (F -value=7.42, p value<0.001), but only for *C. reinhardtii* and *E.*
282 *gracilis* there was evidence that the stronger interactors predicted these target species with smaller forecast errors.
283 For *C. reinhardtii* the slopes were -0.315 and -0.336 and for *E. gracilis* they were -0.284 and -0.523, respectively in
284 the constant and in the fluctuating temperature setting (see Figs. 4B,E and Table S4).

285 **Discussion**

286 We found that the forecast skill of the abundance of a species increases the more interactions the species has within
287 the system (e.g. with other species) but also that it increases the weaker these interactions are on average. We
288 found that these two measures — the number of interactions and their mean strength — are negatively correlated
289 resulting in the abundance of species with many but on average weak interactions to be predicted the most skilfully.
290 While the fluctuating temperatures did not influence these findings, they lowered the median skill of forecasting the
291 abundances of three out of eight target species.

292 Previous studies reported contrasting results as in some cases predictions improved with increasing system
293 complexity (Berlow et al., 2009; Dornelas et al., 2011; Iles and Novak, 2016; Mougi, 2017), while in others the
294 opposite was the case (e.g. Doak et al., 2008; Novak et al., 2011; Jonsson et al., 2018). The latter led to the statement
295 that ecological forecasting is limited by the low intrinsic predictability of real-world systems due to their great
296 complexity (Beckage et al., 2011). In our study, the addition of complexity (i.e. the fluctuating temperature) to the
297 system lowered forecast skill for some but not all species. While this result is evidence for a negative relation
298 between system complexity and forecast skill, it also indicates that this relation can be species-specific. Thus, an
299 universal association between increasing system complexity and the predictability of the abundance of its
300 components is less likely to exist.

301 Contrasting this, the more connected a species was in the system, the better it was forecasted. The negative
302 correlation between the number of interactions and the mean interaction strength could explain this as it might
303 indicate that species with many but weak interactions were less dependent on the state of individual system
304 components and more dependent on the state of the whole system. Regardless, this result provides a first insight
305 into why certain aspects of ecological systems are more predictable than others (e.g. some species abundances more
306 than others; Harris, 1994). It suggests that species with few, strong interactions should be sampled more frequently
307 than those with many weak interactions to achieve a comparable forecast skill. Thus, it has the potential of improving
308 the monitoring of species in real-world ecosystem, which can be a costly endeavor (e.g. Manley et al., 2004; Jones,
309 2011).

310 Yet, it remains unclear why a species with few strong interactions is not predicted more skilfully than a species with
311 many weaker interactions. However, this result is corroborated by the finding that interaction strength is not a good
312 indicator of how well an interacting variable predicts the abundance of the target species. Knowledge about good
313 predictors of species abundances could help our understanding of ecological forecasting and our skill in carrying it
314 out (Petchey et al., 2015). Based on these results it is likely that interaction strength can be excluded as selection
315 criteria for good predictors of species abundances.

316 In this context, we found that, as expected, median forecast skill increased the more system variables we included
317 as predictors in the forecast models. However, in most cases we achieved the best forecast skill already with few
318 predictors included (i.e. between two and four predictors in approximately 90% of forecast models), regardless of
319 how many interactions the forecast species had. This suggests that if it is known *a priori* which system components
320 are good predictors of the abundance of a specific species, then data collection can potentially be streamlined by
321 focusing on these variables rather than on the whole system.

322 In our experiment we compared fluctuating and constant temperatures as the former are more truthful to natural
323 conditions. The lower predictability of the abundance of some species in the fluctuating temperature setting when
324 compared to the constant temperature setting suggests that prediction skill might be overestimated in experiments
325 in which temperatures are constant. In fact, in laboratory-based or simulated time series experiments the
326 temperature is usually kept at one or more constant levels (e.g. in Yeo et al., 2003; Ferguson and Ponciano, 2014;
327 Daugaard et al., 2019) and only rarely fluctuating temperatures are used (e.g. Descamps-Julien and Gonzalez, 2005;
328 Jiang and Morin, 2007). Given that temperature is a strong driver of species metabolic rates (Brown et al., 2004) and
329 thus also of their dynamics (e.g. Lee et al., 2007; Bernhardt et al., 2018), fluctuating temperatures should more
330 frequently be considered to better reconcile results from laboratory or simulation experiments with real-world
331 insights.

332 The distribution of interaction strengths in a system is known to be right-skewed, with the bulk of the interactions
333 being weak and only comparably few interactions being strong (e.g. Paine, 1992; Wootton, 1997; Wootton and
334 Emmerson, 2005). This was also the case in our study (supplementary Fig. S9), while the number of interactions
335 remained comparable within species across replicates (supplementary Fig. S8). Moreover, our finding that the
336 number of interactions and the average interaction strength are strongly negatively correlated represents novel
337 empirical evidence of the theoretical finding that generalists have predominately weak interactions while specialist
338 are responsible for the right-skew of the interaction strength distribution (Wootton and Stouffer, 2016). Given that
339 weak interactions have been identified as systemstabilizing (e.g. McCann and Hastings, 1997; Neutel et al., 2002;
340 Otto et al., 2007), our results support previous observations stating that generalist species have a stabilizing

341 function due to the weak interactions they engage in (e.g. Mougi and Nishimura, 2007; Chakraborty, 2015; Brechtel
342 et al., 2019, note, however, that we did not carry out a stability analysis of the system in this study).

343 Several robustness analyses confirm that the results are not sensitive to the specifications of the experimentation
344 and analyses. Noticeably, the analyses of the potential influence of the different measurement methods
345 (videography, manual count, flowCam, and flow cytometry) on the results revealed that the main results most often
346 still occurred within measurement methods (Section S6.2.7). However, the 95% confidence intervals often
347 overlapped zero due to small sample sizes. The analyses also showed that any effect of measurement method was
348 not due to anything as simple as differences in measurement error across methods (Section S6.2.6). Since
349 measurement method is confounded with species identity, we cannot tease apart their possible influences on the
350 main results, and therefore cannot completely rule out that measurement method has, for some reason, some
351 power in explaining the main results reported.

352 In conclusion, we provide novel insights into why the abundance of some species are better predictable than others
353 in the same system. The dependency of forecast skill on the number and the strength of species interactions not
354 only improves our knowledge of ecological forecasting. It has also the potential of improving the resource allocation
355 for the sampling and monitoring of species, as comparable forecast skill across species likely requires varying
356 amounts of data per predicted species based on how much and how strongly this species interacts. We also shed
357 further light on the relationship between elements of system complexity and forecast skill, showing that the
358 relationship can be both species-specific and of different sign within and across systems. Thus, forecasting skill may
359 deteriorate with increasing complexity, but this cannot be taken for granted and can depend on whether one is
360 comparing across or within systems.

361 **Acknowledgments**

362 The research was supported by Swiss National Science Foundation Project 310030_188431. We thank Yves Choffat
363 for his help with the experiment and the Predictive Ecology Group of the University of Zurich for the provided
364 feedback on an early draft of this article.

365 **References**

- 366 Altermatt, F., Fronhofer, E.A., Garnier, A., Giometto, A., Hammes, F., Klecka, J., *et al.* (2015). Big answers from small
367 worlds: a user's guide for protist microcosms as a model system in ecology and evolution. *Methods in Ecology and*
368 *Evolution*, 6, 218–231.
- 369 Bascompte, J., Melián, C.J. & Sala, E. (2005). Interaction strength combinations and the overfishing of a marine food
370 web. *Proceedings of the National Academy of Sciences*, 102, 5443–5447.
- 371 Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of*
372 *Statistical Software*, 67, 1–48.
- 373 Beckage, B., Gross, L.J. & Kauffman, S. (2011). The limits to prediction in ecological systems. *Ecosphere*, 2, art125.
- 374 Benincà, E., Huisman, J., Heerkloss, R., Jöhnk, K.D., Branco, P., Van Nes, E.H., *et al.* (2008). Chaos in a long-term
375 experiment with a plankton community. *Nature*, 451, 822–825.
- 376 Berlow, E.L., Dunne, J.A., Martinez, N.D., Stark, P.B., Williams, R.J. & Brose, U. (2009). Simple prediction of interaction
377 strengths in complex food webs. *Proceedings of the National Academy of Sciences*, 106, 187–191.
- 378 Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., *et al.* (2004). Interaction
379 strengths in food webs: issues and opportunities. *Journal of animal ecology*, 73, 585–598.
- 380 Bernhardt, J.R., Sunday, J.M. & O'Connor, M.I. (2018). Metabolic theory and the temperature-size rule explain the
381 temperature dependence of population carrying capacity. *The American Naturalist*, 192, 687–697.
- 382 Blonder, B., Moulton, D.E., Blois, J., Enquist, B.J., Graae, B.J., Macias-Fauria, M., *et al.* (2017). Predictability in
383 community dynamics. *Ecology Letters*, 20, 293–306.
- 384 Borrett, S.R. & Patten, B.C. (2003). Structure of pathways in ecological networks: relationships between length and
385 number. *Ecological Modelling*, ISEM The third European Ecological Modelling Conference, 170, 173–184.
- 386 Bradbury, J.W. & Vehrencamp, S.L. (2014). Complexity and behavioral ecology. *Behavioral Ecology*, 25, 435–442.

387 Brechtel, A., Gross, T. & Drossel, B. (2019). Far-ranging generalist top predators enhance the stability of meta-
388 foodwebs. *Sci Rep*, 9, 12268.

389 Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M. & West, G.B. (2016). Toward a metabolic theory of ecology.
390 *Ecology*, 1771–1789.

391 Chakraborty, S. (2015). The influence of generalist predators in spatially extended predator–prey systems. *Ecological*
392 *Complexity*, 23, 50–60.

393 Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.A., *et al.* (2001). Ecological forecasts: an
394 emerging imperative. *Science*, 293, 657–660.

395 Clark, T.J. & Luis, A.D. (2020). Nonlinear population dynamics are ubiquitous in animals. *Nat Ecol Evol*, 4, 75–81.

396 Clements, W.H. & Rohr, J.R. (2009). Community responses to contaminants: using basic ecological principles to
397 predict ecotoxicological effects. *Environmental Toxicology and Chemistry*, 28, 1789–1800.

398 Cohen, J.E., Newman, C.M., Cohen, A.E., Petchey, O.L. & Gonzalez, A. (1999). Spectral mimicry: a method of
399 synthesizing matching time series with different Fourier spectra. *Circuits Systems and Signal Process*, 18, 431–442.

400 Dagaard, U., Petchey, O.L. & Pennekamp, F. (2019). Warming can destabilize predator–prey interactions by shifting
401 the functional response from Type III to Type II. *Journal of Animal Ecology*, 88, 1575–1586.

402 Descamps-Julien, B. & Gonzalez, A. (2005). Stable coexistence in a fluctuating environment: an experimental
403 demonstration. *Ecology*, 86, 2815–2824.

404 Deyle, E.R., May, R.M., Munch, S.B. & Sugihara, G. (2016). Tracking and forecasting ecosystem interactions in real
405 time. *Proceedings of the Royal Society B: Biological Sciences*, 283, 20152258.

406 Dietze, M. (2017). *Ecological forecasting. Ecological Forecasting*. Princeton University Press.

407 Dietze, M.C., Fox, A., Beck-Johnson, L.M., Betancourt, J.L., Hooten, M.B., Jarnevich, C.S., *et al.* (2018). Iterative near-
408 term ecological forecasting: needs, opportunities, and challenges. *PNAS*, 115, 1424–1432.

409 Doak, D.F., Estes, J.A., Halpern, B.S., Jacob, U., Lindberg, D.R., Lovvorn, J., *et al.* (2008). Understanding and predicting
410 ecological dynamics: are major surprises inevitable. *Ecology*, 89, 952–961.

411 Dornelas, M., Phillip, D.A.T. & Magurran, A.E. (2011). Abundance and dominance become less predictable as species
412 richness decreases. *Global Ecology and Biogeography*, 20, 832–841.

413 Ferguson, J.M. & Ponciano, J.M. (2014). Predicting the process of extinction in experimental microcosms and
414 accounting for interspecific interactions in single-species time series. *Ecology Letters*, 17, 251–259.

415 Gauzere, P., Iversen, L.L., Barnagaud, J.-Y., Svenning, J.-C. & Blonder, B. (2018). Empirical predictability of community
416 responses to climate change. *Frontiers in Ecology and Evolution*, 6, 186.

417 Godfray, H.C.J. & May, R.M. (2014). Open questions: are the dynamics of ecological communities predictable? *BMC*
418 *Biology*, 12, 22.

419 Harris, G. (1994). Pattern, process and prediction in aquatic ecology - a limnological view of some general ecological
420 problems. *Freshwater Biology*, 32, 143–160.

421 Hattab, T., Leprieur, F., Lasram, F.B.R., Gravel, D., Le Loc’h, F. & Albouy, C. (2016). Forecasting fine-scale changes in
422 the food-web structure of coastal marine communities under climate change. *Ecography*, 39, 1227–1237.

423 Hayes, K.R. & Barry, S.C. (2008). Are there any consistent predictors of invasion success? *Biol Invasions*, 10, 483–506.

424 Holmes, E., E., Ward, E., J. & Wills, K. (2012). MARSS: multivariate autoregressive state-space models for analyzing
425 time-series data. *The R Journal*, 4, 11.

426 Houlahan, J.E., McKinney, S.T., Anderson, T.M. & McGill, B.J. (2017). The priority of prediction in ecological
427 understanding. *Oikos*, 126, 1–7.

428 Iles, A.C. & Novak, M. (2016). Complexity increases predictability in allometrically constrained food webs. *American*
429 *Naturalist*, 188, 87–98.

430 Jiang, L. & Morin, P.J. (2007). Temperature fluctuation facilitates coexistence of competing species in experimental
431 microbial communities. *J Anim Ecology*, 76, 660–668.

432 Jones, J.P.G. (2011). Monitoring species abundance and distribution at the landscape scale. *Journal of Applied*
433 *Ecology*, 48, 9–13.

434 Jonsson, T., Kaartinen, R., Jonsson, M. & Bommarco, R. (2018). Predictive power of food web models based on body
435 size decreases with trophic complexity. *Ecology Letters*, 21, 702–712.

436 Kadoya, T., Gellner, G. & McCann, K.S. (2018). Potential oscillators and keystone modules in food webs. *Ecology*
437 *Letters*, 21, 1330–1340.

438 Leary, D.J. & Petchey, O.L. (2009). Testing a biological mechanism of the insurance hypothesis in experimental
439 aquatic communities. *Journal of Animal Ecology*, 78, 1143–1151.

440 Lee, K.-S., Park, S.R. & Kim, Y.K. (2007). Effects of irradiance, temperature, and nutrients on growth dynamics of
441 seagrasses: a review. *Journal of Experimental Marine Biology and Ecology*, The Biology and Ecology of Seagrasses,
442 350, 144–175.

443 Manley, P.N., Zielinski, W.J., Schlesinger, M.D. & Mori, S.R. (2004). Evaluation of a multiple-species approach to
444 monitoring species at the ecoregional scale. *Ecological Applications*, 14, 296–310.

445 McCann, K. & Hastings, A. (1997). Re-evaluating the omnivory–stability relationship in food webs. *Proceedings of*
446 *the Royal Society of London. Series B: Biological Sciences*, 264, 1249–1254.

447 McCann, K., Hastings, A. & Huxel, G.R. (1998). Weak trophic interactions and the balance of nature. *Nature*, 395,
448 794–798.

449 McCarthy, J.K., Mokany, K., Ferrier, S. & Dwyer, J.M. (2018). Predicting community rank-abundance distributions
450 under current and future climates. *Ecography*, 41, 1572–1582.

451 Mitchell, M. (2009). *Complexity: a guided tour*. Oxford University Press.

452 Mougi, A. (2017). Spatial complexity enhances predictability in food webs. *Scientific Reports*, 7, 1–5.

453 Mougi, A. & Nishimura, K. (2007). A resolution of the paradox of enrichment. *Journal of Theoretical Biology*, 248,
454 194–201.

455 Neutel, A.-M., Heesterbeek, J.A. & De Ruiter, P.C. (2002). Stability in real food webs: weak links in long loops. *Science*,
456 296, 1120–1123.

457 Novak, M., Wootton, J.T., Doak, D.F., Emmerson, M., Estes, J.A. & Tinker, M.T. (2011). Predicting community
458 responses to perturbations in the face of imperfect knowledge and network complexity. *Ecology*, 92, 836–846.

459 O’Gorman, E.J. & Emmerson, M.C. (2009). Perturbations to trophic interactions and the stability of complex food
460 webs. *Proceedings of the National Academy of Sciences*, 106, 13393–13398.

461 Otto, S.B., Rall, B.C. & Brose, U. (2007). Allometric degree distributions facilitate food-web stability. *Nature*, 450,
462 1226–1229.

463 Paine, R.T. (1992). Food-web analysis through field measurement of per capita interaction strength. *Nature*, 355,
464 73–75.

465 Park, J., Smith, C., Sugihara, G. & Deyle, E. (2021). *rEDM: empirical dynamic modeling* (‘EDM’).

466 Pennekamp, F., Schtickzelle, N. & Petchey, O.L. (2015). Bemovi, software for extracting behavior and morphology
467 from videos, illustrated with analyses of microbes. *Ecology and Evolution*, 5, 2584–2595.

468 Petchey, O.L. (2000). Environmental colour affects aspects of single–species population dynamics. *Proceedings of*
469 *the Royal Society of London. Series B: Biological Sciences*, 267, 747–754.

470 Petchey, O.L., Pontarp, M., Massie, T.M., Kéfi, S., Ozgul, A., Weilenmann, M., *et al.* (2015). The ecological forecast
471 horizon, and examples of its uses and determinants. *Ecology Letters*, 18, 597–611.

472 Planque, B. (2016). Projecting the future state of marine ecosystems, “la grande illusion”? *ICES J Mar Sci*, 73, 204–
473 208.

474 Romanuk, T.N., Zhou, Y., Brose, U., Berlow, E.L., Williams, R.J. & Martinez, N.D. (2009). Predicting invasion success in
475 complex ecological networks. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 364, 1743–
476 1754.

477 Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., *et al.* (2012). Detecting causality in complex ecosystems.
478 *Science*, 338, 496–500.

479 Takens, F. (1981). Detecting strange attractors in turbulence. In: *Dynamical Systems and Turbulence, Warwick 1980*,
480 Lecture Notes in Mathematics (eds. Rand, D. & Young, L.-S.). Springer, Berlin, Heidelberg, pp. 366–381.

481 Thompson, P., Hürlemann, S. & Altermatt, F. (2021). Species interactions limit the predictability of community
482 responses to environmental change. *The American Naturalist*, 198, 694–705.

483 Ward, B.A. & Follows, M.J. (2016). Marine mixotrophy increases trophic transfer efficiency, mean organism size, and
484 vertical carbon flux. *Proceedings of the National Academy of Sciences*, 113, 2958–2963.

485 Wootton, J.T. (1997). Estimates and tests of per capita interaction strength: diet, abundance, and impact of
486 intertidally foraging birds. *Ecological Monographs*, 67, 45–64.

487 Wootton, J.T. (2002). Indirect effects in complex ecosystems: recent progress and future challenges. *Journal of Sea*
488 *Research, Structuring Factors of Shallow Marine Coastal Communities, Part I*, 48, 157–172.

489 Wootton, J.T. & Emmerson, M. (2005). Measurement of interaction strength in nature. *Annual Review of Ecology,*
490 *Evolution, and Systematics*, 36, 419–444.

491 Wootton, K.L. & Stouffer, D.B. (2016). Many weak interactions and few strong; food-web feasibility depends on the
492 combination of the strength of species' interactions and their correct arrangement. *Theor Ecol*, 9, 185–195.

493 Ye, H., Beamish, R.J., Glaser, S.M., Grant, S.C.H., Hsieh, C., Richards, L.J., *et al.* (2015). Equation-free mechanistic
494 ecosystem forecasting using empirical dynamic modeling. *Proc Natl Acad Sci USA*, 112, E1569.

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

497 Yeo, H., Pell, J.K., Alderson, P.G., Clark, S.J. & Pye, B.J. (2003). Laboratory evaluation of temperature effects on the
498 germination and growth of entomopathogenic fungi and on their pathogenicity to two aphid species. *Pest*
499 *Management Science*, 59, 156–165.

500 Yodzis, P. (1988). The indeterminacy of ecological interactions as perceived through perturbation experiments.
501 *Ecology*, 69, 508–515.